

GESELLSCHAFT
FÜR INFORMATIK



Maike Klein, Daniel Krupka, Cornelia Winter, Volker
Wohlgemuth (Hrsg.)

INFORMATIK 2023

Designing Futures: Zukünfte gestalten

26. – 29. September 2023

Berlin

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings

Series of the Gesellschaft für Informatik (GI)

Volume 337

ISBN 978-3-88579-731-9

ISSN 1617-5468

Volume Editors

Maike Klein (maike.klein@gi.de)

Gesellschaft für Informatik, Anna-Louisa-Karsch-Straße 2, 10178 Berlin

Daniel Krupka (daniel.krupka@gi.de)

Gesellschaft für Informatik, Anna-Louisa-Karsch-Straße 2, 10178 Berlin

Cornelia Winter (cornelia.winter@gi.de)

Gesellschaft für Informatik, Ahrstraße 45, 53175 Bonn

Volker Wohlgemuth (volker.wohlgemuth@htw-berlin.de)

Hochschule für Technik und Wirtschaft Berlin (htw), 10313 Berlin (Postfach)

Series Editorial Board

Andreas Oberweis, KIT Karlsruhe

Torsten Brinda, Universität Duisburg-Essen

Dieter Fellner, Technische Universität Darmstadt

Ulrich Frank, Universität Duisburg-Essen

Michael Goedicke, Universität Duisburg-Essen

Barbara Hammer, Universität Bielefeld

Ralf Hofestädt, Universität Bielefeld

Wolfgang Karl, KIT Karlsruhe

Michael Koch, Universität der Bundeswehr München

Heiko Roßnagel, Fraunhofer IAO Stuttgart

Andreas Thor, HTWK Leipzig

Ingo Timm, Universität Trier

Karin Vosseberg, Hochschule Bremerhaven

Maria Wimmer, Universität Koblenz-Landau

Dissertations: Rüdiger Reischuk, Universität zu Lübeck

Thematics: Agnes Koschmider, Universität Kiel

Seminars: Judith Michael, RWTH Aachen

© Gesellschaft für Informatik, Bonn 2023

printed by Köllen Druck+Verlag GmbH, Bonn



This book is licensed under a Creative Commons BY-SA 4.0 licence.

Vorwort

Wie gelingt es uns, die Zukunft so zu gestalten, dass sie für die großen gesellschaftlichen und wirtschaftlichen Herausforderungen wie Dekarbonisierung oder demografischer Wandel gewappnet ist? Eine Frage, auf die es viele unterschiedliche Antworten gibt. Fest steht jedoch: Ohne Informatik wird es nicht gehen. Deswegen haben wir die INFORMATIK 2023, die offizielle Jahrestagung der Gesellschaft für Informatik e.V. (GI), unter das Motto „Designing Futures: Zukünfte gestalten“ gestellt. Dabei stehen die Gestaltung digitaler Technologien und informatischer Systeme sowie deren Gestaltungskraft für gesellschaftliche, kulturelle, soziale und wirtschaftliche Prozesse im Fokus.

Dieser Tagungsband führt vor Augen, wie vielseitig sowohl die Fragen, als auch die Antworten sind, vor denen die Informatik heutzutage steht. In den mehr als 200 Beiträgen finden Sie innovative Ansätze, neue Perspektiven und viele kluge Gedanken aus unterschiedlichen Fachrichtungen und Anwendungsgebieten.

Denn für die deutschsprachige Informatik-Community ist die INFORMATIK jedes Jahr ein wichtiger Treffpunkt und eine gute Gelegenheit, sich fachübergreifend auszutauschen und neue Kontakte zu knüpfen. Dieses Jahr fand sie vom 26. bis 29. September 2023 an der Hochschule für Wirtschaft und Technik in Berlin statt. Um noch mehr Aspekte der Informatik einzubinden, konnten die Teilnehmenden parallel noch zwei Partnerkonferenzen besuchen: die KI2023 sowie die SKILL.

Schon seit 1975 beschäftigt sich ein Fachbereich der GI mit dem Thema Künstliche Intelligenz. Und schon zum 46. Mal trafen sich dessen Mitglieder sowie KI-Interessierte aus Forschung, Wirtschaft und Gesellschaft zur KI2023. Die Konferenz hatte auch in diesem Jahr einige Highlights auf dem Programm, ob Keynotes, Workshops oder Tutorials. Diese drehen sich um aktuelle Debatten und Herausforderungen der KI-Forschung: von Chatbots über Knowledge Representation and Reasoning bis hin zu den ethischen Fragen, die sich rund um KI stellen.

Die SKILL, die Studierendenkonferenz der Gesellschaft für Informatik, richtet sich an alle, die Informatik sowie angrenzende Disziplinen studieren. Im Mittelpunkt der Konferenz stehen studentische Arbeiten und Projekte. Einreichungen von Doktorand*innen sowie Einreichungen mit Ko-Autorenschaften von Professor*innen sind nicht zugelassen. Angenommene Beiträge werden in einem eigenen Tagungsband publiziert. Für herausragende Beiträge werden Preise ausgelobt.

Neben den Preisen der SKILL wurden im Rahmen der INFORMATIK 2023 auch wieder eine Vielzahl an weiteren Preisen verliehen: So wurde der GI-Dissertationspreis zusammen mit der Österreichischen Computerergesellschaft (OCG) und der Schweizer Informatik Gesellschaft (SI) vergeben. Im Rahmen des festlichen Konferenzdiners wurde zudem ein neuer Jahrgang an Fellows und Junior-Fellows benannt sowie die Konrad Zuse

Medaille verliehen, mit der die GI und die Konrad Zuse Stiftung herausragende Persönlichkeiten auszeichnen, die sich in besonderer Weise um die Informatik verdient gemacht haben.

Zum zweiten Mal wurde in diesem Jahr auch der von Helmut und Heide Balzert gestiftete Preis für einen herausragenden Beitrag zur Vermittlung von Inhalten der Informatik und zur Ausbildung/Didaktik der Informatik (IAD) vergeben. Ziel des von GI in Kooperation mit der OCG und der SI verliehenen Preises ist es, die Aufmerksamkeit für die Bedeutung der „digitalen Didaktik“ bei der Vermittlung von komplexen Sachverhalten in der Informatik zu fördern und zu unterstützen.

Sie merken schon: Auf der INFORMATIK kamen auch 2023 viele unterschiedliche Perspektiven und Personen zusammen. Besonders in den Workshops ging es thematisch in die Tiefe. Die knapp 210 schriftlichen Beiträge aus fast 50 Workshops und Tutorials der INFORMATIK 2023 sind in diesem Tagungsband zusammengefasst.

Nun wünschen wir Ihnen viel Spaß bei der Lektüre dieses Tagungsbands!

Berlin im September 2023

Maike Klein, Daniel Krupka, Cornelia Winter und Volker Wohlgemuth

Sponsoren

Wir danken den nachfolgenden Institutionen und Unternehmen herzlich für die Unterstützung des INFORMATIK Festivals 2023.

Platin



Silber



Bronze



Partner



Tagungsleitung

Gesamtleitung:	Volker Wohlgemuth, HTW Berlin Daniel Krupka, Gesellschaft für Informatik
Koordination Hauptprogramm:	Julia Meisner, Gesellschaft für Informatik Frithjof Nagel, Gesellschaft für Informatik
Koordination Workshops:	Markus Durst, Gesellschaft für Informatik
Koordination Tagungsband:	Maike Klein, Gesellschaft für Informatik

Organisationsteam Konferenz

Alexander Scheibe	Gesellschaft für Informatik
Maximilian Weinl	Gesellschaft für Informatik
Markus Durst	Gesellschaft für Informatik

Organisationsteam Tagungsband

Maike Klein	Gesellschaft für Informatik
Karl-Heinz Künkel	Gesellschaft für Informatik
Ludger Porada	Gesellschaft für Informatik
Maithu Bui	Gesellschaft für Informatik
Maja Denisova	Gesellschaft für Informatik
Alexandra Resch	Gesellschaft für Informatik
Cornelia Winter	Gesellschaft für Informatik

Workshop-Chairs

Andreas Abecker	disy Informationssysteme GmbH
Alexa Becker	Universität Leipzig
Hans Brandt-Pook	FH Bielefeld
Carsten Brockman	SAP Deutschland SE & Co. KG
Johannes Buchner	ZUG gGmbH, IPE Berlin, FU Berlin
Manuel Burghardt	Universität Leipzig
Christina B. Class	Ernst-Abbe-Hochschule Jena
Benjamin Çoban	FZI Forschungszentrum Informatik
Jan deMeer	smartspacelab.eu GmbH
Jens Dörpinghau	BIBB Bundesinstitut für Berufsbildung
Daniel Eberz-Eder	Dienstleistungszentr. Ländl. Raum Rheinland-Pfalz
Stephan Fahrenkrog-Petersen	HU Berlin, Weizenbaum-Institut
Matthias Goeken	Hochschule der Deutschen Bundesbank
Gesa Götte	Fraunhofer IFF
Rüdiger Grimm	Universität Koblenz, Fraunhofer SIT
Liane Haak	Hochschule Osnabrück
Christine Hennig	Fraunhofer FOKUS
Jens Henningsen	Fraunhofer IESE
Wilfried Honekamp	Deutsche Hochschule der Polizei
Andreas Jedlitschka	Fraunhofer IESE
Ulrich John	IU Internationale Hochschule
Reinhard Kahle	Universität Tübingen
Frank Fuchs-Kittowski	HTW Berlin
Maike Klein	Gesellschaft für Informatik
Michael Koch	Universität der Bundeswehr München
Sebastian Kreideweiß	CPS GmbH, cmsensus.eu, TYPO3 GmbH
Julian Kunkel	GWDG
Ralf Laue	Westsächsische Hochschule Zwickau
Julia Meisner	Gesellschaft für Informatik
Vera G. Meister	TH Brandenburg
Lydia Müller	Universität Leipzig
Daniel Müssig	Fraunhofer IOSB-AST
Stefan Naumann	HS Trier, Umwelt-Campus Birkenfeld
Frank Neumann	HTW Berlin
David A. Plecher	TU München
Gudrun Schiedermeier	Hochschule Landshut
Johannes Schleiss	Otto-von-Guericke-Universität Magdeburg
Sonja Schimmmler	Fraunhofer FOKUS, TU Berlin
Andreas Schmitz	Universität Koblenz
Thorsten Schoormann	Universität Hildesheim
Annika Selzer	Fraunhofer SIT
Franz Sitzmann	TüftelAkademie

Christine Skupsch
Vera Spitzer
Christoph Sorge
Lena Steinmann
Stefan Voget
Doris Weißels
Maria A. Wimmer
Jochen Wittmann
Volker Wohlgemuth

Schule am Ried, Frankfurt am Main
Universität Koblenz
Universität des Saarlandes
Universität Bremen
Continental Automotive Technologies
Fachhochschule Kiel
Universität Koblenz
HTW Berlin
HTW Berlin

Inhaltsverzeichnis

Young Scientists and early-stage research in Data Science Workshop (YSDS-23)

Gesa Götte

The Effect of Adversarial Debiasing on Model Performance 39

Sandro Hartenstein, Sidney Leroy Johnson

An Approach for a Fast Cost Estimation of Software Projects supported by Sentiment AI Analysis 45

Jasmin Antonia Riebel

User-centered Design Methods in Data-Intensive Software Development Processes: A State-of-the-Art Review 49

Claire Ponciano, Markus Schaffert, Jean-Jacques Ponciano

Deep Learning Datasets Challenges For Semantic Segmentation - A Survey 57

Bildung - Aktuelle Entwicklungen und Perspektiven (an Hochschulen) im Bereich Data Science

Lena Steinmann, Dirk Nowotka, Lea Oberländer, Helen Pfuhl, Heiner Stuckenschmidt, Rolf Drechsler

Workshop: „Aktuelle Entwicklungen und Perspektiven (an Hochschulen) im Bereich Data Science“ 73

Thomas Gautschi, Thomas Gschwend, Lea Oberländer, Simone Ponzetto, Heiner Stuckenschmidt

Der Mannheim Master in Social Data Science 81

Markus Herklotz, Wiebke Weber

BERD Academy: Data Science Kompetenzen für Forschende in Business, Economics and Related Sciences 89

Daniela Merten, Heike Neuroth, Carsten Schneemann, Kathrin Woywod <i>Zertifikatskurs „Forschungsdatenmanagement für Studierende“ der Landesinitiative für Forschungsdatenmanagement in Brandenburg</i>	95
Katharina Weiß, Reinhold Decker <i>Data Science und Forschungsdaten als Treiber für wissenschaftliche Interdisziplinarität</i>	107
Johannes Putzke, Thomas Schimmer, Manuela Richter, Lucas Hamel <i>Towards the Establishment of Data Science Centers at Higher Education Institutions</i>	117
Chris Biemann, Tilo Böhmann, Sebastian Gerling Louis, Martin Semmann <i>Das Konzept des House of Computing and Data Science als Blaupause zur Digitalisierung der Forschung</i>	123
Michael Panitz, Claus Spiecker, Ian Wolff, Tanja von Fransecky, Carina Schiller <i>Von der Bedarfserhebung zum Beratungsangebot</i>	129
Canan Hastik, Frank Lange, Jan-Michael Haugwitz, Peter Pelz <i>Entwicklung einer Wissensbasis für Lehr- und Lernmaterialien zu FAIRem Forschungsdatenmanagement und Data Science im Projekt DALIA (Data Literacy Alliance)</i>	137

Bildung - Wirtschaftsinformatik an Hochschulen angewandter Wissenschaften

Christian Czarnecki, Thomas Barton, Vera G. Meister, Alexander Lübke <i>Wirtschaftsinformatik an Hochschulen Angewandter Wissenschaften (Workshop)</i>	147
Verena Majuntke, Frederik Obermaier <i>Towards Sustainable Software</i>	149

Regina Schuktomow, Hubertus von Tippelskirch, Margit Scholl <i>Informationssicherheit in den Arbeitsalltag nachhaltig integrieren</i>	165
-----------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Bildung - Interdisziplinäres Forschen und Lernen in der Ingenieurinformatik

Rasmus Rettig, Maximilian Weltz <i>Virtuous, multi-level integration of research and education within the project “European Digital Dynamic Mapping” (EDDY)</i>	185
Frank Neumann, Christina Kratsch <i>Agilität, Praxisbezug und Teamarbeit im Softwareentwicklungsprojekt der Ingenieurinformatik</i>	197
Bernhard J. Berger, Goerschwin Fey <i>Informatik für Ingenieure an der Technischen Universität Hamburg</i>	211
Carsten Burmeister, Bonke Möller, Dominik Miller, Andreas Haase, Hendrik Lam <i>Interdisziplinärer Industrie 4.0 Demonstrator (II4D)</i>	223
Luise Stromeyer, Frank Burghardt <i>Informatische Bildung: zwischen Nebenfach und Grundlagenbildung</i>	231
Christina Dieckhoff, Christophe Barlieb, Christian Groth, Thomas Linner, Florian Weininger <i>Erfolgsfaktor Interdisziplinarität: Das Lehrformat Digitalisierungskollegs an Bayerischen Hochschulen</i>	241
Dorian Zwanzig, Anja Kahl, Ute Dietrich <i>SIENA: Sprachmodellbasierte Identifikation und Extraktion von Nutzeranforderungen</i>	253
Pauline Hergersberg, Judith Lippold, Bastian Nahrstedt, ThomasBaar <i>Visualization Support for Contracts in VeriFast</i>	265
Nils T. Siebel <i>On the Use of Process Containers for Running Vulnerable Programs more Securely</i>	277

Léon Niclas Dodrimong, Ute Dietrich <i>Vertikale Integration im Kontext von Industrie 4.0</i>	289
Christian Lehmann, Björn Wudka, Volker Wohlgemuth, Frank Burghardt <i>Entwicklung von Löschdrohnen im Projekt "Feuerwehr der Zukunft für praxisorientierten Ausbildung von Ingenieuren</i>	295
Kai Schauer <i>RoboTwins</i>	299

Bildung - Serious XR - Wissenstransfer mit Augmented und Virtual Reality (SXR)

Tobias Zengerle, David A. Plecher, Salome Flegr, Jochen Kuhn, Martin R. Fischer <i>Teaching Optical Principles in XR</i>	313
Sebastian Walchshäusl, Christian Eichhorn, David A. Plecher, Tim Simecek, Gudrun Klinker, Atsushi Hiyama, Masahiko Inami <i>Generating an Environment for Socializing Between Older Adults in a VR Supermarket</i>	325

Bildung - KI-Bildung - Ein Workshop zu Aus- und Weiterbildung über Künstliche Intelligenz

Anna Faust, Martin Dröge, Carolin Odebrecht <i>Assessment of AI literacy – Development and testing of a customizable set of items</i>	341
Ivan Iovine <i>Integrating Artificial Intelligence and Robotics into Art Curriculum</i>	347

Ricardo Knauer, Raphael Wallsberger, Marvin Grimm, Marian Bookhahn, Frank Neumann, Stephan Matzka, Erik Rodner <i>Foundations for applied artificial intelligence: enabling and supporting AI teaching and research</i>	353
Florian Weber, Thiemo Wambsganss, Matthias Söllner <i>Unterstützung des juristischen Schreibens durch ein KI- basiertes Schreibsystem</i>	357
André Selmanagić, Juliane Ahlborn <i>Interdisciplinary Perspectives on the Use of Competency-Based Learning and AI in Personalized Learning Environments</i>	363
Franz Walgenbach, Lukas Wachter <i>Entscheidungsbäume in der Mittelstufe</i>	369
Caroline Berger-Konen, Jessica Felgentreu, Bettina Ötvös <i>Der KI-Campus 2.0 an der FernUniversität in Hagen – lernen, vernetzen, entdecken!</i>	371
Ingo Siegert, Stefan Hillmann, Philine Thalia Görzig, Matthias Busch, Jan Nehring, Xenia Klinge <i>Die Chatbot-Challenge – Spielend mit KI von der Idee zum Dialogsystem</i>	377
Sebastian Lang, Ingo Siegert, Viktor Artiushenko, Johannes Schleiss <i>AI Engineering als interdisziplinäres Einführungsmodul zwischen Künstlicher Intelligenz und Ingenieurwesen</i>	381
Martin Hänel, Matthias Söllner <i>Harnessing Conversational AI in Higher Education: The Education Buddy - An Innovative Intervention for International Students Facing Study Difficulties</i>	385
Susanne Podworny, Yannik Fleischer, Rolf Biehler <i>Entscheidungsbäume erstellen ab Klasse 6</i>	389
Andrea Beyer, Konstantin Schulz <i>DAIalos: Forschen und Lernen zugleich?</i>	391
Theresa Kruse, Marc Berges, Jonas Betzendahl, Michael Kohlhase, Dominic Lohr, Dennis Müller <i>Learning with ALeA: Tailored experiences through annotated course material</i>	395

Sven Baumer, Franz Jetzinger, Tilman Michaeli <i>Künstliche Intelligenz im verpflichtenden Informatikunterricht: Ein skalierbares Fortbildungsangebot für Informatiklehrkräfte</i>	399
Gia Minh Vo, Nils Pancratz <i>Künstliche Intelligenz in der universitären Lehrerinnen- und Lehrerbildungsphase</i>	403
Marc Egloffstein, Kristina Kögler <i>Design Review von offenen Online-Kursen zum Thema Künstliche Intelligenz</i>	409
Silvia Joachim, Martin Hennecke <i>Enaktive Bestimmung der Hyperparameter beim Entscheidungsbaum- und k-nächste-Nachbarn-Algorithmus</i>	415
Eric Marx, Nadine Bergner <i>Einstieg in den ML-Workflow durch Entwicklung einer intelligenten Museumsapp</i>	419
Katharina Schüller, Florian Rampelt, Henning Koch, Johannes Schleiss <i>Better ready than just aware: Data and AI Literacy as an enabler for informed decision making in the data age</i>	425
David Weigert, Ulrike Scorna, Fabian Behrendt <i>KI-Kompetenzen in ingenieurwissenschaftlichen Studiengängen entwickeln und fördern</i>	431
Maria Freese, Birgit Zürn, Julia Arlingaus <i>Artificial Intelligence meets Serious Gaming</i>	437
Lisa Hopf, Maria Hinsche, Christina B. Class <i>Integration ethischer Kompetenzen in einen Kurs zu Grundlagen Künstlicher Intelligenz</i>	439
Selbsteinschätzung zu KI-Kompetenzen und -Funktionen <i>Erhebung einer Selbsteinschätzung zu KI-Kompetenzen sowie zu genutzten KI-Funktionen und Fortbildungsinteressen von Lehrkräften</i>	445

Stefan Seegerer, Michaela Gläser-Zikuda, Florian Hofmann, Ralf Romeike <i>KI-Technologieentwicklung und Bildung als miteinander in Beziehung stehende Aufgaben für die Lehrerbildung</i>	447
Yannik Fleischer, Sven Hüsing, Rolf Biehler, Lukas Höper, Susanne Podworny, Carsten Schulte <i>Ein Projektkurs Data Science und Künstliche Intelligenz für die Sekundarstufe II</i>	453
Malte Prieß, Kerstin Prechel <i>KI lehren, KI lernen an der DSHS – Potentiale für die Bildungslandschaft</i>	457
Stefanie Krause, Simon Adler, Johannes Bühl, René Schenkendorf, Kerstin Schneider, Frieder Stolzenburg, Fabian Transchel <i>Entwicklung interdisziplinärer Module in der Hochschulbildung</i>	461

Cybersecurity Privatsphäre - 4. Interdisciplinary Privacy Security at Large Workshop

Annika Selzer, Ingo J. Timm <i>Digitalisierung: Chancen nutzen, Risiken vermeiden</i>	467
Annika Selzer, Ingo J. Timm <i>Post Pandemic Follow-Up</i>	469
Tanya Gärtner <i>Towards a Taxonomy of Cyber Defence in International Law</i>	477
Jessica Kriegel, Alina Boll <i>Der geplante US Active Cyber Defense Certainty Act</i>	489
Ines Geissler <i>Datenschutzrechtliche Rollen in Metaversen und im virtuellen Weiterleben</i>	497
Ines Geissler <i>Leben in Metaversen und im Virtual Afterlife</i>	511

K. Valerie Carl	
<i>Data Privacy and Security in the Context of Corporate Digital Responsibility: A Scoping Review</i>	523
Suvi Lehtosalo, Daniel W. Woods	
<i>Deceptive patterns in consent dialogs on children’s websites</i>	537
Lucas Lange, Borislav Degenkolb, Erhard Rahm	
<i>Privacy-Preserving Stress Detection Using Smartwatch Health Data . . .</i>	549
Marian Eleks, Jonas Rebstadt, Henrik Kortum, Oliver Thomas	
<i>Privacy Aware Processing</i>	561

Cybersecurity Privatsphäre - 3. International Workshop on Digital Forensics / IWDF3

Florian Heinke, Marie-Luise Heuschkel, Dirk Labudde	
<i>Analysing Distributions of Feature Similarities in the Context of Digital Anthropometric Pattern Matching Probability</i>	573
Stefan Kellermann, Dirk Labudde	
<i>Vorbetrachtung zur Entwicklung einer Analysesoftware zur Auswertung von videografierten Beschuldigtenvernehmungen im Strafermittlungsverfahren</i>	583
Dominic Helfer, Felix Rothe, Ronny Bodach	
<i>Post-mortem path correlation based on the NT Object Manager in Windows 1x systems</i>	597
Eileen Köhler, Daniel Spiekermann	
<i>Geschwindigkeitsanalyse bei der Sicherung von verschiedenen Speichermedien</i>	607
Elena Pistorius, Sabine Richter, Dirk Labudde	
<i>The digital skeleton in modern video analysis - inter- and intraspecific comparison of individual rigs</i>	611
Dirk Volkmann, Sabine Schildein, Roman Povalej	
<i>Digital Twin – Potenziale für die Polizei</i>	623

Mirijam Steinert, Sabrina Herchel, Dirk Labudde <i>Möglichkeiten, Grenzen und Konzepte für eine digitalbasierte Täterbewertung im Bereich der §§ 174 – 184</i>	635
Martin Morgenstern, Wilfried Honekamp <i>Requirements for a public digital forensics cloud</i>	649
Johannes Fährdrich, Lars Mechler <i>Discussion of Basic Concepts of digital Forensics on the example of Volatility and Manipulability</i>	659

Cybersecurity Privatsphäre - Recht und Technik. Datenschutz im Diskurs (RuT2023)

Rüdiger Grimm, Gerrit Hornung, Christoph Sorge, Indra Spiecker genannt Döhmann <i>Recht und Technik – Datenschutz im Diskurs</i>	675
Matthias Kohn, Merle Freye, Mehrdad Bahrini, Alexander Herbst <i>Gesundheits-Apps auf dem Prüfstand –Überprüfung der Angaben in Datenschutzerklärungen zur Datenweitergabe</i>	677
Kai Julian Kemmler <i>Verantwortlichkeit im Unternehmen und Verschuldensmaßstab im Rahmen des Art. 83 DSGVO</i>	689
Annika Selzer <i>Umbruch im Datenschutz</i>	705
Jan-Philipp Stroscher, Hendrik Link <i>Verarbeitung personenbezogener Daten in KI-Reallaboren nach dem KI-VO-E als Herausforderung für den datenschutzrechtlichen Zweckbindungsgrundsatz</i>	715

Kritisches Denken - Konzeptionelle Herausforderungen für die KI - Zukünfte

Klaus Mainzer

Zukunft durch hybride IT- und KI-Innovationen 735

Reinhard Kahle

Zukunft der KI: Verantwortung und Vertrauen 749

Kritisches Denken - Zukunftswerkstatt Informatik Gesellschaft (IUG 2023)

Patrick Liedtke, Hans-Knud Arndt

Materialität und Immaterialität – Begriffsbestimmungen und Herausforderungen für die Informatik 759

Andrea Kienle, Gabriele Kunau

Aufnahme läuft! Partizipative, sozio-technische Gestaltung digitaler Kommunikationssysteme 765

Nils Pancratz

„Ich denke, dass ethische Probleme nicht in der Entwicklung, sondern in der Anwendung auftreten“ 771

Paula Bräuer, Lucas Schwarz, Isabella Peters

Transdisziplinäre Entwicklung eines WebGIS für die Atommüll-Endlagerstandortsuche in Deutschland 777

Kultur Design - Digital Cultures Cultural Analytics (InfDH 2023)

Meinard Müller, Frank Zalkow

FMP Notebooks 785

Thomas Schmidt, Fabian Schiller, Mathias Götz, Christian Wolff <i>A Corpus of Memes from Reddit: Acquisition, Preparation and First Case Studies</i>	795
Christof Weiß, Meinard Müller, Stephanie Klauk, Rainer Kleinertz <i>Neue Wege für die Musikforschung</i>	805
Andrew McLeod <i>An Analysis of Automatically Generated Music</i>	815
Michael Achmann, Christian Wolff <i>Computergestützte Bildtypenanalyse durch Zero-Shot Klassifikation mit CLIP</i>	821
Bernhard Liebl, Manuel Burghardt <i>Designing a Prototype for Visual Exploration of Narrative Patterns in News Videos</i>	831
Stefan Balke, Adrian Maiworm <i>Digitalisierung in der Amateurmusik</i>	841
Torsten Roeder <i>Diskettenmagazine als Frühes Kulturerbe: Erschließung und Re-Digitalisierung</i>	851
Moritz Feichtinger <i>Annotation, Simulation und Analyse eines historischen Datenbanksystems</i>	863
Peter Meier, Simon Schwär, Gerhard Krump, Meinard Müller <i>Evaluating Real-Time Pitch Estimation Algorithms for Creative Music Game Interaction</i>	873
Stephanie Klauk, Rainer Kleinertz, Pascal Schmolenzky, Christof Weiß, Meinard Müller <i>Perspektiven computergestützter harmonischer Analyse: Beethovens op. 14 Nr. 1 als Gegenstand gattungsübergreifender Korpusanalyse</i>	883

Kultur Design - Designing alternative future home stories

Benedikt Haupt, Christian Pentzold, Alexa Becker, Arne Berger
Designing alternative future home stories 893

Öffentliche Infrastruktur - Research Data Infrastructures for Data Science and AI (RDI4DataScience)

Sonja Schimmler, Christine Hennig
NFDI4DS at a Glance 901

Anna-Lena Lorenz, Maria Christoforaki, Christine Hennig, Angelie Kraft, Stephanie von Maltzan, Sonja Schimmler
Community and Training in NFDI4DS 905

Saurav Karmakar, Matthäus Zloch, Fidan Limani, Benjamin Zapilko, Sharmila Upadhyaya, Jennifer D’Souza,, Leyla Jael Castro, Georg Rehm, Marcel R. Ackermann, Harald Sack, Zeyd Boukhers, Sonja Schimmler, Danilo Dessí, Peter Mutschke, Stefan Dietze
Research Knowledge Graphs in NFDI4DS 909

Sonja Schimmler, Bianca Wentzel, Arnim Bleier, StefanDietze, Saurav Karmakar, Peter Mutschke, Angelie Kraft, Tilahun A. Taffa, Ricardo Usbeck, Zeyd Boukhers, Sören Auer, Leyla Jael Castro, Marcel R. Ackermann, Thomas Neumuth, Daniel Schneider, Ziawasch Abedjan, Atif Latif, Fidan Limani, Raia Abu Ahmad, Georg Rehm, Sima Attar Khorasani, Matthias Lieber
NFDI4DS Infrastructure and Services 919

Ekaterina Borisova, Raia Abu Ahmad, Georg Rehm, Ricardo Usbeck, Jennifer D’Souza, Markus Stocker, Sören Auer, Judith Gilsbach, Anastasia Wolschewski, Johannes Keller, Daniel Schneider, Thomas Neumuth, Sonja Schimmler
NFDI4DS Transfer and Application 925

**Raia Abu Ahmad, Ekaterina Borisova, Georg Rehm, Stefan Dietze,
Saurav Karmakar, Wolfgang Otto, Jennifer D’Souza, Fidan Limani,
Ricardo Usbeck**
NFDI4DS Shared Tasks 931

**Amanda Wein, Jan Reinkensmeier, Anke Weidlich, Johan Lilliestam,
Veit Hagenmeyer, Sebastian Lehnhoff**
NFDI4Energy Task Area 4: FAIR Data for Energy System Research 937

Öffentliche Infrastruktur - IT-Governance und Strategisches Informationsmanagement (ITG- SIM)

Daria Goscinska, Till J. Winkler
*Operationalizing Enterprise Service Management Capability – the OATIP
Model* 947

Mona Möstl, Patrick Stoll
*Ein vereinfachter Ansatz zur Entwicklung einer Digitalstrategie in
Unternehmen* 957

Lena Schlosser, Helmut Beckmann
Von Data Fabrics zu Smart Data Fabrics – eine Literaturanalyse 967

Lisa Otten, Matthias Goeken
*Wertbeitrag von Big Data, KI und Maschinellem Lernen. Bezugsrahmen
und Demonstration für Prüfungs- und Aufsichtsprozesse* 977

Betül Demirözer, Helmut Beckmann
Agile Strategieentwicklung 991

Vanessa Wegner, Helmut Beckmann
Digital Business Technology Platforms 1003

Öffentliche Infrastruktur - IT-Kompetenzen für die digitale Verwaltung der Zukunft

- Michael Koddebusch, Paul Brützke, Holger Koelmann, Jörg Becker**
The Public Official's Selection Parameters for E-Competence Continuous Education 1017
- Uta Katja Schlichte, Lisa Schröder-Puls, Carol Marmor-Drews, Maryna Butara, Sabine Schlick, Martin Kohls**
Vernetzung der Behörden-IT mit Lehre, Wissenschaft und Forschung 1029
- Leila Feddou, Maximilian Raupach, Felicitas Löffler, Samira Babalou, Jonas Hoyer, Marianne Mauch, Birgitta König-Ries**
On which legal regulations is a public service based? Fostering transparency in public administration by using knowledge graphs 1035
- Marc Egloffstein, Philipp Kuscher, Dirk Ifenthaler**
Modularisierung von digitalen Weiterbildungsangeboten für die öffentliche Verwaltung 1041
- Nassrin Hajinejad, Alinka Rother, Jana Plomin, Dorian Grosch**
Wegbereiter: Konzeptuelle Arbeitsinstrumente zur Unterstützung interdisziplinären Denkens in der digitalen Verwaltungspraxis 1049
- Tim Pidun, Silko Grellmann, Ludwig Schönthier, Pascal Thielemann, Stefan Handke**
(K)ein funktionierender Markt? 1057

Öffentliche Infrastruktur - Interoperabilität und Standardisierung der digitalen öffentlichen Beschaffung

- Cedric Pauken, Andreas Schmitz, Maria A. Wimmer**
Nutzung von elektronischen Katalogformaten in öffentlichen Beschaffungsverfahren 1071

Matthias Hartmann, Andreas Schmitz, Maria A. Wimmer
Datenmanagementkonzepte und -frameworks im öffentlichen Beschaffungswesen 1085

Gerhard Klassen, Raphael Palombo, Luca T. Bauer, Björn Niehaves
Overcoming Inefficiency in Public Procurement: An OpenData Approach 1097

Öffentliche Infrastruktur - CMS-Verteilung im öffentlichen Bereich in Deutschland - Erkenntnisse und Vorschläge für Konsolidierung

Julian Fastnacht, Sebastian Kreideweiß, Phoenix Lußky
CMS-Verteilung im öffentlichen Bereich in Deutschland - Erkenntnisse und Vorschläge für Konsolidierung 1105

GI Quantum Computing Workshop

Elisabeth Lobe
quark: QUantum Application Reformulation Kernel 1115

Marika Federer, Steve Lenk, Daniel Müssig, Markus Wappler, Jörg Lässig
Constrained Grover Adaptive Search for Optimization of the Bidirectional EV Charging Problem 1121

Domenik Eichhorn, Maximilian Schweikart, Nick Poser, Tobias Osborne, Ina Schaefer
Providing Quantum Readiness: The Vision of the ProvideQ Toolbox . . . 1129

Volker Reers
Towards Performance Benchmarking for Quantum Reinforcement Learning 1135

Ökologische Nachhaltigkeit - Umweltinformatik zur Gestaltung einer nachhaltigen Zukunft

Stefanie Nina Steinbichl, Janosch Grellner, Nico Hannoun, Frank Fuchs-Kittowski <i>SeaSaver. Serious Game für die Bewusstseinsbildung bei Kindern zur Verschmutzung von Ozeanen mit Plastikmüll</i>	1149
Lucas Tiedemann, Jochen Wittmann, Michaela Zoll <i>Konzeption einer webbasierten Anwendung zur Unterstützung und Akzeptanzsteigerung des Photovoltaikausbaus in Betrieben</i>	1169
Malina Bergmann, Sebastian Rauch, Josephine Brömme, Jochen Wittmann <i>Analyse und Entwicklung eines Prototyps zur Optimierung des Verbrauchs von Ökostrom in Privathaushalten</i>	1181
Dennis Junger, Volker Wohlgemuth <i>Design and implementation of a lecture for teaching current Green Coding approaches and practices at the HTW Berlin</i>	1197
Jennifer McClelland, Tanja Riedel, Florian Beyer, Heike Gerighausen, Burkhard Golla <i>State of the Art Open Access Remote Sensing with ESA Sentinel 1 SAR Data</i>	1207
Eva Wein, Fabian Bressel, Julian Berndt, Hellen Siewert, Rosemarie Bähne, René Krüger, Frank Fuchs-Kittowski <i>Informationsbedarf von Bürger*innen bezüglich Niedrigwasser, Trockenheit und Dürre - Ergebnisse einer Umfrage</i>	1221

Ökologische Nachhaltigkeit - 11. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel (UINW 2023)

Michael Pleger, Ina Schiering <i>Digital Transformation in Forestry - Stakeholders and Data Collection in German Forests</i>	1245
--------------------------------------------------------------------------------------------------------------------------------------------------	------

Dusan Dokic, Hannah Stein, Sabine Janzen, Wolfgang Maaß <i>Towards Energy-Efficient Large-Scale Artificial Intelligence for Sustainable Data Centers</i>	1255
Florian Fehring, Grit Behrens, Jannis Thiel <i>Smartmonitoring – eine webbasierte Plattform für das Monitoring von Umweltdaten aus verschiedensten Schnittstellen und Sensoren zur Nutzung für studentische Projekte und Forschungsprototypen</i>	1267
Stefan Naumann, Achim Guldner, Sebastian Weber, Max Westing <i>Was weiß ChatGPT über Nachhaltige Software-Entwicklung und Green Coding? Erste Tests und Bewertungen</i>	1277
Dennis Junger, Max Westing, Christopher Freitag, Achim Guldner, Konstantin Mittelbach, Sebastian Weber, Stefan Naumann, Volker Wohlgemuth <i>Potentials of Green Coding - Findings and Recommendations for Industry, Education and Science</i>	1289

Ökologische Nachhaltigkeit - Nachhaltige Wertschöpfungssysteme (NaWerSys) III

Thorsten Schoormann, Friedemann Kammler, Paul Christoph Gembarski, Simon Hagen <i>3. Workshop: Nachhaltige Wertschöpfungssysteme</i>	1303
Stephanie Winkelmann, Julia Schweihoff, Ilka Jussen, Frederik Möller <i>Turning Old into New – The Lane Change to a Circular Economy in the Automotive Industry</i>	1309
Christoph Hoppe, Robert Schmelzer, Frederik Möller, Thorsten Schoormann <i>Data Spaces as Enablers for Sustainability</i>	1323
Stefan Neubig, Dominik Rebholz, Andreas Hein, Robert Kelle, Helmut Kremer <i>To Graph or Not to Graph: The Missing Pieces for Knowledge Graphs in Sustainable Tourism</i>	1335

Marco Di Maria, David Walter¹ and Ralf Knackstedt
*Promoting Sustainable Ecosystems through Interorganizational
Unlearning – A Call for Research* 1351

Christoph Heinbach, Henning Gössling, Oliver Thomas
The Hitchhiker's Guide to Urban Spaces 1359

Ökologische Nachhaltigkeit - Zukunft nachhaltig gestalten durch digitalisierte Wertschöpfungs- prozesse (DigiWe)

**Arne Seipolt, Ralf Buschermöhle, Maximilian Höfinghoff,
Goy-Hinrich Korn, Marcel Schumacher**
*Technology Readiness Levels of Reinforcement Learning methods for
simulation-based production scheduling* 1375

Justus Leskow, Steffen Greiser, Goy-Hinrich Korn
Sensor-based modelling of carbon dioxide emissions for trailer traffic . . . 1391

**Marcel Schumacher, Ralf Buschermöhle, Liane Haak, Max
Höfinghoff, Arne Seipolt, Goy-Hinrich Korn**
*Enhancing Digital Twins for Production through Process Mining
Techniques: A Literature Review* 1399

Marco Barenkamp, Dany Moualeu-Ngangue
Process Mining und Künstliche Intelligenz in der Beschaffung 1407

**Maximilian Höfinghoff, Ralf Buschermöhle, Goy-Hinrich Korn,
Marcel Schumacher, Arne Seipolt**
*Assessing the performance of Neural Networks in Recognizing Manual
Labor Actions in a Production Environment* 1421

Stefan Gudenkauf, Javier Franke, Janek Behrens
*Features of Event-Driven Message Queuing Architectures in
Manufacturing: A Reference Model for Comparison* 1435

Ökologische Nachhaltigkeit - KI und Kreislaufwirtschaft

Anna Hoffmann

Zukunftsstrategien partizipativ entwickeln 1455

Ökologische Nachhaltigkeit - KIU-2023

Andreas Abecker, Julian Bruns, Stefan Naumann

4. Workshop Künstliche Intelligenz in der Umweltinformatik 1473

Robert Dehghan, Sebastian Schmidt, Julian Schwierzy

3D Printing and Sustainability: A Web Analysis Approach 1483

Jonathan Vogl, Désirée Hilbring, Divas Karimanzira

Nutzung offener Standards für die Integration von Pegel- und Abflussdaten zur Verbesserung der Frühwarnung bei Sturzfluten 1493

Sebastian Weber, Achim Guldner, Lejla Begic Fazlic, Guido Dartmann, Stefan Naumann

Sustainability in Artificial Intelligence - Towards a Green AI Reference Model 1503

Alexander Münzberg, Christian Troost, Nils Reinosch, Daniel Martini, Liv Seuring, Alexander Niehus, Rajiv Srivastava, Thilo Streck, Thomas Berger, Ansgar Bernardi

Landwirtschaftliche Ertragsvorhersage im Kontext begrenzter realer Trainingsdatensätze: ein Transfer-Learning-Ansatz unter Verwendung tiefer neuronaler Netze 1515

Ökologische Nachhaltigkeit - Resilient Smart Farming Lab (RSFLab) 2.0 - Edge Computing mit der HofBox erleben

Franz Kuntke, Daniel Eberz-Eder, Matthias Trapp, Christian Reuter <i>RSF-Lab'23: Konzepte und Anwendungen zur resilienten digitalen Landwirtschaft</i>	1529
Paul Schulze, Frank Fuchs-Kittowski, Tim Hafemeister, Mario Berndl, Christian Simmet, Martin Schulze <i>Mobiles Assistenzsystem zum Monitoring und Bewertung des Zustandes von Ebersperma</i>	1535
Martin Weis, Sebastian Bökle, Christian Bauer <i>The Hofbox as a decentralised solution for agricultural operations</i>	1551
Jonas Franken, Marco Zivkovic, Nadja Thiessen, Jens Ivo Engels, Christian Reuter <i>Das Netz hat Geschichte</i>	1563
Denis Orlov, Franz Kuntke, Christian Reuter <i>Optimierte Messenger-Applikation zur Notfallkommunikation via LoRaWAN-DTN</i>	1575

Ökologische Nachhaltigkeit - Kolloquium Landwirtschaft der Zukunft - Ist KI ein wesentlicher Schlüssel zur nachhaltigeren Landwirtschaft?

Jens Dede, David Wewetzer, Anna Förster <i>AI in the Wild: Challenges of Remote Deployments</i>	1583
Ronald Böck, Siddarth Venkateswaran, Thi Nguyen, Dominik Durner <i>Textual Descriptions Used for Classification of Oaked vs Unoaked Wines</i>	1591

Armin Wolf, Silke Cuno <i>Emission-Reducing Vehicle Routing in Food Logistics</i>	1599
Vanshika Bawa, Ibrahim Baroud, Stefan Schaffer <i>“HalloBzar”: A German chatbot for accessing the regional digital marketplace</i>	1607
Bonito Thielert, Patrick Menz, Sebastian Warnemünde, Katharina Holstein, Lauritz Klein, David Kiliyas, Miriam Runne, Wolfgang Jaraus, Uwe Knauer <i>Drohnenbasiertes Verfahren zur Detektion geschädigter Obstbäume in Obstbaumplantagen</i>	1615
F. Thürkow, C. Lorenz, J. Ramstetter, I. Hoppe, M. Haase <i>KI-basierte Detektion von Feldhamsterbauen auf landwirtschaftlichen Nutzflächen mittels multi-sensoraler UAS-Daten</i>	1623
Philipp Flierl, Alexander Zimmermann, Michael Niedermeier, Faryal Noori, Erich Fuchs, Anton Schmailzl <i>Vertical-Farming-System für die Bilddatengenerierung zur KI-gestützten Identifikation des Wachstumszentrums von Beikräutern</i>	1633
Ahmad Kadi, Daniel Martini, Ansgar Bernardi, Daniel Eberz-Eder <i>Mobile semantische Dokumentation als Basis für KI-gestützte Beratungsdienste: Das GeoBox Buchungsjournal</i>	1639
Maksim Kukushkin, Matthias Enders, Reinhard Kaschuba, Martin Bogdan, Thomas Schmid <i>Canola seed or not? Autoencoder-based Anomaly Detection in Agricultural Seed Production</i>	1645
Mortesa Hussaini, Anthony Stein <i>Federated Learning in Agriculture: Potential and Challenges</i>	1653

Ökologische Nachhaltigkeit - Workshop on Systems to Support Renaturation Projects 2023 (SyRePro23)

Kim Bürgl, Lydia Müller

Semi-automatic extraction of metadata from old geological maps 1665

Lydia Müller, Daniel Wiegrefe

Nachnutzungskonzept für Daten in der Altlastenverwaltung 1675

Kenn Kaubukowski, Baldwin Nsonga, Vanessa Kretzschmar, Yves Annanias, Daniel Wiegrefe, Mario Hlawitschka

Classification of Large and Heterogeneous LiDAR Data Sets 1683

Christian Bender, Mario Hesse

Geodaten in der öffentlichen Verwaltung 1693

Wirtschaft, Management Industrie - Joint Workshop IntDig 2023 MOC 2023; Intelligente Digitalisierung, (KI-basiertes) Management und Optimierung komplexer Systeme

Ulrich John

Intelligente Digitalisierung 1709

Joel Lehmann, Andreas Lober, Tim Häußermann, Alessa Rache, Hartwig Baumgärtel, Julian Reichwald

Domain-agnostic Intelligent Digital Twins: Merging of Application-near Knowledge Representations with the Proactive Internet of Digital Twins (IoDT) 1725

Sven Löffler, Ilja Becker, Petra Hofstedt, André Nitze, Silvia Hennig, Alexander Klinge

Planung des Ländlichen On-Demand-Verkehr - Probleme, Analyse und Algorithmen 1739

Reinhold von Schwerin, Alexander Hafner <i>DASU – Transferzentrum für Digitalisierung, Analytics Data Science Ulm</i>	1751
Klaus-Jürgen Meier <i>Die Hürden der Digitalisierung in KMUs erfolgreich nehmen</i>	1763
Lukas-Walter Thié, Felix Krieger, Burkhardt Funk <i>Extraction of Information from Invoices – Challenges in the Extraction Pipeline</i>	1777
Florian Allwein, Maik H. Wagner <i>Kriterien und Framework für intelligente Digitalisierung aus soziotechnischer Perspektive – zur Diskussion</i>	1793
Maik Günther, Guido Bayard <i>Wärmewende in einer Großstadt</i>	1801
Zurana Mehrin Ruhi, Hannah Stein, Wolfgang Maaß <i>A Proposal for Physics-Informed Quantum Graph Neural Networks for Simulating Laser Cutting Processes</i>	1809

Wirtschaft, Management Industrie - Künstliche Intelligenz für kleine und mittlere Unternehmen (KI-KMU 2023)

Frederik Simon Bäumer, Hans Brandt-Pook, Falk Maoro, Sergej Schultenkämper, Benjamin Stecker <i>Von der Theorie zur Praxis: Erfahrungen bei der akademischen Begleitung von KI-Projekten in KMUs</i>	1815
Bastian Sirvend, Alexander Stahlkopf, Frederik Simon Bäumer <i>Synergien zwischen KI und menschlicher Expertise: Online-Sichtbarkeitsoptimierung durch Text-Transformation</i>	1827

Lymperis Perakis, Julian Balling, Frank Binder, Gerhard Heyer, Franz Kreupl
Classifying figures and illustrations in electronics datasheets: A comparative evaluation of recent computer vision models on a custom collection of 4000 technical documents 1833

Wirtschaft, Management Industrie - (Agiles) Enterprise Architecture Management in Forschung und Praxis

Gergana Vladova, André Ullrich, Eldar Sultanow, Marinho Tobolla, Sebastian Sebrak, Christian Czarnecki, Carsten Brockmann
Visual analytics for knowledge management 1851

Carsten Brockmann
Process view on interface development of ERP and legacy systems 1871

Wirtschaft, Management Industrie - Workshop zum Stand, den Herausforderungen und Impulsen des Geschäftsprozessmanagements (ZuGPM 2023)

Ralf Laue
ChatGPT – Modelliere mir meinen Geschäftsprozess! 1877

Felix Holz, Dennis Vogel, Michael Fellmann
Specification and Application of a Domain Specific Modeling Language for Social Services 1879

Leonard Nake, Stephan Kuehnel, Laura Bauer, Stefan Sackmann
Towards Identifying GDPR-Critical Tasks in Textual Business Process Descriptions 1895

Thomas Bauer

Verhalten und Ausführungssemantik optionaler Kanten in Geschäftsprozessen 1909

Peter A. François, Ralf Plattfaut

The Reuse of Business Process Automation Artefacts 1923

Wirtschaft, Management Industrie - Skills and Qualifications for a Digitalized Future (SKILLS'23)

Felix Derksen, Jens Dörpinghaus

Digitalization and Sustainability in German Continuing Education 1945

Matthias Meister

Ethical Challenges of Digitalization in Education 1955

Gian-Luca Gücük, Dejan Simic, Stephan Leible, Tom Lewandowski, Emir Kučević

Enhancing educational insights: A real-time data analytics stack for project-based learning 1963

Ali Vahdatnia, Danoosh Peachkah

Monitoring of Digitization and Sustainability on Twitter 1973

Michael Tiemann, Stefan Udelhofen, Lisa Fournier

What social media can tell us about essential occupations 1983

Wirtschaft, Management Industrie - 8th Industrial Automation and Control Systems Standardization Workshop (IACS 2023)

Erkin Kirdan, Josef Schindler, Karl Waedt

Optimizing OPC UA Deployments on Node.js through Advanced Logging Techniques 1995

Jan Sudeikat, Michael Köhler-Bußmeier <i>Controlled Run-Time Adaptivity in Industrial Agent Systems - Challenges and Research Prospects</i>	2003
Christoph Legat, Uwe Seebacher, Dominik Brunner <i>Reasoning about Causal Effects of Regulation and Legislation on Interconnected Markets</i>	2015
Christoph Legat, Marvin Böll <i>Towards a Standardization Ecosystem for Industry 4.0</i>	2025
Natasha Edeh, Robert Altschaffel, Karl Waedt <i>Generation of Plausible Synthetic Data for Stego-Malware Detection for Inter-zone IACS Protocols</i>	2041
Romarick Yatagha, Karl Waedt, Josef Schindler, Erkin Kirdan <i>Security challenges and best practices for resilient IIoT Networks: Network Segmentation</i>	2051
Jan deMeer <i>Data Spaces as the Distributed Communication Means for Industrial Automation and Control Systems</i>	2071
Wael Alsabbagh, Chaerin Kim, Peter Langendörfer <i>No Attacks Are Available: Securing the OpenPLC and Related Systems</i> . .	2085

Autorenverzeichnis

Young Scientists and early-stage
research in Data Science Workshop
(YSDS-23)

The Effect of Adversarial Debiasing on Model Performance

Gesa Götte ¹

Abstract: This paper explores the effect of adversarial debiasing on the performance of machine learning models. As concerns about fairness in algorithmic decision-making grow, techniques for detecting and mitigating biases in ML models have been developed. However, there is a trade-off between fairness and model performance. This study investigates the impact of using adversarial debiasing on model performance in different scenarios of potential sampling biases and target distributions. Simulated data with varying structural and sampling parameters is used to evaluate the models' performance. The results show that while adversarial debiasing can lead to significant improvements in certain scenarios, it can also result in impairments or no significant difference in performance compared to the normal models.

Keywords: Fair AI; Debiasing

1 Introduction

Since machine learning (ML) models infiltrate sensitive areas in our everyday life, concerns about their fairness have emerged. Discrimination by ML models, as revealed in cases like the biased crime recidivism algorithm in the United States [La16], has sparked discussions on preserving constitutional rights and fairness in algorithmic decision-making. In order to detect or mitigate such unfair biases, different fairness definitions have been employed, such as "demographic parity" and "equality of odds" [Be17] and techniques for data manipulation, algorithm adaptation, and model adaptation have been developed. While these techniques aim to achieve fairness, there is a trade-off between this fairness and the model's performance, measured by its accuracy on a validation set. Biased training data, resulting from non-representative sampling or mislabeling, can lead to skewed models [SG19] [GOW20]. Model evaluations on validation sets, derived from the same biased data, may not accurately reflect real-world performance. The assumed trade-off for using debiasing techniques may occur on the validation data but not in application for certain biases in the data. The presented work aims to examine, in a toy scenario, how the effect on the models' performance by using a debiasing technique for training varies in different constellations of potential sampling biases and target distributions.

¹ Fraunhofer Institute for Factory Operation and Automation IFF, Sandtorstraße 22, 39104 Magdeburg, Germany
gesa.goette@iff.fraunhofer.de

2 Methods

This experiment examines the effect of the debiasing method called adversarial debiasing [ZLM18] on a model’s performance. It aims to determine if there are data and sampling structures where adversarial debiasing improves the model’s performance on the application data. The analysis focuses on overall performance (Acc) and subgroup/label performances ($TPR_0, TPR_1, TNR_0, TNR_1$). Simulated data is used, varying structural and sampling parameters to simulate different populations and sampling biases. The approach is explorative in order to understand the impact of the implemented biases on performance metrics.

2.1 Data Simulation

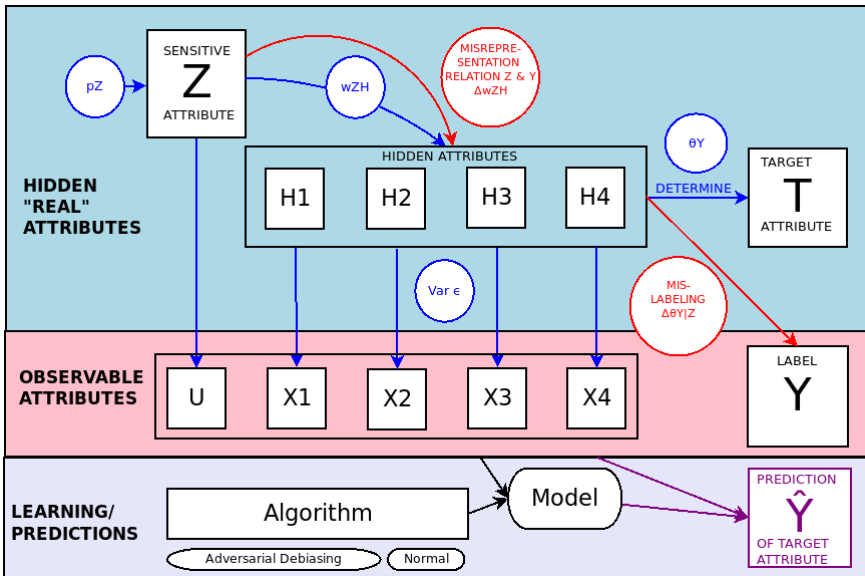


Fig. 1: Data Generation for the Experiment

The data and sampling structure for the toy data are divided into hidden variables (blue background), observable/ measurable variables (red background), and the learning and prediction on this data (violet background). The circles represent the variation factors of the experiment. Blue circles: Application Population (Data Structure) factors. Red circles: Sampling factors.

The simulated data contains a binary target attribute T with $T = 1$ if $\sum H_i \geq \theta_T$, and $T = 0$ otherwise. The values for θ_T are set to 0 (common positives), 1 (few positives) or 2 (rare positives). \mathbf{H} are the subjects’ properties, i.e. different abilities, that determine the target attribute and they are potentially dependent on a binary grouping variable Z (e.g. the subjects’ gender) with $P(Z = 1) = p_Z$, where the subjects with $Z = 0$ (“Z0”) represent the (potentially) disadvantaged group. The two groups are either equally represented ($p_Z = 0.5$) or the Z0 are highly underrepresented ($p_Z = 0.95$) in the data. The summed mean of the

subjects' properties H is either equal for both groups ($|w_{ZH}|=0$), slightly lower ($|w_{ZH}|=1$), or much lower for $Z0$ ($|w_{ZH}|=2$) with $H|Z = z \sim \mathcal{N}(-w_{ZH} + 2 * z * w_{ZH}, I_4)$.

For the data sampling for model training, the grouping variable is only measured indirectly and noisy by a correlate U with $U|Z = z \sim \mathcal{N}(z - 0.5, 1)$. The subjects' properties are also observed by correlates, e.g. test scores, denoted as \mathbf{X} , such that $X_{i,j} = H_{i,j} + \epsilon_{i,j}$ with $\epsilon_{i,j} \sim^{iid} \mathcal{N}(0, Var(\epsilon))$ for all properties $i = 1, \dots, 4$ and subjects $j = 1, \dots, N$. The noise of the properties' measurement $Var(\epsilon)$ is either small (0.1), medium (0.3) or large (0.5). The distribution from which the properties are sampled for training $\mathbf{w}_{ZH_{sample}}$ potentially differs from the true distribution of \mathbf{w}_{ZH} . $|\mathbf{w}_{ZH_{sample}}|$ is also 0, 1 or 2, resulting in scenarios where the effect of the grouping variable on the subjects' properties is either smaller, equal or bigger. The subjects' label Y in the training data is observed with (potential) systematic mislabeling for $Z0$, represented by $\Delta\theta_Y|Z$. The label y_i is 1 (positive) if the subject's score s_i is greater than or equal to θ_T , and 0 (negative) otherwise. The score s_i is calculated as the sum of the properties H_i diminished by $\Delta\theta_Y|Z$ for all subjects from $Z0$. There is either no mislabeling ($\Delta\theta_Y|Z = 0$), some ($\Delta\theta_Y|Z = 0.4$) or a lot of false negative labels for $Z0$, while there is no mislabeling for $Z1$.

The sampling parameters mentioned above are combined in a full-factorial design, resulting in 486 scenarios. The structure of the generated data is illustrated in Figure 1.

2.2 Implementation of Adversarial Debiasing

For each scenario, an application dataset with 10,000 samples and 50 training sets with 1,000 samples each were created. Four models were trained on each training set, including one "normal" model and three "debaised" models to increase convergence probability.

The "normal" model was a Multilayer Perceptron (MLP) with two layers, each containing 200 neurons. It was trained using the ADAM algorithm and cross-entropy loss. The "debaised" models incorporated an additional adversary model based on adversarial debiasing [ZLM18]. The adversary model predicted the grouping variable based on the label and the main model's output. The main model's weights were updated to minimize the target prediction loss and increase the adversary's loss, while the adversary's weights were optimized to predict the grouping variable.

During training, we used exponential decay for the learning rate (0.001 initial, 0.96 decay). The adversary's loss weight in the main model was set to 0.1 based on hyperparameter tuning. Training stopped at min 100 epochs (loss < 0.3) or max 400 epochs.

2.3 Evaluation

"Debaised" models with a validation set accuracy more than 0.1 lower than the corresponding "normal" model, or a larger accuracy bias than the "normal" model, were excluded. The best

debiased model (from at most three trials) was selected from each of the 50 training sets of the 486 scenarios for further analysis.

For each scenario, the accuracy (Acc), true positive rate for Z_0 (TPR_0), true negative rate for Z_0 (TNR_0), true positive rate for Z_1 (TPR_1), and true negative rate for Z_1 (TNR_1) were compared between the "normal" and "debiased" model on the application dataset. The performance differences were statistically analyzed using t-tests or Wilcoxon-Mann-Whitney tests. The significance level for the resulting 2430 tests was adjusted using the Holm-Bonferroni method [Ho79]. Additionally, linear models were constructed for each performance metric difference, with the sampling parameters as predictors.

3 Results

Only 20.5% (14947) of all trained "debiased" models met the specified criteria and were considered for further analysis. Out of the 486 scenarios, a total of 14,868 models from 483 scenarios were included in the analysis. It is worth noting that for three scenarios, the adversarial debiasing technique failed, resulting in no suitable "debiased" models.

3.1 Performance Difference on the Application Data

In Table 1, you can find a summary of the results from the performance difference tests conducted on the "debiased" and the "normal" models using application data. The term "significant improvement" indicates that the adversarial debiasing resulted in models with significantly better performance compared to the "normal" models. On the other hand, "significant impairments" suggests that the debiasing process led to models with significantly worse performance. In one of the 483 analysed scenarios ($p_Z = 0.5$, $|\mathbf{w}_{ZH}| = 0$, $\theta_T = 0$,

Performance Metric	Significant Improvement	Significant Impairment	No Significant Difference
Acc	1	121	341
TPR_0	10	106	347
TPR_1	120	1	342
TNR_0	60	25	378
TNR_1	120	1	342

Tab. 1: Summary of the performance analysis comparing the "debiased" model to the "normal" model on the application datasets for each scenario.

(ϵ) = 0.1, $\Delta\theta_Y|Z = 0.8$, $|w_{ZH_{sample}}| = 0$), the application of adversarial debiasing resulted in a significant improvement in overall accuracy ($\hat{m} = 0.005$), with a p-value of 0.000003. In the scenario with $p_Z = 0.5$, $|\mathbf{w}_{ZH}| = 0$, $\theta_T = 2$, (ϵ) = 0.1, $\Delta\theta_Y|Z = 0.8$ and $|w_{ZH_{sample}}| = 2$, the "debiased" model exhibited the highest significant improvement in TPR_0 ($\hat{m} = 0.174$) with a p-value of 0.000000004. In the scenario with $p_Z = 0.95$, $|\mathbf{w}_{ZH}| = 0$, $\theta_T = 0$, (ϵ) = 0.5, $\Delta\theta_Y|Z = 0.8$ and $|w_{ZH_{sample}}| = 1$, the "debiased" model exhibited the highest significant improvement in TNR_0 ($\hat{m} = 0.126$) with a p-value of 0.011.

3.2 Main Effects of the Sampling Parameters

Table 2 provides insights into the performance effects of adversarial debiasing through linear models on application data. The term $\Delta\text{metric}_{z_{app}}$ represents the difference between the metric of the "adversarial" model and the metric of the "normal" model on the application data for subjects with $z = 0$. Positive values indicate an improvement in the metric achieved through the application of adversarial debiasing. Several noteworthy results emerge from the analysis. Firstly, scenarios involving $|\mathbf{w}_{ZH_{sample}}| = 2$ demonstrate a significant improvement in the True Positive Rates (ΔTPR_0 and ΔTPR_1), indicating enhanced detection rates for both groups. However, both TNR_0 and TNR_1 significantly impair for $|\mathbf{w}_{ZH_{sample}}| = 2$, implying diminished detection rates for subjects with $y = 0$. Moreover, scenarios involving $Var(\epsilon) = 0.3$ or $Var(\epsilon) = 0.5$ exhibit significant increases in both TNR_0 and TNR_1 , suggesting improved classification performance for both groups. Conversely, scenarios characterized by $p_Z = 0.95$, $\theta_T = 1$, $\theta_T = 2$, $Var(\epsilon) = 0.3$ or $Var(\epsilon) = 0.5$ showcase significant decreases in both TPR_0 and TPR_1 , indicating decreased detection rates.

Tab. 2: Performance Effect of Adversarial Debiasing on Application Data

	<i>Dependent variable (DV):</i>			
	$\Delta TPR_{0_{app}}$	$\Delta TPR_{1_{app}}$	$\Delta TNR_{0_{app}}$	$\Delta TNR_{1_{app}}$
$p_Z = 0.95$	-0.099***	-0.067***	0.012***	-0.003**
$ \mathbf{w}_{ZH_{sample}} = 1$	0.026***	0.002	0.003***	0.026***
$ \mathbf{w}_{ZH_{sample}} = 2$	0.080***	0.072***	-0.017***	-0.019***
$\theta_T = 1$	-0.024***	-0.034***	0.007***	-0.00005
$\theta_T = 2$	-0.081***	-0.098***	0.004***	-0.012***
$Var(\epsilon) = 0.3$	-0.041***	-0.045***	0.013***	0.030***
$Var(\epsilon) = 0.5$	-0.070***	-0.086***	0.017***	0.048***
$\Delta\theta_Y Z = 0.4$	-0.001	-0.006*	0.003***	0.0003
$\Delta\theta_Y Z = 0.8$	-0.001	-0.012***	0.005***	0.003*
$ \mathbf{w}_{ZH} = 1$	-0.026***	-0.001	0.003***	0.007***
$ \mathbf{w}_{ZH} = 2$	-0.034***	0.003	0.001	0.007***
Intercept	0.032***	0.004	-0.014***	-0.004**

Note: 10, 384 observations for all DV; *p < 0.1; **p < 0.05; ***p < 0.01

4 Discussion

The results demonstrate that adversarial debiasing can improve model performance, particularly when considering specific performance metrics relevant to the application. Notably, scenarios with a high prevalence of subjects with $z = 0$ ($p_Z = 0.95$) showed significant enhancements in TPR_0 and TPR_1 , potentially leading to an overall accuracy improvement. In one scenario, the debiasing even had a positive effect on overall performance. Furthermore, the findings highlight the importance of accurately measured attributes (low $Var(\epsilon)$),

balanced data in terms of the grouping variable ($p_Z = 0.5$), and a target attribute that is not excessively rare within subgroups (small θ_T) for the effectiveness of adversarial debiasing in practical applications. Surprisingly, systematic mislabeling had minimal influence.

However, it should be noted that the implemented debiasing technique faced issues of instability and convergence in many scenarios, suggesting limitations in its adaptability to certain types of data.

The transferability of the identified dependencies between data structures, sampling biases, and model performance to other debiasing methods is questionable. Therefore, further verification is necessary to determine the most suitable debiasing method for different scenarios. Additionally, it is important to acknowledge that the obtained results are limited to the specific binary classification problem with debiasing for a binary grouping variable and the implemented representation and measurement biases. Other scenarios, such as regression problems and different types of data biases (e.g., historical bias), were not considered within the scope of this study.

5 Conclusions

The study shows that adversarial debiasing can improve the performance for the given classification task, especially for specific data subsets and based on the data and sampling structure. But it also highlights that debiasing can have a negative performance impact depending on the characteristics of the data. The results provide guidance for checking the suitability of training data for applying adversarial debiasing, such as measurement accuracies or an adequate representation of relevant data subsets. Further research is needed to validate the found effects using other debiasing methods and scenarios.

Bibliography

- [Be17] Beutel, Alex; Chen, Jilin; Zhao, Zhe; Chi, Ed H.: Data Decisions and Theoretical Implications when Adversarially Learning Fair Representations. 7 2017.
- [GOW20] Goette, Gesa; Oefele, Marcel; Wolters, Viviane: , Beware of the Bias - How the Representations of Subgroups in Training Data Sets affects Binary Classifier Decisions, 2020.
- [Ho79] Holm, Sture: A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 6(2):65–70, 1979.
- [La16] How we analyzed the COMPAS recidivism algorithm.
- [SG19] Suresh, Harini; Gutttag, John V.: A Framework for Understanding Unintended Consequences of Machine Learning. CoRR, abs/1901.10002, 2019.
- [ZLM18] Zhang, Brian Hu; Lemoine, Blake; Mitchell, Margaret: Mitigating Unwanted Biases with Adversarial Learning. 1 2018.

An Approach for a Fast Cost Estimation of Software Projects supported by Sentiment AI Analysis

A Position Paper

Sandro Hartenstein^{1,2,3}, Sidney Leroy Johnson⁴

Abstract: In this short paper, an approach for rapid cost estimation of software projects is presented. Compared to traditional cost estimation, requirements are directly evaluated by an AI sentiment analysis instead of examining a generic target architecture.

Keywords: Cost Estimation, Software Engineering, Sentiment AI Analysis

1 Motivation

A fast cost estimation of software development projects with AI support based on requirements sentiment analysis can help software development managers estimate costs using the best possible techniques. It can save time and resources and improve the accuracy of software cost estimation. This leads to better project planning and resource allocation. Sentiment analysis can help to get a better understanding of the project's requirements. Potential problems in the development process can be more easily identified and addressed.

2 Literature

The requirement to determine the costs of software projects quickly and precisely has existed for a long time, but for an accurate prediction many parameters have to be taken into account. Nagarajappa et al. suggests a learning-based cost estimation model that leverages relational databases to improve accuracy [NS23]. The proposed approach estimates project cost based on the effort required to complete software development, which is a key driver of the project cost. The proposed model is designed to address the challenges posed by the variability in open-source development, including variable team

¹ Berlin School of Economics and Law, Alt-Friedrichsfelde 60, 10315 Berlin, Germany

² Otto-von-Guericke-University Magdeburg, Universitätspl. 2, 39106 Magdeburg, Germany

³ sandro.hartenstein@hwr-berlin.de

⁴ Berlin School of Economics and Law, Alt-Friedrichsfelde 60, 10315 Berlin, Germany; s_johnson@stud.hwr-berlin.de

sizes, working hours and expertise. The proposed model is evaluated against 100 open source software repositories and shows its effectiveness in accurately estimating development costs. 2021 Zakaria et al. involved machine learning algorithms into software effort estimation. The objective of this research is to use several algorithms of machine learning to estimate the effort of software project development [Za21]. The studies examined the performance of AI support in software project effort estimation. Our aim is to identify and operationalize the best practices.

3 Approach

Our approach is to estimate software development costs based on requirements catalogs and optimized AI models. The idea is to use a self-learning model that is trained with a variety of project metadata, software requirements, and designs. Figure 1 illustrates the approach schematically. We use sentiment analysis for classification because we had already successfully tested a similar study using sentiment analysis of requirements for software reengineering projects [SHJ23]. An important finding is that the quality of the results is very much dependent on the training data. For this reason, our initial focus is on good requirements catalogues. In perspective, we expect a more accurate estimation through the additional evaluation of individual sentiment in the requirement description. This has already been investigated for other domains of software development requirements analysis and found to be useful [ZZ19].



Figure 1 Our approach for a fast cost estimation with sentiment ai classification

4 Next Steps


The next step is to evaluate the approach prototypically. This includes the development of the prototype and the training of the model, as well as the validation. The results and findings are intended to demonstrate the capabilities of AI-assisted sentiment analysis in the context of effort and cost estimation in an applicable manner.

Bibliography

- [NS23] Nagarajappa, R. K. B.; Suresh, Y.: Software Development Effort Estimation Using Relational Database and Optimized Learning Mechanism. *Journal of Computer Science* 4/19, pp. 540–553, 2023.

- [SHJ23] Schmietendorf, A.; Hartenstein, S.; Johnson, S. L.: KI-gestützte Modernisierung von Altanwendungen: (Sentiment-) Analysen im Diskurs des Anforderungsmanagements. In (GI-FG SRE Ed.): WSRE, pp. 20–21, 2023.
- [Za21] Zakaria, N. A. et al.: Software Project Estimation with Machine Learning. *International Journal of Advanced Computer Science and Applications* 6/12, 2021.
- [ZZ19] Zhao, L.; Zhao, A.: Sentiment Analysis Based Requirement Evolution Prediction. *Future Internet* 2/11, p. 52, 2019.

User-centered Design Methods in Data-Intensive Software Development Processes: A State-of-the-Art Review

Jasmin Antonia Riebel  ¹

Abstract: As intelligent products and services become our new normal, discussions about user needs gain traction. This paper explores the state-of-the-art literature on the intersection of design methodologies and data-intensive software development. Considering 16 papers out of 182 from 2017-2023, the extent to which user-centered approaches have already found their way into data-intensive software development processes was examined. Here, Explainable Artificial Intelligence (XAI), Human-Centered Explainable Artificial Intelligence (HCXAI), and Human-Centered Artificial Intelligence (HCAI) were identified. Second, existing frameworks and case studies that combine Design Thinking (DT) and AI development were examined. Although DT has been found to be valuable in software development due to its focus on empathy, problem definition, and interdisciplinarity, this paper argues that user-centeredness still falls short and should be considered more consequently, especially in the early stages of development where empathy for the user, the understanding of the problem, and concrete ideas are formed.

Keywords: design methods, AI development, design thinking, user-centered design, values, ethics, ideation, creative process

1 Introduction

As the use of data in software systems becomes increasingly important for the development of innovative products and services [EG19], it is equally important to consider both individual user needs and broader values such as transparency and safety during development. While being part of the HYBRIDA project², the present research presents a review of the state-of-the-art literature on the intersection of design methodology and data-intensive software development, focusing on two research questions: *To which extent have user-centered approaches³ already found their way into data-intensive software development?* and *What role do design methods play in the development process of data-intensive software?*

¹ University of Applied Sciences, Mainz, School of Design, Holzstraße 36, 55116 Mainz, Germany, jasmin.riebel@hs-mainz.de

² Funded over the course of two years (2023-2025) by the Carl Zeiss Foundation, HYBRIDA aims to develop methods and tools to support interdisciplinary collaboration in development teams to foster creativity and communication, resulting in products and services that are more focused on user needs and values.

³ In the context of this research, user-centered approaches describe any framework that prioritizes human needs during the development cycle of an application

2 Methodology

For the literature review, a slightly simplified process derived from Kitchenham guidelines for performing systematic literature reviews in software engineering [Ki07] was used (Fig. 1). The literature search was conducted in two phases. First, a meta-search engine was used to combine keywords related to design methodologies (e.g., tools, techniques) with AI vocabulary, focusing on widely established design frameworks that are known to be user-centered, such as Design Thinking. 2018-2023 was chosen as the time frame for the search. The results were limited to peer-reviewed articles in English. Subject headings suggested by the search engine (e.g., user-centered, user experience, ethics, etc.) were applied to refine the search. Based on title and abstract screening, 13 out of 166 articles were selected. The papers chosen for inclusion were expected to meet the following criteria: (1) either report on practical projects and/or refer to case studies applying design methods in data-intensive software development, or (2) introduce a framework or method for user-centered approaches to software development processes. After screening the full texts, four papers were found to be out of scope and excluded. For the second phase of the search, the reference lists of the included papers were used to back-track further relevant sources. Here, another 16 publications, including conference papers, published in a similar timeframe (2017-2023) were selected, again based on titles and abstracts. After reading the papers in full, seven papers that meet the above-mentioned criteria were included.

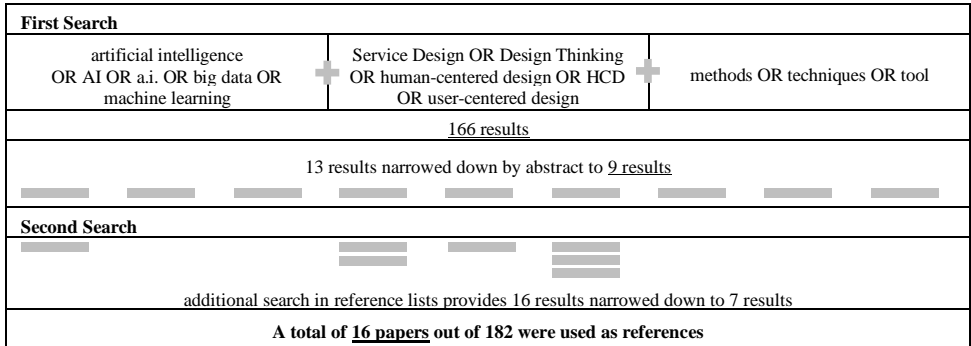


Fig. 1: Simplified Literature Search Process Schema

3 Results

For the first research question, the following user-centered approaches to data-intensive software development were identified:

XAI – One approach to making AI applications more user-friendly is to make them explainable, with the goal of increasing trustworthiness, fairness, accessibility, and privacy awareness, among others. Explainable Artificial Intelligence (XAI) is a discipline that has gained popularity among academics and industry players in recent years. By now,

XAI is "widely acknowledged as a crucial feature for the practical deployment of AI models" [Ar20, p82]. The cornerstone of this concept is understandability, which includes both transparency and interpretability. These two terms open two different and equally essential layers of XAI: On the one hand, the model needs to be transparent and, consequently, understandable to a human – usually the developer of the model. On the other hand, understandability also refers to how comprehensible the decisions made by the model are to a human – usually the user of the application. Therefore, it is particularly important for development teams to know their target audience well, as the purpose of explainability may differ according to different audience profiles [Ar20].

H CXAI – However, explainability alone may not be sufficient. Schoonderwoerd et al. emphasize the need for methods that explicitly focus on human-centered research, design, and development of XAI (HCXAI). As Kirsch critically notes, end users are still hardly considered in explainability research [Ki17]. HCXAI goes a step further than XAI by emphasizing the need to not only address the technical side of the system, but to ground explanations "in an understanding of the primary purpose of the AI system, its users, and its intended use, to learn if, why, what, and when explanation is required" [Sc21, p1]. User-centered design methods can help determine what information to present, when to present it, and how to present it, since "the frequency, form and implementation of explanations depend on the specific application and context" [Ki17, p4].

HCAI – Shneiderman, who developed the Human-Centered Artificial Intelligence (HCAI) framework, agrees that user-centered design approaches are needed. HCAI could "liberate" [Sh20, p1] Design Thinking and empower stakeholders to innovatively apply and refine the systems created. With the HCAI framework, Shneiderman aims to challenge interdisciplinary teams to ask "fresh questions" and develop solutions that "promote human self-efficacy, mastery, and responsibility" [Sh20, p5]. Ultimately, reliable, safe, and trustworthy systems can be achieved by designing technologies that "give users appropriate control while providing high levels of automation" [Sh20, p8].

Looking at the second research question, some attempts have already been made to apply design methods to data-intensive software development. First, the findings and lessons learned from various case studies that have worked with DT in data-intensive software systems will be reported, followed by the evaluation of different frameworks that combine DT or equivalent methods with data-intensive software development.

Case Studies – Novak et al. define DT as "a systematic innovation approach that emphasizes empathy and engagement of end users in data collection and analysis, and rapid, iterative prototyping to identify design solutions" [No20, p1859]. Consequently, understanding both the problem and potential users is critical for developing a comprehensive solution. In their study of DT in applied informatics, the authors describe the discomfort participants initially felt when being asked to reconsider their values and beliefs, in contrast to simply implementing a solution to a predefined task. Ultimately,

participants rated the experience positively, emphasizing the interdisciplinary and collaborative nature of the process while highlighting DT's ability to address complex societal issues, interdisciplinarity, and problem structuring as characteristics especially applicable in an AI context. Pham et al., who explored the role of DT in big data innovation, found DT to be particularly helpful in identifying the project's rationale to be developed and, consequently, in selecting the most appropriate data to achieve the intended outcomes. Moreover, design methodologies are particularly important to ensure that data-driven products and services are customer-centric. As a result, big data is more likely to reach its full potential when combined with DT [PMD21]. By using examples from Netflix and Airbnb, Verganti et al. explain how AI gives way to solutions that are potentially more user-centered since AI designs not only for a specific target group but also personalizes the solution for every single individual. Consequently, AI "reinforces the principles of Design Thinking, namely: being people-centered, abductive, and iterative" [VVI20, p212]. Therefore, AI changes the way design is practiced. Giaccardi et al. even question whether common user-centered design frameworks have reached their limits due to the "constant becoming" [Gi20, p33] of contemporary technology such as AI, and suggest exploring the idea of designing not for technologies, but with them.

Frameworks – The changing nature of the DT process in relation to AI [VVI20, Gi20] could as well be observed when examining different frameworks proposed in the publications reviewed. All authors based their frameworks on the DT process and mindset but also made some adaptations to better serve AI purposes (Fig. 2). As Chebabi et al. state, one of the biggest challenges was that clients didn't always have a clear vision of how AI could benefit their businesses – they just wanted to use AI "because it's a buzz word" [CA20, p212]. The difficulty with this approach is that the process does not start with an open problem statement, but with a predefined technology. To better support the process, the authors extended the Ideation phase by adding three sub-steps. "Empathy" focuses on user needs and the customer journey, while "Data" presents various use cases as examples of how AI could be part of the solution. In the "Solutions" phase, the team examines what data the client already has and what would be needed. Only then brainstorming is used to create ideas that are ideally already viable. The authors agree with Chhabra and Williams' call for cross-functional teams [CW19] since for success it is crucial to include specialists with data knowledge in the DT team who can answer questions on topics such as data security.

Like Chebabi et al., Bhalla emphasizes empathy in his framework [Bh19]. With "The 3S Process," he suggests making AI strategy a part of the business school curriculum to better prepare students for careers as leaders who will eventually deal with the adoption of AI in their organizations. In his framework, DT helps the students to gain a better understanding of the user's needs, to allow for different perspectives on the problem, and to be clear about ethical concerns, such as potential privacy and security issues.

Coming more from a technologist perspective, Kurti et al. address the challenge of AI systems lacking transparency by using DT to increase the understandability and explainability of ML and AI solutions since too little is known yet about “the way of communicating the explanations to users such as that they enhance the user’s trust” [Ku21, p38]. To achieve user-centered explanations that ultimately lead to acceptance of and collaboration with the system, the authors adapt the standard DT process. For example, during "Explore", the developers iteratively examine the given dataset to gain an understanding of what features might help explain the ML and AI results. The authors conclude that DT widened their understanding of the problem space and helped them “to identify alternative possibilities that might not be instantly apparent with our initial level of understanding” [Ku21, p40].

According to Zhou et al., current design methods are not tailored to guide designers in working with the constantly growing nature of ML. Therefore, they developed a design method called Material Lifestyle Thinking (MLT) and an accompanying canvas for the conceptual phase of the process, with the goal of bridging the gap between ML and UX. MLT consists of five iterative phases (Fig. 2) and works with practical design guidelines while respecting the principles of DT [Zh20].

DT Standard Process	Empathize		Define		Ideate	Prototype	Test		
The 3S Process (Bhalla, 2019)	Story		Strategy			Solution			
	Scenario	Research	Empathy	Define	Ideate	Prototype	Test	Deploy	Feedback
BlueJourney (Chebabi et al., 2020)	Investigation		Ideation			Prototype	Pitches and Selection		
			Empathy	Data	Solutions				
DT for ML & AI (Kurti et al., 2021)	Explore		Define		Select	Tune		Evaluate	
ML Lifecycle Canvas (Zhou et al.)	Understand crucial features; describe design goal		Identify co-creators	Identify touchpoints where co-creators exchange resources		Envision possible UX by considering tradeoffs		Monitoring	

Fig. 2: Comparing the Examined Frameworks to the Standard Design Thinking Process

4 Conclusion

For the first research question, XAI, HCXAI, and HCAI were identified as user-centered approaches to data-intensive software development. Since the present research was conducted by a single researcher, a possible limitation of this review might be its reliability although inclusion and exclusion decisions have been discussed with fellow researchers.

Addressing the second research question, it was found that frameworks such as DT have a lot to offer for a problem-based and user-centered approach. Firstly, empathy needs to be at the core of the design process and is crucial for understanding the user’s needs and examining the problem at hand from multiple and diverse viewpoints. According to Weller, the DT framework helps development teams to stay open-minded throughout the

whole process and consider alternative possibilities when needed [We19]. Moreover, ‘interdisciplinary’ was mentioned often in the reviewed papers as another cornerstone of the DT approach since “diversity of perspective has been shown time and again to foster fresh thinking to solve tough problems” [CW19, p2]. Moreover, the case studies show that another strength of the DT process lies in problem definition and questioning of the status quo, or as Liedtka puts it, DT fosters “enhanced learning” [Li20, p54].

On the other hand, some shortcomings could be identified when DT is applied to data-intensive software development processes. According to Zhou et al., established design frameworks like DT fall short in considering the growable nature of data-driven applications [Zh20]. Therefore, design processes need to become anticipatory [GR20] instead of reactive to be able to accommodate users’ needs in a complex environment. DT processes usually sketch the next near future and then react to changes made to, e.g., the new software system through iteration. In the case of data-intensive solutions, DT should anticipate further consequences of the use of data, especially ethical aspects, and situations where the use of data can potentially generate negative consequences for individuals and society. This is especially important since genuine ethical values might go beyond the wishes and needs of users and, therefore, need to be anticipated by the development team.

In a nutshell, there is still a gap between what research knows and what practitioners do, since “common approaches to deploying AI tools are not improving outcomes”, [LS20, pE1]. To develop useful AI tools, the process needs to begin with “asking what system change the AI tool is expected to precipitate” since being intentional is crucial when it comes to “matching the algorithm to the problem, and not the other way around” [pE2]. When data science connects with the stakeholders’ needs early, a positive influence and, therefore, the desired change will be reached faster, leading to a more effective process of deploying AI tools [LS20]. However, user-centeredness still seems to be treated more as an ‘add-on’ that comes into play after technical questions have been addressed. Therefore, methods that integrate the human factor already in the research stage [Sc21] are needed.

Consequently, this paper calls for a user-centered design process that starts with problem-finding and value-centered ideation as opposed to the more common technology or business-driven processes. User-centeredness should be introduced continuously from the very beginning and take on a more holistic perspective instead of solely focusing on distinctive parts such as user-centered explainability. Therefore, concentrating more on value-related “should we” questions rather than technology-driven “could we” questions is necessary, especially at the beginning of the process in phases such as ‘Define’ and ‘Ideate’ (Fig. 2). The empathy phase of DT seems to be ideal for anticipating and addressing potential ethical issues early in development [Bh19]. To get deeper into these promising topics, it is necessary to understand how ethical values are already considered in user-centered methods, and how these values can be built into data-intensive software systems. Moreover, it is crucial to verify the findings of the present research in a real-world setting, using qualitative research methods such as semi-structured interviews to gain further insides into industry best practices when it comes to user-centered approaches towards data-intensive software development.

Bibliography

- [Ar20] Arrieta, A. B. et al.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* 58, C (Jun 2020), 82–115, 2020.
- [Bh19] Bhalla, N.: The 3S Process: A Framework for Teaching AI Strategy in Business Education. *Technology Innovation Management Review*, 9(12), 36-42, 2019.
- [CW19] Chhabra, A.; Williams, S.: Fusing data and design to supercharge innovation – in products and processes, <https://www.mckinsey.com/capabilities/quantumblack/our-insights/fusing-data-and-design-to-supercharge-innovation-in-products-and-processes>, accessed: 17/05/2023
- [CA20] Chebabi, R. Z; von Atzingen Amaral, H.: BlueJourney for AI – A Study Beyond Design Thinking to Develop Artificial Intelligence Solutions. In: Marcus, A., Rosenzweig, E. (eds) *Design, User Experience, and Usability. Design for Contemporary Interactive Environments. HCII 2020. Lecture Notes in Computer Science*, 12201, 212-221. Springer, Cham, 2020.
- [EG19] Engels, B.; Goecke, H.: Big Data in Wirtschaft und Wissenschaft: Eine Bestandsaufnahme (Nummer 130), IW-Analysen, <https://www.econstor.eu/bitstream/10419/201760/1/1670732177.pdf>, accessed: 10/07/2023.
- [GR20] Giaccardi, E.; Redström, J.: Technology and More-Than-Human Design. *Design Issues: history/theory/criticism*, 36(4), 33-44, 2020.
- [Ki07] Kitchenham, B.: Guidelines for performing Systematic Literature Reviews in Software Engineering, Version 2.3., https://www.elsevier.com/_data/promis_misc/525444systematicreviewsguide.pdf, accessed: 11/07/2023.
- [Ki17] Kirsch, A.: Explain to whom? Putting the User in the Center of Explainable AI. *Proceedings of the First International Workshop on Comprehensibility and Explanation in AI and ML 2017 co-located with 16th International Conference of the Italian Association for Artificial Intelligence (AI*IA 2017)*, Bari, Italy, 2017.
- [Ku21] Kurti, A. et al.: Increasing the Understandability and Explainability of Machine Learning and Artificial Intelligence Solutions: A Design Thinking Approach. In: Ahram, T., Taiar, R., Groff, F. (eds) *Human Interaction, Emerging Technologies and Future Applications IV. IHMET-AI 2021. Advances in Intelligent Systems and Computing*, 1378, 37-42. Springer, Cham, 2021.
- [Li20] Liedtka, J.: Putting Technology in Its Place: Design Thinking’s Social Technology at Work. *California Management Review*, 62, 53-83, 2020.
- [LS20] Lindsell, C. J.; Stead, W. W.; Johnson, K. B.: Action-Informed Artificial Intelligence-Matching the Algorithm to the Problem. *JAMA*, 323(21), 2141–2142, 2020.
- [No20] Novak, L. L. et al.: Design thinking in applied informatics: what can we learn from Project HealthDesign? *Journal of the American Medical Informatics Association: JAMIA*, 28(9), 1858–1865, 2020

- [PMD21] Pham, C.T.; Magistretti, S.; Dell’Era, C.: The role of design thinking in Big Data innovations. *Innovation*, 24, 290 – 314, 2021.
- [Sc21] Schoonderwoerd, T. A. J. et al.: Human-centered XAI: Developing design patterns for explanations of clinical decision support systems. *International Journal of Human Computer Studies*, 154, 1-25. [102684], 2021.
- [Sh20] Shneiderman, B.: Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction*, 36, 1-10, 2020.
- [VVI20] Verganti, R.; Vendraminelli, L.; Iansiti, M.: Innovation and Design in the Age of Artificial Intelligence. *Journal of Product Innovation Management*, 37, 212-227, 2020.
- [We19] Weller, A. J.: Design thinking for a user-centered approach to artificial intelligence. *She Ji: The Journal of Design, Economics, and Innovation*, 5(4), 394-396, 2019.
- [Zh20] Zhou, Z. et al.: ML Lifecycle Canvas: Designing Machine Learning-Empowered UX with Material Lifecycle Thinking. *Human-Computer Interaction*, 35, 362 – 386, 2020.

Deep Learning Datasets Challenges For Semantic Segmentation - A Survey

Claire Ponciano,¹ Markus Schaffert, Jean-Jacques Ponciano

Abstract: This survey offers a comprehensive analysis of challenges encountered when employing large-scale datasets for deep learning-based semantic segmentation, an area with significant implications for industries such as autonomous driving, precision agriculture, and medical imaging. Through a systematic review of 94 papers from *Papers with Code*, we identified 32 substantial challenges, which we categorized into six key areas: Data Quality and Quantity, Data Preprocessing, Resource Constraints, Data Management and Privacy, Generalization, and Data Compatibility. By identifying and explicating these challenges, our research provides a crucial reference point for future studies aiming to address these issues and enhance the performance of deep learning models for semantic segmentation. Future work will focus on leveraging AI and semantic technologies to provide solutions to these challenges.

Keywords: Deep Learning; Deep Learning challenges; Semantic segmentation; Data quality; Resource constraints; Generalization; Data management; Data privacy; Data compatibility

1 Introduction

Deep learning has revolutionized computer vision, particularly semantic segmentation, which involves pixel or point classification in image or point cloud respectively. Among various emerging deep learning models, our focus is on semantic segmentation due to its versatility across fields like BIM [PPP21], Optimal Recording and Modelling [Po19, PPB21, Po21], precision agriculture [An21], and medical imaging [Is18b]. While these models are effective, they encounter challenges, often discussed in terms of the models themselves [Sh19]. Large-scale datasets, used in deep learning, present their own unique set of hurdles, which curtail their use and model effectiveness, despite their transformational potential [Xi22]. For example, in drone-based semantic segmentation [Ta21], issues like low-resolution images, class imbalance, varying lighting and weather conditions, and the presence of sensitive information pose significant challenges to model accuracy and data privacy. Overcoming these challenges is crucial for enhancing deep learning models' performance in semantic segmentation, which necessitates a thorough understanding of these hurdles and the development of robust, ethical strategies for dataset utilization. This paper examines these challenges by answering the research question: What are the dataset challenges

¹ Mainz University of Applied Sciences, i3mainz – Institute for Spatial Information and Surveying Technology, Lucy-Hillebrand-Straße 2 55128 Mainz, Germany (claire.ponciano,markus.schaffert,jean-jacques.ponciano)@hs-mainz.de

negatively impacting deep learning-based semantic segmentation models? We conduct a survey on 47 datasets (from a pool of 267) and 94 related papers from *Papers with Code*². These selected datasets and the survey of their related papers aim to identify obstacles in large-scale datasets for semantic segmentation based on deep learning, thus promoting the development of more reliable and robust models. This endeavor constitutes a significant step towards understanding and mitigating the challenges linked to deep learning datasets for semantic segmentation, with potential far-reaching benefits across industries. The paper subsequently presents related work in dataset challenge research, outlines our method for identifying challenges, and discusses the 32 identified challenges and their implications. Then, this survey concludes that it unveils a diverse array of challenges in large-scale datasets for deep learning-based semantic segmentation, laying a comprehensive groundwork for future endeavors aimed at harnessing artificial intelligence and semantic technologies to address these issues and advance the field of computer vision.

2 Related Work

Research on constructing high-quality datasets is a nascent field, where works like [Sh19] delve into complexities of deep learning networks, and others such as [Mu19] probe issues linked to these datasets, but bypass semantic segmentation. However, in the paper [Mu19], we identify applicable challenges for creating semantic segmentation datasets, which include hurdles like metadata scarcity and diversity deficiency, which can hamper training and evaluation. In addition, the *heterogeneity in data* signifies a delicate balance between diversity and consistency, pivotal for dataset utility. *Data Quality* is critical as inferior data can degrade model performance, and *Data sources and Distribution* underscores the necessity for varied data sources and balanced classes to avert model bias. These challenges elucidate the diligence required in forming quality semantic segmentation datasets and also highlight the potential for enhanced deep learning outcomes. Despite these challenges identified in related works, our understanding remains limited. Hence, we propose an exhaustive survey to identify dataset-based challenges, leading us to our research question. The subsequent section outlines the method employed to survey these dataset challenges.

3 Method

Semantic segmentation, powered by a myriad of available datasets, is a burgeoning area of research. *Papers with Code* is a hosting platform providing datasets, papers, and code. It provides an invaluable resource for our survey.

² <https://paperswithcode.com/>

Datasets	Selected Papers
2D-3D-S	[Ke17, RFB15]
ADE20K	[Li19, WH21]
ApolloScape	[Ca20, Hu19]
AVE	[Zh22b, Zh23]
BDD100K	[He17, Wa20a]
BraTS 2015	[Is18a, Ka17]
CamVid	[BKC17, Yu21]
Cityscapes	[MWA18a, Ch17c]
COCO	[Ho17, Wa22]
COCO-Stuff	[Ki19, Fa22]
DAVIS	[Ha22, Vo19]
DAVIS2017	[St21, Vo19]
EuroSAT	[GD22, Ne19]
GTA5	[He16, SZ14]
HAM10000	[Ge20, Jh20]
Helen	[De20, Zh22a]
IDD	[CHS22, MV21]
KITTI	[Ca19, MWA18b]
KITTI-360	[Qi17b, Qi17a]
Kvasir	[RFB15, Zh18b]
LabelMe	[Gh15, Zh19]
LIP	[Su19, Zh18a]
Make3D	[Go19, GVZ16]
Mapillary Vistas Dataset	[Bu20, Ch20]
Medical Segmentation Decathlon	[Is18b, Ji22]
NYU Depth V2	[LSD15a, Oq23]
Objects365	[Fa22, Ku22]
Pascal-5i	[LE22, Zh20c]
PascalContext	[Wa20b, Ch17b]
Pascal VOC 2007	[Ch17b, Re16]
Pascal VOC 2012	[Re16, Ch17c]
PartNet	[Gu20, Mo19a]
Promise12	[MNA16, Is19]
ReferItGame	[Yu16, Ac21]
RefCOCO	[Li21, Zo23]
ScanNet	[Pa16, He17]
Semantic3D	[Hu20, Gu20]
SegTrack-v2	[Oh18, Su23]
ShapeNet	[Ch15, Su18]
S3DIS	[Gu17, Qi17a]
SUN3D	[Ch17a, Sa20]
SUNCG	[ESL19, Mo19b]
SUN RGB-D	[LSD15b, BKC17]
Synthia	[Zh20a, Hu18]
VisDA-2017	[Yu20, Zh20b]
Virtual KITTI	[Hu22, Re19]
YouTube-VIS 2019	[Go21, WBP17]

Tab. 1: Selected datasets with surveyed top papers

The survey presented in this paper is based on 47 datasets out of 267. These 47 datasets have been selected according to two criteria showing widespread use and thus enabling us to highlight the relevance of the challenges identified by the survey for practical applications. The first criterion for selecting a dataset is that it should have been used in more than 50 papers³. The second criterion is that the dataset has been used by the most *starred* approaches on GitHub. Once these 47 datasets had been identified, the two most cited papers related to these 47 datasets available on *Papers with Code* were studied in order to (i) identify the dataset challenges highlighted by the authors, and (ii) correlate the issues of the approaches related to the same dataset with the characteristics of the dataset in order to deduce the problematic characteristics of a dataset and therefore its challenges. The 47 datasets and the 94 related papers studied for this paper are shown in Table 1. This survey spans a wide range of data types, problem contexts, and application domains. From autonomous driving to precision agriculture to medical imaging, each of them has unique data characteristics and inherent challenges. For example, autonomous driving datasets like KITTI and ApolloScape often present high-resolution urban scenes with many object classes, leading to issues such as class imbalance and extensive computational requirements. Precision agriculture datasets like EuroSAT may comprise multispectral or hyperspectral data, adding complexity to data preprocessing and management. Medical imaging datasets like PROMISE12 and BraTS 2015 are generally marked by class imbalance and stringent privacy constraints, thereby posing a distinct set of challenges. The identified challenges were then grouped together where they related to a similar problem. This method enabled us to identify 32 challenges, which are critical challenges in the development of deep learning models for semantic segmentation. These challenges have been classified into six groups and are described in the following section.

4 Results

The analysis of the comprehensive assortment enabled us to compile an exhaustive array of challenges associated with deep learning-based semantic segmentation as seen in Table 2, organized into six categories: (i) Data Quality and Quantity, (ii) Data Preprocessing, (iii) Resource Constraints, (iv) Data Management and Privacy, (v) Generalization, and (vi) Data Compatibility. In the category of **Data Quality and Quantity**, challenges include ensuring accurate annotations (1), incorporating diverse objects into the dataset (2), achieving class balance (3), accounting for data variability (4), managing data resolution (5), handling class overlap (6), ensuring overall data quality (7), and determining the level of detail in annotations (8). The **Data Preprocessing** category involves challenges such as applying augmentation techniques for data diversification (9), aligning data through registration (10), normalizing data for scale invariance (11), standardizing data for consistency (12), scaling data to desired sizes (13), and resolving ambiguity in class labels (14). Within the **Resource Constraints** category, challenges include dealing with lengthy training times

³ at the time of writing, i.e., April 2023

(15), optimizing computational resources (16), and managing the cost associated with data annotation (17). In the realm of **Data Management and Privacy**, challenges encompass ethical considerations (18), privacy protection (19), anonymization techniques (20), data security (21), access control (22), and effective data storage and management (23). The **Generalization** category involves challenges related to data bias (24), data distribution (25), outlier detection (26), rare events (27), and long-tail distribution (28).

Category	Challenges	Datasets
Data Quality and Quantity	Annotation Quality Object Diversity Data Imbalance Data Variability Data Resolution Class Overlap Data Quality Annotation Granularity	COCO, ImageNet, Pascal VOC 2012 ImageNet, Pascal VOC 2012, Objects365 COCO, ImageNet, Pascal VOC 2012, Cityscapes, BDD100K NYU Depth V2, KITTI, SUN RGB-D, SUN3D KITTI, ImageNet, Cityscapes COCO, ImageNet, Pascal VOC 2012, Cityscapes KITTI, Cityscapes, BDD100K ShapeNet, PartNet
Data Preprocessing	Data Augmentation Data Registration Normalization Standardization Scaling Class Ambiguity	All datasets KITTI, NYU Depth V2, SUN RGB-D, SUN3D All datasets All datasets ImageNet, COCO, Cityscapes COCO, ImageNet, Pascal VOC 2012, Cityscapes
Resource Constraints	Training Time Resource Availability Annotation Cost	ImageNet, COCO, BDD100K ImageNet, COCO, BDD100K COCO, ImageNet, Pascal VOC 2012, Cityscapes
Data Management and Privacy	Ethical Considerations Data Privacy Data Anonymization Encryption Access Control Data Storage and Management	AVE, YouTube-VIS 2019, BDD100K, IDD AVE, YouTube-VIS 2019, BDD100K, IDD AVE, YouTube-VIS 2019, BDD100K, IDD All datasets All datasets ImageNet, COCO, BDD100K
Generalization	Data Bias Data Distribution Presence of Outliers Rare Events Long-tail Distribution	ImageNet, COCO, Pascal VOC 2012, Cityscapes ImageNet, COCO, Pascal VOC 2012, Cityscapes ImageNet, COCO, Pascal VOC 2012 BDD100K, Cityscapes, KITTI ImageNet, COCO, Pascal VOC 2012, Cityscapes
Data Compatibility	Data Format Interoperability Hardware Compatibility Data Versioning	All datasets All datasets All datasets All datasets

Tab. 2: Summary of challenges in semantic segmentation deep learning datasets with examples

Finally, challenges within the **Data Compatibility** category include handling varying data formats (29), ensuring interoperability between systems (30), optimizing hardware compatibility (31), and managing data versioning (32). These challenges collectively shape the landscape of data-driven research, prompting scientists and scholars to develop innovative solutions and strategies to overcome them and unlock the full potential of data-driven discoveries. Moreover, Table 2 shows that some challenges were identified across two to five datasets, but others (i.e., Data Augmentation, Normalization, Standardization, Encryption, Access Control, Data Format, Interoperability, Hardware Compatibility, and Data Versioning) were observed in all datasets. The challenges observed in the 47 datasets can therefore be seen as priorities, since they are the most widespread, whatever the application domain. Consequently, Data Compatibility, Data Preprocessing, Data Management and Privacy seem to be the most pressing areas to address in order to improve deep learning-based semantic segmentation.

5 Conclusion

This survey has embarked on a comprehensive journey to identify the key obstacles in employing large-scale datasets for deep learning-based semantic segmentation. From 94 papers available in *Papers With Code* repository, we have identified 32 prominent challenges. These challenges span six categories: Data Quality and Quantity, Data Preprocessing, Resource Constraints, Data Management and Privacy, Generalization, and Data Compatibility. We have delved into each category, shedding light on the nuances and intricacies of the challenges therein. Issues pertaining to the quality, volume, and management of data, as well as the computational resources required, have all been examined. Moreover, we have scrutinized the hurdles that affect the model's ability to generalize to unseen data and the difficulties posed by data compatibility issues. This exploration has enabled us to better understand the landscape of semantic segmentation challenges, and to identify the most pressing challenges to be solved. The identified challenges paves the way for future work aimed at devising effective solutions. Looking ahead, our future work will be geared towards developing a solution leveraging artificial intelligence and semantic technologies to overcome these challenges. By building upon the comprehensive understanding of the challenges that this survey offers, we aim to create a framework that is capable of addressing these issues in a holistic and efficient manner. In summary, the survey has laid a solid foundation for future research in this domain, setting the stage for the development of solutions that can truly harness the power of large-scale datasets in semantic segmentation. This represents a significant stride towards advancing the field of computer vision, with potential implications extending across various industries and domains.

Bibliography

- [Ac21] Achlioptas, Panos; Ovsjanikov, Maks; Haydarov, Kilichbek; Elhoseiny, Mohamed; Guibas, Leonidas J: Artemis: Affective language for visual art. In: Proceedings of the

- IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11569–11579, 2021.
- [An21] Anand, Tanmay; Sinha, Soumendu; Mandal, Murari; Chamola, Vinay; Yu, Fei Richard: AgriSegNet: Deep aerial semantic segmentation framework for IoT-assisted precision agriculture. *IEEE Sensors Journal*, 21(16):17581–17590, 2021.
- [BKC17] Badrinarayanan, Vijay; Kendall, Alex; Cipolla, Roberto: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.
- [Bu20] Buslaev, Alexander; Iglovikov, Vladimir I; Khvedchenya, Eugene; Parinov, Alex; Druzhinin, Mikhail; Kalinin, Alexandr A: Alumentations: fast and flexible image augmentations. *Information*, 11(2):125, 2020.
- [Ca19] Casser, Vincent; Pirk, Soeren; Mahjourian, Reza; Angelova, Anelia: Depth prediction without the sensors: Leveraging structure for unsupervised learning from monocular videos. In: *Proceedings of the AAAI conference on artificial intelligence*. volume 33, pp. 8001–8008, 2019.
- [Ca20] Caesar, Holger; Bankiti, Varun; Lang, Alex H; Vora, Sourabh; Liong, Venice Erin; Xu, Qiang; Krishnan, Anush; Pan, Yu; Baldan, Giancarlo; Beijbom, Oscar: nuscenes: A multimodal dataset for autonomous driving. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 11621–11631, 2020.
- [Ch15] Chang, Angel X; Funkhouser, Thomas; Guibas, Leonidas; Hanrahan, Pat; Huang, Qixing; Li, Zimo; Savarese, Silvio; Savva, Manolis; Song, Shuran; Su, Hao et al.: Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015.
- [Ch17a] Chang, Angel; Dai, Angela; Funkhouser, Thomas; Halber, Maciej; Niessner, Matthias; Savva, Manolis; Song, Shuran; Zeng, Andy; Zhang, Yinda: Matterport3d: Learning from rgb-d data in indoor environments. *arXiv preprint arXiv:1709.06158*, 2017.
- [Ch17b] Chen, Liang-Chieh; Papandreou, George; Kokkinos, Iasonas; Murphy, Kevin; Yuille, Alan L: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.
- [Ch17c] Chen, Liang-Chieh; Papandreou, George; Schroff, Florian; Adam, Hartwig: Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.
- [Ch20] Chen, Liang-Chieh; Lopes, Raphael Gontijo; Cheng, Bowen; Collins, Maxwell D; Cubuk, Ekin D; Zoph, Barret; Adam, Hartwig; Shlens, Jonathon: Naive-student: Leveraging semi-supervised learning in video sequences for urban scene segmentation. In: *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX* 16. Springer, pp. 695–714, 2020.
- [CHS22] Caron, Mathilde; Houlsby, Neil; Schmid, Cordelia: Location-Aware Self-Supervised Transformers. *arXiv preprint arXiv:2212.02400*, 2022.
- [De20] Deng, Jiankang; Guo, Jia; Ververas, Evangelos; Kotsia, Irene; Zafeiriou, Stefanos: Retinaface: Single-shot multi-level face localisation in the wild. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 5203–5212, 2020.

- [ESL19] Eysenbach, Ben; Salakhutdinov, Russ R; Levine, Sergey: Search on the replay buffer: Bridging planning and reinforcement learning. *Advances in Neural Information Processing Systems*, 32, 2019.
- [Fa22] Fang, Yuxin; Wang, Wen; Xie, Binhui; Sun, Quan; Wu, Ledell; Wang, Xinggang; Huang, Tiejun; Wang, Xinlong; Cao, Yue: Eva: Exploring the limits of masked visual representation learning at scale. *arXiv preprint arXiv:2211.07636*, 2022.
- [GD22] Gesmundo, Andrea; Dean, Jeff: An evolutionary approach to dynamic introduction of tasks in large-scale multitask learning systems. *arXiv preprint arXiv:2205.12755*, 2022.
- [Ge20] Gessert, Nils; Nielsen, Maximilian; Shaikh, Mohsin; Werner, René; Schlaefer, Alexander: Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. *MethodsX*, 7:100864, 2020.
- [Gh15] Ghifary, Muhammad; Kleijn, W Bastiaan; Zhang, Mengjie; Balduzzi, David: Domain generalization for object recognition with multi-task autoencoders. In: *Proceedings of the IEEE international conference on computer vision*. pp. 2551–2559, 2015.
- [Go19] Godard, Clément; Mac Aodha, Oisín; Firman, Michael; Brostow, Gabriel J: Digging into self-supervised monocular depth estimation. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 3828–3838, 2019.
- [Go21] Gong, Tao; Chen, Kai; Wang, Xinjiang; Chu, Qi; Zhu, Feng; Lin, Dahua; Yu, Nenghai; Feng, Huamin: Temporal ROI align for video object recognition. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. volume 35, pp. 1442–1450, 2021.
- [Gu17] Gupta, Saurabh; Davidson, James; Levine, Sergey; Sukthankar, Rahul; Malik, Jitendra: Cognitive mapping and planning for visual navigation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 2616–2625, 2017.
- [Gu20] Guo, Yulan; Wang, Hanyun; Hu, Qingyong; Liu, Hao; Liu, Li; Bennamoun, Mohammed: Deep learning for 3d point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(12):4338–4364, 2020.
- [GVZ16] Gupta, Ankush; Vedaldi, Andrea; Zisserman, Andrew: Synthetic data for text localisation in natural images. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 2315–2324, 2016.
- [Ha22] Hao, Yuying; Liu, Yi; Chen, Yizhou; Han, Lin; Peng, Juncai; Tang, Shiyu; Chen, Guowei; Wu, Zewu; Chen, Zeyu; Lai, Baohua: EISeg: An Efficient Interactive Segmentation Tool based on PaddlePaddle. *arXiv e-prints*, pp. arXiv–2210, 2022.
- [He16] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian: Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 770–778, 2016.
- [He17] He, Kaiming; Gkioxari, Georgia; Dollár, Piotr; Girshick, Ross: Mask r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. pp. 2961–2969, 2017.
- [Ho17] Howard, Andrew G; Zhu, Menglong; Chen, Bo; Kalenichenko, Dmitry; Wang, Weijun; Weyand, Tobias; Andreetto, Marco; Adam, Hartwig: Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

- [Hu18] Huang, Xun; Liu, Ming-Yu; Belongie, Serge; Kautz, Jan: Multimodal unsupervised image-to-image translation. In: Proceedings of the European conference on computer vision (ECCV). pp. 172–189, 2018.
- [Hu19] Huang, Xinyu; Wang, Peng; Cheng, Xinjing; Zhou, Dingfu; Geng, Qichuan; Yang, Ruigang: The apolloscape open dataset for autonomous driving and its application. *IEEE transactions on pattern analysis and machine intelligence*, 42(10):2702–2719, 2019.
- [Hu20] Hu, Qingyong; Yang, Bo; Xie, Linhai; Rosa, Stefano; Guo, Yulan; Wang, Zhihua; Trigi, Niki; Markham, Andrew: Randla-net: Efficient semantic segmentation of large-scale point clouds. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11108–11117, 2020.
- [Hu22] Hu, Hou-Ning; Yang, Yung-Hsu; Fischer, Tobias; Darrell, Trevor; Yu, Fisher; Sun, Min: Monocular quasi-dense 3d object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):1992–2008, 2022.
- [Is18a] Isensee, Fabian; Kickingereder, Philipp; Wick, Wolfgang; Bendszus, Martin; Maier-Hein, Klaus H: Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3. Springer, pp. 287–297, 2018.
- [Is18b] Isensee, Fabian; Petersen, Jens; Klein, Andre; Zimmerer, David; Jaeger, Paul F; Kohl, Simon; Wasserthal, Jakob; Koehler, Gregor; Norajitra, Tobias; Wirkert, Sebastian et al.: nnu-net: Self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486*, 2018.
- [Is19] Isensee, Fabian; Jäger, Paul F; Kohl, Simon AA; Petersen, Jens; Maier-Hein, Klaus H: Automated design of deep learning methods for biomedical image segmentation. *arXiv preprint arXiv:1904.08128*, 2019.
- [Jh20] Jha, Debesh; Riegler, Michael A; Johansen, Dag; Halvorsen, Pål; Johansen, Håvard D: Doubleu-net: A deep convolutional neural network for medical image segmentation. In: 2020 IEEE 33rd International symposium on computer-based medical systems (CBMS). IEEE, pp. 558–564, 2020.
- [Ji22] Jiao, Rushi; Zhang, Yichi; Ding, Le; Cai, Rong; Zhang, Jicong: Learning with limited annotations: a survey on deep semi-supervised learning for medical image segmentation. *arXiv preprint arXiv:2207.14191*, 2022.
- [Ka17] Kamnitsas, Konstantinos; Ledig, Christian; Newcombe, Virginia FJ; Simpson, Joanna P; Kane, Andrew D; Menon, David K; Rueckert, Daniel; Glocker, Ben: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical image analysis*, 36:61–78, 2017.
- [Ke17] Ke, Guolin; Meng, Qi; Finley, Thomas; Wang, Taifeng; Chen, Wei; Ma, Weidong; Ye, Qiwei; Liu, Tie-Yan: Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.
- [Ki19] Kirillov, Alexander; Girshick, Ross; He, Kaiming; Dollár, Piotr: Panoptic feature pyramid networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 6399–6408, 2019.

- [Ku22] Kuo, Weicheng; Cui, Yin; Gu, Xiuye; Piergiovanni, AJ; Angelova, Anelia: F-VLM: Open-Vocabulary Object Detection upon Frozen Vision and Language Models. arXiv preprint arXiv:2209.15639, 2022.
- [LE22] Lüddecke, Timo; Ecker, Alexander: Image segmentation using text and image prompts. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7086–7096, 2022.
- [Li19] Liu, Chenxi; Chen, Liang-Chieh; Schroff, Florian; Adam, Hartwig; Hua, Wei; Yuille, Alan L; Fei-Fei, Li: Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 82–92, 2019.
- [Li21] Li, Junnan; Selvaraju, Ramprasaath; Gotmare, Akhilesh; Joty, Shafiq; Xiong, Caiming; Hoi, Steven Chu Hong: Align before fuse: Vision and language representation learning with momentum distillation. *Advances in neural information processing systems*, 34:9694–9705, 2021.
- [LSD15a] Long, Jonathan; Shelhamer, Evan; Darrell, Trevor: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431–3440, 2015.
- [LSD15b] Long, Jonathan; Shelhamer, Evan; Darrell, Trevor: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431–3440, 2015.
- [MNA16] Milletari, Fausto; Navab, Nassir; Ahmadi, Seyed-Ahmad: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). Ieee, pp. 565–571, 2016.
- [Mo19a] Mo, Kaichun; Zhu, Shilin; Chang, Angel X; Yi, Li; Tripathi, Subarna; Guibas, Leonidas J; Su, Hao: Partnet: A large-scale benchmark for fine-grained and hierarchical part-level 3d object understanding. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 909–918, 2019.
- [Mo19b] Mousavian, Arsalan; Toshev, Alexander; Fišer, Marek; Košecká, Jana; Wahid, Ayzaan; Davidson, James: Visual representations for semantic target driven navigation. In: 2019 International Conference on Robotics and Automation (ICRA). IEEE, pp. 8846–8852, 2019.
- [Mu19] Munappy, Aiswarya; Bosch, Jan; Olsson, Helena Holmström; Arpteg, Anders; Brinne, Björn: Data management challenges for deep learning. In: 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE, pp. 140–147, 2019.
- [MV21] Mohan, Rohit; Valada, Abhinav: Efficienttps: Efficient panoptic segmentation. *International Journal of Computer Vision*, 129(5):1551–1579, 2021.
- [MWA18a] Mahjourian, Reza; Wicke, Martin; Angelova, Anelia: Unsupervised learning of depth and ego-motion from monocular video using 3d geometric constraints. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 5667–5675, 2018.

- [MWA18b] Mahjourian, Reza; Wicke, Martin; Angelova, Anelia: Unsupervised learning of depth and ego-motion from monocular video using 3d geometric constraints. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 5667–5675, 2018.
- [Ne19] Neumann, Maxim; Pinto, Andre Susano; Zhai, Xiaohua; Houlsby, Neil: In-domain representation learning for remote sensing. arXiv preprint arXiv:1911.06721, 2019.
- [Oh18] Oh, Seoung Wug; Lee, Joon-Young; Sunkavalli, Kalyan; Kim, Seon Joo: Fast video object segmentation by reference-guided mask propagation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7376–7385, 2018.
- [Oq23] Oquab, Maxime; Darcet, Timothée; Moutakanni, Théo; Vo, Huy; Szafranec, Marc; Khaidov, Vasil; Fernandez, Pierre; Haziza, Daniel; Massa, Francisco; El-Nouby, Alaaeldin et al.: DINOv2: Learning Robust Visual Features without Supervision. arXiv preprint arXiv:2304.07193, 2023.
- [Pa16] Paszke, Adam; Chaurasia, Abhishek; Kim, Sangpil; Culurciello, Eugenio: Enet: A deep neural network architecture for real-time semantic segmentation. arXiv preprint arXiv:1606.02147, 2016.
- [Po19] Ponciano, J-J; Karmacharya, A; Wefers, S; Atorf, P; Boochs, F: Connected Semantic Concepts as a Base for Optimal Recording and Computer-Based Modelling of Cultural Heritage Objects. In: Structural Analysis of Historical Constructions: An Interdisciplinary Approach. Springer, pp. 297–304, 2019.
- [Po21] Ponciano, Jean-Jacques; Roetner, Moritz; Reiterer, Alexander; Boochs, Frank: Object Semantic Segmentation in Point Clouds—Comparison of a Deep Learning and a Knowledge-Based Method. ISPRS International Journal of Geo-Information, 10(4):256, 2021.
- [PPB21] Ponciano, Jean-Jacques; Prudhomme, Claire; Boochs, Frank: From Acquisition to presentation—The potential of semantics to support the safeguard of cultural heritage. Remote Sensing, 13(11):2226, 2021.
- [PPP21] Plaß, B; Prudhomme, C; Ponciano, JJ: Bim on Artificial Intelligence for Decision Support in E-Health. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 43:207–214, 2021.
- [Qi17a] Qi, Charles R; Su, Hao; Mo, Kaichun; Guibas, Leonidas J: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 652–660, 2017.
- [Qi17b] Qi, Charles Ruizhongtai; Yi, Li; Su, Hao; Guibas, Leonidas J: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Advances in neural information processing systems, 30, 2017.
- [Re16] Redmon, Joseph; Divvala, Santosh; Girshick, Ross; Farhadi, Ali: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779–788, 2016.
- [Re19] Ren, Zhile; Gallo, Orazio; Sun, Deqing; Yang, Ming-Hsuan; Sudderth, Erik B; Kautz, Jan: A fusion approach for multi-frame optical flow estimation. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 2077–2086, 2019.





- [RFB15] Ronneberger, Olaf; Fischer, Philipp; Brox, Thomas: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer, pp. 234–241, 2015.
- [Sa20] Sarlin, Paul-Edouard; DeTone, Daniel; Malisiewicz, Tomasz; Rabinovich, Andrew: Superglue: Learning feature matching with graph neural networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4938–4947, 2020.
- [Sh19] Sharma, Ochin: Deep Challenges Associated with Deep Learning. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). pp. 72–75, 2019.
- [St21] Stone, Austin; Ramirez, Oscar; Konolige, Kurt; Jonschkowski, Rico: The Distracting Control Suite—A Challenging Benchmark for Reinforcement Learning from Pixels. arXiv preprint arXiv:2101.02722, 2021.
- [Su18] Suwajanakorn, Supasorn; Snavely, Noah; Tompson, Jonathan J; Norouzi, Mohammad: Discovery of latent 3d keypoints via end-to-end geometric reasoning. *Advances in neural information processing systems*, 31, 2018.
- [Su19] Sun, Ke; Zhao, Y; Jiang, B; Cheng, T; Xiao, B; Liu, D; Mu, Y; Wang, X; Liu, W; Wang, J: High-resolution representations for labeling pixels and regions. arXiv 2019. arXiv preprint arXiv:1904.04514, 2019.
- [Su23] Su, Yukun; Deng, Jingliang; Sun, Ruizhou; Lin, Guosheng; Su, Hanjing; Wu, Qingyao: A unified transformer framework for group-based segmentation: Co-segmentation, co-saliency detection and video salient object detection. *IEEE Transactions on Multimedia*, 2023.
- [SZ14] Simonyan, Karen; Zisserman, Andrew: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
- [Ta21] Tan, Haobin; Chen, Chang; Luo, Xinyu; Zhang, Jiaming; Seibold, Constantin; Yang, Kailun; Stiefelhagen, Rainer: Flying guide dog: Walkable path discovery for the visually impaired utilizing drones and transformer-based semantic segmentation. In: 2021 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, pp. 1123–1128, 2021.
- [Vo19] Voigtlaender, Paul; Chai, Yuning; Schroff, Florian; Adam, Hartwig; Leibe, Bastian; Chen, Liang-Chieh: Feelvos: Fast end-to-end embedding learning for video object segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9481–9490, 2019.
- [Wa20a] Wang, Jingdong; Sun, Ke; Cheng, Tianheng; Jiang, Borui; Deng, Chaorui; Zhao, Yang; Liu, Dong; Mu, Yadong; Tan, Mingkui; Wang, Xinggang et al.: Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 43(10):3349–3364, 2020.
- [Wa20b] Wang, Jingdong; Sun, Ke; Cheng, Tianheng; Jiang, Borui; Deng, Chaorui; Zhao, Yang; Liu, Dong; Mu, Yadong; Tan, Mingkui; Wang, Xinggang et al.: Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 43(10):3349–3364, 2020.

- [Wa22] Wang, Jianfeng; Yang, Zhengyuan; Hu, Xiaowei; Li, Linjie; Lin, Kevin; Gan, Zhe; Liu, Zicheng; Liu, Ce; Wang, Lijuan: Git: A generative image-to-text transformer for vision and language. arXiv preprint arXiv:2205.14100, 2022.
- [WBP17] Wojke, Nicolai; Bewley, Alex; Paulus, Dietrich: Simple online and realtime tracking with a deep association metric. In: 2017 IEEE international conference on image processing (ICIP). IEEE, pp. 3645–3649, 2017.
- [WH21] Wang, Weijun; Howard, Andrew: MOSAIC: Mobile Segmentation via decoding Aggregated Information and encoded Context. arXiv preprint arXiv:2112.11623, 2021.
- [Xi22] Xiao, Aoran; Huang, Jiaying; Guan, Dayan; Zhan, Fangneng; Lu, Shijian: Transfer learning from synthetic to real lidar point cloud for semantic segmentation. In: Proceedings of the AAAI Conference on Artificial Intelligence. volume 36, pp. 2795–2803, 2022.
- [Yu16] Yu, Licheng; Poirson, Patrick; Yang, Shan; Berg, Alexander C; Berg, Tamara L: Modeling context in referring expressions. In: Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer, pp. 69–85, 2016.
- [Yu20] Yu, Chaohui; Wang, Jindong; Liu, Chang; Qin, Tao; Xu, Renjun; Feng, Wenjie; Chen, Yiqiang; Liu, Tie-Yan: Learning to match distributions for domain adaptation. arXiv preprint arXiv:2007.10791, 2020.
- [Yu21] Yu, Changqian; Gao, Changxin; Wang, Jingbo; Yu, Gang; Shen, Chunhua; Sang, Nong: Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation. International Journal of Computer Vision, 129:3051–3068, 2021.
- [Zh18a] Zhao, Jian; Li, Jianshu; Cheng, Yu; Sim, Terence; Yan, Shuicheng; Feng, Jiashi: Understanding humans in crowded scenes: Deep nested adversarial learning and a new benchmark for multi-human parsing. In: Proceedings of the 26th ACM international conference on Multimedia. pp. 792–800, 2018.
- [Zh18b] Zhou, Zongwei; Rahman Siddiquee, Md Mahfuzur; Tajbakhsh, Nima; Liang, Jianming: Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer, pp. 3–11, 2018.
- [Zh19] Zhou, Bolei; Zhao, Hang; Puig, Xavier; Xiao, Tete; Fidler, Sanja; Barriuso, Adela; Torralla, Antonio: Semantic understanding of scenes through the ade20k dataset. International Journal of Computer Vision, 127:302–321, 2019.
- [Zh20a] Zhao, Sicheng; Yue, Xiangyu; Zhang, Shanghang; Li, Bo; Zhao, Han; Wu, Bichen; Krishna, Ravi; Gonzalez, Joseph E; Sangiovanni-Vincentelli, Alberto L; Seshia, Sanjit A et al.: A review of single-source deep unsupervised visual domain adaptation. IEEE Transactions on Neural Networks and Learning Systems, 33(2):473–493, 2020.
- [Zh20b] Zhao, Sicheng; Yue, Xiangyu; Zhang, Shanghang; Li, Bo; Zhao, Han; Wu, Bichen; Krishna, Ravi; Gonzalez, Joseph E; Sangiovanni-Vincentelli, Alberto L; Seshia, Sanjit A et al.: A review of single-source deep unsupervised visual domain adaptation. IEEE Transactions on Neural Networks and Learning Systems, 33(2):473–493, 2020.

- [Zh20c] Zhao, Yanan; Price, Brian; Cohen, Scott; Gurari, Danna: Objectness-aware few-shot semantic segmentation. arXiv preprint arXiv:2004.02945, 2020.
- [Zh22a] Zheng, Qingping; Deng, Jiankang; Zhu, Zheng; Li, Ying; Zafeiriou, Stefanos: Decoupled Multi-task Learning with Cyclical Self-Regulation for Face Parsing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4156–4165, 2022.
- [Zh22b] Zhou, Jinxing; Wang, Jianyuan; Zhang, Jiayi; Sun, Weixuan; Zhang, Jing; Birchfield, Stan; Guo, Dan; Kong, Lingpeng; Wang, Meng; Zhong, Yiran: Audio–Visual Segmentation. In: Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXVII. Springer, pp. 386–403, 2022.
- [Zh23] Zhou, Jinxing; Shen, Xuyang; Wang, Jianyuan; Zhang, Jiayi; Sun, Weixuan; Zhang, Jing; Birchfield, Stan; Guo, Dan; Kong, Lingpeng; Wang, Meng et al.: Audio-Visual Segmentation with Semantics. arXiv preprint arXiv:2301.13190, 2023.
- [Zo23] Zou, Xueyan; Yang, Jianwei; Zhang, Hao; Li, Feng; Li, Linjie; Gao, Jianfeng; Lee, Yong Jae: Segment everything everywhere all at once. arXiv preprint arXiv:2304.06718, 2023.

**Bildung - Aktuelle Entwicklungen
und Perspektiven (an Hochschulen)
im Bereich Data Science**

Workshop: „Aktuelle Entwicklungen und Perspektiven (an Hochschulen) im Bereich Data Science“


Lena Steinmann ¹, Dirk Nowotka ², Lea Oberländer³, Helen Pfuhl⁴, Heiner Stuckenschmidt ⁵ und Rolf Drechsler ⁶


Abstract: In einer zunehmend von Daten geprägten Welt spielt Data Science eine zentrale Rolle, um insbesondere auch hochdimensionale Daten effektiv zu nutzen und daraus wertvolles Wissen zu generieren. Auch das Hochschulsystem hat auf diese Entwicklung reagiert und in den letzten Jahren wurden verschiedene Forschungsstrukturen, wie zum Beispiel Data Science Center, Beratungsstellen, neue Studiengänge sowie Aus- und Weiterbildungsangebote aufgebaut. Der Workshop fördert einen interdisziplinären und standortübergreifenden Austausch zwischen Vertreter:innen verschiedener Data-Science-Initiativen und legt so die Grundlage für eine zukünftige überregionale Zusammenarbeit. Durch den Austausch rund um innovative Ansätze, bewährte Praktiken und Erfahrungen werden Synergien geschaffen und die Weiterentwicklung des Feldes vorangetrieben. Dabei stehen die Herausforderungen und Chancen der datenintensiven Forschung im Fokus der Diskussionen. Der Workshop bietet eine Plattform, um den Dialog zu fördern, Netzwerke zu stärken und Kompetenzen auf nationaler Ebene zu bündeln.

Keywords: Data Science, Forschungsdatenmanagement, Data Literacy, Hochschulen

1 Hintergrund


Data Science wird als Schlüsseldisziplin des 21. Jahrhunderts angesehen und ist ein zentraler Bestandteil der digitalen Transformation. Mit zukunftsweisenden Verfahren wie maschinellem Lernen, können selbst aus komplexen Daten neue Erkenntnisse gewonnen werden, die als Grundlage für datenbasierte Entscheidungen dienen können. Eine


¹ Universität Bremen, Data Science Center, Bibliothekstraße 5, 28359 Bremen, lena.steinmann@uni-bremen.de,  <https://orcid.org/0000-0001-5443-0581>

² Digital Science Center, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, 24118 Kiel, dn@zs.uni-kiel.de,  <https://orcid.org/000-0002-5422-2229>

³ Mannheim Center for Data Science, Universität Mannheim, L5, 68161 Mannheim, lea.oberlaender@uni-mannheim.de

⁴ Digital Science Center, Christian-Albrechts-Universität zu Kiel, Fraunhoferstr. 13, 24118 Kiel, hpfuhl@email.uni-kiel.de

⁵ Mannheim Center for Data Science / Fakultät für Wirtschaftsinformatik und Wirtschaftsmathematik, Universität Mannheim, L5 / B6, 68161 Mannheim, heiner.stuckenschmidt@uni-mannheim.de,  <https://orcid.org/0000-0002-0209-3859>

⁶ Universität Bremen, Data Science Center, Bibliothekstraße 5, 28359 Bremen, drechsler@uni-bremen.de,  <https://orcid.org/0000-0002-9872-1740>

Grundvoraussetzung dafür sind qualitätsgesicherte, maschinenlesbare Daten, die mittels eines nachhaltigen Forschungsdatenmanagements (FDM) gewährleistet werden.

Auch das Wissenschaftssystem hat auf die wachsende Bedeutung von Data Science als interdisziplinäres Wissenschaftsfeld mit vielfältigen Anwendungsfeldern reagiert. Ein ganz wesentlicher Aspekt dabei ist der steigende Bedarf an Datenkompetenzen in Wissenschaft und Wirtschaft, dem durch neue Qualifizierungswege und Studiengänge nachgekommen werden muss, wie unter anderem von der „Expert Group on FAIR“ der Europäischen Kommission [Eu18], dem Stifterverband [Ki18] und dem Rat für Informationsinfrastruktur (RfII) [Ra19] empfohlen. Die Gesellschaft für Informatik e. V. (GI) hat sich ebenfalls seit 2018 mit der Task-Force "Data Science/Data Literacy" und einhergehenden Workshops intensiv mit diesem Thema befasst und entsprechende Empfehlungen erarbeitet [Ge19, Ge23]. Auch der Fakultätentag Informatik e.V. (FTI) widmete im November 2022 einen Workshop speziell dem Thema Data-Science-Studiengänge. Der Wissenschaftsrat formulierte bereits in 2020 acht Leitlinien für die erfolgreiche Umsetzung datenintensiver Forschung und empfiehlt dabei konkret die Errichtung von Data Science Centern als interdisziplinäre Forschungsstrukturen für die datenintensive Wissenschaft, die auch Möglichkeiten zur Weiterbildung und Kompetenzentwicklung bieten [Wi20].

Viele deutsche Hochschulen haben auf diese Entwicklungen und Empfehlungen bereits reagiert, indem sie verschiedene Forschungsstrukturen, wie Data Science Center, aber auch innovative Studiengänge, Qualifizierungsangebote und Weiterbildungsprogramme geschaffen haben. Gleichzeitig wurden bedeutende Fortschritte im Bereich der Dateninfrastrukturen und des FDM erzielt, darunter der Aufbau der European Open Science Cloud (EOSC) [Eu23], von Gaia-X [Ga23] und der Konsortien der Nationalen Forschungsdateninfrastruktur (NFDI) [Na23]. Diese Initiativen ermöglichen einen verbesserten Zugang zu Forschungsdaten und fördern ihre Nachnutzbarkeit. Darüber hinaus wurde im Jahr 2022 mit der Verstetigung von fünf durch den Bund geförderten KI-Kompetenzzentren durch den Bund die KI-Forschung im deutschen Wissenschaftssystem gestärkt [Bu23]. Ab Dezember 2023 plant das Bundesministerium für Bildung und Forschung (BMBF) zudem die Förderung des Aufbaus von Datenkompetenzzentren in der Wissenschaft [Bu23b]. Diese Zentren dienen als Orte des Lernens, Forschens und Vernetzens und tragen zur Erweiterung der datenwissenschaftlichen Kompetenzen bei. Das Data-Science-Ökosystem der deutschen Wissenschaftslandschaft ist somit vielfältig aufgestellt und bietet ein hohes Synergiepotenzial.

2 Motivation und Ziele des Workshops

Der Workshop „Aktuelle Entwicklungen und Perspektiven (an Hochschulen) im Bereich Data Science“ knüpft an die zuvor beschriebenen Entwicklungen an und hat das Ziel, verschiedene Akteure im Bereich Data Science, insbesondere der Data Science Center, zusammenzubringen und auf operativer Ebene zu diskutieren, wie zukünftige

Zusammenarbeit gefördert und das Feld der Data Science im wissenschaftlichen Kontext weiter vorangetrieben werden kann. Er bietet ein Forum für den interdisziplinären Austausch zwischen Vertreter:innen unterschiedlicher Data-Science-Initiativen. Auf Grund der engen Verbindung von Data Science und FDM sind auch Organisationen und Projekte aus diesem Bereich herzlich willkommen. So wird eine nachhaltige Vernetzung innerhalb der Community ermöglicht und der Grundstein für zukünftige Kooperationen gelegt. Es soll ein Überblick über erfolgreiche Konzepte und Umsetzungsstrategien gegeben und über zukünftige Herausforderungen und Chancen des digitalen Zeitalters diskutiert werden. Dabei soll auch die Rolle von Data Science im Informationszeitalter und ihr Potenzial zur Bewältigung gesellschaftlicher Herausforderungen beleuchtet werden.

3 Thematische Schwerpunkte

Der Workshop bietet eine Plattform für die Vorstellung von Aktivitäten, Konzepten, Erfahrungsberichten, Lösungsansätzen und Best Practices aus verschiedenen Data-Science-Initiativen. Dies geschieht in Form von Impulsvorträgen und Postern. Die eingereichten Beiträge sollen zur Diskussion anregen und die Schaffung von Synergien fördern.

Es sind Beiträge unter anderem, aber nicht ausschließlich, zu folgenden Themenfeldern eingeladen:

1. Data Science Einrichtungen an Hochschulen, z. B. Data Science Center:
 - a) Tätigkeitsschwerpunkte und Ausrichtung (z. B. in den Bereichen Forschung, Trainings, Services, Beratung, Vernetzung, technische Infrastruktur)
 - b) Erfahrungen aus Aufbau und Etablierung der Einrichtung
 - c) Rolle von Datenmanagement und der Nationalen Forschungsdateninfrastruktur (NFDI)
 - d) Besondere Herausforderungen und deren Bewältigung
 - e) Zukünftige Entwicklungen
2. Data Science in Lehre und Weiterbildung:
 - a) Neue und innovative Data-Science-Studiengänge
 - b) Qualifizierungsangebote für Forschende unterschiedlicher Statusgruppen und Disziplinen (z. B. Zertifikatskurse, Micro-Degrees, Graduiertenschulen, modulare Angebote)
 - c) Einbindung anderer Data-Literacy-Aspekte (Datenmanagement, Critical Thinking, ethische, rechtliche und soziale Aspekte)

- d) Lerninhalte und -formate
- e) Nachnutzbarkeit von Konzepten (z. B. als Open Educational Resources)
- f) Rolle von BMBF-Datenkompetenzzentren
- g) Besondere Herausforderungen und deren Bewältigung
- h) Zukünftige Entwicklungen

4 Ausrichtende Institutionen

Der Workshop wird gemeinsam von dem Data Science Center der Universität Bremen (DSC@UB), dem Digital Science Center der Universität Kiel (DSC@CAU) und dem Mannheim Center for Data Science der Universität Mannheim (MCDS) ausgerichtet. Im Folgenden werden diese drei Einrichtungen als Good-Practice-Beispiele kurz vorgestellt.

4.1 Data Science Center der Universität Bremen (DSC@UB)

Das Data Science Center (DSC@UB) wurde 2019 mit Unterstützung vom Land Bremen als interdisziplinäres Institut an der Universität Bremen gegründet und dient als Knotenpunkt für die datenintensive Forschung. Die übergeordneten Ziele des DSC sind die wertschöpfende Nutzung von Daten in allen Forschungsbereichen nachhaltig zu ermöglichen und den kulturellen Wandel im Sinne einer „FAIR-Data-Kultur“⁴⁷ voranzutreiben. Dafür werden die Aspekte FDM und Data Science eng verzahnt und abgestimmte Angebote entwickelt. So ist es möglich, Forschende im gesamten Datenlebenszyklus zu unterstützen und eine bestmögliche Wertschöpfung aus Daten zu erzielen [SD21].

Das DSC@UB bietet umfassende Services für datenintensiv Forschende, darunter Rechenkapazitäten wie GPU Server, finanzielle Unterstützung durch „DSC Seed Grants“, Data-Literacy-Trainings, Unterstützung bei der Implementierung von Data-Science-Methoden wie maschinelles Lernen und Datenmanagement-Beratung durch Data Stewards (siehe auch [St23] für eine detaillierte Beschreibung der Aktivitäten und Services). In seinem interdisziplinären Expert:innen-Netzwerk werden Kompetenzen gebündelt, wodurch die Vernetzung und der Transfer von Wissen über fachliche Grenzen hinweg gefördert und disziplinäre Silos aufgebrochen werden. Dies schafft neue Wege für die kooperative Forschung.

Somit repräsentiert das DSC@UB einen wichtigen strukturellen Pfeiler für die datenintensive Forschung an der Universität Bremen und fungiert als Bindeglied zwischen Forschenden und Infrastruktureinrichtungen wie Bibliotheken oder Rechenzentren.

⁷ Gemäß der FAIR-Prinzipien sollen Forschungsdaten auffindbar (findable), zugänglich (accessible), interoperabel (interoperable) und nachnutzbar (reusable) sein [Wi16].

Außerdem ist es an der Strategieentwicklung an der Universität Bremen und in der U Bremen Research Alliance - dem 2016 gegründeten Kooperationsnetzwerk der Universität Bremen und zwölf außeruniversitärer Forschungsinstitute - beteiligt. Als Beispiel sei hier ein Whitepaper erwähnt, das den aktuellen Stand sowie zukünftige Handlungsfelder für die disziplinübergreifende Etablierung eines kooperativen FDM skizziert [Pi21].

4.2 Digital Science Center der Universität Kiel (DSC@CAU)

Das Digital Science Center an der Universität Kiel wurde 2022 gegründet mit dem Ziel die digitalen Kompetenzen an der Universität zu vernetzen und Methoden gemeinsam mit den Anwendern weiter zu entwickeln. Projekte und Programme rund um datenintensive Forschung finden hier ein gemeinsames Dach zum Netzwerken und Erfahrungsaustausch, sodass Methoden und Erkenntnisse aus der Informatik und der Mathematik zu den Fachbereichen und Anwendungen gebracht werden können. Die intensive und gemeinschaftliche Weiterentwicklung der Datenwissenschaft wird in die Lage versetzt, einen Mehrwert in der Forschung, in der Lehre und für den Wissens- und Technologietransfer zu schaffen.

Auch die Ausbildung der Fachkräfte von morgen ist ein wichtiges Ziel des Digital Science Centers. Hier werden künftig Expertinnen und Experten ausgebildet, die sowohl Kompetenzen in der Methodik der Datenwissenschaft wie auch in einem gewählten Anwendungsfeld haben. Durch die Verschränkung von Methodik und Anwendung in Fachdisziplinen aller Fakultäten werden die Potenziale der Datenwissenschaften in die gesamte Universität und darüber hinaus getragen. Studiengänge nach dem Prinzip der 2-Fächer-Kombination von Datenwissenschaften und Anwendungsfach mit einer Profilierung im Bereich Künstlicher Intelligenz, werden ergänzt durch Angebote für Ausgründungen und Start-ups mit einem KI-Fokus, genauso wie für Unternehmen mit einem Interesse an KI-Anwendungen.

4.3 Mannheim Center for Data Science der Universität Mannheim (MCDS)

Das Mannheim Center for Data Science (MCDS) bündelt die Expertise der Universität Mannheim in den Wirtschafts-, Sozial- und Verhaltenswissenschaften. Es wurde mit Beschluss des Senats vom 4. März 2020 als Institut der Universität Mannheim eingerichtet. Mitglieder des Gründungsdirektoriums sind Prof. Dr. Frauke Kreuter, Prof. Dr. Florian Stahl, Prof. Dr. Heiner Stuckenschmidt und Prof. Dr. Jutta Mata. Die Profilprofessur „Data Science in den Wirtschafts- und Sozialwissenschaften“, die Prof. Dr. Markus Strohmaier seit Januar 2022 innehat, kooperiert ebenfalls eng mit dem MCDS.

Die Mission des MCDS ist es, ein besseres Verständnis für Daten und ihren Wert für die Erforschung der digitalen Transformation von Individuum, Wirtschaft und Gesellschaft

zu schaffen. Dabei müssen Regeln für den verantwortungsvollen Umgang mit Daten und Analysemethoden zum Nutzen der Gesellschaft identifiziert werden. Die Ermöglichung innovativer, interdisziplinärer Forschung, die gesellschaftliche Prozesse und globale Herausforderungen in den Blick nimmt, die auch für lokale Institutionen und Unternehmen relevant sind, ist ein weiteres Anliegen des MCDS.

Das MCDS versteht sich als Plattform zur Vernetzung und Stärkung der Forschung zu und mit Data Science an der Universität Mannheim und darüber hinaus. In dieser Funktion organisiert es Lehrveranstaltungen, Vernetzungsevents und Vorträge und leistet einen Beitrag zur Wissenschaftskommunikation, indem es regelmäßig über neue Projekte und Fortschritte im Bereich Data Science an der Universität Mannheim informiert. Es ist dabei eng in den Struktur- und Entwicklungsplan der Universität eingebunden, der die Stärkung von Data Science in Lehre und Forschung verankert.

5 Workshop-Beiträge

Der Workshop umfasst drei geladene Keynotes zu den Themenbereichen (1) Data Science Einrichtungen an Hochschulen (Prof. Dr. Philipp Wieder), (2) Data Science in der Lehre (Prof. Dr. Peer Kröger) und (3) Data Science und Datenmanagement / NFDI (Dr. Sonja Schimmler). Die dazugehörigen Sessions werden mit thematisch passenden Impulsvorträgen ausgestaltet. Für den weiteren Austausch wird es eine Poster-Session geben.

Insgesamt wurden acht Beiträge zur Publikation und weitere Beiträge zur Beteiligung ohne Veröffentlichung eingereicht. Dabei ist eine Vielzahl von Data-Science-Einrichtungen vertreten, die wir hier in alphabetischer Reihenfolge auflisten:

- Bielefeld Center for Data Science, Universität Bielefeld
- Campus-Institut Data Science, Georg-August-Universität Göttingen
- Center for Interdisciplinary Digital Sciences, Technische Universität Dresden
- Data Science Center, Universität Bremen
- Digital Science Center, Universität Kiel
- House of Computing and Data Science, Universität Hamburg
- Mannheim Center for Data Science, Universität Mannheim
- Munich Data Science Institute, Technische Universität München

Darüber hinaus sind vielfältige Beiträge aus den Bereichen Lehre und Weiterbildung dabei, z. B. zu der BERD (*Business, Economic, and Related Data*) Academy, dem Projekt DALIA (*Data Literacy Alliance*), der Helmholtz School for Marine Data Science (*MarDATA*), dem Mannheim Master in Social Sciences, dem Zertifikatskurs

„Forschungsdatenmanagement für Studierende“, der NFDI4Earth Academy, dem Master of Science „Data Science“ der FU Berlin und der interdisziplinären Promovierenden-Ausbildung „Data Train – Training in Forschungsdatenmanagement und Data Science“ der U Bremen Research Alliance. Auch aus der Schnittmenge zwischen Data Science und Datenmanagement sind mit NFDI4DataScience, dem Projekt „Forschungsdatenmanagement an Hochschulen für angewandte Wissenschaften in Rheinland-Pfalz“ (FDM@HAW.rlp) und dem Projekt „institutionalisiertes und nachhaltiges Forschungsdatenmanagement in Brandenburg“ (IN-FDM-BB) Initiativen vertreten.

6 Literaturverzeichnis

- [Eu18] European Commission, Directorate-General for Research and Innovation, Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data, Brüssel, 2018.
- [Ki18] Kirchherr, J. et al.: Future Skills: Welche Kompetenzen in Deutschland fehlen, Stifterverband für die Deutsche Wissenschaft, Essen, 2018.
- [Ra19] Rat für Informationsinfrastrukturen: Digitale Kompetenzen - dringend gesucht! Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt Wissenschaft, Göttingen, 2019.
- [Ge19] Gesellschaft für Informatik e.V. (GI), Data Science: Lern- und Ausbildungsinhalte, Berlin, 2019.
- [Ge23] Gesellschaft für Informatik e.V. (GI), <https://gi.de/datascience>, Stand: 14.06.2023.
- [Wi20] Wissenschaftsrat: Zum Wandel in den Wissenschaften durch datenintensive Forschung, Köln, 2020.
- [Eu3] European Open Science Cloud (EOSC), <https://eosc-portal.eu/>, Stand: 14.06.2023.
- [Ga23] Gaia-X, <https://gaia-x.eu/>, Stand: 14.06.2023.
- [Na23] Nationale Forschungsdateninfrastruktur (NFDI), <https://www.nfdi.de/>, Stand: 14.06.2023.
- [Bu23a] Bundesministerium für Bildung und Forschung (BMBF), <https://www.bmbf.de/bmbf/shareddocs/kurzmeldungen/de/2022/06/50-millionen-foerderung-fuer-ki-kompetenzzentren.html>, Stand: 14.06.2023.
- [Bu23b] Bundesministerium für Bildung und Forschung (BMBF), https://www.bildung-forschung.digital/digitalezukunft/de/wissen/Datenkompetenzen/datenkompetenzzentren_f%C3%BCr_die_wissenschaft_ordner/datenkompetenzzentren_fuer_die_wissenschaft_node.html, Stand: 14.06.2023.
- [Wi16] Wilkinson, M.D. et al.: The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data 3, 2016. DOI: 10.1038/sdata.2016.18.

- [SD21] Steinmann, L., Drechsler, R.: Verzahnung von Data Stewardship und Data Science – Wege und Perspektiven, Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managern Nr. 3, S. 82-91, 2021. DOI: 10.17192/bfdm.2021.3.8342.
- [St23] Steinmann, L. et al.: Das Data Science Center an der Universität Bremen: Interdisziplinärer Knotenpunkt und Service-Infrastruktur für die datenintensive Forschung, in: Heuveline, Vincent und Bisheh, Nina (Hrsg.): E-Science-Tage 2023: Share Your Research Data, Heidelberg: heiBOOKS, (in review).
- [Pi21] Pigeot I. et al.: Etablierung eines kooperativen Forschungsdatenmanagements in der U Bremen Research Alliance, Zenodo, 2021, DOI: 10.5281/zenodo.4775371.

Der Mannheim Master in Social Data Science

Thomas Gautschi¹, Thomas Gschwend², Lea Oberländer³, Simone Ponzetto⁴ und Heiner Stuckenschmidt⁵

Abstract: In den empirisch geprägten Sozialwissenschaften besteht zunehmender Bedarf nach wissenschaftlichem Nachwuchs mit starken Methodenkompetenzen. Hierbei stehen nicht mehr nur klassische Umfragemethoden und parametrische Statistik im Vordergrund. Zunehmen werden auch neue Datentypen wie Texte und Bilder sowie vereinzelt auch Video und Audio verwendet. Hierdurch entsteht ein Bedarf nach Methoden aus den Bereichen Data Science und Maschinellem Lernen. Der Mannheim Master in Social Data Science soll diesen Bedarf gezielt bedienen und die nächste Generation von Sozialwissenschaftlerinnen und Sozialwissenschaftlern mit fundierten Data Science Kenntnissen ausbilden. Hierbei wird besonderer Wert auf die enge Verzahnung von inhaltlichen Fragestellungen und klassischen Forschungsmethoden gelegt.

Keywords: Data Science Ausbildung, Social Data Science


1 Hintergrund


Die Universität Mannheim weist von ihrer Entstehung, Entwicklung und aktuellen Struktur her ein besonderes Profil auf. Dieses Profil wird geprägt von den weithin anerkannten Wirtschafts- und Sozialwissenschaften und ihrer Vernetzung mit leistungsstarken Geistes- und Rechtswissenschaften sowie mit Mathematik und Informatik. In Mannheim existiert ein interdisziplinär strukturiertes System in Forschung und Lehre. Die wechselseitige Ergänzung aller Fächer ist das Alleinstellungsmerkmal der Universität und sichert den anhaltenden Erfolg ihrer

¹ Universität Mannheim, Professur für Methoden der Empirischen Sozialforschung, A5, 6, 68159 Mannheim, gautschi@uni-mannheim.de

² Universität Mannheim, Professur für Politikwissenschaft, quantitative sozialwissenschaftliche Methoden, A5, 6, 68159 Mannheim, gschwend@uni-mannheim.de,  <https://orcid.org/0000-0002-8656-9622>

³ Universität Mannheim, Mannheim Center for Data Science, L5,5, 68161 Mannheim, datascience@uni-mannheim.de

⁴ Universität Mannheim, Lehrstuhl für Informationssysteme III: Unternehmensdatenanalyse, B6, 26, 68159 Mannheim, simone@informatik.uni-mannheim.de,  <https://orcid.org/0000-0001-7484-2049>

⁵ Universität Mannheim, Lehrstuhl für Praktische Informatik II: Künstliche Intelligenz, B6, 26, 68159 Mannheim, heiner.stuckenschmidt@uni-mannheim.de,  <https://orcid.org/0000-0002-0209-3859>

Wissenschaftlerinnen und Wissenschaftler und Absolventinnen und Absolventen im nationalen und internationalen Wettbewerb. Im Rahmen der Struktur- und Entwicklungsplanung 2024 – 2028 hat sich die Universität das Ziel gesetzt, die empirische Forschung, speziell in den Wirtschafts- und Sozialwissenschaften weiter zu stärken und hierbei gezielt den Einsatz neuer Methoden der Datenanalyse aus dem Bereich des Maschinellen Lernens zu fördern, um international, wo ähnliche Initiativen zu beobachten sind, weiterhin auf höchstem Niveau konkurrenzfähig zu bleiben.

Ein wesentlicher Faktor ist hierbei die gezielte methodische Ausbildung des wissenschaftlichen Nachwuchses. Der Ausbau des Lehrangebots im Bereich „Social and Economic Data Science“ auf Master Niveau ist eine wichtige Maßnahme zur Sicherstellung dieser Ausbildung. Das Land Baden-Württemberg stellt der Universität Mannheim zu diesem Zweck jährliche Mittel von etwa 800.000 Euro zur Verfügung, mit deren Hilfe der bestehende Studiengang Mannheim Master in Data Science aus- und der beantragte Studiengang aufgebaut werden soll. Aus diesen Mitteln wurden neue Lehrstühle für Maschinelles Lernen (Fakultät für Wirtschaftsinformatik und Wirtschaftsmathematik (WIM)) und Social Data Science (Fakultät für Sozialwissenschaften (SOWI)) geschaffen. Letzterer wurde thematisch so ausgeschrieben, dass er den geplanten Studiengang optimal unterstützen kann, und soll eine zentrale Rolle bei der Einrichtung und Durchführung des Studiengangs bilden. Mit Marc Ratkovic konnte ein international ausgewiesener Experte aus Princeton für diese Stelle gewonnen werden.

Aktuell laufen Planungen für eine stärkere Integration des Themas Data Science in die Promovierendenausbildung im Rahmen der Graduiertenschule CDSS (Center for Doctoral Studies in Social and Behavioral Sciences). Ziel ist es, Absolventinnen und Absolventen des geplanten Studiengangs eine attraktive Perspektive für eine wissenschaftliche Karriere zu bieten und so methodisch umfassend ausgebildete Nachwuchswissenschaftler und Nachwuchswissenschaftlerinnen aus den Sozialwissenschaften an die Universität Mannheim zu binden.

2 Qualifikationsziele

2.1 Die GI Empfehlung

Die Gesellschaft für Informatik hat als Reaktion auf die zunehmende Nachfrage nach Data Science Inhalten in der universitären Lehre eine Vorstands Task Force 'Data Science / Data Literacy' gegründet, die sich seit inzwischen mehr als fünf Jahren mit dem Thema Data Science Ausbildung beschäftigt. Aktuelles Ergebnis ist eine Empfehlung für die Gestaltung von Data Science Studiengängen [Ab21]. Da diese jedoch explizit Studierende der Informatik und Wirtschaftsinformatik als Zielgruppe hat, haben wir uns entschlossen, das zwei Jahre zuvor erschienene Arbeitspapier Lern-

und Ausbildungsinhalten [Vo19] als Grundlage für die Entwicklung des Studiengangs zu verwenden, der sich an Sozialwissenschaftler und Sozialwissenschaftlerinnen richten soll. Dieses hat den Vorteil, dass zwischen Personas mit unterschiedlichem Vorwissen und fachlichem Hintergrund unterschieden wird.

2.2 Profil des MMDS

Gemäß den oben beschriebenen strategischen Zielen der Universität ist das Qualifikationsziel des geplanten Studiengangs die frühe methodische Ausbildung des wissenschaftlichen Nachwuchses in den Sozialwissenschaften, speziell der Soziologie und der Politikwissenschaft. Hierbei sollen traditionelle methodische Ansätze, die in Mannheim in der Regel bereits in den entsprechenden Bachelorstudiengängen unterrichtet werden, um fortgeschrittene Methoden aus den Bereichen Statistik und Maschinelles Lernen erweitert werden. Zentrale Qualifikationsziel ist die Fähigkeit, Methoden des Maschinellen Lernens aus der Informatik mit den Prinzipien kausaler Inferenz zu vereinen und in Fachkontexten anwenden zu können, um praktische Fragestellungen in Forschung und Praxis beantworten zu können. Dieses Qualifikationsziel spiegelt sich in dem zentralen Pflichtkurs „Machine Learning and Causal Inference“ wider. Das restliche Curriculum wurde um diesen zentralen Kurs herum geplant. Die Struktur des Studiengangs orientiert sich hierbei an den Empfehlungen der Gesellschaft für Informatik für die Data Science Ausbildung und adressiert die dort aufgeführte Persona B: „Persona B hat Kompetenzen eines Bachelors in einer Domänenwissenschaft (dies kann ein technisches oder naturwissenschaftliches Fach sein, aber auch ein Fach im Bereich der Geistes- bzw. Kulturwissenschaften) und will Data-Science-Kompetenzen für die Domäne erwerben.“ Da der geplante Studiengang gemeinsam von den Fakultäten WIM und SOWI getragen wird, entsprechen die Inhalte einer Mischung aus den Empfehlungen für die Personas B1 (Master in Data Science) und B2 (Master in Domäne mit Data Science Schwerpunkt). Es ergibt sich das folgende Raster bzgl. der vermittelten Kompetenzen:

Kompetenzfelder	Pflicht- Wahlinhalte	Lernziel
(1) Grundlagen Mathematik	Statistik	Voraussetzung
(2) Fortgeschrittene Mathematik und Statistik	Pflicht	L2-L3
(3) Grundlagen der Informatik	Pflicht	L2
(4) Fortgeschrittene Informatik	Wahl	L1
(5) Kryptographie und Sicherheit	-	-

(6) Datenethik und Data Privacy	Pflicht	L1
(7) Data Governance	Pflicht	L2
(8) Datenintegration	Wahl	L2
(9) Datenvisualisierung	-	-
(10) Data Mining	Wahl	L2
(11) Maschinelles Lernen/Deep Learning	Pflicht Wahl	L2-L3
(12) Business Intelligence	-	-
(13) Domänenspezifische Anwendungen	Pflicht	L2-L3
(14) Data Science in der Organisation	-	-

Tab. 1: Profil des MMSDS Studiengangs

Der klare Schwerpunkt liegt hierbei auf der Entwicklung eines tiefen Verständnisses für fortgeschrittene Statistik und deren Verbindung zu Methoden des Maschinellen Lernens sowie deren Anwendung in domänenspezifischen Anwendungen. Um dieses Anspruchs- volle Ziel erreichen zu können, beinhaltet der Studiengang einen hohen Anteil an Pflichtveranstaltungen, um die notwendigen mathematischen und statistischen Kenntnisse aufzubauen, die hierfür notwendig sind. Diese werden ergänzt durch ein breites Wahlangebot im Bereich Datenanalysemethoden für spezielle Datentypen wie Text, Bilder oder Netzwerke, die je nach Anwendung notwendig sind, sowie Kurse zu Privacy und Open Science. Da sich der Studiengang wie im Abschnitt zur Zielgruppe primär an den wissenschaftlichen Nachwuchs richtet werden Inhalte, die sich mit der Anwendung von Data Science Methoden in der Wirtschaft beschäftigen nur am Rand behandelt.

2.3 Das Curriculum

Das Curriculum besteht aus vier Teilen plus abschließender Masterarbeit [Fa23b].

Data Science Grundlagen In diesem Teilbereich werden grundlegende Fähigkeiten und Kenntnisse aus Sozialwissenschaften und Informatik, sowie Regeln guter wissenschaftlicher Praxis vermittelt. Er besteht aus den folgenden Pflichtkursen:

Kurs	ECTS	Anbieter
Statistics for Social Scientists	9	Sozialwissenschaften
Database Technologies	6	Informatik
Programming for Data Scientists	6	<i>Informatik</i>
Legal and Ethical Aspects of Privacy	3	Rechtswissenschaften
Open Science and Reproducibility	3	Open Science Office

Tab. 2: Data Science Grundlagen

Hierbei wird Wert daraufgelegt, Grundlagen aus unterschiedlichen Bereichen zu vermitteln und den Grundstein für eine interdisziplinäre Sicht auf das Thema zu vermitteln.

Data Science Methoden - Pflichtbereich Im Methoden Pflichtteil liegt der Schwerpunkt auf neusten Methoden aus dem Bereich der Sozialwissenschaften. Wie erwähnt steht hierbei die Verbindung von Maschinellem Lernen und kausaler Inferenz als Schlüsselveranstaltung im Mittelpunkt.

Kurs	ECTS	Anbieter
Sampling and Data	9	Sozialwissenschaften
Machine Learning and Causal Inference	9	Sozialwissenschaften

Tab. 3: Data Science Methoden - Pflichtbereich

Der thematisch unterspezifizierte Kurs 'Current Topics' soll es ermöglichen, neue Entwicklungen im Methodenbereich und aktuelle Forschungsthemen der Universität in den Studiengang einfließen zu lassen, ohne das Curriculum zu ändern.

Data Science Methoden - Wahlbereich Während der Methoden-Pflichtbereich sozialwissenschaftliche Methoden der Datenanalyse als Schwerpunkt hat, bietet der Wahlbereich die Möglichkeit, Kurse aus dem bestehenden Masterstudiengang 'Mannheim Master in Data Science' [Fa23a] zu belegen. Die Studierenden können hierbei drei beliebige Kurse aus dem Bereich Datenanalysemethoden wählen, besonders empfohlen werden jedoch die Kurse Machine Learning, Deep Learning, Text Analytics, Computer Vision und Network Science aus dem Lehrangebot der Informatik. Zusätzlich können auch Data Mining Kurse oder Kurse zum Thema Datenmanagement und Datenintegration belegt werden. Studierende sollen hier idealerweise eine Kombination wählen, die für die von ihnen gewählten Anwendungskurse bzw. die angestrebte Masterarbeit nützlich sind.

Data Science Anwendungen Im Anwendungsbereich sollen die Studierenden lernen, die Verbindung zwischen Methoden und wissenschaftlichen Fragestellungen in den Sozialwissenschaften herzustellen. Hierzu besteht eine breite Auswahl von Kursen vor allem aus den Fächern Soziologie, Politikwissenschaften sowie Medien- und Kommunikationswissenschaft, von denen drei Kurse belegt werden müssen. Grundsätzlich sind hier auch substanzielle Kurse aus anderen Mannheimer Fächern denkbar, so dass sich das Angebot sicher weiterentwickeln wird.

3 Zielgruppe

Mit der Einführung des Studiengangs wird das Angebot der Data Science Ausbildung neu strukturiert. Das Angebot besteht dann aus den Studiengängen

- Wirtschaftsinformatik mit einem Schwerpunkt in Data Science für Absolventen und Absolventinnen von Bachelorstudiengängen in Wirtschaftswissenschaften und Wirtschaftsinformatik,
- den Mannheim Master in Data Science für Absolventinnen und Absolventen der (Wirtschafts-) Informatik und (Wirtschafts-) Mathematik, sowie
- den geplanten Mannheim Master in Social Data Science für Absolventinnen und Absolventen von Bachelorstudiengängen in Sozialwissenschaften, speziell den in Mannheim vertretenen Fächern Soziologie, Politikwissenschaft, Psychologie und Medien- und Kommunikationswissenschaft.

Hierbei zielt der geplante Studiengang primär auf den wissenschaftlichen Nachwuchs und soll auf eine Promotion in den Sozialwissenschaften vorbereiten, die in Mannheim traditionell stark empirisch und quantitativ geprägt ist. Dementsprechend ist der Studiengang mit einer sehr kleinen Kohorte von 15 Studierenden pro Jahr geplant und stellt relativ hohe Anforderungen an die Bewerberinnen und Bewerber, speziell bezüglich methodischer Vorkenntnisse. Die Anforderungen sind so definiert, dass sich Absolventinnen und Absolventen der Studiengänge in den Mannheimer Sozialwissenschaften und vergleichbare Studiengänge anderer Universitäten bewerben können, sofern sie eine Bachelorarbeit mit quantitativer Ausrichtung vorweisen können. Diese Absolventinnen und Absolventen stellen die primäre Zielgruppe für den geplanten Studiengang dar.

Ein wesentliches Merkmal ist die interdisziplinäre Ausrichtung des Studiengangs. Der Studiengang wird von den Fakultäten WIM und SOWI gemeinsam getragen. Dementsprechend bestehende kooperative Governance Strukturen, insbesondere eine gemeinsame Studienkommission sowie ein gemeinsamer Prüfungsausschuss, bei dem der oder die Vorsitzende aus den Sozialwissenschaften kommt, um die Interessen und Bedürfnisse von Studierenden mit sozialwissenschaftlichem Hintergrund optimal adressieren zu können. Aus dem gleichen Grund wird die Mehrzahl der Pflichtveranstaltungen von Lehrenden der SOWI Fakultät durchgeführt. Trotz der

starken Einbindung der SOWI Fakultät wurde beschlossen, den Studiengang an der WIM Fakultät zu verorten, um die Data Science Ausbildung der Universität Mannheim an einer Fakultät zu bündeln und so eine bessere Abstimmung der verschiedenen relevanten Studiengänge zu gewährleisten.

4 Herausforderungen

Die Einführung des MMDS ist ein ambitioniertes Projekt und sieht sich einer Reihe von Herausforderungen gegenüber. Diese betreffen zunächst den Anspruch, einen forschungszentrierten Studiengang zu etablieren, dessen klares Ziel die Ausbildung der nächsten Generation von Professorinnen und Professoren ist. Dies kann nur gelingen, wenn das Niveau des Studiengangs von Anfang an sehr hoch ist. Hierzu muss zunächst sichergestellt werden, dass nur Studierende aufgenommen werden, die die notwendigen Qualifikationen mitbringen. Erfahrungen mit anderen Studiengängen zeigen, dass dies innerhalb des gegebenen rechtlichen Rahmens oft nur sehr schwer möglich ist. Bachelornoten sind trotz aller Bemühungen einer Standardisierung kaum vergleichbar, zudem wird die Kohorte zwangsläufig heterogen sein, da diese sich aus verschiedenen Fächern rekrutiert. Um dieses Problem etwas zu entschärfen, wurde das Konzept der unterschiedlichen Studiengänge für verschiedene Zielgruppen eingeführt, so dass sich die Bewerberinnen und Bewerber ganz überwiegend aus den Sozialwissenschaften rekrutieren werden.

Die Unterteilung in verschiedene Studiengänge für unterschiedliche Zielgruppen kann das Problem unterschiedlicher Anforderungen der Zielgruppen allerdings nur teilweise lösen. Aufgrund der begrenzten Ressourcen der Fakultät sind die meisten Methodenkurse aus der Informatik Teil aller drei Data Science Programme, so dass die Heterogenität zwar in den Programmen etwas reduziert ist, in den einzelnen Vorlesungen jedoch wieder voll durchschlägt. Zudem unterscheiden sich die Ausbildungsziele auch konzeptionell. So bildet der Master Wirtschaftsinformatik vor allem Nachwuchs für die Wirtschaft aus – der Hauptsitz von SAP ist nur eine halbe Stunde Fahrt von Mannheim entfernt. Eine weitere Herausforderung ist die Tatsache, dass die Programme, in denen die gleichen Veranstaltungen durchgeführt werden, wiederum in Zusammenarbeit mit unterschiedlichen Fakultäten, insbesondere der Fakultät für Betriebswirtschaftslehre auf der einen und der Sozialwissenschaftlichen Fakultät auf der anderen Seite, betrieben werden. Hierdurch entsteht nicht nur die Situation, dass unterschiedliche inhaltliche Anforderungen bedient werden müssen, diese müssen auch noch mit Vertreterinnen und Vertretern unterschiedlicher Fakultäten abgestimmt werden. Es bestehen zwar langjährige Erfahrungen in dieser Richtung, die konkrete Umsetzung stellt allerdings immer aufs Neue eine Herausforderung dar.

Insgesamt bleibt die Frage zu beantworten, ob sich die gewählte Aufteilung in unterschiedliche, auf verschiedene Zielgruppen zugeschnittene Aufteilung bewährt,

oder ob ein einzelner Data Science Studiengang mit unterschiedlichen Wahlfächern nicht das bessere Konzept wäre.

5 Fazit

Mit dem Mannheim Master in Social Data Science planen wir einen Studiengang, der in dieser Form mit Ausnahme des Master Social and Economic Data Science an der Universität Konstanz in Deutschland wenig Konkurrenz haben dürfte. Er verbindet hier den quantitativen Fokus der Mannheimer Sozialwissenschaften mit dem Datenanalyse Schwerpunkt der Mannheimer (Wirtschafts-)Informatik und passt daher ideal in das Portfolio der Universität. Ein Blick ins Ausland zeigt, dass wir hiermit keineswegs Vorreiter sind, sondern die Einführung eher einen notwendigen Schritt darstellt, um international konkurrenzfähig zu bleiben.

Literaturverzeichnis

- [Ab21] Abedjan, Z.; Bendig, T.; Brefeld, U.; Bürkle, J.; Desel, J.; Edlich, S.; Eppler, T.; Goedicke, M.; Hachmeister, N.; Heidrich, J.; Höppner, S.; Kast, S. M.; Krupka, D.; Lang, Klaus; Liggesmeyer, P.; Meisner, J.; Scholtes, I.; Tropmann-Frick, M.: Empfehlungen für Masterstudiengänge „Data Science“ – auf Basis eines Bachelors in (Wirtschafts-)Informatik oder Mathematik. Bericht, 2021.
- [Fa23a] Fakultät für Wirtschaftsinformatik und Wirtschaftsmathematik: Master of Science 'Mannheim Master in Data Science'. <https://www.wim.uni-mannheim.de/studium/studienorganisation/mannheim-master-in-data-science/>, 2023.
- [Fa23b] Fakultät für Wirtschaftsinformatik und Wirtschaftsmathematik: Master of Science Mannheim Master in Social Data Science'. tbd., 2023.
- [Vo19] Vorstands Task Force Data Science Data Literacy: Data Science Lern- und Ausbildungsinhalte. Arbeitspapier, Gesellschaft für Informatik, 2019.

BERD Academy: Data Science Kompetenzen für Forschende in Business, Economics and Related Sciences

Markus Herklotz¹ und Wiebke Weber²

Abstract: Mit der BERD Academy wird seit 2022 ein auf die speziellen Charakteristika von (unstrukturierten) wirtschafts- und sozialwissenschaftlichen Daten abgestimmtes Trainingsprogramm entwickelt. Da insbesondere neuere Datenquellen und deren Methoden noch nicht immer Teil der universitären Curricula sind, spricht das Programm vor allem Forschende und Nachwuchswissenschaftler*innen an, die ihr Kompetenzportfolio damit erweitern möchten. Dadurch soll die BERD Academy die wissenschaftlich fundierte Anwendung von Data Science Methoden in den entsprechenden Disziplinen stärken. Das auf die geringen Zeitressourcen von Forschenden ausgerichtete, flexible Angebot rangiert von kleineren Einheiten wie Coffee Lectures bis hin zu mehrwöchigen Flipped Classroom Kursen.

Keywords: Data Science Education, Forschungsdatenmanagement, Social Data Science, Economics and Business Sciences, Unstructured Data, Open Science, Nationale Forschungsdateninfrastruktur

Data Science ist eines der am stärksten wachsenden Berufsfelder. Auf dem US-Arbeitsmarkt beispielsweise liegt der Zuwachs an prognostizierten Stellen nicht nur deutlich über dem Durchschnitt aller anderen Berufsfelder insgesamt, sondern auch höher als wissenschaftliche mathematische Berufe im Speziellen.³ Diese wachsende Nachfrage nach Data Science Kompetenzen offenbart eine Lücke im Angebot, das heißt beim qualifizierten wissenschaftlichen Nachwuchs. Gerade in den Sozial- und Wirtschaftswissenschaften ist eine quantitative Grundausbildung ein etablierter Bestandteil des Studiums, aber dieser beruht “meist immer noch auf dem Arbeiten mit Daten aus traditionellen Erhebungsmethoden wie Befragungen und Beobachtung“ [KK20, S. 14]. Neuere Datentypen wie bspw. von Social Media und Smartphone-Sensoren können ein hohes wissenschaftliches Wertschöpfungspotential bergen, finden aber weniger Beachtung. Zwar passen Hochschulen ihre Curricula an, aber im volatilen und wachsenden Gebiet Data Science ist es eine Herausforderung, mit den neusten Entwicklungen Schritt zu halten. Gleichzeitig wirft die Anpassung der Studiengänge Fragen nach der Übertragbarkeit von empirischen Standards (bspw. in der Stichprobenziehung) auf Data Science Methoden sowie deren ethischen und rechtlichen Aspekte auf.

Neben der Modernisierung der Data Science Grundausbildung in den Sozial- und Wirtschaftswissenschaften für neue Kohorten von Nachwuchswissenschaftler*innen, kann vor allem ein Weiterbildungsbedarf bei bereits mit Daten arbeitenden Forschenden ausgemacht

¹ Ludwig-Maximilians-Universität München, Institut für Statistik, Ludwigstr. 33, 80539 München, Deutschland
m.herklotz@lmu.de

² Ludwig-Maximilians-Universität München, Institut für Statistik, Ludwigstr. 33, 80539 München, Deutschland
wiebke.weber@lmu.de

³ <https://www.bls.gov/ooh/math/data-scientists.htm> (Stand: 26.05.2023)

werden. Diese haben die traditionelle Grundausbildung abgeschlossen und vermutlich einen großen Teil ihres nun benötigten Kompetenzportfolios 'on the job' erworben. Andere Sozial- und Wirtschaftswissenschaftler*innen arbeiten ausschließlich mit den traditionellen Methoden der Umfrageforschung und sind sich über das etwaige Potenzial von Big Data und Machine Learning für die eigene Forschung nicht bewusst.

An diese Gruppen richtet sich das Weiterbildungsangebot der BERD Academy. Das Trainingsprogramm des Konsortiums für Business, Economics and Related Data in der Nationalen Forschungsdateninfrastruktur (BERD@NFDI⁴) hat das Ziel, Forschende aus den Wirtschafts- und Sozialwissenschaften für die Arbeit mit – insbesondere – unstrukturierten Daten zu befähigen. Damit werden sowohl Personen angesprochen, die bereits mit diesen Daten arbeiten und sich mit wissenschaftlich fundierten Angeboten weiterentwickeln wollen, als auch jene, die damit noch keinen Kontakt hatten. Ein weiterer wichtiger Aspekt, der beide Gruppen betrifft, ist nicht nur Wert auf Erhebungs- und Auswertungsmethoden zu legen, sondern vor allem auch das Forschungsdatenmanagement zu fördern. Dies bietet die direkte Schnittstelle zu den Forschungsdatenzentren mit der Expertise auf diesem Feld und gewinnt mit zunehmendem Umfang und Vielfalt der Daten noch zusätzlich an Bedeutung. In jedem Falle sollen so Barrieren bei der Arbeit mit unstrukturierten Daten abgebaut und bisher nicht genutzte Datenpotentiale ausgereizt werden. Um dies zu erreichen, fokussieren die Angebote der BERD Academy auf ein anwendungsorientiertes Training, damit die Teilnehmenden die Inhalte und Kompetenzen direkt auf tatsächliche Aufgaben in ihrer Forschung übertragen können.

Um den limitierten Zeit-Ressourcen der berufstätigen Forschenden gerecht zu werden, bietet die BERD Academy eine Vielfalt von Formaten an, die den jeweiligen Inhalten und Zielgruppen angepasst sind. Es gibt sowohl vollständig asynchrone Kurse, welche die Teilnehmenden selbstständig in ihrem eigenen Rhythmus absolvieren, als auch synchrone Workshops, bei denen das Hauptaugenmerk auf gemeinsamen Austausch und Interaktion liegt. Eine weitere wichtige Säule sind Flipped Classroom Kurse, die asynchrones und synchrones Lernen verbinden. Dabei arbeiten die Teilnehmenden in ihrem eigenen Rhythmus durch die wöchentlichen Materialien (Videos, Literatur, Aufgaben) und treffen sich für ein etwa einstündiges Online-Meeting mit den Lehrenden, um die aktuellen Themen zu diskutieren. Diese Umsetzung der Flipped Classroom Kurse konnte bereits für die gleichen Zielgruppen im BMBF-geförderten Weiterbildungsangebot International Program in Survey and Data Science (IPSDS) umfangreich erprobt werden und hat sich für berufstätige Lernende als besonders geeignet herausgestellt [Ha21].

⁴ Die Nationale Forschungsdateninfrastruktur (NFDI) soll die langfristige Nutzbarkeit von Daten stärken, indem die Verfügbarkeit und der Austausch von Forschungsdaten gefördert werden. Die von Bund und Ländern finanzierte Initiative bildet sich durch Konsortien, welche die verschiedenen Fachbereiche vertreten und durch den Aufbau von Services den Datenaustausch in ihren jeweiligen Disziplinen stärken. BERD@NFDI ist das Konsortium für wirtschafts- und sozialwissenschaftliche Daten mit einem Fokus auf unstrukturierten Daten und entwickelt eine Plattform für die Sammlung, Verarbeitung, Analyse und Aufbewahrung von Daten. Das Konsortium besteht aus einer Reihe an Institutionen mit unterschiedlichen Zuständigkeiten (<https://www.berd-nfdi.de/consortium/>). Für die hier vorgestellte BERD Academy zeichnen sich insbesondere die LMU München und die UB Mannheim verantwortlich.

Das für Teilnehmende kostenlose Angebot der BERD Academy ist thematisch vielfältig und breit gefächert (siehe Abb. 1). Als Lehrende werden sowohl Mitarbeitende des Projekts als auch externe Expert*innen eingebunden. So wurden im ersten Jahr einzelne Workshops sowohl zur Analyse von Netzwerkdaten als auch zur Auffindbarkeit von Daten mittels AI Tools angeboten. Die Workshop-Reihe “Statistics for the Public Goods“ beschäftigte sich mit dem möglichen Übergang von einer statistischen Ära zu einer Epoche von Daten und diskutierte die damit verbundenen ethischen Aspekte und Auswirkungen auf die Produktion von Daten. Der vierwöchige Flipped Classroom Kurs “How to Make Use of Machine Learning & Microsimulation in Official Statistics“ vermittelte Teilnehmenden die Anwendung von Machine Learning und Microsimulation Verfahren auf Daten aus der amtlichen Statistik. Die Veranstaltungen stehen grundlegend allen Interessierten offen. “Statistics for the Public Good“ und “How to Make Use of Machine Learning and Microsimulation in Official Statistics“ konnten darüber hinaus auch von Studierenden der LMU München für ECTS credits besucht werden. Auf diese Weise sollen Nachwuchswissenschaftler*innen bereits früh mit den Themen vertraut gemacht und spezifische Anwendungsgebiete aufgezeigt werden. Eine Herausforderung der damit häufig heterogenen Teilnehmendengruppe sind die unterschiedlichen Vorkenntnisse. Da insbesondere Probleme mit dem Einstieg in die Programmiersprachen festgestellt werden können, bietet die BERD Academy self-paced Kurse zu R und Python an, die den Teilnehmenden grundlegende Kenntnisse vermitteln und ihnen somit eine solide Grundlage für fortgeschrittenere Kurse bieten.

2022 Training



Workshops

Democratizing Data:
AI Tools for Discovery

Sep 29

AI tools for discovery:
Finding Datasets and Software in
Publications (Internal)

Oct 28

A Connected World:
Data Analysis for Real-World
Network Data

Dec 08

Workshop Series

Statistics for the Public Good

The Public Good
Statistics:
A Reflective
Introduction

Oct 10

The Public Good
Statistics:
Let's Talk About
Data Culture!

Oct 11

Should
Government Data
Concern or Serve
us?

Oct 11

Values, Ethics,
and what they
mean for Data
Quality

Oct 12

Data 4 Policy: Is
the Statistical
Era Replaced by
an Era of Data?

Oct 13

Flipped Classroom Course

How to use Machine Learning & Microsimulation in Official Statistics

Nov 14 - Dec 05

Coffee Lectures

Data Literacy Essentials

Daten verstehen	Oct 13
Daten erheben	Oct 20
Daten organisieren	Oct 27

Data Literacy Snacks

Data Protection Basics	Nov 09
GDPR Applicability	Nov 23

Self-Paced Course

Data Science with R	Permanent
Data Science with Python	Permanent

<https://www.berd-nfdi.de/berd-academy/>

Abb. 1: Angebotenes Training der BERD Academy im Kalenderjahr 2022

Die Angebote der BERD Academy sind auf eine langfristige Nutzung ausgelegt. Bereits erprobte Kurse und Workshops sollen erneut angeboten werden und dabei festgestellte Lücken entweder durch die Weiterentwicklung des jeweiligen Moduls oder durch Neuentwicklungen geschlossen werden. So ist am Trainingsprogramm für das aktuelle Jahr

2023 zu erkennen, dass einige der Kurse und Workshops aus dem Jahr 2022 erneut besucht werden können (siehe Abb. 2). Darüber hinaus werden aber auch einige neue Module angeboten. Dabei liegt für das Jahr 2023 ein besonderer Fokus auf eine Erweiterung der Inhalte um Forschungsdatenmanagement-spezifische Angebote. Dafür wird es zum einen eine Train the Trainer Workshop-Reihe für Fachreferent*innen in den Sozial- und Wirtschaftswissenschaften geben, um die Angebotsseite von Datendiensten zu stärken. Da diese Gruppe direkte Ansprechpartner*innen für Forschende beim Zugang zu Daten sind, ist die Stärkung der auf diesem Gebiet vorhandenen Kompetenzen zentral, um sowohl die Nutzung als auch Bereitstellung vorhandener Daten zu fördern. Zum anderen werden mit einem eigens entwickelten Flipped Classroom Kurs “Make your research reproducible – a hands-on course“ konkret Forschende angesprochen. Dabei sollen nicht nur bereits häufig adressierte Grundlagen um bspw. die FAIR Prinzipien⁵ vermittelt, sondern spezifisch gezeigt werden, mit welchen Software-Lösungen und inhaltlichen Ansätzen eine konkrete Umsetzung erreicht werden kann. Auf diese Weise sollen die Open Science Prinzipien von der abstrakten, ideellen Ebene in die Disziplinen übertragen werden. Dazu findet die Entwicklung des Kurses reziprok mit dem Fortschritt des Konsortiums und der entstehenden BERD Platform für den Austausch von Daten und Algorithmen statt.

2023 Training



Workshops

Research Data Management (Internal) Mar 16

A Connected World: Data Analysis for Real-World Network Data Jul 19

Workshop Series

Statistics for the Public Good Oct 10 Oct 11 Oct 12

Train the Trainer for Subject Librarians Sep 25 - Oct 06

Flipped Classroom Courses

How to use Machine Learning & Microsimulation in Official Statistics Nov - Dec

Make your research reproducible – a hands-on course Oct 17 - Nov 14

Coffee Lectures

Data Literacy Essentials Mar 09 Mar 16 Mar 23 Mar 30

FAIR Data Week May 30 - Jun 2

Webinars

Quo Vadis, Digital Privacy? Perceptions, Practices and Policies Jan 26

IAB SMART Webinar Series

The IAB-SMART Study: Collecting Behavioral Smartphone Sensor Data for Social Research Jun 28

What do geolocation smartphone data add to a survey panel? Available indicators from the IAB-SMART Project Jul 05

How to tidy and anonymize raw smartphone geolocation data: Code and Practitioner's Examples from the IAB-SMART Project Jul 12

Events and Cooperations

DataFEST Germany Apr 14 - 16

Women in Data Science Munich June 19

Data Science for Social Good Munich Aug - Sep

<https://www.berd-nfdi.de/berd-academy/>

Abb. 2: Geplantes Training der BERD Academy im Kalenderjahr 2023

Mit der BERD Academy⁶ wird seit Mitte 2022 ein auf die speziellen Charakteristika von (unstrukturierten) wirtschafts- und sozialwissenschaftlichen Daten abgestimmtes Trainingsprogramm entwickelt. Es spricht sowohl Einsteiger*innen als auch fortgeschrittene Forschende an und soll damit die breitflächige wissenschaftliche Anwendung von Data

⁵ Die Abkürzung FAIR steht für *F*indable, *A*ccessible, *I*nteroperable, and *R*e-usable. Diese Grundsätze sollen Forschungsdaten erfüllen, um nachhaltig nutzbar zu sein.





⁶ Vollständiges, aktuelles Angebot unter: <https://www.berd-nfdi.de/berd-academy/> (Stand: 26.08.2023)

Science Methoden in den betreffenden Disziplinen stärken. Das praxisorientierte, flexible Angebot von kleineren Einheiten wie Coffee Lectures bis hin zu mehrwöchigen Flipped Classroom Kursen zielt explizit darauf ab, auch Forschenden mit geringen Zeitressourcen die Weiterbildung zu ermöglichen. Gleichzeitig ermöglicht diese Flexibilität eine ständige Weiterentwicklung, um neue Themen, Methoden und Daten zeitgemäß in attraktive Trainingsangebote zu überführen.

Literaturverzeichnis

- [Ha21] Haensch, Anna-Carolina; Herklotz, Markus; Keusch, Florian; Kreuter, Frauke: The International Program in Survey and Data Science (IPSDS): A modern study program for working professionals. *Statistical Journal of the IAOS*, 37:1–13, Juli 2021.
- [KK20] Keusch, Florian; Kreuter, Frauke: Zukunft der Aus- und Weiterbildung in der Markt- und Sozialforschung. In (Keller, Bernhard; Klein, Hans-Werner; Wachenfeld-Schell, Alexandra; Wirth, Thomas, Hrsg.): *Marktforschung für die Smart Data World*, S. 3–25. Springer Fachmedien Wiesbaden, Wiesbaden, 2020.


Zertifikatskurs „Forschungsdatenmanagement für Studierende“ der Landesinitiative für Forschungsdatenmanagement in Brandenburg


Daniela Merten ¹, Heike Neuroth ², Carsten Schneemann ³ und Kathrin Woywod ⁴


Abstract: Ein professionelles Forschungsdatenmanagement verbessert die Nachvollziehbarkeit und Nutzbarkeit von Forschungsdaten. Forschende können eine wachsende Zahl an Schulungs- und Qualifizierungsangeboten zum Forschungsdatenmanagement nutzen, um den professionellen Umgang mit Daten zu erlernen. Schulungsangebote richten sich meist an Promovierende, wissenschaftliches und forschungsunterstützendes Personal. Für Bachelor- und Masterstudierende gibt es kaum Angebote, obwohl Forschungsdaten häufig bereits im Rahmen von Abschlussprojekten und -arbeiten Grundlage ihrer Fragestellungen sind. Die Landesinitiative Forschungsdatenmanagement in Brandenburg (FDM-BB), ein Verbund der acht forschenden staatlichen Hochschulen Brandenburgs, entwickelte daher im Rahmen des Projekts IN-FDM-BB ein Konzept für einen Zertifikatskurs zum Forschungsdatenmanagement für Studierende. Der erste Zertifikatskurs wurde im März 2023 als einwöchige Online-Spring School mit 30 Studierenden und elf Dozierenden mit unterschiedlicher beruflicher Expertise (Forschung und Lehre, Forschungsservice, Bibliothek) durchgeführt und ist, je nach Prüfungsleistung, mit 2 bis 4 ECTS / Leistungspunkten anrechenbar. Der Artikel beschreibt organisatorische, formale, technische, didaktische und inhaltliche Aspekte des Zertifikatskonzepts und liefert einen ersten Erfahrungsbericht.

Keywords: Forschungsdaten, Forschungsdatenmanagement, Qualifizierung, Zertifikat, Schulung, Spring School, Studierende, Brandenburg, Landesinitiative

¹ Universität Potsdam, Universitätsbibliothek, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, daniela.mertzen@uni-potsdam.de,  <https://orcid.org/0000-0003-4471-9255>

² Fachhochschule Potsdam, Fachbereich Informationswissenschaften, Kiepenheuerallee 5, 14469 Potsdam, heike.neuroth@fh-potsdam.de,  <https://orcid.org/0000-0002-3637-3154>

³ Fachhochschule Potsdam, Fachbereich Informationswissenschaften, Kiepenheuerallee 5, 14469 Potsdam, carsten.schneemann@fh-potsdam.de,  <https://orcid.org/0000-0002-2683-5853>

⁴ ehemals Brandenburgische Technische Universität Cottbus-Senftenberg, jetzt Technische Universität Berlin, Universitätsbibliothek, Fasanenstraße 88, 10623 Berlin, kathrin.woywod@tu-berlin.de,  <https://orcid.org/0000-0002-9774-326X>

1 Einleitung

In Deutschland setzt sich nach und nach das Konzept des Data Stewardships (DSS) durch. Data Stewards (DS) unterstützen Wissenschaftler*innen im Forschungsdatenmanagement und gelten als Fachkräfte der Zukunft. Schätzungen gehen davon aus, dass allein für die European Open Science Cloud (EOSC⁵) in den nächsten Jahren rund 500.000 Data Stewards gebraucht werden, wenn ein Data Steward für rund 20 wissenschaftliche Daten-Produzierende verantwortlich sein soll [Mo16]. Die wegweisende Publikation von Wilkinson et al. (2016) mit dem Titel *The FAIR Guiding Principles for scientific data management and stewardship* formulierte zum ersten Mal 15 FAIR-Kriterien zur verbesserten Nachnutzbarkeit von Forschungsdaten und führte den Begriff Data Stewardship international ein [Wi16].

Bundesweit existiert ein noch überschaubares Angebot an zertifizierten Qualifizierungen und Schulungen zum Forschungsdatenmanagement (FDM), das sich an Promovierende sowie an wissenschaftliches und forschungsunterstützendes Personal richtet (z. B. [Ko22]). Doch auch Bachelor- und Masterstudierende erheben bereits häufig eigene Forschungsdaten oder nutzen diese nach, beispielsweise für Projekte oder Abschlussarbeiten. Masterstudierende legen mit ihrem Abschluss den ersten wissenschaftlichen Grad ab und sind potenzielle Nachwuchswissenschaftler*innen, die von bedarfsgerechten FDM-Schulungen profitieren können, um so den professionellen Umgang mit Forschungsdaten von Beginn an als festen Bestandteil ihrer Arbeit in den Forschungsalltags integrieren zu können. Die Lehre von FDM, z. B. in Form von Seminaren, Projekten oder Vorlesungen in einem Studiengang, hat bisher allerdings noch keinen systematischen Eingang in deutsche Curricula bzw. Studiengänge gefunden (z. B. [Ha23]). Der Weiterbildende Masterstudiengang Digitales Datenmanagement der Fachhochschule Potsdam und der Humboldt-Universität zu Berlin ist bisher noch ein einzigartiges Studienangebot⁶ (z. B. [Ko22]): „Als einer der ersten seiner Art adressiert er explizit die Bedarfe künftiger DS und deckt die Aufgaben bzw. erforderlichen Kompetenzen des fächerübergreifenden DSS weitgehend ab.“ ([Se23], S. 38). Nur langsam gewinnen Studierende als Zielgruppe von FDM-Schulungen und Weiterbildungen an Aufmerksamkeit. Beispielsweise zielt das Projekt FDM@Studium.nrw⁷ darauf ab, disziplinübergreifende OER (Open Educational Resources)-Materialien für die Hochschullehre zu entwickeln. Breit aufgestellte Angebote, welche sich spezifisch den FDM-Bedarfen Studierender widmen, gibt es bislang nicht. Qualifikationsangebote dieser Art sind jedoch wichtig, um die gute wissenschaftliche Praxis im Umgang mit Forschungsdaten frühzeitig zu erlernen (z. B.

⁵ European Open Science Cloud, <https://eosc-portal.eu/>, Stand: 23.05.2023.

⁶ Digitales Datenmanagement (M.A.), <https://www.fh-potsdam.de/studium-weiterbildung/studiengaenge/digitales-datenmanagement-m>, Stand: 23.05.2023.

⁷ Gefördert von Oktober 2022 bis März 2024 durch das Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen

[De21, De22]). Nur so können Transparenz, Nachvollziehbarkeit und Nachnutzbarkeit von Forschungsdaten sowie die Integrität von Forschungsergebnissen gesichert werden.

Die Landesinitiative Forschungsdatenmanagement in Brandenburg (FDM-BB)⁸ konzipierte seit dem Frühjahr 2022 einen bundeslandweiten FDM-Zertifikatskurs, der sich vorrangig an Masterstudierende, aber auch an Bachelorstudierende aller brandenburgischen forschenden Hochschulen richtet.⁹ Innerhalb der Landesinitiative, im Rahmen des vom Bundesministerium für Bildung und Forschung (BMBF) und Ministerium für Wissenschaft, Forschung und Kultur (MWFK) geförderten Verbundprojektes *Institutionalisiertes und nachhaltiges Forschungsdatenmanagement in Brandenburg* (IN-FDM-BB, [In23]) aller acht staatlichen Hochschulen Brandenburgs entstand eine hochschulübergreifende Arbeitsgruppe zur Entwicklung eines FDM-Zertifikatskurses für Studierende, die kooperativ das didaktische und inhaltliche Konzept für den Zertifikatskurs entwickelte.

2 Konzept

Im Folgenden werden die Parameter erläutert, die bei der Entwicklung des Konzepts für den Zertifikatskurs berücksichtigt wurden bzw. werden mussten.

Einige *organisatorische* Rahmenparameter waren dabei von Beginn an gesetzt, da der Zertifikatskurs alle Studierenden der brandenburgischen Hochschulen adressieren soll und für alle interessierten Studierenden zugänglich ist:

- *Termin:* Der Zertifikatskurs soll ein- bis zweimal jährlich in der vorlesungsfreien Zeit stattfinden, so dass allen interessierten Studierenden eine Teilnahme möglich ist. Eine Herausforderung dabei war, eine freie Woche zu identifizieren, da die vorlesungsfreien Zeiten der acht Hochschulen nicht identisch sind und Studierende häufig in der vorlesungsfreien Zeit Prüfungen ablegen.
- *Teilnahmegebühren:* Der Zertifikatskurs soll für die Studierenden kostenfrei angeboten werden.
- *Format:* Der Zertifikatskurs soll als einwöchige, rein digitale *Spring School* durchgeführt werden, sodass Studierenden keine Anfahrts- oder Übernachtungskosten entstehen und weite Anfahrtswege für Studierende und Dozierende kein Hindernis darstellen.
- *Teilnehmende:* Die maximale Anzahl an Teilnehmenden soll auf 30 Teilnehmende begrenzt werden, um eine angemessene Betreuung durch die Dozierenden gewährleisten zu können und um Austausch, Diskussion und Reflexion im Rahmen praktischer Übungen im digitalen Raum zu fördern.

⁸ Forschungsdatenmanagement in Brandenburg (FDM-BB), <https://fdm-bb.de/>, Stand: 23.05.2023.

⁹ Private Hochschulen und Verwaltungshochschulen werden beispielsweise nicht adressiert.

- *Dozierende*: Dozierende müssen dafür gewonnen werden, sich kostenfrei für ihren Teil der Lehrthemen zu engagieren. Ein Honorarvertrag darf auf Grund der Zuwendungsrichtlinien des BMBF für das Vorhaben IN-FDM-BB nicht ausgestellt werden, auch eine Bezahlung über z. B. Sponsorengelder ist seitens des Förderers ausgeschlossen, auch nach dem Förderende.
- *Struktur der Spring School*: Feste Anfangs-, Pausen- und Endzeiten sind Grundvoraussetzung für einen reibungslosen Ablauf. Eine feste Struktur ermöglicht auch eine Teilnahme von Studierenden, die z. B. für Care-Aufgaben verantwortlich sind. Sie konnten sich auf eine klare zeitliche Wochenstruktur verlassen.

Formale Kriterien müssen ebenfalls während der Konzeptionsphase berücksichtigt werden, um die Vielfalt des studentischen Alltags und den spezifischen Anforderungen der eigenen Hochschule (Stichwort Hochschulautonomie) gerecht werden zu können:

- *Vergabe von ECTS / LP¹⁰*: Der Zertifikatskurs soll mit 2 bis 4 ECTS-Punkten an der jeweiligen Hochschule anrechenbar sein. Zwei ECTS setzen eine aktive Kursteilnahme inklusive vorbereitender Selbstlernphase zum Einlesen in die grundlegende Literatur und die Teilnahme an einem Quiz quer über alle Lehrinhalte voraus. Der zeitliche Gesamtumfang wurde mit 50 Stunden kalkuliert. Für Studierende, die eine zertifizierte Teilnahmebescheinigung über 3 oder 4 ECTS benötigen¹¹, wird die Leistungsanforderung in mehreren Aktivitätsbereiche gegliedert: Die vorbereitende Selbstlernphase (10 Stunden) sowie die aktive Teilnahme an der einwöchigen Spring School (40 Stunden) und eine nachbereitende Arbeit in Form einer speziellen Prüfungsleistung, die einem Arbeitsaufwand von 1 (25 Stunden) oder 2 (50 Stunden) ECTS entspricht. Es müssen Prüfungsleistungen definiert werden, die den Interessen der Studierenden gerecht werden, aber auch formal den Ansprüchen der eigenen Hochschule genügen.

Technische Voraussetzungen für die Durchführung der digitalen *Spring School*:

- *Videokonferenz-Tool*: Da der Kurs als digitale Veranstaltung konzipiert wurde, braucht es ein stabiles und vielseitiges Videokonferenzsystem, welches z. B. über Umfragemöglichkeiten, Break-Out Räume für Gruppenarbeiten, ein Whiteboard etc. verfügt. Als geeignetes Videokonferenzsystem haben sich die Verantwortlichen für Zoom entschieden, auch weil es von der Universität Potsdam datenschutzkonform nach DSGVO angeboten werden kann.¹²
- *Lernmanagementsystem (LMS)*: Ein geeignetes LMS für den Zertifikatskurs muss mehrere Voraussetzungen erfüllen: Ausgewiesenes Rollen- und Rechtesystem, Zulassung von extern Nutzenden (Studierende und Dozierende) ohne Beantragung von Hochschulaccounts bei der anbietenden Hochschule und eine technische

¹⁰ European Credit Transfer Point (ECTS) bzw. Leistungspunkt (LP), im weiteren Verlauf nur noch ECTS

¹¹ z. B. weil das entsprechende Wahlpflichtfach im eigenen Studiengang mehr als 2 ECTS-Punkte verlangt

¹² ZIM - Zentrum für Informationstechnologie und Medienmanagement, <https://www.uni-potsdam.de/de/zim/angebote-loesungen/webconferencing/zoom>, Stand: 23.05.2023.

Umgebung, die weniger LMS-affinen Dozierenden die Nutzung ohne größere Einarbeitungen erlaubt. Hier kann auf die Moodle-Lernplattform für digitale Projekte der BTU Cottbus zurückgegriffen werden.¹³

- *Anmeldesystem:* Für ein digitales Anmeldesystem, welches datenschutzkonform alle relevanten Daten auf der Anmeldeseite¹⁴ abfragen kann, wurde ein spezifisches Plugin gekauft.

Folgende *inhaltliche* Rahmenparameter spielten bei der Konzeptentwicklung eine Rolle:

- *Auswahl der Inhalte:* Zunächst wurde analysiert, welche Inhalte bei anderen relevanten FDM-Schulungen angeboten werden, in welchem Rahmen und in welchem Umfang. Die DINI/nestor-AG Forschungsdaten¹⁵ hat einen umfangreichen und stets aktualisierten Materialkatalog zu FDM-Schulungsmaterialien zusammengestellt, der über das Medien-Repository der Humboldt-Universität zu Berlin¹⁶ zugänglich und durchsuchbar ist [FD18]. Auch das Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement [Bi21] bietet einen guten Einstiegspunkt in das Thema FDM. Als Basis für die Themenauswahl konnte ebenfalls der Zertifikatskurs Forschungsdatenmanagement der TH Köln dienen.¹⁷ Hilfreich waren auch laufende Projekte wie z. B. *Forschung unterstützen - Empfehlungen für Data Stewardship an akademischen Forschungsinstitutionen: Ergebnisse des Projektes DataStew* [Se23] oder der *Workshopbericht zum Community Workshop: Kartierung der Datenkompetenzlandschaft* [LDZ23] des NFDI-Projekts Datenkompetenz¹⁸. Die *Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards* [Pe22] bot ebenfalls einen guten Einstiegspunkt, um Inhalte und Lernziele des Zertifikatskurses festzulegen. Auf europäischer Ebene bietet der *EOSCPillar Training & Support*¹⁹ im Rahmen der European Open Science Cloud (EOSC) ein Nachweissystem relevanter Trainings- und Schulungsmaterialien an, welches in die Konzeption des Kurses mit einbezogen wurde. Die zu vermittelnden Inhalte wurden anschließend durch die Dozierenden mit der entsprechenden Expertise im jeweiligen Themenschwerpunkt einer Lehrinheit und in Rücksprache mit denjenigen Dozierenden, die in der

¹³ Plattform für digitale Projekte der BTU, <https://www.b-tu.de/elearning/projekte/>, Stand: 23.05.2023.

¹⁴ Die Anmeldung für den Zertifikatskurs im März 2023 war und wird voraussichtlich für folgende Zertifikatskurse möglich sein über die Website des Verbundprojekts FDM-BB: <https://fdm-bb.de/zertifikatskurse/>, Stand: 23.05.2023.

¹⁵ Dini/nestor-AG Forschungsdaten, <https://dini.de/ag/dininestor-ag-forschungsdaten/>, Stand: 23.05.2023.

¹⁶ Forschungsdatenmanagement HU Berlin, <https://rs.cms.hu-berlin.de/dataman/pages/home.php?login=true/>, Stand: 23.05.2023.

¹⁷ Zertifikatskurs Forschungsdatenmanagement TH Köln, https://www.th-koeln.de/weiterbildung/zertifikatskurs-forschungsdatenmanagement_82048.php, Stand: 23.05.2023.

¹⁸ Projekt „Datenkompetenz in NFDI“, <https://www.nfdi.de/datenkompetenz/>, Stand: 23.05.2023.

¹⁹ EOSC Pillar Gateway, <https://eosc-pillar.d4science.org/web/eoscpillartrainingandsupport/catalogue>, Stand: 23.05.2023.

Hochschullehre tätig sind oder waren, ausgewählt, um das Kursniveau bedarfsgerecht zu gestalten.

- *Disziplinübergreifende Relevanz der Inhalte:* Bei der inhaltlichen Planung muss die große Fächervielfalt der brandenburgischen Hochschulen berücksichtigt werden. Diese reicht von Studiengängen im Bereich Design oder Soziale Arbeit, über Europäische Medienwissenschaften, Judaistik, Bibliothekswissenschaft, Literatur- und Sprachwissenschaften bis hin zu ingenieurwissenschaftlichen, technischen und naturwissenschaftlichen Studiengängen. Es musste also ein Angebot entwickelt werden, welches Lehrinhalte abbildet, die disziplinübergreifend relevant sind, jedoch jeweils fachspezifische Standards berücksichtigen. Innerhalb des Zertifikatskurses können punktuell konkrete Beispiele aus verschiedenen Disziplinen präsentiert werden und durch Gruppenarbeiten oder Übungsaufgaben mit Fokus auf bestimmte Fächer die Fächervielfalt berücksichtigt werden.
- In Übereinstimmung mit dem aktuellen State-of-the-Art (s.o.), deckt der Zertifikatskurs folgende *disziplinübergreifende Lerninhalte* (s. Abb. 1) ab: (1) Theoretische Einführung in das FDM inkl. Definition und Motivation, FD-Policies, FD-Lebenszyklus, FAIR und CARE Prinzipien, Open Science Praktiken, (2) Datendokumentation und Metadaten-Standards, Datenmanagementpläne inkl. Anfertigung eines Datenmanagementplans mit dem Tool RDMO (Research Data Management Organiser)²⁰, (3) technische FD-Infrastrukturen und Tools inkl. Dateiformate, Datenspeicherung, Backup, Datensicherheit, Langzeitarchivierung, FD-Repositoryn, Versionsverwaltungsdienste, kollaborative Tools, (4) FD-Publikationen inkl. persistente Identifizierung, Lizenzierung, rechtliche Aspekte des FDM inkl. Urheberrecht und Datenschutzrecht, (5) gute wissenschaftliche Praxis inkl. Forschungsethik und wissenschaftliche Integrität, sowie Projektmanagement im Kontext von Forschungsdaten.
- Für die Vermittlung bestimmter Lehrinhalte müssen *geeignete Dozierende* zur Verfügung stehen. Die am ersten Zertifikatskurs beteiligten Dozierenden haben sich über Monate auf Inhalte, didaktische Herangehensweisen, einheitliche Strukturen etc. beim Aufbau der Lehrskripte geeinigt.

²⁰ RDMO, <https://rdmo.aip.de/>, Stand: 23.05.2023

FDM Zertifikatskurs – Spring School 06. März - 10. März 2023

	Montag, 6.3.2023	Dienstag, 7.3.2023	Mittwoch, 8.3.2023	Donnerstag, 9.3.2023	Freitag, 10.3.2023
Uhrzeit	Modulkurs 1: Einführung in das FDM	Modulkurs 2: Grundlagen Datenmanagementpläne	Modulkurs 3: Grundlagen technischer FD-Infrastruktur & Tools	Modulkurs 4: Datenpublikation u. rechtliche Aspekte des FDM	Modulkurs 5: Gute wiss. Praxis u. Projektmanagement
09:00 - 10:30	1.1 Organisatorisches, Einführung in das Forschungsdaten- management	2.1 Datendokumentation, Metadaten	3.1 Aktives Datenmanagement	4.1 Datenpublikation, Persistente Identifizierung, Zitation	5.1 Gute wissenschaftliche Praxis
10:30 - 10:45	Pause				
10:45 - 12:15	1.2 Der Forschungsdaten- Lebenszyklus, FAIR Data Principles	2.2 Datenmanagementpläne (DMP)	3.2 Langzeitarchivierung	4.2 Lizenzen, Re3Data	5.2 Projektmanagement mit Fokus auf Datenmanagement
12:15 - 13:30	Mittagspause				
13:30 - 15:00	1.3 Open Science	2.3 DMP mit dem Research Data Management Organiser (RDMO)	3.3 Kollaborative Tools und Versionierung	4.3 Rechtliche Aspekte des FDM	5.3 Übung, Evaluation
15:00 - 15:30	Pause				
15:30 - 17:00	1.4 Übung	2.4 Übung	3.4 Übung	4.4 Übung	5.4 Individuelle Beratung zu Prüfungsleistungen

Abb. 1: Struktur der einwöchigen digitalen Spring School im März 2023

Ein Zertifikatskurs, der über eine Woche jeweils ganztägig durchgeführt wird und eine sehr heterogene Zielgruppe bezogen auf Studiengang, Semesterzahl, angestrebtes Abschlussniveau (Bachelor, Master) adressiert, bringt einige Herausforderungen mit sich. Ein ausgefeiltes *didaktisches* Konzept soll helfen, dass die Studierenden nicht mental “abschalten”, sich hinter sog. “schwarzen Kacheln” im Zoom-Raum verbergen oder den Kurs abbrechen. Komplexe Themen sollen so aufbereitet werden, dass auch Anfänger*innen für sich Lernerfolge und Kompetenzaufbau verbuchen können. Folgende didaktische Überlegungen liegen dem Zertifikatskurs zugrunde:

- *Didaktisches Vorgehen:* Generell steht weniger das Vermitteln von reinem Wissen im Vordergrund, stattdessen soll das didaktische Konzept das individuelle Problembewusstsein für die verschiedenen Themenbereiche schärfen. Somit steht das problembasierte Lernen im Fokus, welches die Interaktivität und die Lösungskompetenz fördert und der *Ermöglichungsdidaktik* folgt [Ar12] unter Berücksichtigung des sog. *Shifts from Teaching to Learning* [WG05]. Theoretische Einführungen werden jeweils durch weiterführende Literatur, praktische Übungen, Konsultationen und Self-Assessments (z. B. in Form eines Quiz) flankiert, um didaktisch durch verschiedene, interaktive Instrumente den spezifischen Bedarfen der Studierenden nachzukommen. In diesem Zertifikatskurs werden abwechslungsreiche Lehr- und Lernmethoden angewendet:
 - *Theoretische Einführungen* in die identifizierten (s.o.) Lerninhalte in kurzen inhaltlichen Inputformaten mit Hinweisen zu relevanter und – bei Interesse für das Selbststudium weiterführender – Literatur

- Konkrete *praktische Beispiele* aus dem FDM-Alltag inklusive *moderierter Diskussionsrunden* zu bestimmten (z. B. ethischen) Fragestellungen
- Vielfalt an *unterschiedlichen Formaten, Medien und Tools* für eine abwechslungsreiche Vermittlung von Inhalten, z. B. Text, strukturierter Text in Tabellenform, Video, Audio, kollaborative Tools etc.
- *Live-Anwendungen* von FDM-relevanten Diensten, sowohl durch die Dozierenden als auch in Gruppenarbeit zum selbstständigen Erfahren
- *Kurze Übungseinheiten* inklusive Überprüfung des individuellen Lernerfolgs und Kompetenzaufbaus. In der Selbstlernphase nahmen die Studierenden an einem interaktiven Quiz (multiple choice) mit Fragen zu jedem Themenbereich teil, um sich einen Überblick über Inhalte und Anforderungsniveau zu verschaffen. Es konnte jederzeit wiederholt werden, um eigene Lernfortschritte zu evaluieren
- *Kleinere Umfragen*, um in anonymisierter Form jederzeit prüfen zu können, ob die Studierenden den Inhalten folgen können
- *Gamification*, z. B. ein “Glücksrad-Spiel”, um gelernte Terminologien zu festigen

3 Durchführung des Zertifikatskurses

Der erste Zertifikatskurs für Studierende fand als einwöchige, digitale *Spring School* im März 2023 statt und war, je nach Umfang der Prüfungsleistung, mit 2 bis 4 ECTS anrechenbar. Ein Modulhandbuch²¹ wurde erstellt, welches den Regeln der Modularisierung der Verordnung zur Regelung der Studienakkreditierung entspricht²²; es beinhaltet die Modulbeschreibung samt Qualifikationszielen, Informationen zu Lehr- und Lernformen, Teilnahmevoraussetzungen, Voraussetzungen für die Vergabe von ECTS inklusive Prüfungsart, -umfang und -dauer, ECTS und Benotung, Häufigkeit des Angebots des Moduls, Arbeitsaufwand und Dauer des Moduls. Zusätzlich enthält das Modulhandbuch wichtige organisatorische Informationen und einen Code of Conduct. Die Lehrmaterialien des Zertifikatskurses stehen als OER-Materialien zur Verfügung.²³

Unter den knapp 100 Bewerber*innen wurden 30 Plätze nach festgelegten Auswahlkriterien (Voranmeldung, Hochschule etc.) vergeben. Die übrigen Bewerber*innen wurden auf eine Warteliste gesetzt und sind bei Absage angenommener Bewerber*innen nachgerückt. Von den Teilnehmenden waren ca. 25%

²¹ Zertifikatskurs „Forschungsdatenmanagement für Studierende“, <https://doi.org/10.5281/zenodo.7936966>, Stand: 27.05.2023.

²² Verordnung zur Regelung der Studienakkreditierung (Studienakkreditierungsverordnung - StudAkkV), <https://bravors.brandenburg.de/verordnungen/studakkv#9>, Stand: 23.05.2023.

²³ Zertifikatskurs „Forschungsdatenmanagement für Studierende“, <https://doi.org/10.5281/zenodo.7936966>, Stand: 23.05.2023.

Bachelorstudierende und 75% Masterstudierende. Das Fächerspektrum war vielfältig und deckte u. a. die Bereiche Bioinformatik, Biotechnologie, BWL, Chemie, Informationswissenschaften, Landschaftsnutzung und Naturschutz, Linguistik, Mathematik, Psychologie und Soziale Arbeit ab. Ca. 75% der Teilnehmenden gaben vor Kursbeginn an, noch FDM-Anfänger*innen zu sein. 60% der Teilnehmenden gaben an, bereits mit Daten zu arbeiten. Als Motivation zur Kursteilnahme, wurden u. a. folgende Gründe angegeben (Freitextfeld): FDM in den Forschungsalltag integrieren; Daten auffindbar, nachvollziehbar und nachnutzbar aufzubereiten; als Vorbereitung auf die Promotion; Umgang mit Daten für die Masterarbeit; das Wissen erweitern, um beruflich in diesem Feld zu arbeiten.

Insgesamt elf Dozierende der BTU Cottbus, der FH Potsdam und der Universität Potsdam mit unterschiedlicher beruflicher Expertise (Bibliothek, Forschungsservice und Forschung und Lehre) vermittelten die in Abb. 1 dargestellten Lehrinhalte. Am Ende jedes Veranstaltungstages wurde eine kurze Lehr-Evaluation durchgeführt. Die Evaluationen wurden von 60% bis 80% der Studierenden ausgefüllt und zeigten insgesamt eine hohe bis sehr hohe Zustimmung bezüglich Qualität, Struktur und Verständlichkeit der Kursinhalte sowie Angemessenheit des Anforderungsniveaus. Vereinzelt gab es kritische Kommentare zum Anforderungsniveau und einige Studierende wünschten sich mehr Zeit für Austausch und Diskussion. Die Studierenden gaben an, dass dies ihre erste Erfahrung einer hochschulübergreifenden Lehrveranstaltung war; der Austausch mit Studierenden und Dozierenden anderer Hochschulen und die Fächervielfalt wurden als sehr bereichernd empfunden.

Nach Abschluss des Zertifikatskurses nahmen 67% der Teilnehmenden an einer Gesamtevaluation teil. Diese ermittelte die Einschätzung der Studierenden zur Relevanz und Schwierigkeit der einzelnen Modulkurse sowie eine Einschätzung zum Lernzuwachs, zu den Lehrformen, zur Lernatmosphäre und zum Vorwissen. Im Durchschnitt, über alle fünf Modulkurse hinweg, wurden die Lerninhalte als sehr relevant angesehen ($\bar{x} = 1.8$, Skala: 1 = höchst relevant, 7 = gar nicht relevant), der Schwierigkeitsgrad als angemessen bewertet ($\bar{x} = 2.98$, 1 = gar nicht schwierig, 7 = extrem schwierig) und der Lernzuwachs als hoch ($\bar{x} = 3$, 1 = extrem hoch, 7 = extrem gering) beurteilt. Die Online-Lehrform inklusive Zusammensetzung aus theoretischem Input, Übungen, Q & A-Sessions und vorbereitender Literatur und Quiz sowie die Lernatmosphäre wurden im Durchschnitt positiv bewertet. Besonders vielversprechend war die Rückmeldung, dass 100% der Evaluierenden FDM als wichtiges Thema erachteten; 95% gaben an, FDM in ihr weiteres wissenschaftliches Arbeiten zu integrieren. In den Freitextfeldern gab es vereinzelte Kritikpunkte, z. B. „Vorbereitendes Quiz war etwas zu tief in einzelnen Themengebieten und hat sehr in den Kurs vorgegriffen.“ Positiv wurde z. B. vermerkt: „(...) die brandenburgweite Ausrichtung und der Kontakt zu Studierenden außerhalb der eigenen Hochschule/des eigenen Fachbereichs waren interessant und bereichernd.“. Insgesamt lässt sich feststellen, dass ein ausgezeichnetes Arbeitsklima und ein reger Austausch entstehen, wenn das Curriculum Freiheitsgrade für die Studierenden vorsieht, sich auch außerhalb der eigenen Hochschule Lerninhalte freiwillig aussuchen und anrechnen lassen zu können.

4 Fazit und Ausblick

Insgesamt war die erste Iteration des Zertifikatskurses ein großer Erfolg: Der erste brandenburgische Zertifikatskurs für Studierende stieß auf großes Interesse; es gab über dreimal so viele Anmeldungen wie verfügbare Plätze. Eine einwöchige, rein digital durchgeführte *Spring School* mit ganztägigen Veranstaltungen verlangt Studierenden und Dozierenden einiges ab; das positive Feedback der Studierenden unterstützt die Entscheidung für dieses Konzept jedoch. In einem nächsten Schritt werden die gesamten Evaluationen (pro Tag und bezogen auf den kompletten Zertifikatskurs) herangezogen, um die Konzeption inhaltlich und didaktisch weiterzuentwickeln. Auch wird geprüft, ob neue Themenfelder, wie z. B. ein stärkerer Fokus auf ethische Aspekte des FDM, zugunsten bisher gelehrter Themenfelder in die *Spring School* integriert werden sollen.

Im Rahmen des Projekts IN-FDM-BB werden weitere Qualifizierungsprogramme entwickelt, so auch für die Zielgruppen Forschende und FDM-Verantwortliche. Der Zertifikatskurs für Studierende wird weiterhin mindestens einmal jährlich in der vorlesungsfreien Zeit stattfinden.

Literaturverzeichnis

- [Ar12] Arnold, R.: Ermöglichungsdidaktik - die notwendige Rahmung einer nachhaltigen Kompetenzreifung. *Berufsbildung in Wissenschaft und Praxis* 41/12, S. 45-48, 2012.
- [Bi21] Biernacka, K. et.al.: Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement (Version 4) [Computer software]. Zenodo, 2021.
- [De21] Deutsche Forschungsgemeinschaft: Umgang mit Forschungsdaten: Checkliste für Antragstellende zur Planung und zur Beschreibung des Umgangs mit Forschungsdaten in Forschungsvorhaben, 2021. https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_de.pdf, Stand: 23.05.2023.
- [De22] Deutsche Forschungsgemeinschaft: Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex (Version 1.1), 2022. <https://doi.org/10.5281/zenodo.6472827>, Stand: 23.05.2023
- [FD18] FDMentor; DINI/nestor-AG Forschungsdaten: Materialkatalog zum Forschungsdatenmanagement (1.0) [Data set], 2018.
- [Ha23] Haase, C.: Forschungsdatenmanagement – Eine Schlüsselkompetenz in der Hochschulbildung? Eine Analyse des Ist-Zustandes an der Brandenburgischen Technischen Universität Cottbus-Senftenberg und Zukunftsperspektiven. Master's Thesis DDM 2023. Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft, in press.
- [In23] IN-FDM-BB/FHP; IN-FDM-BB/UP: IN-FDM-BB Förderantrag: Institutionalisiertes und nachhaltiges Forschungsdatenmanagement in Brandenburg. Zenodo, 2023.
- [Ko22] Koch, M. et.al.: Forschungsdatenmanagement in Aus- und Weiterbildung: Überblick über Studiengänge und Kurse in Deutschland (02/22) [Data set], 2022.

- [LDZ23] Lehmann, A.; Day, S.; Zinke, W.: Workshopbericht zum Community Workshop: Kartierung der Datenkompetenzlandschaft, 2023.
- [Mo16] Mons, B.: We need 500.000 respected data stewards to operate the European Open Science Cloud. E-Infrastructures Reflection Group – News blog, 2016.
- [Pe22] Petersen, B. et.al.: Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards (Version 1). Zenodo, 2022.
- [Se23] Seidlmayer, E. et.al.: Forschung unterstützen - Empfehlungen für Data Stewardship an akademischen Forschungsinstitutionen: Ergebnisse des Projektes DataStew. Forschungsergebnisse der HAW Hamburg, 2023.
- [WG05] Welbers, U.; Gaus, O. (Eds.): The shift from teaching to learning: Konstruktionsbedingungen eines Ideals; für Johannes Wildt zum 60. Geburtstag. W. Bertelsmann, 2005.
- [Wi16] Wilkinson, M.D. et.al.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3/16, S. 160018, 2016.

Data Science und Forschungsdaten als Treiber für wissenschaftliche Interdisziplinarität

Vom Data Science Center zur skalierbaren Serviceeinrichtung für datenbezogene Forschung

Katharina Weiß¹ und Reinhold Decker ²

Abstract: Aktuelle, hochkomplexe gesellschaftliche Herausforderungen, wie die Klimakrise oder die Corona-Pandemie, unterstreichen die Bedeutsamkeit von fachübergreifender, interdisziplinärer Forschung. Data Science und datenbezogene Diskurse als Treiber für interdisziplinäre Forschung zu nutzen, ist der grundlegende Ansatz des *Bielefeld Center for Data Science* (BiCDaS). Ein erfolgreiches, am BiCDaS entwickeltes Format zur Förderung des datenbezogenen Diskurses sind die DataLabs. DataLabs sind agile, interdisziplinäre Kooperationsformate in der Forschung, die anhand eines Themenschwerpunkts oder anhand von Datentypen konzipiert sind und die den wissenschaftlichen Dialog fördern. Bei der Unterstützung datenbezogener Forschung in ihrer gesamten Breite sind dem BiCDaS dennoch Grenzen gesetzt. Deshalb wurde mit den *Scientific Data Services* (SDS) ein interdisziplinäres und skalierbares Unterstützungsformat für datenbezogene Forschung entwickelt, das es erlaubt, heterogene Einrichtungen sichtbarer zu machen, deren Angebote in komplementärer Weise zu bündeln und das in der Lage ist, sich flexibel an aktuelle Entwicklungen anzupassen.


Keywords: Data Science Center, Datenservices, Forschungsdaten, Interdisziplinarität

1 Bielefeld Center for Data Science: Tätigkeitsschwerpunkte und Ausrichtung

Aktuelle, hochkomplexe gesellschaftliche Herausforderungen, wie die Klimakrise oder die Corona-Pandemie, unterstreichen die Bedeutung von Fachbereichsgrenzen überschreitenden Kooperationen in Politik, Gesellschaft und Wissenschaft. Das *Bielefeld Center for Data Science* (BiCDaS) [Bi23a] begreift Data Science als einen wichtigen Treiber für wissenschaftliche Interdisziplinarität. Es ist darauf ausgerichtet, mit seinen Aktivitäten, Daten- und Analysepotenziale zu erschließen und den daten(analyse)zentrierten, wissenschaftlichen Dialog zu fördern.

Das BiCDaS ist eine stark interdisziplinär ausgerichtete, horizontale Einrichtung und

¹ Universität Bielefeld, Scientific Data Services, Universitätsstraße 25, 33615 Bielefeld, katharina.weiss@uni-bielefeld.de.

² Universität Bielefeld, Lehrstuhl für BWL, insb. Marketing, Universitätsstraße 25, 33615 Bielefeld, rdecker@uni-bielefeld.de,  <https://orcid.org/0000-0003-2972-6182>.

Netzwerk an der Universität Bielefeld. Die horizontale Struktur des BiCDaS wird dadurch deutlich, dass es seit seiner Gründung im Jahr 2017 Forschende aller Fakultäten der Universität Bielefeld – von den Geisteswissenschaften bis hin zu den Naturwissenschaften und der Medizin – sowie Mitarbeiter:innen zentraler Serviceeinrichtungen, zu den Themen Data Science und Forschungsdaten, vernetzt. Die Tätigkeitsschwerpunkte des BiCDaS umfassen im Wesentlichen Aktivitäten in fünf Kernbereichen:

1. *Unterstützung von Data-Science-Forschung:* Diese Form der Unterstützung geschieht beispielsweise durch die Organisation bzw. Co-Organisation von wissenschaftlichen Veranstaltungen auf nationaler und internationaler Ebene. So hat das BiCDaS z.B. das 12. Symposium des Center for Biotechnology (CeBiTec) „Big Data in Medicine and Biotechnology“, sowie mehrere Summerschools der European Association for Data Science (EuADS) [Eu22], mitorganisiert.
2. *Bewerben von Data Science und Schaffung von Datenbewusstsein:* Die betreffenden Aktivitäten richten sich hierbei zum einen an bereits am Thema Data Science interessierte Personen, etwa mit einer *Lecture Series Data Science*, zum anderen aber auch an Personen, die bislang mit dem Thema Daten bzw. Data Science wenig Berührung hatten. Diese Gruppe wurde z.B. mit dem *Data Science Day* adressiert, einem Ausstellungs- und Austauschformat bei dem Bielefelder Forschende ihre datenbezogenen Arbeiten in Form von Postern und/oder praktischen Demonstrationen der universitätsweiten Öffentlichkeit vorstellen.
3. *Unterstützung von Data Science in der akademischen Forschung:* Das BiCDaS unterstützt Forschende u.a. bei der Einwerbung von (großen) datenbezogenen Drittmittelprojekten, wie etwa dem Leibniz Wissenschaftscampus SOEP RegioHub [LS23].
4. *Förderung der Data-Science-Lehre:* Das BiCDaS fördert Data-Science-Lehre sowohl im engeren Sinn, etwa durch die Unterstützung bei der Umsetzung eines Masterstudiengangs Data Science [Ze23], als auch im weiteren Sinn im Bereich der Data-Literacy-Lehre, die Studierenden aller Fachrichtungen Datenkompetenzen vermittelt [Ha21]. Hierzu wurde z.B. ein eigenes Modul „*Data Literacy - Kulturtechnik des 21. Jahrhunderts*“ im individuellen Ergänzungsbereich geschaffen, der allen Studierenden der Universität Bielefeld offensteht. Studierende können in diesem Modul eine Ringvorlesung, speziell entwickelte Vertiefungsveranstaltungen (z.B. zu den Themen Data Storytelling und digitale Kompetenzen) sowie dafür geöffnete fachspezifische, datenbezogene Veranstaltungen belegen.
5. *Data Science als „Brücke zur Welt“:* Hierbei werden Forschende mit regionalen, überregionalen und internationalen Stakeholdern aus Gesellschaft, Wirtschaft und Wissenschaft vernetzt und Kooperationen mit diesen angebahnt. So hat das BiCDaS z.B. im Rahmen des Forum Wissenschaftskommunikation im Jahr 2022 eine Podiumsdiskussion zum Thema „Von hier wech! – Wissenschaft trifft Bürgerschaft an der regionalen Datenquelle“ mit regionalen und nationalen gesellschaftlichen

Stakeholdern organisiert.

Neben diesen fünf Kernbereichen verfolgt das BiCDA_S noch zwei zu den Kernbereichen horizontal liegende Querschnittsthemen, nämlich

(1) den *datenzentrierten wissenschaftlichen Dialog* und (2) das Thema *Datenethik*. Das BiCDA_S fördert Ersteres u.a. mit dem Format der *BiCDA_S DataLabs*, das in Abschnitt 2 ausführlich dargestellt wird. Zum Thema Datenethik organisiert das BiCDA_S u.a. Veranstaltungen wie etwa eine Paneldiskussion zum Thema „*Legal and Ethical Aspects of Big Data in Medicine and Biotechnology*“ im Rahmen des vom BiCDA_S mitorganisierten CeBiTec-Symposiums „*Big Data in Medicine and Biotechnology*“ im Jahr 2018.

2 BiCDA_S DataLabs: Ein agiles Kollaborationsformat für interdisziplinäre Forschungskoperationen

Exemplarisch für eine BiCDA_S-Aktivität, die sich besonders förderlich für die interdisziplinäre Zusammenarbeit erwiesen hat, steht das Format der BiCDA_S DataLabs [Bi23a]. DataLabs sind ein agiles Format für Forschungskoperationen. Sie sind sowohl themenzentriert als auch bedarfsorientiert konzipiert. Ein DataLab kann ohne formalen Aufwand (wie ihn z.B. der Aufbau einer zentralen wissenschaftlichen Einrichtung nach sich ziehen würde) sehr schnell zu einem gegebenen Forschungsinteresse eingerichtet werden, da der dafür erforderliche administrative Unterbau bereits durch das BiCDA_S geben ist. Inhaltlich kann ein DataLab beispielsweise auf eine spezifische Dateninfrastruktur, ein Werkzeug, einen konkreten Datensatz oder bestimmte Methoden und Modelle fokussiert sein.

DataLabs verfolgen zwei wesentliche Ziele: (1) die Förderung des interdisziplinären Austausches der Forschenden untereinander und (2) die Bündelung der Expertise zu einem ausgewählten datenbezogenen Thema, einschließlich deren Sichtbarkeit nach außen. Zudem bieten DataLabs den Vorteil, schnell auf thematisch passende Förderaufträge reagieren zu können. Dies ist in Anbetracht der stark verkürzten Intervalle zwischen Förderauftrag und Einreichungstermin, insbesondere für interdisziplinär angelegte Ausschreibungen, relevant.

Zum Zeitpunkt der Erstellung dieses Beitrags bestanden drei BiCDA_S DataLabs: (1) das DataLab „Digital Humanities“ (2) das DataLab „HPC and Advanced Simulations“ sowie (3) das DataLab „Corona – Data, Impact and Solutions“. Des Weiteren befanden sich zwei weitere DataLabs zu den Themen „Nachhaltigkeit“ und „Inklusion“ in der Entwicklungsphase. Die Konzeption der BiCDA_S DataLabs war von verschiedenen Initiativen des *Berkley Institute for Data Science* [Bi23b] inspiriert, innerhalb derer Forschende über Disziplingrenzen hinweg zu bestimmten Datentypen arbeiten (z.B. ImageXD, TextXD, GraphXD). Eine fokussierte Ausrichtung der Bielefelder DataLabs an häufig genutzten Datentypen, wie z.B. Bilddaten oder Textdaten, wurde von Bielefelder Forschenden bislang nicht favorisiert.

Die BiCDaS DataLabs bieten vielfältige Anreize, insbesondere in Bezug auf wissenschaftliche Interdisziplinarität, die eine Beteiligung für Forschende lohnenswert machen: (1) Sie vernetzen Forschende zu einem spezifischen Thema, erleichtern somit (2) die Stellung von datenbezogenen (insb. auch interdisziplinären) Drittmittelförderanträgen und unterstützen (3) die Erstellung von gemeinsamen interdisziplinären Publikationen. Eine finanzielle Unterstützung als Anreiz erhalten die DataLabs nicht. Dies war bislang auch nicht erforderlich, da die Mitglieder der DataLabs i.d.R. bereits finanzierten Strukturen, wie z.B. einer Fakultät, angehören.

3 Vom BiCDaS zu den Scientific Data Services

Trotz der vielfältigen Data-Science-Aktivitäten in Forschung und Lehre stößt das BiCDaS in Bezug auf die Unterstützung von datenbezogener Forschung und Lehre in der gesamten, an der Universität Bielefeld vorhandenen und erwünschten Breite, an seine Grenzen. Gleichzeitig existieren an der Universität Bielefeld neben dem BiCDaS zahlreiche weitere Einrichtungen, Netzwerke und Initiativen, die komplementäre Angebote zur Unterstützung datenbezogener Forschung bereitstellen (z.B. das Zentrum für Statistik, das Kompetenzzentrum Forschungsdaten, das Servicezentrum Medical Data Science und das Data-Champions-Netzwerk). Diese Vielfalt und Heterogenität, die sowohl seitens der vorhandenen Unterstützungsangebote und Services als auch seitens der Anbieter:innen besteht, stellt einerseits eine große Herausforderung für die effiziente Unterstützung von datenbezogener Forschung und Lehre dar. So erschwert sie auf der Ebene der Nutzer:innen die Sichtbarkeit und Auffindbarkeit und damit schlussendlich auch die Nutzung geeigneter Angebote. Auf der Ebene der Anbieter:innen hemmt sie die Identifikation und Nutzung von Synergiepotenzialen, die gemeinsame Weiterentwicklung vorhandener Services sowie die Entwicklung neuer universitätsweiter Services. Andererseits bietet das Nebeneinander von Angeboten und Anbieter:innen aber auch praktische Vorteile, etwa in Form von Bereicherung durch unterschiedliche (fachliche) Perspektiven, Vielfalt der vorhandenen Angebote und hoher intrinsischer Motivation der Mitarbeiter:innen in den verschiedenen Einrichtungen.

Um den im vorliegenden Kontext bestehenden Herausforderungen zu begegnen und gleichzeitig die Chancen der Heterogenität sowie des Nebeneinanders von Anbieter:innen und Angeboten optimal zu nutzen, wurden die *Scientific Data Services* (SDS) [SD23] als lehr- und forschungsnaher Dienstleistungseinrichtung gegründet. Die SDS verbinden an der Universität Bielefeld vorhandene datenbezogenen Einrichtungen, Initiativen und Netzwerke unter einem gemeinsamen Dach. Sie machen datenbezogene Kompetenzen sichtbar(er) und aktivieren Synergiepotenziale. Die SDS können aus drei unterschiedlichen Perspektiven betrachtet werden, und zwar der (1) datenbezogenen Einrichtungen, der (2) datenbezogenen Kompetenzen und der (3) datenbezogenen Services. Diese Unterteilung ermöglicht eine Strukturierung der Heterogenität bereits vorhandener Services und Einrichtungen, um auf diese Weise zum einen, das Auffinden benötigter Unterstützung zu erleichtern und zum anderen, die Zuordnung und Integration

von (neuen) Einrichtungen und Projekten zu den SDS zu vereinfachen. Die Perspektive der *datenbezogenen Einrichtungen* erleichtert hierbei die Zuordnung von datenbezogenen Einrichtungen und Projekten zu den SDS. Die Perspektive der *datenbezogenen Kompetenzen* vereinfacht für Nutzer:innen das Auffinden gesuchter Unterstützung anhand benötigter *datenbezogener Kompetenzen*, z.B. im Bereich Datendarstellung. Die Perspektive der *datenbezogenen Services* vereinfacht hingegen das Auffinden von Unterstützung anhand benötigter *Services*, z.B. Beratung zu Fragen des Datenschutzes. Diese drei Strukturierungsperspektiven wurden auch beim Aufbau der SDS-Website genutzt (s. Abb. 1).

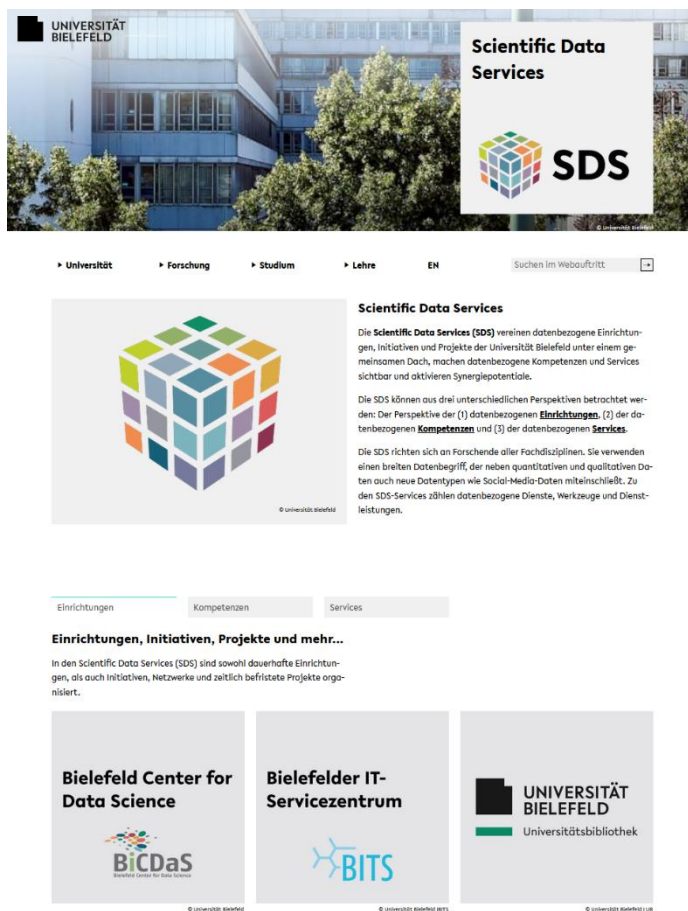


Abb. 1 zeigt einen Screenshot der aktuellen SDS-Website

Zu den unter dem Dach der SDS vereinten datenbezogenen Einrichtungen, Initiativen und Netzwerken zählen zum Zeitpunkt der Erstellung dieses Beitrags, neben dem BiCDaS,

auch zentrale Einrichtungen, wie das Bielefelder IT-Servicezentrum, die Universitätsbibliothek und das Zentrum für Statistik, sowie fachspezifische Einrichtungen, wie das Servicezentrum Medical Data Science, Netzwerke, wie das Open-Science-Netzwerk Bielefeld und verschiedene Data-Literacy-Initiativen. Aufgrund seiner horizontalen Netzwerkstruktur wird das BiCDaS gegenwärtig zu einem zentralen Element der SDS weiterentwickelt. Unter der Perspektive der Kompetenzen sind Services und Angebote grob entlang des Datenlebenszyklus von „Daten sammeln“ bis hin zu „Daten archivieren“ geordnet. Unter der Perspektive der datenbezogenen Services werden folgende Aspekte erfasst: (1) Beratung, (2) Vernetzung, (3) Datenzugang, (4) Infrastruktur, (5) Drittmittel, (6) Werkzeuge und (7) Veranstaltungen. Die SDS sind bewusst flexibel erweiterbar und skalierbar konzipiert, um sich schnell an neue wissenschaftliche Entwicklungen sowie die sich kontinuierlich verändernde datenbezogene Projekt-, Angebots- und Servicelandschaft anpassen zu können. Die SDS richten sich an Forschende aller Fachdisziplinen und verwenden deshalb bewusst einen breiten Datenbegriff, der neben quantitativen und qualitativen Daten auch neue Datentypen, wie multimodale Social-Media-Daten, miteinschließt.

Die Ausstattung der SDS lässt sich in drei Bereiche unterteilen, (1) die Ausstattung der *Dachorganisation*, (2) die Ausstattung der *Untereinheiten* sowie (3) die Ausstattung mit *Drittmitteln*. Als *Dachorganisation* verfügen die SDS über eine Dauerstelle, die das Management der SDS als Ganzes übernimmt. Des Weiteren verfügt die Dachorganisation über zusätzliche Mittel, die für studentische und wissenschaftliche Hilfskräfte, aber z.B. auch für die Durchführung von Veranstaltungen eingesetzt werden können. Die zu den SDS gehörenden *Untereinrichtungen* wiederum sind i. d. R. ihrerseits mit finanziellen Mitteln ausgestattet und finanzieren auch die dort tätigen Mitarbeiter:innen. Aufgrund der Heterogenität in Bezug auf die Organisationsform und den Aufbau der SDS-Untereinrichtungen unterscheiden sich auch deren personelle Ausstattungen mitunter erheblich. Das Spektrum reicht von Einrichtungen mit vielen dauerhaft finanzierten Stellen bis hin zu Einrichtungen mit nur wenigen befristeten Stellen. Exemplarisch für eine SDS-Untereinrichtung soll hier die Ausstattung des BiCDaS als zentrales Element der SDS detaillierter dargestellt werden.

In der Aufbauphase (von 2017 bis 2022) verfügte das BiCDaS über eine Geschäftsführungsstelle, die aufgrund des Projektcharakters befristet eingerichtet worden war. Ferner verfügte das BiCDaS über eine drittmittelfinanzierte Stelle im Bereich Data Literacy Education. Mit der Integration als zentrale Untereinrichtung in die SDS im Jahr 2023 wurde die Geschäftsführungsfunktion im BiCDaS vom SDS-Management mit übernommen. Nach Beendigung der BiCDaS-Aufbauphase und den sich damit verändernden Aufgaben lag dies unter organisationsökonomischen Gesichtspunkten nahe. Mit dem Übergang in den Regelbetrieb fallen im BiCDaS nunmehr hauptsächlich Managementaufgaben und weniger Gestaltungsaufgaben an. Zudem können Routineaufgaben im Zusammenhang mit etablierten BiCDaS-Formaten nun von wissenschaftlichen und studentischen Hilfskräften der SDS übernommen werden. Zum BiCDaS-Kernteam gehört bis zum Ende der betreffenden Drittmittelförderung auch weiterhin eine Stelle im Bereich von Data Literacy Education, für die sich gegenwärtig um eine Anschlussfinanzierung bemüht

wird. Die sonstigen Mitglieder des BiCDaS werden über andere, vom BiCDaS unabhängige universitäre Strukturen, insb. Fakultäten oder zentrale Einrichtungen finanziert.

Sowohl bei der SDS-Dachorganisation als auch bei etlichen Unterorganisationen der SDS werden für spezielle Themen regelmäßig *Drittmittel* eingeworben. So wird beispielsweise, wie bereits oben kurz erwähnt, am BiCDaS das Thema Data Literacy bislang durch Drittmittelprojekte abgedeckt (z.B. DataLiteracySkills@OWL, Bi_DaNte). Ferner besteht für Drittmittelprojekte auch die Möglichkeit, zeitlich auf ihre Förderdauer begrenzt, Untereinrichtung der SDS zu werden. Gegenwärtig trifft dies etwa auf den Leibniz WissenschaftsCampus SOEP RegioHub zu.

Aufgrund der Heterogenität in den Organisationsformen der SDS-Untereinrichtungen sind derzeit auch die Incentivierungen der angebotenen Services ganz unterschiedlich gestaltet. Diese reichen von indirekten, eher ideellen Incentives wie etwa der Vernetzung mit an einem Thema interessierten Kolleg:innen, der Erhöhung der Sichtbarkeit eigener Kompetenz und Forschung (z.B. über das Data-Champions-Netzwerk), dem Zugang zu Daten aus verschiedensten Anwendungsgebieten (z.B. über das Zentrum für Statistik) bis hinzu Co-Autor:innenschaften und stellenmäßiger Beteiligung an beantragten Drittmittelprojekten (z.B. beim Statistik-Beratungs-Centrum). Für die SDS als Ganzes befinden sich derzeit spezifische Incentivierungsmodelle in der Entwicklungsphase.

4 Erste Erfahrungen beim Aufbau von SDS

Zu unseren wesentlichen Erfahrungen beim Aufbau der SDS zählt, dass die Unterstützung zentraler Hochschulgremien für die Akzeptanz in der Breite und die Schnelligkeit der Umsetzung von großem Nutzen ist. Förderlich erwiesen und erweisen sich überdies die folgenden Aspekte:

- eine flexible und skalierbare (Dach-)Struktur, die sich schnell an Änderungen bei Einrichtungen, Projekten und Serviceangeboten anpassen lässt,
- die Strukturierung heterogener, sich z.T. kontinuierlich verändernder Einrichtungen und Angebote anhand der oben erwähnten drei Perspektiven,
- die aktive Förderung von interdisziplinärem datenbezogenem Austausch und von fachgebietsübergreifender Vernetzung,
- ein breiter Datenbegriff, der neben quantitativen und qualitativen Daten auch neuere Datentypen wie multimodale Social-Media-Daten miteinschließt,
- und die partizipativ organisierte Weiterentwicklung vorhandener Formate und Angebote.

5 Ausblick und Weiterentwicklungsmöglichkeiten

Um die SDS-Angebote weiterzuentwickeln, sollen zukünftig drei bereits in der BiCDaS-Projektphase erfolgreich bearbeitete Arbeitsfelder weiter ausgebaut werden, und zwar (1) *datenbezogene Drittmittelservices*, (2) *datenbezogene Forschungsservices* sowie (3) *datenbezogene Forschungsaktivitäten*. Die *datenbezogenen Drittmittelservices* sollen die ganze Bandbreite der inhaltlichen Unterstützung datenbezogener, interdisziplinärer Drittmittelprojekte abdecken (resp. die Initiierung, Unterstützung und Durchführung von Drittmittelanträgen). Die *datenbezogenen Forschungsservices* sollen zum einen eine umfassende Unterstützung von Forschenden bei konkreten datenbezogenen Projekten und Problemstellungen (z.B. beim Textmining oder komplexen Datendarstellungen), zum anderen aber auch Hilfestellungen beim Auffinden bereits vorhandener Forschungsdaten und Forschungswerkzeuge bieten. Für Letzteres sind als Ergänzung der SDS-Webseite zwei Online-Portale geplant: Das erste Portal soll systematisch die an der Universität Bielefeld verfügbaren Daten aus (großen) Forschungsprojekten identifizieren, im zulässigen Rahmen sichtbar machen und ggf. Zugangsmöglichkeiten darstellen. Das zweite Portal soll einen Überblick über an der Universität entwickelte Forschungswerkzeuge (wie z.B. Textmining-Algorithmen und -Codes), bieten. *Datenbezogene Forschungsaktivitäten* schließlich sollen gezielt aktuelle, interdisziplinär zu adressierende Themen (z.B. Data Storytelling) und Forschungsfragen im Umfeld von Data Science, Forschungsdatenmanagement und Data Literacy Education bearbeiten und Forschungs Kooperationen zu diesen fördern (u.a. in Form von Überblicksartikeln, Strategiepapieren oder konkreten (Drittmittel-)Forschungsprojekten).

Abschließend lässt sich feststellen, dass sich die Scientific Data Services mit ihrer aktuellen Struktur und ihrem inhaltlichen Profil auch sehr gut in das Forschungsprofil der Universität Bielefeld einfügen, das dem Grundsatz der „Transcending Boundaries“ folgt: „Forschung an der Universität Bielefeld heißt Grenzen zu überwinden – zwischen Disziplinen, zwischen Menschen und zwischen Wissenschaft und Gesellschaft.“

Danksagung

Unser Dank gilt an dieser Stelle insb. Dr. Nils Hachmeister, der als früherer Geschäftsführer das BiCDaS von seiner Gründung im Jahr 2017 bis zum Jahr 2022 in maßgeblicher Weise mit aufgebaut und geprägt hat.

Literaturverzeichnis

[Bi23a] BiCDaS, Bielefeld Center for Data Science, <https://www.uni-bielefeld.de/einrichtungen/bicdas/>, Stand: 06.04.2023.

[Bi23b] Berkley Institute for Data Science, <https://bids.berkeley.edu/research/image-xd>, Stand:

04.06.2023.

- [Eu22] EuADS, European Association for Data Science <https://www.euads.org/fjkdlasjdiglsmdgkcxjhvckh/euads-summer-school/>, Stand: 04.06.2023.
- [LS23] LSC, Leibniz ScienceCampus SOEP RegioHub, <https://lsc-soep-regiohub.com/>, Stand:04.06.2023.
- [Ha21] Hachmeister, N.; Weiß, K.; Theiß, J.; Decker, R. Balancing plurality and educational essence: higher education between data-competent professionals and data self empowered citizens. [Data](#), 6/10, S. 1-15. 2021.
- [SD23] SDS, Scientific Data Services, <https://www.uni-bielefeld.de/einrichtungen/sds/index.xml>, Stand: 06.04.2023.
- [Ze23] ZeSt, Zentrum für Statistik, <https://www.uni-bielefeld.de/einrichtungen/zest/master-data-science/>, Stand:04.06.2023.

Towards the Establishment of Data Science Centers at Higher Education Institutions

Poster Session

Johannes Putzke¹, Thomas Schimmer², Manuela Richter³ and Lucas Hamel⁴

Abstract: Numerous Higher Education Institutions (HEI) are considering the launch of data science centers, partly induced by a call from the German Federal Ministry of Education and Research (BMBF). However, data science centers at HEIs substantially differ from their counterparts in industry. In this brief paper, we highlight some of these differences and identify a challenge in setting up data science centers at HEIs using the case of an initiative at eight universities of applied sciences in Germany.

Keywords: Data Science, FAIR Data, Research Data Management, RDM, skilled labor shortage, Universities of Applied Sciences, UAS

1 Introduction

More than a decade has passed since Davenport and Patil's [DP12] seminal Harvard Business Review article, "Data Scientist: The Sexiest Job of the 21st Century". During this decade, many scholars, practitioners, and policymakers alike acknowledged the need for qualified data scientists. Consequently, many universities have established data science centers or are considering doing so.⁵

However, launching data science centers at higher education institutions (HEIs) presents certain challenges. This paper reports one of these challenges, using the case of the German project "Research Data Management at Universities of Applied Sciences in the State of Rhineland-Palatinate – FDM@HAW.rlp" and proposes a solution. The structure of the paper is as follows: First, we highlight two particular characteristics of data science centers

¹ Trier University of Applied Sciences, Umwelt-Campus Birkenfeld, Campusallee, 55768 Hoppstädten-Weiersbach, Germany, joh.putzke@hochschule-trier.de

² Mainz University of Applied Sciences, FDM@HAW.rlp project office, Lucy-Hillebrand-Str.2, 55128 Mainz, Germany, thomas.schimmer@hs-mainz.de

³ Mainz University of Applied Sciences, FDM@HAW.rlp project office, Lucy-Hillebrand-Str.2, 55128 Mainz, Germany, manuela.richter@hs-mainz.de

⁴ Koblenz University of Applied Sciences, RheinAhrCampus, Joseph-Rovan-Allee 2, 53424 Remagen, Germany, hamel@hs-koblenz.de

⁵ In Germany, for example, the universities of Bremen, Kiel or Mannheim have established data science centers. Furthermore, the German Federal Ministry of Education and Research (BMBF) selected 22 data science initiatives that intend to build up data competence centers. From these 22 initiatives, up to 10 initiatives will be financed for building up the data competence centers for a period of three years.

at HEIs compared to those at private enterprises. Next, we describe the “FDM@HAW.rlp” project, which aims to establish data science centers for research data management. Finally, we outline an unresolved challenge associated with these and other related projects, and propose a potential solution.

2 Particularities of data science centers at HEIs

Before discussing the particularities of data science centers at HEIs, it would be reasonable to provide a definition of “data science”. However, defining data science is challenging as the term constantly evolves “with the continuous advancement in technological trends” [Re22, p. 226]. Nonetheless, “Data Science” encompasses components such as “Artificial Intelligence (AI), Machine Learning (ML), Big Data (BD), Big Data Analytics (BDA), Visualization, and Business Intelligence & Analytics (BI&A), Mathematics, Computer Science & Programming Skills, and Domain Knowledge.” [Re22, p. 226]. Moreover, it is generally agreed that a deep understanding of statistics is a prerequisite for a data scientist. However, it is also crucial to gather the data science community in this workshop to explore their common understanding of “data science”.

One of the key distinguishing features of data science centers at HEIs, as opposed to those at private enterprises, could be the nature of the data that HEI data scientists work with.⁶ Over the past few years, the open science movement has led to numerous requests for this data to be “FAIR” (see [Wi16]), an acronym for (1) findable, (2) accessible, (3) interoperable, and (4) reusable. To ensure “FAIR” data, HEIs and research institutes have increasingly established research data management systems (refer to the literature reviews by [Do22] and [Pe17]). Consequently, they have created the role of “data stewardship”. Numerous authors have explored the role model and tasks of data stewardship (for instance, [Cu21, Gr21, Pe18, Ro19, Ve19, WAS23, Wh18]). As will be shown in the next section, the project under consideration also established a “RDM / data stewardship” role.

Another distinguishing feature of data science centers at HEIs is that the shortage of skilled labor is more pronounced than in industry.⁷ This discrepancy can be attributed to the fact that the salaries of data scientists working in industry are often significantly higher than those of their counterparts at HEIs. This disparity initiates what we term the “vicious cycle of double shortage of skilled labor”. Data scientists from HEIs (often with PhD degrees) transition to industry due to the higher salaries. Consequently, these professionals are absent at HEIs, leaving a gap in the education of future data scientists who could help alleviate the skilled labor shortage.

⁶ It should be noted that the type of data handled in data science centers can widely vary, regardless of the private or public sector.

⁷ Furthermore, there may be significant differences in the expectations for the specific services offered by data centers at private enterprises and HEIs. For instance, while HEIs may have a strategic goal of educating data scientists who can apply their knowledge across various industries, this may not necessarily be the case for private enterprises (exceptions may apply).

3 The project “FDM@HAW.rlp”

In this section, we present the German project “Research Data Management at Universities of Applied Sciences in the State of Rhineland-Palatinate – FDM@HAW.rlp”. This project involves eight HEIs aiming to ensure ‘FAIR’ data within their institutions. To accomplish this, they are working to develop competences, services, and technical infrastructure for quality assured Research Data Management (RDM). The project receives funding from the German Federal Ministry of Education and Research (BMBF) and is financed within the Recovery and Resilience Facility of the European Union for a period of three years. The project’s primary objective is not data analyses (as may be the case in data science centers in industries), but rather ensuring “FAIR” data.

The project is organized as follows: An “RDM scout” is appointed at each of the eight Universities of Applied Sciences (UAS). These RDM scouts serve as the main point of contact for researchers at their respective institutions. In addition, there are four ‘RDM/Data stewards’ responsible for the following scientific domains: (1) engineering, (2) informatics/computer science, (3) life sciences and natural sciences, and (4) humanities and social sciences. These “RDM stewards” are officially based at four of the UAS but are tasked with offering services to researchers at all eight UAS. There is also a “central coordination unit” (equivalent to one full-time position) officially located at one of the UAS. After three years, the roles of “data scouts”, “RDM stewards”, and the “central coordination unit” are expected to evolve into eight sustainable RDM officer positions, with one at each of the eight UAS.

4 Unresolved challenges

The project exhibits several strengths, such as offering discipline-specific services to researchers despite limited resources and funding. However, certain unresolved challenges persist. For instance, it remains uncertain where the sustainable RDM officer positions will ultimately be located. Recent research driven by library personnel suggests that the RDM centers should be situated within libraries. However, alternative placements of these centers can also be envisaged. For instance, infrastructure specialists from the technical side might advocate for a location within the universities’ data centers. Additionally, data scientist or scholars with robust artificial intelligence skills or domain knowledge might argue that these roles should be part of the computer science departments or demand specialized data science positions within the domain departments. Consequently, we contend that aligning data science positions within an existing structure is misguided. Instead, data science roles should be housed in newly created, independent data science centers. These new data science centers should include data scientists, data stewards, and possibly researchers with domain knowledge (for an explanation of the different roles see [SD21]). Ideally, individuals could even embody several of these roles.

Acknowledgements

The authors thank Raghvendra Jain for proofreading the paper. The authors would like to acknowledge the partner institutions of the project FDM@HAW.rlp: Mainz University of Applied Sciences (funding number 16FDFH104A), Bingen Technical University of Applied Sciences (funding number 16FDFH104B), Kaiserslautern University of Applied Sciences (funding number 16FDFH104C), Koblenz University of Applied Sciences (funding number 16FDFH104D), Ludwigshafen University of Business and Society University of Applied Sciences (funding number 16FDFH104E), Trier University of Applied Sciences (funding number 16FDFH104F), Hochschule Worms University of Applied Sciences (funding number 16FDFH104G), Catholic University of Applied Sciences Mainz (funding number 16FDFH104H).

Funding





The project FDM@HAW.rlp is funded by the German Federal Ministry of Education and Research (BMBF) within the funding measure "Reuse and Management of Research Data at Universities of Applied Sciences" (funding number 16FDFH104) and financed within the Recovery and Resilience Facility of the European Union. The authors are responsible for the content of this publication.

Bibliography

- [Cu21] Curdt, C.; Dierkes, J.; Helbig, K.; Lindstädt, B.; Ludwig, J.; Neumann, J.; Parmaksiz, U.: Data Stewardship im Forschungsdatenmanagement - Rollen, Aufgabenprofile, Einsatzgebiete: Überblick: 11. DINI/nestorWorkshop, 16. und 17.11.2020. Bausteine Forschungsdatenmanagement, (3):70–81, 2021.
- [Do22] Donner, E. K.: Research data management systems and the organization of universities and research institutes: A systematic literature review. *Journal of Librarianship and Information Science*, 55(2):261–281, 2022.
- [DP12] Davenport, T. H; Patil, DJ: Data scientist: The sexiest job of the 21st century. *Harvard business review*, 90(5):70–76, 2012.
- [Gr21] Gruber, A.; Schranzhofer, H.; Knopper, S.; Stryeck, S.; Hasani-Mavriqi, I.: Kompetenzen von Data Stewards an österreichischen Universitäten. *Mitteilungen der VÖB*, 74(1):12–32, 2021.
- [Pe17] Perrier, L.; Blondal, E.; Ayala, A. P.; Dearborn, D.; Kenny, T.; Lightfoot, D.; Reka, R.; Thuna, M.; Trimble, L.; MacDonald, H.: Research data management in academic institutions: A scoping review. *PLOS ONE*, 12(5):1–14, 2017.
- [Pe18] Peng, G.: The State of Assessing Data Stewardship Maturity - An Overview. *Data Science Journal*, 17(7):1–12, 2018.

- [Re22] Reddy, R. C.; Bhattacharjee, B.; Mishra, D.; Mandal, A.: A systematic literature review towards a conceptual framework for enablers and barriers of an enterprise data science strategy. *Information Systems and e-Business Management*, 20(1):223–255, 2022.
- [Ro19] Rothfritz, L.: *Data Stewardship als Boundary-Work*, 2019.
- [SD21] Steinmann, L.; Drechsler, R.: Verzahnung von Data Stewardship und Data Science – Wege und Perspektiven. *Bausteine Forschungsdatenmanagement*, (3):82–91, 2021.
- [Ve19] Verheul, I.; Imming, M.; Ringerma, J.; Mordant, A.; Ploeg, J.-L. v. d.; Pronk, M.: *Data Stewardship on the map: A study of tasks and roles in Dutch research institutes*, 2019.
- [WAS23] Wendelborn, C.; Anger, M.; Schickhardt, C.: What is data stewardship? Towards a comprehensive understanding. *Journal of Biomedical Informatics*, 140:104337, 2023.
- [Wh18] Whyte, A.; de Vries, J.; Thorat, R.; Kuehn, E.; Sipos, G.; Cavalli, V.; Kalaitzi, V.; Ashley, K.: *Skills and Capability Framework*, 2018.
- [Wi16] Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; Bouwman, J.; Brookes, A. J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, C. T.; Finkers, R.; Gonzalez-Beltran, A.; Gray, A. J. G.; Groth, P.; Goble, C.; Grethe, J. S.; Heringa, J.; 't Hoen, P. A. C.; Hooft, R.; Kuhn, T.; Kok, R.; Kok, J.; Lusher, S. J.; Martone, M. E.; Mons, A.; Packer, A. L.; Persson, B.; Rocca-Serra, P.; Roos, M.; Schaik, R. v.; Sansone, S.; Schultes, E.; Sengstag, T.; Slater, T.; Strawn, G.; Swertz, M. A.; Thompson, M.; Lei, J. v. d.; Mulligen, E. v.; Velterop, J.; Waagmeester, A.; Wittenburg, P.; Wolstencroft, K.; Zhao, J.; Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018, 2016.

Das Konzept des House of Computing and Data Science als Blaupause zur Digitalisierung der Forschung

Chris Biemann ¹, Tilo Böhmann², Sebastian Gerling ³, Jan Louis ⁴ und Martin Semmann ^{5 6}

Abstract: Forschung auf internationalem Niveau kommt nicht mehr ohne die strukturelle Nutzung der Potenziale der Digitalisierung, dabei insbesondere in den Bereichen Data Science und Künstlicher Intelligenz, aus. Gleichzeitig ist der Brückenschlag in die wissenschaftliche Anwendung aufgrund sehr heterogener Fächerkulturen, Vorerfahrungen und der hohen Methodenvielfalt nicht trivial. Ein vielversprechender Ansatz, mit diesen Herausforderungen umzugehen, ist das Konzept des House of Computing and Data Science. Als zentrale Einrichtung wird das Ziel eines reziproken, gemeinsamen Erkenntnisgewinns und Ausbau von Methodenkompetenz verfolgt. Dabei werden Impulse in die gesamte Universität gegeben, gleichzeitig aber auch Erfahrungen und Kompetenzen aus den dezentralen Bereichen gebündelt.

Keywords: Data Science, Künstliche Intelligenz, Digitalisierung, Organisationsstruktur, Research Software Engineering.

1 Einleitung


Die Digitalisierung ist ein zentraler Treiber von Veränderungen für Unternehmen, die Zivilgesellschaft, aber auch die Forschung. Damit einher gehen sich verändernde Anforderungen, um erfolgreiche, sichtbare und nicht zuletzt auch wirkungsvolle Forschung durchzuführen. So sind mittlerweile bspw. über 2/3 der Forschenden in ihrer Arbeit abhängig von Forschungssoftware [He18]. Zudem werden Forschungsaktivitäten zunehmend kollaborativer innerhalb der Disziplinen, was ebenfalls zu einer Dynamisierung von Wissenschaft geführt hat und immer noch führt [WJU07]. Die Digitalisierung erfordert vor dem Hintergrund heterogener Disziplinen eine Anpassung in

¹ Universität Hamburg, HCDS & Fachbereich Informatik, Vogt-Kölln-Str. 30, 22527 Hamburg,

chris.biemann@uni-hamburg.de,  <https://orcid.org/0000-0002-8449-9624>

² Universität Hamburg, Fachbereich Informatik, Vogt-Kölln-Str. 30, 22527 Hamburg, tilo.boehmann@uni-hamburg.de


³ Universität Hamburg, CDO, Mittelweg 177, 20148 Hamburg, sebastian.gerling@uni-hamburg.de,

 <https://orcid.org/0000-0001-6047-3481>

⁴ Universität Hamburg, Fachbereich Physik, Luruper Chaussee 149, 22761 Hamburg, jan.louis@uni-

hamburg.de  <https://orcid.org/0000-0001-7254-6641>

⁵ Universität Hamburg, HCDS, Vogt-Kölln-Str. 30, 22527 Hamburg, martin.semmann@uni-hamburg.de,

 <https://orcid.org/0000-0001-5316-3696>

⁶ Die Reihung der Autoren erfolgt alphabetisch.

drei zentralen Bereichen.

1. Die Adaption und Nutzung digitaler Methoden in den Disziplinen
2. Die konsequente Weiterentwicklung bestehender Methoden sowie Forschung an neuen Methoden, insbesondere auch durch Impulse aus der Anwendung
3. Der kritische Diskurs der Folgen von Digitalisierung auf die einzelnen Fachdisziplinen, Forschung generell, aber auch auf Gesellschaften

Dieses Spannungsfeld auszutariieren stellt eine Herausforderung, vor allem für Volluniversitäten dar, da ein gemeinsames Handeln und miteinander lernen erforderlich ist. Dies ist insbesondere der Breite der potenziellen Anwendungsdisziplinen und den damit einhergehenden umfassenden methodischen Herausforderungen und Anpassungsbedarfen deutlich. Neben solchen relevanten weichen Faktoren bedarf es aber auch eines Veränderungsprozesses in der Wahrnehmung digitaler Infrastrukturen [KMY21]. Wissenschaftliche Software stellt einen Treiber und Ermöglicher von Erkenntnisgewinn dar, sollte daher auch entsprechend als relevanter Forschungsbeitrag beurteilt werden [a]. Dieser Beitrag manifestiert sich in der technologischen Unterstützung von Verfahren innerhalb der Forschungskontexte, wodurch sowohl Effizienzgewinne, aber auch eine grundsätzliche Durchführbarkeit von Forschungstätigkeiten ermöglicht wird. So könnten bspw. durch automatisierte Analysemöglichkeiten großer Textkorpora Veränderungen im Zeitverlauf beforscht werden, die ohne die entsprechenden Technologien nicht bzw. nicht in diesem Umfang möglich wären. Zudem gelten insbesondere Anforderungen des wissenschaftlichen Arbeitens auch für wissenschaftliche Software [AIRe13, Ch22, Ha20, HePa15]. Die dazu nötige Professionalisierung der Entwicklung von Forschungssoftware kann allerdings nicht in allen Disziplinen gleichermaßen durch Promovierende, die die Treiber der (Weiter-)Entwicklung von Forschungssoftware sind, gesichert werden. Dazu bedarf es eines Brückenschlags in die Informatik, als (auch) ingenieurwissenschaftlich geprägte Fachdisziplin, die gerade in der systematischen Entwicklung von Software einen starken Ankerpunkt hat.

2 Der Lösungsansatz: Das House of Computing and Data Science

Ausgehend von diesen Rahmenbedingungen, wurde das House of Computing and Data Science (HCDS) an der Universität Hamburg Ende 2021 als eine zentrale Einrichtung der Hochschule gegründet. Die Initialisierungsphase wurde mit insgesamt knapp 8 VZÄ bestritten, die neben der wissenschaftlichen Leitung, der Geschäftsführung auch (Senior) Softwareentwickler und die Koordination von Digitalkompetenzen umfassen. Das HCDS dient als Blaupause für den Ausbau, Weiterentwicklung und Transfer digitaler Methodenkompetenz (Ebene 1) und gleichermaßen für interdisziplinäre Forschung mit deutlichem Beitrag auch in den Bereichen Data Science und künstlicher Intelligenz (Ebene 2). Neben diesen fachlichen Schwerpunkten, zielt das Leitbild des House of Computing and Data Science auf eine Open Science-Kultur ab, wobei parallel insbesondere die

Grundsätze einer professionellen Softwareentwicklung, im Sinne eines wissenschaftlichen Artefaktes, dem Aufbau und der Pflege einer Community von Forschenden mit Affinität für digitale Methoden und Digitalisierung, aber auch der Weiterbildung in den Disziplinen [Co21, HePa15].

Das Konzept des HCDS verbindet diese Aspekte über die Ebenenstruktur und übergreifende Elemente. Aufgrund der Etablierung als eine zentrale Hochschuleinheit, stellt eine wichtige Rolle die Vernetzung und Etablierung als erster Kontaktpunkt dar. Dies geschieht auf Ebene 1, dem Methodenkompetenzzentrum. Hier werden Anfragen aus der Breite der Hochschule aufgenommen, Beratungsformate etabliert und bei Bedarf Weitervermittlung zu spezifischeren Expertisen. Damit wird ein niedrigschwelliges Angebot realisiert, dass es Forschenden ermöglicht Unterstützung in der Digitalisierung und Data Science zu erhalten. Dabei wird die Digitalisierung als Treiber exzellenter Forschung sehr umfassend verstanden, weshalb verschiedene Kompetenzprofile innerhalb des HCDS vorgehalten werden müssen. Zentrale Anforderung ist dabei eine ausgeprägte Kompetenz in der Informatik, insb. Methoden und Technologien der künstlichen Intelligenz und Digitalisierung, gleichzeitig aber auch einschlägige Erfahrungen in weiteren Fachdisziplinen. Hiermit kann ein schneller Transfer und Adaption von digitalen Innovationen gesichert werden. Insbesondere Anfragen, die Standardisierungspotentiale bieten oder mit bereits bestehenden Methoden und Lösungen realisiert werden können, werden durch Mitarbeitende der zentralen Einheit direkt umgesetzt. Beispielsweise zeigen sich bereits Überschneidungen bei den Anforderungen aus den Geisteswissenschaften. Hier bietet sich die Chance eine Toolsuite zu entwickeln, die insbesondere für Zeitreihenanalysen verschiedener Textkorpora dient. Eine Weiterentwicklung kann dabei dann zu potenziellen Mehrwerten bei allen Anwendenden führen. Komplexere Anfragen, die auch ein Potential für gemeinsame Forschungstätigkeiten bergen, werden auf Ebene 2 weiter spezifiziert. Hierfür wurde ein Förderprogramm initiiert, dass sogenannte Cross-Disciplinary Labs (CDL) fördert. Zielsetzung ist, interdisziplinäre Verbünde zu fördern, die durch die Kooperation wissenschaftliche Beiträge in den beteiligten Disziplinen erarbeiten. Also Forschungsvorhaben, die als Tandem zwischen Informatik und Anwendung positioniert werden. Hiermit kann die interdisziplinäre Kooperation initiiert und ausgeweitet werden, um eine Stärkung der digitalen Methodenwissenschaft und der Fachwissenschaft zu erreichen.

Das Scharnier der beiden Ebenen bildet der systematische Austausch von Methodenwissen, damit einerseits Wissen über die Einführung und Adaption in Fachwissenschaften vorgehalten und nachgenutzt werden kann. Gleichermaßen werden aber auch methodische Weiterentwicklungen aus den CDLs zurückbezogen und für weitere Nutzungskontexte aufbereitet.

Für das 2-Ebenenmodell des House of Computing and Data Science stehen mittel- und langfristige Forschungsk Kooperationen im Vordergrund, wobei als Ankerdisziplin Informatik bzw. informatische Methoden dienen. Solche interdisziplinären Verbünde bieten ein großes Potential zur Digitalisierung in der Breite einer Volluniversität beizutragen, da unterschiedliche Disziplinen mit ihren spezifischen Anforderungen und

Methoden, aber insbesondere auch Vorerfahrungen gezielt unterstützt werden können. Gleichzeitig können Impulse in die Weiterentwicklung von Methoden der Informatik geleistet werden.

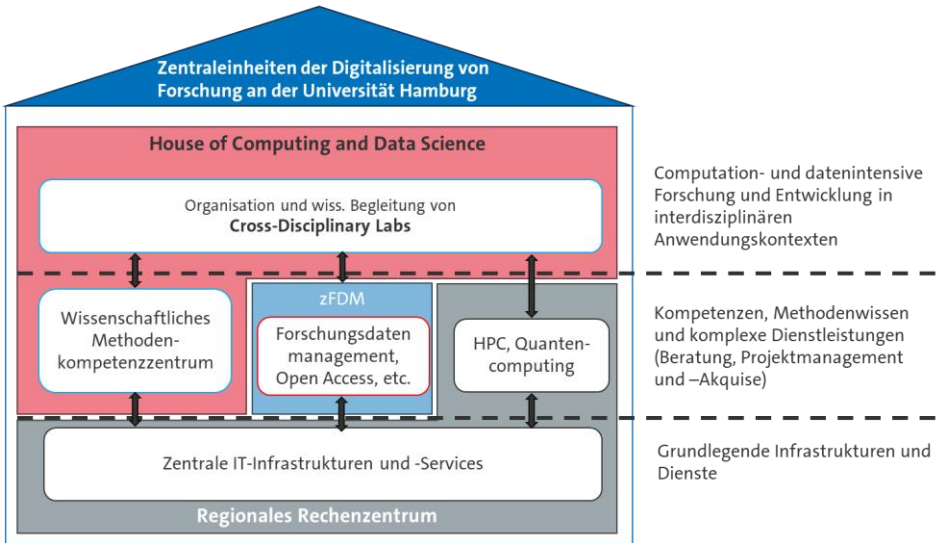


Abb. 1: 2-Ebenenmodell des House of Computing and Data Science

Über beide Ebenen hinaus besteht der generelle Bedarf nach Austausch, Professionalisierung der Softwareentwicklung im Forschungskontext und nach Bildungsangeboten. Hier kann das House of Computing and Data Science durch die Positionierung als zentrale Einheit, verortet zwischen Chief Digital Officer und Vizepräsident für Forschung, sowohl aus sich heraus, als auch aus den an Verbänden beteiligten Disziplinen ein Angebot schaffen, dass in die Breite der Hochschule wirkt und spezifische Lücken der Curricula schließt. Konkret werden bereits vereinzelt Lehrveranstaltungen positioniert, es ist aber angestrebt, dies perspektivisch und unter der Bereitstellung weiterer Ressourcen systematisch auszubauen. Dies sorgt in Konsequenz zu weiteren Potentialen für die Bildung, Forschung, aber auch den generellen Transfer in die Gesellschaft [Ho23, Pi22].

3 Beispielhafte Anwendungsfälle

Analog zur Bandbreite der Fachdisziplinen, ist das Portfolio an Unterstützungsleistungen auch umfassend und je in, sofern kapazitär möglich, maßgeschneiderter Form. Zur

Verdeutlichung werden drei beispielhafte Anwendungsfälle vorgestellt.

Die Beratung (1) durch das wissenschaftliche Methodenkompetenzzentrum ist ein niedrigschwelliges Angebot das zunächst mit einer fokussierten Klärung des Forschungsgegenstandes und Reflektion des methodischen Status quo beginnt. Dabei werden die Gespräche von Mitarbeitenden des HCDS geführt, die einschlägige, relevante Vorerfahrungen bzw. potenziell übertragbare Vorerfahrungen haben. Basierend auf den initialen Informationen können dann Potenziale digitaler Methoden bzw. vorhandener Lösungen vorgestellt werden. Ziel ist es, effizient zu einer hinreichend nutzbaren und nutzenstiftenden Lösung zu kommen. Gleichmaßen können aber nicht alle Anfragen in diesem Schritt gelöst werden. Insbesondere komplexere, zeitkritische oder entwicklungsintensive Vorhaben werden dann systematisch dahingehend geprüft, ob eine co-Finanzierung oder Akquise von Drittmitteln realistisch und zweckmäßig ist. Es resultieren Projekte (2), die als klassische Forschungsprojekte kooperativ initiiert und bearbeitet werden. Beispielhaft kann die Kooperation im Rahmen des D-WISE Projektes zur digitalen Wissenssoziologischen Diskursanalyse benannt werden [Fi23]. Abschließend gibt es das Konstrukt der Cross-Disciplinary Labs (3) als interdisziplinärer Rahmen für langfristige Kooperationen mit dem Ziel sowohl einen wissenschaftlichen Beitrag in der Methodenwissenschaft, aber auch in der Anwendungsdisziplin zu schaffen. Zeitgleich erfolgt ein systematischer Rückbezug auf das Methodenkompetenzzentrum, um die Resultate nachzuhalten. Dadurch wird sichergestellt, dass die involvierten Parteien jeweils forschend eingebunden sind und keine Anbieter-Kunde-Relation entsteht.

4 Fazit

Das vorgestellte Konzept des HCDS stellt eine gut übertragbare Blaupause dar, um dem steigenden Bedarf nach digital- und Data Science-Kompetenzen in einer Volluniversität umzusetzen. Die besondere Herausforderung sind dabei die stark dezentrale Organisation und gleichzeitig heterogenen Vorerfahrungen. Durch die zentrale Positionierung des HCDS und dem Anerkennen und Ausbauen der dezentralen Kompetenzen, kann eine Digitalisierung der Forschung zügig gesichert bzw. initiiert werden. Ziel ist damit perspektivisch ein sich erweiterndes Portfolio an Methoden und Tools bereitzustellen, die flexibel in verschiedene Anwendungskontexte übertragen werden können. Um dies zu sichern ist ein weiterer Ausbau der Schnittstellen zwischen der zentralen Einheit und dezentralen Einheiten zu Themen der Digitalisierung, aber auch dezentralen Forschungseinheiten angestrebt. Zudem sollte perspektivisch ein Angebot entwickelt werden, dass auch Master- bzw. Promotionsniveau Lehrangebote schafft, die systematisch eine data- und digital-literacy unterstützen.

Referenzen






[AIRe13] Alden, K.; Read, M.: Scientific software needs quality control. Nature 448/2013, S. 448–

448, 2013.

- [Co21] Cohen, J. et al.: The Four Pillars of Research Software Engineering. *IEEE Software* 1/2021, S. 97–105, 2021.
- [Ch22] Chue Hong, N. et al.: FAIR Principles for Research Software (FAIR4RS Principles). <https://doi.org/10.15497/RDA00068>, Stand Version 1 24.05.2022, 2022.
- [Fi23] Fischer, T. et al.: D-WISE - Digitale Wissenssoziologische Diskursanalyse. In (Busch, A.; Trilcke, P. Hrsg.): *DHd 2023 Open Humanities Open Culture. 9. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum"*, Trier, S. 388-389, 2023.
- [Ha20] Hasselbring, W. et al.: From FAIR research data toward FAIR and open research software. *it - Information Technology* 1/2020, , S. 39–47, 2020.
- [Ho23] Horsfall, D. et al.: Research software engineering accelerates the translation of biomedical research for health. *Nature Medicine*, 6/2023, S. 1313–1316, 2023.
- [HePa15] Hey, Tony ; Payne, Mike C.: Open science decoded. *Nature Physics* 5/2015, S. 367–369, 2015.
- [He18] Hettrick, Simon: Software in Research Survey. <https://doi.org/10.5281/zenodo.1183562>, Stand Version 1, 23.02.2018.
- [Ka21] Katz, D. et al.: Recognizing the value of software: a software citation guide, *F1000Research* 9/2021, 2021.
- [KMY21] Knowles, R.; Mateen, B.; Yehudi, Y.: We need to talk about the lack of investment in digital research infrastructure. *Nature Computational Science* 3/2021, S. 169–171, 2021.
- [Pi22] Pierce, M. et al.: The science gateway community institute’s consulting services program: Lessons for research software engineering organizations. *arXiv preprint, arXiv:2209.03958*, 2022.
- [WJU07] Wuchty, Stefan ; Jones, Benjamin F. ; Uzzi, Brian: The Increasing Dominance of Teams in Production of Knowledge. *Science* 5827/2007, S. 1036–1039, 2007.

Von der Bedarfserhebung zum Beratungsangebot

Kooperative Durchführung einer hochschulübergreifenden Umfrage zum Forschungsdatenmanagement


Michael Panitz ¹, Claus Spiecker ², Ian Wolff ³, Tanja von Fransecky ⁴
und Carina Schiller ⁵


Abstract: Forschungsdatenmanagement (FDM) gewinnt aufgrund der Anforderung nach wissenschaftlicher Nachvollziehbarkeit und Nachnutzung von Daten an Hochschulen an Bedeutung. Die Bedarfserhebung in Brandenburg ist die erste, die sich an alle Forschenden sowohl an Fachhochschulen als auch an Universitäten eines ganzen Bundeslandes richtet. Erhoben werden Kenntnisstand und Bedarfe für Beratung, Schulung und technische Unterstützung im Umgang mit Daten als Basis für den lokalen und landesweiten Aufbau von Infrastruktur und Serviceangeboten. Der Fragenkatalog basiert auf einer bundesweiten Erhebung an Fachhochschulen zu FDM. Erste Zwischenergebnisse der Erhebung deuten darauf hin, dass die Kenntnisse zu FDM sehr gering und die Bedarfe zu Beratung, Schulungen und unterstützender technischer Infrastruktur hoch sind.


Keywords: Forschungsdatenmanagement, Landesinitiative, Brandenburg, Umfrage, Hochschule


1 Einleitung


Aktives Forschungsdatenmanagement (FDM) gewinnt zunehmend an Bedeutung im wissenschaftlichen Alltag. In Repositorien abrufbare Forschungsdaten erhöhen die Nachvollziehbarkeit und damit die Transparenz der Entstehungsprozesse von Forschungsergebnissen. Dieser Umgang mit Forschungsdaten wird als selbstverständlicher Teil der Sicherung guter wissenschaftlicher Praxis angesehen, etwa von der Deutschen Forschungsgemeinschaft („DFG-Kodex“ [De22]). Ohne ein systematisches FDM werden an Forschungseinrichtungen künftig kaum noch

¹ Technische Hochschule Wildau, Zentrum für Forschung und Transfer, Hochschulring 1, 15745 Wildau, panitz@th-wildau.de,  <https://orcid.org/0000-0001-7497-632X>

² Fachhochschule Potsdam, Zentrale Einrichtung Forschungs- und Transferservice, Kiepenheuerallee 5, 14469 Potsdam, claus.spiecker@fh-potsdam.de,  <https://orcid.org/0000-0001-9947-8810>

³ Hochschule für nachhaltige Entwicklung Eberswalde, InnoSupport, Schicklerstraße 5, 16225 Eberswalde, ian.wolff@hnee.de,  <https://orcid.org/0000-0002-0413-0035>

⁴ Europa-Universität Viadrina, Referat Forschung und Wissenschaftlicher Nachwuchs, Auditorium Maximum, Logenstr. 4, 15230 Frankfurt (Oder), fransecky@europa-uni.de,  <https://orcid.org/0009-0004-4729-4453>

⁵ Technische Hochschule Brandenburg, Hochschulbibliothek, Magdeburger Straße 50, 14770 Brandenburg an der Havel, carina.schiller@th-brandenburg.de,  <https://orcid.org/0009-0001-4440-1311>

Projektförderungen möglich sein.

An vielen Hochschulen ist FDM im Aufbau begriffen. Um Synergieeffekte zu erzeugen sowie gemeinsame Standards und Infrastrukturen zu schaffen, haben sich in Brandenburg acht Fachhochschulen und Universitäten in einem Verbundprojekt zusammengeschlossen.⁶ Die beteiligten Hochschulen decken dabei nicht nur verschiedene Hochschultypen ab, auch die lokale FDM-Kompetenz ist unterschiedlich ausgeprägt.

Diese Aspekte und die definierte Zielgruppe „aktiv Forschende“ verleihen der gemeinsam konzipierten Bedarfserhebung zum FDM einen herausragenden Status. Außerdem bilden unter den bisher vorliegenden Umfragen institutionsübergreifende Erhebungen eher die Ausnahme.⁷ Bereits abgeschlossene lokale Umfragen in Brandenburg wiederum konzentrierten sich auf die generellen Anforderungen der jeweiligen Hochschule.

Ausgehend von den aktuellen und vergleichbaren Ergebnissen ist eine flächendeckende Infrastruktur für das gesamte Bundesland geplant. Neben der Implementierung des FDM an der jeweiligen Hochschule werden zusätzlich hochschulübergreifende Dienste entwickelt und bereitgestellt. Folglich geht die Bedarfserhebung über eine reine Bestandsaufnahme hinaus. Das FDM-Forschungsprojekt wendet die FAIR-Prinzipien [WDA16] an und wird die Ergebnisse dokumentiert verfügbar machen.

2 Forschungsinteresse

In der Erhebung werden die Bedarfe der Forschenden an allen Hochschulen in Brandenburg im Umgang mit Forschungsdaten ermittelt. Es geht darum, Gemeinsamkeiten und Unterschiede zwischen den Hochschulen und Hochschulformen zu erkennen und passgenaue Angebote – landesweit und lokal – zu entwickeln. Folgende Fragen stehen daher im Fokus:

- Welchen Kenntnisstand besitzen die Forschenden der Hochschulen im FDM?
- Zu welchen Themen werden Fortbildungen und Unterstützungsangebote benötigt?
- Welche technischen Systeme sind für die Institutionalisierung relevant?
- Welche speziellen Bedarfe lassen sich für einzelne Fachdisziplinen ableiten?

3 Zielgruppe

Die Umfrage zur Bedarfserhebung richtet sich an alle Forschenden der beteiligten

⁶ Projektwebseite: <https://fdm-bb.de/>, Stand: 14.07.2023.

⁷ Umfragen zum Umgang mit Forschungsdaten – Übersicht vgl. [forschungsdaten.org](https://www.forschungsdaten.org/), https://www.forschungsdaten.org/index.php/Umfragen_zum_Umgang_mit_Forschungsdaten_an_wissenschaftlichen_Institutionen, Stand: 14.07.2023.

Hochschulen des Landes Brandenburg. Um die Bedarfe sowohl für die gesamte Forschungslandschaft des Bundeslands als auch für jede einzelne Hochschule zu ermitteln, müssen die Ergebnisse der Umfrage für jede Hochschule separat auswertbar sein. Fragen, die nur für bestimmte Hochschulen relevant sind, werden daher nicht gestellt. Der Zielgruppe wurden alle haupt- oder nebenberuflich in Voll- oder Teilzeit beschäftigten akademischen Mitarbeiter*innen und Professor*innen zugerechnet, was unmittelbar vor dem Start der Erhebung Ende März 2023 insgesamt 4.387 Personen (siehe Tab. 1) umfasst. Da es sich bei akademischen Mitarbeiter*innen nicht per se um Forschende handelt und nicht alle Professor*innen forschend tätig sind, ist die Personenanzahl der Zielgruppe der aktiv Forschenden geringer anzusetzen. Da eine genauere Bestimmung der Anzahl der Forschenden nur an einigen Hochschulen möglich und zugleich mit einer zusätzlichen, schwer einzuschätzenden Ungenauigkeit verbunden wäre, ist diese Bezugsgröße im Sinne der Vergleichbarkeit der Hochschulen untereinander für die Rücklaufquote maßgeblich [Pa23]. Angestrebt wird eine Rücklaufquote von insgesamt mindestens 10 % unter der Berücksichtigung der hohen Ausgangszahl.

Hochschule	Anzahl Professor*innen	Anzahl akad. Mitarbeiter*innen
BTU ⁸	179	675
EUV ⁹	103	468
FBKW ¹⁰	72	70
UP ¹¹	328	1.607
FHP ¹²	100	101
HNEE ¹³	61	212
THB ¹⁴	56	88
THW ¹⁵	76	191
Gesamt	975	3.412

Tab. 1: Zielgruppe der Bedarfserhebung

4 Methodisches Vorgehen

Eine projektinterne Arbeitsgruppe hat die Befragung des Projekts EVER_FDM¹⁶, die 2022-2023 an allen Fachhochschulen in Deutschland durchgeführt wurde [We23],

⁸ Brandenburgische Technische Universität Cottbus-Senftenberg.

⁹ Europa-Universität Viadrina Frankfurt (Oder).

¹⁰ Filmuniversität Babelsberg KONRAD WOLF.

¹¹ Universität Potsdam.

¹² Fachhochschule Potsdam.

¹³ Hochschule für nachhaltige Entwicklung Eberswalde.

¹⁴ Technische Hochschule Brandenburg.

¹⁵ Technische Hochschule Wildau.

¹⁶ “Entwicklung und Verbreitung von Forschungsdatenmanagement an Fachhochschulen und Hochschulen für Angewandte Wissenschaften”, siehe dazu: https://fzdw.de/projekte/ever_fdm/, Stand: 14.07.2023.

aufgrund der ähnlichen Ausrichtung als erste Orientierung genutzt. Da sich aber die Bedarfserhebung nicht nur an Fachhochschulen, sondern auch an Universitäten des Bundeslandes richtet und einige Aspekte nicht, wenig oder anders betrachtet wurden, fand eine intensive und kritische Prüfung und Überarbeitung statt. Hierbei wurden 17 Fragen ohne bzw. nur mit leichten Anpassungen übernommen, 5 inhaltlich angepasst, 11 komplett entfernt und 10 neu erstellt.

Nach der Überarbeitung wurde der Fragebogen in acht Kategorien mit den folgenden Inhalten unterteilt:

1. *Allgemeines I:* Abfrage zur Forschungsaktivität an einer Hochschule.
2. *Informationen zu den Daten:* Angaben zur Generierung der Forschungsdaten, sowie zum Umfang und zu Dateiformaten.
3. *Informationen zur Speicherung:* Planung von Forschungsvorhaben und laufenden Projekten mit Datenmanagementplänen (DMP). Datenmanagement während und nach einem Projekt.
4. *Nutzung und Publikation:* Nachnutzungsverhalten beim Umgang mit bereits veröffentlichten Datensätzen. Publikationsverhalten bei eigenen Datenpublikationen. Gründe und Anreize zum Publizieren von Daten.
5. *Organisation der Forschung:* Verantwortlichkeiten für das Datenmanagement im Projektkontext. Kooperationen mit Partnerinstitutionen.
6. *Schulungs- und Unterstützungsbedarf:* Kenntnisstand zu Themen des FDM wie Datendokumentation, Recht und Ethik, Datenpublikation und Langzeitarchivierung, Open-Science-Praktiken, technische Dienste und Infrastruktur.
7. *Vorgaben, rechtliche und ethische Aspekte:* Abfrage zu Richtlinien und Organisationen im FDM, sowie rechtlichen Grundlagen und dem Umgang mit ethischen Fragestellungen. Data Literacy in der eigenen Lehrtätigkeit. Beantragung von zusätzlichen Ressourcen für das Datenmanagement.
8. *Allgemein II:* Abfragen zu den Statusgruppen, Fachgebieten, Altersgruppen und der Affiliation.

Abb. 1 verdeutlicht den strukturellen Aufbau des Fragebogens und wie viele einzelne Antwortoptionen und gemessene Merkmale (Items) in jeder der Kategorien mit welchen Fragearten abgefragt werden. Nach der Auswertung der Ergebnisse werden in der Analysephase durch Berechnung verschiedener Variablen die Anforderungen und Bedarfe unterschiedlicher Fachdisziplinen an das FDM ermittelt. Woraufhin Maßnahmen in den Bereichen der technischen Bedarfe, Informationsmaterialien, Informations- und Weiterbildungsveranstaltungen, stakeholder-spezifischen Qualifizierungsprogrammen, Beratungsangebote und First Level Support im Projekt IN-FDM-BB besprochen und dann an den Hochschulen umgesetzt werden.

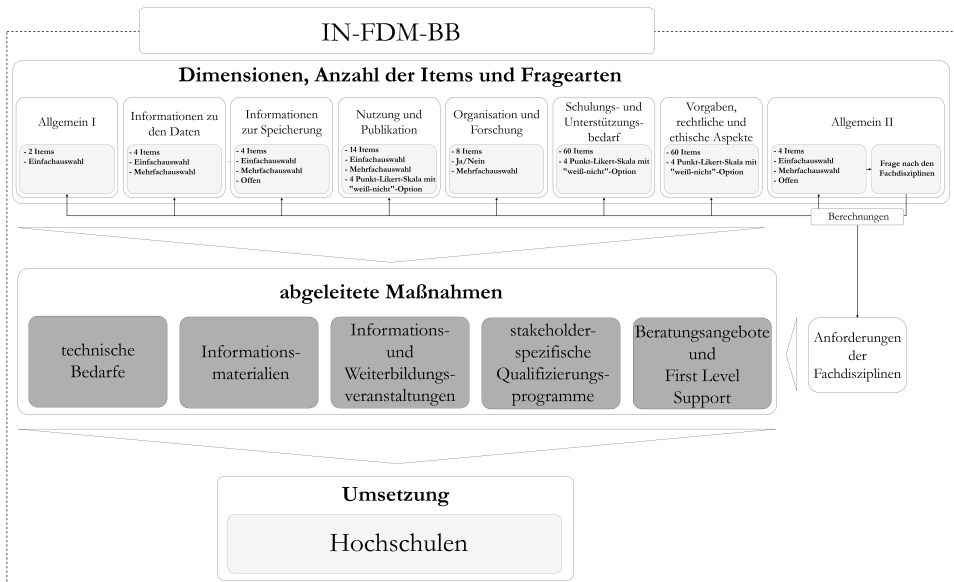


Abb. 1: Bedarfserhebung im Projekt IN-FDM-BB

Wie auch schon bei der 2017 an der FH Potsdam durchgeführten Bedarfserhebung [Ar18] kam auch bei der landesweiten Bedarfserhebung “Quamp” von Sociolutions¹⁷ als Umfrage-Tool zum Einsatz.

Im Zuge des Roll-outs wurden die Einladungen zur Teilnahme an der Umfrage über die internen Verteiler der Hochschulen verschickt. Es zeigte sich, dass die Verfügbarkeit, Pflege und Zugang zu E-Mail-Verteilern, die sich an die Zielgruppe der Forschenden richten, von Hochschule zu Hochschule deutlich variieren. Insofern war die Identifizierung und gezielte Adressierung “forschender Wissenschaftler*innen” nicht überall problemlos möglich. Im Vorfeld wurde diskutiert, durch die Verlosung eines Gewinns unter allen Teilnehmenden den Anreiz zur Teilnahme zu erhöhen. Dies stellte sich aber in datenschutzrechtlicher Hinsicht als problematisch und nicht zufriedenstellend lösbar dar, so dass die Einrichtung einer Gewinnoption verworfen wurde.

Nach dem Versand der Einladungen zur Teilnahme an der Umfrage gab es an den einzelnen Hochschulen verschiedene Strategien, um die Anzahl der Teilnehmenden zu erhöhen. Dazu zählten beispielsweise später versandte Reminder-E-Mails mit aktualisierten Anschriften, die Einbindung der Bedarfserhebung in thematisch passende Newsletter sowie das Verteilen von Flyern mit QR-Code zum schnellen Aufrufen der Umfrage. Eine weitere Maßnahme war die direkte Ansprache von Multiplikator*innen – wie Vizepräsident*innen für Forschung und Lehre, Forschungsprofessor*innen und

¹⁷ Produktseite „Quamp Survey“ von Sociolutions, <https://www.sociolutions.de/article/solutions-surveys-feedback-questionnaires.html>, Stand: 14.07.2023.

Forschungsgruppenleiter*innen – sowie der Austausch mit Forschenden auf dem Campus. Die gezielte und persönliche Ansprache von Multiplikator*innen mit der Bitte um Weiterleitung der Umfrage und Motivation zur Teilnahme lieferte die besten Resultate in puncto Steigerung der Rücklaufquote. Hilfreich war hier der intensive Austausch von Erfahrungswerten innerhalb der Arbeitsgruppe, welche Maßnahmen direkt sichtbare Erfolge erzielen konnten.

5 Zwischenergebnisse

Die vorläufigen Ergebnisse beruhen auf der Rückmeldung von rund 100 Personen; so viele hatten sich an der Bedarfserhebung innerhalb der ersten Woche¹⁸ beteiligt. Auf dieser Grundlage lassen sich noch keine belastbaren Vergleiche zwischen den verschiedenen Hochschulstandorten treffen, es zeichnet sich aber insgesamt ein geringer Kenntnisstand zu FDM ab. Es wird daher in der zu einem späteren Zeitpunkt folgenden Gesamtauswertung darauf zu achten sein, ob der geringe Kenntnisstand – insbesondere bei der Beantwortung der Fragen zu den beiden Kategorien zu Datentypen und Speicherung – möglicherweise zu unrealistischen Einschätzungen bei der Beantwortung geführt hat.

Bei den bisherigen Ergebnissen ist in (2) *Informationen zu den Daten* eine Tendenz zur Verwendung von Datenarten zu erkennen, deren Beschreibung und Publikation unkompliziert einzuordnen sind (bspw. Textdokumente und Grafiken). Bei der Kategorie (3) *Speicherung von Informationen* zeigt sich, dass überwiegend leicht verfügbare zentrale oder mobile Speicherlösungen bevorzugt werden. Eine Berücksichtigung der FAIR-Prinzipien [WDA16] ist kaum erkennbar. Ähnlich sieht es auch im Bereich der (4) *Nutzung von Publikationen* aus. Aktuell werden Forschungsdaten nur selten veröffentlicht. Anreize wie etablierte Standards, Workflows zur Publikation, passende Speicherorte oder auch Anreize finanzieller Art könnten eine Veröffentlichung von Daten fördern. Bisher werden Konzepte zum Datenmanagement, wie das Festlegen von Verantwortlichkeiten, oder das Erstellen eines Datenmanagementplans (DMP) bei der (5) *Organisation der Forschung* nicht mitgedacht. Wohingegen Kollaborationen zwischen verschiedenen räumlich getrennten Forschungsstandorten als die Regel angesehen werden können und dementsprechende Lösungen zum Datensharing geschaffen werden sollten. Insgesamt lässt sich ein geringer (6) *Wissensstand zu Open-Science-Praktiken* und den angebotenen technischen Diensten zum FDM entlang des gesamten Forschungsdatenlebenszyklus erkennen. Es lassen sich hieraus Bedarfe im Bereich der Data Literacy ableiten. Entsprechende Fortbildungen können eng an den Bedarfen entlang konzipiert oder angepasst werden, bzw. eine Verankerung des FDM in die Curricula perspektivisch mitgedacht werden. Dienstleistungen, die zentral an einer Hochschule eingerichtet sind, werden als wichtig angesehen, wobei besonderer Wert auf persönliche Beratung gelegt wird. Hieraus kann der Bedarf nach einer Stärkung und Verstetigung des

¹⁸ Stand: Ende April 2023.

FDM an den Hochschulen abgeleitet werden. Auch sind die Initiativen zum FDM bisher relativ unbekannt und der Kenntnisstand zu (7) *rechtlichen und ethischen Rahmenbedingungen* eher niedrig, worauf mit angepassten Dienstleistungen reagiert werden wird.

6 Ausblick

Der Bedarf an einem aktiven FDM zeigt sich auch in den Ergebnissen der Erhebung unter den Forschenden der acht brandenburgischen Hochschulen. Die Forschenden konnten durch verschiedene Strategien über Mailinglisten, direkte Ansprache unmittelbar oder über Multiplikatoren erreicht werden. Beim Abschluss der Umfrage wurde die angestrebte Rücklaufquote von 10 % nicht nur erreicht, sondern mit rund 13 % deutlich übertroffen.

Bereits zu einem frühen Zeitpunkt der Bedarfserhebung war erkennbar, dass FDM eher als ein Randphänomen an den Hochschulen wahrgenommen wird und noch nicht als integraler Bestandteil einer neuen Forschungslandschaft. Basierend auf den ersten Antworten von rund 100 Personen, lassen sich bereits Erkenntnisse hinsichtlich der Bedarfe und Inhalte für Beratung, Qualifizierung und Weiterbildung, Informationsmaterialien sowie zur technischen Unterstützung ableiten. Die Ergebnisse fließen nach ausgiebiger Analyse aller Daten in die im Rahmen des Projektes IN-FDM-BB geplanten Maßnahmen [In23] ein. So werden beispielsweise an allen acht Hochschulen Beratungsstellen eingerichtet oder ausgebaut und Qualifizierungsmaßnahmen für Forschende und FDM-Verantwortliche entwickelt. Speziell für die Forschenden wird auch ein zielgruppenspezifischer Wissensspeicher mit einschlägigen, aktuellen und relevanten Materialien eingerichtet. Zur Unterstützung der Forschenden ist im Bereich Infrastruktur die Implementierung landesweiter Dienste zur Erstellung von Datenmanagementplänen und zur Publikation bzw. Archivierung von Forschungsdaten geplant. Alle Angebote werden über die Projekt- und die jeweiligen Hochschulwebseiten, Newsletter, ggf. auch über Printprodukte bekannt gemacht. Ziel ist es, FDM durch Institutionalisierung an allen brandenburgischen Hochschulen als Teil des Forschungsalltags zu etablieren.

Danksagung

Finanziert durch das Bundesministerium für Bildung und Forschung (BMBF) und die Europäische Union (NextGenerationEU), sowie das Ministerium für Wissenschaft, Forschung und Kultur (MWFK) des Landes Brandenburg.

Literatur

[Ar18] Arndt, O.; Glatz, L.; Hummel, B.; Porst, M.; Schabalowski, W.; Skubatz, S.: Umfrage

zum Forschungsdatenmanagement an der FH Potsdam : Projektbericht, Zenodo, DOI: <https://doi.org/10.5281/zenodo.1161792>, Stand: 14.07.2023.

- [De22] Deutsche Forschungsgemeinschaft: Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex, Version 1.1, DFG, Bonn, 2022, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf, Stand: 14.07.2023.
- [In23] IN-FDM-BB/FHP, & IN-FDM-BB/UP. (2023). IN-FDM-BB Förderantrag: Institutionalisiertes und nachhaltiges Forschungsdatenmanagement in Brandenburg. Zenodo. <https://doi.org/10.5281/zenodo.7737224>, Stand: 14.07.2023.
- [Pa23] Panitz, M.: W 1.2.1 Konzept der Bedarfserhebung. Werkstattbericht. Zenodo, 30. März 2023. <https://doi.org/10.5281/zenodo.7870896>, Stand: 14.07.2023.
- [We23] Werth, R.: Präsentation „Entwicklung und Verbreitung von FDM an Fachhochschulen“ (EVER_FDM), https://indico.desy.de/event/37011/contributions/132893/attachments/80170/104661/2023-02-14_EVER_FDM%40RDA-DE_v1.1.pdf, Stand: 14.07.2023.
- [WDA16] Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J. et. al.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific data 3, 2016.

Entwicklung einer Wissensbasis für Lehr- und Lernmaterialien zu FAIRem Forschungsdatenmanagement und Data Science im Projekt DALIA (Data Literacy Alliance)

Canan Hastik¹, Frank Lange², Jan-Michael Haugwitz³ und Peter Pelz⁴

Abstract: DALIA ist ein Projekt zur Entwicklung einer Wissensbasis für Data Literacy, die über einen semantischen Knowledge-Graph zugänglich ist. Die Plattform basiert auf etablierten Technologien des Semantic Web und wird aktuelle Methoden der künstlichen Intelligenz (KI) implementieren, um die effiziente Bereitstellung von Lehr- und Lernmaterialien in personalisierten Lernpfaden sicherzustellen. Suchergebnisse und Empfehlungen können generisch oder disziplinspezifisch sein und sich auf alle Kompetenz- und Erfahrungsstufen beziehen. Ferner ist die Integration von modernen und nutzungsfreundlichen Technologien, wie Frage-Antwort-Interaktion, vorgesehen. Knowledge-Graphen verbessern nicht nur die Interoperabilität und Sichtbarkeit der Ressourcen gemäß den FAIR-Prinzipien, sondern ermöglichen auch eine Vernetzung mit anderen fachspezifischen Modellen innerhalb der NFDI. Die Entwicklung erfolgt in enger Zusammenarbeit mit Expertengruppen und zukünftigen Nutzenden - Lernende und Lehrpersonen aller Disziplinen.

Keywords: Datenkompetenz, Wissensbasis, Open Educational Resources, Semantic Web, KI

1 Einleitung

In diesem Beitrag wird die DALIA Wissensbasis des vom BMBF bis 2025 geförderten Projektes „Knowledge-Graph der Data Literacy Alliance (Dalia) für FAIRE Datennutzung und -bereitstellung auf der Basis von Semantic-Web-Technologie“ [Da23] hinsichtlich nachhaltiger Vernetzung von Anbietenden- und Nutzenden-Communitys von Lehr- und Lernmaterialien vorgestellt. Sie dient der disziplinübergreifenden Förderung der Datenkompetenz (Data Literacy) mittels Nutzung etablierter Technologien des Semantic Web wie Resource Description Framework (RDF), Ontologien und fachspezifischen Vokabularen, sowie Techniken des maschinellen Lernens, wodurch Datenaustausch durch die Anreicherung mit Metadaten maschinenlesbar und

¹ Technische Universität Darmstadt, Institut für Fluidsystemtechnik, Karolinenplatz 5, 64289 Darmstadt, canan.hastik@fst.tu-darmstadt.de, <https://orcid.org/0000-0003-1729-4642>

² RWTH Aachen University, IT Center, Seffenter Weg 23, 52074 Aachen & Institut für Anorganische Chemie, Landoltweg 1A, 52072 Aachen, f.lange@itc.rwth-aachen.de, <https://orcid.org/0000-0002-9346-6031>

³ RWTH Aachen University, IT Center, Seffenter Weg 23, 52074 Aachen, haugwitz@itc.rwth-aachen.de, <https://orcid.org/0009-0007-3576-3947>

⁴ Technische Universität Darmstadt, Institut für Fluidsystemtechnik, Karolinenplatz 5, 64289 Darmstadt, peter.pelz@fst.tu-darmstadt.de, <https://orcid.org/0000-0002-0195-627X>

kontextsensitiv gestaltet wird. Das Projekt arbeitet eng mit den Communitys der NFDI-Sektion Training & Education zusammen und soll nach Abschluss an den NFDI-Verein übergeben werden.

2 Was ist DALIA?

Datenkompetenz von Anfang an. Mit diesem Paradigma betonen Prof. Dr. Peter Pelz und Prof. Dr. rer. nat. Sonja Herres-Pawlis die Relevanz der Datenkompetenzförderung und damit verbunden der nachhaltigen Nutzung und Bereitstellung von Forschungsdaten [Pe21]. Das von beiden initiierte Projekt hat zum Ziel, einen nennenswerten Beitrag bei der Umsetzung dieses Paradigmas durch die Entwicklung einer Plattform für Lehr- und Lernmaterialien als Wissensbasis zu leisten. Durch die enge Verbindung zum NFDI e.V. und den Sektionen werden starke Synergien bei der Entwicklung von DALIA ermöglicht. Das Projekt leistet einen wichtigen Beitrag zur Steigerung der Datenkompetenz und der Etablierung einer Datenkultur in Forschung und Lehre. Darüber hinaus fördert DALIA den Kulturwandel, indem von den neuen Technologien und Trends Gebrauch gemacht wird. Dadurch werden Microservices für unterschiedliche Community-Bedarfe adressierbar und möglich.

Die verfügbaren Inhalte und Plattformen für Bildungsmaterialien zur FAIRer Datennutzung und Bereitstellung sind durch große Heterogenität gekennzeichnet. Materialien sind oft generisch und es mangelt an inhaltlich-fachlich-didaktischen Qualitätskriterien sowie formalen Kriterien, beispielsweise hinsichtlich der Nachnutzbarkeit dieser. Zudem sind verwendete Metadaten in Beschreibungsumfang und -tiefe, aber auch Qualität nicht ausreichend, um für Suchende für deren individuelle und ebenso heterogenen Bedarfe passende Materialien zu finden. DALIA soll als Single Point of Entry den einheitlichen Zugang zu einer Vielzahl von Repositorien ermöglichen und durch Kurationsprozesse zu einer Anreicherung von Metadaten und damit zu Erfolgen auch bei hoch spezifischen Suchanfragen führen. In DALIA kuratierte Metadaten können im Sinne eines Metadaten-Ökosystems wieder in die Repositorien zurückfließen und die Qualität für lokale Suchen verbessern. DALIA wird als Plattform konzipiert, die Lehr- und Lernmaterialien rund um den Themenkomplex Data Literacy in Form einer Knowledge-Base sichtbar, verfügbar und zugänglich macht. Diese wird, entsprechend zu aktuellen Entwicklungen in der Nationalen Forschungsdateninfrastruktur (NFDI) und dem FAIR-Konzept [Wi16] passend, als semantischer verknüpfter Knowledge-Graph umgesetzt. Zielsetzung ist es, bestehende edukative Materialien der NFDI-Konsortien, Datenkompetenzzentren, der FDM-Landesinitiativen und internationaler Stakeholder wie EOSC, OpenAIRE und RDA zu vernetzen, besser sichtbar und zugänglich zu machen. Die Verweise und Vorschläge der Knowledge-Base sind fachspezifisch und abhängig von Interesse und Bedarf der Lehrenden und Lernenden am relevanten Thema. Dabei kommen verschiedene Methoden der künstlichen Intelligenz, wie Semantic Reasoning, Machine Learning und Natural Language Processing (NLP) zum Einsatz. Diese leiten die

Interaktion der Nutzenden an und verbessern Suchergebnisse durch Analyse von Kriterien wie fachspezifische Data-Literacy-Kompetenzprofile [Le20; Pe23], Disziplin- und Zielgruppenspezifik.

2.1 Fragen wir ChatGPT...

Durch Klartextfragen kann aus der Menge von vorgegebenen Daten durchaus Neues entstehen. Fragen wir ChatGPT [Op23] (Wissensstand 2021), „Was ist Dalia?“, bekommen wir derzeit folgende Antworten:

- „DALIA steht für "Data Analysis and Logical Inference Assistant". Es handelt sich dabei um eine computergestützte Analyse- und Inferenzplattform, die speziell für den Umgang mit großen Datenmengen entwickelt wurde. DALIA kombiniert verschiedene Techniken der Datenanalyse, des maschinellen Lernens und der logischen Inferenz, um Muster, Zusammenhänge und Erkenntnisse in den Daten zu identifizieren.“

und

- „DALIA basiert auf fortgeschrittenen Algorithmen und Modellen des maschinellen Lernens, einschließlich Deep Learning und künstlicher Intelligenz. Die Plattform ermöglicht es den Benutzern, komplexe Abfragen und Analysen durchzuführen, um tiefere Einblicke in ihre Daten zu gewinnen. Sie bietet auch Funktionen zur Visualisierung der Ergebnisse und zur interaktiven Exploration der Daten.“

ChatGPT und ähnliche Angebote liefern auf Trainingsdaten basierende logisch konstruierte Aussagen. Die Fragestellung bedingt dabei die Musterantworten, die aus statistischen Modellen und Algorithmen generiert werden. Die beiden oben zitierten spekulativen Aussagen zeigen eine Überbetonung von klangvollen Technologiebegriffen. Inhaltlich sind die generierten Antworten recht weit entfernt von DALIA, auch wenn die Antworten sehr überzeugend aussehen - zumindest, wenn Kenntnisse fehlen, Information und Daten kritisch zu hinterfragen und zu bewerten. DALIA ist keine spezialisierte Datenanalyseplattform für große Datenmengen, wie ChatGPT behauptet, stattdessen nutzt DALIA Ressourcen aus verschiedenen Quellen, um Nutzenden bei der individuell optimierten Suche passende Ergebnisse zu Data-Literacy-Themen, wie Forschungsdatenmanagement, Datenethik oder Datenpräsentation zu liefern.

3 Was hat DALIA vor

DALIA wird als Single Point of Entry, also als zentraler Einstiegspunkt konzipiert. Zielsetzung von DALIA ist es, bestehende oder in Entwicklung befindliche verteilte Plattformen und heterogene Materialsammlungen, wie das Training und Education Repositories der NFDI-Konsortien, z.B. das NFDI4Ing Education Repository [Ed23],

oder kuratierte Materialien auf Plattformen wie Zenodo [Mo15], über standardisierte Schnittstellen mit der Wissensbasis zu verbinden. Die Materialien aus verschiedenen Quellen umfassen Standards und Best Practices, Selbstlern- und Lehrmaterialien unterschiedlicher Medienformate, die durch Lehrende nachgenutzt werden können. Damit leistet DALIA einen wichtigen Beitrag zur Vernetzung und Zugänglichmachung heterogener Quellen für fachspezifische Lehr- und Lernmaterialien in den jeweiligen Community-eigenen Plattformen. Darüber hinaus sollen Mehrwerte für die Anbietenden von Lehr- und Lernmaterialien geschaffen werden, u.a. indem Inhalte verteilter Quellen durch hochwertige Metadaten und Vokabulare angereichert und Werkzeuge zur Qualitätsbewertung von Materialien entwickelt werden.

3.1 Quellen und Materialien

Als Quelle für Open Educational Resources (OER) dienen verschiedene Repositorien und vielfältige Quellen unterschiedlicher Anbieter, wie z. B. Audio- und Video-Portale von Hochschulen, welche grundlegende Qualitätsstandards erfüllen. Qualitätsmerkmale sind u. A. eine Mindestverfügbarkeit und die Möglichkeit der Nachnutzbarkeit. Außerdem wird der Diskurs zur Anwendung der FAIR-Prinzipien auf OER verfolgt und berücksichtigt. In DALIA zugänglichen Materialien umfassen vielfältige Formate, darunter sind Standards und Best Practices, Selbstlern- und Lehrmaterialien unterschiedlicher Medientypen (Text, Bild, Audio, Video) und vielfältigen Formaten wie u. A. Slides, Poster, Podcasts, Lehrfilme oder Tutorial-Videos.

3.2 Graphentechnologie und Wissensbasis

Für die Implementierung der Knowledge-Base wird auf etablierte Technologien des Semantic Web zurückgegriffen, insbesondere wird der Knowledge-Graph mittels RDF und Vokabularen/Ontologien modelliert. SPARQL dient als standardisierte Schnittstelle sowohl für das DALIA-Portal zur Kuratierung und zum Lernen, als auch für einen öffentlichen Endpunkt zur Datenabfrage des Knowledge-Graphen. Die Umsetzung als Knowledge-Graph dient nicht nur als Datengrundlage für Empfehlungs-Algorithmen der DALIA-Plattform, sondern bringt auch Vorteile für Ersteller und Kuratoren der Lehr- und Lernmaterialien. So werden Ressourcen im Knowledge-Graphen mit maschinenlesbaren Metadaten angereichert, was insbesondere deren Interoperabilität verbessert. Durch Zurverfügungstellung über den öffentlichen SPARQL-Endpunkt werden die Ressourcen Teil der Linked Open Data Cloud und erhalten somit eine bessere Sichtbarkeit und Nachnutzbarkeit im Sinne der FAIR-Prinzipien.

Um das edukative Material in den verteilten Repositorien zu beschreiben und auffindbar zu machen, wird ein ontologiebasiertes Informationsmodell (DALIA Core Model) mit einem Maximum an semantischer Expressivität und disziplinübergreifenden Interoperabilität entwickelt. Die Verwendung von RDF-basierten Technologien und des graphenbasierten Modells bieten die Repräsentation des Wissens über die Domäne und

der Beziehungen zwischen diesen, was wiederum den Such- und Empfehlungsdienst unterstützt.

3.3 Nutzerbedarfe stehen im Vordergrund

Als Community-zentrierter Ansatz legt DALIA Wert darauf, aktuelle Trends wie Chatbots zur Interaktion von Nutzenden bei der Entwicklung angemessen einzusetzen und den Angebotscharakter (Affordance) zu berücksichtigen. Mit dieser Technologie ist es möglich, Fragen und Aufgabenstellungen in Prosa zu formulieren aber auch Programmieren, Testen und Konvertieren von Quellcode in der Softwareentwicklung zu unterstützen, aus Prosa Quelltext zu generieren und diesen in eine Vielzahl von Programmiersprachen zu konvertieren. Die KI ist darüber hinaus beim Generieren von API-Aufrufen, Erzeugen von Datenbankabfragen, der Fehlersuche und dem Testen recht fortgeschritten.

Das DALIA Core Model wird sich sukzessive mit weiteren fach- und disziplinspezifischen Modellen innerhalb der NFDI vernetzen. Perspektivisch können NutzerInnen über eine Vorauswahl der zur Verfügung gestellten Lernangebote in Form von Kursen und Lernpfaden einen individuellen Zugang zur Plattform und somit DALIA zu einem maßgeschneiderten personalisierten Angebot konfigurieren. Darüber hinaus sind noch weitere Anwendungsmöglichkeiten vorgesehen, wie beispielsweise die Integration von Abfrage-Wizards oder rudimentäres Natural Language Question Answering, sodass Forschende, Lehrende und Lernende individuelle Anfragen gemäß ihrem jeweiligen Bedarf formulieren können, wie

- „Ich möchte meine Messdaten FAIR ablegen, bitte liefere mir eine kommentierte Best-Practice-Anleitung“
- „Ich möchte Daten anderer nutzen, bitte liefere mir Metriken und Werkzeuge zur Beurteilung der formalen und inhaltlichen Datenqualität“
- „Ich möchte meinen Lernenden einen Lernpfad in DALIA vorschlagen, den Lernfortschritt tracken und für einzelne Lerneinheiten Leistungsnachweise vergeben“
- „Ich möchte für meinen Beruf relevante Kompetenz entwickeln, bitte schlage mir einen Lernpfad vor“
- „Bitte schlage mir weitere Themen vor, die für mich relevant sind.“

4 Nächste Schritte

Nach der allgemeinen Erhebung und Analyse der bestehenden Angebote im Bereich Data Literacy wird der Ansatz verfolgt, zielgruppenorientiert die Nutzendenbedarfe zu identifizieren. Um eine bestmögliche Nutzerorientierung zu erreichen, werden in

aufeinanderfolgenden Fokusgruppen Workshops zusammen mit der NFDI-Sektion Training & Education kontinuierlich weitere Anforderungen und Nutzenperspektiven systematisch erhoben und zielgerichtet in die Entwicklung von DALIA integriert. Ein Vorteil dieser Herangehensweise ist, dass die Community aktiv teilnehmen und teilhaben kann und selbst Einfluss hat, das „Endprodukt“ DALIA mitzugestalten. Daher ist das Community-Building und -management ein wesentlicher Aspekt hinsichtlich des Monitorings von Zielgruppen, Bedarfen und Beständen. Außerdem werden aus den NFDI-Communities heraus generische und fachspezifische modulare skalierbare Vermittlungs- und Schulungskonzepte, aber auch Zertifikatskurse entwickelt, die in DALIA Anwendung finden sollen.

5 DALIA Kommunikationskanäle

Aktuelles über das DALIA Projekt, zu technologischen Entwicklungen und Community-Themen aber auch Terminankündigungen werden im DALIA Projekt Newsletter [Ne23], in dem NFDI-DALIA-Rocket Chat Kanal und in Social Media veröffentlicht und verbreitet.

6 Danksagung

Dieses Projekt mit dem Förderkennzeichen 16DWWQP07A wurde gefördert vom Bundesministerium für Bildung und Forschung (BMBF) und der Fördermaßnahme aus der Aufbau- und Resilienzfazilität der EU.

7 Bibliographie

- [Da23] DALIA: Knowledge-Base für „FAIR data usage and supply“ als Knowledge-Graph, https://www.fst.tu-darmstadt.de/forschung_fst/zusammenarbeit_in_der_forschung/dalia/dalia_ueberblick.de.jsp, Stand: 29.05.2023.
- [Ed23] Education Repository, NFDI4Ing, <https://git.rwth-aachen.de/nfdi4ing/education>, Stand: 29.05.2023.
- [Le20] Lemaire, M. et al. (2020), Das DIAMANT-Modell 2.0: Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-ServiceLandschaft, <https://doi.org/10.25353/ubtr-xxxx-f5d2-fffb>.
- [Mo15] Molloy, L. (2015), Research data management (RDM) open training materials, <https://zenodo.org/communities/dcc-rdm-training-materials/?page=1&size=20>, Stand: 29.05.2023.
- [Ne23] DALIA Newsletter, <https://dalia.pages.rwth-aachen.de/newsletter>.
- [Op23] OpenAI Introducing ChatGPT, <https://openai.com/blog/chatgpt>, Stand: 29.05.2023.

- [Pe21] Pelz, P. et al. (2021), Sektionskonzept Training & Education zur Einrichtung einer Sektion im Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V., <https://zenodo.org/record/5599770>.
- [Pe23] Petersen, B. et al. (2023), Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards, <https://doi.org/10.5281/zenodo.8010617>.
- [Wi16] Wilkinson, M.D. et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3, 160018, 2016.

Bildung - Wirtschaftsinformatik an
Hochschulen angewandter
Wissenschaften

Wirtschaftsinformatik an Hochschulen Angewandter Wissenschaften (Workshop)

Christian Czarnecki¹, Thomas Barton², Vera G. Meister³, Alexander Lübke⁴

Abstract: Der Workshop wird in Kooperation mit dem Arbeitskreis Wirtschaftsinformatik an Hochschulen für Angewandte Wissenschaften (AKWI) organisiert. Es werden die besten Beiträge der diesjährigen Jahrestagung des AKWI in einer erweiterten Form vorgestellt. Daneben findet ein Austausch zu aktuellen Themen in Lehre und Forschung statt.

Keywords: Wirtschaftsinformatik, Hochschulen Angewandter Wissenschaften, anwendungsorientierte Forschung.

Vorwort

Der Workshop Wirtschaftsinformatik (WI) an Hochschulen Angewandter Wissenschaften (HAW) wird in Kooperation mit dem Arbeitskreis WI an HAW (AKWI) organisiert. Der AKWI ist der Dachverband der Fachbereiche mit deutschsprachigen WI-Studiengängen und/oder -Studienschwerpunkten an HAW. Er versteht sich als fachkompetenter und hochschulpolitischer Ansprechpartner zu Themen des Studiums der WI an HAW sowie der WI als anwendungsbezogene Wissenschaft. Der AKWI ist als Arbeitskreis des Fachbereichstags Informatik (siehe www.fbti.de) und zugleich als Fachgruppe WI-AKWI der Gesellschaft für Informatik (GI) organisiert (siehe www.akwi.de).

Der AKWI organisiert jährlich eine Jahrestagung, die aus einem wissenschaftlichen und hochschulpolitischen Teil besteht. Die 36. Jahrestagung fand vom 11.-13.09.2023 in Wildau statt. Für den wissenschaftlichen Teil wurden insgesamt 17 Forschungsbeiträge eingereicht, aus denen in einem doppelblinden Begutachtungsverfahren die vorliegenden zwei Artikel für den Workshop auf der INFORMATIK 2023 empfohlen wurden. Diese Beiträge wurden bereits auf der AKWI-Jahrestagung präsentiert und werden im vorliegenden Workshop in überarbeiteter und erweiterter Form vorgestellt.

Die Beiträge zeigen die inhaltliche und methodische Breite der WI-Forschung an HAW. Es wird dargestellt, wie mit Serious Games das Informationssicherheitsbewusstsein in Un-

¹ FH Aachen, Fachbereich Elektrotechnik und Informationstechnik, Eupener Str. 70, 52066 Aachen, czarnecki@fh-aachen.de

² Hochschule Worms, Fachbereich Informatik, Erenburgerstraße 19, 67549 Worms, barton@hs-worms.de

³ Technische Hochschule Brandenburg, Fachbereich Wirtschaft, Magdeburger Straße 50, 14770 Brandenburg an der Havel, vera.meister@th-brandenburg.de

⁴ Technische Hochschule Wildau, Fachbereich Wirtschaft, Informatik, Recht, Hochschulring 1, 15745 Wildau, luebbe@th-wildau.de

ternehmen erhöht werden kann. Darüber hinaus wird die Nachhaltigkeit von Software diskutiert und mögliche Maßnahmen werden strukturiert. Neben diesen beiden wissenschaftlichen Beiträgen bietet der Workshop einen Austausch zu aktuellen Themen der Lehre und Forschung an.

Programmkomitee

Wolfgang Alm (TH Aschaffenburg)
Gunnar Auth (HS Meißen)
Thomas Barton (HS Worms)
Christian Czarnecki (FH Aachen)
Christian Drumm (FH Aachen)
Ingo Elsen (FH Aachen)
Heinz Faßbender (FH Aachen)
Dorothee Feldmüller (Hochschule Bochum)
Andreas Heberle (Hochschule Karlsruhe)
Frank Herrmann (OTH Regensburg)
Jürgen Karla (HS Niederrhein)
Norbert Ketterer (HS Fulda)
Ute Klotz (Lucerne University of Applied Sciences and Arts)
Claudia Lemke (HWR Berlin)
Olga Levina (TH Brandenburg)
Alexander Lübbe (TH Wildau)
Verena Majuntke (HTW Berlin)
Birte Malzahn (HTW Berlin)
Vera Meister (TH Brandenburg)
Christof Mosler (HFT Stuttgart)
Christian Müller (TH Wildau)
Kerstin Prechel (Duale Hochschule Schleswig-Holstein)
Jörg Puchan (HS München)
Stefanie Regier (HS Karlsruhe)
Margit Scholl (TH Wildau)
Thomas Specht (HS Mannheim)

Towards Sustainable Software

A Structured Map of Techniques and Best Practices for Organizations and Development Teams to Enhance Sustainability

Verena Majuntke¹ and Frederik Obermaier¹

Abstract: As the ecological impact of human activities becomes an increasing concern, the need for sustainable practices extends to the software industry. This paper presents a structured map of techniques and best practices for enhancing software sustainability, defined as software that minimizes environmental impact and maximizes resource efficiency while considering economic and social factors. The map is aimed at organizations and development teams ready to implement sustainable measures, offering a collection of well-known techniques and best practices. The insights were gathered through a literature review and further enriched by open interviews with four experts from software development, providing both theoretical and practical perspectives. By adopting and implementing these sustainable software development practices, organizations can contribute to the global effort to mitigate the environmental impact of technology and promote responsible resource usage.

Keywords: software, software development, sustainability, sustainability overview

1 Introduction

The growing concern of the ecological impact of human activities has led to a greater emphasis on sustainability in many industries [KPF15]. The increasing reliance on software applications in our daily lives, coupled with the rapid pace of technological innovation, has led to an increase in energy consumption and carbon emissions [La21]. However, the impact of software development, operation and usage on the environment is often overlooked [CFS12]. In recent years, organizations have recognized the need for action regarding sustainable software which is software that minimizes environmental impact and maximizes resource efficiency in its design, development, operation and use while considering economic and social factors. Although it might seem tempting to wait for comprehensive guides or standardized processes for achieving software sustainability, this approach may inherently impede progress. Instead, organizations must actively take pragmatic steps to enhance sustainability in their software. By adopting and implementing practices for sustainable software development, organizations can contribute meaningfully

¹ HTW Berlin, Fachbereich 4, Treskowallee 8, 10318 Berlin, verena.majuntke@htw-berlin.de, frederikobermaier@me.com

to the global effort to mitigate the environmental impact of technology and promote responsible resource usage.

The contribution of this paper is a structured map of existing techniques and best practices to enhance the sustainability of software. The target group of the map are organizations and development teams who are ready to implement respective measures. It is not designed as a step-by-step guide but rather a structured overview of well-known techniques and best practices along the software lifecycle phases design, development, operation and use. The entities in the map have mainly been identified through a literature review. The findings were then discussed in interviews with 4 experts from software development. The interviews also yielded current practical and organizational insights which were integrated into the map.

The paper is structured as follows: Section 2 discusses related work. A definition for sustainable software is given and the map of techniques and practices to enhance software sustainability is presented in Section 3. Section 4 comprises the summaries of the expert interviews followed by a summary and outlook on future work in Section 5.

2 Related Work

A general sustainability model has been defined by Penzenstadler and Femmer [PF13]. The model aims at a continuous improvement of sustainability of software products. It consists of three levels with sub-aspects, where the first level describes a goal with dimensions that potentially influence each other. The second level represents these dimensions with values, potentially supported by regulatory requirements. The third level approximates them with measurable indicators and respective activities. The model uses two alternating phases to improve the sustainability of software. The model provides a generalized framework to concretize sustainability in a specific context. However, specific techniques and practices are not part of the model. Similarly, Johann et al. [Jo12] introduce a reference model for green and sustainable software and identify criteria with sustainability relevant effects. The model incorporates guidelines and checklists to navigate developers and stakeholder to produce more sustainable software. However, in contrast to our work, the model is a theoretical model and guideline and checklists need to be filled with content. Another approach that is more general is the *GREENSOFT*-Model [Na11] that defines three areas, namely *development*, *distribution* and *end of life*. Divided into three subareas – first/second/third-order effects – the model provides examples for criteria and indicators but lacks practical content. Overviews also exist with the focus on energy efficiency. Georgiou et al. [GRS20] provide an overview of existing work in the area of energy-aware software development comprising approaches, tools and techniques. The overview focuses on energy-efficient programming and does not discuss the operation and use of software. Pinto et al. [PC17] discuss the energy consumption of software for different lifecycle phase. In fact, energy efficiency is a part of sustainability but does not consider topics like longevity of software, maintainability etc.

3 Map of techniques and practices for enhancing sustainability of software

In this section, we first define the term *sustainable software* in Section 3.1. Subsequently, the map of techniques and practices to enhance software sustainability is presented in Section 3.2.

3.1 Definition Sustainable Software

In literature, a variety of definitions for the term *sustainable software* exist. First of all, the term *sustainable software* can be interpreted in two different ways according to Penzenstadler et. al. [Pe14a]: (1) the software supports sustainability goals, or (2) the software code and the development process itself is sustainable. In this research, we focus on the latter, software whose code and the process to obtain the code itself are sustainable.

The United Nations (UN) Brundtland Report defines *sustainability* as the ability to “meet the needs of the present without compromising the ability of future generations to satisfy their own needs” [Un87]. According to the UN, sustainability addresses three dimensions, namely society, the economy, and the environment. Applied to software, Dick et. al. [Di13] define sustainable software as “software, whose impacts on economy, society, human beings, and environment that result from development, deployment, and usage of the software are minimal and/or which have a positive effect on sustainable development”. Moreover, Nauman et. al. [Na11] state that sustainability correlates to resources used and energy consumed during the production and use of the software. In [Ca20], the authors link software sustainability with the longevity of software only using resources which are actually needed. Calero et. al. [CMP21] define sustainable software as software which “is energy-efficient, minimizes the environmental impact of the processes it supports, and has a positive impact on social and/or economic sustainability. These impacts can occur direct (energy), indirect (mitigated by service) or as a rebound effect”. While existing definitions often include three dimensions, the authors also showed that most literature regarding sustainable software focuses on the environmental dimension, addressing how software development and its use affect energy consumption and the use of other resources. In the context of this research, we focus on the environmental dimension as well.

We define sustainable software as *software that minimizes environmental impact and maximizes resource efficiency in its design, development, operation and use (while considering economic and social factors)*. This software is also referred to as *Green Software* [CMP21].

3.2 Map of techniques and practices

We conducted a comprehensive literature review to examine prevalent techniques and practices for enhancing software sustainability. Additionally, we interviewed four experts,

whose insights are summarized in Section 4, to gather feedback on these methods and incorporate practical knowledge and organizational expertise. Our findings are organized in a map of techniques and best practices, tailored for organizations and software developers focused on bolstering software sustainability. This map assists in pinpointing areas where improvements can be made. While the methods identified are widely recognized and employed in software development, this study does not evaluate their relative effectiveness. Instead, it underscores the significance of implementing practical measures over a comprehensive strategy. The structured overview aids in promoting the adoption of pragmatic approaches to boost software sustainability. However, it serves as an overview, not a step-by-step guide.

<p>Design</p> <ul style="list-style-type: none"> • Design Patterns • Microservices Architectures • Minimized network load <ul style="list-style-type: none"> • Edge Computing • Library and Plugin imports • Caching 	<p>Development</p> <ul style="list-style-type: none"> • Resource efficient Programming <ul style="list-style-type: none"> • High (downward compatibility) • Maintainability • Selection of energy-efficient modes • Transparency • Choice of datastructures and algorithms • Open Source • Source Code Analysis • Refactoring • Choice of Programming Language • Programming Practices
<p>Operation</p> <ul style="list-style-type: none"> • Hosting <ul style="list-style-type: none"> • Tactics for Public Clouds • Demand Shaping 	<p>Use</p> <ul style="list-style-type: none"> • Offline use • Sustainable pre-settings • Visibility

Fig. 1: Map of techniques and practices for software sustainability

Figure shows the map of techniques and practices. It is organized into four dimensions: (1) design, (2) development, (3) operation and (4) use according to software lifecycle phases. Each dimension comprises measures for the respective phase. In the following, the techniques and practices for (1) through (4) are discussed. In addition, organizational measures which have mainly been extracted from the expert interviews are addressed. Subsequently, further solutions and approaches which could not be related to a lifecycle phase are presented.

Design - In the optimal case, sustainability should already be considered when software is designed as the energy consumption may vary significantly depending on software components and their interaction for instance.

- **Design Patterns** are reusable solutions to a commonly occurring problem in software design. Sahin et al. [Sa12] compared the energy consumption of applications running on an embedded system in two variations, with the use of design patterns and without it. Overall, 15 of the 23 design patterns of Gamma et al. [Ga95] were analyzed showing that the *Flyweight*, *Mediator*, and *Proxy* pattern could reduce energy consumption while especially the *Decorator* pattern increased energy consumption significantly. Similarly, Bunse et al. [BS13] analyzed design patterns in the context of Android applications proving that *Facade*, *Abstract Factory*, *Observer*, *Decorator*, *Prototype* and *Template Method* increased energy consumption in applications.
- **Microservices Architecture** refers to an architectural design pattern in which an application is built as a set of loosely-coupled, independently deployable, and scalable services. Compared to monolithic architectures, microservices offer easy interchangeability of services and great potential for dynamically scaling with the current demand in a cloud environment [SA22]. Combined with a cloud infrastructure, microservices allow for the automatic scaling of only those parts of an application experiencing increased load, resulting in optimal hardware utilization and granular measurement of emissions intensity. Additionally, the ability to easily shut down unused services can increase the ecological sustainability of the software. However, a software map using microservices involves higher network loads and a greater focus on effective deployment management which also needs to be considered.
- **Minimized network load** Web-based software applications often perform computing processes on the server infrastructure or in dedicated data centers, with data transmitted through interfaces like HTTP. To minimize resource consumption, edge computing can be used, where more computing operations are performed on the edge, near the resource being used [Be16]. Regular evaluation of imported libraries and plugins also reduces data transfer over the internet, with small to medium savings being of potential impact, as exemplified by a plugin which was used on over two million websites, saving a total of 708 tons of carbon per year [Da22]. Caching also reduces the number of transmitted data packets and required energy on both device and server sides, with additional benefits such as faster loading times and more robust behavior during spikes in traffic [Er11].

Development - Since software is evolving, the development of software in an ongoing task. The measures collected in this category can be implemented during development.

- **Resource efficient programming** Bitkom's "Resource-Efficient Programming" guide from 2021 [Bi21] provides a comprehensive collection of aspects to consider for sustainable software development. The guide draws its aspects from a discussion with the German Federal Environment Agency. Among the listed aspects, some key ones are *high (downward) compatibility*, *maintainability*, *selection of energy-efficient modes*, *default sustainable settings*, *low hardware stress*, and *transparency*. Also, choosing the "right" data structure has shown to improve energy efficiency.

The tool *GreenC5* for instance, predicts the optimal data structure among the Copenhagen Comprehensive Collection Classes for C#. The decision support system SEEDS [MPC14] dynamically evaluates Java collection types and proposes modifications.

- **Open source software** is software whose source code is publicly visible and can be partly modified. It is a widely used method for developing and acquiring software leveraging transparency, interoperability, and open knowledge exchange [Fu03]. These aspects have parallels with the previously mentioned concept of maintainability or longevity. It can be assumed that open source software has the potential to be a more environmentally sustainable software, as it is designed to be more open, maintainable, and durable than proprietary or closed software.
- **Source code analysis** is the process of examining the source code of a software application to identify potential security, performance, or quality issues. *GreenAdvisor* [AHS15] identifies energy centers based on system calls assisting developers to comprehend and evaluate the consequences of their choices through analysis and guidance. *PEEK* [Hö14] analyzes source code during development suggesting code modifications for energy efficiency.
- **Refactoring** source code is a fundamental way to optimize software's energy efficiency. Refactoring involves restructuring the code internally while maintaining the external functionality. Through successful refactoring, the energy efficiency of code can be improved [PSC15], [GJW13], [Se22]. For Android applications, Anwar et al. showed that eliminating code smells could reduce the device's energy consumption by up to over 10% [APS19]. The previously discussed framework SEEDS by Manotas et al. [MPC14] analyzes source code during development and suggests refactorings for energy savings.
- The **choice of a programming language** may have a natural consequence for the energy consumption of the developed code. Pereira et al. [Pe21] compare 27 programming languages with respect to energy efficiency. Their measurements show interesting findings such that faster languages often consume more energy and how memory usage influences energy efficiency.
- **Programming practices** refer to a set of guidelines, standards, and conventions that developers follow to write high-quality code. Li and Halfond [LH14] ran empirical evaluations in the context of Android application development. They concluded that bundling network packets up to a certain size and using programming practices – which they collected from the Androids developer forum – as accessing class fields, performing invocations led to a decreased energy consumption. Similarly, Tonini et al. [To13] found out that a proper use of a for-loop and getters and setter can improve energy efficiency. Singh et al. [SNM15] showed that choosing energy-efficient APIs with an optimal set of parameters when implementing file operations in Java can have an energy saving impact when the software is run on large scale data servers.

Operation - The operation of software can be very energy consuming. By intelligent hosting and/or the use of demand shifting, sustainability in software operation can be improved.

- **Hosting** Leading cloud providers like Amazon, Microsoft, and Google, offer scalable and flexible solutions for hosting applications making it a competitive option for ecological sustainability. These "hyperscale" providers enable energy-efficient operation compared to traditional on-premise hosting, as measured by the "power usage efficiency" (PUE) indicator. Google Cloud – for instance – reported a PUE of 1.10 in 2022 [Go22], while traditional hosting solutions often have a PUE of > 2.0 [Jo18]. Using the public cloud, Vos et al. [Vo22] identified a set of 18 reusable tactics to optimize energy efficiency. However, the use of a hyperscale provider has implications for other aspects, such as data privacy, cost structure, and IT support, which need to be considered in a business decision. Further factors to consider when selecting a host include improving hosting efficiency, providing emission monitoring dashboards, and offering green energy.
- **Demand shaping** Technical potentials for improving the ecological sustainability of software mainly exist in the field of emission efficiency. One aspect in the area of emission awareness is the so-called "Demand Shaping". The core of this approach is the inclusion of the currently available green electricity in the power grid. When the proportion of renewable energies is high, more computing operations are performed faster, as both the power consumption of the devices and the power consumption of the networks required for data transmission cause relatively low emissions at the time of implementation. The most prominent example of demand shaping is the automatic adjustment of video players in the streaming area. The adjustment of quality due to poor network connection of the end-user is widespread, but the adjustment of quality due to insufficient green electricity in the power network is not [Asa].

Use - Upon software release and operation, energy consumption hinges on end-users' actual usage. That is, while software must accommodate offline use, its offline utilization ultimately relies on the end-user. Likewise, sustainable presets do not ensure corresponding configurations by the user.

- **Offline use:** Allowing applications to be used offline reduces the energy and emissions consumed by the end device. Additional optimizations may further reduce energy consumption, such as Edge Computing or Computing-on-device [Mo18]. However, the potential benefits may not outweigh the infrastructure and architectural costs in certain contexts. Depending on the maturity and energy efficiency of the server infrastructure and expected computational intensity, server-side operation may be preferred.
- **Sustainable (pre)settings:** Sustainable default settings are effective in enhancing the ecological sustainability of software applications. Different development platforms offer approaches to reduce the resource intensity of applications. For example, Microsoft's EcoQoS [Mi] automatically reserves and consumes fewer

hardware resources, making the application more energy-efficient and emission-reducing. In [Si15], for example, Sivitos et al. discuss pre-settings for washing programs and while energy efficiency is often claimed based on the “eco-modes”, in practice such programs are often not accepted by the end user. Applications should also offer preconfigured settings that balance performance, user experience, and emission reduction. Music streaming applications can serve as a reference for configuring streaming quality.

- **Visibility:** Often, the energy consumption of software is not visible to end users. As a consequence, users cannot assess whether or not their use of software has a high or low energy consumption. In [CTG08] Chetty et al. – for instance – discuss the energy consumption in households. They conducted a qualitative study with 15 households and 33 participants showing that most participants are not aware of their energy consumption. They concluded that making real time energy consumption visible would have a positive impact on sustainability.

Organizational - In addition to focusing on the lifecycle phases of software, the software development process itself requires a respective approach to produce sustainable software. While the measures for the former were identified in literature, the organizational measures have mainly been drawn from the expert interviews reported in Section 4.

According to expert C, it is crucial to make the emissions of software visible. This includes emissions generated during design, development, operation and use. On one hand this comprises the ability of the software to know its emissions and on the other hand to make this information publicly available. For instance, The GREENSOFT-Model [Ke13] provides criteria that can be used to assess significant indicators.

In order to integrate sustainability in the long run, experts A, B and C found that sustainability considerations should be incorporated in decision making processes on team-level. For instance, non-functional requirements could be defined with sustainability in mind [Pe14b], [Di13], the product vision should align with sustainability goals and teams should regularly review and optimize their sustainability efforts. Especially, when it comes to agile development, agile methods can inherently reduce "waste" and improve efficiency. Optimizing measures could be defined during the artifact's conception phase, and ecological sustainability could be integrated as an acceptance criterion in existing evaluations, such as Scrum's Sprint Reviews minimizing waste through lean thinking, and producing more sustainable outcomes. In this context, Dick et al. [Di13] proposed extending Scrum based on the GREENSOFT model, which involves continuously inspecting and adapting software development towards environmental, social, and economic impacts. This approach includes Sustainability Review & Preview events that reflect on past activities and propose solutions to existing sustainability issues. A designated Sustainability Executive is responsible for identifying optimization and savings opportunities. Interim Sustainability Presentations and Final Sustainability Presentations are also included to demonstrate the impact of sustainability practices. In order to reinforce organizational changes and to consider sustainability throughout the

entire software lifecycle, experts A, C and D stated that company-wide commitment is necessary.

Other approaches and measures - In addition to the discussed measures, there are solutions available to optimize energy efficiency and minimize resource consumption. An extension for the Kubernetes [Ku] container management system allows tasks to be executed on nodes in regions with a high percentage of renewable energy in their power mix [ASW19] reducing emissions. Microsoft's previously mentioned EcoQoS for Windows enables developers to save up to 90% of CPU energy consumption by selecting modes. With over 1.4 billion devices running on Windows [Mi22], such APIs have a great potential. AWS provides a "Carbon Footprint Tool" that makes emitted emissions visible and measurable for its customers [Am22].

Moreover, interest groups have emerged for the development of ecologically sustainable software, such as the "GreenTech Alliance" which aims at building a network of companies that are 100% committed to fighting climate change [Gr]. In addition to such groups, there are also defined value foundations. The "Sustainable Web Manifesto" refers to six principles for sustainable development of web applications, some of which are clearly related to ecological sustainability [Su]. Other less abstract principles are the 8 "Principles of Green Software Engineering," which are based on two fundamental approaches: everyone has a role to play in the fight against climate change, and sustainability alone is sufficient to justify work based on these principles [Asb].

4 Expert Interviews

In this section, we summarize the 4 expert interviews which were conducted to explore the topic of sustainable software (Section 3.1 to Section 3.4). The aim of the interviews was to extract insights and discuss techniques and practices identified in the theoretical part. The interviews were part of a Master's thesis [Ob22] and were carried out in May and June 2022 via video conference. The interviews followed a systematic approach for knowledge maximization but the guide was adapted to the flow of each conversation. All experts have a broad experience in software development. Their names have been anonymized, referring to Expert A/B/C/D. A short introduction of each expert is given at the beginning of each summary.

Expert A is a Product Development Coach who currently works at a large retail company. The interviewee has extensive experience as a Product Owner and is involved in coaching and training in the area of agile product development through a group of coaches. Additionally, the interviewee has recently become more involved in organizational development. In the interview, Expert A noted that ecological sustainability is often viewed as a *mean to increase economic success*, rather than as an actual end goal. He expects the hype around the topic to decrease over time, making it vital to introduce additional motivators to incorporate sustainability into software development. Second, the expert distinguishes between *sustainable software and software for sustainability*, stating

that the latter is more important. He suggests that software with an ecologically sustainable purpose is more effective than trying to make software ecologically sustainable, as the effort is not justifiable. Third, Expert A believes that *agile development has the potential to be more emission-efficient* due to its focus on reducing waste. However, he cautions that ecological sustainability is not automatically considered in agile development, and the responsibility for integrating it lies with management which has to promote awareness throughout the organization. Fourth, the interviewee states that *top management should drive and embody ecological sustainability* in software development, allowing awareness to cascade organization-wide. Although this holistic approach extends beyond software development, it should be discussed. However, he notes that management often pursues sustainability for ulterior motives, like brand image or monetary savings. Finally, the expert agrees that ecological sustainability can be considered a *non-functional requirement* but emphasizes its limited connection with security as a requirement. Security is more tangible and has immediate, measurable (economic) consequences. Therefore, it's essential to include metrics and increased visibility in the framework to ensure sustainability is adequately addressed.

Expert B is a former IT Manager and Principal Software Engineer who previously worked at a large automotive company. Additionally, he is involved as a member of the advisory board for a non-profit organization that promotes a framework for managing the business of IT. According to Expert B, a significant factor in software emissions could be counteracted by *data minimization*. The expert cites a project in the automobile industry, where each car sends unprocessed usage data daily, amounting to 130 terabytes per car per day. This data volume would cause high network load and emissions. He suggests processing data "on the edge" to reduce transmission and sees potential in promoting offline usage for sustainable software. Second, the expert believes *politics should take responsibility* to incentivize sustainable behavior and penalize harmful actions in software development. In a specific context, the business side should push the technical side towards sustainable development integrating sustainability into the software lifecycle. He also mentions "ownership" in agile development, suggesting developers should be more responsible and involved in designing functionalities, countering current mechanisms where developers lose sight of the bigger picture. As a third aspect, the expert identifies that poor *software quality and architecture* often lead to non-sustainable software. Frequent functionality releases driven by business pressure contribute to weak IT architectures, inefficiency, and inadequate performance. He cites examples like unused microservices in cloud infrastructure, which may necessitate cleanup projects. The expert connects the lack of "ownership" to software quality issues and suggests that product decision-makers should adjust their goals to promote sustainable and valuable development in both software quality and architecture.

Expert C is a current CTO at a medium-sized consultancy, which provides technology adoption solutions driven by behavioral research. The interviewee has previously worked in various roles in software development, consulting, and (agile) product management. Expert C emphasizes the *importance of visibility* in resource consumption and impact for integrating ecological sustainability in software development. He compares it to making

sustainable choices at a supermarket, which is often challenging. In the software industry, specific metrics are needed to help teams and decision-makers comprehend the emission-related effects of their choices. Second, the expert believes *cost is the driving force for businesses* to adopt emission awareness. Financial incentives can promote resource-efficient software development. Utilizing hyperscale providers like AWS reduces the ecological footprint through efficient system utilization, but cost-saving is the primary motivation, not climate-friendliness. Third, the expert sees *potential in agile software development for integrating ecological sustainability*. Agile methods aim to reduce waste, which could lead to more sustainable software. However, agile practices are not a guarantee, as their implementation varies. In contrast, the waterfall model allows incorporating sustainability from the project's start, but offers little chance to add it later if overlooked initially making the right balance the key.

Expert D is a consultant and product owner at a large international IT consulting company. The interviewee has deep knowledge in Green IT and advises clients on software sustainability. The expert criticizes the *offsetting of CO2 emissions through the purchase of green energy certificates* or support for climate projects, citing a lack of standardization that ensures actual effectiveness. He refers to a Nature paper [101] showing that only 58% of the targeted savings were effective. While compensation is necessary, he believes that emission reduction should be the main focus. Throughout the interview, the *motivation of decision-makers* was discussed. Expert D mentioned that improving public perception through eco-friendly initiatives is a common driver. Another motivation is meeting environmental, social and governance (ESG) standards – criteria which are used to assess investments. Meeting the criteria can positively impact stock prices, confirming that economic factors often play a role in environmentally sustainable motivations. Third, the expert emphasizes the *importance of management's role* in integrating ecological sustainability into the software lifecycle. He calls for stricter regulations to make software efficiency improvements attractive to decision-makers. The expert compares this to the introduction of seatbelt laws, which became second nature over time. A similar goal should be pursued for ecological sustainability in software. Fourth, the expert sees the greatest potential for ecological sustainability in a mixture of measures which need to be taken. He suggests a pyramid-like approach: *quick wins include implementing emission awareness in software, optimizing backup creation with green energy usage, and leveraging "time shifting" (demand shifting)*. The next step should be optimizing resource utilization, particularly by migrating to cloud providers. Lastly, he considers code efficiency optimization, which is less impactful and harder to prioritize in operational businesses. Without stricter regulations, code efficiency may remain a secondary focus. As already discussed, the expert sees the primary responsibility for sustainability lying with management. However, long-term success requires *those implementing measures to build knowledge and actively integrate sustainability*. Agile methods can be advantageous as agile projects may make it easier to justify decisions. Expert D also sees potential in establishing a sustainability committee, similar to an architectural committee, where experts can pool knowledge and make it accessible to all involved.

The input of the 4 interviews was analyzed and merged with the findings from the literature review into the map of techniques and practices for software sustainability described in Section 3.2. In addition to discussions about the findings from the literature review, the interviews provided new practical and organizational insights which were also integrated into the map.

5 Summary and Future Work

In this paper, we addressed the problem of enhancing software sustainability. A definition of sustainable software was given and a map of existing techniques and best practices for enhancing software sustainability was presented as the primary contribution. The target audience for this map are organizations and development teams eager to implement sustainable measures. It is important to note that the map is not a step-by-step guide, but rather an overview of techniques and best practices structured across the software lifecycle phases, namely design, development, operation, and use. It provides practical starting points that organizations and development teams can pick to enhance software sustainability.

Currently, we are looking into assessing the efficiency of different measures to allow a comparison between them. In the future, we plan on developing a step-by-step guide on how to increase software sustainability during software development. Furthermore, the social and economic dimension for software sustainability have to be incorporated.

Remark: This paper has originally been published in [MO23]

Bibliography

- [AHS15] Aggarwal, K.; Hindle, A.; Stroulia, E.: GreenAdvisor: A tool for analyzing the impact of software evolution on energy consumption: 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME). IEEE, pp. 311–320, 2015.
- [Am22] Amazon Web Services: New Customer Carbon Footprint Tool. <https://aws.amazon.com/blogs/aws/new-customer-carbon-footprint-tool/>, accessed 03/01/2023.
- [APS19] Anwar, H.; Pfahl, D.; Srirama, S. N.: Evaluating the Impact of Code Smell Refactoring on the Energy Consumption of Android Applications: 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE, pp. 82–86, 2019.
- [Asa] Assim, H. Green Cloud Advocacy Lead at Microsoft: Demand Shaping. <https://principles.green/principles/demand-shaping/>, accessed 15/04/2023.

- [Asb] Assim, H. Green Cloud Advocacy Lead at Microsoft: Principles of Green Software Engineering. <https://principles.green/>, accessed 15/04/2023.
- [ASW19] Aled, J.; Schien, D.; Wolff, A.: A Low Carbon Kubernetes Scheduler. In (CEUR Workshop Proceedings Ed.): 6th International Conference on ICT for Sustainability. ICT4S. CEUR-WS, 2019.
- [Be16] Beck, M. T. et al.: Mobile Edge Computing. *Informatik-Spektrum* 2/39, pp. 108–114, 2016.
- [Bi21] Bitkom: Leitfaden Ressourceneffiziente Programmierung. <https://www.bitkom.org/Bitkom/Publikationen/Ressourceneffiziente-Programmierung>, accessed 15/09/2023.
- [BS13] Bunse, C.; Stiemer, S.: On the Energy Consumption of Design Patterns. *Softwaretechnik-Trends* 2/33, pp. 4–5, 2013.
- [Ca20] Calero, C. et al.: 5Ws of green and sustainable software. *Tsinghua Science and Technology* 3/25, pp. 401–414, 2020.
- [CFS12] Capra, E.; Francalanci, C.; Slaughter, S. A.: Is software “green”? Application development environments and energy efficiency in open source applications. *Information and Software Technology* 1/54, pp. 60–71, 2012.
- [CMP21] Calero, C.; Moraga, M. Á.; Piattini, M.: Introduction to Software Sustainability. In (Calero, C.; Moraga, M. Á.; Piattini, M. Eds.): *Software Sustainability*. Springer International Publishing, Cham, pp. 1–15, 2021.
- [CTG08] Chetty, M.; Tran, D.; Grinter, R. E.: Getting to green. In (Youn, H. Y. et al. Eds.): *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, New York, NY, USA, pp. 242–251, 2008.
- [Da22] Danny van Kooten: CO2 emissions on the web. <https://www.dannyvankooten.com/blog/2020/website-carbon-emissions/>, accessed 31/03/2023.
- [Di13] Dick, M. et al.: Green software engineering with agile methods: 2013 2nd International Workshop on Green and Sustainable Software (GREENS). *IEEE*, pp. 78–85, 2013.
- [Er11] Erman, J. et al.: To Cache or Not to Cache: The 3G Case. *IEEE Internet Computing* 2/15, pp. 27–34, 2011.
- [Fu03] Fuggetta, A.: Open source software—an evaluation. *Journal of Systems and Software* 1/66, pp. 77–90, 2003.
- [Ga95] Gamma, E. et al.: *Design Patterns: Elements of Reusable Object-oriented Software*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1995.
- [GJW13] Gottschalk, M.; Jelschen, J.; Winter, A.: Energy-Efficient Code by Refactoring. *Softwaretechnik-Trends* 2/33, pp. 23–24, 2013.
- [Go22] Google Rechenzentren: Effizienz - Datenzentren. <https://www.google.com/intl/de/about/datacenters/efficiency/>, accessed 12/03/2023.

- [Gr] Greentech Alliance: What we do. <https://www.greentech.earth/what-we-do>, accessed 01/03/23.
- [GRS20] Georgiou, S.; Rizou, S.; Spinellis, D.: Software Development Lifecycle for Energy Efficiency. *ACM Computing Surveys* 4/52, pp. 1–33, 2020.
- [Hö14] Hönig, T. e. a.: Proactive energy-aware programming with PEEK: Conference on Timely Results in Operating Systems (TRIOS'14), 2014.
- [Jo12] Johann, T. et al.: How to measure energy-efficiency of software: Metrics and measurement results: 2012 First International Workshop on Green and Sustainable Software (GREENS). *IEEE*, pp. 51–54, 2012.
- [Jo18] Jones, N.: How to stop data centres from gobbling up the world's electricity. *Nature* 7722/561, pp. 163–166, 2018.
- [Ke13] Kern, E., Dick, M., Naumann, S., Guldner, A., & Johann, T.: Green software and green software engineering—definitions, measurements, and quality aspects.: *Proceedings of the First International Conference on Information and Communication Technologies for Sustainability (ICT4S2013)*, 2013.
- [KPF15] Kordestani, A.; Peighambari, K.; Foster, T.: Emerging trends in sustainability research: a look back as we begin to look forward. *International Journal of Environment and Sustainable Development* 2/14, p. 154, 2015.
- [Ku]: Kubernetes. <https://www.kubernetes.io>, accessed 15/04/2023.
- [La21] Lancaster University: Emissions from computing and ICT could be worse than previously thought. www.sciencedaily.com/releases/2021/09/210910121715.htm, accessed 15/04/2023.
- [LH14] Li, D.; Halfond, W. G. J.: An investigation into energy-saving programming practices for Android smartphone app development. In (Müller, H. A. et al. Eds.): *Proceedings of the 3rd International Workshop on Green and Sustainable Software*. ACM, New York, NY, USA, pp. 46–53, 2014.
- [Mi] Microsoft: Introducing Microsoft EcoQos. <https://devblogs.microsoft.com/performance-diagnostics/introducing-ecoqos/>, accessed 15/09/2023.
- [Mi22] Microsoft: By the Numbers. <https://news.microsoft.com/bythenumbers/en/windowsdevices>, accessed 1/03/2023.
- [Mo18] Mocnej, J. et al.: Impact of Edge Computing Paradigm on Energy Consumption in IoT. *IFAC-PapersOnLine* 6/51, pp. 162–167, 2018.
- [MO23] Majuntke, V.; Obermaier, F.: Towards Sustainable Software: A Structured Map of Techniques and Best Practices for Organizations and Development Teams to Enhance Sustainability: *Angewandte Forschung in der Wirtschaftsinformatik: Tagungsband zur 36. Jahrestagung des Arbeitskreises Wirtschaftsinformatik*, 2-17, 2023.

- [MPC14] Manotas, I.; Pollock, L.; Clause, J.: SEEDS: a software engineer's energy-optimization decision support framework: Proceedings of the 36th International Conference on Software Engineering. ACM, New York, NY, USA, 2014.
- [Na11] Naumann, S. et al.: The GREENSOFT Model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 4/1, pp. 294–304, 2011.
- [Ob22] Obermaier, F.: *Integration von Nachhaltigkeit in die Konzeption und Entwicklung von Software-Projekten*. Masterarbeit, Berlin, 2022.
- [PC17] Pinto, G.; Castor, F.: Energy efficiency. *Communications of the ACM* 12/60, pp. 68–75, 2017.
- [Pe14a] Penzenstadler, B. et al.: Systematic mapping study on software engineering for sustainability (SE4S). In (Shepperd, M.; Hall, T.; Myrtveit, I. Eds.): *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. ACM, New York, NY, USA, pp. 1–14, 2014.
- [Pe14b] Penzenstadler, B. et al.: Safety, Security, Now Sustainability: The Nonfunctional Requirement for the 21st Century. *IEEE Software* 3/31, pp. 40–47, 2014.
- [Pe21] Pereira, R. et al.: Ranking programming languages by energy efficiency. *Science of Computer Programming* 205, p. 102609, 2021.
- [PF13] Penzenstadler, B.; Femmer, H.: A generic model for sustainability with process- and product-specific instances. In (Malakuti, S. et al. Eds.): *Proceedings of the 2013 workshop on Green in/by software engineering*. ACM, New York, NY, USA, pp. 3–8, 2013.
- [PSC15] Pinto, G.; Soares-Neto, F.; Castor, F.: Refactoring for Energy Efficiency: A Reflection on the State of the Art: 2015 IEEE/ACM 4th International Workshop on Green and Sustainable Software. IEEE, 2015.
- [Sa12] Sahin, C. et al.: Initial explorations on design pattern energy usage: 2012 First International Workshop on Green and Sustainable Software (GREENS). IEEE, pp. 55–61, 2012.
- [SA22] Shaikh, K. A.; Agaskar, S. S.: *Microservices: Architecting and Design Considerations*. In (Shaikh, K. A.; Agaskar, S. S. Eds.): *Azure Kubernetes Services with Microservices*. Apress, Berkeley, CA, pp. 25–60, 2022.
- [Se22] Sehgal, R. et al.: Green software: Refactoring approach. *Journal of King Saud University - Computer and Information Sciences* 7/34, pp. 4635–4643, 2022.
- [Si15] Sivos, S. et al.: Don't just press the button! Why appliance settings increasingly matter for efficiency delivery and rulemaking: *Proceedings of Eceee Summer Study 2015*.
- [SNM15] Singh, J.; Naik, K.; Mahinthan, V.: Impact of Developer Choices on Energy Consumption of Software on Servers. *Procedia Computer Science* 62, pp. 385–394, 2015.
- [Su] Sustainable Web Manifesto: Sustainable Web Manifesto. <https://www.sustainablewebmanifesto.com/>, accessed 22/03/2023.

- [To13] Tonini, A. R. et al.: Analysis and Evaluation of the Android Best Practices Impact on the Efficiency of Mobile Applications: 2013 III Brazilian Symposium on Computing Systems Engineering. IEEE, pp. 157–158, 2013.
- [Un87] United Nations World Commission on Environment and Development Ed.: Report of the World Commission on Environment and Development: our common future, 1987.
- [Vo22] Vos, S. et al.: Architectural Tactics to Optimize Software for Energy Efficiency in the Public Cloud: 2022 International Conference on ICT for Sustainability (ICT4S). IEEE, pp. 77–87, 2022.

Informationssicherheit in den Arbeitsalltag nachhaltig integrieren:

Informationssicherheitskultur verstehen, mit Serious Games sensibilisieren und das Informationssicherheitsbewusstsein der Mitarbeitenden erhöhen

Regina Schuktomow¹, Hubertus von Tippelskirch² und Margit Scholl³

Abstract: Um das Informationssicherheitsniveau in Unternehmen zu erhöhen, muss Informationssicherheit gelebt werden. Dabei ist das Verständnis einer mannigfaltigen Informationssicherheitskultur in kleinen und mittelständischen Unternehmen (KMU) unabdingbar. Regelmäßige Sensibilisierung der Mitarbeitenden durch verzahnte Methoden wie analoge und digitale Serious Games führen tendenziell zur Steigerung des Informationssicherheitsbewusstseins sowie zur nachhaltigen Einbettung der Informationssicherheit in den Arbeitsalltag. Es wird gezeigt, wie KMU ihr Informationssicherheitsniveau erhöhen und die Informationssicherheitskultur verbessern können. Resultierende Erkenntnisse aus dem Forschungsprojekt „Awareness Labor KMU (ALARM) Informationssicherheit“ führen zu praktischen Hilfestellungen für KMU. Die Sicherheitskultur wird näher beleuchtet und Modelle zum Verständnis dazu aufgezeigt. Formen der kreativ und praxisnah gamifizierten Instrumente, wie sie in dem vorliegenden Projekt entwickelt, in der Praxis erprobt und kostenfrei durch die Förderung zur Verfügung gestellt werden, können erfolgreich zur nachhaltigen Sensibilisierung von Mitarbeitenden und Führungskräften genutzt werden. Damit wird die Notwendigkeit von Informationssicherheitsmaßnahmen in Unternehmen anschaulich diskutierbar und eine entsprechende, gezielte Kompetenz- und Personalentwicklung motiviert.

Keywords: Informationssicherheit und Awareness, Informationssicherheitskultur, Serious Games, Sensibilisierung, KMU

1 Einleitung

Bereits der Bericht „Die Lage der IT-Sicherheit in Deutschland“ [BSI21] konstatiert einen erheblichen Bedarf an mehr Cybersicherheit in der deutschen Wirtschaft. So haben im Jahr 2021 z.B. die Meldungen von Sicherheitsvorfällen an das Bundesministerium für Sicherheit und Informationstechnik (BSI) mit 14,8 Millionen einen bisherigen Höchstwert erreicht. Die meisten Vorfälle werden der unzureichenden Sensibilisierung von Mitarbeitenden zugeschrieben [BSI21] [BMI21]. KMU sind von besonderer Bedeutung, denn sie tragen maßgeblich zum wirtschaftlichen Erfolg Deutschlands bei. Ohne ein ausreichendes

¹ Technische Hochschule Wildau (TH Wildau), Fachbereich Wirtschaft, Informatik, Recht (WIR), Hochschulring 1, 15745 Wildau, regina.schuktomow@th-wildau.de

² TH Wildau, Fachbereich WIR, Hochschulring 1, 15745 Wildau, hubertus.tippelskirch@th-wildau.de

³ TH Wildau, Fachbereich WIR, Hochschulring 1, 15745 Wildau, margit.scholl@th-wildau.de

und zeitgemäßes Informationssicherheitsniveau in KMU ist dieser Erfolg gefährdet. Viele Unternehmen würden gerne ihre Informationssicherheit verbessern, allerdings erscheint es ihnen als eine große Herausforderung [BSI22]. Es ist daher notwendig, die Informationssicherheitskultur zu verstehen, Hilfe zur Selbsthilfe sowie praxisnahe und einfach umsetzbare Konzepte zur fortlaufenden Erhöhung des Informationssicherheitsniveaus und der Informationssicherheitskultur in KMU bereitzustellen, die auch das Bewusstsein für Informationssicherheit aller Mitarbeitenden steigern (oft als Awareness = bewusste Wahrnehmung bezeichnet). Denn beispielsweise besagt der Data Breach Investigations Report von Verizon, dass 82 Prozent der Verstöße durch unbewusstes Verhalten und/oder aktive (Fehl-) Entscheidungen von Menschen ermöglicht wurden, die den Diebstahl von Anmeldeinformationen, Phishing, Missbrauch oder Fehler beinhalteten [Ve22]. Um Informationssicherheitsniveau zu erhöhen muss offenbar ein Grundverständnis an Informationssicherheitskultur in Unternehmen vorhanden sein und vor allem gelebt werden. Dabei ist es für KMU wichtig, Hilfsangebote niederschwellig zu gestalten, um eine praxisorientierte Umsetzung von technisch-organisatorischen Maßnahmen unter Berücksichtigung der ISO/IEC Normfamilie 2700x bzw. der BSI-Standards 200-1 bis 200-4 zu initiieren, wozu auch Sensibilisierungs- und Schulungsmaßnahmen gehören.

Dienstleistungen von Unternehmen, die sich auf aufwändig simulierte Softwareangriffe spezialisiert haben, um Mitarbeitende zum Thema Cyberbedrohungen zu schulen, können vielversprechend für KMU sein. Doch es wurde mehrmals nachgewiesen, dass der isolierte Einsatz von ausschließlich digitalen Tools und Software wie z. B. simulierte Phishing-Angriffe keine nachhaltige und langfristige Wirkung haben und vor allem nicht zur gewünschten und notwendigen Verhaltensänderung beitragen [EN18] [VSB20] [Sa22] [BSN16]. Mitarbeitende ermüden von rein digitalen Schulungskonzepten, bei denen es in den meisten Fällen darum geht, nicht auf den Fehler hereinzufallen [Sa22]. Vor diesem Hintergrund verfolgt das Projekt „Awareness Labor KMU (ALARM) Informationssicherheit“ das Ziel, KMU dabei zu unterstützen, ihre Sicherheitskultur – einhergehend mit einer Personal- und Kompetenzentwicklung der Mitarbeitenden – durch miteinander verzahnte Maßnahmen zu erhöhen. Im Projekt wird iterativ in drei Phasen agil und partizipatorisch ein innovatives Gesamtszenario für Informationssicherheit mit analogen und digitalen erlebnisorientierten Serious Games, „Vor-Ort-Simulationen“ und weiteren Überprüfungen wie Quiz und Awareness-Messungen entwickelt und erprobt. Abbildung 1 skizziert das Gesamtszenario und den Prozess des Forschungsprojektes.

Im dreijährigen Forschungsprojekt „ALARM Informationssicherheit“ wurden u.a. sieben analoge und sieben digitale Serious Games entwickelt. Die Inhalte der Serious Games beruhen auf vorab durchgeführten Umfragen und tiefenpsychologischen Interviews, die mit Pilotunternehmen stattfanden. Es ging dabei darum, den konkreten Bedarf in KMU festzustellen und die relevanten Informationssicherheitsthemen zu definieren. Jedes Unternehmen hat verschiedene Aufgabenbereiche, die von einer unterschiedlichen Anzahl der Mitarbeitenden durchgeführt werden. Jede der Tätigkeiten hat eigene Anforderungen und Regeln die beachtet werden müssen. Die Informationssicherheit an sich als Thema spielt hierbei eine große Rolle. Der Forrester-Umfrage von 2020, bei der 315 IT-Entscheidungsträger abgestimmt haben, ist zu entnehmen, dass die Cyberangriffe von Unternehmen

heute als Bedrohung anerkannt werden und 87 Prozent aller Entscheidungsträger besorgt über einen solchen externen Angriff sind [Fo20]. Fast 80 Prozent der befragten Unternehmen haben innerhalb des letzten Jahres einen Sicherheitsverstoß erlebt, die zum Verlust von Kundendaten, Verlust von sensiblen Unternehmensdaten und vor allem finanziellen Verlusten führten. Nach Angaben von Forrester Research kostet der durchschnittliche Datenverstoß bis zu 7 Millionen Dollar pro Vorfall, von der Reaktion und Benachrichtigung, Produktivitätsverluste, potenzielle Klagen, Bußgelder, Geldstrafen und anderen Verbindlichkeiten. Im Allgemeinen nehmen auch die Verstöße gegen die Datenschutzrichtlinien zu [Fo20]. Um das Risiko einer nicht korrekten Reaktion auf z.B. externe Angriffe zu minimieren, müssen Unternehmen vorausschauend handeln und Präventionsmaßnahmen treffen. Dafür langen technische Maßnahmen allein nicht aus. Vielmehr müssen die eigenen Mitarbeitenden nachhaltig für das Thema der Informationssicherheit sensibilisiert werden.

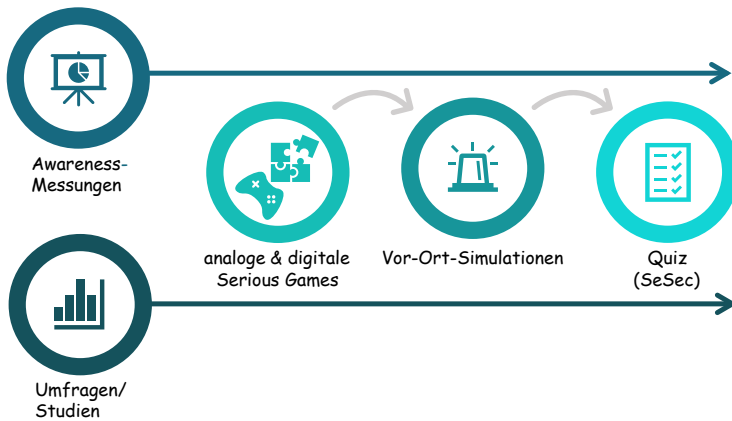


Abb. 1: Gesamtszenario und Prozess der Sensibilisierung

Das Projekt „ALARM Informationssicherheit“ befasste sich dazu mit Tätigkeitsprofilen [GA13] in Unternehmen. Aus den Tätigkeitsprofilen wurden Sicherheitsprofile abgeleitet und daraus Kompetenzprofile erstellt. Die Sicherheitsprofile stellen eine Bewertung dar, ob z.B. eine bestimmte Tätigkeit ein hohes Sicherheitsrisiko und hohen Angriffspunkt für das Unternehmen darstellt und in welcher Weise. So hat z.B. das Tätigkeitsprofil *Allgemeine Büroverwaltung* wie z.B. Sekretariat zur Verwaltung von Betriebsmittel ein hohes Sicherheitsrisiko wegen der zahlreichen verfügbaren internen Informationen und enge Arbeit mit Kunden, da es den zentralen Punkt für Informationsmanagement darstellt. Zum Sekretariat gehören neben der allgemeinen Büroverwaltung Administration und Organisation wie z.B. Terminkoordination oder Datenbankpflege. Unter Berücksichtigung des IT-Grundschutz-Kompendium nach BSI-Standard [BSI20] können konkrete Anforderungen für die Erweiterung der Kompetenzen der Mitarbeitenden im Bereich der Informationssicherheit gestellt (Kompetenzprofile) und z.B. konkrete Ziele für Schulungen und Weiterbildungsmöglichkeiten konzipiert werden.

Parallel wurden "Vor-Ort-Simulationen" („Angriffe“) entwickelt, die Teil der integrativen Sensibilisierungsstrategie sind, um den Reifegrad der Unternehmen zu überprüfen. Die Ergebnisse ermöglichen den Aufbau einer IT-Sicherheitsstrategie und dienen als Grundlage für niederschwellige Sicherheitskonzepte, um ebenfalls die Sensibilisierung der Mitarbeitenden zu erhöhen. Die Themen der "Vor-Ort-Angriffe" leiten sich aus den Tätigkeits- und Sicherheitsprofilen der Pilotunternehmen ab und sind in Tabelle 1 dargestellt. Aus den durchgeführten „Vor-Ort-Simulationen“ sind praktische Handlungsempfehlungen sowie niederschwellige Sicherheitskonzepte für Unternehmen entstanden.

1. Phishing	Beschaffung persönlicher Daten anderer Personen mit gefälschten E-Mails.
2. Passwort Breach Service	Mithilfe der geschäftlichen E-Mailadressen wird geprüft, ob persönliche Identitätsdaten bereits im Internet veröffentlicht wurden und missbraucht werden könnten.
3. Smishing	Unter Smishing versteht man das Phishing nach sensiblen Daten per SMS.
4. Tailgating	Physischer „Einbruch“ in das Unternehmen, um sensible Daten zu stehlen.
5. "BlackOut Szenario"/Incident Response Prozess	Simulierter Ransomware-Angriff mit dem Ziel den Incident Response Prozess in den Unternehmen zu aktivieren.
6. CEO Fraud	Betrugsmethode über E-Mail als Kommunikationsmittel, bei der sich Angreifende als Geschäftsführer, Manager oder Chef eines Unternehmens ausgeben.
7. Live-Hacking	Bildungs- und Informationsveranstaltung zur Entwicklung von persönlichem Risikobewusstsein und zur Sensibilisierung der Durchführung von Sicherheitschecks der IT-Infrastruktur auf organisatorischer Ebene.

Tab. 1: Durchgeführte „Vor-Ort-Angriffe“ in Pilotunternehmen

Ein weiterer Bestandteil des Gesamtszenarios ist der „Security Self Check (SeSec)“, ein Selbsttest als niederschwellige Sensibilisierungsmaßnahme, der dazu dient, Daten zum Wissenstand eines Probanden zu ermitteln. Er ermöglicht auch den Vergleich mit anderen Teilnehmenden des Selbsttests und erweitert oder aktualisiert den Wissenstand der Probanden durch eine sofortige Auswertung. Das zentrale Ziel bestand darin, Indikatoren zu ermitteln und basierend darauf, Lernpfade und Empfehlungen für gezielte Sensibilisierungsmaßnahmen zu entwickeln, die sich aus den erkannten Wissenslücken durch den Selbsttest ergeben.

Begleitend zur Entwicklung der Maßnahmen und Durchführung von Awareness-Trainings wurden Awareness-Messungen in Pilotunternehmen und mit Studierenden durchgeführt. Awareness-Messungen sind eine Methode, um das Bewusstsein und die Wahrnehmung von Mitarbeitenden in Bezug auf Themen der Informationssicherheit zu messen. Das Ziel

dieser Messungen besteht darin, den Kenntnisstand und das Verständnis der Mitarbeitenden in Bezug auf die relevanten Themen zu erfassen. Awareness-Messungen sind individuell gestaltbar und können verschiedene Formen annehmen, einschließlich Fragebögen, Online-Tests, Schulungen oder praktischen Tests wie Phishing-Simulationen. Sie können quantitative Daten liefern, indem sie beispielsweise den Prozentsatz der richtigen Antworten auf Fragen oder den Anteil der Mitarbeitenden, die eine bestimmte Richtlinie oder Sicherheitspraxis kennen, erfassen. Darüber hinaus können sie qualitative Erkenntnisse liefern, indem sie beispielsweise die Einstellungen, Wahrnehmungen und Haltungen der Mitarbeitenden durch offene Fragen oder Interviews erfassen. Aufgrund von geringer Datenmenge im Projektverlauf konnte nur eine positive Tendenz der Messungen festgestellt werden. Allerdings erwies sich das Messkonstrukt als erfolgreich und sollte weiterhin zu Forschungszwecken eingesetzt werden, um qualitativ und quantitativ fundierte Aussagen machen zu können.

2 Informationssicherheitskultur verstehen und fördern

Anhand der aktuellen Forschung im Rahmen des Forschungsprojektes „ALARM Informationssicherheit“ entsteht 2023 ein Report zu Informationssicherheitskulturen in KMU, in dem das Ziel verfolgt wird, nicht nur eine Bewertungshilfe zur Einschätzung Stärken und Schwächen der Akteure und ihrer Organisation zu entwickeln, sondern darüber hinaus das soziale und organisatorische Arbeitsumfeld für aktuelle Maßnahmen und den Nährboden für zukünftige Veränderungen zu erfassen. Wichtig ist dabei die Informationssicherheitskultur zu verstehen und zu fördern. Ein solches Hilfsmittel bildet die Grundlage für eine strategische Steigerung der Informationssicherheit in den untersuchten Unternehmen, welche auch von anderen KMU zur Selbsteinschätzung herangezogen und Teil eines bewussten Veränderungsmanagements werden kann. Eine kulturelle Betrachtungsweise scheint bei derart verschiedenen Aspekten, wie den Akteuren, dem Zwischenmenschlichen, dem persönlichen und allgemeinen Ethikverständnis, etablierten Prozessen und Organisationsebenen gefordert.

2.1 Informationssicherheitskultur – Teil der Unternehmenskultur

Ethos bedeutet Sitte, Gewohnheit oder eine Einstellung und die zugehörige Wissenschaft wird Ethik genannt und von den Zusammenhängen vom richtigen Handeln und Wollen angetrieben [SM16]. Im menschlichen Verhalten ist das „Ethos“ die Grundlage, wobei „Ethik“ daraus die Leitlinien begründet, die bestimmen, wie das Ethos gelebt wird. In der Moral definiert sich das Spezifische, artikuliert als konkrete Regeln [Sc20]. Arbeitsethos soll die unterbewussten, verinnerlichten und benennbaren, unterstützten Werte beschreiben. Arbeitsmoral ist die angewandte Form des Ethos, definiert richtige und falsche Verhaltensweisen [AN10], und wird im Folgenden als verhaltenssteuernde Kraft im „Machen“ interpretiert. Arbeitsethik mit konkreten ethischen Normen, Regeln und Kriterien für die angemessene Ausübung des Arbeitsethos und als systemischer Ort der Moral

[AN10] wird im Folgenden als „Sollen“ interpretiert. Geraten Moral im „Machen“ und Ethik im „Sollen“ aus dem Gleichgewicht, ergeben sich Sicherheitslücken. Damit ist bereits das Spannungsfeld einer Informationssicherheitskultur beschrieben. Bei der Informationssicherheitskultur handelt es sich um einen Teil der Unternehmenskultur. Ein theoretischer Ansatz zur Modellierung einer Informationssicherheitskultur in Unternehmen findet sich in dem Konzept zur Organisationskultur von Edgar Schein [Sc85]. Modellierung der Unternehmenskultur stellt dabei Scheins Kulturebenen-Model aus 1) beobachtbaren Artefakten, 2) Gefühlen und unterstützten und benannten Werten und 3) verinnerlichten Grundannahmen [Sc90]. Eine zusammenfassende Definition einer gesunden Informationssicherheitskultur beinhaltet demnach aus wissenschaftlicher und unternehmerischer Perspektive nach da Veiga et al. 2020 [Da20] folgende Bestandteile, die leicht modifiziert auch dem entwickelten Ansatz zugrunde liegen:

Die Informationssicherheitskultur steht im Zusammenhang mit dem Verhalten von Menschen in einem organisatorischen Kontext, wie in einem KMU, zum Schutz von Informationen. Innerhalb der Organisation werden Informationen unter Einhaltung von Informationssicherheitsrichtlinien und -verfahren verarbeitet. Es besteht ein Verständnis für die Umsetzung der Anforderungen in einer aufmerksamen Art und Weise, die durch regelmäßige Kommunikation, Bewusstseinsbildung, Schulung und Ausbildungsinitiativen eingebettet ist. Individuelle Verantwortlichkeiten und Rollen sind den Mitarbeitenden klar und bewusst. Das erlernte und routinierte Verhalten wird mit der Zeit Teil der etablierten Arbeitsweise. Dieses ist das Ergebnis der Annahmen, Werte und Überzeugungen der Mitarbeitenden, sowie ihres Wissens, ihrer Einstellung und ihrer Wahrnehmung bezüglich des Schutzes von Datenbeständen. Die Informationssicherheitskultur wird durch die Vision der Geschäftsleitung zusammen mit den unterstützenden Führungskräften in Übereinstimmung mit den Informationssicherheitsrichtlinien gesteuert. Interne und externe Faktoren beeinflussen diesen Steuerprozess. Ein angemessenes Umfeld aus Informations- und Kommunikationstechnik (IKT) unterstützen die Umsetzung, welche in den Artefakten der Organisation und dem Verhalten der Mitarbeitenden sichtbar wird. Dadurch wird ein vertrauensvolles Umfeld für alle Beteiligten geschaffen und Integrität etabliert. Ergänzende Definition nach da Veiga et al. 2020 [Da20].

2.2 Informationssicherheitskultur als Erkenntnisökonomie

Anhand des Ansatzes von van Niekerk und Solms [NS10] Informationssicherheitskultur durch eine Angebot-Nachfrage-Beziehung darzustellen, den Dienstleistungsphasen, und den Produktionsfaktoren nach Gutenberg [Gu83] und dem Konzept des Compliance Budgets von Beutement, Sasse und Wonham [BSW08] wurde die Informationssicherheitskultur als Erkenntnisökonomie (Information Security Culture as Economy of Cognition - ISCEC) modelliert (Ausprägung als Dienstleistungsprozess s. Abb. 2).

Zu van Niekerk und Solms [NS10] gibt es vier wesentliche Unterscheide in der Betrachtung. Die Erkenntnisvektoren „Womit“, „Was“, „Wie“ und „Warum“ beschreiben im neuen Modell in der Perspektive des ausführenden Individuums die Kräfte, welche die

Kultur zusammenhalten und verändern lassen. Tauscht man das „Warum“ durch ein „Wozu“ in der Perspektive der Entscheidungstragenden oder durch ein „Welche Folgen“ in der Nachhaltigkeitsperspektive werden alle Kulturkreise abgedeckt. Dazu einige konkrete Beispiele im Dienstleistungsprozess (vgl. Abb. 2).

Informationssicherheitskultur als Erkenntnisökonomie (ISCEC)

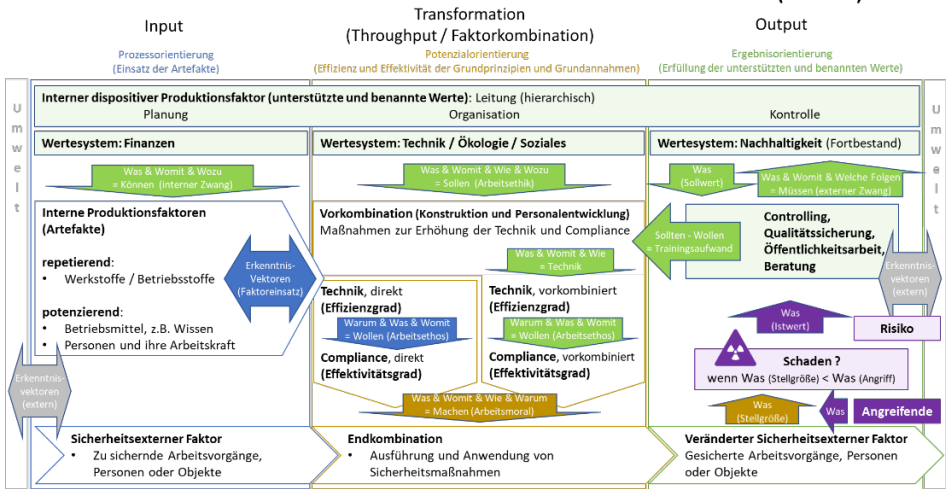


Abb. 2: Interne Dienstleistung Informationssicherheit – Phasendarstellung, Produktionsfaktoren und Informationssicherheitskultur (Erkenntnisvektoren) – vereinfachte Darstellung

Das „Was“ beschreibt den Output. Aus Sicht der Unternehmensleitung bildet das „Was“ die Ziele, daher benannte Werte als Inhalte. Werden diese in Richtlinien niedergeschrieben bilden sie als Artefakt ein Betriebsmittel, das neben Kapital und Arbeitszeit als Input eingesetzt werden kann und beschreibt „Womit“ ein Arbeitsvorgang gesichert wird. Zusammen bedingen sie mit der Motivation („Wozu“) der planenden Entscheidungstragenden den Faktoreinsatz und somit das Handlungsbudget, das „Dürfen“. Zur Umsetzung der Richtlinie muss das „Wie“ vermittelt und in einem Vorkombinationsprozess mit dem „Was“ und „Womit“ als (allgemeine) Technik das „Können“ trainiert werden, wodurch die Unternehmensleitung die Effizienz und das Informationssicherheitsangebot steuern kann.

Wird eine Maßnahme im Vorkombinationsprozess mit dem „Was“ und „Womit“ durch das „Warum“ nicht genügend begründet, verankert oder automatisiert, obwohl aufwendig trainiert wurde, wird die Maßnahme nicht ausgeführt und ist damit nicht effektiv. Ein geringes Compliance Budget bedingt eine geringe Informationssicherheitsnachfrage. Diese Handlungsgrenze beschreibt aus Perspektive des Individuums das „Wollen“ und dessen

Arbeitsethos. Die Effektivität wird daher maßgeblich über Sensibilisierung des „Warums“ gesteuert. Dort wo Technik und Compliance im Ist-Zustand ihr Gleichgewicht finden entsteht Informationssicherheit in Form des „Machens“ als Ausdruck der Arbeitsmoral.

Durch gute Organisation kann somit dieses in Einklang mit dem angestrebten Gleichgewicht, dem „Sollen“, und der verbundenen Arbeitsethik gebracht werden. Damit das Unternehmen nachhaltig bestehen kann ist eine Kontrollinstanz zwingend notwendig, die das Budgetziel „Müssten“ und das Handlungsgleichgewicht „Sollten“ definiert, indem sie alle Faktoren abgleicht, auch externe, zu denen im besonderen Maße die Angreifenden und das verbundene Schadensrisiko gehören (vgl. Abb. 2).

Während auf rein finanziell und leistungsorientierte Wertesysteme basierte Ansätze oft scheitern Informationssicherheitskultur positiv zu verändern [JG07], und die von der Wissenschaft vorgegebenen Ansprüche und technische Entwicklungen KMU überfordern [JG07], bietet die Erkenntnisökonomie einen Ansatz, der neben Rotstift und Zeigefinger die Stärke als familiär verbundene „Trutzburg“ und Schwäche für „Verstrickungen“ [Po21] der KMU berücksichtigt. Auch wenn Erkenntnisökonomie umfangreiche und komplexe Zusammenhänge umfasst, bietet sie für alle Beteiligten einen einfachen Einstieg und Umgang im Alltag, da man nur vier Fragen mitdenken muss: „Was“, „Womit“, „Wie“ und „Warum“. Dieses bildet ein intuitives Bewusstsein für die Kultur und die dahinter wirkenden Mechanismen. Informationssicherheitskultur lässt sich als Subkultur der Organisationskultur eines Unternehmens durch Edgar Scheins Kulturebenen-Model aus 1) beobachtbaren Artefakten, 2) unterstützten und benannten Werten und 3) verinnerlichten Grundannahmen erfassen. Die Kernaussagen der Informationssicherheitskultur als Erkenntnisökonomie, denen auch Unternehmen folgen sollten, können folgend zusammengefasst werden:

1. Informationssicherheitskultur bedarf eines umfassenden Ansatzes.
2. Drei Kulturebenen (Artefakte, unterstützte Werte und Grundannahmen bzw. Grundprinzipien) und deren vier Erkenntnisvektoren („Womit“, „Was“ und „Warum“ bzw. „Wie“) müssen bei allen Maßnahmen mitgedacht und ins richtige Verhältnis gesetzt werden, da die Menschen in der Informationssicherheitskultur größtenteils entscheidungsfreie Individuen sind.
3. Drei Rollen, Funktionen, Orientierungen sowie Bewertungssysteme korrespondieren zu den drei Kulturebenen und es müssen sowohl zugehörige Kennzahlen als auch Detektoren geschaffen werden.
4. Kombinationen von Erkenntnisvektoren beschreiben Steuervektoren, Expansionspfade (Effizienz durch Technik), Handlungsgrenzen (Effektivität durch Compliance) und deren Gleichgewichte (Informationssicherheitskulturniveaus).
5. Kommunikation spiegelt eine Kultur wider. Wissensvermittlung ist Kommunikation. Erkenntnisse können durch Wissensvermittlung beeinflusst werden. Wissen allein langt jedoch nicht. Die Kräfte die eine Kultur zusammenhalten und verändern sind gelebte Erkenntnisse. Der Niederschlag dessen findet sich in den drei Kulturebenen.
6. Wirtschaftlichkeit ist fester Bestandteil der Informationssicherheitskulturen fast aller Unternehmen. Die Forschung muss das anerkennen und bewusste Entscheidungen

gegen Informationssicherheit akzeptieren. Anpassbarkeit der Komplexität an das KMU, Nutzung von ökonomischen Logiken und Transfer in wirtschaftstheoretische Modelle ist die Konsequenz.

7. Leitung ist Führung durch Entscheidung und Organisation durch Kommunikation. Ein Verständnis der Verhältnismäßigkeit von Erkenntnisvektoren befähigt Entscheidungstragende im Unternehmen oder im Staat ein Unternehmen so zu versorgen, dass dessen Risiken besser verwaltet werden können. Zu wenig erkanntes und umgesetztes Wissen gefährdet somit den Selbsterhalt in der Zukunft und damit die Nachhaltigkeit des Unternehmens. Kontrollmechanismen sichern die Leitungsebene ab. Die Berücksichtigung der Informationssicherheitskultur gehört somit zu einem ganzheitlichen, ökonomischen Optimierungs- und Managementprozess.

2.3 Zusammenhang von Veränderungsmanagement, Informationssicherheitskultur und Serious Games

Als umfassender Ansatz erfasst die Erkenntnisökonomie ebenfalls Konzepte zum Veränderungsmanagement. Hierbei bietet sich im Blick auf Informationssicherheitskultur die Verwendung eines ADKAR-Change-Management-Modells (Awareness-Desire-Knowledge-Ability-Reinforcement for change) an, wie es da Veiga [Da18] in Anlehnung an den Change-Management-Dienstleister Prosci durch Kombination mit einem Umfragebogen zur Bestandsaufnahme (as-is assessment) und Erfolgskontrolle weiterentwickelt haben. Die Erfolgskontrolle gilt dabei als Vorbereitung der verstärkenden Anschlussmaßnahmen (reinforcing the change), zur Diagnose von Lücken, Widerständen und Durchführung von Korrekturmaßnahmen. Der Veränderungsprozess dazwischen besteht aus zwei Bestandteilen, nämlich der Veränderungsvorbereitung und des eigentlichen Veränderungsmanagements. Sowohl die vier genannten Veränderungsmanagement-Prozessschritte als auch die in den letzten beiden Prozessschritten integrierten ADKAR-Elemente lassen sich durch eine Kombination der Erkenntnisökonomie als Verständnishilfe und Serious Games als impactbezogenes Verständnisvehikel maßgeblich begleiten.

Eine umfangreiche, wissenschaftlich sauber konzipierte und erprobte Umfrage, wie da Veigas Information Security Culture Assessment, kann ein geeignetes Mittel für die Bestandsaufnahme und Erfolgskontrolle sein, muss aber sprachlich, kulturell und inhaltlich an das KMU angepasst, aufwendig begleitet, kommuniziert und ausgewertet werden. Dieses zieht Ressourcen vom Compliance Budget sowohl der Unternehmensleitung als auch der Mitarbeitenden. Adrian Field beschrieb gerade im Zuge der Pandemie und zunehmende Digitalisierung eine Überforderung durch Umfragen und Feedbackaufforderungen in der Gesellschaft, eine sogenannte Umfragemüdigkeit. Dieses gilt, da die Bereitschaft zur Beantwortung ein knappes öffentliches Gut sei [Fi20]. Im Umfeld der konkurrierenden Werte eines Unternehmens, in dem die eigentliche Tätigkeit in Konkurrenz mit der Informationssicherheit steht [KKÅ17] verschärft sich dieses Problem. Bestandsaufnahme und Erfolgskontrolle können im Blick auf eine Stärke von KMU, der oft familiären Struktur und Überschaubarkeit [Po21] daher alternativ zu Umfragen auch auf ihre empathischen

Detektoren zurückgreifen, geeignete Personen in der Rolle von Beobachtern und Beobachterinnen innerhalb ihres spezifischen Wertesystems. Gute Gelegenheiten bieten dafür gemeinschaftliche Aktivitäten bei denen das Thema Informationssicherheit zur Sprache kommt und kommuniziert wird, Informationssicherheitsvorfälle und Simulationen und insbesondere moderierte und beobachtete Serious Games. Denn diese leben zum einen in einer kontrollierbaren, moderierbaren Sandkastenumgebung zum anderen durch die Diskussion. Auf der ersten, vorbereitenden Vorkombinationsstufe kann mit Hilfe der Erkenntnisökonomie die Informationssicherheitskulturmanagementstrategie mitgeformt sowie ein internes Patenschaftsmodell entworfen, und das Veränderungsmanagementteam anhand von Serious Games sensibilisiert und als Moderatoren ausgebildet werden. Durch diesen Perspektivwechsel profitieren nicht nur Menschen in den Unternehmen, sondern auch Trainierende und Wissen schaffende Menschen. Auf der zweiten, Vorkombinationsstufe können Führungskräfte und Mitarbeitende anhand der vorgestellten Theorien und Serious Games durch die vier Veränderungsmanagementelemente geführt werden. Veränderungsbewusstsein (A-wareness for change) wird durch Aufklärung auf allen Unternehmensebenen und durch erfolgreiche Kommunikation mit Blick auf alle vier Kulturebenen in ihren vier Ausprägungen erzeugt. Veränderungswillen (D-esire for change) wird durch Notwendigkeiten geschaffen, die sich aus den drei benannten ökonomischen, sozial-ökologischen und nachhaltigkeitsorientierten Wertesystemen ergeben. Beim Wissen (K-nowledge) kommt wieder das Trainingsinstrumentarium durch die Serious Games zum Tragen, wie den Lernszenarien, aber auch Selbsttests und Sicherheitskonzepten, welche im Projekt „ALARM Informationssicherheit“ entwickelt und kostenlos bereitgestellt wurden. Letztendlich werden noch das „Können“ (A-bility) auf Basis erkenntnisökonomisch optimierter Produktionsfaktoren, wie beispielsweise genügend zur Verfügung gestellter Betriebsmittel oder Zeit, und anhand Serious Games unter anderem verbesserter Selbstwirksamkeitserwartungen erweitert und damit alle beteiligten Personen stärker befähigt. Nach der eingangs beschriebenen Erfolgskontrolle bezüglich der Werteziele wiederholt sich der Prozess dann durch angepasste Anschlussmaßnahmen (R-einforcement). Es zeigt sich deutlich, dass im Veränderungsmanagementprozess für Informationssicherheitskultur die Brücke zwischen den Konzepten der Erkenntnisökonomie und den Instrumenten, wie Serious Games, geschlagen wird. Bilden erstere die Wertesysteme und Detektoren, so letztere das Vehikel und die Plattform für eine erfolgreiche und teils für KMU spezifische pragmatische Anwendungsmöglichkeit, auf die im Weiteren näher eingegangen werden soll.

3 Einsatz der Serious Games in Unternehmen zur Steigerung der Informationssicherheit

Der gamifizierte Ansatz in jeweils sieben analogen und digitalen Serious Games soll wissensbasierte Schulungen ergänzen, um die Mitarbeitenden stärker zu involvieren, zu emotionalisieren und um schließlich dadurch ihr Bewusstsein für Informationssicherheit zu erhöhen. Game-based Learning ist eine Form des Lernens und Lehrens, das Wissen auf unterhaltsame und motivierende Art den Teilnehmenden zu vermitteln [LA09]. Durch den

spielerischen Ansatz in Kombination mit Narration (sinngewandte Erzählung) wird der Lerninhalt mit eigenen Erfahrungen verknüpft und regt dazu an, sich neues Wissen zu erschließen [FioJ].



Sicher zuhause wohnen & arbeiten
Dieses Serious Game gibt einen Überblick über die wichtigsten betrieblichen und privaten Informationssicherheits- und Datenschutzrisiken in der eigenen Wohnung bzw. im eigenen Haus sowie über zugehörige Präventionsmaßnahmen, um Risiken zu minimieren.



Cyber Pairs
Dieses Serious Game bricht mögliche Barrieren auf und führt zu mehr Sicherheit im Umgang mit Begriffen bzw. Beziehungen von gängigen bzw. neuartigen Cybercrime-Angriffen, indem es dabei unterstützt, diese auch im Detail zu verstehen und in Bezug auf mögliche Präventionsmaßnahmen unterscheiden zu können – stets verbunden mit der Fragestellung, was jede/r Einzelne von uns tun kann, um Risiken zu minimieren.



Multi-Faktor-Authentifizierung (MFA)
Dieses Serious Game vereint Aspekte von Passwortschutz und der Multi-Faktor-Authentifizierung (MFA) und demonstriert, dass der Schutz von Informationen in einem großen Maße von einer sicheren Authentifizierung abhängt. Es zeigt, wie ein „starkes“, weil sicheres, Passwort gebildet wird und dass ein (!) Faktor zum Schutz sehr sensibler Informationen nicht ausreichend ist.



Daten- und Informationsschutz
Der Schutz von Informationen und Daten von Kund/innen, Mitarbeitenden und anderen Parteien ist Teil des Geschäftes jedes Unternehmens. Dieses Serious Game unterstützt dabei, Daten- und Informationsschutz zu gewährleisten, indem der Umgang mit den wichtigsten Schutzstrategien rekapituliert und eingeübt wird.



Die fünf Phasen des CEO Frauds
Dieses Serious Game gibt einen Überblick über den Gesamtprozess von CEO Fraud und über Präventionsmaßnahmen – insbesondere auch für das oft übersehene „Vorspiel“ der Vorbereitungen.



Infoklassen-Roulette
Der Zweck von Informationsklassifizierung ist der Schutz von wertvollen Informationen jeder Organisation. Die „richtigen“ Klassen hängen von den potenziellen Auswirkungen auf Verfügbarkeit, Beschädigung oder Verlust von Informationen ab. Dieses Serious Game unterstützt beim Verständnis von Informationsklassifizierung und deren Notwendigkeit.



Mobile Kommunikation, Apps & Co.
Dieses Serious Game sensibilisiert in Bezug auf Risiken und Präventionsmaßnahmen, die die potenziellen Gefahren mobiler Kommunikation bzw. bei Nutzung von Apps verringern.

Abb. 3: Auflistung analoger Serious Games mit Lernziel

Die digitalen und analogen Serious Games enden mit Lessons learned bzw. mit Goldenen Regeln, die den Spielenden die wichtigsten Lerninhalte des Spiels nochmals ins Bewusstsein rufen. Um alle Mitarbeitenden „abzuholen“ und zu deren Kompetenzentwicklung beizutragen, werden im vorliegenden Projekt Awareness-Trainings eingesetzt, welche analoge und digitale Serious Games beinhalten. Diese behandeln auf die Bedarfe von KMU abgestimmte Informationssicherheits-Themen, z. B. Passwortsicherheit, CEO-Fraud, Homeoffice, Informationsklassifizierung, Social Engineering, Apps, Messenger-Dienste [Po21]. In Abbildungen 3 und 4 sind alle digitalen und analogen Serious Games mit deren Lernzielen dargestellt. Die analogen Serious Games sind mit Moderationsanleitungen ausgestattet, so dass eine für Informationssicherheit zuständige Person die Moderation übernehmen kann. Die analogen Serious Games werden als Lernstationen aufgebaut. Die damit verbundene Spieldauer ist skalierbar, so dass sie von 15 Minuten bis zu einer Stunde Lernmaterial bieten. Die digitalen Serious Games können als Vertiefung und Ergänzung

zu den analogen eingesetzt werden oder unabhängig von diesen. Die Mischung von analogen und digitalen Serious Games verfolgt den ganzheitlichen Ansatz und führt zur Abwechslung und wirkt auf Teilnehmende belebend und motivierend. Mitarbeitende können die Serious Games, deren Themen sie interessieren, unabhängig von Zeit und Ort spielen. Bei den digitalen Serious Games handelt es sich um immersive Geschichten, die Alltagssituationen aus dem Berufsleben in KMU darstellen [Pr22]. In Kombination mit einer auditiven und visuellen Ansprache ermöglichen sie das Eintauchen der Spielenden in die Situationen, so dass die Spielenden Bezüge zu ihrem Arbeitsalltag herstellen können [Yp14].

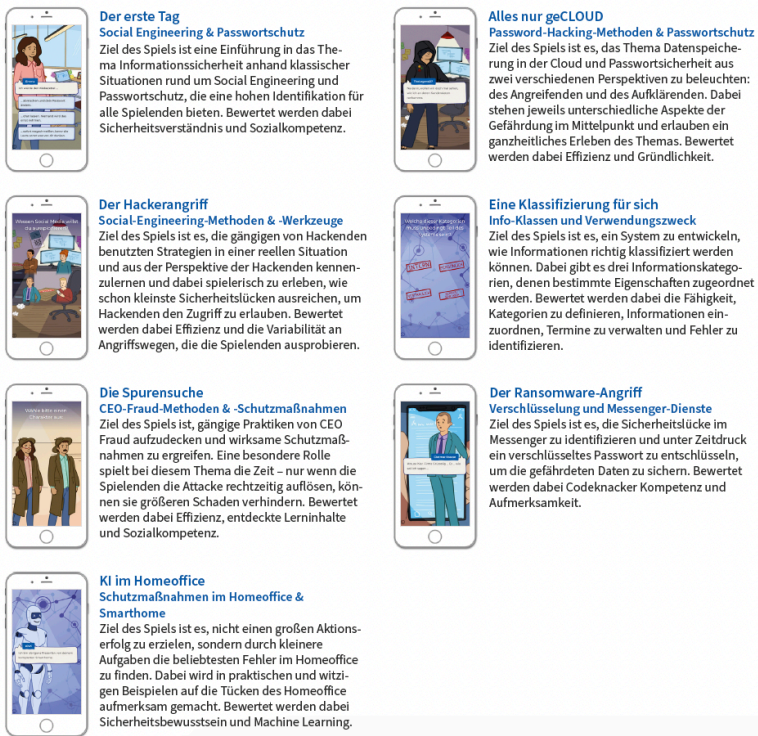


Abb. 4: Auflistung digitaler Serious Games mit Lernziel

Der Abbildung 5 ist ein möglicher Ablauf einer Sensibilisierungsmaßnahme zu entnehmen. Das Einhalten der einzelnen Schritte wird empfohlen, um möglichst ganzheitlich vorzugehen. Es empfiehlt sich, an erster Stelle eine oder mehrere Personen mit der Aufgabe der Sensibilisierung für Informationssicherheit im Unternehmen zu beauftragen. Es folgt eine Bestandsaufnahme des Informationssicherheitsniveaus. Dabei ist wichtig zu erfassen, welche Schwachstellen es in Unternehmen gibt, welche Bereiche bereits ausreichend geschützt sind und wo noch Bedarf an Schutz- und Sensibilisierungsmaßnahmen

besteht. Eine kurze Umfrage an alle Mitarbeitende zu deren Sicherheitsverhalten im Unternehmen sowie zur Kenntnis über Schutzmaßnahmen gegen Cyberbedrohungen wird Aufschluss über die interne Sicherheitslage im Unternehmen geben. Nachdem die Zielgruppe(n) definiert wurden sollten die Maßnahmen und der Einsatz genau geplant werden. Zur Auswahl der Maßnahmen gehören *Security Kompakt Infoblätter*, die ebenfalls aus dem Forschungsprojekt entstanden sind. Diese können als Impuls während des Awareness-Trainings eingesetzt werden. Serious Games als Bestandteil der Maßnahmen regen zum Erfahrungsaustausch an, der eine wichtige Rolle beim Erlernen neuer Inhalte darstellt. Um das ausgewählte Sensibilisierungsthema abzurunden, wird empfohlen das thematisch passende niederschwellige Sicherheitskonzept bzw. Handlungsanweisung mit den Teilnehmenden im Anschluss zu besprechen. Um den Erfolg von Sensibilisierungs- und Schulungsmaßnahmen zu erfassen, sollten Awareness-Messungen durchgeführt werden und Feedback zum Serious Game eingeholt werden. Dies ermöglicht Wünsche der Teilnehmenden zu berücksichtigen und sie intrinsisch zu motivieren und somit höheres Verständnis für Cyberrisiken zu erlangen. Hier können die im Projekt entwickelten Umfragen zu Informationssicherheitskultur und zur Messung des Bewusstseins für Informationssicherheit unterstützen. Es ist notwendig Sensibilisierungsmaßnahmen, wie den Einsatz von Serious Games, in regelmäßigen Abständen einzusetzen, um kontinuierliche Verbesserung des Informationssicherheitsniveaus zu erreichen.



Abb. 5: Ablauf einer Sensibilisierungsmaßnahme mit Einsatz von Serious Games

4 Mehrwert für Wirtschaft und Gesellschaft

Informationssicherheit muss gelebt werden. Dies kann nur funktionieren, wenn Menschen sich der Informationssicherheitsgefahren bewusst sind und präventiv aktiv dagegenhandeln und das eigene Verhalten im Arbeits- sowie Privatumfeld nachhaltig ändern. Eine vertrauenswürdige Informationssicherheits-, Fehler- und Kommunikationskultur im Unternehmen ist dabei essentiell. Mitarbeitenden sollen Ängste und Bedenken im Falle einer falschen Entscheidung genommen werden. Es ist wichtig, über Fehler und Erfolge miteinander zu sprechen, um das eigene und das Verhalten anderer zu reflektieren und gemeinsam über mögliche Lösungen nachzudenken. In diesem Zusammenhang darf Sensibilisierung für Informationssicherheit nicht als eine einzelne Maßnahme betrachtet werden. Alle Maßnahmen des vorgestellten innovativen und praxisnahen Gesamtszenarios können und sollten immer wieder eingesetzt werden, denn Informationssicherheit ist ein fortlaufender Prozess. Befragungen (Sicherheitskultur, Awareness-Messung) sollten wiederholt werden, analoge und digitale Serious Games können in geeigneten Situationen zur Erinnerung absolviert oder für neue Mitarbeitende als Einstieg genutzt werden. Sicherheitskonzepte, Handlungsempfehlungen, Goldene Regeln sollten präsent platziert, immer wieder diskutiert und ggf. aktualisiert werden. Damit Mitarbeitende sich leichter sicherheitskonform verhalten können, bedarf es der Unterstützung des Top-Managements des Unternehmens [Sa22]. Durch den Einsatz von vielfältigen kreativen Methoden und gebündelten Maßnahmen entsteht durch das vorliegende Projekt ein Instrument zur Selbsthilfe für KMU, welches die Mitarbeitenden befähigt, achtsamer im Umgang mit möglichen Risiken zu sein, und so zur Steigerung des Informationssicherheitsniveaus des gesamten Unternehmens beiträgt.

5 Ausblick

Um eine erfolgreiche und nachhaltige Sensibilisierung zu erreichen, muss diese interaktive und abwechslungsreiche Methoden beinhalten, Austausche im diskursiven Setting fördern, kurzweilig sein und in wiederkehrenden Abständen zu aktuellen Themen stattfinden. Das innovative und praxisnahe Gesamtszenario ist als ganzheitliches Konzept zu betrachten, das zu der dringend notwendigen Sensibilisierung von Führungskräften und Mitarbeitenden und zu einer gezielten Personalentwicklung in KMU beiträgt. Informationssicherheit wird durch die verschiedenen Maßnahmen des Gesamtszenarios im Zusammenhang mit den zunehmend digitalen Arbeitsprozessen konkret (be-)greifbar gemacht, gleichzeitig werden die Mitarbeitenden der Unternehmen emotional berührt und aktiv in die Maßnahmen einbezogen. Eine nachhaltige und unternehmensweite Informationssicherheitskultur soll damit aufgebaut und etabliert werden. Bei dem im vorliegenden Projekt entwickelten Gesamtszenario steht der Diskurs und der Austausch von Erfahrungen und Wissen im Vordergrund. Für eine deutschlandweite Nutzung und Erhöhung des Informationssicherheitsniveaus in Unternehmen, müssen Strukturen geschaffen werden, die

breitenwirksamen Einsatz von Moderatorenschulungen und Awareness-Trainings ermöglichen, um Multiplikatoren für Serious Games und das Gesamtszenario des Forschungsprojektes auszubilden und Mitarbeitende sensibilisieren zu können.

Danksagung

Wir danken dem Bundesministerium für Wirtschaft und Klimaschutz (BMWK) für die Förderung des Projekts Awareness Labor KMU (ALARM) Informationssicherheit. Wir möchten uns auch bei den beteiligten Pilotunternehmen, Unterauftragnehmer und assoziierten Partnern sowie bei der gesamten Forschungsgruppe für ihre hilfreichen kritischen Kommentare und Austausch bedanken.

Referenzhinweis: Dieser Beitrag stellt eine Erweiterung für die INFORMATIK 2023 dar, der im Rahmen der 36. AKWI Jahrestagung 2023 in Proceedings veröffentlicht wird: *Schuktomow, R., von Tippelskirch, H. & Scholl, M. (2023). Informationssicherheit in den Arbeitsalltag nachhaltig integrieren: Informationssicherheitskultur verstehen, mit Serious Games sensibilisieren und das Informationssicherheitsbewusstsein der Mitarbeitenden erhöhen. 36. AKWI Jahrestagung 2023.*

Literaturverzeichnis

- [AN10] Aßländer, M. S.; Nutzinger, H. G.: Der systematische Ort der Moral ist die Ethik! Einige kritische Anmerkungen zur ökonomischen Ethik Karl Homanns. *Zeitschrift für Wirtschafts- und Unternehmensethik* 3/11, S. 226–248, 2010.
- [BMI21] Bundesministerium des Innern, für Bau und Heimat (BMI): *Cybersicherheitsstrategie für Deutschland 2021*. Berlin, 2021.
- [BSI20] Bundesamt für Sicherheit in der Informationstechnik (BSI): *IT-Grundschutz-Kompendium*. Edited by Bonn Bundesamt für Sicherheit in der Informationstechnik. Reguvis Fachmedien, Bonn, 2020.
- [BSI21] Bundesamt für Sicherheit in der Informationstechnik (BSI): *Die Lage der IT-Sicherheit in Deutschland 2021*. Bonn, 2021.
- [BSI22] Bundesamt für Sicherheit in der Informationstechnik (BSI): *Cyber-Sicherheit für KMU. Die TOP 14 Fragen*. Berlin, 2022.
- [BSN16] Bada, M., Sasse, A. & Nurse, J.: *Cyber Security Awareness Campaigns: Why do they fail to change behaviour?: International Conference on Cyber Security for Sustainable Society*, 2016.
- [BSW08] Beautement, A.; Sasse, M. A.; Wonham, M.: The compliance budget. In (Bishop, M. Hrsg.): *Proceedings of the 2008 workshop on New security paradigms*. ACM, New York, NY, S. 47–58, 2008.

- [Da18] da Veiga, A.: An approach to information security culture change combining ADKAR and the ISCA questionnaire to aid transition to the desired culture. *Information and Computer Security* 5/26, S. 584–612, 2018.
- [Da20] da Veiga, A. et al.: Defining organisational information security culture—Perspectives from academia and industry. *Computers and Security* 92, 2020.
- [EN18] ENISA (European Union Agency for Network and Information Security): *Cybersecurity Culture Guidelines: Behavioural Aspects of Cybersecurity*, 2018.
- [Fo20] Forrester. Consulting Thought Leadership Paper Commissioned By Palo Alto Networks 2020.
- [Fi20] Field, A.: Survey fatigue and the tragedy of the commons: Are we undermining our evaluation practice? *Evaluation Matters—He Take Tō Te Aromatawai*, S. 1–11, 2020.
- [FioJ] Fingerhut, K. (o. J.): Narration als Lernform im Fachunterricht und die Erweiterung von Sprachkompetenz im Fachunterricht. <https://docplayer.org/63632962-Narration-als-lernform-im-fachunterricht-und-die-erweiterung-von-sprachkompetenz-im-fachunterricht-gekuerzte-fassung.html>, Stand: 25.04.2022.
- [GioJ] Gittermann, A.: Wie erstelle ich ein Tätigkeitsprofil? Engagiert! Mitgestalten - Engagementförderung der EKKW. https://engagiert-mitgestalten.de/magic/show_image.php?id=300392, Stand: 01.06.2023.
- [Gu83] Gutenberg, E.: *Die Produktion*. Springer, Berlin, Heidelberg, 1983 (1951).
- [JG07] Johnson, M. E.; Goetz, E.: Embedding Information Security into the Organization. *IEEE Security & Privacy Magazine* 3/5, S. 16–24, 2007.
- [KKÅ17] Karlsson, F.; Karlsson, M.; Åström, J.: Measuring employees' compliance – the importance of value pluralism. *Information and Computer Security* 3/25, S. 279–299, 2017.
- [LA09] Linek, S. B. & Albert, D.: Game-based Learning: Gender-Specific Aspects of Parasocial Interaction and Identification. Conference: International Technology, Education and Development Conference (INTED), 2009.
- [Po21] Pokoyski, D., Matas, I. & Haucke, A.: Qualitative Wirkungsanalyse Security Awareness in KMU: Tiefenpsychologische Grundlagenstudie im Projekt Awareness Labor KMU (ALARM) Informationssicherheit. Scholl, M. (Hrsg), Technische Hochschule Wildau, Wildau, 2021. <https://alarm.wildau.biz/sttic/d6490e49f8d31adfa35259134b8d1b9d/220316-alarm-studie-final.pdf>, letzter zugriff 25.04.2022.
- [Pr22] Prott, F., Küchler, U., Schuktomow, R. & Scholl, M.: Serious Games als Lernmethode zur Steigerung der Informationssicherheit. AKWI-Tagungsband zur 35. AKWI-Jahrestagung, 2022, 325–334, 2022.
- [Sa22] Sasse, A., Hielscher, J., Friedauer, J., Menges, U. & Peiffer, M.: Warum IT-Sicherheit in Organisationen einen Neustart braucht. In: 18. Deutscher IT-Sicherheitskongress des BSI, 2022.
- [Sc20] Scholl, M.: (How) Can Directive (EU) 2019/1937 on whistleblowers be used to build up a security and safety culture in Institutions? *International Journal of Information Security* 2/7, S. 40–57, 2020.

- [Sc85] Schein, E. H.: *Organizational Culture and Leadership*. Jossey-Bass Publishers, San Francisco, 1985.
- [Sc90] Schein, E. H.: *Organizational culture*. *American Psychologist* 2/45, S. 109–119, 1990.
- [SM16] Schanz, R.; Müller-Vorbrüggen, M.: *Ethik. (Ethics)*. In (Gessler, M. Hrsg.): *Kompetenzbasiertes Projektmanagement (PM 3)*. Handbuch für die Projektarbeit, Qualifizierung und Zertifizierung auf Basis der IPMA Competence Baseline Version 3.0. GPM Deutsche Gesellschaft für Projektmanagement e.V, Nürnberg, S. 1091–1103, 2016.
- [Ve22] Verizon: *Data Breach Investigations Report (DBIR) 2022*. <https://verizon.com/business/resources/reports/2022/dbir/2022-data-breach-investigations-report-dbir.pdf>, Stand 28.06.2022.
- [VSB20] Volkamer, M. Sasse, M. A. & Boehm, F.: *Analysing Simulated Phishing Campaigns for Staff*. In: *Computer Security. ESORICS 2020. Lecture Notes in Computer Science*, Vol 12580. Springer, Cham, 312–328, 2020.
- [NS10] van Niekerk, J. F.; Solms, R. von: *Information security culture: A management perspective*. *Computers & Security* 4/29, S. 476–486, 2010.
- [Yp14] Ypsilanti, A.; Vivas, A.B.; Räisänen, T.; Viitala, M.; Ijäs, T.; Ropes, D.: *Are Serious Video Games Something More than a Game? A Review on the Effectiveness of Serious Games to Facilitate Intergenerational Learning*. *Education and Information Technologies*, 19, S. 515–529, 2014.

Bildung - Interdisziplinäres
Forschen und Lernen in der
Ingenieurinformatik

Virtuous, multi-level integration of research and education within the project “European Digital Dynamic Mapping” (EDDY)

Excellent applied research as an enabler for excellent education in the context of Project-Based Learning in Higher Education Institutions


Rasmus Rettig ¹ and Maximilian Weltz ²


Abstract: Project- and Problem-Based Learning has been subject to research and implementation in engineering and computer science curricula for about the last 20 years. However, the projects and problems focused on in education have hardly been questioned for their relevance and effect. Considering current global and financial challenges, the authors propose to revise project- and problem-based approaches in education in the context of relevance and excellence in applied research. Besides coping with current cost-cutting, existing control mechanisms could ensure both relevance and excellence. Example projects based on more than 10 years of hands-on experience in Project-Based Learning in the context of externally funded research, for various classroom situations are included.

Keywords: Project Based Learning (PBL), Research and Education, Funded Projects, Mobility, Autonomous Vehicles, Data Analysis

1 Introduction

Over the last decades, education in engineering and computer science shifted from strict fact- and method-based, chalk-and-talk [MT03] learning towards problem- and project-based education. Main drivers for this paradigm shift were substantial changes in the engineering environment with increased levels of volatility, uncertainty, complexity, and ambiguity (VUCA) [Ci23]. New skills were required, like agile development methods, deriving information based on the analysis of huge data pools, Machine Learning (ML), and many more. In parallel, the future employers of engineering graduates identified a lack of required skills for the modern workplace, including communication and cooperation capabilities in an internationally shared work environment. At the same time,

¹ Hamburg University of Applied Sciences, Faculty of Engineering and Computer Science, Department of Information and Electrical Engineering, Berliner Tor 7, 20099 Hamburg, Rasmus.Rettig@haw-hamburg.de  <https://orcid.org/0000-0003-1524-4844>

² Hamburg University of Applied Sciences, Faculty of Engineering and Computer Science, Department of Information and Electrical Engineering, Berliner Tor 7, 20099 Hamburg, Maximilian.Weltz@haw-hamburg.de  <https://orcid.org/0009-0001-0941-9241>

universities faced a lack of students in the fields of Science, Technology, Engineering and Mathematics (STEM). Specifically, many Universities of Applied Sciences, subject to constant cost-cutting, did not keep up with the innovation pressure and limited their education program development to managing the status quo at controlled or minimized effort. Processes, like the update of curricula, were slow and did not reflect consequences of VUCA and the high speed of development of new requirements and competencies. As a measure, Project- and Problem-Based Learning (PBL) was introduced to foster the development of the new skillset. This led to the introduction of many toy-based projects without real-life relevance. However, without a funding source beyond the universities own budget, cost pressure led to a reduction of project-based education. Teaching lacking skills was partially left to industry to educate in the final semesters during internships or within dual degrees. However, this approach led to raising the question of the future role of Universities of Applied Sciences, if they were neither providing the relevant education, which is outsourced to industry, nor generating research results advancing society, which was left to (technical) universities with significant scientific ambitions and focus.

2 Related Work

PBL in engineering has been subject to research since about the turn of the century and has been introduced in many engineering and computer science curricula over the last decade. Recent publications include the idea of a Curriculum 4.0 to address COVID and VUCA requirements [Ci23]. The authors describe case studies from two different European universities. Other relevant publications include [MT03, PL11, MSH22]. The authors consistently focus on the learning outcomes of the students, but face difficulties in measuring results on a standardized scale. The authors of [BK22] analyze students' motivation and find that "Students live and learn in the real world". Thus, they need to know "that their work matters", and it is not sufficient to setup a "toy based" PBL. Consequently, project outcomes must be connected to real-life, beyond the state-of-the-art problems.

3 Questions of Research, Methods, and Objectives

Based on a mixed qualitative and quantitative assessment, the authors review their experience in project-based education for proven real-live projects, which withstand scientific scrutiny. Furthermore, a strategy for implementation within courses at Universities of Applied Sciences should allow readers to make best use of the author's experiences. The following core questions address the perspectives of "prerequisites" and "implementation":

- How to create PBL experiences based on relevant, real-live problems? What are the prerequisites? How to prove relevance?

- How to bridge the gap of missing financing?
- What are strategies to be considered for successful implementation in the context and limitations of research and teaching at a University of Applied Sciences?

4 Virtuously connecting excellent research and excellent teaching within the project EDDY

The project “European Digital Dynamic Mapping” (EDDY), started 2021, was funded by the Federal Ministry for Digital and Transport within the innovation initiative mFUND [FM21, FM23] for a total duration of 36 months. The main goal of the project is the development and implementation of Local Dynamic Maps (LDMs) [ET11, Ei19] to enable autonomous vehicles, reduce traffic emissions, and protect vulnerable road users. Specifically, the roles and responsibilities of cities and municipalities as the owner of public roads are to be revisited. The total project volume is almost € 3.6 million with work packages shared among seven project partners including the Free and Hanseatic City of Hamburg, the German Aerospace Center (DLR), the Hamburg University of Applied Sciences and four industry partners.

The Urban Mobility Lab at the Hamburg University of Applied Sciences contributes in multiple ways to reach the project goals: Main task is the prototypical implementation of an LDM-based visual localization system, which provides means for landmark based localization in areas with systematically wrong or insufficient GNSS signals. This complex subproject requires a combination of multiple skills in engineering and computer science disciplines, like visual, camera-based object classification in real time, multi-dimensional data-analysis, and complex, Kalman-based filtering. On top of that, strong application knowledge in the field of highly automated vehicles is a prerequisite to successfully tackle the project. It is subject to significant research over the last decade, as indicated by multiple publications in the field [Zh11, Sp15, Da22].

On top of the main task within EDDY, a significant number of additional ideas were developed during the project. Since localization based on point-objects was the base, additional tasks were taken over to classify and collect data for objects relevant for the project partners, which were not necessarily suitable for localization. Thus, as an example, the automated detection and localization of construction site beacons was identified as a value add for the city.

In total, all projects initiated since the beginning of EDDY can be classified into four groups, as summarized in Table 1.

Project Level	Name	Semester (typ.)	Duration (months)	Scientific Relevance
1	Standard Curriculum	4-6	3	0
2	Single-student and Team Projects	4-5	3	0
3	Bachelor Thesis	7	3	1
4	Master Thesis	10	6	2
5	PhD	16	36	3

Tab. 1: Overview of project classifications; “scientific relevance”: Potential for peer reviewed publications 0: none to 3: 100 % (must)

4.1 Level 1 – Integration into the Standard Curriculum

Since 2013, the author regularly teaches “Sensors in Mechatronics” in a one semester course in the 6th semester based on a classroom-focused, project-based approach with changing assignments as indicated in Table 2.

Semester	Project description	No. of students / No. of teams	Real-live relevance	Key-findings in formal students’ evaluations and direct feedback
SS13	Automated Parking	6/6	Toy-Robot	+ First time release from frontal-lecture-jail
WS13/14	Collision Avoidance	35/11	Toy-Robot	- Too little technical input from the lecturer + No frontal lectures, students’ initiative counts
WS14/15	HAW Luftschiiff [HA13]	34/11	Toy-Robot	
WS15/16	HAW Seaglider	31/10	Toy-Robot	+ Organization +/- Demanding project - Too few direct technical inputs from the lecturer
WS16/17	Spherobot [HA16]	21/7	Toy-Robot	

Semester	Project description	No. of students / No. of teams	Real-live relevance	Key-findings in formal students' evaluations and direct feedback
WS17/18	Magnet finding submarine [HA17]	18/6	Toy-Robot	
WS 18/19	Indoor Mapping [HA18]	15/5	Toy-Robot	
WS 20/21	Automated Platooning	22/8	Toy-Robot	+ Extremely good learning atmosphere + Ratio of technical inputs and project work
WS 21/22	Various, related to funded projects	25/7	Real, Projects contribute to EDDY	
WS 22/23	The safe-bike-box (EDDY)	23/7	Real, Projects contribute to EDDY	+ Student's feedback: "First time real-life mechatronics projects"

Tab. 2: Overview of level 1 project-based education.

Starting in the winter semester 21/22 the projects were directly associated with the project EDDY. Specifically, in the following winter semester, all teams were assigned the task to develop their "safe-bike-box" (Figure 1), a sensor and communication system to be integrated into an electric bike to measure the street quality and automatically detect evasive maneuvers by the cyclist. The data and events were to be sent to the local dynamic map server of the EDDY project and visualized with a geographic reference (Figures 2 and 3). Industry partners of the project supported by providing server access and by personally introducing visualization frameworks to the students. The direct input from industry partners was a significant motivator for the students and underlined the relevance of the assignment. The course was tightly organized based on the semester schedule with weekly mixed lecture and lab appointments. Technical inputs, for example on data structures or server communication, were included in the lecture segment. The progress of each team was closely monitored, with biweekly, team-specific feedback including technical aspects and the team progress. All teams reached the assigned goal with various creative add-ons and demonstrated their individual solution in a drive test at the end of the semester.

Students' evaluations consistently indicate a "very good" overall impression of the lecture with chances for improvement concerning the students' overall time effort and the balance of project work with technical content directly delivered by the lecturer.

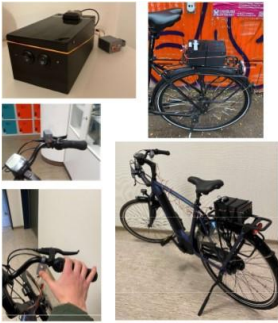


Fig. 1: Safe-Bike-Box mechanical setup [Ba23]

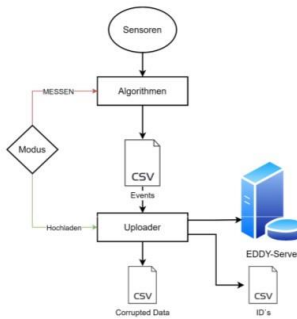


Fig. 2: Safe-Bike-Box data flow [Ba23]

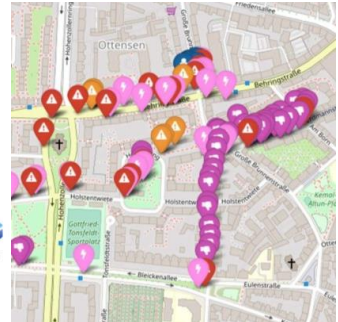


Fig. 3: Safe-Bike-Box measurement results [Ba23]

The main characteristics of the level 1 projects are:

- Tight, regular schedule for all teams (like lecture and lab).
- Mix of classroom, lab, and free working on the assigned project.
- Regular in-plenum technical inputs, joint progress reviews (technical, process).

The main learnings and observations are:

- A common task for all teams supports the effort and adds aspects of competition and cooperation.
- A strong start with a clear goal for the semester is required.
- Variable roles of the educator: Starting as the "customer setting the goals" to "providing technical inputs" to "observing each teams progress and providing support as a coach" and finally as the "examiner providing a grade for the result".

4.2 Level 2 – Single-Student and Team-Projects

One of the main measures to reflect PBL in the curriculum of Mechatronics was the integration of a bachelor project in the 4th semester. The project is scheduled for a duration of three months and runs in conjunction with an introduction to project management. Within the first months of the project EDDY, the need for inline anonymization during the test-drives for visual localization became evident and a jetson-nano-based system was developed by a team of three students (Figure 4). The system was integrated into the test-vehicle measurement system. Furthermore, the use of local dynamic maps for the

coordination of construction sites was discovered as a side-use, not planned in the original project (Figure 5). Another team of three students successfully pursued this task and integrated a detector for construction site beacons into the test vehicle setup.



Fig. 4: Inline anonymizer [SVB23]



Fig. 5: Detector for construction-site beacons [GWH23]

The main characteristics of the level 2 projects are:

- Start with a defined, written goal for the project and work-products including their quality.
- Open lab, access to the Urban Mobility Lab at the university.
- Biweekly, on-demand reviews on progress and results.

The main learnings and observations are:

- A detailed, written project charter is a must to avoid the project running off-track.
- These projects typically run into phases of low priority, which lead to delays.
- Due to the variable timing and availability of the students, the best practice was to connect to the project EDDY content-wise but to keep the timing disconnected.
- Good documentation is a priority to be able to use results in follow-up projects.

4.3 Level 3 – Bachelor Thesis Level

The duration of a bachelor thesis in Mechatronics at the Hamburg University of Applied Sciences is three months. During this time, the student is supposed to be fully focused on the bachelor thesis. Figures 6 and 7 shows a result, conducted in conjunction with EDDY. A measurement system for the classification of vertical acceleration and jerk was developed and connected to the EDDY-server to measure and document insufficient street quality.



Fig. 6: System setup [PI22]

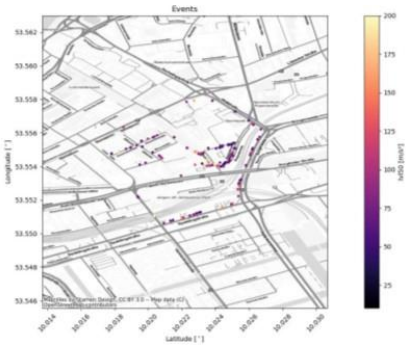


Fig. 7: Street quality events [PI22]

The main characteristics of the level 3 projects are:

- Start with a defined, written goal for the project and work-products including their quality.
- Full integration into the lab and research team.
- Biweekly, on-demand reviews on progress and results.

The main learnings and observations are:

- A detailed, written project charter is a must to avoid the project running off-track.
- Full dedication of the student is a must.
- External funding allows financing of tasks with direct project relevance (student tutor, part-time research associate, payment within industry standards) and required material.
- Partially, close coupling with external projects is possible.
- Good documentation is a priority to be able to use results in follow-up projects.
- Typically, only about 5 % of the bachelor thesis are of sufficient scientific quality to create peer-reviewed publications.

4.4 Level 4 – Master Thesis Level

Due to the longer duration of a master thesis of six months, and the advanced technical and methodological competencies of the students, a significant increase in complexity is expected. The master candidate was working fully integrated into the research project EDDY and paid within industry standards for students in their master thesis. A localization system for the test vehicle (Figure 8) to complement other localization techniques in

locations where they do not work (Figure 9), was developed. The system used landmark-based localization (Figures 10 and 11).



Fig. 8: Test vehicle of the Urban Mobility Lab [We23]



Fig. 9: GNSS-performance along the test track for automated and connected driving (TAVF) [We23]



Fig. 10: Basic functionality for landmark-based localization [We23]

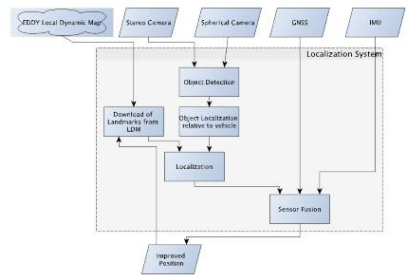


Fig. 11: System developed within a Master Thesis [We23]

The main characteristics of the level 4 projects are:

- Start with a defined, written goal for the project and work-products including their quality.
- Full integration into the lab, the research team, and the project EDDY.
- Regular, weekly review on progress and results.

The main learnings and observations are:

- A detailed, written project charter is a must. However, due to the length and complexity, adjustments may be required.
- Full dedication of the candidate.
- External funding allows financing of tasks with direct project relevance (student tutor, part-time research associate, payment within industry standards).

- Good documentation is a priority to be able to use results in follow-up projects.
- Currently, only about 10 % of the Master Thesis are of sufficient scientific quality to create peer-reviewed publications.

4.5 Level 5 –PhD-Thesis Level

Up to writing this publication, no PhD was completed within the project EDDY. However, the author successfully lead a candidate up to reaching a “Dr. phil.” within the funded project “X-eptance Impulse” [XE21].

The main characteristics of the level 5 projects are:

- Initial finding-phase to identify relevant questions of research.
- Full scale scientific discovery process, incl. dead ends.
- Full integration into the research team and project (full time research associate).

The main learnings and observations are:

- PhD-level long-term involvement is a must to reach a sufficient level of competencies and knowledge to reach international acceptance.
- Due to the lack of permanent mid-level scientist positions for research at Universities of Applied Sciences, external funding is a must.

5 Results and Discussion

Based on the analysis of level 1 to level 5 projects, the keys to successfully integrate research and education at Universities of Applied Sciences are:

- Consistent and continuous acquisition of external funding for several years.
- Paid research associates (Master or PhD level) focusing on the project for longer periods are required to reach a sufficient competence level to successfully compete and contribute to science.
- Shorter project types must be smartly structured and well controlled to contribute to funded projects.

6 Summary and Outlook

Virtuous integration of research and education is the key for Universities of Applied Sciences to stay relevant and competitive in the development of science and education.

Funding is a must and successful acquisition is a valid indicator to demonstrate relevance for industry, science, and society. By smartly integrating existing strengths, system-immanent deficiencies, like the lack of permanent midlevel positions for scientists, can be overcome.

7 Acknowledgements

Funding of the project EDDY by the German Federal Ministry for Digital and Transport within the mFUND initiative is thankfully acknowledged (FKZ: 19F2208C). The authors furthermore acknowledge the cooperation and helpful discussions with Heilbronn University of Applied Sciences.

8 Author's Contributions

Rasmus Rettig: Conceptualization, Methodology, Writing, Funding Acquisition, Supervision. Maximilian Weltz: Investigation, Visualization, Writing, Review & Editing.



References

- [Ba23] Bahr, M. et al.: SafeBikeBox. HAW Hamburg, 17/01/2023.
- [BK22] Boss, S.; Krauss, J.: Reinventing Project-Based Learning, Your Field Guide to Real-World Projects in the Digital Age. Third Edition, International Society for Technology in Education, 2022.
- [Ci23] Ciolacua, M. I. et. al.: Fostering Engineering Education 4.0 Paradigm Facing the Pandemic and VUCA World. In (Longo F., ed.): 4th International Conference on Industry 4.0 and Smart Manufacturing, Procedia Computer Science 217, pp. 177–186, 2003.
- [Da22] Dauplain, X. et al.: Conception of a High-Level Perception and Localization System for Autonomous Driving. In (Passaro, V.M.N., ed.): Sensors 22-24, 9661, <https://doi.org/10.3390/s22249661>, 2022.
- [Ei19] Eiter, T. et al.: Towards a Semantically Enriched Local Dynamic Map. In (Prathombutr, P., ed.): International Journal of Intelligent Transportation Systems Research 17, <https://doi.org/10.1007/s13177-018-0154-x>, pp. 32–48, 2019.
- [ET11] ETSI: TR 102 863 V1.1.1 (2011-06) – Technical Report, Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Local Dynamic Map (LDM); Rationale for and guidance on standardization. https://www.etsi.org/deliver/etsi_tr/102800_102899/102863/01.01.01_60/tr_102863v010101p.pdf, accessed 20/05/2023.

- [FM21] Federal Ministry for Digital and Transport: European Digital Dynamic Mapping – EDDY, <https://bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/eddy.html>, accessed 20/05/2023.
- [FM23] Federal Ministry for Digital and Transport: mFUND, <https://bmdv.bund.de/DE/Themen/Digitales/mFund/Ueberblick/ueberblick.html>, accessed 20/05/2023.
- [GWH23] Günther, M.; Wagner, S.; Heidenreich, S.: Baustellenbakenerkennung. HAW Hamburg, 04/03/2023.
- [HA13] HAW Hamburg: Mechatronik: Studierende der HAW Hamburg bauen Luftschiffe. <https://youtu.be/kVGofntR6T4>, accessed 20/05/23.
- [HA16] HAW Hamburg: Mechatronik: Studierende entwickeln autonom fahrende Systeme. <https://youtu.be/Zsswcag7I8U>, accessed 20/05/23.
- [HA17] HAW Hamburg: Autonome Unterwasserfahrzeuge: Mechatronik im 6. Semester an der HAW Hamburg. <https://youtu.be/RKMuo8KgU0s>, accessed 20/05/23.
- [HA18] HAW Hamburg: HAW Autonomous Mapping Vehicle. <https://youtu.be/BooiXiIHMyM4>, accessed 20/05/23.
- [MSH22] Mursid, R.; Saragih, A. H.; Hartono, R.: The effect of the blended project-based learning model and creative thinking ability on engineering students' learning outcomes. In (Sahin, I., ed.): International Journal of Education in Mathematics, Science, and Technology (IJEMST), 10-1, <https://doi.org/10.46328/ijemst.2244>, pp. 218-235, 2022.
- [MT03] Mills, J. E.; Treagust, D. F.: Engineering education-Is problem-based or project-based learning the answer?. Australasian Journal of Engineering Education 3, pp. 2-16, 2003.
- [PL11] Pucher, R.; Lehner, M.: Project Based Learning in Computer Science – A Review of More than 500 Projects. In (Bekirogullari, Z., ed.): Procedia - Social and Behavioral Sciences, Volume 29, ISSN 1877-0428. <https://doi.org/10.1016/j.sbspro.2011.11.398>, pp. 1561-1566, 2011.
- [Pl22] Plaggemeyer, J.: Entwicklung eines Straßenqualitätsindikators mit Anbindung an einen Geoserver. Bachelor-Thesis, HAW Hamburg, 05/12/2022.
- [Sp15] Spangenberg, R.: Landmark-based Localization for Autonomous Vehicles. PhD-Thesis, FU Berlin, 2015.
- [SVB23] Sauer, R.; Vogler, T.; Blessin, R.: Inline Anonymizer. HAW Hamburg, 18/04/2023.
- [We23] Weltz, M.: Entwicklung eines Multi-Sensor Systems zur Lokalisierung im urbanen Umfeld auf Basis einer lokalen dynamischen Karte. Master-Thesis, HS Heilbronn, 14/04/2023.
- [XE21] X-Eptance Impulse: Project Website. <https://www.haw-hamburg.de/forschung/forschungsprojekte-detail/project/project/show/x-eptance-impulse/>, accessed 20/05/2023.
- [Zh11] Zhu, Z. et al.: High-precision localization using visual landmarks fused with range data. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), Piscataway, NJ, IEEE, ISBN 978-1-4577-0394-2, pp. 81–88, 2011.

Agilität, Praxisbezug und Teamarbeit im Softwareentwicklungsprojekt der Ingenieurinformatik

Wie können fachliche Aspekte zusammen mit sozialer, Methoden- und Organisationskompetenz erfolgreich im Projektformat vermittelt werden?

Frank Neumann ¹ und Christina Kratsch ²


Abstract: Projektmodule haben sich in Informatik-Studiengängen als eine Lehrform etabliert, mit der Studierenden die Gelegenheit gegeben wird, Lösungen für Aufgabenstellungen mit hohem Praxisbezug im Team zu konzipieren und zu implementieren. Hierbei werden die in anderen Modulen erworbenen Fachkenntnisse aktiviert und zur Anwendung gebracht. Für das Softwareentwicklungsprojekt im 4. Fachsemester des Studiengangs Ingenieurinformatik der HTW Berlin wird ein Konzept entwickelt, mit dem wichtige Anforderungen und Trends der heutigen Arbeitswelt adressiert werden sollen. Dazu gehören neben der Teamarbeit auch die Anwendung von agilen Methoden sowie ein für die HAWs typischer starker Praxisbezug. Im Rahmen des vorliegenden Artikels wird der für die Umsetzung dieses Lehrkonzepts notwendige organisatorischen Rahmen sowie dessen iterative Weiterentwicklung beschrieben. Für eine abschließende Bewertung des Projektmoduls werden die Evaluierungsergebnisse der in das Modul involvierten Studierenden in Hinblick von Kriterien wie Praxisbezug, Teamarbeit und Projektmanagementfähigkeiten ausgewertet.


Keywords: Ingenieurinformatik, Projektmodul, Lehrkonzept, Praxisbezug, Teamarbeit, Agiles Projektmanagement, Scrum

1 Einleitung

1.1 Komplexe Anforderungen an die Lehre

Informatik-Studiengänge sollen nicht nur Fachwissen vermitteln, sondern auch praxisorientierte Fähigkeiten und Kompetenzen, die im frontalen Unterricht nicht umsetzbar sind. Studierende benötigen für die Entwicklung moderner Softwarelösungen natürlich primär Kenntnisse in Programmierplattformen und den dazugehörigen

1 HTW – Hochschule für Technik und Wirtschaft Berlin, Fachbereich 2: Ingenieurwissenschaften - Technik und Leben, Wilhelminenhofstr. 75a, 12459 Berlin, frank.neumann@htw-berlin.de,  <https://orcid.org/0000-0002-8530-3283>

2 HTW – Hochschule für Technik und Wirtschaft Berlin, Fachbereich 2: Ingenieurwissenschaften - Technik und Leben, Wilhelminenhofstr. 75a, 12459 Berlin, christina.kratsch@htw-berlin.de,  <https://orcid.org/0000-0003-0565-1112>

Werkzeugen. Die Anforderungen der Praxis gehen aber weit darüber hinaus und ergeben sich sowohl aus technischer, organisatorischer als auch sozialer Perspektive:

A1: Zeit- und Teamkonflikte managen: Softwareprojekte sind langfristige Arbeiten von Wochen bis Jahren, welche in Teams umgesetzt werden. Soziale Konflikte gehören zum Alltag und können Einfluss auf Produktivität und Projekterfolg haben. Fähigkeiten in Selbstorganisation und Kommunikation gehören ebenso zum Repertoire eines Softwareentwicklers wie Kenntnisse von Team-Dynamiken und der Umgang damit.

A2: Professionelle Kundenkommunikation: Ein Softwareentwickler muss in der Lage sein, empathisch zuzuhören, Kunden zu verstehen und zu antizipieren, sowie zielsicher zwischen dem Gewünschten und dem technisch Machbaren zu vermitteln. Dies erfordert den reflektierten Umgang mit Sprache sowie Kenntnisse strukturierter Fragemethoden.

A3: Zielkonflikte managen: Softwareprojekte führen oft zu Zielkonflikten zwischen Projektumfang, Aufwand und Zeit. Ein erfolgreiches Team kennt diese Konflikte, und kann damit nachhaltig und kreativ umgehen, um die Konflikte professionell aufzulösen.

A4: Typische Muster des agilen Arbeitens verstehen: Softwareentwicklung erfolgt meist in agilen Frameworks wie SCRUM oder Kanban. Diese gehören zum Basis-Skill eines Informatikers. Die zugehörigen Verantwortlichkeiten ebenso wie typische Konfliktmuster im Teamalltag lassen sich durch direktes Erleben gut vermitteln. Auch Probleme des agilen Arbeitens das Finden eines guten Sprint-Tempos oder der Umgang mit technischer Schuld sind nur in der gelebten Praxis gut nachvollziehbar.

A5: Kontinuierliches Lernen und Umgang mit neuesten Technologien: Wie viele Berufsbilder unterliegt die Softwareentwicklung einer rasanten Entwicklung und erfordert die Bereitschaft zum kontinuierlichen Lernen bis ins hohe Alter. Je nach Fachrichtung lassen sich Details dieses Fortschritts nicht im üblichen Lehrbetrieb abbilden.

Um all diesen Anforderungen gerecht zu werden, haben sich Projektmodule als effektive Lehrform etabliert. Diese ermöglichen den Studierenden die Konzeption und Implementierung praxisbezogener Lösungen für komplexe Aufgabenstellungen. Insbesondere im Bereich der Softwareentwicklung bieten Projektmodule eine geeignete Plattform, um die erworbenen Fachkenntnisse aus anderen Modulen aktiv anzuwenden.

1.2 Softwareentwicklungsprojekt als praxisnahes Konzept

Das vorliegende Paper widmet sich dem Lehrkonzept des Softwareentwicklungsprojektes (SWE) im 4. Fachsemester des Studiengangs Ingenieurinformatik an der HTW Berlin. In diesem Projekt entwickeln Studierende in Teams von 3-4 Teilnehmern über die Dauer eines Semesters eine Softwarelösung für einen unternehmerischen Auftraggeber und werden von Dozenten bei der Umsetzung begleitet. Ziel dieses Lehrkonzepts ist es, wichtige Trends und Anforderungen der heutigen Arbeitswelt, insbesondere Agilität, Praxisbezug und Teamarbeit, in den Fokus zu rücken. Durch die Integration agiler

Methoden und den Einsatz von Teamarbeit soll den Studierenden eine realitätsnahe Lernumgebung geboten werden, die sie optimal auf ihre beruflichen Herausforderungen vorbereitet. Agile Methoden ermöglichen es dabei, flexibel auf neue Anforderungen zu reagieren, iterative Entwicklungsprozesse anzuwenden und eine enge Zusammenarbeit im Team zu pflegen. Diese Fähigkeiten sind von entscheidender Bedeutung, um in der heutigen dynamischen Arbeitswelt erfolgreich zu sein.

Darüber hinaus legt das Lehrkonzept einen starken Fokus auf Praxisbezug. Studierende haben die Möglichkeit, erlerntes Wissen und Fähigkeiten in konkreten Projektsituationen anzuwenden und so Kompetenzen weiterzuentwickeln. Durch die Zusammenarbeit mit externen Stakeholdern wird eine enge Verknüpfung zur realen Arbeitswelt geschaffen, was den Studierenden wertvolle Einblicke und Kontakte ermöglicht.

Im Rahmen dieses Artikels wird der organisatorische Rahmen des Lehrkonzeptes und dessen stufenweise Verbesserung beschrieben. Zudem werden Evaluierungsergebnisse von involvierten Studierenden ausgewertet in Hinblick auf Kriterien wie Praxisbezug, Teamarbeit und Projektmanagementfähigkeiten. Diese Ergebnisse der Evaluierung dienen dazu, das Lehrkonzept kontinuierlich an die Bedürfnisse der Studierenden und der Arbeitswelt anzupassen. Insgesamt bietet das Paper einen detaillierten Einblick in das Lehrkonzept des Softwareentwicklungsprojektes. Es zeigt die Relevanz der Integration von Agilität, Praxisbezug und Teamarbeit in den Unterricht und bietet wertvolle Erkenntnisse für andere Bildungseinrichtungen.

2 Kompetenzfelder des Softwareentwicklungsprojekts als Lehr- und Lernziele

Die Ingenieurinformatik ist ein Wissensgebiet, das eng mit der Digitalisierung und der Anwendung von Informatik-Kenntnissen in Ingenieurdisziplinen verbunden ist. Je nach Ausprägung im jeweiligen Studiengang stehen IT-Systeme in Anwendungsbereichen wie z.B. Maschinenbau, Elektrotechnik oder Verfahrenstechnik im Vordergrund. Die Gesellschaft für Informatik (GI) bezeichnet interdisziplinäre Studiengänge wie die Ingenieurinformatik als *Typ 3* und grenzt sie damit von reinen Informatikstudiengängen (*Typ 1*) und solchen Curricula ab, bei denen es nur eine spezielle Anwendungsdomäne gibt (*Typ 2*) [Zu16]. In *Typ 3* Studiengängen liegen der Informatikanteil und die kumulierten Anteile der anderen Fachdisziplinen (z.B. Mechanik, Elektronik, Elektrotechnik, Regelungstechnik) annähernd auf dem gleichen Niveau.

Aus der Einordnung der Ingenieurinformatik als interdisziplinärem Studiengang ergibt sich die Notwendigkeit, dass fachübergreifende und verstärkt überfachliche Kompetenzen zu vermitteln sind. Absolvent:innen solcher Studiengänge werden in heterogenen Anwendungsfeldern tätig sein und benötigen eine interdisziplinäre Sicht, um mit den Stakeholdern im jeweiligen Anwendungskontext erfolgreich zusammen-zuarbeiten.

Die GI schlägt konkrete Ausbildungsziele für das Informatikstudium in den jeweiligen Studiengangs-Typen vor, die in folgende Bereiche unterteilt wird [Zu16]:

- Formale, algorithmische und mathematische Kompetenzen
- Analyse-, Entwurfs-, Realisierungs- und Projektmanagement-Kompetenzen
- Technologische Kompetenzen
- Fachübergreifende Kompetenzen
- Methoden- und Transferkompetenz
- Soziale Kompetenzen und Selbstkompetenzen

Bei der Vermittlung von Kompetenzen sollte auch Art und Komplexität des jeweiligen Vermittlungskontextes betrachtet werden. Die GI orientiert sich hierbei an der Anderson Krathwohl Taxonomy [LD01] und vereinfacht diese zu einer reduzierten AKT-Matrix. Dort werden folgende Abstufungen von Vermittlungskontexten aufgeführt [Zu16]:

- K1: ohne Kontextualisierung
- K2: kleine Beispiele
- K3: komplexere Beispiele
- K4: interne Projekte
- K5: betriebliche Projekte

Die für das SWE typischen Projekte bewegen sich hier mit den Abstufungen K4 und K5 durchweg im Bereich starker Kontextualisierung und hoher Komplexität. Aufbauend auf der o.g. Kompetenz-Taxonomie soll nun nachfolgend betrachtet, welche Qualifikationen in den jeweiligen Inhaltsbereichen durch das SWE vermittelt werden sollen. Abschließend wird die dem SWE zugrunde liegende Modulbeschreibung vorgestellt.

2.1 Formale, algorithmische und mathematische Kompetenzen

Die fachlichen Kompetenzen in diesem Inhaltsbereich werden in den ersten drei Fachsemestern in den Modulen *Mathematik*, *Programmierung*, *Modellierung* und *Algorithmen* erworben. Hierbei fokussiert das Fach *Modellierung* im 2. Fachsemester auf der Vermittlung von Konzepten zum Entwurf objektorientierter Software mittels UML [HT23]. Diese Fertigkeiten werden im SWE für die Analyse und den Entwurf eines komplexen Projektes angewendet, indem eine technische Spezifikation des zu entwerfenden Systems erstellt werden muss.

2.2 Analyse-, Entwurfs-, Realisierungs- und Projektmanagement-Kompetenzen

Bei den Analyse-Kompetenzen steht insbesondere die Fähigkeit im Vordergrund, die Anforderungen des Projektes anhand des Lastenheftes zu verstehen, es z.B. hinsichtlich etwaiger Inkonsistenzen zu analysieren und mit den Stakeholdern zu diskutieren. Basierend auf der Analyse des Lastenheftes erarbeiten die Studierenden initial ein Kontextdiagramm des zu implementierenden Systems und identifizieren die Schnittstellen zu den eingebundenen Fremdsystemen. Nachfolgend wird das System in Subsysteme und Module partitioniert und dessen logische Systemarchitektur als Komponentendiagramm beschrieben. Die Studierenden identifizieren außerdem nicht-funktionale Anforderungen und diskutieren hierzu mit den Stakeholdern. Anschließend werden drei bis fünf dieser Anforderungen ausgewählt und als Kriterien für die Wahl der Programmiersprache, Entwicklungsumgebung und Software-Bibliotheken angewendet.

Für die Gestaltung der Benutzeroberfläche werden ebenfalls Entwurfskompetenzen benötigt. Hierbei werden Mockups mit industrietüblichen Werkzeugen entwickelt.

Für die softwaretechnische Umsetzung werden den Studierenden Workshops zu Themen angeboten, die für die bearbeiteten Projekte relevant sind, jedoch im bisherigen Studienverlauf nicht vermittelt wurden. Exemplarische Themen sind

- der Einstieg in das Unit-Testing sowie der
- Einsatz von Continuous Integration and Deployment in *gitlab*.

Um in Teamarbeit erfolgreich eine Softwareprojekt durchführen zu können, sind Kompetenzen im Bereich Projektmanagement notwendig. Da die entsprechende Lehrveranstaltung parallel zum Projektmodul läuft, wird im Rahmen des Projekts Scrum als agiles Vorgehensmodell mit der höchsten industriellen Verbreitung vorgestellt und in drei Sprints praktiziert. Vom Unternehmen *ditigal.ai* wurden für die praktischen Umsetzung von Scrum 200 Lizenzen der Software *agility* gesponsert.

2.3 Technologische Kompetenzen

Im technologischen Bereich können alle in den ersten drei Semestern vermittelten Qualifikationen für das Projekt zur Anwendung gebracht werden. In den ersten Durchgängen des SWE herrschte eine gewisse Unsicherheit, wie mit Projektthemen umzugehen ist, für die bisher keine technologischen Kompetenzen vermittelt wurden. Hierzu ist über die Jahre die Überzeugung gewachsen, dass Studierende mit der entsprechenden Unterstützung in der Lage sind, sich im Rahmen des SWE in neue Technologien einzuarbeiten. Letztlich spiegelt das ein Erfordernis der modernen Arbeitswelt wider, in der sich Technologien sehr dynamisch weiterentwickeln, andere jedoch auch an Bedeutung verlieren. Die Einarbeitung in neue Technologien unterstützen die Lehrenden bzw. die Projektpartner durch Lehrangebote in folgenden Bereichen:

- Einstieg in neue Programmiersprachen wie z.B. Python oder Kotlin
- Erlernen der dazugehörigen integrierten Entwicklungsumgebungen
- Vermittlung von Kenntnissen zu RESTful Webservices
- Einführung zu spezifischen Software Development Kits
- Grundlagen der Künstlichen Intelligenz

2.4 Fachübergreifende Kompetenzen

Bei den fachübergreifenden Kompetenzen hat sich ein Fokus auf der Vermittlung von Kenntnissen zum Urheberrecht und zu Lizenzmodellen in der Softwareentwicklung als sehr hilfreich herausgestellt. Hierbei erlernen die Studierenden für ihr konkretes Projekt, welche Lizenz für die zu entwickelnde Software unter den Randbedingungen der gewählten Softwarebibliotheken, Programmiersprachen und Datenbanken anwendbar ist und den Kooperationspartnern ausreichende Nutzungsrechte einräumt.

2.5 Soziale Kompetenzen und Selbstkompetenzen

In diesem Inhaltsbereich ist die Kommunikationskompetenz der Studierenden hervorzuheben, da sie ihre Konzepte und jeweiligen Sprintergebnisse in regelmäßigen Abständen darstellen müssen. In die Benotung des Projekts fließen sowohl formale Kriterien der technischen Spezifikation als auch der Abschlusspräsentation ein. Hierbei werden die Kriterien kommuniziert und zur Anwendung gebracht, die für die Niveaustufe der Bachelorarbeit und des Bachelorkolloquiums relevant sind.

Im Bereich des Konfliktmanagements hat sich herauskristallisiert, dass die Studierenden durch vorhergehende Projekte im schulischen Bereich mit Konflikten meist adäquat umgehen können. Hierzu tragen auch die klaren Regeln des Scrum-Prozesses hinsichtlich Planung und Arbeitsverfolgung bei. In der ersten Projektveranstaltung wird hinsichtlich dieser Thematik explizit sensibilisiert und ein einfaches Regelwerk vorgestellt. Für weitergehende Maßnahme zum Konfliktmanagement hat sich bisher keine Notwendigkeit ergeben.

Da der Studiengang Ingenieurinformatik hinsichtlich der Herkunftsländer sehr divers zusammengesetzt ist, versuchen die Lehrenden, diese Diversität auch in den einzelnen Projektgruppen abzubilden. Dadurch soll u.a. sichergestellt werden, dass es in jeder Projektgruppe deutsche Muttersprachler gibt.

2.6 Modulbeschreibung

Die im Modulhandbuch des Studiengangs Ingenieurinformatik der HTW Berlin veröffentlichte Modulbeschreibung des SWE fasst die dargestellten Kompetenzbereiche treffend zusammen [HT23]:

„Die Studierenden sind befähigt, ein Softwareentwicklungsprojekt im Aufgabenbereich des Ingenieurwesens selbständig zu analysieren, zu strukturieren, zu entwerfen und umzusetzen. Sie beherrschen die Adaption des in den ingenieurwissenschaftlichen Grundlagenfächern erworbenen Wissens. Neben der fachlichen Qualifikation haben sie soziale, Methoden- und Organisationskompetenzen erworben. Sie haben erste Erfahrungen mit einer Software-Entwicklungsmethodik erworben.“

3 Konzept der Projektaktivitäten

In der beschriebenen Form startete das SWE bereits im Sommersemester 2015 mit einem Lehrkonzept, das aus den gewonnenen Erfahrungen, den Hinweisen der Unternehmen und Studierenden sukzessive weiterentwickelt wurde.

3.1 Ablauf der Projektaktivitäten

Für das SWE stehen während eines typischen Semesters ca. 13 Wochen zur Verfügung. Diese Zeitspanne wird in fünf Phasen unterteilt, welche in Abb. 1 für ein exemplarisches Projekt und mit beispielhaften User Stories abgebildet sind. Ziel des Projektes war die Erstellung einer Festpunktfelddatenbank für ein Logistik- und Mobilitätsunternehmen mit zugehöriger Benutzerverwaltung.

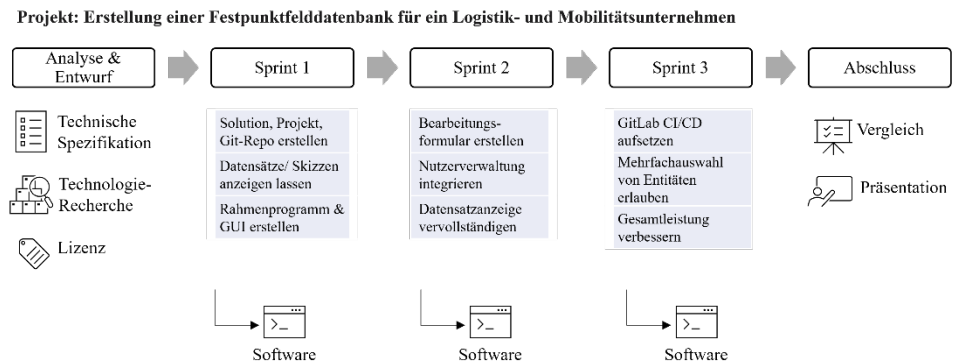


Abb. 1: Ablauf und Inhalte eines Projektes zur Erstellung einer Festpunktfelddatenbank für ein Logistik- und Mobilitätsunternehmen.

In Tab. 1 werden ergänzend dazu die von den Studierenden zu erstellenden Ergebnisse detaillierter ausgeführt.

Jeweils am Ende der ersten vier Abschnitte liefern die Projektgruppen ihre Arbeitsresultate ab und erhalten dafür durch die Lehrenden eine Bewertung in Punkten, in die das Feedback der Product Owner einfließt. Die als Ergebnis der jeweiligen Sprints entwickelte partiell umgesetzte Software wird dafür auf der Grundlage der umgesetzten User Stories getestet und bewertet. Durch dieses frühe und regelmäßige Feedback von Dozenten und Product Owner wird so versucht, eine möglichst wertstiftende und qualitativ hochwertige Software zu erreichen. Die Bewertung am Ende jedes Sprints geht auf Umfang und Qualität des entwickelten Codes ebenso ein wie auf die strukturierte Planung der Phasen und die Effizienz bei der Umsetzung.

Projektphase	Inhalte und Ergebnisse
Technische Spezifikation (Analyse und Entwurf)	<ul style="list-style-type: none"> • Use Case Diagramm, Beschreibung der Anwendungsfälle • Nicht-funktionale Anforderungen & Lösungsansätze dazu • Fachliche Strukturen und Modelle • Kontext, Schnittstellen (Level 0) • Mockups der Benutzeroberfläche • Systemarchitektur (Level 1), Technologie-Stack • Festlegung/Begründung für Lizenzmodell
Sprint 1-3	<ul style="list-style-type: none"> • Sprintziel und Weekly Scrums • Sprintbacklog bestehend aus User Stories und Aufgaben • Sprint Review und Abgabe einer funktionsfähigen Software
Abschlusspräsentation	<ul style="list-style-type: none"> • Katalog der zu realisierenden Anforderungen • Umsetzung der Anforderungen in der Applikation anhand von typischen Anwendungsszenarien • Vergleich der Spezifikation mit dem umgesetzten Stand

Tab. 1: Projektphasen des Softwareentwicklungsprojekts und deren Ergebnisse

3.2 Initiales Konzept der Projektaktivitäten

Initial startete das SWE mit der Idee, Stakeholder aus Unternehmen und Forschungseinrichtungen als Product Owner einzubeziehen. Aufgrund der zeitlichen Constraints für die Vorbereitung konnte zunächst nur eine Firma zur Mitarbeit gewonnen werden. Dadurch stand nur ein gemeinsames Thema für alle Projektgruppen zur Verfügung. Das Thema wurde vom Product Owner in einem Lastenheft beschrieben. Bei der Gruppengröße wurde auf vier Studierende orientiert. Bereits initial wurde der in Tab. 1 beschriebene Ablauf adoptiert. Allerdings wurden damals nur die technischen Spezifikationen und Sprintergebnisse bewertet. Scrum wurde aufgrund der erkennbaren Durchdringung der IT-Industrie als Vorgehensmodell gewählt, wobei die Dozierenden als Scrum Master agieren. Anfangs konnten die Konsultationstermine von den einzelnen

Teams frei gewählt werden, wodurch sich ein wöchentliches Scrum nur schwer sicherstellen ließ.

3.3 Weiterentwicklung des Konzepts

Ausgehend von eigenen Erfahrungen und Feedback der Studierenden wurden in den Folgesemestern die Rahmenbedingungen so angepasst, dass die Lernziele bestmöglich erreicht werden. Damit jede Gruppe ein individuelles Thema bekommt, wurden deutlich mehr Firmen und Forschungseinrichtungen als Product Owner gewonnen. Um eine Vergleichbarkeit der Bewertungen bei unterschiedlichen Themen zu gewährleisten, wird der Zeitaufwand für die Umsetzung der User Stories durch jedes Team so geschätzt und gestaltet, dass jeder Studierende ca. 10 Stunden Projektarbeit investieren muss. Drei Studierende pro Team haben sich als optimale Gruppengröße etabliert.

Anfangs erfolgte die Wahl des Themas noch analog und per Moderation der Lehrenden, was sich als zeitaufwändig und diskussionsintensiv herausstellte. Aufbauend auf den Erfahrungen aus der Pandemiezeit wird nun die entsprechende Gruppenwahl-Aktivität aus dem Moodle mit deutlich reduziertem Aufwand verwendet. Dies ermöglicht es den Studierenden auch, Projekthemen nach eigenen Interessen und individuellen Fähigkeiten zu wählen. Die Dozenten sind darüber hinaus bemüht, Themen auf ähnlichem Schwierigkeitslevel auszuwählen und unvorhergesehene Schwierigkeit in der individuellen Betreuung im Semester auszugleichen. Mindeststandards in der Bewertung tragen ebenfalls zur Fairness und Vergleichbarkeit der Gruppenthemen bei.

Durch die verbesserte Verfügbarkeit von Lösungen für Videokonferenzen können die Product Owner nun remote an Sprint Reviews teilnehmen. Wahlweise wird dies auch für Weekly Scrums praktiziert, welche nun verpflichtend mit festgelegten Zeiten sind.

Da Abschlusspräsentationen ohne Bewertung teilweise recht lieblos vorbereitet wurden, werden diese nun in die Bewertung einbezogen. Zur Orientierung stehen außerdem Checklisten für die Aufgaben pro Woche während der Spezifikationsphase bereit.

Über die Jahre hat es sich als sinnvoll herausgestellt, meist einstündige Workshops zu Themen anzubieten, die für die Projekte relevant sind, jedoch im bisherigen Studium nicht vermittelt wurden (vgl. Abschnitt 2.3). Die Workshops finden üblicherweise vor den Konsultationsterminen statt und bestehen aus Vorträgen zu technisch-organisatorischen Themen (z.B. REST-API Design oder agile Projektplanung), gefolgt von individueller Unterstützung der Projektteams zum Thema nach Bedarf.

Im Unterschied zu „klassischen“ Projektformaten gliedert sich die Betreuung durch die Lehrenden in vielfältigere Tätigkeiten auf. Hierbei hervorzuheben ist der signifikante organisatorische Aufwand beim Einwerben von Projekthemen sowie die Veranstaltung von Sprintreviews und Abschlusspräsentation. Durch insgesamt vier Abgaben während des Projekts ist der auch Bewertungsaufwand hoch. Dies sichert jedoch zeitnahes Feedback für die Gruppen für Folgesprints und steigert dadurch den Lernerfolg. Eine gute

Tradition ist außerdem die gemeinsame Reflektion mit Product Ownern zu den gemachten Erfahrungen und Verbesserungsmöglichkeiten am Ende des Semesters.

4 Evaluierung durch Studierende

An der HTW Berlin finden regelmäßig Evaluationen der Lehrveranstaltungen statt, bei denen auch für Projektmodule spezifische Themen durch die Studierenden bewertet werden. Relevante Teilbereiche der Evaluationsergebnisse aus vier Semestern sollen nachfolgend ausgewertet werden. In den Evaluationen werden Fragestellungen zu Kriterien wie Praxisbezug, Teamarbeit und Projektmanagementfähigkeiten bewertet.

Der *Praxisbezug* wird bei den Studierenden durch folgende Fragestellungen im Bereich zwischen „trifft zu“ und „trifft nicht zu“ bewertet.

- „Die Lehrende/der Lehrenden stellt Aufgaben mit Bezug zur Praxis.“

Für den Bereich *Teamarbeit* sind folgende Kriterien relevant:

- „... unterstützt eine konstruktive Gruppen- und Partnerarbeit.“
- „... gibt konstruktive Rückmeldungen zum Stand des Projektes.“
- „... habe ich meine Arbeitstechniken verbessert.“
- „... kann ich meine Stärken / Schwächen bei der Teamarbeit besser einschätzen.“
- „... kann ich mich besser für eine konstruktive Atmosphäre innerhalb von Teams einsetzen.“

Die Entwicklung im Bereich *Projektmanagement* wird durch folgende Frage bewertet:

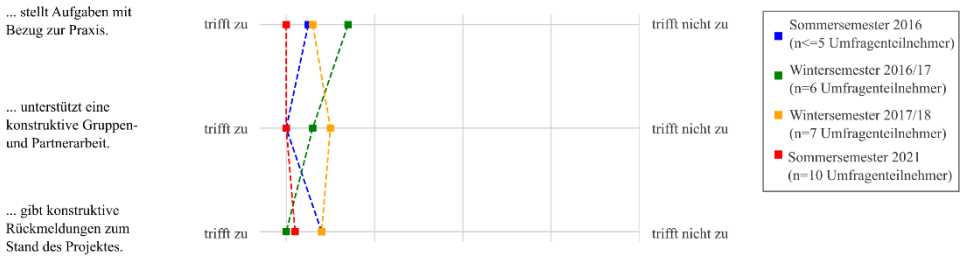
- „... sehe ich mich eher in der Lage, Projekte eigenständig zu managen.“

Abb. 2 zeigt diese Fragestellungen sowie die Bewertungen der Studierenden zu den Kriterien in ihrer zeitlichen Entwicklung durch die vier Stichproben.

Die Bewertung der *Praxisorientierung* liegt mit 1,0 im Sommersemester 2021 am höchsten. Gegenüber früheren Semestern ist hier eine moderate Verbesserung sichtbar, jedoch mit signifikanten Standardabweichungen in früheren Semestern.

Bei der Kompetenzentwicklung für die *Teamarbeit* sind die Ergebnisse im Sommersemester 2021 (Mittelwert 1,24) besser als in früheren Semestern, jedoch nicht in jedem einzelnen Bereich. Bei den Kriterien zu Stärken und Schwächen in der Teamarbeit und dem persönlichen Einsatz für eine konstruktive Teamatmosphäre gibt es hohe Abweichungen zwischen den einzelnen Bewertungen, die sich in einer signifikanten Standardabweichung widerspiegeln.

Stimmen Sie zu? Die Lehrende / der Lehrende ...



Wie beurteilen Sie Ihren Lernfortschritt?

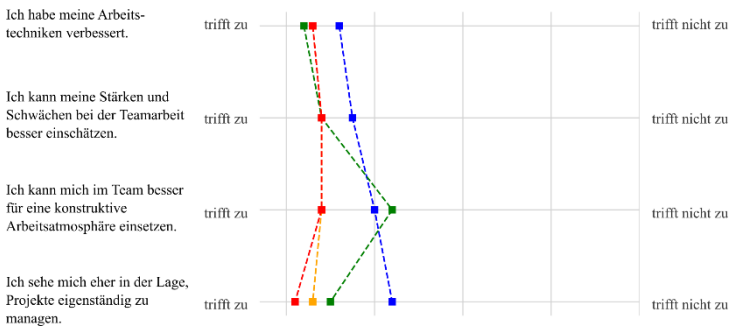


Abb. 2: Auswertung der Evaluationsbefragung unter Teilnehmenden des Softwareentwicklungsprojektes über mehrere Semester.

Im Bereich des *Projektmanagements* kann man hingegen eine klare Verbesserung konstatieren. Von einem Wert von 2,2 im SS 2016 hat sich hier die Einschätzung auf 1,1 im SS 2021 sehr deutlich und kontinuierlich verbessert.

Die subjektiven Evaluierungen der Studierenden in den diskutierten Bereichen stellen deutlich überdurchschnittliche Werte im Vergleich zum Fachbereich dar und spiegeln den Erfolg des Moduls und die Wertschätzung der Studierenden wider.

5 Zusammenfassung und Ausblick

Im Sommersemester 2023 absolvieren die Teilnehmenden bereits die 17. Runde des SWE in der dargelegten Form. Gleichzeitig hat das Projekt mit 35 Studierenden einen neuen Teilnehmerrekord erzielt. Dies resultiert in 12 Projektgruppen, die nur durch das Engagement von 2 Dozierenden sinnvoll zu betreuen sind. Damit ist das Format an eine gewisse Skalierungsgrenze hinsichtlich der Studierendenzahlen gestoßen, da sich im Laufe der Zeit herausgestellt hat, dass Lehrende höchstens 5-6 technologisch diverse Projektthemen parallel betreuen können. Da die Aufnahmezahlen des Studiengangs

Ingenieurinformatik in der Höhe von 44 Studierenden liegen, sollte das auch in Zukunft kein Problem darstellen.

Hervorzuheben sind das anhaltend große Interesse und die Unterstützung von Firmen und Forschungseinrichtungen. Über die Jahre haben sich hier Product Owner aus folgenden Unternehmen engagiert: 3YOURMIND GmbH, Airbus Defence and Space, BSS Business Solutions for Services Ost GmbH, DB Netz AG, Detecon Consulting, Deutsche Gesellschaft für Sonnenenergie e.V., eemagine Medical Imaging Solutions GmbH, Freudenberg FST GmbH, GEA Refrigeration Germany GmbH, GFal e.V., Gräbert GmbH, HTW Motorsport, InMediasP GmbH, PACE Aerospace und IT GmbH, PI Informatik GmbH, SAP AG und SENTECH GmbH. Die Autoren danken den teilnehmenden Unternehmen für ihr kontinuierliches Engagement und freuen sich darauf, die Zusammenarbeit auch in Zukunft fortzusetzen.

Das Softwareentwicklungsprojekt soll auch in Zukunft weiterentwickelt und an aktuelle Herausforderungen adaptiert werden. Zukünftige geplante Schwerpunkte umfassen:

- ein verstärkter Anwendungsschwerpunkt im Bereich Künstlicher Intelligenz,
- Einsatz und die Evaluierung von KI-gestützten Entwicklertools,
- ein verstärkter Fokus auf Entwicklungsframeworks, Tool Stacks, IDEs und
- eine optionale Unterstützung der Studierenden mit Cloud-Frameworks zum Entwickeln oder Deployen von (Teil-)Lösungen ihrer Arbeit.

Darüber hinaus planen wir, Studierende verstärkt zu ermutigen, die eigenen Lernerfolge und persönlichen Stolpersteine während des Projektes reflektieren. Dazu wollen wir in Zukunft passende Formate entwickeln, die die Studierenden bei der Standortbestimmung ihrer individuellen Stärken, Schwächen und Lernfortschritte unterstützen.

Danksagung

Unser Dank gilt allen Firmen und Forschungseinrichtungen, welche das Softwareentwicklungsprojekt kontinuierlich mit neuen Projektideen und persönlichem Engagement der *Product Owner* unterstützen. Wir danken darüber hinaus *digital.ai* für die Bereitstellung einer modernen Projektumgebung sowie einer Vielzahl an Lehrenden und Kolleg:innen, die das Projekt in der Vergangenheit durch persönliche Beiträge unterstützt haben. Abschließend danken wir allen Studierenden für ihren Einsatz, ihre Kreativität und Bereitschaft, sich jedes Semester auf das Projektabenteuer einzulassen!

Literaturverzeichnis

- [HT23] HTW Berlin Studiengang Ingenieurinformatik, Modulhandbuch Sommersemester 2023, lsf.htw-berlin.de, Stand: 03.07.2023.
- [LD01] Anderson, L. W.; Krathwohl, D. R.: A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives. Addison Wesley, 2001.
- [Zu16] Zukunft, O.: Empfehlungen für Bachelor- und Masterprogramme im Studienfach Informatik an Hochschulen (Juli 2016), 2016.

Informatik für Ingenieure an der Technischen Universität Hamburg


Bernhard J. Berger¹ und Goerschwin Fey²


Abstract: Diese Veröffentlichung stellt die Grundlagenveranstaltung *Informatik für Ingenieure* der Technischen Universität Hamburg vor, die für mehrere Ingenieurstudiengänge als Einführung in die Informatik und das Programmieren dient. Dabei ist es wichtig, in welcher Tiefe und Zusammenstellung die einzelnen Themeninhalte vermittelt werden. Die Veranstaltung befindet sich – wie viele *Hilfsveranstaltungen* – im Zielkonflikt zwischen Vermittlung zum tiefgehenden Verständnis weniger Inhalte und breit angelegtem Überblick vieler Inhalte. Somit gibt es zwei Ziele: (1) Studierende sollen mit Informatik-Fachbegriffen umgehen können, (2) Studierende sollen grundlegende Programmierkenntnisse erwerben. Der finale Leistungsnachweis mit theoretischen und praktischen Inhalten wird mithilfe einer Online-Klausur mit automatisierter Auswertung absolviert. Semesterbegleitende Tutorien und Zwischenprüfung bereiten die Studierenden auf die abschließende Klausur vor. Durch die Herausforderungen für die Veranstaltung ist diese in einem ständigen Wandel. Dieses Papier gibt einen Überblick über die Konzeption der Veranstaltung, die in der Diversität der beteiligten Studiengänge begründeten Herausforderungen und stellt eine automatisierte nicht-klausurspezifische Klausurauswertung zur Verfügung.

Keywords: Informatik, Ingenieure, Programmieren, Klausur, statische Auswertung

1 Einleitung

In vielen Ingenieurberufen ist es heutzutage selbstverständlich, ein grundlegendes Maß an Programmierkenntnissen vorauszusetzen. Vergleichbares lässt sich über ein prinzipielles Verständnis der Informatik und ihrer Teilgebiete sagen. Dies liegt in unterschiedlichen Veränderungen der letzten Jahrzehnte begründet: Zum einen ist die Etablierung des Prinzips der Industrie 4.0 [Mo22] ein entscheidender Grund, zum anderen aber auch der erhöhte Einsatz von eingebetteten Systemen, sowie das allgegenwärtige Paradigma Cyber-Physischer Systeme. Ein weiterer Aspekt ist die immer stärker verbreitete Anwendung von Machine Learning in allen Bereichen unseres alltäglichen Lebens, was somit auch diejenigen betrifft, die die Technik, die wir alltäglich nutzen, konzipieren und realisieren. Je nach Ausprägung des Ingenieurberufes ist das fundierte Wissen in unterschiedlichen Teilgebieten der Informatik von Vorteil oder sogar essenziell, um die Möglichkeiten eines Systementwurfes einzuschätzen. Diese Entwicklung hat dazu geführt, dass in den meisten ingenieurbezogenen Studiengängen mindestens ein Modul mit Informatikbezug und der Vermittlung von

¹ Technische Universität Hamburg, Institut für Eingebettete Systeme, Schwarzenberg-Campus 3, 21073 Hamburg, Germany, bernhard.berger@tuhh.de,  <https://orcid.org/0000-0001-6093-9229>

² Technische Universität Hamburg, Institut für Eingebettete Systeme, Schwarzenberg-Campus 3, 21073 Hamburg, Germany, goerschwin.fey@tuhh.de,  <https://orcid.org/0000-0001-6433-6265>

Programmierkenntnissen in den jeweiligen Studienverlaufsplänen vorhanden ist, um die Studierenden möglichst umfassend auf ihre späteren Berufe vorzubereiten.

Hier setzt auch das Modul *Informatik für Ingenieure – Einführung & Überblick* der Technischen Universität Hamburg an. Dieses Modul ist eine Pflichtveranstaltung für Studierende unterschiedlicher Studiengänge im Bereich der Ingenieurwissenschaften, die in dieser Form seit dem Wintersemester 2021/22 besteht.

Ursprünglich existierten in unterschiedlichen ingenieurbezogenen Studiengängen unterschiedliche Informatik-Veranstaltungen, welche jedoch zum Wintersemester 2021/22 zusammengefasst wurden, um einheitliche Lerninhalte der an der Technischen Universität Hamburg in Bezug auf die Informatik sicherzustellen. Die Veranstaltung in ihrer aktuellen Form ist für die Studiengänge Maschinenbau, Schiffbau, Logistik und Mobilität, Elektrotechnik, Allgemeine Ingenieurwissenschaften und General Engineering Science im Bachelorstudium verpflichtend. Da manche dieser Studiengänge an der Technischen Universität Hamburg als englischsprachige Studiengänge angeboten werden, wird das Modul *Informatik für Ingenieure* sowohl auf Deutsch als auch auf Englisch angeboten. Dies bezieht sich sowohl auf die Vorlesung als solche, als auch auf jegliches Übungsmaterial und die Klausur. Die Modulbeschreibung ist online verfügbar [Te21]. Ein optionales teilweise konsekutives Modul *Informatik für Ingenieure – Programmierkonzepte und Data Handling* wird in einigen Studiengängen genutzt, aber im Rahmen der vorliegenden Arbeit nicht diskutiert.

Das Modul besteht aus einer Vorlesung mit Übungsbetrieb. Die Vorlesung findet im Jahresturnus jedes Wintersemester statt, da sie in den meisten Studiengängen im Studienverlaufsplan als Modul für das erste Studiensemester ausgewiesen wird. Dies führt zu einer durchschnittlichen Teilnehmendenzahl von 600–700 Studierenden in jedem akademischen Jahr. Um eine gerechte und vergleichbar nachvollziehbare Bewertung der Semesterendprüfung bei dieser Größenordnung der Teilnehmendenzahl zu gewährleisten, wurde die konzeptionelle Entscheidung für eine automatisierte Auswertung getroffen. Die Klausur prüft unterschiedliche Bereiche des Modulinhalt in unterschiedlichen Komplexitätsstufen ab. Da pro Gesamtprüfung mehrere Klausuren geschrieben werden (sowohl mit unterschiedlichen Aufgaben, als auch als Wiederholung der gleichen Klausur mit unterschiedlichen Teilnehmenden), wird neben der konventionellen Auswertung der Klausuren ein statistischer Vergleich der jeweiligen Resultatverteilungen durchgeführt, um ein vergleichbares Anforderungsniveau sicherzustellen. Gleichzeitig liefert die innere Analyse der Klausurergebnisse Anregungen für mögliche Verbesserungen der Veranstaltung, sowie einen Vergleich über die Jahre hinweg (zeitliche Analyse). Die gesammelte statistische Auswertung führen wir über die Statistik-Software *R* durch und bereiten die Resultate graphisch auf. Partiiell verwenden wir diese Darstellungen auch innerhalb von Evaluationsgesprächen mit den Studierenden.

Die vorliegende Veröffentlichung umfasst die folgenden Beiträge:

1. Beschreibung der Lehr- und Lerninhalte der Veranstaltung *Informatik für Ingenieure*,

2. Beschreibung einer lehrveranstaltungsunabhängigen und quelloffenen statistischen Klausurauswertung zur besseren Evaluation und besserem Vergleich von Klausuren,
3. Identifikation von aktuellen Herausforderungen der Veranstaltung und Diskussion der Entwicklung der Veranstaltung zur Begegnung dieser Herausforderungen.

2 Organisation der Lehrveranstaltung

Die Vorlesung zum Modul *Informatik für Ingenieure* findet in jedem Wintersemester mit begleitendem Übungsbetrieb statt. Es ist aber gemäß Regelungen an der TUHH in jedem Semester möglich, eine Prüfung abzulegen. Daher wird im Sommersemester ein Repetitorium angeboten, das sich aber in der Struktur auf Grund unserer Erfahrung von der eigentlichen Vorlesung unterscheidet. Außerdem bieten wir eine eintägige Veranstaltung im Rahmen des Brückenkurses zum Studieneinstieg an.

In Rahmen der eintägigen Veranstaltung vor Beginn der Vorlesung bieten wir eine erste Einführung in das Programmieren an [Te]. Ziel des Brückenkurses ist es, bei den Teilnehmenden die Entwicklungsumgebung, aktuell verwenden wir Visual Studio Code mit Microsofts C/C++-Erweiterung und die *GNU Compiler Collection*, zu installieren und ein erstes einfaches Programm zu programmieren. Der Hauptteil des Brückenkurses besteht aus praktischer Arbeit in Kleingruppen, die durch drei plenare Zusammenkünfte eingerahmt werden. Nach einer generellen Vorstellung und einer anekdotischen Einführung in die Programmierung im Plenum werden die Teilnehmenden in jeweils von einem/einer Tutor*in betreute Gruppen aufgeteilt, um den Rest des ersten Blocks in Kleingruppen die Installation der eingangs erwähnten Software abzuschließen. Der bessere Betreuungsschlüssel erlaubt hier eine bessere Konzentration auf die individuellen Probleme der einzelnen Studierenden. Nach einer Mittagspause gibt es im Plenum erneut die Möglichkeit, aufgetretene Probleme anzusprechen und kleinere Aufgaben zu motivieren, welche dann in den Kleingruppen am Nachmittag bearbeitet werden. Den Abschluss bildet wieder eine gemeinsame Plenumsveranstaltung.

Die Veranstaltung im Wintersemester besteht aus einer Vorlesung im Umfang von 3 SWS und Tutorien im Umfang von 2 SWS in Kleingruppen. Hier ist das Ziel, dass jedes Tutorium maximal 20 Teilnehmende umfasst, was bei 600 bis 700 Teilnehmenden ungefähr 30 angestrebten Tutorien entspricht. Die Vorlesung konzentriert sich auf die theoretischen Aspekte der Veranstaltung und ergänzt diese mit praktischen Anteilen und Programmierbeispielen. Die Tutorien vertiefen auf Basis von Übungsblättern die theoretischen Aspekte nochmals, beschäftigen sich aber auch ausführlich mit den Programmierinhalten der Veranstaltung. Um die aktive Teilnahme zu fördern, gibt es semesterbegleitend drei Zwischenprüfungen, die die Teilnehmenden online ablegen können. Diese Zwischenprüfungen werden in der abschließenden Klausur als Bonuspunkte mit bis zu 10 % der erreichbaren Gesamtpunkte beim Bestehen der Klausur angerechnet. Die Zwischenprüfungen finden im gleichen System wie die eigentliche Klausur statt und bieten den Studierenden damit die Option sich mit

dem Klausursystem vertraut zu machen. Für die persönliche Vorbereitung der Studierenden stellen wir zudem zwei bis drei Wochen vor der Klausur Altklausuren online, die von den Teilnehmenden gerechnet werden können und alle zwei Tage ausgewertet werden. Teilnehmende erhalten die Ergebnisse per E-Mail in einem Bericht. Im Rahmen der Veranstaltung unterstützen wir auch empirische Studie, die untersuchen, wo für Programmieranfänger die Hürden beim Erlernen des Programmieren liegen [ETK22]. Hierdurch unterstützen wir die Forschung im Bereich der Lehrdidaktik und könne unsere eigene Veranstaltung verbessern.

Aus den Erfahrungen der letzten Klausuren heraus fokussiert sich das Repetitorium verstärkt auf den Programmiereteil der Veranstaltung. Auswertungen der letzten Jahre haben gezeigt, dass Aufgaben aus diesem Bereich im Schnitt schlechter abschneiden als Aufgaben aus den anderen Bereichen. Daher gibt es im Repetitorium keine einfache Wiederholung der Vorlesung, sondern eine Hörsaalübung im Umfang von 2 SWS und Tutorien in Kleingruppen im Umfang von 2 SWS. Im Rahmen des Repetitoriums gibt es eine fortlaufende Liste von Programmieraufgaben mit steigender Komplexität. Hierbei wird bewusst darauf verzichtet, diese den Veranstaltungswochen zuzuordnen, um allen Teilnehmenden ein individuelles Tempo zu ermöglichen. Gegen Ende des Semesters werden nochmals Altklausuren gerechnet, um nochmal die Themen der anderen Bereiche aufzufrischen. Zur Klausurvorbereitung gibt es diese nochmals im Klausursystem als Onlineklausur.

3 Lehrinhalte

Die Veranstaltung *Informatik für Ingenieure* hat vier wesentliche Komponenten im Bereich der zu vermittelnden Fachkompetenzen, die in Abbildung 1 dargestellt sind. Erstens sollen die Studierenden einen Überblick über die Grundlagen der Rechnerarchitektur erhalten. Zweitens sollen sie die (theoretischen) Grundlagen der Informatik erlernen und drittens einen Einblick die Themen der Softwaretechnik erhalten. Dies alles soll dazu dienen, die grundlegenden Programmierfähigkeiten, die sie erlernen sollen, mit einem theoretischen Verständnis zu versehen. Dabei ist zu differenzieren zwischen wissensbasierten Kompetenzen, die die Studierenden sich aneignen sollen und praktischen Kompetenzen. Die drei Basissäulen tragen hauptsächlich zum Erwerb der wissensbasierten Kompetenzen bei, während die grundlegenden Programmierfähigkeiten den Großteil der praktischen Kompetenzen ausmachen. Die folgenden Abschnitte beschreiben diese vier grundlegenden Komponenten genauer.

In der ersten Säule der Veranstaltung wird die Funktionsweise heutiger Rechner systematisch eingeführt. Zunächst werden Binärzahlen sowie Zahlen- und Zeichencodierungen eingeführt. Diese Basis wird genutzt, um mit Hilfe des Konzepts einer booleschen Algebra digitale Schaltungen einzuführen. Hierzu gehören einfache Schaltungen wie Multiplexer, Addierer und Multiplizierer. Diese Basisbausteine werden genutzt, um einen einfachen Prozessor und dessen zugehörigen Steuerungsautomaten herzuleiten. Daran kann schließlich die von-Neumann-Architektur erklärt werden. Hierbei wird das Konzept von deterministischen und nicht-deterministischen endlichen Automaten eingeführt, um den Steuerungsautomat auch

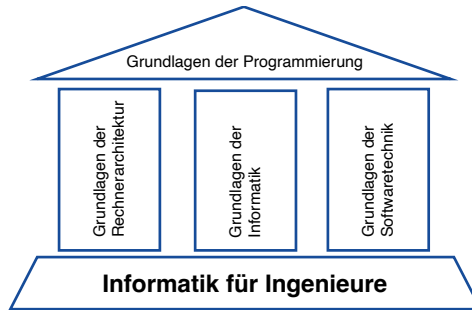


Abb. 1: Übersicht über die Modulnhalte der Veranstaltung *Informatik für Ingenieure*

auf konzeptioneller Ebene erläutern zu können. Die Grundlagen der Hardwareebene werden genutzt, um Maschinencode als unterste Ebene zur Programmierung von Rechnern einzuführen. Hierfür wird ein einfacher reduzierter Befehlssatz genutzt, um Speicherorganisation und indirekte Speicherzugriffe zu erklären.

In der zweiten Säule der Veranstaltung werden grundlegende (theoretische) Konzepte aus der Informatik eingeführt. Zu Beginn werden grundlegende Konzepte aus der Komplexitätstheorie zusammen mit der Landau-Notation vorgestellt, um den Studierenden ein Werkzeug zur Betrachtung von Algorithmen an die Hand zu geben. Diese werden anhand konkreter Beispiele in Bezug zur Rechnerarchitektur gesetzt. Anschließend werden dynamische Datenstrukturen anhand von Listen vorgestellt und deren Implementierungskonzepte präsentiert. Die einzelnen Operationen werden auf ihre Komplexität hin untersucht, um den Studierenden die Bedeutung der Wahl geeigneter Algorithmen aufzuzeigen. Auch das Sortieren von und Suchen in Feldern wird in diesem Kontext diskutiert.

In der letzten Säule wird auf die Grundlagen der objektorientierten Modellierung eingegangen. Hierbei werden die Unterschiede zwischen Daten und Operationen herausgearbeitet. Um den Studierenden ein Werkzeug zur Beschreibung objektorientierter Modellierung mitzugeben, werden in diesem Rahmen UML-Klassendiagramme eingeführt. Als wichtige Konzepte für die Modellierung werden Abstraktion, Modularisierung und Kapselung vorgestellt. Außerdem gibt es eine kurze Einführung in Softwareentwicklungsprozesse und in diesem Rahmen auch eine Einführung in die Konzepte des systematischen Testens wie Unit-Tests und Integrationstests.

Die Veranstaltung führt das Programmieren mit den Sprachen C und C++ ein, da diese die Möglichkeit bieten, zunächst mit imperativer Programmierung zu starten und die grundlegende Syntax und Semantik der Sprache C einzuführen. Außerdem sind diese Programmiersprachen im Bereich der Hardwaresteuerung noch immer weit verbreitet und für Studierende von ingenieurbezogenen Studiengängen mit hoher Wahrscheinlichkeit später relevant. Die Studierenden erhalten eine Einführung in den Umgang mit den Programmen Compiler, Debugger und Profiler, um Fehler selbständig analysieren und beheben zu

können. Zu Beginn der Veranstaltung werden zunächst die elementaren Datentypen, Arrays, Zeiger und die dynamische Speicherverwaltung eingeführt (auf die Grundlagen der zweiten Säule zurückgreifend). Später werden dann auf der Basis von C++ Klassen und die objektorientierten Konzepte aufgegriffen, was den Bogen zur dritten Säule schlägt.

In einer Klausur können Aufgaben im Allgemeinen nach drei Kriterien eingeordnet werden: Aufgabenform, Aufgabenthematik und Aufgabenniveau. Die Aufgabenthematik beschreibt hierbei den thematischen Bereich, aus dem die Aufgabe stammt. Im Rahmen der hier vorgestellten Veranstaltung werden Aufgaben thematisch zu den vier oben genannten Bereichen zugeordnet. Das Aufgabenniveau beschreibt das Anforderungsfeld der Aufgabe zwischen Reproduktion, Analyse und Transfer. In einer Klausur streben wir eine Verteilung von 60-30-10 bei einer Bestehensgrenze von 50% an. Die Aufgabenform beschreibt die Art der Aufgabenstellung. Durch die elektronische Form der Klausur und die automatisierte Auswertung bieten sich nur bedingt viele Möglichkeiten der Aufgabenstellung an. Im Rahmen dieser Veranstaltung wurden bisher Rechenaufgaben, Multiple-Choice Aufgaben, Drag-and-Drop Aufgaben und Programmieraufgaben verwendet. Hierfür verwenden wir das YAPS Prüfungssystem [BF21], ein erweiterbares digitales Prüfungssystem, das die genannten Aufgabentypen im Web-Browser ermöglicht. Für die Programmieraufgaben wird hierfür die verfügbare Serverinfrastruktur verwendet, um die Lösung der Studierenden zu bauen und auszuführen. Die Bearbeitungszeit der Klausur beträgt 90 Minuten. Die Klausur wird auf mehrere Durchführungen aufgeteilt, da die Kapazität, um die notwendige Anzahl an Klausuren gleichzeitig zu schreiben, nicht vorhanden ist. Das Klausurergebnis erhalten Teilnehmende in einem übersichtlichen Bericht per Email über den die Bewertung übersichtlich transparent gemacht wird. Eine Klausureinsicht vor Ort entfällt dementsprechend, außerdem wird die Schwelle für die Klausureinsicht drastisch gesenkt. Bei etwaigen Einsprüchen werden gegebenenfalls alle Klausurbewertungen angepasst, somit werden gleiche Bedingungen für alle Studierenden erreicht.

Durch den Schwerpunkt der Veranstaltung auf der Kompetenz Programmieren umfassen Programmieraufgaben stets 50% der erreichbaren Klausurpunkte (allerdings in unterschiedlichen Anforderungsniveaus). Die verbleibenden 50% der erreichbaren Klausurpunkte verteilen sich auf die anderen inhaltlichen Themen. Bei diesen Aufgaben werden unterschiedliche Aufgabentypen verwendet und es ist das Ziel, den Anteil der Multiple-Choice Aufgaben möglichst gering zu halten. Hierfür eignen sich Aufgaben, wie zum Beispiel das Umrechnen zwischen Zahlensystemen und das Ausführen eines Moore- oder Mealey-Automaten als Aufgaben mit dynamisch generierten Werten pro Studierendem. Mit Drag-und-Drop-Aufgaben stellen wir Aufgaben zu Besuchsreihenfolgen in Graphen oder lassen aus einer beschränkten Anzahl an vorgegebenen Code-Snippets in der Vorlesung vorgestellte Algorithmen zusammenbauen. Als letztes gibt es in dieser Kategorie noch Multiple-Choice Aufgaben, um Fragen im reproduktiven Aufgabenbereich zu stellen. Der Anteil der reinen Multiple-Choice Fragen ist pro Klausur auf circa 25% der erreichbaren Gesamtpunkte beschränkt. In der Klausur gibt es zwei Programmieraufgaben mit unterschiedlichem Schwierigkeitsgrad, die aktuell zu gleichem Teil die Gesamtpunkte des Programmiererteils ergeben. Die erste

Programmieraufgabe beschränkt sich auf imperative Programmierkonzepte, also den Umfang der Sprache C einschließlich dynamischer Speicherverwaltung. In dieser Aufgabe sind mehrere Funktionsrümpfe vorgegeben, die die Studierenden zu füllen haben, aber die einen sehr geringen Anteil an selbst zu entwickelnder Algorithmik enthalten. Hier müssen zum Beispiel Arrays auf dem Heap angelegt und initialisiert, Werte in Arrays quadriert oder die Summe aller Elemente eines Arrays berechnet werden. Für eine Funktion gibt es in dieser Aufgabe einen vorgegebenen natürlichsprachigen Pseudo-Code, den die Studierenden in C-Code übersetzen müssen. Daher fallen die Funktionen dieser Aufgabe in die Reproduktions- bzw. Analyse-kategorie. In der zweiten Programmieraufgabe nutzen wir zusätzlich noch Klassen und die STL von C++. Hierbei erhalten die Studierenden aber eine Dokumentation aller eingesetzten STL-Typen. Diese Aufgabe fällt in die Analyse- und Transferkategorie, da sie deutlich mehr selbstständiges Denken erfordert, um diese erfolgreich lösen zu können.

4 Statistische Auswertung der Klausuren

Im Folgenden wird die statistische Auswertung der Klausuren beschrieben und ihre allgemein anwendbare Struktur erläutert, sowie die Interpretation ihrer Visualisierungen dargelegt. Bei der Auswertung einer Klausur sind prinzipiell drei unterschiedliche Bereiche von Interesse: (1) Wie ist ein einzelner Durchlauf einer einzelnen Klausur zu bewerten? (2) Wie ist ein gesamter Durchlauf eines Semesters mit potentiell mehreren Klausuren und Wiederholungen einzelner Klausuren zu bewerten? (3) Wie ist die Entwicklung der Klausurergebnisse im zeitlichen Kontext zu bewerten? Die einzelnen Fragen werden in den jeweiligen Abschnitten diskutiert.

4.1 Auswertung eines einzelnen Klausurdurchlaufs

Hier wird das Ergebnis einer zeitgleich geschriebenen Klausur zwischen allen Teilnehmenden analysiert.

Ein schlichter Ausgangspunkt für die Analyse ist ein Histogramm [Wa06] oder Boxplot [Tu77] der Punkteverteilung. Ob die Punkte in diesem Fall relativ oder absolut dargestellt werden, ist nicht von Bedeutung. Ein Boxplot visualisiert gleichzeitig den Median (das Lagemaß), Minimum und Maximum (Spannweite) und unteres/oberes Quartil (Inter-Quartils-Abstand). Es ist möglich, auch noch den Mittelwert mit einzuzeichnen. Die Analyse kann für die einzelnen Aufgaben wiederholt werden, sodass aufgabenspezifische Effekte in aufgabenspezifischen Histogrammen und Boxplots identifiziert werden können.

Um die Güte bzw. Eignung einzelner Aufgaben in Bezug auf das Gesamtklausurergebnis zu beschreiben, können die einzelnen Klausurergebnisse entsprechend ihrer Gesamtpunktzahl gruppiert werden und ihr Resultat innerhalb einer spezifischen Aufgabe als Boxplot

repräsentiert werden. Eine Aufgabe, die exakt die Ergebnisse der Gesamtklausur repräsentiert, erhielte so eine Boxplotkurve, die der kumulativen Normalverteilung entspricht. Aufgaben, die durchschnittlich schlechtere Ergebnisse lieferten, haben eine nach rechts verschobene Kurve, Aufgaben, die durchschnittlich bessere Ergebnisse liefern, haben sie eine nach links verschobene Kurve. Abbildung 2 zeigt zwei Aufgaben in der erläuterten Darstellung.

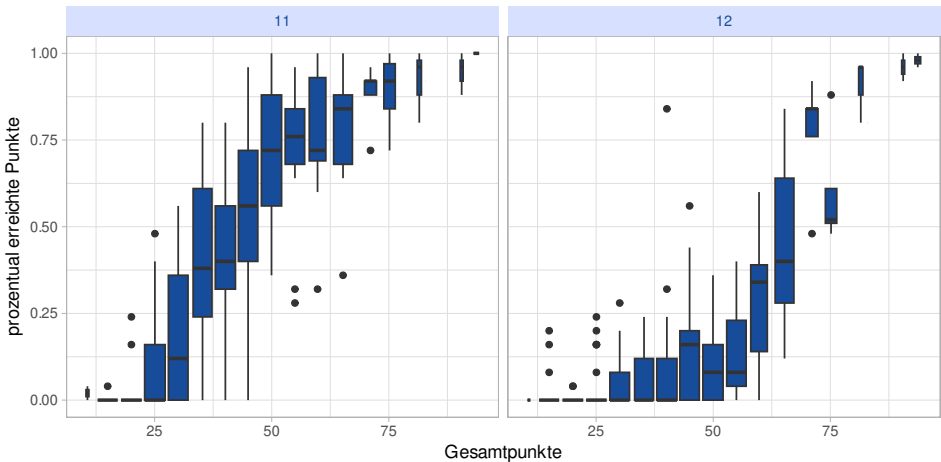


Abb. 2: Zwei Beispielprogrammieraufgaben als Box-Plots, die die Gesamtpunkte in der Klausur gegen die Punkte der jeweiligen Aufgabe abbildet. Die erste Aufgabe stammt aus dem Reproduktions- und Analysekatégorie. Die zweite Aufgabe ist die schwierigere Programmieraufgabe aus der Analyse- und Transferkatégorie. Somit entsprechen die Darstellung im wesentlichen den Erwartungen.

Die Trennschärfe [Fi04] einer Aufgabe bestimmt die Fähigkeit dieser Aufgabe, das Gesamtklausurergebnis vorherzusagen. Eine hohe Trennschärfe zeigt an, dass das Ergebnis in dieser Aufgabe das Gesamtergebnis gut vorhersagt, eine niedrige Trennschärfe sagt, dass diese Aufgabe kaum Einfluss auf das Gesamtergebnis hat. Eine negative Trennschärfe bedeutet, dass gute Ergebnisse bei dieser Aufgabe auf ein schlechtes Gesamtergebnis hindeuten, was eigentlich eher unerwartet ist.

Zunächst werden die Resultate für themen- und niveaugleiche Aufgaben addiert, falls mehrere vorhanden sind. Diese kumulativen Ergebnisse werden dann für die entsprechenden Themen- und Niveaustufen als Boxplots (oder Histogramme) dargestellt. Mit dieser Grafik kann zum Einen untersucht werden, ob unterschiedliche Themen unterschiedlich gut verstanden wurden, und zum Anderen, ob die Einordnung der Aufgaben hinsichtlich ihres Niveaus mit den Resultaten der Studierenden in der Klausur übereinstimmt.

In dieser Veranstaltung werden innerhalb des Semesters Testate geschrieben, in denen es den Studierenden möglich ist, Bonuspunkte für ihre finale Modulnote zu sammeln. Ein Punktdiagramm zwischen erreichten Bonuspunkten und erreichten Punkten in der Klausur kann Aufschluss darüber geben, ob eine (erfolgreiche) Teilnahme (oder Nicht-Teilnahme) an

den Testaten, die nicht verpflichtend sind, einen Einfluss auf die Leistung in der Klausur hat. Damit wird die Korrelation zwischen erzielter Bonuspunkte und Klausurpunkte dargestellt.

4.2 Auswertung eines Gesamtdurchlaufs

Für den Gesamtüberblick über alle aktiven Läufe wiederholen wir alle obigen Diagramme zur Gesamtlage (also Boxplot und Histogramme) mit den Gesamtdaten. Um Unterschiede zu visualisieren, werden die unterschiedlichen Klausurdurchläufe mit unterschiedlichen Farben symbolisiert.

Um einen Vergleich der Resultate von Klausurwiederholungen sichtbar zu machen vergleichen wir die Verteilungen – entweder nicht-parametrisch (wenn die Voraussetzungen nicht erfüllt sind) oder parametrisch. Dieser kann sowohl auf Aufgabenebene als auch auf Klausurebene durchgeführt werden.

Für den Vergleich zwischen unterschiedlichen Klausurvarianten verwenden wir einen nicht-parametrische Test. Um diese auf Aufgabenebene durchführen zu können, ist allerdings die Klassifizierung anhand der oben genannten Kriterien wichtig, da sich die Aufgaben in den unterschiedlichen Klausurvarianten unterscheiden.

4.3 Auswertung der Entwicklung der Veranstaltung

Eine einfache Möglichkeit, die Entwicklung der Klausurergebnisse (bei gleichen Verteilungen der Aufgaben in Bezug auf die oben genannten Kriterien) über die Zeit zu vergleichen, sind die Boxplots über die Gesamtergebnisse eines Durchlaufs für die einzelnen Jahre darzustellen. Anhand der Visualisierung kann dann beobachtet werden, ob die Streuung oder Lage sich über die Zeit verändert.

4.4 Implementierung

Der statistische Auswertung der Klausur liegt eine R-Implementierung zugrunde. R ist eine Programmiersprache, die sich auf statistische Analyse von Daten fokussiert [Da15]. Die Auswertung ist unabhängig von der Veranstaltung *Informatik für Ingenieure* und kann daher für die Auswertung beliebiger Klausuren verwendet werden. Die Daten müssen in einem einfachen CSV-Format vorliegen, sodass das Skript automatisch die beschriebenen Informationen berechnen kann. Dieses Skript ist öffentlich zugänglich [BF].

5 Herausforderungen

Lehre ist stets mit Herausforderungen verbunden. Interdisziplinäre Lehre der unterstützenden Form ist allerdings neben den *normalen* Herausforderungen noch mit zusätzlichen Herausforderungen konfrontiert.

Eine Herausforderung besteht in der Konzeption der Veranstaltung, die durch die Modulstruktur vorgegeben ist: Sowohl die theoretischen Grundlagen als auch Programmierfähigkeiten innerhalb eines Semesters zu vermitteln, ist didaktisch fordernd. Obwohl die theoretischen Grundlagen aus Sicht eines erfahrenen Programmierers wichtig sind für die Entwicklung ausreichender Programmierkompetenzen, ist dies für Programmieranfänger oft nicht in Gänze nachzuvollziehen. So entsteht für manche Studierende eine Disparität zwischen den Lerninhalten, die ihren Lernerfolg erschweren.

Der Zeitpunkt der Veranstaltung im Studienverlaufsplan (erstes Semester) ist eine weitere Herausforderung. Zu diesem Zeitpunkt wirken die Lerninhalte für Studierende noch nicht mit ihren späteren beruflichen Aufgaben motiviert. Sie beginnen erst, sich mit ihrer Fachrichtung auseinander zu setzen und sehen noch nicht die Aufgaben, in denen grundlegende informatische Kenntnisse von Vorteil sind und Programmierfähigkeiten notwendig. Dies ist sowohl problematisch im Hinblick auf die Motivation der Studierenden als auch auf die Anwerbung von studentischen Tutor*innen. Der Bedarf an studentischen Tutor*innen ist in der Veranstaltung hoch und es ist nicht immer möglich, die benötigte Menge an Tutor*innen zu finden. Da die Technische Universität Hamburg auch den Studiengang *Computer Science* anbietet, suchen wir auch hier aktiv nach möglichen Tutor*innen.

Ein weiterer Aspekt ist die mangelnde Akzeptanz der Wichtigkeit der vollständigen Bearbeitung von Übungsaufgaben. Wie viele Fächer reicht auch (insbesondere im Programmieren) in dieser Veranstaltung ein rein theoretisches Verständnis nicht aus. Um eine ausreichende Kompetenzstufe im Bereich des Programmierens zu erlangen, ist konstantes und ausdauerndes Üben an unterschiedlichen Programmieraufgaben von Nöten. Daher beinhaltet diese Veranstaltung umfangreiches Übungsmaterial und legt Wert auf Präsenzübungen in den Tutorien. Nichtsdestotrotz führt die Tatsache, dass die Bearbeitung der Übungszettel nicht verpflichtend ist, sondern nur als Angebot vorliegt, zu einer geringeren Anzahl an Studierenden, die diese vollständig bearbeiten und auch bei Herausforderungen sich diesen stellen. Ein größerer Anteil an Studierenden sieht die Übungsaufgaben als optional und erreicht so nicht den nötigen Erfahrungsschatz, der für einen erfolgreichen Erwerb von Kompetenzen im Bereich der Programmierung notwendig wäre.

Es ist durchgängig ein Rückgang an bereits vorhandenen Kompetenzen bei den Studierenden zu beobachten, sowie eine hohe Streuung, was ihre Vorkenntnisse betrifft. Für einen größeren Anteil an Studierenden ist beispielsweise die Formulierung der Übungsaufgaben oder partiell Klausuraufgaben auf einem Niveau, das nicht ihrer LeseverständnisKompetenz entspricht. Ähnliches gilt für grundlegende mathematische Kompetenzen, wie Rechenregeln oder logisches Denken. Gleichzeitig gibt es jedoch einen kleinen Anteil an Studierenden,

die enorme Vorkenntnisse und Kompetenzen mitbringen, die durch eine Unterforderung im Kurs demotiviert werden. Im universitären Kontext ist jedoch eine binnendifferenzierte Leistungserbringung bisher noch nicht vorgesehen, sowie eine binnendifferenzierte Lehrstruktur (wie sie in Schulen umgesetzt wird) in ihrer Umsetzung sehr herausfordernd. In Bezug auf Programmierkenntnisse wird dies partiell in den Übungsgruppen umgesetzt. So gibt es explizit Übungsgruppen für Studierende mit Vorkenntnissen und Übungsgruppen für Studierende ohne Vorkenntnisse. Ein Nachteil dieser Lösung ist allerdings die dadurch mangelnde Durchmischung unterschiedlicher Gruppen, ein Lernen durch Lehren (in Gruppenstrukturen bspw.) ist so nicht mehr möglich.

Eine weitere, eher unerwartete Herausforderung ist das fehlende Wissen im grundlegenden Umgang mit Computern. Grundlegend ist hier die Beobachtung, dass moderne Betriebssysteme mittels App-Stores Installationsvorgänge weitgehend vereinfachen und die entsprechenden Apps den tatsächlichen Speicherort von Daten nicht mehr offensichtlich für Nutzende darstellen. Während die Studierenden sich mit der Nutzung von Standardanwendungen sehr gut auskennen, wird es bei Programmen, die manuell installiert werden müssen, wie einem Compiler deutlich schwieriger. Vielen Studierenden ist hierbei zum Beispiel unklar, wo Programme installiert werden, wie installierte Dateien auf dem System gefunden werden können. Ebenso ist der Umgang mit Archiven, die entpackt werden müssen, oder das Auffinden eigener gespeicherter Daten eine Herausforderung.

6 Ausblick

Konstant progressierende Anforderung an die Veranstaltung sowie sich diesen Änderungen anpassende Lehrkonzepte seitens der Veranstalter führen zu einem steten Wandel in der Veranstaltungsform. Eine standardisierte statistische Auswertung soll bei der Evaluierung dieser Änderungen unterstützen und wird im Rahmen dieser Publikation auch für andere Veranstaltungen zur Verfügung gestellt.

Die Veranstaltung lehrt die Grundlagen der Rechnerarchitektur, (theoretischen) Informatik und Softwaretechnik, um den Studierenden das Grundgerüst für den Erwerb von Programmierkompetenzen zu geben, der den Schwerpunkt der Veranstaltung ausmacht. Die richtige Balance zwischen beiden Aspekten zu finden, ist eine Herausforderung, der sich entsprechend der Rückmeldung der Studierenden von Jahr zu Jahr durch neue Konzepte in der zeitlichen Abfolge gestellt wird.

Neben diesen inhaltlichen Herausforderungen sind jedoch auch Unterschiede in der Vorbildung und ein Rückgang im grundlegenden technischen (rechnerbasierten) Verständnis eine Herausforderung. Durch den großen Anteil der Zeit, der in Tutorien durch Vorarbeiten wie Installationsprozesse gebunden wird, ist eine geplante Änderung für den nächsten Durchlauf die Durchführung einer Lerneinheit diese grundlegenden rechnerbasierten Kenntnisse (über z.B. Betriebssysteme) aufgreift.

Unserer Meinung nach ist eine ganzheitliche, grundlegende Bildung im informatischen Sinne für die Ingenieur*innen im 21. Jahrhundert von essentieller Bedeutung. Daher die Weiterentwicklung unseres Lehr- und Lernkonzeptes von dem Grundgedanken geprägt, die thematische Breite in jedem Fall beizubehalten, jedoch durch Zusatzveranstaltungen oder binnendifferenzierte Übungsaufgaben zu Beginn der Veranstaltung die unterschiedlichen Vorbildungsstände der Studierenden aufzufangen, um ihnen einen bildungsgerechten Start in ihr Studium zu ermöglichen sowie eine erfolgreiche berufliche Laufbahn.

Literaturverzeichnis

- [BF] Auswertungsskript für Klausuren. <https://collaborating.tuhh.de/es/ce/public/exam-statistics>, Stand: 31.08.2023.
- [BF21] Bahnsen, F. H.; Fey, G.: YAPS - Your Open Examination System for Activating and emPowering Students. In: 2021 16th International Conference on Computer Science & Education (ICCSE). IEEE, 2021.
- [Da15] Daroczi, G.: Mastering Data Analysis with R. Packt Publishing, 2015.
- [ETK22] Eckert, D.; Timmermann, D.; Kautz, C.: Student Misconceptions about Loops in Introductory Programming Courses and the Influence of Representations. In: 2022 IEEE Frontiers in Education Conference (FIE). IEEE, 2022.
- [Fi04] Fisseni, H. J.: Lehrbuch der psychologischen Diagnostik. Hogrefe Verlag, Göttingen, Germany, 3. Auflage, 2004.
- [Mo22] Mohl, E.: Veränderte Kompetenzbereiche durch Industrie 4.0 in der Ausbildung von Techniker*innen an Höheren Technischen Lehranstalten in Österreich. Springer Fachmedien Wiesbaden, 2022.
- [Te] Homepage Brückenkurs. <https://www.tuhh.de/rzt/beratung/brueckenkurse/programmieren>, Stand: 31.08.2023.
- [Te21] Technische Universität Hamburg: Modulhandbuch – Bachelor of Science (B.Sc.) – Maschinenbau. Bericht, Technische Universität Hamburg, 2021.
- [Tu77] Tukey, J. W.: Exploratory Data Analysis. Addison-Wesley series in behavioral science. quantitative methods. Pearson, Upper Saddle River, NJ, 1977.
- [Wa06] Wasserman, L.: All of Nonparametric Statistics. Springer New York, 2006.

Interdisziplinärer Industrie 4.0 Demonstrator (II4D)

Carsten Burmeister¹, Bonke Möller², Dominik Miller³, Andreas Haase⁴ und Hendrik Lam⁵

Abstract: Das Thema Industrie 4.0 ist an der Fachhochschule Wedel in fast allen Studiengängen vertreten, wie zum Beispiel im IT-Ingenieurwesen, in der Informatik, im Wirtschaftsingenieurwesen oder der Betriebswirtschaftslehre. Dabei konzentrieren sich die Studiengänge auf unterschiedliche Bereiche des großen Feldes Industrie 4.0. Wir erwarten, dass sich für Studierende ein tieferes Verständnis und eine höhere Lernmotivation ergibt, wenn die bisher häufig getrennten Industrie 4.0-Bereiche zusammengeführt werden und zusätzlich durch interdisziplinäre, praktische Übungen und Erfahrungen ergänzt werden können. Im Rahmen einer Projektförderung der Stiftung für Innovation in der Hochschullehre haben wir die Möglichkeit, diesen Ansatz umzusetzen, indem wir einen Industrie 4.0 Demonstrator und didaktische Konzepte zur interdisziplinären Nutzung entwickeln. In diesem Artikel erläutern wir unsere Motivation zur Entwicklung des Demonstrators. Wir beschreiben das zugrundeliegende Konzept und unsere Erwartungen. Außerdem zeigen wir den aktuellen Stand des Projekts und erste Erfahrungen.

Keywords: Industrie 4.0, Demonstrator, Interdisziplinarität, Hochschullehre

1 Einleitung

Die vierte industrielle Revolution, auch bekannt als Industrie 4.0, hat weitreichende Auswirkungen auf verschiedenste Berufsfelder und verändert grundlegend die Art und Weise, wie wir produzieren und arbeiten [BTV14]. Sie präsentiert eine Fülle von neuen Konzepten und Technologien, die die Grenzen zwischen physischen und digitalen Systemen verwischen. Infolgedessen sehen wir eine signifikante Bedeutung von Industrie 4.0 für Studierende verschiedenster Disziplinen.

An der FH Wedel integrieren wir Industrie 4.0 Inhalte in eine Vielzahl von Studiengängen, darunter IT-Ingenieurwesen, Informatik, Wirtschaftsingenieurwesen und Betriebswirtschaftslehre (BWL). Dabei wird der Fokus auf unterschiedliche Aspekte der Industrie 4.0 gelegt, abhängig vom jeweiligen Studiengang. Grundlegende Technologien der Automatisierungstechnik, Themen des Industrial Internet of Things (IIoT), Konzepte wie Predictive Maintenance im Ingenieurwesen, IT-Sicherheit, Big Data, Cloud- und Edge-Computing in der Informatik, sowie neue Geschäftsmodelle in der BWL sind nur einige der Bereiche, die vertieft behandelt werden.

¹ FH Wedel, Fachbereich Technik, Feldstraße 143, 22880 Wedel, carsten.burmeister@fh-wedel.de

² FH Wedel, Fachbereich Technik, Feldstraße 143, 22880 Wedel, bonke.moeller@fh-wedel.de

³ FH Wedel, Fachbereich Technik, Feldstraße 143, 22880 Wedel, dominik.miller@fh-wedel.de

⁴ FH Wedel, Fachbereich Technik, Feldstraße 143, 22880 Wedel, andreas.haase@fh-wedel.de

⁵ FH Wedel, Fachbereich Technik, Feldstraße 143, 22880 Wedel, hendrik.glowatzki@fh-wedel.de

Diese Heterogenität von Industrie 4.0 zeigt sich nicht nur in unseren Studiengängen, sondern auch in der Literatur und der aktuellen Forschung. Verschiedene Definitionen von Industrie 4.0 heben unterschiedliche Aspekte dieses umfangreichen Feldes hervor. So definiert das Netzwerk *Plattform Industrie 4.0* des Bundesministeriums für Wirtschaft und Klimaschutz: „Industrie 4.0 bezeichnet die intelligente Vernetzung von Maschinen und Abläufen in der Industrie mit Hilfe von Informations- und Kommunikationstechnologie.“ [BM23]. Götz und Jankowska betonen in ihrer Definition hingegen auch die wirtschaftlichen Einflüsse: „It indicates the profound transformation of business models by enabling the fusion of virtual and real worlds and the application of digitization, automation, and robotics in manufacturing.“ [GJ17].

Doch trotz der theoretischen Breite in der Lehre besteht bei Studierenden oft eine Zurückhaltung, praxisnahe Abschlussarbeiten zu wählen, die eine hohe Innovationshöhe aufweisen. Es scheint, dass das Studium, obwohl es viele Teilbereiche von Industrie 4.0 abdeckt, keinen ausreichend praxisnahen Blick auf das Gesamtphänomen Industrie 4.0 nicht vermittelt.

In diesem Paper beschreiben wir unsere Bemühungen, diese Lücke zu schließen und sowohl eine praxisnähere Ausbildung im Bereich der Industrie 4.0 als auch einen interdisziplinären Blick auf das gesamte Themengebiet zu bieten. Unser Hauptinstrument ist ein Industrie 4.0 Demonstrator, der es den Studierenden ermöglicht, Industrie 4.0 in Gänze auf einer sehr praktischen Ebene zu erleben. Nach einer gründlichen Evaluierung kommerzieller Demonstratoren stellten wir fest, dass diese für unsere Zwecke ungeeignet waren: Sie waren zu starr und wenig flexibel, konzentrierten sich häufig auf isolierte Bereiche von Industrie 4.0 und ihre Erweiterbarkeit war begrenzt.

Daher entwickeln wir in einem Industrie 4.0 OpenLab mit dem Interdisziplinären Industrie 4.0 Demonstrator (II4D), ein innovatives, flexibles und erweiterbares System, das verschiedene Aspekte der Industrie 4.0 abdeckt. Dieses Projekt wird durch die Stiftung Innovation in der Hochschullehre gefördert.

2 Anforderungen an einen Interdisziplinären Industrie 4.0 Demonstrator

Für die Entwicklung eines effektiven Interdisziplinären Industrie 4.0 Demonstrators (II4D) sind eine Reihe von Anforderungen zu berücksichtigen. Diese beziehen sich auf verschiedene Aspekte, einschließlich der technischen Fähigkeiten, der Benutzerfreundlichkeit, der Flexibilität und der Erweiterbarkeit des Demonstrators.

2.1 Technische Anforderungen

Der II4D soll eine breite Palette von Industrie 4.0 Technologien und Konzepten abbilden. Dazu gehören unter anderem moderne Kommunikationstechnologien und Sensorik,

Automatisierung, Maschinelles Lernen (ML), Cloud- und Edge-Computing und Predictive Maintenance. Der Demonstrator soll in der Lage sein, sowohl die Zusammenhänge zwischen diesen Technologien als auch deren praktische Anwendung zu veranschaulichen. So sollen die Studierenden die Möglichkeiten bekommen, sowohl an technischen Details zu forschen und zu entwickeln als auch Konzepte umzusetzen und deren Wirkungsweise im Gesamtsystem kennenzulernen.

2.2 Benutzerfreundlichkeit

Da der II4D als Lehrmittel für Studierende verschiedener Studiengänge dient, muss er besonders benutzerfreundlich gestaltet sein. Eine intuitive Benutzeroberfläche auf einer hohen Abstraktionsebene und leicht verständliche Dokumentationen sind entscheidend, um den Studierenden den Zugang zum Demonstrator und den darin abgebildeten Technologien zu erleichtern. Auf der anderen Seite müssen auch tiefer liegende Schnittstellen für Studierende technischer Studiengänge leicht zugänglich und gut dokumentiert sein.

2.3 Flexibilität

Ein interdisziplinärer Demonstrator muss flexibel genug sein, um eine Vielzahl von Einsatzszenarien und Experimente zu unterstützen. Die Studierenden sollten in der Lage sein, verschiedene Parameter zu ändern und die Auswirkungen dieser Änderungen auf das System zu beobachten. Darüber hinaus soll der Demonstrator in der Lage sein, sich an neue Technologien und Trends anzupassen, die im Rahmen von Industrie 4.0 auftauchen. Auch sollen verschiedene Arten von Produktionsorganisationsformen konfigurierbar sein.

2.4 Erweiterbarkeit

In Anbetracht der dynamischen Natur von Industrie 4.0 soll der Demonstrator so konzipiert sein, dass er erweitert und angepasst werden kann, um neuen Entwicklungen Rechnung zu tragen. Dies umfasst nicht nur technologische Erweiterungen, sondern auch die Möglichkeit, den Demonstrator an verschiedene Lernziele und pädagogische Kontexte anzupassen.

Die oben genannten Anforderungen sind entscheidend für die Entwicklung eines II4D, der das breite Spektrum von Industrie 4.0 abdeckt und den Studierenden einen praxisnahen Einblick in dieses wichtige Gebiet bietet. In den folgenden Abschnitten werden wir diskutieren, wie wir diese Anforderungen in der Konzeption und Entwicklung unseres II4D berücksichtigt haben.

3 Konzept

3.1 Pädagogisches Konzept

Unser Hauptziel beim Entwurf des II4D ist es, Studierenden eine praxisnahe Erfahrung mit den Technologien und Prozessen von Industrie 4.0 zu ermöglichen. Hierbei legen wir besonderen Wert auf die Nutzung der vorhandenen Infrastruktur, darunter CNC-Fräsen, 3D-Drucker und Platinenfertigungs- und -bestückungseinrichtungen. Durch diese Strategie können wir Studierenden zeigen, dass mit vorhandenen Mitteln komplexe Aufgabenstellungen gelöst werden können. Eines unserer pädagogischen Ziele besteht darin, den Studierenden die Scheu vor praktischen Arbeiten zu nehmen und ihre Motivation für eigene Projekte zu stärken.

3.2 Technisches Konzept

Das technische Grundkonzept des II4D besteht in der Nachbildung einer stationsbasierten Fließfertigung. Ein Werkstück wird mithilfe eines Transportroboters durch die verschiedenen Fertigungsstationen geführt. Als beispielhafte Anwendung kann der Prozess der Zubereitung eines Tees dienen, wobei der Prozess über mehrere Fertigungsstationen verteilt ist.

Die Stationen können frei im Raum verteilt werden. In einer ersten Ausbaustufe folgt der Roboter vorgegebenen Linien. In einer geplanten Erweiterung soll sich der Roboter mithilfe von Ultra-Wideband-Technik (UWB) im Raum orientieren, wie z.B. in [FSBP21] beschrieben. Jeder Fertigungsschritt kann von mehreren Stationen durchgeführt werden, was die Nutzung verschiedener Load-Balancing-Methoden ermöglicht.

Die Kommunikation auf Protokollebene erfolgt über das Message Queuing Telemetry Transport (MQTT) Protokoll. Als Kommunikationstechnologie wird zunächst WLAN genutzt. Geplant sind Erweiterungen für Feldbusse zur Kommunikation mit speicherprogrammierbaren Steuerungen (SPS), sowie Zigbee, LoRa und UWB als weitere Funkstandards.

Eine Integration von Cloud-Diensten wie Amazon Web Services (AWS), Google Cloud, Azure und Siemens Mindsphere ist ebenfalls geplant. Vorerst wird jedoch nur Node-RED implementiert um z.B. ein Dashboard zu realisieren.

Der II4D nutzt fast ausschließlich ESP32-Mikrocontroller und Raspberry Pi-Computer, da diese in unseren Lehrveranstaltungen bereits behandelt werden. Diese Wahl erleichtert den Studierenden den Einstieg in die Erweiterung des Systems.

Um sicherzustellen, dass die Studierenden auch Erfahrung mit klassischer Automatisierungstechnik sammeln, soll in einem späteren Schritt auch eine SPS integriert werden.

4 Stand der Umsetzung

Die Arbeit am Interdisziplinären Industrie 4.0 Demonstrator (II4D) im Rahmen des von der Stiftung Innovation in der Hochschullehre geförderten Projekts hat Ende 2022 begonnen. Mehrere Schlüsselkomponenten des Systems konnten bereits vollständig oder teilweise implementiert werden. Wir konnten bei allen Entwicklungen das pädagogische Konzept der In-house Fertigung nutzen. Zum Beispiel wurden alle entwickelten Platinen auch an der FH-Wedel gefräst und bestückt. Wir achten während der Entwicklung bereits darauf, standardisierte Prozesse zu nutzen und die Entwicklungen digital zu dokumentieren. In der folgenden Abb. 1 ist ein Bild des 3D Modells des Transportroboters gezeigt.

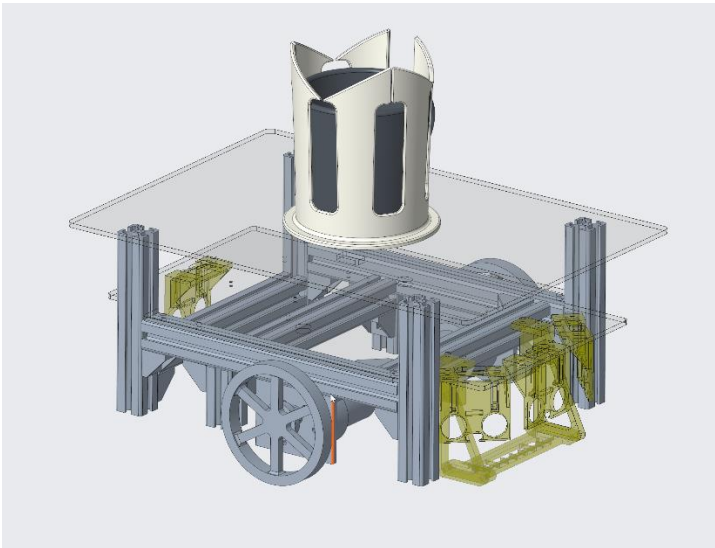


Abb.1: 3D Modell des Transportroboters

4.1 Transportroboter

Der Transportroboter, ein zentraler Bestandteil des II4D, ist vollständig umgesetzt. Der Roboter nutzt einen ESP32-Mikrocontroller und einen Raspberry Pi. Für die Motorsteuerung wurde eine Platine entwickelt, die bereits die Möglichkeit zur Integration von UWB-Chips bietet.

Der Roboter ist mit Abstandssensoren zur Kollisionsvermeidung und Infrarotsensoren zur Linienverfolgung ausgestattet. Ein Modul auf dem Raspberry Pi verwaltet aktuelle Aufträge und plant Routen, während der ESP den Antrieb steuert, Linien folgt und in vorgegebene Richtungen abbiegt.

Die Software des Roboters ist modular aufgebaut, sodass Studierende in kleinen Projekten Teilaspekte des Roboter-Verhaltens ändern können. Darüber hinaus bietet die Hardware Erweiterungsmöglichkeiten, was die Anpassungsfähigkeit und zukunftssichere Gestaltung des Roboters gewährleistet.

4.2 Stationen

Die geplanten Fertigungsstationen sind modular konzipiert. Das Basismodul ist mechanisch und elektrisch bereits als Prototyp umgesetzt. Die aufgabenspezifische Funktionalität jeder Station steht als nächster Schritt in der Umsetzung an.

4.3 Zentraler Steuerungscomputer

Der zentrale Steuerungscomputer, auf dem ein MQTT-Broker und eine einfache Auftragsverwaltung laufen, ist bereits in Betrieb. Zukünftige geplante Erweiterungen beinhalten eine grafische Benutzeroberfläche, einen einfachen Web-Zugriff und eine Dashboard-Anwendung. Es wird erwartet, dass diese Elemente in zukünftigen Studierendenprojekten entwickelt und implementiert werden.

5 Beispielhafte Integration in die Lehre

Die Einbindung des Interdisziplinären Industrie 4.0 Demonstrators (II4D) in das Curriculum der Hochschule bietet Studierenden eine einzigartige Gelegenheit, ihre theoretischen Kenntnisse in die Praxis umzusetzen und damit zu vertiefen. Nachfolgend werden einige der Konzepte vorgestellt, wie der II4D in verschiedene Studiengänge integriert werden könnte.

5.1 Wirtschaftsingenieurwesen: Industrie 4.0 Projekt

Studierende des Wirtschaftsingenieurwesens verfügen bis zum Start ihres Industrie 4.0 Projekts bereits über die notwendigen Programmierkenntnisse, einschließlich der Arbeit mit Mikrocontrollern, sowie elektrotechnisches Grundlagenwissen. Ein typisches Projekt könnte beispielsweise die Integration von Sensoren und die Übertragung von Sensordaten beinhalten. Dabei könnten Konzepte wie digitaler Zwilling oder Predictive Maintenance praktisch erprobt und angewendet werden.

5.2 Data Science und Artificial Intelligence: Prognose und Simulation

Im Modul "Prognose und Simulation" des Studiengangs Data Science und Artificial Intelligence könnten die Studierenden ihre bereits erworbenen Kenntnisse im Data

Engineering und Machine Learning nutzen, um Predictive-Maintenance-Verfahren für eine bestimmte Station oder den Transportroboter zu entwickeln, zu implementieren und zu bewerten.

5.3 IT-Ingenieurwesen: Technisches Softwareprojekt

Studierende des IT-Ingenieurwesens haben bis zum Beginn ihres technischen Softwareprojekts umfangreiche Programmierkenntnisse und Detailwissen im Bereich Elektrotechnik und Maschinenbau erworben. Zudem sind Grundlagen der Industrie 4.0, des Machine Learning und der Regelungstechnik vorhanden. Typische Projekte könnten eine optimierte Linienfolgeregulierung, die Integration von UWB-Lokalisierung oder die Erweiterung des Stationenportfolios umfassen.

5.4 Smart Manufacturing: Production Optimisation

Im Masterstudiengang Smart Manufacturing lernen die Studierenden im Modul Production Optimisation verschiedene Methoden kennen, Produktionsprozesse zu modellieren und zu optimieren. Insbesondere Line-Balancing Methoden sollen hier untersucht werden. In der Theorie werden Warteschlangenmodelle genutzt, die dann simulativ überprüft werden. Der finale Schritt ist dann die Umsetzung im Demonstrator, indem eine modellierte Fertigungsstraße praktisch umgesetzt wird und mit Messungen die theoretischen Ergebnisse betätigt werden.

6 Fazit

Insgesamt bietet der II4D eine wertvolle Ressource, um praktische Projekte in verschiedenen Studiengängen zu unterstützen und den Studierenden so einen realitätsnahen Einblick in die Welt von Industrie 4.0 zu ermöglichen. Diese Integration in die Lehre steht in vollem Einklang mit dem interdisziplinären Ansatz des II4D und trägt dazu bei, den Studierenden ein tiefgehendes Verständnis der Industrie 4.0-Technologien und -Prozesse zu vermitteln.

Literaturverzeichnis

- [Bu23] Bundesministerium für Wirtschaft und Klimaschutz, Plattform Industrie 4.0, <https://www.plattform-i40.de/IP/Navigation/DE/Industrie40/WasIndustrie40/was-ist-industrie-40.html>, Stand 25.05.2023
- [BTV14] Bauernhansl, T., Ten Hompel, M., Vogel-Heuser, B.: Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung, Technologien, Migration. Springer-Vieweg, Wiesbaden, 2014

- [Fo21] Fontanelli, D., Shamsfakhr, F., Bevilacqua, P. und Palopoli, L.: UWB Indoor Global Localisation for Nonholonomic Robots with Unknown Offset Compensation. 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 2021
- [GJ17] Götz, M., Jankowska, B.: Clusters and Industry 4.0—do they fit together?, *European Planning Studies*, 25(9), 1633-1653, 2017

Informatische Bildung: zwischen Nebenfach und Grundlagenbildung

Die Entwicklung eines generischen Musterkurses im Fach Informatik als Pilotprojekt für die Weiterentwicklung der Curricula von Grundlagenfächern

Luise Stromeyer¹ und Frank Burghardt²

Abstract: Im Projekt "Innovative digitale MINT-Curricula"³ erfolgt die Weiterentwicklung der Curricula der Grundlagenfächer Informatik und Mathematik unter Berücksichtigung eines interdisziplinären Ansatzes und die lernendenzentrierte Vermittlung mithilfe digitaler Angebote sowohl für die Lehre als auch das Lernen. Dabei werden digitale Hilfsmittel entwickelt, die einen Mehrwert für die Studierenden und das Lernen bieten. Die Bausteine ermöglichen eine optimale Vor- und Nachbereitung von Lehrveranstaltungen. Digitale Angebote erfolgen auf der Basis unterschiedlicher Moodle-Aufgaben und der Nutzung von VPL bzw. JupyterHub. Ein weiteres Ziel des Projektes ist die wissenschaftliche Evaluation der Maßnahmen hinsichtlich der Zielerreichung. In diesem Beitrag werden das bis dato entwickelte Lehrkonzept mit den zugrundeliegenden didaktischen Überlegungen für das Fach Informatik dargestellt, sowie das Evaluationsvorhaben skizziert.

Keywords: MINT-Curricula, Grundlagenausbildung, Moodle, digitale Lehre

1 Einleitung

Einige ingenieurtechnische Studiengänge, die an der HTW Berlin angeboten werden, haben ihren Ausbildungsschwerpunkt nicht in der Informatik. Da die Informatik in ihren unterschiedlichsten Facetten jedoch zum Berufsalltag gehört, ist in den Curricula der betreffenden Studiengänge eine mindestens einsemestrige Lehrveranstaltung Informatik verpflichtend verankert. Diese findet in der Regel im 1. Semester des Bachelorstudiums statt.

Da derzeit bei den Studierenden auf Grund der fehlenden obligatorischen Informatikausbildung in der Schule [Di22, S. 10] keine Grundlagen bezüglich des Faches vorausgesetzt werden können, reicht die Spanne allein bezüglich des fachlichen Vorwissens von „keine oder wenig Erfahrung“ bis „Leistungskurs Informatik“ und darüber hinaus. Dieser Umstand muss in einem Grundlagenkurs im 1. Semester eines Studiums beachtet werden.

¹ HTW Berlin, Fachbereich 1, Wilhelminenhofstraße 75A, 12459 Berlin, luise.stromeyer@htw-berlin.de

² HTW Berlin, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin, frank.burghardt@htw-berlin.de

³ <https://projekte.htw-berlin.de/hochschule/curriculum-innovation-hub/innovative-digitale-mint-curricula/>

Im Rahmen des Projektes „Innovative MINT-Curricula“ werden Instrumente und Konzepte entwickelt und erprobt, die die Weiterentwicklung der digitalen Lehr-Lern-Elemente gerade in solchen Grundlagenfächer, wie sie hier beschrieben wurden, ermöglichen. Als Pilotkurs dient die Informatikvorlesung im Studiengang Maschinenbau an der Hochschule für Technik und Wirtschaft (HTW) Berlin.

Die Arbeit erfolgt unter Beachtung fach- und fachdidaktischer Anforderungen, der Berücksichtigung der großen Heterogenität der Studierenden, aber auch der der Lehrenden, sowie dem Aspekt der Nachnutzbarkeit der erstellten Materialien. Als weiterer Aspekt wird auch die Möglichkeit der Digitalisierung und Automatisierung der Klausur angestrebt.

2 Das Konzept

Zu Beginn des Projektes wurden eine Bestandsaufnahme durchgeführt, mögliche Problemfelder identifiziert und das Ziel festgesteckt als greifbaren Output einen generischen Musterkurs in dem Lernmanagementsystem (LMS) Moodle zur Verfügung zu stellen. Dieser Musterkurs soll allen interessierten Lehrenden zur freien Nutzung bereitstehen und auch in kommenden Jahren, insbesondere durch weitere Erkenntnisse aus der Nutzung, erweitert und ergänzt werden können.

Entstanden ist ein modular aufgebauter Musterkurs, dessen Inhalte und Materialien Lehrende nach ihren Bedürfnissen und Anforderungen individuell zusammenstellen und in ihre eigene Lehre integrieren können, ergänzt mit Handreichungen für den didaktischen reflektierten Einsatz der Materialien. Inhaltlich und didaktisch war neben dem fachlichen Rahmen das leitende Element für den Gesamtaufbau immer die Lernendenperspektive, die Perspektive der Lehrenden wurde jedoch immer mitbedacht.

2.1 Fachlicher Inhalt der Lehrveranstaltung

Im Sinne des lebenslangen Lernens und einer ganzheitlich ausgerichteten Bildung soll die informatische Grundbildung, die in den hier betrachteten Kursen aufgebaut wird, sich nicht auf das Programmieren beschränken, sondern die Lernenden an die große Bandbreite der Informatik heranführen [G19]. Der reflektierte Einsatz der digitalen Technologien erfordert vielfältige Kenntnisse und Kompetenzen. In diesem Sinne umfassen die fachlichen Inhalte, die bereitgestellt werden, auch Themen wie Digitalisierung, IT-Sicherheit, Datenverarbeitung und Datentransfer u.s.w.

2.2 Effektivierung der Lernzeit

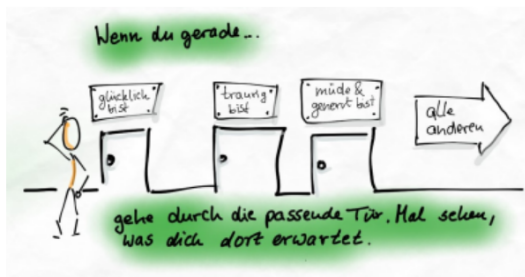
Eine als wichtig identifizierte Herausforderung dieser Nebenfachkurse ist der Punkt, dass die kontinuierliche Beschäftigung der Studierenden mit den Inhalten über das laufende

Semester zu selten gegeben ist. Es gibt keine verpflichtenden Hausaufgaben und die Note geht nicht in die Abschlussnote ein.

Im Sinne eines gewollten Wechsels weg von der Lehrendenzentrierung hin zu einer Lernendenzentrierung [Go19, S. 5] soll der Anteil der effektiv von den Lernenden nutzbaren Lernzeit innerhalb der Lehrveranstaltung erhöht werden.

Insbesondere im Bereich der Übung wurden hierfür im Projekt Selbstlernaufgaben entwickelt, die durch ein individualisiertes Feedback das Lernen im individuellen Tempo ermöglichen. Diese Aufgaben wurden in dem LMS Moodle mit Hilfe der Plugins STACK [St23] und VPL [Vp23] entwickelt.

Übungen if...else...








-  if ... else ... Niveau 1: richtigen Code erkennen
-  if ... else ... Niveau 2: Code ergaenzen
-  if ... else ... Niveau 3 : Code vollständig selbst schreiben, einfach
-  if ... else ... Niveau 4: Code vollständig selbst schreiben, schwieriger
-  Musterlösung if ... else ...

Abb. 1: VPL (Moodle) Programmierübungen auf verschiedenen Niveaustufen

Neben dem angesprochenen Feedback sind insbesondere für das Programmierenlernen binnendifferenzierte Aufgaben [vgl. Kl08, S. 102 ff.] entwickelt worden um der großen Heterogenität innerhalb der Lerngruppen zu entsprechen und allen Lernenden die Möglichkeit zu geben sich mit dem zentralen Lerngegenstand entsprechend ihres persönlichen Wissensstandes auseinanderzusetzen. Beispielsweise gibt es ausgehend vom Programmierkonzept der if-else-Verzweigung eine zentrale Aufgabenstellung, die für alle Niveaustufen gilt. Die notwendige Binnendifferenzierung erfolgt nun über das Erweitern dieser Aufgabenstellung, indem auf einer unteren Niveaustufe z.B. Teile vom zu schreibenden Code schon vorgegeben werden, auf einer höheren Stufe muss der gesamte Code vom Lernenden selbst geschrieben werden. Der Unterschied besteht darin, dass das Maß an Unterstützung innerhalb der Aufgaben sehr unterschiedlich ist. Lernende mit wenig Programmierkenntnissen können auf einer unteren Niveaustufe beginnen, fortgeschrittene Lernende schreiben den gesamten Code für diese Aufgabe allein. Auf

diese Weise arbeiten alle Lernenden am gleichen Lerngegenstand, also hier mit dem gleichen Programmierkonzept, jedoch ihren Fähigkeiten und (Vor)Kenntnissen entsprechend und nicht im Gleichschritt.

Neben den klassischen Rückmeldungen, die auch viele IDEs geben, ermöglicht VPL die Hinterlegung unterschiedlicher Testcases durch die Lehrenden, sodass hier eine zusätzliche Rückmeldung an die Studierenden darüber erfolgt, welche Fälle in ihrem Code noch nicht bedacht sind (Abb. 2).

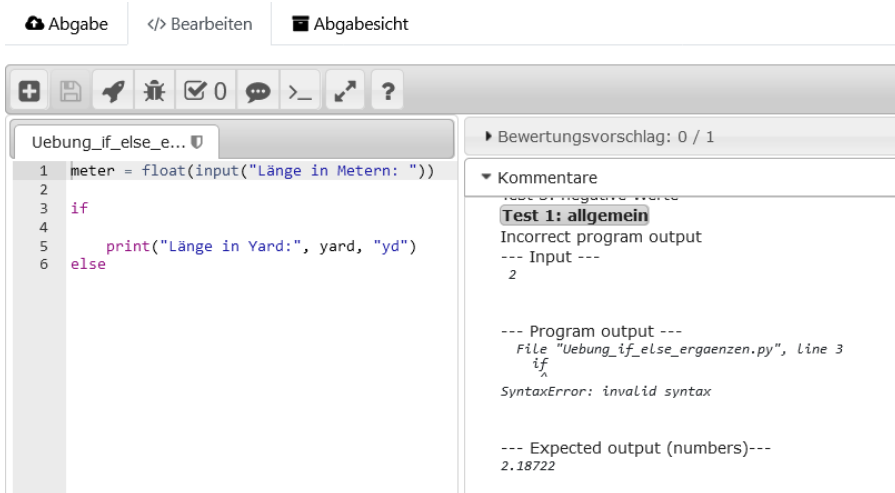


Abb. 2: Rückmeldung und Bewertung durch hinterlegte Testfälle

Pate für die Aufgabestellung stehen unter anderem Problemfelder aus anderen Fächern wie der Mathematik und der Physik. Hier soll auf lange Sicht noch eine stärkere inhaltliche Verzahnung erfolgen.

Daneben werden Möglichkeiten der Aktivierung der Studierenden in eher klassischen Vorlesungssituationen erprobt, wie z.B. die Nutzung von Quizzes im Plenum, aber auch analoge Elemente wie Murmelrunden etc. Diese Elemente, sowie themenbezogene Wissenstests in Moodle, geben den Studierenden schon während des Semesters regelmäßig eine Rückmeldung über ihren Wissensstand. Durch die Digitalisierung vieler Aufgaben, der Nutzung von VDI (Virtual Desktop Integratin) und dem Angebot der Nutzung von Notebooks in einem Jupyterhub können Studierenden auch ohne Installation umfangreicher Softwarepakete viele Aufgaben zuhause auf dem eigenen Laptop ausführen.

2.3 Weitere didaktische Entscheidungen

Aus lernpsychologischer Sicht wurde bewusst auf eine Reduzierung der Toolvielfalt auf das LMS Moodle mit den Plugins Stack und VPL geachtet. Jede Methode und jedes Tool, das verwendet wird, muss von den Studierenden erst erlernt werden. Dementsprechend müssen alle verwendeten Tools auf ihre Nützlichkeit und Notwendigkeit hin geprüft werden, insbesondere, wenn ihre Bedienung komplex ist und neben dem fachlichen Inhalt erst erlernt werden muss. Da das Fach für die Studierenden Nebenrollencharakter hat, muss damit gerechnet werden, dass die Bereitschaft sich immer wieder in neue Tools und Methoden einzuarbeiten, die nicht primär mit dem fachlichen Inhalt zu tun haben, schnell als zusätzlich Belastung empfunden wird.

Die Binnendifferenzierung, die durch die Heterogenität der Lerngruppen gefordert ist, sowie die Möglichkeiten des automatisierten, individualisierten Feedbacks in Übungsphasen sind wichtige Grundbausteine für die Lernendenzentrierung. Darüber hinaus bietet die Darbietung aller Inhalte innerhalb eines LMS die Möglichkeit, die Lerninhalte sowohl fachlich, als auch fachdidaktisch sinnvoll zu bündeln und für die Lernenden zu strukturieren.

2.4 Nachnutzbarkeit und Verbreitung

Sämtliche Übungen und Materialien stehen in einem frei innerhalb der HTW zugänglichen Musterkurs in Moodle zur Verfügung und können nachgenutzt und angepasst werden.

Darüber hinaus werden Handreichungen zu technischen Fragen, sowie zum Gesamtkonzept und weiterführenden didaktischen Ideen und Ansätzen erarbeitet, sodass die individualisierte Nachnutzung und Erweiterung ermöglicht wird.

3 Die Evaluation der binnendifferenzierten Selbstlernaufgaben

Die innerhalb dieses Projektes erstellten Selbstlernaufgaben sind in ihrer Handhabung nicht trivial. Neben der fachlichen Fragestellung müssen die Studierenden auch die Eingaben richtig vornehmen, die Rückmeldungen aufnehmen und verarbeiten, Hinweise und Hilfe selbständig nutzen und auch das Feedback, das an unterschiedlichen Stellen zu bekommen ist, einholen.

Da das Erstellen solch komplexer Aufgaben sehr aufwendig ist, müssen diese nicht nur fehlerfrei laufen, sondern sollen den Lernenden auch wirklich die Möglichkeit zum Lernen geben. Sollte die technische Seite der Eingabe zu komplex sein und von der inhaltlichen Aufgabe ablenken, wären die Aufgaben zum Lernen ungeeignet.

Neben der regelmäßigen Erprobung und Anpassung der Aufgaben wird deshalb innerhalb des Projektes eine Evaluation in Form einer Vorstudie zum Nutzungsverhalten

durchgeführt. Hierfür werden einzelne Studierende dazu angehalten während der Bearbeitung der Aufgaben all ihre Gedanken laut auszusprechen und all ihr Handeln zu kommentieren (sogenannte think-aloud-Interviews oder Methode des lauten Denkens, vgl. [BD14 S.369 ff.]). Die so entstehenden Interviews werden aufgenommen (nur Ton) und anschließend transkribiert. Die dann in schriftlicher Form vorliegenden Aussagen in Folgenden kategorisiert, um so zu filtern, womit die Lernenden während der Bearbeitung der Aufgaben vor allem beschäftigt sind, wo Verbesserungsbedarf in Bezug auf die erstellten Aufgaben besteht und welche anderen Hürden und Hindernisse einer erfolgreichen Bearbeitung der Aufgaben durch Lernende gegebenenfalls im Wege stehen.

Zusätzlich werden für die Evaluation die Nutzungsdaten von Moodle in die Auswertung mit einbezogen. Die Auswertung der Evaluation ist für Herbst 2023 geplant.

3.1 Voraussetzungen und Durchführungsbedingungen der Interviews

Im Rahmen der Evaluation wurden bis Juli 2023 auf der Basis der Freiwilligkeit 6 Interviews mit Studierenden aus dem oben genannten Studiengang Maschinenbau durchgeführt. Die Studierenden waren zum Zeitpunkt der Interviews im 1. Fachsemester immatrikuliert und belegten die Pflichtveranstaltung Informatik, eine seminaristische Lehrveranstaltung mit Übung und Tutorium, die für dieses Projekt, wie in Kapitel 1 beschrieben, als Pilotveranstaltung dient. Die Aufgaben, die während der Interviews bearbeitet wurden, waren alle aus dem Bereich der binnendifferenzierten Programmieraufgaben, die mit dem Plugin VPL in Moodle realisiert wurden.

Zum SoSe 2023 konnte zum ersten Mal ein gesamter Satz dieser binnendifferenzierten Aufgaben genutzt werden, das heißt, es stehen zu den Programmierkonzepten „while-Schleife“, „for-Schleife“, „if-else-Verzweigung“ und „Funktionen“ für die Programmiersprache Python entsprechende Aufgabensets zur Verfügung. Diese Aufgaben wurden im Rahmen der zur Veranstaltung gehörenden Übung von den Studierenden bearbeitet, stehen aber auch bis zur Klausur zur freien Bearbeitung zur Verfügung. In der Vorlesung wurde das Plugin einmal zentral vorgestellt, eine Kurzanleitung stand den Studierenden in Moodle zur Verfügung. In der Übung wurde das Plugin mit einer einfachen Programmieraufgabe gemeinsam ausprobiert und auf Besonderheiten wie Hilfen, Hinweis und Rückmeldungen, hingewiesen, bevor in den folgenden Wochen mit Hilfe der VPL-Aufgaben die Konzepte gezielt geübt wurden. Die Konzepte selbst wurden in den Vorlesungen eingeführt.

Die Interviews wurden zeitlich parallel zur Übung durchgeführt. Jedes Interview wurde als Audio-Datei aufgezeichnet. Da die teilnehmenden Studierenden keinerlei Nachteile durch die Interviewsituation, insbesondere in Bezug auf die Klausur haben sollten, konnten sie während des Interviews genauso Fragen stellen und alle erlaubten Hilfen nutzen, wie die anderen Lernenden während der Übung.

Zu beachten ist, das mit fortlaufendem Semester die interviewten Studierenden über immer mehr Kenntnisse und Erfahrung im Umgang mit den VPL-Aufgaben verfügen können.

3.2 Die Interviews

Die think-aloud-Interviews wurden nach folgendem Schema durchgeführt:

1. Damit die Interviewten mit der Methode des lauten Denkens vertraut werden, wird ihnen eine einfache Aufgabe ohne Bezug zur Programmierung gestellt und sie versuchen diese mit lautem Denken, aber noch ohne Aufnahme, zu lösen.
2. Die eigentliche Bearbeitung der Aufgabe wird mit zwei Frageabschnitten umrahmt. Diese offenen Fragen holen direkte Rückmeldungen in Bezug auf den Beobachtungsgegenstand ein.

Interviewfragen

vor der Bearbeitung:

Wie gut schätzen Sie Ihr Können in Bezug auf das ... (hier behandelte Thema eintragen/nennen) ein?

nach der Bearbeitung:

(icebreaker: Was nehmen Sie persönlich, für sich, aus der heutigen Aufgabe/Bearbeitung/Übung mit?)

1. Wie schwer fanden Sie die Aufgabe?
2. Fassen Sie kurz in 1-2 Sätzen zusammen, was Sie mit dieser Aufgabe gelernt haben.
3. Hat es sich für Sie gelohnt diese Aufgabe zu bearbeiten?
4. Was hat Ihnen an der Aufgabe gefallen und warum?
5. Was würden Sie bei der Aufgabenstellung/Aufgabe anders machen/verändern/verbessern und wozu? Womit sind Sie nicht so gut zurechtgekommen und warum?
6. Schätzen Sie bitte ein, mit welchen Aspekten der Aufgabe Sie während der Bearbeitung am meisten beschäftigt waren. Z.B.: Suche mit..., Rechnen, Eingabetechnik, Programmierthema, ...?

Abb. 3: Die einbettenden Interviewfragen

3. Die Studierenden bearbeiten eine selbstgewählte Übungsaufgabe aus dem gegebenen Set, dass zu einem Programmierkonzept gehört und sprechen alle Gedanken laut aus und kommentieren, was sie gerade tun.

3.3 Erste Eindrücke aus den Interviews

Grundsätzlich wird das Angebot von den Studierenden laut Selbstaussagen gut angenommen und als nützlich bewertet. Eine Person, die erst einige Wochen nach Semesterbeginn das Studium aufnehmen konnte, war motiviert mit diesen Aufgaben die entstandenen Lücken selbständig aufzuarbeiten.

Unabhängig von der Selbsteinschätzung in Bezug auf den eigenen Lernstand haben alle interviewten Personen mit den Aufgaben des Niveaus 1 angefangen. In der Regel wurde dies mit mangelnden Kenntnissen begründet, aber auch mit Unsicherheit. Die Zeit, die für die Bearbeitung der Aufgabe auf Niveau 1 benötigt wurde, war sehr unterschiedlich. Inwieweit für die Wahl des Niveaus die Interviewsituation mit verantwortlich war und die interviewten Personen lieber Aufgaben gewählt haben, die sie in Gegenwart der interviewenden Person eher bewältigen können, kann nicht sicher festgestellt werden.

Alle Studierenden haben sich während der Bearbeitung kontinuierlich und den größten Teil der Zeit mit dem Lerngegenstand beschäftigt. Schwierigkeiten gab es in Bezug auf das Nutzen der Hinweise, Rückmeldungen und Hilfen. Konnten diese nicht zielführend genutzt werden, half jedoch oft die Intervention durch die interviewende Person in Form einer Erklärung worauf man z.B. bei der Rückmeldung achten kann. Dies wirft die weiterführende Frage auf, über welches Wissen und welche Kompetenzen Lernende, die erfolgreich mit den erstellten VPL-Aufgaben arbeiten wollen, verfügen müssen. Rückmeldungen wie „line 3: if ^ SyntaxError: invalid syntax“ müssen gelesen werden können, damit man mit ihnen zielführend arbeiten kann.

Auf die Interviewfrage „Fassen Sie kurz in 1-2 Sätzen zusammen, was Sie mit dieser Aufgabe gelernt haben.“ haben alle Interviewteilnehmenden antworten können. Die Angaben bezogen sich immer auf das Programmieren lernen, waren aber individuell sehr unterschiedlich. Sie reichten von konkreten Aussagen bezüglich des individuellen Wissenszuwachses in Bezug auf den konkreten Lerngegenstand der bearbeiteten Aufgabe, über allgemeingültigere Aussagen, z.B., dass die Syntax bei Programmieren offensichtlich wichtig ist, bis hin zum Reflektieren des eigenen Lernens.

Obwohl die interviewten Personen mit fortschreitendem Semester teilweise schon über Erfahrungen mit der Bearbeitung von VPL-Aufgaben verfügten, kann über die Gewöhnung an dieses Aufgabenformat im Rahmen dieser Vorstudie keine Aussage getroffen werden. Die Fallzahlen sind mit 6 Teilnehmenden dafür zu klein. Um die Heterogenität der Lernenden in diesem Rahmen dennoch bestmöglich zu berücksichtigen, wurden die Interviews mit 6 unterschiedlichen Personen geführt, Längsschnittdaten liegen deshalb nicht vor.

4 Fazit und Ausblick

Die geforderte und notwendige Weiterentwicklung bestehender Curricula umfasst mehrere Facetten [Go19]. Eine davon ist die verpflichtende Verankerung digitaler Lehr-Lern-Inhalte. Im Rahmen des Teilprojektes „Innovative digitale MINT-Curricula“ wird exemplarisch gezeigt, welche Möglichkeiten digitale Lehr-Lern-Inhalte bieten und wie sie in die bestehende Lehre eingebunden werden können. Gleichzeitig wird deutlich, dass das Erstellen solcher Inhalte sehr zeit- und ressourcenaufwendig ist, wenn man die Möglichkeiten der digitalen Aufgaben umfangreich nutzen möchte. Solche Möglichkeiten

sind z.B. Formen des individualisierten Feedbacks, punktuell abgestimmte Hinweise und Aufgabenquantität durch Randomisierung von Aufgabeninhalten.

Das Interesse und der Bedarf an solchen Inhalten sind groß. Dies zeigen Gespräche mit Lehrenden. Herausfordernd scheint die Implementierung in die eigene Lehre zu sein, denn digitalen Aufgaben sieht man oft auf den ersten Blick nicht an, welche „Fähigkeiten“ sie haben. Das „Erkunden“ der Aufgaben und Materialien durch die Lehrenden benötigt viel Zeit, die Qualität der frei verfügbaren Aufgaben muss momentan für jede Aufgaben von den Lehrenden selbst beurteilt werden. Aus diesem Grund liegt neben der Entwicklung und kontinuierlichen Weiterentwicklung der Aufgaben ein weiterer Schwerpunkt im Projekt in der kommenden Zeit auf der Erarbeitung und Bereitstellung von Begleitmaterialien für Lehrende, die die Materialien des Musterkurses nutzen wollen.

Danksagung

Das Projekt „Innovative digitale MINT-Curricula“ ist ein Teilprojekt im Gesamtprojekt „Curriculum Innovation Hub“, welches von der Stiftung Innovation in der Hochschullehre gefördert wird (08/21 – 07/24).

Literaturverzeichnis

- [BD14] Bortz, J., Döring, N.: Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5. Auflage, Springer-Verlag Berlin Heidelberg, 2014.
- [Di22] Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (SWK) (Hrsg.): Digitalisierung im Bildungssystem: Handlungsempfehlungen von der Kita bis zur Hochschule. Zusammenfassung. Bonn, 2022. <http://dx.doi.org/10.25656/01:25274>, Stand 24.05.2023.
- [Go19] Gottburgsen, A. et.al.: Ingenieurausbildung für die digitale Transformation. Zukunft durch Veränderung. VDI-Studie April 2019. https://www.vdi.de/fileadmin/user_upload/VDI-Studie_Ingenieurausbildung_fuer_die_Digitale_Transformation.pdf, Stand 01.06.2023.
- [Kl08] Klippert, H.: Besser lernen. Kompetenzvermittlung und Schüleraktivierung im Schulalltag. Klett Lernen und Wissen GmbH, Stuttgart, 2008.
- [St23] STACK, Online assessment, <https://stack-assessment.org/>, Stand: 06.07.2023.
- [Vp23] VPL, the Virtual Programming lab for Moodle, <https://vpl.dis.ulpgc.es/>, Stand: 06.07.2023.

Erfolgsfaktor Interdisziplinarität: Das Lehrformat Digitalisierungskollegs an Bayerischen Hochschulen

Christina Dieckhoff¹, Christophe Barlieb², Christian Groth³, Thomas Linner⁴ und Florian Weininger⁵

Abstract: Um den zunehmenden Anforderungen an die Beherrschung digitaler Techniken und an die Fähigkeit zur interdisziplinären Zusammenarbeit an Studierende aller Fachrichtungen zu begegnen wurde das interdisziplinäre Lehrformat Digitalisierungskollegs für Studierende entwickelt. Das in vielen Fachbereichen ausbaufähige Angebot von Digitalthemen in der Hochschullehre wird hiermit dauerhaft erweitert. Ein Digitalisierungskolleg besteht aus einer Vorlesungsreihe mit angrenzendem Seminar, in denen Studierende interdisziplinäre Lösungen für Fragen der digitalen Transformation entwickeln. Geleitet werden sie von etablierten Wissenschaftlerinnen und Wissenschaftlern, aktiv betreut und ausgestaltet von ein bis zwei Coaches. Kernelement sowohl des Kollegs als auch der einzelnen Projekte ist die Interdisziplinarität. Eine*r der beteiligten Projektleiter*innen hat einen direkten Bezug zur Technik und kommt aus der Informatik, der Wirtschaftsinformatik, der Elektrotechnik oder vergleichbaren Disziplinen. Zielgruppe der Projekte sind Studierende verschiedener Disziplinen im Masterstudium oder in den letzten Semestern eines Bachelorstudiums. Durch die Teilnahme erwerben auch Studierende aus digitalisierungsfernen Fächern frühzeitig umfangreiche IT-Kenntnisse. Als Begleiteffekt der umfangreichen Vernetzung zwischen den Digitalisierungskollegs (Studierende, Coaches und Projektleitende) entsteht bereits zu Beginn einer wissenschaftlichen Karriere eine große digitale Community. Alle Teilnehmenden lernen frühzeitig die interdisziplinäre Zusammenarbeit und verbessern erheblich ihre Karrierechancen innerhalb und außerhalb der Wissenschaft.

Keywords: Digitalisierung, Digitalisierungskompetenz, Digitalisierungskollegs, Interdisziplinarität

1 Einleitung

Eine interne und noch unveröffentlichte Studie des bidt (Bayerisches Forschungsinstitut für digitale Transformation) in Kooperation mit dem Lehrstuhl für Software Engineering an der TU München hat die Auswertung der Digital Literacy in der Hochschullehre in

¹ Christina Dieckhoff, Bayerisches Forschungsinstitut für Digitale Transformation (bidt), Gabelsbergerstr. 4, 80333 München, christina.dieckhoff@bidt.digital

² Christophe Barlieb, Fakultät für Architektur, OTH Regensburg, Seybothstraße 2 93053 Regensburg; christophe.barlieb@oth-regensburg.de

³ Christian Groth, Fakultät für Informatik, Hochschule Hof, Alfons-Goppel-Platz 1, 95028 Hof, christian.groth@hof-university.de

⁴ Thomas Linner, Fakultät Bauingenieurwesen, OTH Regensburg, Seybothstraße 2 93053 Regensburg; thomas.linner@oth-regensburg.de

⁵ Florian Weininger, Fakultät Bauingenieurwesen, OTH Regensburg, Seybothstraße 2 93053 Regensburg, florian.weininger@oth-regensburg.de

Bayern im Sommersemester 2023 der TU München, Uni Augsburg, LMU München, Universität Regensburg und der JMU Würzburg zum Thema. Erste Ergebnisse zeigen, dass Digital Literacy in der Hochschullehre zumeist auf die Verwendung digitaler Tools und Techniken reduziert ist und selten interdisziplinär angegangen wird. Auswirkungen der Digitalisierung auf Prozesse, Strukturen und Forschungsinhalte werden in den Fachbereichen außerhalb der Informatik noch verhältnismäßig selten adäquat adressiert.

Digitalisierungskompetenzen und die Fähigkeit zum interdisziplinären Arbeiten werden in Zukunft in allen wissenschaftlichen Disziplinen ebenso wie am Arbeitsmarkt von überragender Bedeutung sein [Br19; PBB18]. Auf diesen Bedarf wird seit wenigen Jahren auch seitens der Fördergeber*innen reagiert. So schreibt etwa die European Commission in ihren wissenschaftlichen Förderprogrammen zwingend die Erstellung eines Data Management Plans vor. In der Ausschreibung des Akademienprogramms aus dem Jahr 2019 wurde für die Antragseinreichung erstmals eine Aussage zur Verwendung von digitalen Techniken [Au20] und eine Aussage zur Weiterqualifizierung der wissenschaftlichen Mitarbeiter*innen gefordert [Au20], vorrangig im Erwerb digitaler Kompetenzen. In der Ausschreibung ein digitales Bearbeitungs- und Publikationskonzept sowie Überlegungen zur Langzeitarchivierung erwartet sowie der Einsatz von Informationstechnik, besonders hinsichtlich der Publikation der Forschungsergebnisse und der langfristigen Datensicherung. Der Gesamteindruck ist, dass auch 2023 digitale Kompetenzen nur vereinzelt und bei Weitem nicht ausreichend vermittelt werden [Br19]. Somit besteht weiterhin in vielen Fächern ein erheblicher Nachholbedarf, wie es auch in der Dagstuhl-Erklärung „Bildung in der digitalen vernetzten Welt“ von März 2016 beschrieben wird [Da16]. Die noch unveröffentlichte Studie des bidt bestätigt, dass das Thema Interdisziplinarität ausbaufähig ist, wie im Positionspapier des Wissenschaftsrats von 2020 (S. 34-39 und 59- 66) beschrieben wird [Wi20]. Sie zeigt darüber hinaus die oft noch ausstehende Verankerung von Digitalthemen in der Hochschullehre außerhalb der Informatik und den Naturwissenschaften.

Die bisherigen Förderformate setzen digitale Kenntnisse voraus, die im dafür erforderlichen Umfang bisher noch nicht vermittelt werden. Eine 2017 erstmals veröffentlichte Richtlinie zur Messung der digitalen Kompetenzen Lehrender zeigt, dass hier noch Nachholbedarf besteht [Eu19]. Die Lösung zukünftiger Fragen zu Digitalisierungsthemen setzt umfangreiche interdisziplinäre und digitale Kenntnisse voraus, die bereits frühzeitig als ein integraler Bestandteil der Studierendenausbildung gelehrt werden sollten. In digitalen Forschungs- und Arbeitsumgebungen werden neue digitale Methoden wie beispielsweise Prompt Engineering mit ChatGPT zur Entwicklung von Roboter-Codes ebenso wie bei der Entwicklung von Nutzer Interfaces oder Verbesserung der User Experience in der Mensch-Maschine-Kommunikation eine hohe Bedeutung bekommen.

Der frühzeitige Erwerb digitaler Kompetenzen und Kenntnisse im interdisziplinären Kontext, beispielsweise in Programmierertools wie Python oder C++ dienen den Studierenden hierfür als Grundlage für ein Grundverständnis für Digitalisierungsthemen wie beispielsweise die Verwendung von Machine Learning in der Robotik oder die Gestaltung von AR-Umgebungen unter Wahrung ausreichender Distanz für den Benutzer oder die

Gestaltung von Netzwerkumgebungen. Dies kann bereits durch den Erwerb von Grundkenntnissen und handwerklicher Kompetenz oder von Prozesswissen zur Lösung digitaler Aufgaben erfolgen. Eine interdisziplinäre Herangehensweise im Kompetenzerwerb berücksichtigt dabei die zunehmende Komplexität digitaler Fragestellungen, die sich durch rechtzeitige Integration weiterer Fachperspektiven nachhaltiger beantworten lassen.

Die Reaktion auf die zunehmend komplexen Bedarfe ist die Ausbildung einer interdisziplinär denkenden Generation, die eine neue Herangehensweise verfolgt, indem sie Digitalisierungsfragen bereits zu Beginn interdisziplinär denkt und löst, weil sie die hierfür erforderlichen Kompetenzen besitzt. Aus diesem Grund wurde 2021 am bidt in Zusammenarbeit mit dem Bayerischen Staatsministerium für Wissenschaft und Kunst das interdisziplinäre Lehrformat „Digitalisierungskolleg“ entwickelt, das Anfang 2022 ausgeschrieben wurde. Das Bayerische Forschungsinstitut für Digitale Transformation (bidt) koordiniert diese Programme. Das bidt trägt als Institut der Bayerischen Akademie der Wissenschaften (BAW) dazu bei, die Entwicklungen und Herausforderungen des digitalen Wandels besser zu verstehen. Das bidt untersucht Chancen und Herausforderungen der digitalen Transformation und informiert über aktuelle Entwicklungen. Es versteht sich als Schnittstelle zwischen Forschung, Politik, Wirtschaft und Gesellschaft. Durch seine konsequent interdisziplinäre Ausrichtung verfügt das Institut über breite Expertise und Erfahrung, das Schnittstellenthema Digitalisierung schon im Studium zu begleiten und zu unterstützen. Zudem bietet das bidt mit seinen vielfältigen Forschungsprojekten und Dialogaktivitäten ein fruchtbares Umfeld mit zahlreichen Kooperationsmöglichkeiten und Vernetzungspartnern für die Teilnehmenden der Förderprogramme.

In diesem Paper wird die Struktur des Lehrformats und die Zielsetzung der Ausschreibung beschrieben. Anhand der Projekte „CyberCraft Kolleg“ der OTH Regensburg und „digitale Methoden“ der Universität Bayreuth mit der Hochschule Hof wird anhand erster Erfahrungen dargestellt, wie die Interdisziplinarität in der Hochschullehre zum Thema Digitalisierung in der Zusammenarbeit von Ingenieurwissenschaften und Informatik umgesetzt wird.

2 Das Förderformat Digitalisierungskollegs

Die Ausschreibung war an alle staatlichen Hochschulen in Bayern gerichtet. Die Beteiligung mindestens einer Technikwissenschaft bzw. der Informatik war Voraussetzung für die Einreichung eines Antrags. Alle Kollegs sind für eine Dauer von vier Jahren als Projekt aufgesetzt und werden mit Mitteln des Freistaats Bayern gefördert. Eine Verstetigung nach erfolgreicher Zwischenevaluierung an den Hochschulen ist vorgesehen. Der Umfang eines Kollegs wird von den Antragsteller*innen mitbestimmt: Gefördert werden Personal- und Sachkosten in einem Umfang von 100%, eine Programmpauschale ist nicht vorgesehen. Das sind bis zu zwei Stellen für sogenannte „Coaches“, die die Studierenden betreuen, sowie Sachmittel für die Ausstattung der Kollegs. Die Hochschulen werden an der Förderung durch Bereitstellung der Räume und Bereitstellung der für die Arbeit

notwendigen Grundausstattung beteiligt und haben im Antrag eine Verstärkung des Formats nach Ablauf der Förderung zugesagt. Die Digitalisierungskollegs sind daher ein nachhaltiges Förderformat.

Die Resonanz auf die Ausschreibung war hoch: dreizehn Digitalisierungskollegs aus sehr unterschiedlichen Fachbereichen wurden bewilligt, die jeweils sehr innovative Fragestellungen bearbeiten (Abbildung 1). Sechs davon sind Verbundkollegs mehrerer Hochschulen. An jedem Kolleg ist ein*e Projektleiter*in (Professor*in) aus der Informatik, der Wirtschaftsinformatik oder der Elektrotechnik beteiligt. Die inhaltliche Ausrichtung der Kollegs konnte durch die Antragsteller*innen frei bestimmt werden. Es gibt daher ein breites Fächerspektrum an Digitalisierungsthemen. Die thematische Vielfalt der Kollegs kann und soll auch dem koordinierenden Institut bidt als Anregung dienen, um neue Themenfelder und Fragestellungen aufzugreifen.

Digitalisierungskollegs

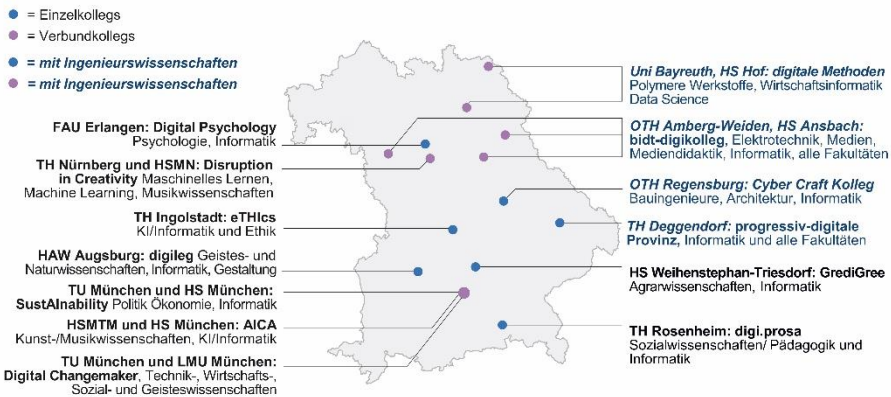


Abb. 1: Lage und Themen der Digitalisierungskollegs
 Grafik: Dieckhoff

2.1 Thematische Ausrichtung des Lehrformats

Verschiedene Vorgehensweisen verhelfen den Studierenden zum Wissenserwerb: Tutorials, Übungen, Projekte, Vorlesungen. Dabei werden unterschiedliche Lerntypen der Studierenden angesprochen und im Digitalisierungskolleg berücksichtigt. Ein Digitalisierungskolleg kann ein Zusatzstudium oder auch ein Modul innerhalb eines bestehenden Lehrangebots sein. Die Einbettung in die Hochschullehre konnten die Antragsteller*innen selber definieren. Zielgruppe sind Studierende unterschiedlicher Fächer im Master-Studium oder in den letzten Semestern eines Bachelor-Studiums. Die Teilnahme dauert ein bis zwei Semester. Die Attraktivität des Lehrformats wird durch den Erwerb von ECTS Punkten erreicht, so profitieren Studierende von der Teilnahme an den Digitalisierungskollegs. Hier sollen im ersten Schritt durch Teilnahme an einem umfangreichen

Lehrprogramm (Vorlesung, Seminar oder Projektveranstaltung) grundlegende Digitalkompetenzen wie beispielsweise Python, Deep Learning, Natural Language Understanding, Intelligent Robotics, LaTeX oder Generative AI erworben werden (Programm des Verbundkollegs der Uni Bayreuth mit der Hochschule Hof).

Im zweiten Schritt wird in eigenen, von Coaches betreuten Projekten, die auch als Abschlussarbeit verwendet werden dürfen, inhaltlich und praxisorientiert gearbeitet. Digitale Fragestellungen aus Bereichen wie beispielsweise Maschinelles Lernen, Künstliche Intelligenz „AI“ und „Applied Deep Learning“ sollen bearbeitet werden und aktuelle Fragen rund um die Entwicklung digitaler Techniken und deren Folgen für die Gesellschaft aufgreifen. Diese können zum Beispiel technische Themen mit gesellschaftsrelevanter Reflexion (z.B. KI-Entwicklung, Cyber Security, Green Tech), das Ermitteln von tatsächlichen Bedarfen und die Entwicklung von Anwendungen (z.B. in den Bereichen Robotik, Gesundheit, Medizin), oder die Digitalisierung kleinteiliger, menschenzentrierter industrieller Prozesse mit einer stark handwerklich geprägten, lokalen Kultur betreffen. „Baukultur“ (Klimagerechtes Bauen) dient in einem Kolleg als Anwendungsfall, da insbesondere das Bauwesen durch seine kleinteilige, manuelle und handwerklich geprägte Arbeitskultur von nachhaltigen Digitalisierungskonzepten via Mikroinnovationen profitieren könnte. Dieser Bereich ist im Besonderen durch den bevorstehenden Generationenwechsel und dem Fachkräftemangel betroffen.

2.2 Struktur des Lehrformats

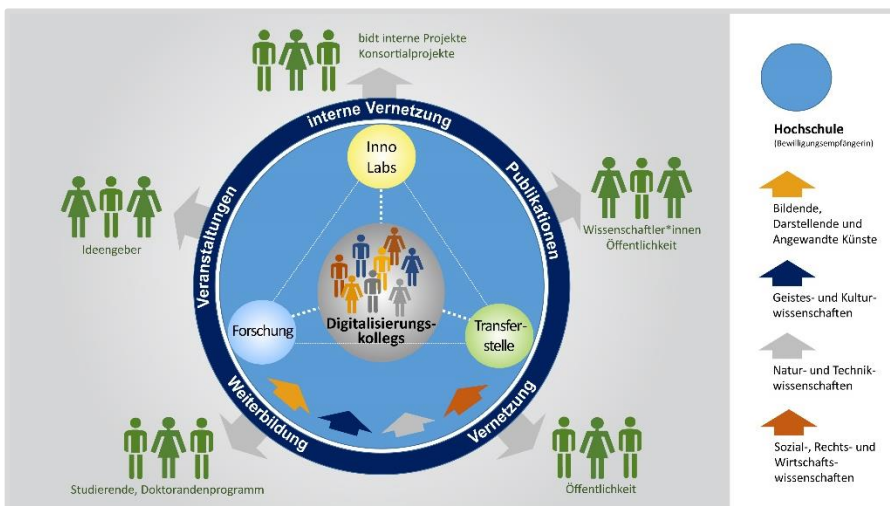


Abb. 2: Aufbau und mögliche Einbettung eines Digitalisierungskollegs
Grafik: Dieckhoff

Kernelement eines Kollegs und der einzelnen Projekte ist die Interdisziplinarität: Die Ausschreibung war offen für alle Fächergruppen (Abbildung 2). Voraussetzung für eine Förderung war die Beteiligung von mindestens zwei unterschiedlichen Fächern. Eine*r der beteiligten Wissenschaftler*innen sollte einen direkten Bezug zur Technik haben, d.h. aus der Informatik, der Wirtschaftsinformatik, der Elektrotechnik oder einer vergleichbaren Disziplin kommen. Die einzelnen Lehr- und Lernformate erstrecken sich auf ein bis zwei Semester, in dem Studierende interdisziplinäre Lösungen zu Fragen der digitalen Transformation entwickeln. Es wird von einer*m oder mehreren Projektleiter*innen geleitet und von ein oder mehreren Coaches aktiv betreut und ausgestaltet.

Alle Beteiligten (Projektleiter*innen, Mitarbeitende wie Coaches und Studierende) haben ein interdisziplinäres Mindset und sind untereinander kooperationsfähig. Daher ist der fachübergreifende Austausch von Methoden zwischen den Digitalisierungskollegs zu erwarten unter der Annahme, dass die Kollegs zwar sehr verschiedene Themenfelder haben, aber vergleichbare Fragestellungen bearbeiten. Das übergeordnete Querschnittsthema Digitalisierung bietet den Beteiligten daher einen Rahmen für einen engen und produktiven Austausch und eine Erweiterung der digitalen Kompetenzen [EMT17].

Die Bildung einer digitalen Community erfolgt durch eine gezielte Unterstützung durch die Programmkoordination mit Schwerpunkt auf Interdisziplinarität in allen Maßnahmen und Veranstaltungen und ein starkes Community- und Netzwerkmanagement. Es unterstützt die Anbahnung neuer Kooperationen und die Einbindung anderer Förderprogramme, v.a. der Nachwuchsförderung am bidt. Die Koordination ermöglicht in virtuellen Meetings und Vierteljahrestreffen in Präsenz an den verschiedenen Standorten der Kollegs einen konkurrenzfreien inhaltlichen Austausch. Darüber hinaus steht die Förderung eines ganzheitlichen Career Developments (innerhalb wie außerhalb der Wissenschaften) für alle Beteiligten im Vordergrund. Die Koordination unterstützt die Digitalisierungskollegs gezielt in folgenden Bereichen:

Schulungen: Für die Coaches der Kollegs gibt es die Möglichkeit, an Schulungen des bidt und des Graduate Centers für Promovierende und Postdocs teilzunehmen.

Veranstaltungen: Die Coaches sowie die Projektleitungen der Kollegs können mit anderen Wissenschaftler*innen am bidt an den zahlreichen Veranstaltungen des Instituts in Austausch treten. Postdocs und Nachwuchsgruppenleiter*innen aus dem bidt Graduate Center geben Fachvorträge auf den Netzwerktreffen der Digitalisierungskollegs.

Interne Vernetzung: die Vierteljahrestreffen ermöglichen den internen Austausch aller Beteiligten zu fachlichen Themen, es wurden mehrere Arbeitskreise gegründet. Auf einem jährlichen Demo-Day werden die besten Studierendenprojekte in verschiedenen Kategorien präsentiert und ausgezeichnet. Am Ende wird eine Prämierung der besten Projekte/Ideen vorgenommen. Die Ergebnisse der Digitalisierungskollegs werden publiziert und zudem in einem Veranstaltungsformat am bidt präsentiert (z.B. bidt-Konferenzen).

Externe Vernetzung: Eine Vernetzung aller Beteiligten mit relevanten Akteuren, wie z.B. aus dem Vorgängerformat Innovationslabore, mit den Postdocs im Graduate Center, mit

den Nachwuchsgruppenleiter*innen, mit den Transferstellen der Hochschulen und ggf. auch mit Unternehmen, wird aktiv durch die Koordination gefördert.

Publikationen: Das bidt steht als Plattform für öffentlichkeitswirksame Präsentationen bereit, die aus den Projekten eines Digitalisierungskollegs oder den Arbeitskreisen der Betreuer*innen hervorgehen können.

3 Zwei Digitalisierungskollegs am Beispiel

Für diesen Beitrag sind zwei Kollegs mit Beteiligung der Ingenieurwissenschaften relevant, die genauer betrachtet werden. Zum einen das „CyberCraft Kolleg“ (CKK) der OTH Regensburg mit den Fakultäten Bauingenieurwissenschaften, Architektur und der Regensburg School of Digital Sciences. Und zum anderen das Verbundkolleg „digitale Methoden“, bestehend aus der Uni Bayreuth und der Uni Hof mit dem Lehrstuhl Polymere Werkstoffe der Fakultät „Ingenieurwissenschaften mit Informatik“.

3.1 Interdisziplinäres Konzept und thematischer Fokus des CyberCraft Kollegs (CCK) und des Verbundkollegs „digitale Methoden“:

CyberCraft beschreibt im Digitalisierungskolleg der OTH Regensburg ein systematisches Vorgehen bei der Umsetzung einer menschenzentrierten digitalen Transformation in Bereichen wie dem Handwerk und dem Bauwesen. Das CCK wird über den Förderzeitraum neue Lernangebote wie Projektkurse, Weiterbildung, Masterstudium etc. an für den Anwendungsbereich hervorbringen, um gezielt im Innovationsprozess neue digitale Kompetenzen aufzubauen.

Der Modus des Digitalisierungskollegs „digitale Methoden“ der Universität Bayreuth mit der Hochschule Hof lässt sich im Wesentlichen auf einen semesterweisen Wechsel von seminaristischen Veranstaltungen und Hands-on-Kursen beschreiben, wobei mit einem seminaristischen Semester begonnen wird. Das erste Semester wurde jedoch als Test aufgesetzt und beide Varianten reduziert aber gleichzeitig angeboten. Die so gewonnenen Erfahrungen und werden im weiteren Verlauf detaillierter beschrieben.

3.2 Zulauf – Studierendenzahlen, Anmeldungen pro Platz:

Im CyberCraft Kolleg wurden im Wintersemester 2022/23 bereits zwei CCK-Kurse prototypisch mit jeweils zehn Plätzen angeboten und konnten mit jeweils 13-14 Teilnehmenden ausgelastet werden. Die initialen Anmeldungen pro Kurs lagen bei ca. 20 Teilnahmebewerber*innen, was ca. 2,0 Anmeldungen pro freien Platz entspricht. Im Sommersemester wurden ebenfalls zwei CCK-Kurse, prototypisch mit jeweils 15 Plätzen, angeboten. Die initialen Anmeldezahlen bei einem der Kurse lagen über 40 Teilnahmebewerber*innen, was 2,6 Anmeldungen pro freien Platz, entspricht. In beiden

Semestern wurden die Kurse der Evaluation unterzogen. Im Wintersemester erzielten die Kurse Spitzenwerte in der Evaluation (Profillinie im Bereich 1-2 auf einer Skala von 1-5) und erhielten gleichzeitig wertvolles Feedback, welches jeweils in Rückkopplungsgesprächen mit den Studierenden verifiziert wurde. Für das Sommersemester stehen die Evaluationen noch aus.



Abb. 3: Studierende im Kolleg fertigen unterstützt durch Augmented Reality handgefertigte Ziegel, um so den durch ein lokales Unternehmen gestellten handwerklichen Anwendungsfall mit neuen Mitteln effizient und digital zu gestalten; Bildautoren: Barlieb und Weininger

Die angebotenen Kurse wurden über die RSDS grundsätzlich allen Fakultäten angeboten und zeigten de facto besonderen Zulauf aus den folgenden Fakultäten: Bauingenieurwesen, Industriedesign, Architektur, Informatik, Elektrotechnik Betriebswirtschaft (insbesondere aus dem Master „Logistik und Organisation“) und der Fakultät „Soziales und Gesundheit“. Die Anzahl der Teilnehmenden aus dem Bereich Informatik und Elektrotechnik liegt bei ca. 25% der Gesamt-Teilnehmendenzahl. Im Wintersemester werden zusätzlich zu den zwei bereits etablierten Kursen drei Projektkurse angeboten, welche in Zusammenarbeit mit der Industrie „hands-on“ drei Handwerks-Anwendungsfälle bearbeiten.

Im Verbundkolleg „digitale Methoden“ meldeten sich 45 Studierende in Bayreuth und 19 Studierende in Hof für die angebotenen Veranstaltungen „Python for non-programmers“, „Latex“ und „Introduction to Deep Learning“ an. Insgesamt handelt es sich um eine starke diversifizierte Personengruppe. So liegt der Frauenanteil je nach Kurs zwischen 25% und 37% und die Studierenden gehören der Informatik, den Ingenieurwissenschaften, den Materialwissenschaften oder auch den Geowissenschaften an. Im aktuellen Semester wird das Angebot deutlich über die offiziellen Anmeldezahlen hinweg genutzt. Die Bereitstellung also Online-Format wird bei den Studierenden gerne angenommen.

3.3 Erwarteter Mehrwert nach Studienende für die Berufs- und Karriereperspektive durch eine Teilnahme an einem Digitalisierungskolleg

Für Studierende im Bauingenieurwesen an der OTH Regensburg gilt: Innovation im Handwerk ist abhängig von einzelnen Innovator*innen. Handwerksbetriebe können sich keine Forschungs- und Entwicklungsabteilungen leisten. Deshalb müssen CCK-Abgänger*innen das gesamte Skillset einer Innovationsabteilung von der Idee bis zur Umsetzung beherrschen. Für Handwerksunternehmen wird so Digital- und Robotik-Innovation erst möglich, für die CCK-Abgänger*innen heißt das, dass sie sehr schnell in leitende, koordinierende Positionen (z.B. die Position der Entwicklungsleitung) treten können und müssen.

Innerhalb dieses Semesters konnte bereits ein erheblicher Mehrwert für alle Teilnehmenden des Verbundkollegs in Bayreuth und Hof festgestellt werden, da im Gegensatz zu Modulhandbüchern und flexible Ausrichtung auf aktuelle Themen ermöglicht wurde.

Nichtinformatiker*innen konnten dadurch digitale Kompetenzen erlangen, die nicht im eigenen Studiengang behandelt werden und Informatiker*innen konnten aktuelle Themen im Bereich KI und Deep Learning behandeln. Dies wurde mit Hilfe von einfachen Selbstevaluierungen zu Beginn und am Ende des Semesters überprüft.

Die Zusammenarbeit in Projektteams von ca. 5 Studierenden ermöglichte den Studierenden einen Aufbau ihrer Sozialkompetenzen und die interdisziplinäre Zusammensetzung der Teams förderte die Einschätzung der Fachkompetenzen und der Planung eines zielgerichteten Einsatzes, z.B. als Spezialist*innen für ein bestimmtes Teilproblem, aber auch die Fachkompetenzübertragung auf andere Teammitglieder. Zusätzlich hatten die Studierenden jederzeit Zugriff auf die Unterstützung eines Coaches, so dass eine zielgerichtete Bearbeitung der Projekte gegeben war.

3.4 Erste Erfahrungen der Projektleiter, Coaches und der Studierenden

In diesem Abschnitt berichten wir über persönliche Rückmeldungen aller Beteiligten. Wie erwartet zeigt sich am CyberCraft Kolleg an der OTH Regensburg, dass der Aufbau der interdisziplinären Kurse und die dabei geforderte Flexibilität (z.B. Eingehen auf einzelnen Studierende, Abstimmung mit den Fakultäten im Vorlauf etc.) signifikant den Aufwand bei fakultätsspezifischen Kursen übersteigt. Mit den Ressourcen und Coaches des bidt-Kollegs ist dieser Aufwand jedoch effizient leistbar. Als zentrales und wichtiges Element wird bei Projektleiter*innen, Coaches und Studierenden gleichermaßen das Arbeiten „hands-on“ und das Arbeiten mit starkem Anwendungsbezug mit den CCK-spezifischen Tools (Roboter, Sensoren, Software etc.) gesehen.

Beim Digitalisierungskolleg „digitale Methoden“ handelt es sich aus Sicht der Mitglieder des Digitalisierungskollegs insgesamt um ein sehr gelungenes Format. Die Inhalte ließen sich sehr schnell, zielgerichtet und flexibel definieren. Die Studierenden nehmen das niederschwellige Angebot sehr gerne wahr. Für die Projektleiter*innen ist insbesondere

auch der Austausch und die Vernetzung mit anderen Digitalisierungskollegs besonders wertvoll. Bereits auf den ersten Treffen konnte ähnliche Problematiken identifiziert diskutiert werden.

Verbesserungspotential besteht insbesondere in der Vergabe von ECTS-Punkten. Da ein Studium bereits auf Vollzeit ausgelegt ist, wurden die Kurse im Digitalisierungskolleg ohne ECTS von den Studierenden niedriger priorisiert. Online zur Verfügung gestellte Lernvideos wiesen jedoch sehr hohe Klickzahlen auf, was auf ein eigenständiges Selbststudium hinweisen könnte. Das Problem der ECTS-Vergabe wurde erkannt und es wird aktuell in beiden Hochschulen gegengesteuert, indem die Kurse in reguläre Studiengänge integriert werden. Für die hochschulübergreifende Vergabe von ECTS werden aktuell u.a. Möglichkeiten über die Hochschule Bayern geprüft.

Ein weiterer offener Punkt der Zusammenarbeit zwischen den Hochschulen betrifft die Verwaltung der Unterrichtsmaterialien. Dies wird aktuell über Moodle Kurse abgewickelt, auf die nur Mitglieder der jeweiligen Hochschule Zugriff haben. Hier wird im Moment eine Möglichkeit zur gemeinsamen Nutzung über Shibolet geschaffen.

Ein weiterer nicht zu unterschätzender Punkt betrifft die unterschiedliche Art der Hochschulen. Während Bayreuth als klassische Universität auftritt, handelt es sich bei Hof um eine Hochschule für angewandte Wissenschaften. Es wurde erkannt, dass für kommende Semester eine entsprechende Vorlaufzeit eingeplant werden muss, um die Kurse synchron zu halten und einen Austausch der Studierenden zu ermöglichen.

Weitere Anmerkungen seitens der Studierenden wurden auch bereits umgesetzt. So wurde z.B. der Latex Kurs wie gewünscht in eine Blockveranstaltung gewandelt und der Unterricht in den weiteren Kursen deutlich häufiger mit kleineren Aufgaben und Challenges durchmischt.

Noch offen ist auch der Punkt, wie eine höhere Beteiligungsquote bei der End-Evaluierung der Studierenden erreicht werden kann. Diese fiel bisher leider deutlich niedriger als zu Beginn aus.

4 Zusammenfassung und Ausblick

Mit den Digitalisierungskollegs wird ein großer Themenbereich der Digitalisierung in der Hochschullehre abgedeckt. Eine Teilnahme ermöglicht den Erwerb umfangreicher digitaler Kompetenzen und kann dadurch den Grundstein für eine wissenschaftliche Karriere im Bereich Digitalisierung legen.

In ersten Evaluationen der Digitalisierungskollegs in Regensburg und Bayreuth/Hof konnte die Wirkung und der positive Effekt bereits bestätigt werden.

So wurden beispielsweise im CyberCraft Kolleg erste Module des Kollegs als interdisziplinäre Projekte zusammen mit Unternehmen und auf deren Anwendungsszenarien hin

durchgeführt und mit konkreten, durch die Studierenden entwickelten innovativen Lösungen abgeschlossen. Das CCK folgt in der Organisation seiner Lehrmodule (Seminare, Tutorials, Projekte etc.) der Entwicklungsmethode V-Modell um ganzheitlich Innovationsprozesse einzuüben welche die Absolvent*innen des Kollegs zu Innovations-treiber*innen bei ihren späteren Arbeitgeber*innen werden lassen.

Im Verbundkolleg der Universität Bayreuth mit der Hochschule Hof werden hochschul-übergreifend grundlegende Kompetenzen für die Durchführung und Dokumentation von KI-getriebenen Projekten im Rahmen von Vorlesungsreihen vermittelt, welche im darauffolgenden Semester in interdisziplinären Gruppen in verschiedenen Domänen angewendet werden.

Die Teilnahme an weiteren Förderprogrammen zu Digitalisierungsthemen ist dadurch leichter möglich. Durch den Besuch erwerben auch Studierende aus digitalisierungsfernen Fächern frühzeitig umfangreiche IT-Kenntnisse und ein digitales Mindset für zukünftige digitale Themenbereiche. Als Begleiteffekt der umfangreichen Vernetzung zwischen den Digitalisierungskollegs (Studierende, Coaches und Projektleitende) entsteht bereits zu Beginn einer wissenschaftlichen Karriere eine große digitale Community, die durch die Koordination und eine gezielte Vernetzung der Studierenden, Coaches und Projektleitungen unterstützt wird. Alle Teilnehmenden wie auch die Mitarbeitenden lernen frühzeitig die interdisziplinäre Zusammenarbeit und verbessern erheblich ihre Karriere-chancen innerhalb und außerhalb der Wissenschaft.

Im Rahmen der Nachhaltigkeit werden die Kollegs verstetigt. So wird in Regensburg beispielsweise auf einen neuen hoch-interdisziplinären Master an der Schnittstelle von Bauwesen und anderen Disziplinen hingearbeitet der als Gesamtprogramm als auch in Light-Versionen (Zertifikatskurs, Weiterbildung) verfügbar sein soll. In Bayreuth und Hof werden die Inhalte im Rahmen von Master-Studiengängen verstetigt, welche sich jeweils an Bachelorabsolvent*innen im Bereich der Informatik und der Ingenieurwissenschaften richten.

Im zukünftigen Verlauf des Projektes werden Evaluierungen durchgeführt und die wesentlichen Erfolgsfaktoren und auch mögliche Schwachstellen identifiziert. Diese Ergebnisse sollen in einem Folgepapier adressiert werden, welches einen vertieften Einblick in die Methoden geben soll und auch als Handreichung zur Nachahmung dienen soll.

Innerhalb der Digitalisierungscommunity sind zudem Folgeprojekte zu erwarten und sehr wahrscheinlich werden vermehrt weitere Drittmittel eingeworben. Diese Prognose stützt sich auf die Beobachtung aus dem Vorläuferformat Innovationslabore (2017 – 2021): dort entstanden in der Projektlaufzeit zahlreiche weitere Anträge und Projekte.

- Insgesamt 23 Drittmittelanträge: am Innovationslabor Würzburg 15.
- Eine Vielzahl an Unternehmensgründungen: Gesamt: 31, v.a. an den Universitäten Bayreuth und FAU Erlangen-Nürnberg, darunter mehrere erfolgreiche Anträge auf Gründungsförderung (z.B. EXiST, Flüge-Bayern).

Darüber hinaus bildeten sich weitere Forschungs- und Kooperationsprojekte und zahlreiche daraus hervorgehende Anträge wurden gestellt.

5 Danksagungen

Mit herzlichem Dank an die Mitverfasser Christophe Barlieb, Christian Groth, Thomas Linner und Florian Weininger für die prototypischen Erfahrungsberichte aus den beiden Projekten in Regensburg, Bayreuth und Hof. Vielen herzlichen Dank an alle (weiteren) Digitalisierungskollegen so wie die Coaches und Koordinatoren welche derzeit mit großem Engagement daran arbeiten die in diesem Beitrage beschriebenen Methoden in vielfältigen Bereichen in der Lehre zu verankern. Eine wissenschaftliche Auswertung der Ergebnisse aller Kollegen erfolgt in weiteren Beiträgen.

6 Quellen- und Literaturverzeichnis

- [Au20] Ausschreibung für das Forschungsprogramm der deutschen Akademien der Wissenschaften (Akademienprogramm) für 2023, https://www.hadw-bw.de/sites/default/files/documents/2023_Ausschreibung_Akademienprogramm.pdf, Stand: 04.07.2023.
- [Br19] Brinda, T. et.al.: Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt. Ein interdisziplinäres Modell. In: (Pasternak, A., Hrsg.): Informatik für alle, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn, S. 25-33, 2019.
- [Da16] Dagstuhl-Erklärung. Bildung in der digitalen vernetzten Welt, https://gi.de/fileadmin/GI/Hauptseite/Themen/Dagstuhl-Erklärung_2016-03-23.pdf, Stand: 04.07.2023.
- [EMT17] Eichhorn, M.; Müller, R.; Tillmann, A.: Entwicklung eines Kompetenzrasters zur Erfassung der "Digitalen Kompetenz" von Hochschullehrenden. In (Igel, C., Hrsg.): Bildungsräume. Proceedings der 25. Jahrestagung der Gesellschaft für Medien in der Wissenschaft, 5. bis 8. September 2017 in Chemnitz. Münster; New York: Waxmann, S. 209-219, 2017.
- [Eu19] Europäischer Rahmen für die digitale Kompetenz Lehrender (DigCompEdu) https://joint-research-centre.ec.europa.eu/system/files/2019-09/digcompedu_german_final.pdf, Stand: 04.07.2023.
- [PBB18] Porat, E.; Blau, I.; Barak, A.: Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. Computers & Education, 126, S. 23-36, 2018.
- [Wi20] Wissenschaft im Spannungsfeld von Disziplinarität und Interdisziplinarität - Positionspapier, <https://www.wissenschaftsrat.de/download/2020/8694-20.pdf?blob=publicationFile&v=5>, Stand: 04.07.2023.

SIENA: Sprachmodellbasierte Identifikation und Extraktion von Nutzeranforderungen

Eine innovative Methode zur automatisierten Anforderungsanalyse

Dorian Zwanzig¹, Anja Kahl² und Prof. Dr.-Ing. Ute Dietrich³

Abstract: In diesem Artikel wird eine innovative Methode zur Identifikation und Extraktion von Nutzeranforderungen aus natürlichsprachlichen Quellen präsentiert. Als Ausgangsbasis dient uns eine Fallstudie zur Entwicklung einer Fachanwendung für Physiotherapeuten. Die Methode nutzt OpenAI's Generative Pretrained Transformer (GPT) Modelle und deren Fähigkeit zur Verarbeitung natürlicher Sprache. Eine quantitative Analyse wurde durchgeführt, um die Wirksamkeit dieser Sprachmodelle bei der Anforderungsanalyse zu bewerten. Die Ergebnisse zeigen deutlich, dass die verwendeten GPT-Modelle eine effektive und kostengünstige Unterstützung bei der Anforderungsanalyse sein können.

Keywords: Anforderungsanalyse, Nutzeranforderungen, Künstliche Intelligenz, Sprachmodell, ChatGPT, Large Language Model, Natürliche Sprache

1 Einführung

Die Identifikation und Extraktion von Nutzeranforderungen stellt einen kritischen und zeitaufwendigen Prozess im Anforderungsmanagement dar, der häufig von speziell geschulten Fachexperten durchgeführt wird und mit erheblichem betriebswirtschaftlichem Aufwand verbunden ist. Ohne eine klare, eindeutige Kenntnis der zu erreichenden Zielstellung wird das Ergebnis an den Bedürfnissen der Stakeholder vorbeigehen und mühevoll und aufwendig nachgebessert werden müssen. In den letzten Jahren haben Large Language Models (LLM) bemerkenswerte Fortschritte in der Analyse und Interpretation von textbasierten Inhalten gezeigt, wie etwa in der Sentiment-Analyse, der Erstellung von Zusammenfassungen oder der Erweiterung von Texten. Angesichts dieser Entwicklungen liegt der Fokus dieser Studie auf einer Untersuchung, in wie weit LLM, insbesondere GPT-3.5 und GPT-4, eingesetzt werden können, um den manuellen Aufwand im Bereich

¹ Hochschule für Technik und Wirtschaft (HTW) Berlin, Fachbereich 4 Informatik, Kommunikation und Wirtschaft – Studiengang Wirtschaftsingenieurwesen, Wilhelmshofstraße 75A, 12459 Berlin, dorian.zwanzig@htw-berlin.de

² Hochschule für Technik und Wirtschaft (HTW) Berlin, Fachbereich 4 Informatik, Kommunikation und Wirtschaft – Studiengang Wirtschaftsingenieurwesen, Wilhelmshofstraße 75A, 12459 Berlin, anja.kahl@student.htw-berlin.de

³ Hochschule für Technik und Wirtschaft (HTW) Berlin, Fachbereich 4 Informatik, Kommunikation und Wirtschaft – Studiengang Wirtschaftsingenieurwesen, Wilhelmshofstraße 75A, 12459 Berlin, ute.dietrich@htw-berlin.de

der Anforderungserhebung zu reduzieren und diesen Prozess zumindest teilweise zu automatisieren.

Die zentrale Forschungsfrage lautet daher: "Inwiefern können Large Language Models, wie GPT-3.5 und GPT-4, effektiv zur Identifikation und Extraktion von Nutzeranforderungen aus natürlichsprachlichen Quellen eingesetzt werden?"

Um diese Frage zu beantworten, wird ein Entwicklungsprojekt für eine Anwendung für Physiotherapeuten als Fallstudie herangezogen. Im Rahmen dieses Projekts wurden semi-strukturierte Interviews mit potenziellen NutzerInnen durchgeführt, um das Nutzerumfeld zu verstehen und Lösungsideen zu sammeln. Die in den Interviews geäußerten Wünsche bezüglich der bereitzustellenden Funktionalitäten und Inhalte wurden vom Projektteam analysiert und in Nutzeranforderungen übersetzt, die als Baseline für den Vergleich in dieser Studie dienen.

Nach Anonymisierung und Freigabe durch die InterviewpartnerInnen wurden relevante Abschnitte aus den Interviews extrahiert und über die OpenAI-API mit den Modellen GPT3.5-turbo-0301 und GPT-4-0314 verarbeitet. Anschließend wurden die extrahierten Anforderungen hinsichtlich ihrer Relevanz, Übereinstimmung mit der Baseline und Qualität analysiert und strukturiert. Abschließend wurden beide Modelle hinsichtlich ihrer Leistungsfähigkeit verglichen um festzustellen, ob und inwieweit LLM zur Identifikation, Extraktion und auch Strukturierung von Nutzeranforderungen beitragen können.

2 Stand der Technik

Natural Language Processing (NLP) im Bereich des Anforderungsmanagement, im Englischen des Requirement Engineerings (RE), ist ein aktives und sich entwickelndes Forschungsfeld. Bis 2019 wurden 404 Studien zu diesem Thema durchgeführt, von denen 67 % Lösungsvorschläge enthielten. Dabei bezogen sich 59 % auf die Analyse und Erhebung von Anforderungen [Zh21]. Aufgrund neuer Techniken wächst das Interesse der RE-Forscher in diesem Bereich. [FZA21] KI-basierte Tools verwenden NLP und Machine Learning (ML) hauptsächlich zur Anforderungsklassifikation, um informelle Texte in formell strukturierte Anforderungen umzuwandeln und zur Anforderungserhebung. [My21; DN20; Be22; Su19]

Aktuelle Ansätze zur Anforderungsextraktion befassen sich mit der Verwendung des BERT-Modells für RE, ChatGPT als Vertreter für Large Language Models und der Entwicklung einer Softwarelösung zur Identifikation von Wissen aus kausalen Zusammenhängen [Zh23; AM21; Fi20]. Hauptfragestellungen im Bereich RE mithilfe von NLP beinhalten die Automatisierung der Anforderungsklassifikation und die automatische Extraktion kausaler Beziehungen aus natürlicher Sprache. Zusätzlich sollen Fragen zur automatischen Generierung von Anforderungen aus Applikations-Reviews und Nutzerrezensionen inklusive der Erkennung und Vermeidung von Mehrdeutigkeiten und /oder Widersprüchlichkeiten beantwortet, Probleme bei der Erstellung formeller

Anforderungen sowie die Beeinflussung der Forschung und Praxis des NLP für das RE durch generative LLMs in den Betrachtungsfokus gerückt werden. [Fi20; DN20; Su19; AM21; Be22]

Aktuell konzentriert sich die durch NLP unterstützte Extraktion von Anforderungen hauptsächlich auf nutzergenerierte Inhalte und Anforderungsspezifikationen. Lediglich vier der bis 2019 veröffentlichten Studien beinhalteten Interview-Transkripte als Eingabedokumente. [Zh21] Die Hauptprobleme, die weitere Forschung erfordern, sind Informalität, Individualität und die in den Texten enthaltenen Rechtschreib- und Grammatikfehler; das Herausfiltern ausschließlich der für den Entwickler relevanten Informationen und der teilweise fehlende Bestand an Datensätzen, um Entwicklungen testen zu können. Insbesondere Mehrdeutigkeiten und der Gebrauch von Synonymen und Homonymen in Aussagen führen zu geringerer Leistung und Fehlern, was einen hohen Aufwand an Übersetzung in formelle Strukturen erfordert. [Be22; Su19; Z23; AM21] Der leistungsstarke ChatGPT im Bereich NLP weist nur begrenztes Wissen im Bereich RE auf. Die dazugehörige empirische Evaluation ergab jedoch, dass in der Nutzung von LLMs für Anforderungsabrufe Potenzial liegt. [Zh23] Eine grundlegende Hürde besteht aktuell darin, dass für alle Sprachmodelle noch menschliche Überwachung erforderlich ist. [Be22]

Angesichts dieser Herausforderungen und Möglichkeiten sollte zukünftige Forschung darauf abzielen, die Leistungsfähigkeit und Anwendbarkeit von NLP- und LLM-Techniken im Bereich RE weiter zu verbessern. Dazu gehört die Untersuchung neuer Ansätze zur Reduzierung von Mehrdeutigkeiten, die Optimierung von Eingabedaten und die Anpassung von Sprachmodellen an RE-spezifische Aufgaben. Darüber hinaus sollte die Forschung auch die Integration verschiedener NLP- und LLM-Tools in bestehende RE-Methoden und -Prozesse in Betracht ziehen, um Synergien zu nutzen und die Effektivität der Anforderungserhebung und -analyse insgesamt zu erhöhen.

3 Methode

Die semi-strukturierten Interviews wurden anhand eines thematisch abgestimmten Leitfadens durchgeführt, der drei Hauptbereiche fokussierte: das Nutzerumfeld inklusive typischer Tätigkeiten und beruflicher Herausforderungen, den Wissenserwerb im beruflichen Kontext mit besonderem Fokus auf medizinischen Leitlinien sowie Ideation d.h. die Entwicklung von Lösungsansätzen für eine digitale Wissensbereitstellung. Für die hier diskutierte Studie ist vor allem der dritte Teil relevant. Ausgewählte Interviewausschnitte zeichneten sich durch ihre Anforderungsdichte aus, das heißt, sie enthielten eine Vielzahl von Anforderungen und konnten auch unabhängig von weiterem Kontext verarbeitet werden. Drei Interviewpartner gaben ihre Zustimmung zur Nutzung ihrer Interviews in dieser Studie und durch Drittsysteme, hier OpenAI, wodurch eine Datenbasis von 16 Interviewabschnitten mit insgesamt etwa 5.000 Wörtern entstand. Jeder Abschnitt umfasste typischerweise ein Frage-Antwort-Paar. Insgesamt wurden 46

Nutzeranforderungen manuell aus diesen Ausschnitten extrahiert, die als Baseline für die Beurteilung der Leistungsfähigkeit der eingesetzten Sprachmodelle dienen.

Die Interviewausschnitte enthielten primär Softwareanforderungen aus der NutzerInnen-Perspektive. Es sollte eine mobile Anwendung entwickelt werden, die es PhysiotherapeutInnen ermöglicht, innerhalb kurzer Zeit auf evidenzbasierte Behandlungsempfehlungen zurückzugreifen. Dabei wurden in der manuellen Analyse unter anderem folgende Anforderungen identifiziert:

- Die Anwendung muss die Eingabe von Diagnosen zur Suche nach Behandlungen ermöglichen.
- Die Anwendung soll eine kurze Übersicht über die Diagnose und Behandlung bereitstellen.
- Die Anwendung soll Informationen als Fließtext bereitstellen.

In einem iterativen Prozess wurde ein ausführlicher Prompt entwickelt, der als Eingabe für die Modelle diente. Bei einem Prompt handelt es sich im Grunde um eine Anweisung in Textform an den ChatGPT oder ein vergleichbares KI-System. Der im Projekt entwickelte Prompt bestand aus zwei Teilen: einem zur Definition des Systems und seiner Aufgabe und einem zur Bereitstellung des Interviewausschnitts. Der erste Teil definierte den Kontext der Aufgabe, hier die Anwendungsentwicklung für Physiotherapeuten, die Rolle des Systems als Expertensystem, dessen Aufgabe, die auf eine Identifikation und Extraktion von strukturierten Anforderungen fokussiert sowie das Ausgabeformat. Der zweite Teil enthielt den jeweiligen Interviewausschnitt.

Im folgenden ist der Systemprompt dargelegt: „Du bist ein System, das Anforderungen in natürlich sprachlichen Interviews identifiziert und extrahiert. Dir wird vom ""user"" ein Interviewausschnitt präsentiert. In dem Interview geht es um die Entwicklung einer Anwendung zur Bereitstellung von medizinischem Fachwissen für PhysiotherapeutInnen. Der Ausschnitt startet mit einer Frage durch einen Interviewer "I". Der Interviewer stellt eine oder mehrere Fragen an den Interviewpartner "B". Der Interviewpartner "B" antwortet auf die Frage. Du analysierst den Interviewausschnitt und versuchst, direkte Hinweise auf die vom Interviewpartner "B" gewünschten Funktionen der Anwendung zu identifizieren. Ignoriere Anforderungen, die nur aus der Aussage des Interviewers "I" resultieren. Du gibst anschließend eine Liste von Anforderungen aus, die du aus dem Interviewausschnitt extrahiert hast. Die Anforderungen sollen in funktionale und nicht-funktionale Anforderungen unterteilt werden. Du kannst auch eine Anforderung als unklar markieren, wenn du dir nicht sicher bist, ob sie eine Anforderung ist. Es sollen keine impliziten Anforderungen ausgegeben werden. Jede Anforderung soll mit einem Zitat aus dem Interviewausschnitt belegt werden. Die Anforderungen soll nach folgendem Schema formuliert werden: Das System muss/kann/soll + <Anforderung>. Gib das Ergebnis deiner Analyse als CSV aus, nutze als Trennzeichen ";" und folgende Struktur: Funktional/nicht funktional/unklar; Anforderung; Beschreibung der Anforderung; Zitat aus Interviewausschnitt.“

Die Interviewausschnitte wurden sequenziell durch die Modelle verarbeitet und die generierten Ausgaben in einer Gesamtliste zusammengefasst. Die Modelle wurden lediglich hinsichtlich des Parameters "Temperature" konfiguriert, welcher den Grad der Zufälligkeit in der Generierung der Ausgaben regelt. Dieser wurde auf 0 gesetzt, um eine Reproduzierbarkeit zu gewährleisten.

Der vollständige Prompt, die Interviewausschnitte und die generierten Anforderungen sind im frei zugänglichen und wiederverwendbaren Projekt-Repository unter <https://github.com/dozwa/siena-001> einsehbar und können frei für eigene Experimente verwendet werden.

Die extrahierten Anforderungen wurden in einem nachfolgenden Schritt manuell auf Relevanz überprüft und Duplikate entfernt. Anschließend wurden die generierten Anforderungen mit der Baseline abgeglichen und die Übereinstimmung quantitativ analysiert.

GPT-3.5 und GPT-4 wurden für diese Studie aufgrund ihrer einfachen Verfügbarkeit und Anwendbarkeit ausgewählt. Die leichte Zugänglichkeit macht diese Modelle besonders relevant für Anwendungen im betrieblichen Kontext. Die umfangreichen Dokumentationen ermöglichen es zudem, diese Modelle mit relativ wenig technischem Wissen zu verwenden. Hier sind lediglich grundlegende Erfahrungen in der Programmierung mit Python erforderlich. Daher sind GPT-3.5 und GPT-4 für den praktischen Einsatz besonders geeignet.

In dieser Studie wurden die statischen Varianten der Modelle, gpt-3.5-turbo-0301 und gpt-4-0314, verwendet. Diese Modelle wurden seit ihrer Veröffentlichung nicht mehr aktualisiert, um die Wiederholbarkeit der Versuche zu gewährleisten und zu verhindern, dass das Modell aus der iterativen Durchführung der Prompt-Entwicklung lernt.

Die abschließende quantitative Analyse basierte auf der Ermittlung verschiedener Werte. Diese umfassen:

- True Positives (TP): Anzahl der korrekt identifizierten relevanten Anforderungen (in der Baseline enthalten und vom System erkannt).
- False Positives (FP): Anzahl der fälschlicherweise als relevant identifizierten Anforderungen (nicht in der Baseline enthalten, aber vom System erkannt).
- False Negatives (FN): Anzahl der relevanten Anforderungen, die vom System nicht erkannt wurden (in der Baseline enthalten, aber vom System nicht erkannt).
- True Negatives (TN): Anzahl der korrekt identifizierten, irrelevanten Anforderungen (nicht in der Baseline enthalten und vom System nicht erkannt).

Anhand dieser Werte wurden folgende Kennzahlen berechnet:

- Genauigkeit (Accuracy): Anteil der korrekt klassifizierten Anforderungen (sowohl relevant als auch irrelevant) an der Gesamtzahl der Anforderungen. Berechnet durch: $(TP + TN) / (TP + FP + FN + TN)$
- Präzision (Precision): Anteil der korrekt identifizierten, relevanten Anforderungen an der Gesamtzahl der vom System als relevant identifizierten Anforderungen. Berechnet durch: $TP / (TP + FP)$
- Sensitivität (Recall): Anteil der korrekt identifizierten, relevanten Anforderungen an der Gesamtzahl der tatsächlich relevanten Anforderungen (Baseline). Berechnet durch: $TP / (TP + FN)$
- F1-Wert (F1-Score): Harmonisches Mittel aus Präzision und Sensitivität, das ein ausgewogenes Maß für die Leistung des Systems bietet, insbesondere bei ungleicher Verteilung der relevanten und irrelevanten Anforderungen. Berechnet durch: $2 * (\text{Präzision} * \text{Sensitivität}) / (\text{Präzision} + \text{Sensitivität})$

4 Ergebnisse

Das GPT-3.5-Modell identifizierte und extrahierte insgesamt 47 relevante Anforderungen aus den Interviews. Davon stimmten 33 (True Positive) genau oder zumindest annähernd mit den Anforderungen aus der Baseline überein. Es ist zu beachten, dass mehrere generierte Anforderungen einer Baseline-Anforderung zugeordnet sein können. Insgesamt konnte das Modell 56% der Baseline-Anforderungen identifizieren. Darüber hinaus wurden 12 relevante Anforderungen erkannt, die nicht in der Baseline enthalten waren (False Positive). Diese waren den Analysten entweder nicht aufgefallen oder wurden nicht als relevant eingestuft. Das Modell schuf insofern einen Mehrwert, als es bisher unbekannte Anforderungen aufdecken konnte. Für 15 Baseline-Anforderungen wurden keine Entsprechungen unter den generierten Anforderungen gefunden (False Negative), und eine generierte Anforderung war völlig irrelevant (True Negative). Es lässt sich feststellen, dass etwa $\frac{3}{4}$ der vom System generierten Anforderungen relevant waren, das heißt, sie wurden korrekt aus den Interviews extrahiert.

Das GPT-4-Modell identifizierte mit 58 Anforderungen deutlich mehr Anforderungen, von denen 39 direkt den Baseline-Anforderungen zugeordnet werden konnten. Die Anzahl der False Positives und False Negatives lag in etwa im Bereich des Vergleichsmodells. Ebenso vergleichbar sind Genauigkeit, Präzision und Sensitivität. Demgegenüber stehen drei irrelevante Anforderungen, die keiner Baseline-Anforderung entsprechen. Dies entspricht der dreifachen Menge an als True Negative klassifizierten Anforderungen im Vergleich zum GPT-3.5-Modell.

Kennzahl	GPT3.5	GPT4
Baseline Anforderungen	46	46
Generierte Anforderungen	47	58
True Positives	33	39
False Positives	12	13
False Negatives	15	17
True Negatives	1	3
Accuracy	56 %	58 %
Precision	73%	75%
Recall	69%	70%
F1-Score	71%	72%

Tab. 1: Vergleich GPT3.5 und GPT4

Die Interpretation der Ergebnisse wird durch zwei Faktoren erschwert: (1) Die Baseline umfasst nicht alle möglichen Anforderungen, und die Modelle haben Anforderungen identifiziert, die nicht in der Baseline enthalten sind. (2) Die Zuordnung der Anforderungen weist eine Kardinalität von 1-n auf, sodass mehrere generierte Anforderungen einer Baseline-Anforderung zugeordnet werden können. Um diese Probleme zu lösen, wurden zwei neue modellspezifische Baselines erstellt, welche um die zusätzlichen, vom System erkannten relevanten Anforderungen erweitert wurden. Zudem wurden generierte Anforderungen mit einer Kardinalität ungleich 1-1 von der Analyse ausgeschlossen.

Nach dieser Bereinigung weisen beide Modelle erwartungsgemäß eine Präzision von 100% auf, das bedeutet, dass alle generierten Anforderungen Teil der Baseline sind. Die Genauigkeit des GPT-4-Modells liegt mit 74% etwas höher als die des GPT-3.5-Modells.

Kennzahl	GPT3.5	GPT4
Baseline Anforderungen	58	59
Generierte Anforderungen	47	58
True Positives	41	43
False Positives	0	0
False Negatives	17	16
True Negatives	1	3
Accuracy	71%	74%
Precision	100%	100%
Recall	71%	73%
F1-Score	83%	84%

Tab. 2: Vergleich GPT3.5 und GPT4 mit erweiterter Baseline:

5 Diskussion

In der quantitativen Analyse wurde gezeigt, dass die beiden Modelle nur geringfügige Unterschiede bei der Erkennung von Anforderungen aufweisen. Dies erscheint zunächst kontraintuitiv, da das GPT-4-Modell deutlich komplexer ist als das GPT-3.5-Modell. Stand Mai 2022 ist die Anwendung des GPT-4-Modells etwa 30-mal teurer als die des GPT-3.5-Modells, was aus wirtschaftlicher Sicht eine Nutzung des GPT-3.5-Modells empfehlenswert erscheinen lässt. Die Kosten für die einfache Verarbeitung der 16 Interviewausschnitte beliefen sich im Juni 2023 unter Verwendung des GPT3.5 Modells auf ca. 4 Euro Cent und mit dem GPT4 Modell auf 86 Euro Cent. Der Umfang der Anfrage belief sich je Interviewausschnitt auf ca. 1000 Token (40% Systemprompt, 60% Interviewausschnitt), die Antworten umfassten durchschnittlich ca. 400 Token. Die gesamte Verarbeitung der Interviews umfasste insgesamt ca. 22.400 Token. Die tatsächlichen API-Kosten für die gesamte Studie betragen jedoch weniger als 5 Euro. Eine wirtschaftliche Betrachtung wäre dennoch bei umfangreicheren Anforderungsanalysen sinnvoll. Würde man das hier diskutierte Vorgehen auf ganze Interviews übertragen, würde die Analyse eines ca. einstündigen Interviews mit einem Transkript von ca. 9.000 Wörtern (ca. 17.000 Token) etwa 1,12 € (GPT4) bzw. 0,05 € (GPT3.5) kosten.

Prompt	Token gesamt	Kommentar	Kosten GPT3.5 je 1000 Token	Kosten GPT4 je 1000 Token
Systemprompt	6.800	40% des Interviews	0,0015 €	0,0300 €
Interviewausschnitt	17.000	1 Std., 9.000 Worte	0,0015 €	0,0300 €
Antwort	6800	40% des Interviews	0,0020 €	0,0600 €
Gesamtkosten			0,0493 €	1,1220 €

Tab. 3: Kostenkalkulation für die Analyse eines einstündigen Interviews

Generell scheint der Einsatz von Large Language Modellen, hier am Beispiel der Modelle von OpenAI, für die Identifikation und Extraktion von Nutzeranforderungen sinnvoll zu sein. Besonders hervorzuheben ist, dass die Modelle in der Lage waren, Anforderungen aufzudecken, die in der manuellen Analyse nicht korrekt herausgearbeitet wurden.

Zukünftige Untersuchungen könnten bereits von einer erweiterten Baseline ausgehen, die auch auf den Ergebnissen aller berücksichtigter Modelle und den manuell identifizierten Anforderungen basieren. Dies würde eine noch bessere Vergleichbarkeit gewährleisten.

Der in dieser Studie verwendete Prompt erwies sich als zielführend, jedoch bedeutet dies nicht, dass er optimal gestaltet ist. Es handelt sich um einen Zero-Shot-Prompt, d.h., dieser ist auf andere Aufgaben übertragbar, ohne das zusätzliche markierte Trainingsbeispiele benötigt werden. Lediglich die Aufgabe muss entsprechend beschrieben werden. Alternativ hätte man einen Multi-Shot-Prompt verwenden können, der neben der Anweisung auch konkrete Beispiele für die Lösung der Aufgabe enthält.

Während in dieser Studie ausschließlich quantitative Aspekte berücksichtigt wurden, sollten zukünftig auch qualitative Aspekte systematisch untersucht werden. Die Qualität der generierten Anforderungen variiert erheblich. Einige Anforderungen wurden in hervorragender Qualität formuliert und berücksichtigen alle relevanten Aspekte einer Anforderungsdefinition. Andere Anforderungen bestanden jedoch aus lediglich zwei Wörtern, ließen viel Interpretationsspielraum und/oder waren strukturell inkorrekt aufgebaut.

Es sollte auch beachtet werden, dass die Modelle die Anweisung hatten, ausschließlich explizit benannte Nutzerwünsche zu berücksichtigen. Die Nutzung der Modelle zur Identifikation von impliziten Anforderungen oder Basis-Anforderungen, z.B. nach dem Kano-Modell, könnte zusätzliche interessante Erkenntnisse und Einblicke liefern.

Ein weiterer diskussionswürdiger Aspekt betrifft die Auswahl der Modelle. Die getroffene Auswahl basiert auf Zugänglichkeit und Verfügbarkeit, ohne umfassende Marktrecherche. Zukünftige Untersuchungen könnten mehr Modelle, insbesondere Modelle unterschiedlicher Anbieter sowie Open-Source-Lösungen, berücksichtigen. Es wäre auch

interessant zu untersuchen, ob ein Open-Source-Modell speziell für diese Aufgabe nachtrainiert werden könnte.

6 Zusammenfassung

Die Untersuchung hat gezeigt, dass der Einsatz der betrachteten Modelle zur Identifikation und Extraktion von Nutzeranforderungen aus natürlichsprachlichen Quellen vielversprechend ist. Beide analysierten Modelle konnten einen Großteil der relevanten Anforderungen aus den Texten extrahieren, wobei das GPT-4-Modell dem weniger komplexen GPT-3.5-Modell leicht überlegen ist. Hervorgehoben werden muss, dass beide Modelle neue, relevante Anforderungen identifiziert haben. Dennoch sollte die Quantität nicht das einzige Kriterium sein, anhand dessen die Leistung gemessen wird. Eine gesonderte Betrachtung der Qualität der Anforderungen wird dringend empfohlen.

Die Autoren sind der Meinung, dass dieser Ansatz positive praktische Auswirkungen haben kann, indem er eine sinnvolle Ergänzung zur manuellen Anforderungsanalyse darstellt. Bei einer exponentiell ansteigenden Anforderungsanzahl komplexer Produkte ist eine manuelle Erhebung schlicht kaum noch zu gewährleisten. Möglicherweise könnten die Modelle auch bei der Identifikation von impliziten Anforderungen nützlich sein. Wirtschaftliche Aspekte sollten bei der Auswahl der Modelle berücksichtigt werden, da die hier betrachteten Modelle signifikant unterschiedliche Kosten verursachen.

Die Leistungsfähigkeit der Modelle könnte durch einen umfangreicheren Modellvergleich weiter objektiviert werden. Durch gezieltes Nach-Training der Modelle oder eine Optimierung der Prompts könnte ihre Leistungsfähigkeit zusätzlich verbessert werden.

Literaturverzeichnis

- [AM21] Araújo, A.; Marcacini, R.: RE-BERT: Automatic Extraction of Software Requirements from App Reviews using BERT Language Model. In (Association for Computing Machinery): SAC'21: Proceedings of the 36th Annual ACM Symposium on Applied Computing, S. 1321-1327, 2021.
- [Be22] Bertram, V. et.al.: Neural Language Models and Few Shot Learning for Systematic Requirements Processing in MDSE. In (Association for Computing Machinery): SLE 2022: Proceedings of the 15th ACM SIGPLAN International Conference on Software Language Engineering. Auckland, S. 260-265, 2022.
- [DN20] Dalpiaz, F.; Niu, N.: Requirements Engineering in the Days of Artificial Intelligence. IEEE Software, Band 37, Ausgabe 4, S. 7-11, 2020.
- [Fi20] Fischbach, J. et.al.: Towards Causality Extraction from Requirements. In (IEEE): 2020 IEEE 28th International Requirements Engineering Conference (RE). Zürich, S. 388-393, 2020.

- [FZA21] Ferrari, A.; Zhao L., Alhoshan W.: NLP for Requirements Engineering: Tasks, Techniques, Tools, and Technologies. In (IEEE): 2021 IEEE/ACM 43rd International Conference on Software Engineering: Companion Proceedings (ICSE-Companion). Madrid, S. 322-323., 2021.
- [My21] Myllynen, S. et.al.: Developing and Implementing Artificial Intelligence-Based Classifier for Requirments Engineering. ASME J of Nuclear Rad Sci. Band 7, Ausgabe 3, S. 1, 2021.
- [Su19] Surana, C. et.al.: Intelligent Chatbot for Requirements Elicitation and Classification. In (IEEE): 2019 4th International Conference on Recent Traneds on Electronics, Information, Communication & Technology (RTEICT). Bangalore, S. 866-870, 2019.
- [Zh21] Zhao, L. et al.: Natural Language Processing for Requirements Engineering: A Systematic Mapping Study. ACM Computing Survey, Band 54, Ausgabe 3, Art. 55, S. 1-41, 2021.
- [Zh23] Zhang, J. et.al.: A Preliminary Evaluation of ChatGPT in Requirements Information Retrieval. 2023.

Visualization Support for Contracts in VeriFast

Pauline Hergersberg¹ Judith Lippold² Bastian Nahrstedt³ Thomas Baar⁴

Abstract: A widespread quality assurance technique for ensuring correctness of software is testing, but relevant test cases might be easily overlooked. Formal methods - an alternative to testing - are applied in software industry only rarely, due to the lack of both widely used verification tools and engineers able to apply such tools effectively. In order to address these problems, our university offers appropriate courses, in which the open-source tool VeriFast is applied to formally verify given contracts for functions implemented in C. While VeriFast is very fast in verifying even larger programs, students often have difficulties when authoring function contracts and other proof arguments formally. To address this problem, we developed a web-based system for visualizing each contract as a graph. In this paper, we describe the architecture and the main features of our system and show on a running example, how our system can support the user of VeriFast.

Keywords: Deductive Verification; Visualization; TripleGraphGrammar (TGG)

1 Introduction

Software has assumed a central and indispensable position in the architecture of virtually all contemporary systems. Especially in safety critical sectors like the automotive industry and spacecraft, software must undergo stringent quality procedures [GLH18] in order to achieve a high level of reliability. A central role in this procedure play unit tests, which systematically evaluate the response of functions across a diverse range of input parameters. However, according to Dijkstra, "program testing can reveal the presence of bugs but cannot guarantee their absence." Hence, despite the usefulness of testing, there is always a possibility of overlooking certain bugs [DDH72].

In some cases, an alternative option is the usage of formal methods in order to prove the correctness of software systems. One widespread formal method technique is *model checking* [CES09, Ho05] in which the system to be verified has to be described in form of its state space. Furthermore, the user has to formulate the desired properties of the system

¹ Hochschule für Technik und Wirtschaft (HTW) Berlin, Department of Engineering I, Wilhelminenhofstraße 75A, 12459 Berlin, Germany Pauline.Hergersberg@Student.HTW-Berlin.de

² Hochschule für Technik und Wirtschaft (HTW) Berlin, Department of Engineering I, Wilhelminenhofstraße 75A, 12459 Berlin, Germany Judith.Lippold@gmx.net

³ Hochschule für Technik und Wirtschaft (HTW) Berlin, Department of Engineering I, Wilhelminenhofstraße 75A, 12459 Berlin, Germany Bastian.Nahrstedt@Student.HTW-Berlin.de

⁴ Hochschule für Technik und Wirtschaft (HTW) Berlin, Department of Engineering I, Wilhelminenhofstraße 75A, 12459 Berlin, Germany Thomas.Baar@HTW-Berlin.de

in form of temporal logic formulas [Pn77, Em90]. In spite of its widespread use, model checking is often not suitable for systems having a large or irregular state space.

A second formal method technique is *deductive software verification* [HH19, Ah05]. Here, implemented functions are annotated with formal contracts following the philosophy of *Design by Contracts* [Me92], describing their expected behavior in terms of pre- and post-conditions. These pre-/post-conditions are expressed as formulas representing a potentially infinite set of program states. Once it is verified that the implementation code of a function obeys the annotated contract, the correct behaviour of the verified code is guaranteed under all circumstances. However, the adoption of this approach in industry is limited due to several reasons, most notably to shortcomings of available verification tools and the high level of expertise required to apply these tools successfully [Fi11].

One powerful deductive verification system is VeriFast [JVP15], introduced by Bart Jacobs, aiming at verifying programming code written in C⁵. Verifying code with VeriFast means to verify every single function separately, i.e. the whole program is verified in a modular way.

VeriFast has been already applied for educational purposes at our university for a couple of years. When teaching deductive verification, one observation is that students often struggle with annotating C code with formal contracts, what is necessary for VeriFast to verify the code. Very soon, the contracts annotated by students can become large, confusing, and erroneous.

The aim of our work is therefore to help users to better understand a concrete proof situation. Research has shown that visualizations can be used for uncovering information to help the user in processing more and possibly complex input [CMS99]. For example, in the field of software engineering, algorithms are already being visualized by animations. We have adopted this general approach of visualizing information. Concretely, we have developed an additional web-based tool to generate graphical representations of contracts based on their textual representation. Our tool should assist the user of VeriFast to spot possible errors more easily and, when used in teaching, might support students in comprehending and applying the concepts of formal verification.

The rest of the paper is organized as follows. Section 2 provides an overview of the features of VeriFast based on a simple example. Section 3 then explains the process of creating a visualization for both pre- and post-condition annotated to a function. In Section 4 the philosophy behind our tool prototype is presented. While Section 5 discusses related work, Section 6 concludes the paper.

⁵ VeriFast is also capable of verifying Java and C++ code, but this is out of scope in this paper.

2 Background

VeriFast is a verifier for programs written in C or other languages. When a C program is available, the user has to annotate the code with so-called assertions, most notably with a contract (i.e. pre-/post-condition) for each function occurring in the code.

A contract consisting of pre-condition P and post-condition Q for function $f()$ claims the following property: Whenever $f()$ is invoked in a state, in which P holds (i.e. P evaluates to *true*), and under the assumption that $f()$ terminates, then, the execution of $f()$ results in a state in which Q holds. Note, that termination itself is not part of this contract. Consequently, the contract is automatically fulfilled in case function $f()$ does not terminate (e.g. because `abort()` has been invoked)⁶.

We explain the functioning of VeriFast more deeply based on a simple but illustrative example, the well-known implementation of the data structure `stack`⁷. The screenshot at the upper half of Fig. 1 shows VeriFast after having successfully verified our `stack` implementation, while the lower half shows the failure of a verification attempt for the very same implementation except code line 22, which has been commented out.

The code lines 4–12 show the declaration of the datastructures `'stack'` and `'node'`. The lines 14–27 show the implementation of function `'createStack()'` together with the additional VeriFast annotations in lines 15–19. A VeriFast annotation can be placed on a single line or across multiple lines. In the former case, the annotation is prefixed with `'/*@'`; in the latter case enclosed by `'/*@'` and `'@*/'`. An annotation is always finished by a semicolon (`;'`). The annotation keyword `'requires'` (line 15) indicates a pre-condition, keyword `'ensures'` (line 16) a post-condition.

The annotation itself is formulated in the *VeriFast annotation language*, an expression language based on C-expressions. Each expression used as pre- or post-condition must be a boolean expression. The language of VeriFast annotations realizes some extensions compared to ordinary C-expressions, most importantly⁸:

- operator `'&*&'` - conjunctive aggregation of two subexpressions
- keyword `'result'` - usable only within a post-condition; represents the return value of annotated function
- `'malloc_block_<S>(<x>)'` - represents the fact that a memory block of struct `S` is available on the heap and can be safely accessed via variable `x`
 - when occurring in the pre-condition but not in the post-condition: memory block has been deleted from the heap during the execution of the function

⁶ There is a mode in VeriFast allowing also the verification of termination, but this is ignored in this paper.

⁷ The example has been taken to a large extend from the VeriFast tutorial [JSP17].

⁸ For a complete list, the reader is referred to <https://verifast.github.io/verifast-docs/> (last access: 2023/08/31).

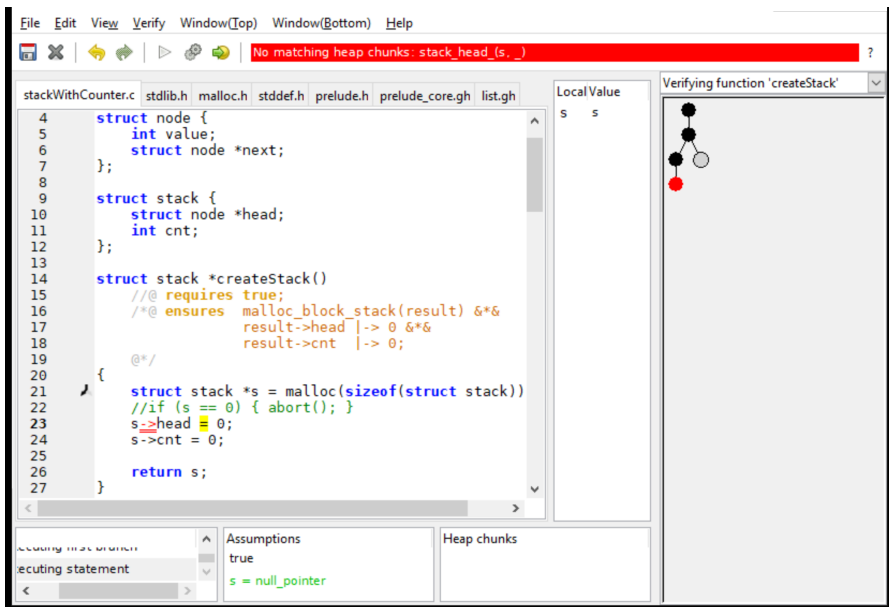
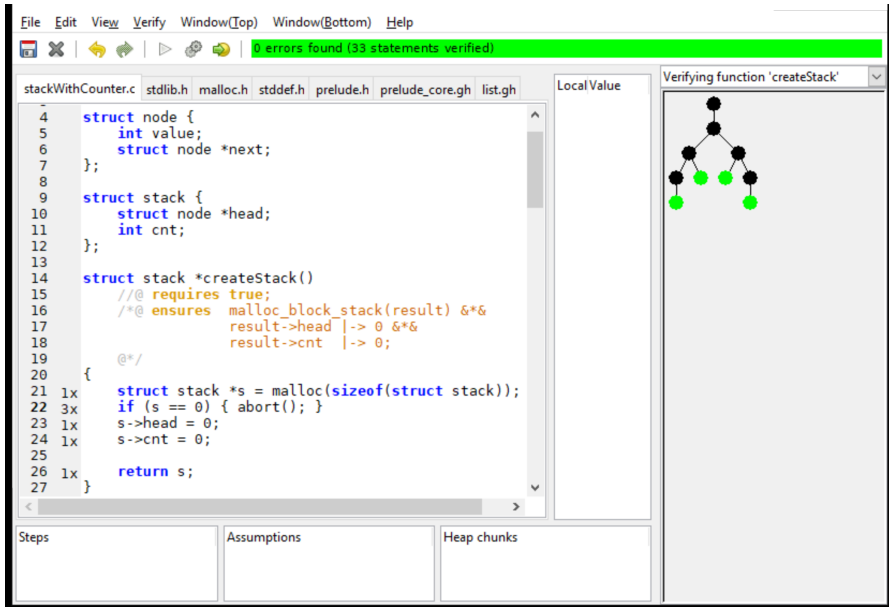


Fig. 1: Successful and Unsuccessful Verification of createStack()

- when occurring in the post-condition but not in the pre-condition: memory block has been created on the heap during the execution of the function
- ' $\langle x \rangle \rightarrow \langle \text{comp} \rangle \mid \rightarrow \langle v \rangle$ ' - represents the fact that component `comp` on object `x` can be safely accessed and the value of `comp` is `v`
- '? $\langle x \rangle$ ' - introduces a new free variable `x`, which can be used in the rest of the annotation (and even in the post-condition, when introduced in the pre-condition)
- '`INT_MAX`' - represents the maximal integer value (important for overflow analysis of arithmetic operations)

The contract annotated to function '`createStack()`' can be read as follows: The function can be invoked in any program state (pre-condition is '`true`') and assures that upon termination a new instance of `stack` has been created, which is returned (keyword '`result`') and whose component '`head`' has value 0 (representing the null pointer) and whose component '`cnt`' has value 0. Note that component '`cnt`' is supposed to encode the current number of '`node`' objects managed by the stack.

When we look at the implementation of '`createStack()`', we see that via '`malloc`' a memory block is allocated on the heap and returned as pointer '`s`'. Then, it is important to check in line 22, whether or not '`malloc`' has returned a null pointer (because, for example, there is no space left on the heap). The the first case, when a null pointer has been returned by '`malloc`', the function '`createStack()`' is finished abruptly via '`abort()`'. Note that this situation is considered to be not a proper termination of the function. Consequently, the constraints formulated in the function's post-condition do not apply.

In the second case, in which '`malloc`' returns a non-null pointer, the components '`head`' and '`cnt`' can be safely accessed and values can be assigned to them (line 23–24). VeriFast offers the possibility to access for every function the corresponding proof tree (see the utmost right part of the tool window). When all leafs of the tree are green, the verification has been successfully finished.

In the lower part of Fig. 1 we see an unsuccessful verification attempt. When line 22 is commented out, the proof attempt of VeriFast reveals that the function execution might stop due to access '`s->head`' in line 23 when '`s`' is a null pointer. Note, that VeriFast explains this error in pane *Assumptions* in the lower part of the window by characterising the erroneous situation with '`s = null_pointer`'.

3 Visualizing Contracts

Fig. 2 shows our tool⁹ displaying the stack example previously introduced. The left side of the screenshot shows the C code (we focus now on function '`push()`') of the stack example

⁹ Our tool is available from <https://languagehub.f1.htw-berlin.de/proofvisualization> (last access: 2023/08/31).

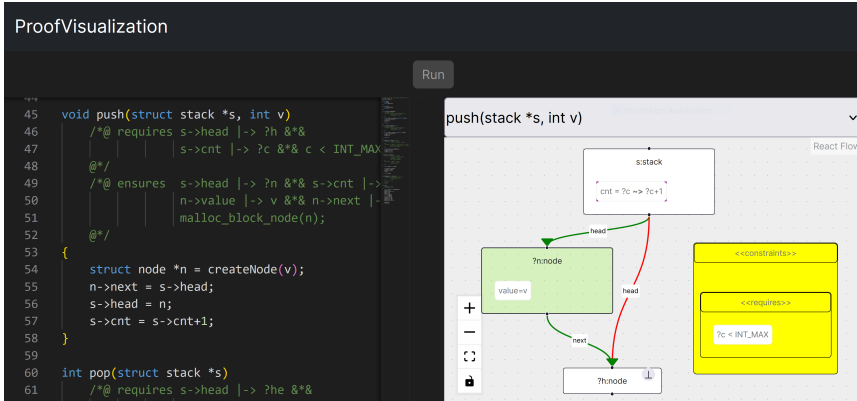


Fig. 2: Visualization of Contract for push()

within the editor, and the right side displays the graphical visualization of function 'push()' based on its VeriFast contracts.

```

void push(struct stack *s, int v)
    /*@ requires s->head |-> ?h &*&
        s->cnt |-> ?c &*& c < INT_MAX;

    @*/
    /*@ ensures s->head |-> ?n &*& s->cnt |-> c+1 &*&
        n->value |-> v &*& n->next |-> h &*&
        malloc_block_node(n);

    @*/
{
    struct node *n = createNode(v);
    n->next = s->head;
    s->head = n;
    s->cnt = s->cnt+1;
}


```

According to the implementation of function 'push()', a new 'node' structure is created and stored in the local variable 'n'. Its component 'value' is set to 'v', and its component 'next' is set to the current 'head' of the stack. The 'head' of the 'stack' is then updated to point to 'n', and the 'cnt' of the 'stack' is incremented by 1.

On the right-hand side, you can see a graphical representation of the 'push()' function, derived from the annotated pre- and post-condition. In this graph:

- There are three boxes labeled with 's:stack', '?n:node', '?h:node' representing the

struct instances involved in the execution of function 'push()'. Note that '?h:node' might also represent a NULL instance. Therefore, this box is marked with \perp .

- There is a yellow box containing additional constraints to be met during the execution of 'push()'.

- The component 'cnt' within 's:stack' changes its value during the execution of 'push()'. The change is reflected by operator '~>' within label 'cnt = ?c ~> ?c+1'.
- The 'head' arrow emanating from 's:stack' pointing to '?h:node' (or possibly a null pointer), is colored in red. This is because a new value is assigned to 'head' in the post-condition.
- A green 'head' arrow emanating from 's:stack' pointing to '?n:node' reflects the new value of 's->head'.
- Another green arrow, labeled 'next', extends from the new '?n' node to the node '?h' (or again, possibly a null pointer), indicating a new relationship formed in the function.
- The constraint within the yellow box is derived from the pre-condition 'c < INT_MAX'.

We color code these components and their transformations to provide an intuitive understanding of the changes that occur within the function based on the VeriFast contracts.

3.1 Algorithmic Generation of the Graph from the Textual Contract

Now that we have walked through a concrete example, we can transition into describing our general algorithm for contract visualization. This algorithm determines how we generate these graphs, what elements are considered objects, and how we color-code the transformations that occur based on the VeriFast contracts.

Starting with the user-entered C code with VeriFast annotations, the first step is to parse this input using a Domain-Specific Language (DSL) realized with Langium¹⁰. The DSL parsing process identifies and segregates the various components of the C code and VeriFast annotations, creating an Abstract Syntax Tree (AST). The AST provides a hierarchical representation of the different elements within the input code, acting as the groundwork for our visualization.

Using the AST, we break down each VeriFast contract into smaller statements, each of which can then be turned into visual objects. The generated visual objects, being derived from the AST, not only represent the VeriFast contracts but also demonstrate their placement and interrelations within the overall code structure. The resulting visualization provides

¹⁰ See <https://langium.org/> (last access: 2023/08/31).

an accurate and intuitive visual aid for understanding VeriFast contracts, with color-coded visual objects signifying the different stages and transformations of the entities involved.

4 Philosophy behind our Tool

The purpose of our tool is to make the work with VeriFast easier for users. This actually means to reduce the cognitive load for the user.

Neuroscience has a great tradition in analyzing how the style of presenting information influences the effectiveness of the consumer of this information. Due to the fact that the human brain can only handle a certain amount of information at the same time, it is a common technique to support thinking processes by visualizing information. Prominent examples mentioned in the work of Card et al., are written multiplication, maps and the periodic table of elements. The diverse mechanisms used to amplify cognition in the above mentioned and more examples, were gathered and summarized by Card et al., in six aspects [CMS99, p. 16]:

Visualizations can amplify cognition by:

- "increasing the memory and processing resources available to the users",
- "reducing the search for information",
- "using visual representations to enhance the detection of errors",
- "enabling perceptual inference operations",
- "using perceptual attention mechanisms for monitoring" and
- "encoding information in a manipulable medium"

To evaluate whether our developed tool actually supports the user in a cognitive way, it is evaluated below using the above cited aspects. Furthermore, additional features that are thinkable but not yet implemented are mentioned.

Increasing memory and processing resources The graph itself presented by the tool acts as outsourcing of information and increasing of human memory. Through the visual representation the user does not have to keep the entirety of the aggregations of subexpressions in mind but can read them off the presented graph.

Reducing the search for information The simplification of the search for information is more true in the visual representation than in the textual representation. For example, if a pointer is missing in the textual form, the user has to create the chain in his head and go

through it in order to find the missing pointer. In the visual representation, it is obvious at first glance where the gap is and which node does not have a next pointer.

Enhancing the detection of patterns The easier discovery of patterns is already provided by the visualisation as a graph. In the visual representation, it is obvious at first glance that it is a concatenation. In the textual representation, the nodes must first be sorted, their meaning learned and internalised in order to be able to discover patterns.

Enabling perceptual inference operations Enabling perceptual inferences is supported in the visual representation by representing the same nodes by the same shapes. Pointers, which are cumbersome and long in textual form, are merely a labelled arrow in the visual representation. The clear arrangement of the individual nodes also facilitates perceptual inferences.

Using perceptual attention mechanisms For this aspect of support a possible but not yet implemented feature could be to highlight differences in the graph as changes have been made to the contract.

Using a manipulable medium Graph visualisation provides users with a medium that can be manipulated. The nodes themselves in the graph can be moved with the mouse. Zooming in and out makes it possible to view the whole or even individual parts that are of particular interest. This allows users to arrange the graphic representation in a way that can support them in their personal process of understanding.

5 Related Work

There are many both commercial (e.g. Dafny) and research environments (e.g. KeY, VeriFast, Why3) for deductive verification available. The basic idea of all these tools is, to extend existing implementation code with specification code and to prove formally, that the implementation code conforms to the specification. All of the above mentioned tools exploit textual languages for formulating the specification (note that implementation languages have traditionally a textual syntax).

Our approach strives to make graphical representations of textual specifications available to the user. A big source of inspiration is the early attempt to extend the textual specification language OCL for UML with a graphical notation: VOCL (Visual OCL) [KTW02] proposes a graphical syntax for OCL together with a tool support (graphical editor) based on Eclipse.

While VOCL aims at visualizing all kinds of OCL expressions (e.g. ordinary invariants), the visualization of the effect of a function execution is the domain of so-called graph grammars and graph transformations ([Ro97]). This technique has been successfully applied especially to specify model transformation rules [GK10]. In contrast to the usual approach to specify a graph transformation by using two graphs - one representing the pre-state and one representing the post-state -, our visualization strives to represent the whole contract with just one graph, encoding the differences of pre- and post-state with colors (e.g. green, red). This idea has been pioneered by Schürr with his concept of Triple-Graph-Grammars (TGGs) [Sc94].

6 Conclusion

In this paper, we described the architecture, functioning and philosophy of a web-based tool, which is designed to assist the user of VeriFast. Our tool offers an editor for C programs enriched with VeriFast-annotations on the left side. Once the code is pasted into this editor and the Run-button has been pressed, our system generates for each annotated function in the C-code a corresponding graphical representation. The layout of this graph can be manually adapted.

So far, the system has been prototypically implemented and deployed on a publicly available URL. We plan to integrate our tool step by step into our courses on formal methods and software verification. The ultimate goal is to support the comprehension and efficiency of VeriFast users when struggling with the formulation of formal contracts. We also plan experiments to actually measure and quantify the support our tool can offer to understand a current proof situation faster and more deeply.

Our tool follows certain design principles derived from cognitive science and neuropsychology to lower the cognitive load for VeriFast users when working on code verification. Visualization is an important tool for reaching this goal. However, there are additional techniques we have in mind. For example, in order to enable a steep learning curve, students can be assisted with automatic explanation of the current situation to find a bug faster and more easily. Another idea is that our tool supports in the future also variant management and problem slicing: When a problem occurs, our tool assists the user to easily create a new variant of code and specification, starts VeriFast to work on the new variant, manages the achieved results, and returns to the original problem smoothly.

Bibliography

- [Ah05] Ahrendt, Wolfgang; Baar, Thomas; Beckert, Bernhard; Bubel, Richard; Giese, Martin; Hähnle, Reiner; Menzel, Wolfram; Mostowski, Wojciech; Roth, Andreas; Schlager, Steffen et al.: The KeY tool: integrating object oriented design and formal verification. *Software & Systems Modeling*, 4:32–54, 2005.

- [CES09] Clarke, Edmund M.; Emerson, E. Allen; Sifakis, Joseph: Model checking: algorithmic verification and debugging. *Commun. ACM*, 52(11):74–84, 2009.
- [CMS99] Card, Stuart K.; Mackinlay, Jock D.; Shneiderman, Ben: Readings in information visualization - using vision to think. Academic Press, 1999.
- [DDH72] Dahl, Ole-Johan; Dijkstra, Edsger W.; Hoare, Charles Antony Richard: Structured programming, volume 8 of A.P.I.C. Studies in data processing. Academic Press, 1972.
- [Em90] Emerson, E. Allen: Temporal and Modal Logic. In (van Leeuwen, Jan, ed.): *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics*, pp. 995–1072. Elsevier and MIT Press, 1990.
- [Fi11] Filliâtre, Jean-Christophe: Deductive software verification. *Int. J. Softw. Tools Technol. Transf.*, 13(5):397–403, 2011.
- [GK10] Greenyer, Joel; Kindler, Ekkart: Comparing relational model transformation technologies: implementing Query/View/Transformation with Triple Graph Grammars. *Softw. Syst. Model.*, 9(1):21–46, 2010.
- [GLH18] Grimm, Tomás; Lettnin, Djones; Hübner, Michael: A survey on formal verification techniques for safety-critical systems-on-chip. *Electronics*, 7(6):81, 2018.
- [HH19] Hähnle, Reiner; Huisman, Marieke: Deductive Software Verification: From Pen-and-Paper Proofs to Industrial Tools. In (Steffen, Bernhard; Woeginger, Gerhard J., eds): *Computing and Software Science - State of the Art and Perspectives*, volume 10000 of *Lecture Notes in Computer Science*, pp. 345–373. Springer, 2019.
- [Ho05] Holzmann, Gerard J.: Software model checking with SPIN. *Adv. Comput.*, 65:78–109, 2005.
- [JSP17] Jacobs, Bart; Smans, Jan; Piessens, Frank: , The VeriFast Program Verifier: A Tutorial, 2017. Available from <https://people.cs.kuleuven.be/bart.jacobs/verifast/tutorial.pdf>.
- [JVP15] Jacobs, Bart; Vogels, Frédéric; Piessens, Frank: Featherweight VeriFast. *Logical Methods in Computer Science*, Volume 11, Issue 3, sep 2015.
- [KTW02] Kiesner, Christiane; Taentzer, Gabriele; Winkelmann, Jessica: Visual OCL: A Visual Notation of the Object Constraint Language. Technical Report 2002/23, TU Berlin, Fakultät IV - Elektrotechnik und Informatik, 2002.
- [Me92] Meyer, Bertrand: Applying "Design by Contract". *Computer*, 25(10):40–51, 1992.
- [Pn77] Pnueli, Amir: The Temporal Logic of Programs. In: 18th Annual Symposium on Foundations of Computer Science, Providence, Rhode Island, USA, 31 October - 1 November 1977. IEEE Computer Society, pp. 46–57, 1977.
- [Ro97] Rozenberg, Grzegorz, ed. *Handbook of Graph Grammars and Computing by Graph Transformations, Volume 1: Foundations*. World Scientific, 1997.
- [Sc94] Schürr, Andy: Specification of Graph Translators with Triple Graph Grammars. In (Mayr, Ernst W.; Schmidt, Gunther; Tinhofer, Gottfried, eds): *Graph-Theoretic Concepts in Computer Science, 20th International Workshop, WG '94, Herrsching, Germany, June 16-18, 1994, Proceedings*. volume 903 of *Lecture Notes in Computer Science*. Springer, pp. 151–163, 1994.

On the Use of Process Containers for Running Vulnerable Programs more Securely

Nils T Siebel¹

Abstract: This article explains how a running program can be fortified against hacker attacks by encapsulating it in a process container. The methodological focus here is on namespaces which reduce or inhibit file system, networking and other activities of a process. This helps to limit potential changes or damage to the file system should the process be under the control of an attacker due to a security flaw.

In the second, practical part of this article is a case study compares the *firejail* and *bwrap* tools that help to configure this namespace functionality on Linux systems. Details are given on how certain major programs like Mozilla Firefox, LibreOffice and the popular PDF viewer Evince can be run more safely and at the same time without practical restrictions.

Keywords: process containers; OS kernel namespaces; security; hardening

1 Introduction

1.1 Motivation

Cyber attacks are very common and can cause a lot of damage in industry and the private sector. While the most common types of cyber attacks change over time, it seems the vulnerabilities enabling these attacks do not change much—and neither do the defense measures become strong enough to reduce successful attacks on PCs/laptops and servers alike.

New methods of defense are needed and the author believes that a powerful approach is the use of process² containers. A practical study has been carried out how best to configure these on an example system and details are presented here.

1.2 Ransomware Attacks

Ransomware attacks are currently one of the most endemic forms of cybercrime. In fact, ransomware was the most common attack type for the manufacturing industry in 2021 [Gr23].

¹ HTW University of Applied Sciences Berlin, Department 2, Wilhelminenhofstr. 75A, 12459 Berlin, Germany
firstname.lastname@htw-berlin.de

² A process is a running program.

A malicious hacker targets a company's computers and/or network with this type of malware. Data within the system is encrypted and at times the system itself is completely locked down, preventing any usage. The attacker then demands a ransom—often to be paid in Bitcoin—in order to release the system and data. This type of attack is particularly dangerous as it primarily affects the widely used Windows 10 operating system by Microsoft, as well as standard software such as Microsoft Office and Adobe Acrobat, thereby putting the data of many businesses at risk.

According to a study conducted by Kaspersky Lab in 2022, 64 % of organisations have so far been the victims of ransomware attacks. Most of them paid the ransom to their attackers [Kas23]. This includes a subsidiary of the US corporation Colonial Pipeline which was attacked by ransomware in May 2021 [Sh21]. The hackers actually took control of the company's network and demanded a ransom of 5 million US dollars. Eventually, the company made the decision to pay the ransom in order to resume operations. It also had to pay a fine of almost \$1 million for letting this attack happen [Ha22].

Examples of companies that likely possess a strong IT department demonstrate that ransomware poses a significant threat under Microsoft Windows 10. Measures taken by these companies to protect themselves against such attacks include installing antivirus software, providing training to employees regarding caution with emails and attachments, regular backups performed by administrative personnel, and automatic and frequent installation of updates.

However, even with these measures in place, data and corporate networks cannot be fully protected, as most ransomware attacks exploit previously unknown vulnerabilities in Microsoft Office, Adobe Acrobat, and the Microsoft Windows operating system to carry out their attacks. Malware components that take advantage of such unknown security flaws is referred to as “zero-day exploits”.

Defense against these attacks is not sufficient with the aforementioned steps alone. In the following we will discuss further technologies for securing against ransomware. One principle in this regard is “defense in depth” which involves implementing multiple layers of protection. This means that overcoming a single security measure is not sufficient to successfully execute an attack. We will discuss the typical steps in an attack as well as countermeasures such as containers and their potential efficacy.

1.3 Structure of this Article

In the first part of the work presented here, research on current attacks primarily aimed at PC systems in the industrial sector was examined, as the majority of attacks and resulting financial losses are prevalent in this area. Ransomware was chosen as the focal point due to its frequent occurrence.

Process containers are introduced and explained, following configuration options for a particular system. Due to the limited support for detailed namespace/container configurations in Microsoft Windows [Ne21, Wi23] all tests were made under Linux though the results can be easily transferred to Windows if and when this functionality is mirrored there.

2 Typical Ransomware Attacks and How to Stop Them

2.1 Steps taken by the Attacking Malware

A typical ransomware attack occurs through an infected PDF or Microsoft Office file, which is sent to the user via email or another method. Once the file is opened the ransomware code is executed on the computer, encrypting (and sometimes, additionally exfiltrating) the data. In many attacks on businesses the hackers demand a ransom of several hundred thousand US dollars to release the data.

The typical sequence of events is as follows:

1. Incoming email with an Office or PDF document.
2. Scanning by antivirus software, but no virus is detected (the ransomware is not included in this program code).
3. Upon reviewing the email, the document is manually or automatically opened using MS Office or Adobe Acrobat, e.g. for looking at or creating a preview/icon of the file.
4. Executable code within the Word/PDF file is executed.
5. The code downloads the actual malware from the Internet, e.g. by accessing a website under control by the hacker. The malware can also be encrypted during download, deceiving the antivirus scanner which only sees unrecognisable data.
6. The downloaded malware is also executed, such as on the Windows Script Host.
7. “Elevation of Privilege” occurs, granting system rights to the malicious code. This needs a flaw in the operating system or software running with administrative privileges [CV23].
8. (Almost) indestructible installation of the malware within the system (persistence).
9. Background operation (possibly undetected for days or weeks):
 - Data encryption
 - Often simultaneously, exfiltration of this data to a server controlled by the attacker

- Simultaneous spreading within the network, infecting and encrypting more data and systems
10. Once the operation is complete original files are removed together with all backups and system restore data. A note is left with a ransom demand and a threat to publish the data if payment is not made within a few days.

Highlighted in the following are core aspects which enable the attack:

- A file containing text and images (i.e., data) can also include instructions (i.e., program code). Although unnecessary for most use cases, this feature has been added to PDF and Microsoft Office formats. The person opening the file may not expect or desire the execution of the code, but it occurs automatically and invisibly in the background.
- MS Office and Adobe Acrobat are fully connected to the internet, even though it is not necessary in practice for viewing and modifying files. This is the reason why program code within the file can fetch the malicious code from the attacker's server.
- The malicious code executes unusual program code, writes and reads unusual files (e.g. overwriting Office documents, vacation photos or scanning the network for other vulnerable computers). These functionalities are unnecessary for a program designed to display a PDF or Office file.

Container technology can address these points and prevent specific necessary steps of the malware or its companion "loader" component. As a result damage can be greatly reduced or perhaps even completely avoided because each subsequent step of the attack depends on the successful completion of the preceding steps.

2.2 Securing Vulnerable Processes with Containers

2.2.1 Definition and Difference to Docker-like Software and Virtualisation

A *process container* is defined as a single process or group of processes that are isolated from other processes and/or computer resources such as storage and network connectivity [Wh22]. Access to personally identifiable information (PII) such as login and computer names can also be restricted.

Docker is a software platform that allows the creation, deployment and management of containers. Docker containers are a special form of containers that are designed as self-contained units which not only contain the application itself but also dependent libraries and other code as well as configuration files. The goal is to achieve consistent and portable execution of a software across different computing environments.

As a core functionality Docker provides a tools for managing containers and accessing “registries” which are storage repositories for Docker images. The most commonly used registry is Docker Hub³ which is a public registry holding a wide variety of pre-built images. Docker Hub is free to use for public images but also offers private repositories for paid accounts.

From the security standpoint the automatic download of code from the Internet (here, in a Docker container) with subsequent execution of this code always holds risks. In order to mitigate these Docker images would need to be built by a trusted and capable user, cryptographically signed and its signature checked before execution. When using Docker one also relies on this user to keep the software, including dependencies, updated and to configure Docker setting for maximum security for the given scenario.

This dependency can often not be guaranteed, which can actually make the use of Docker as a container solution less secure than running a local application which is updated and configured by ourselves.

Virtualisation is a much more complex approach which exceeds both complexity and isolation provided by containers and solutions like Docker. Virtual machines (VMs) are software emulations of complete physical computers that provide a strictly isolated and self-contained environment for running an operating system as well as applications within it. A VM provides a great deal of isolation and independence of the host system.

In this article we focus on simple and lightweight containers, usually around a single process or process group.

2.2.2 Container Technologies in Linux

Containers have had an established use in the Linux kernel for several years [Wh22, Wa18]. They consist mostly mainly of two things: *Linux kernel namespaces* and *control groups* (cgroups).

Namespaces in this context isolate different aspects. The main types are as follows [vK21]:

1. User Namespace: This creates a new set of user and group IDs
2. Process ID (PID) Namespace: Isolates processes, inhibiting the view of and communication with other processes.
3. Network Namespace: A new network stack with its own devices (or lack thereof), routing and firewall rules etc.

³ <https://hub.docker.com/>

4. Mount Namespace: This creates a new, empty file system for the process; one can connect (“mount”) existing folders and files as needed. This enables read-only mounts e.g. of system folders.
5. Interprocess Communication (IPC) Namespace: This enables certain IPC methods such as pipes.
6. UNIX Time-Sharing (UTS) Namespace: Allows to generate distinct host and domain names, creating a separate computer identity.

Control groups are methods to limit and prioritise the use of measurable resources like RAM, I/O and network bandwidth. This is especially useful to defend against attacks on the availability of a system, though in the case of ransomware attacks on desktop systems this is perhaps less relevant.

Cgroups are nevertheless an important and configurable component of Linux process containers.

A fine grained control over the functionality of a running program can furthermore be achieved by *SECCOMP*, the “Secure Computing Mode” which is a further security feature of recent Linux kernels. It provides a mechanism for restricting the system calls that a process can make. Sensitive system calls include execution of other programs (*execve*, *fork*, *clone*), process tracing/debugging (*ptrace*), memory mapping (*mmap*) and network access (e.g. *socket*). The most security relevant SECCOMP functionality therefore overlaps with the effect of using namespaces and will not be further discussed here.

One might wonder whether the use of namespaces and/or cgroups uses system resources itself and may slow a process down. However, all processes in current Linux systems already run in control groups and a common namespace. There is no measurable performance impact of Linux process namespaces on performance.

2.2.3 Container Technologies in Windows 10 und 11

Windows 10 and 11 do not provide a technology similar to Linux kernel namespaces in effect and detail. Microsoft itself explicitly mentions Docker and similar complex systems when talking about containers and seems to exclude smaller solutions that are likely still used within Docker implementations under Windows [Wi23, Ge23]. In fact, any container will need a Windows base image to run (*ibid.*), which is much more similar to the Docker approach than it is to Linux containers.

Unlike Linux kernel namespaces Windows namespaces do not explicitly support filesystem namespaces although it is possible to create an isolated filesystem for a process by creating a virtual drive.

While not done under Windows, containers could be a good way to prevent attacks especially in the case of MS Office or PDF files. In many of these cases attacks through Office or PDF files are executed using exploits that target vulnerabilities within the applications themselves. These are often zero day vulnerabilities/exploits which means that it is impossible to secure the applications themselves by keeping them up to date. In these cases containers could contain an attack even if it cannot be completely prevented.

3 Malware Defense with Containers

Following the attack sequence listed in Section 2.1 the following recommendations can easily be derived:

- Processes should be isolated from the network unless needed
- Processes must be hindered from reading and writing files and (folder) locations that are unnecessary for normal operation
- Especially for Internet-facing programs (which do need network access) collection of personally identifiable information should be limited
- Processes should be isolated from other processes that may run in a less restricted container.

It should be noted that if a process is isolated against the network it can still read files that are on a network share / file system. The reason behind this is that *the operating system* maps the network folder onto a local folder which can then be read locally (subject to our configuration of file system name spaces with regards to that local folder). Therefore the network access (as carried out by the OS itself, outside the container) is not blocked by a network namespace.

3.1 The Firejail Tool

The *Firejail* tool [Fi23] is perhaps the simplest way to run Linux programs in containers. This is due to a large number of pre-defined “profiles” with configuration options for a large number⁴ of popular Linux programs. When called with a program name as a parameter it will load specific (or default) container configuration options, create a container and start the specified program within it.

For example, starting the web browser *firefox* with *firejail* one would call from a terminal (or, change the quickstart menu item to):

⁴ currently, more than 1200

```
$ firejail firefox
```

This will create a container with all files and folders except the `/Download` folder hidden from Firefox. Risky executables like `sudo` are unavailable and system folders are mounted read-only within the container. The temporary folder `/tmp`, shared among all users, seems to be empty within the container, and files put into it will be stored temporarily only for the duration of the Firefox process runtime. Private GPG keys are not readable by Firefox.

Running the PDF viewer Evince

```
$ firejail evince somefile.pdf
```

does not separate the process from the Internet. However, again sensitive files are hidden, basic commands like `ls` are gone. This limits the possibilities, although most files in the home folder can be read and written.

Running without network is easy:

```
$ firejail --net=none evince somefile.pdf
```

which does the same as the above variant but without network functionality.

The profile for the Libre Office suite is similar to that of Evince which makes it much more secure than running without containers. Separation from the network can again be achieved by the `--net=none` option.

3.2 The Bubblewrap Tools

The *bubblewrap* software tools [Bu23] expose all options of Linux kernel namespaces and control groups (and `seccomp`). The *bwrap* command forces strict security by forbidding everything, basically starting with empty namespaces. Then, if you need files like system libraries in `/lib` or `/usr` you need to manually mount them (preferably, read-only). The same goes for own folders and files as needed.

This approach makes the use of graphical programs challenging when aiming for a strict filtering because one needs to mount in folders for fonts and their configurations, libraries, possibly configuration files of audio interface libraries, `openSSL` for checking certificates and so on. An very elaborate example could look like this (command broken across lines):

```
$ bwrap --unshare-user --uid 1000 --gid 100 \  
--ro-bind /tmp/dbus-data-dir.EUqTAc55/etc /etc \  

```

```

--unshare-ipc --unshare-pid --unshare-uts --unshare-cgroup \
--hostname laptop \
--ro-bind /lib64 /lib64 --ro-bind /sbin /sbin --ro-bind /usr /usr \
--ro-bind /bin /bin \
--die-with-parent --new-session --dev /dev --tmpfs /tmp --proc /proc \
--setenv GVFS_DISABLE_FUSE 1 \
--ro-bind /tmp/dbus-data-dir.EUqTAc55/usr/lib/gvfs /usr/lib/gvfs \
--ro-bind /tmp/dbus-data-dir.EUqTAc55/usr/share/dbus-1/services \
  /usr/share/dbus-1/services \
--bind /tmp/dbus-data-dir.EUqTAc55 /tmp/dbus-data \
--setenv XDG_RUNTIME_DIR /tmp/dbus-data \
--bind /tmp/dbus-data-dir.EUqTAc55/etc/machine-id \
  /var/lib/dbus/machine-id \
--ro-bind /home/nils/.Xauthority /home/nils/.Xauthority \
--bind /tmp/.X11-unix /tmp/.X11-unix \
--ro-bind /var/lib/ca-certificates/openssl /var/lib/ca-certificates/openssl \
--bind-try /home/nils/.fontconfig /home/nils/.fontconfig \
--bind-try /home/nils/.cache/fontconfig /home/nils/.cache/fontconfig \
--ro-bind /home/nils/Research/Articles/II-WS2023 \
  /home/nils/Research/Articles/II-WS2023 \
--ro-bind /usr/share/texmf/doc/latex/lni /usr/share/texmf/doc/latex/lni \
--unshare-net dbus-run-session evince /usr/share/texmf/doc/latex/lni/lni.pdf

```

This actually also starts a *dbus* process which is an IPC (inter process communication) system in order to achieve a higher degree of isolation. The program will not start without *dbus*, however *dbus* allows for the sending of commands to other programs. Therefore a new *dbus* instance is started with only our new *evince* process in it. *Evince* will therefore run but not be able to find any other process on that IPC bus.

Naturally no-one wants to type a command like that, this is of course installed in a shell script created within this project with a few simple options. It will set the *bwrap* options, create helper files and folders, then spawn the process.

3.3 Comparison and Alternative Tools

The two tools *Firejail* and *bubblewrap* are similar in that they use the same – OS kernel provided – container technologies. Therefore the effect is identical and has the same power against attacks. The usage and handling therefore is the main difference in daily life. *Firejail* is much simpler to use as it provides profiles readily configured for very many executables. For a beginner this is the better tool. *Bubblewrap* on the other hand has a very restrictive base configuration and you need to actively configure the resources your process needs to use. This avoids using defaults which may not be as restrictive as they could be in your particular use case, increasing security. The necessity for explicit configuration without profiles makes this tool more suitable for advanced users.

An alternative software tool, more similar perhaps to *bwrap* than to *firejail* is the *LXC* software from the Linux Container Projects group, sponsored by Canonical Ltd[Li23]. It is again using the Linux kernel namespaces and cgroups but with differing tools for configuration.

3.4 Remaining Attack Surface

One remaining problem for defending against malware by the use of containers is perhaps the complexity of the configuration. Providing very many options for namespace and control group configuration and making all of these adaptable the Linux OS also relies on individuals to configure these options correctly to both run successfully and well isolated from the system. The *bwrap* software needs more configuration but gives more flexibility while *firejail* will provide “sane” default configurations for a variety of popular programs. Therefore the isolation security relies on the quality of these defaults.

Security provided by containers naturally relies on the correct implementation of the namespace and control group implementations in the Linux kernel. Should there be issues in the implementation the security may be impacted.

3.5 Comparison to other Approaches

Other approaches to defend against malware exploiting software flaws exist and are in use – according to the statistics cited above⁵, with disappointing success levels.

These approaches may include

- Antivirus software scanning files and downloads – defeated by the “loader” concept and encrypted/obfuscated downloads mentioned above.
- Sandbox approaches (usually within web browsers) – similar to containerisation but within the targeted software itself. Malware exploits weaknesses in the vulnerable software which may well include its sandbox, see e.g. [CV22]. Also, the conflict about functionality desired by the software vendor (online access for tracking, code execution in PDF and office files) and ideal security concepts still remains. This would have to be enforced by the security focused user, externally to the software.

4 Conclusions

In this article container technologies for securing a single process or group of processes from cyber attacks was discussed. Using containers could hinder many of the popular malware attacks by providing isolation from file systems, networks and other resources.

⁵ e.g. 64 % of organisations successfully attacked by ransomware

After detailing a typical ransomware attack specific defensive measures through containers were given. Since Microsoft Windows containers are not (yet?) quite as capable and flexible the tests were run under Linux. Two software tools were demonstrated and compared with recommendations to run standard programs under both.

A remaining wish of the author is the wide availability, flexible configurability and good documentation of process namespaces under Windows. It does not, however, look very likely given the current material provided by Microsoft on this subject.

Bibliography

- [Bu23] Bubblewrap. Website, <https://github.com/containers/bubblewrap>.
- [CV22] CVE Details: Policy Bypass in COOP in Google Chrome Prior to 98.0.4758.80 Allowed a Remote Attacker to Bypass iframe Sandbox via a Crafted HTML Page. Website, <https://www.cvedetails.com/cve/CVE-2022-0461/>.
- [CV23] CVE Details: Microsoft Security Vulnerabilities (Gain Privilege). Website, https://www.cvedetails.com/vulnerability-list.php?vendor_id=26&opgpriv=1.
- [Fi23] Firejail Security Sandbox. Website, <https://firejail.wordpress.com/>.
- [Ge23] Get started: Prep Windows for containers. Website, <https://learn.microsoft.com/en-us/virtualization/windowscontainers/quick-start/set-up-environment>.
- [Gr23] The Latest 2023 Ransomware Statistics (updated May 2023). Website, <https://aag-it.com/the-latest-ransomware-statistics/>.
- [Ha22] Colonial Pipeline faces nearly \$1m fine one year after ransomware attack. Website, https://www.theregister.com/2022/05/09/in_brief_security/.
- [Kas23] How Business Executives Perceive Ransomware Threat. Website, 2023. <https://www.kaspersky.com/blog/anti-ransomware-day-report/>.
- [Li23] Linux Containers – LXC – Introduction. Website, <https://linuxcontainers.org/lxc/introduction/>.
- [Ne21] Network isolation and security. Website, 2021. <https://learn.microsoft.com/en-us/virtualization/windowscontainers/container-networking/network-isolation-security>.
- [Sh21] US declares emergency after ransomware shuts oil pipeline that pumps 100 million gallons a day. Website, https://www.theregister.com/2021/05/10/colonial_pipeline_ransomware/.
- [vK21] What Are Namespaces and cgroups, and How Do They Work? Website, <https://www.nginx.com/blog/what-are-namespaces-cgroups-how-do-they-work/>.
- [Wa18] A Sysadmin’s Guide to Containers. Website, <https://opensource.com/article/18/8/sysadmins-guide-containers>.

- [Wh22] What's a Linux container? Website, <https://www.redhat.com/en/topics/containers/whats-a-linux-container>.
- [Wi23] Windows and containers. Website, 2023. <https://learn.microsoft.com/en-us/virtualization/windowscontainers/about/>.

Vertikale Integration im Kontext von Industrie 4.0

Das DFT-Labor als Showcase für eine Smarte Fabrik

Léon Niclas Dodrimong¹ und Ute Dietrich²

Abstract: Vertikale Integration von Maschinendaten zu übergeordneten Unternehmensprozessen ist ein wichtiges Kennzeichen der Industrie 4.0 und hilft dabei, Fabriken intelligent zu vernetzen. Die sogenannten smarten Fabriken beinhalten eine Vielzahl von Industrie 4.0 Technologien. In diesem Paper soll sich theoretisch und praktisch am Beispiel des DFT-Labors der HTW mit der vertikalen Integration im Bereich Smart Factory - Industrie 4.0 auseinandergesetzt werden.

Keywords: Industrie 4.0, Smart Factory, vertikale Integration, Cyber Physical Systems, DFT-Labor, autonome Systeme, Internet der Dinge

1 Einleitung

Zur Umsetzung von sogenannten smarten‘ Fabriken sind jede Menge Kommunikationen auf unterschiedlichsten Ebenen basierend auf unterschiedlich aggregierten Daten nötig. Die Anforderungen an Standards sowohl bezogen auf die Protokolle als auch auf die zugrunde gelegten Datenmodelle sind in den letzten Jahren deutlich gestiegen. Die im Industrie 4.0 fokussierte und durch Cyber-physischen Systems geforderte Vernetzung von Objekten führt zu Integrationsanforderungen sowohl auf horizontaler als auf vertikaler Ebene. Vertikale Integration ist ein somit ein wichtiges Kennzeichen der Industrie 4.0. Im Weiteren soll am Beispiel des Digital Factory Technology (DFT)-Labors der HTW mit der vertikalen Integration im Bereich Smart Factory auseinandergesetzt und eine prototypische Implementierung vorgestellt werden.

2 Industrie 4.0

Die Industrie erfuhr einen stetigen Wandel über die vier industriellen Revolutionen hinweg, welche zu einem grundlegenden Paradigmenwechsel in der Produktion und zum Aufkommen des globalen Informationszeitalters geführt haben. In der mit Industrie 4.0 bezeichneten vierten Industriellen Revolution verschmelzen physische und digitale Technologien, um intelligente Vernetzung und neue produktionstechnologische Innovationen zu ermöglichen.

¹ Hochschule für Technik und Wirtschaft, FB 4, Treskowallee 8, 10318 Berlin, Leon.Dodrimong@student.htw-berlin.de

² Hochschule für Technik und Wirtschaft, FB 4, Treskowallee 8, 10318 Berlin, Ute.Dietrich@htw-berlin.de

2.1 Technologien

Die im Kontext von Industrie 4.0 entwickelten Technologien ermöglichen verbesserte Business-Modelle, einschließlich vernetzter Produktionssysteme, Echtzeitüberwachung, datengetriebener Entscheidungsfindung und verbesserter Produktqualität. Durch das IIoT aggregierte Daten werden mithilfe von Big Data Methoden verarbeitet. Modelle aus der dem Bereich der KI werden zur Analyse verwendet, Cloud Computing dient zur unabhängigen Bereitstellung skalierbarer Ressourcen und vertikale und horizontaler Integration zur Vernetzung von cyber-physischen Systemen. [ITI21]

2.2 Horizontale und vertikale Integration

Horizontale Integration im Sinne einer Vernetzung zwischen einzelnen Komponenten einer Fabrik untereinander dient u.a. zur Automatisierung von Prozessen basierend auf ähnlich aggregierten Datenmodellen [Die22]. Maschinen können miteinander kommunizieren, interagieren und sich so adaptiv an geänderte Bedingungen anpassen. Eine effektive Kommunikation erfordert die einheitliche Nutzung von Protokollen sowie die Schaffung von plattform- bzw. auch Sprach-unabhängigen Schnittstellen.

Die vertikale Integration umfasst dagegen die Integration von Maschinendaten zu übergeordneten Unternehmensprozessen. Eine effizientere, vertikale Integration soll Unternehmen bei der Erreichung von Nachhaltigkeitszielen, z.B. durch digitale Produktpässe, unterstützen und helfen, ressourcensparend zu produzieren. Hinzu kommt eine verbesserte Reaktion auf Nachfrageänderungen, aufgrund direkten Zugriffs auf alle Stufen der Produktionskette.

3 DFT-Labor

Die Zielstellung des Das DFT-Labors im Fachbereichs 4 beinhaltet sowohl die Umsetzung von Demonstrationsbeispielen zur Smarten Fabrik, als auch die Bereitstellung verschiedener Test-Cases für den Einsatz in der Lehre. Studierende sollen einen Einblick in die Vernetzung von Produktionsressourcen durch Cyber-physische Systeme bekommen und befähigt werden, selbständig Szenarien in diesem Umfeld umzusetzen.

3.1 Technologien

Im DFT-Labor stehen verschiedene kooperative Roboter, Fahrerlose Transportsysteme, eine 5-Achsen-NC-Pocket-Maschine, mehrere 3D Drucker und eine VR-Anlage zur Verfügung. Eine vertikale Vernetzung der Komponenten wird mithilfe der Siemens Cloud MindSphere als IoT-as-a-Service Anbieter umgesetzt. Für einen plattformunabhängigen Datenaustausch im Rahmen der industriellen Automatisierung

wird OPC UA, eine serviceorientierte Architektur und Weiterentwicklung des OPC-Standards, verwendet. Ziel ist es, Maschinen und Komponenten so zu vernetzen, dass ein Showcase für smarte Fabriken entsteht.

3.2 Vertikale Integration am Beispiel

Im DFT-Labor erfolgt beispielhaft eine vertikale Integration von Maschinendaten, wobei Daten des Franka Roboterarms bis zur IoT-Cloud MindSphere übertragen werden. Die dahinterliegende Zielstellung bezieht sich auf die Umsetzung vorausschauender Wartungsszenarien durch die Erfassung von Gelenkpositionen inklusive der Geschwindigkeiten und Beschleunigungsparameter. Im ersten Prototyp wurden Daten der Franka Roboter über einen OPC-UA Server mit der MindSphere verbunden [Do21]. Dazu wurde zuerst ein Datenmodell entworfen, welches Daten und Funktionalitäten des Franka Roboters im OPC UA Standard darstellt. Die Companion Specification OPC UA Robotics, ein standardisiertem Informationsmodell für Roboter [Fo23], wurde verwendet, wodurch eine leichtere Integration in andere Umgebungen, welchen ebenfalls die OPC UA Robotics verwenden, möglich ist.

Die Kommunikation mit den Franka Robotern erfolgt mithilfe einer C++ Bibliothek, welche Funktionen zur Datenabfrage, Parametrisierung und Steuerung bereitstellt. Diese Daten und Funktionalitäten werden innerhalb des OPC UA Servers entsprechend des entwickelten Datenmodells zur Verfügung gestellt. Die Implementierung des OPC UA Servers erfolgt mithilfe des OPC UA SDK's, die Umsetzung des Datenmodells mithilfe des UA-Modellierers der Unified Automation GmbH. Somit können Daten und Funktionen über den OPC UA Standard zugänglich gemacht und gegebenenfalls auch von anderen Komponenten genutzt werden.

Ausgewählte Daten wurden in MindSphere als Vorlage, sogenannte Aspekte, modelliert. Dieses Template kann genutzt werden, um einzelne reale Franka Roboter als digitalen Zwilling zu Instanzieren. Um die realen Daten aus dem OPC UA Server mit dem digitalen Abbild zu verknüpfen, erfolgt ein Mapping-Prozess. Hierzu wird ein Edge Device-Gerät als Gateway benötigt, welches die Daten vom OPC UA erhält und an die MindSphere sendet. Genutzt wurde in dem umgesetzten Prototypen die Siemens MindConnect IoT2040-Box, die aktuell eine Virtuelle Maschine mit einem installierten MindConnect OS verwendet. Über das entsprechend konfigurierte Edge Device-Gerät können nun Datenpunkte aus dem OPC UA Datenmodell mit den Daten des digitalen Zwillings in der die MindSphere- Cloud gemappt werden.

Die in der Cloud für beide Roboter aufgesammelten Daten können, wie schon erwähnt, für verschiedene Anwendungsfälle genutzt werden. So lassen sich entsprechende prädiktive Analysen implementieren, die z.B. aufgrund der Gelenkbewegungen verbunden mit zusätzlichen Verschleiß-Daten eine vorausschauende Wartung oder eine Echtzeitsteuerung des Produktionsprozesses ermöglichen.

4 Fazit

Die vertikale Integration spielt eine Schlüsselrolle bei der Umsetzung von smarten Fabriken. Sie ermöglicht, alle Aspekte des Produktionsprozesses zu verbinden und zu koordinieren, vom Rohmaterial bis zum Endprodukt, über die komplette Lieferkette hinweg. In Smarten Fabriken sind Maschinen sowohl horizontal als auch vertikal mit den darüber agierenden Informationssystemen vernetzt. Dies ermöglicht eine kontinuierliche Datenübertragung und Kommunikation, die die Überwachung und Anpassung der Produktion in Echtzeit erlaubt.

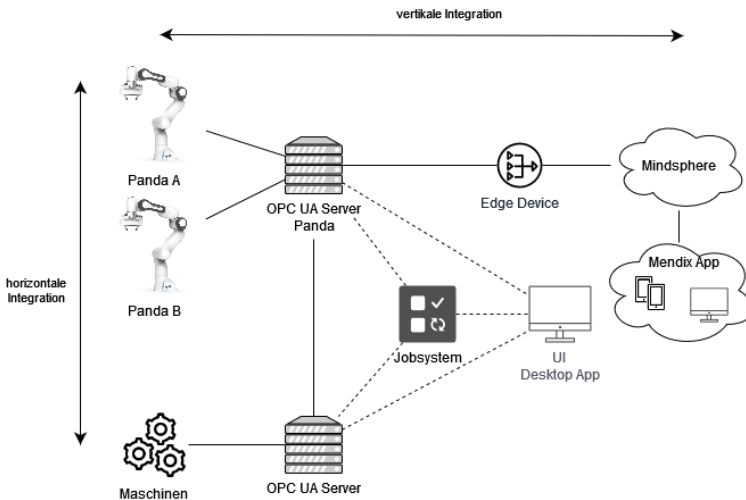


Abb. 1: Erweiterungen im DFT-Labor

Abb. 1 stellt eine potenzielle Erweiterung der aktuellen Implementierung dar. So können zusätzliche Applikationen in MindSphere bereitgestellt werden, die eine verbesserte ad-hoc-Steuerung der angebotenen Ressourcen erlauben. Erweiterte Analysetools z.B. für ein verbessertes Business Intelligence, aber auch unterschiedliche Optimierungsszenarien mittels Einbeziehung von Methoden der Künstlichen Intelligenz zur Umsetzung vorausschauender Service-Angebote sind ebenfalls möglich [Die22]. Die Implementierung weiterer OPC UA Server kann durch horizontale Integration vernetzt werden, wodurch Maschinen miteinander interagieren können. Zudem können Jobsysteme Aufträge entlang eines digitalen Prozess-Zwillings verwalten und über die OPC UA Server Zustände überwachen.

Literaturverzeichnis

- [Di22] Dietrich, U.: Digitaler Zwilling als Schlüsselkomponente für die Automatisierung von Systemen, Digitaler Zwilling in der Automatisierungstechnik, Automatisierungsforum 2022, FH Westküste, 2022.
- [Do21] Dodrimong, L. N.: Entwicklung eines OPC UA Servers für den Franka Emika Roboter mit digitalem Abbild in der MindSphere Cloud, Hochschule für Technik und Wirtschaft in Berlin, 2021.
- [Fo23] Foundation, O.: OPC 40010-1: Robotics - Vertical Integration, 2023, url:<https://reference.opcfoundation.org/v104/Robotics/v100>, Stand: 20. 05. 2023.
- [ITI21] İsa, G.; Tarık, Z. Ç.; İbrahim, B.: Industry 4.0 and Smart Production. TEM Journal/, Oktober, 2021.

Entwicklung von Löschdrohnen im Projekt "Feuerwehr der Zukunft" zur praxisorientierten Ausbildung von Ingenieuren

Christian Lehmann¹, Björn Wudka ¹, Volker Wohlgemuth¹ und Frank Burghardt¹

Abstract: Im Rahmen eines Verbundforschungsprojektes erfolgt die Entwicklung eines Löschdrohnenkonzepts für den Einsatz im Brand- und Katastrophenschutz. Hierzu werden Hard- und Softwareelemente für eine Bodenstation als auch für den Einsatz eines Löschdrohnenschwarms entwickelt. Dabei fließen diverse Kenntnisse aus der Ausbildung der Ingenieure in die Entwicklung ein. Angefangen vom Einsatz von Bus- und remote Kommunikationssystemen, Softwareentwicklung der Firmware und Bedienstation, Bildverarbeitung, Design Thinking und Konstruktion bis hin zu spezifischen Aufgaben der Umweltinformatik werden viele Aspekte autonomer Systeme in diesem Projekt abgedeckt. Das große studentische Interesse spiegelt sich auch in einer sehr hohen Zahl von Abschlussarbeiten wider.

Keywords: Drohnen, autonome Systeme, Schwarmsteuerung, Brand- und Katastrophenschutz

1 Einleitung

Das WIR!-Projekt (Wandel durch Innovation in der Region) zur Konzeption, pilothaften Entwicklung und feuerwehrtechnischen Erprobung eines Löschdrohnenschwarms zur direkten Vegetationsbrandbekämpfung verbindet die Forschung mehrerer Hochschulen aus der Region Brandenburg/Berlin mit einem KMU (Klein- und Mittelständische Unternehmen) und der Feuerwehr. Das Vorhaben führt ein vorher gestartetes ZIM-Projekt (Zentrales Innovationsprogramm Mittelstand) weiter [E123, E122] und hat das Ziel, das Flug- und Brandlöschverhalten einzelner Löschdrohnen in einem Schwarm konzeptionell zu erfassen, zu entwickeln und zu untersuchen. Einen besonderen Schwerpunkt stellen die Erforschung der Löschwirkung bei Wald- und Vegetationsbränden sowie die Erprobung von realitätsnahen Einsatzszenarien im engen Zusammenwirken mit der Feuerwehr dar. Das Löschen von Wald- und Vegetationsbränden ist bis heute sehr personal- und zeitintensiv. Vor allem der große und lang andauernde Einsatz von Feuerwehrkräften stellt für die freiwilligen Feuerwehren der ländlichen Regionen eine besondere Herausforderung dar. Darüber hinaus wird die Brandbekämpfung durch z.B. unzugängliches Gelände [RN22] oder Munitionsbelastung des Brandgebietes [Sp22] erschwert. Zu diesem Zweck werden vermehrt Löschhubschrauber als auch Löschpanzer

¹ Hochschule für Technik und Wirtschaft Berlin, Fachbereich2 Ingenieurinformatik, Wilhelminenhofstraße 75A, 12459 Berlin, vorname.nachname@htw-berlin.de

eingesetzt, die eine schnelle und sichere Brandbekämpfung ermöglichen. Dennoch ist der Einsatz durch Ruhezeiten der Bediener zeitlich begrenzt. Gleichzeitig bedeutet der Einsatz dieser Systeme einen hohen Kostenaufwand der Kommunen und Gemeinden. Dem gegenüber erlaubt die Brandbekämpfung durch lasttragende Drohnen einen kontinuierlichen, kostengünstigen und kräfteschonenden Einsatz. Die Einsatzkräfte der Feuerwehren werden dadurch erheblich bei der Bekämpfung unterstützt und der Schaden für Natur und Forstwirtschaft deutlich reduziert.

Zur Realisierung Projekts werden die Löschdrohnen weiterentwickelt. Hierbei soll ein automatisierter Flugbetrieb eines Drohnenschwarms im Zusammenwirken mit einer speziellen Bodenstation und Einsatzsteuerung erfolgen. Ziel ist es, einen effektiven Löschmitteltransport und -einsatz zur Bekämpfung von Vegetationsbränden zu erreichen. Darüber hinaus werden neben der Betriebssicherheit und Technik der Drohnen sowie der Entwicklung der automatisierten Bodenstation und Einsatzsteuerung in erster Linie die Grundlagen für einen Löscheinsatz durch die theoretische wie praktische Erforschung von Löschwirkung, Einsatztaktik und Einsatzwert und ferner die Klärung des komplexen rechtlichen Rahmens für unbemannte Luftfahrzeuge im BOS-Einsatz (Behörden und Organisationen mit Sicherheitsaufgaben) geschaffen.

Die beteiligten Hochschulen können damit ihren Mitarbeitern und Studierenden sehr attraktive Ingenieurinformatik-fokussierte Themen bieten, welches die Gebiete Konstruktion, interne und externe Kommunikationssysteme, Weiterentwicklung von Firmware und Anwendungssoftware, Automatisierung sowie Entwicklung autonomer Systeme beinhaltet.

2 Fachlicher Überblick

Die zu entwickelnden technischen Hauptkomponenten des Projektes sind

- der Drohnenschwarm
- eine Bodenstation sowie
- eine Bedienstation

Durch das Zusammenwirken der drei Einzelkomponenten wird das in Abb.1 aufgezeigte Schema zur Brandbekämpfung aufgebaut. Das zentrale System bildet dabei die Bodenstation. Diese übernimmt zum einen die Befüllung der Drohnen mit Löschmitteln und der Wechsel des Energieträgers der Drohnen. Zum anderen erfolgt durch sie die Koordination des Drohnenschwarms durch Zuweisung von Landplätzen und Strategien. Die Strategieermittlung erfolgt dabei durch den Drohnenführer der im engen Austausch mit dem Einsatzleiter der Feuerwehr die Strategie des Drohnenschwarms anpasst. Die Strategie wird mit Hilfe der die Bedienstation an das System übermittelt. Der Drohnenschwarm, als ausführende Einheit des Löschangriffs, besteht aus bis zu

40 Drohnen die das Löschmittel kontinuierlich entsprechend der definierten Strategie am Einsatzort abwerfen. Folgend fliegen sie zur Bodenstation zurück.

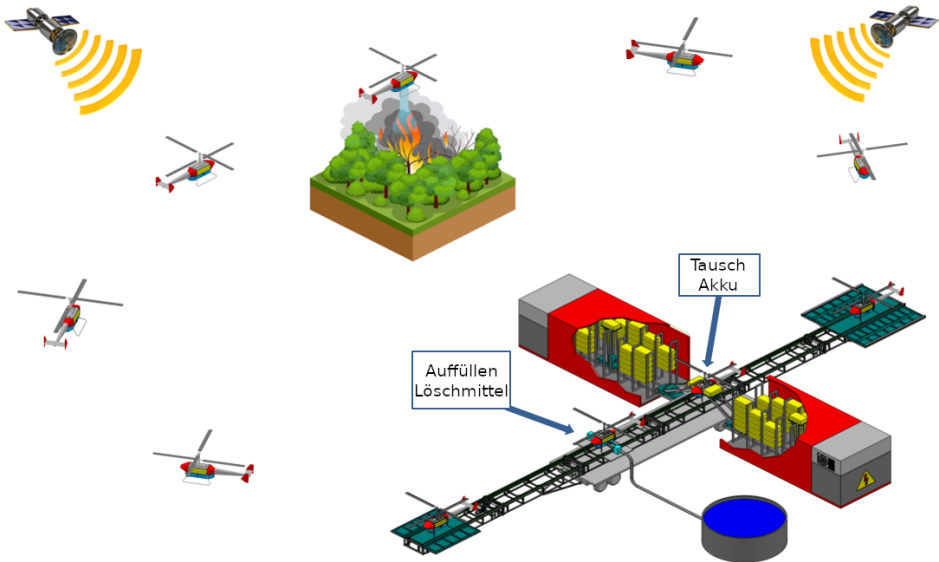


Abb. 1 Schematischer Aufbau des Einsatzkonzepts der Löschdrohnen

Die Entwicklung der drei Komponenten zeigt eine Vielzahl von Problemstellungen auf. Lasttragende Drohnen mit hohen Traglasten sind wenig verbreitet. Leistungsstarke elektrische angetriebene Drohnen mit Lithium-Ionen-Akkusystemen hoher Energiedichte existieren nur vereinzelt und sind zumeist für einen speziellen Anwendungszweck konfiguriert. Das wiederholte automatisierte Landen in Verbindung mit einem automatischen Wechsel der Akkus und der automatischen Aufnahme einer Last sowie das Anfliegen eines entfernten Zieles mit Lastabwurf und Rückkehr zur Bodenstation befindet sich aktuell bei Großunternehmen wie z.B. Amazon in der Erprobung, ist jedoch noch nicht serienreif. Die gegenwärtig bekannten Transportlasten sind zudem recht gering und genügen nicht den Ansprüchen eines Feuerlöschesatzes.

Die Steuerung eines Drohnenschwarms ist ebenfalls differenziert zu betrachten. Im Consumerbereich mit den kleinen, leichten Multikoptern werden kompakte, oft Smartphone-basierte Steuerungen angeboten. Hier bedarf es einiger Übung im Umgang, die Nutzung ist dann jedoch in gewissen Grenzen möglich, der Schaden bei Unfällen ist bedingt durch das geringe Gewicht und die Fluggeschwindigkeiten vertretbar. Formationsflüge als Showeffekte erfolgen verhältnismäßig langsam und entlang vorher definierter Trajektorien. Kollisionen sind demnach selten und die Formationen agieren oft über einem freien Platz, von dem aus auch gestartet wird. Schwerlastdrohnen mit einem wesentlich höheren Eigengewicht erfordern eine Steuerung, die die Sicherheitsanforderungen des Einsatzzwecks erfüllen. Speziell bei hohen Geschwindigkeiten (hier mindestens 60km/h) sowie dem Flug außerhalb des

Sichtbereiches ist ein fortgeschrittenes System zur Kollisionsvermeidung der Drohnen sowie mit anderen Flugobjekten (Rettungshubschrauber) und statischen Hindernissen (Waldbrandwachtürme, Strommasten, Schornsteine) notwendig. Hier existieren noch keine kommerziellen Lösungen, die sich als Grundlage anbieten.

3 Einfluss auf die Lehre

Das Projekt bietet eine Fülle an Forschungsthemen die in studentischen Abschlussarbeiten als auch Dissertationen bearbeitet werden sollen. Weiterhin haben die Themen auch einen Einfluss auf die Lehrinhalte bestimmter Fächer. So wurde zum Beispiel die Vermittlung und der Einsatz von ROS (Robot Operating Systems) vertieft oder eine grundsätzliche Einführung über Akkutechnik in den Mechatronik-Fächern gegeben.

Die bisherigen Abschlussarbeiten wurden vorwiegend an Bacheloranten vergeben und decken ein sehr großes Spektrum ab. Das zeigt sich zum einen an den in mindestens vier Studiengängen (Ingenieurinformatik, Umweltinformatik, Maschinenbau, Industrial Design) entstandenen Abschlussarbeiten. Die dabei bearbeiteten Themen sind zum Beispiel, Bildverarbeitung, elektronische Bussysteme wie UAV-CAN, Konstruktion der Wassertankanlage, SPS- Steuerung zur Automatisierung der Bodenstation, Design Thinking Aspekte zur Gestaltung der Bedieneroberfläche, Konzepte zur Steuerung des Drohnenschwarms, Kollisionsvermeidung etc.


Die Motivation der Studierenden für das Projekt ist sehr groß, da sie selbst aktiv einen Beitrag zur Bekämpfung der Folgen des Klimawandels leisten können. Es gab auch schon Teilnehmer die selbst in einer Katastrophe involviert waren und damit unbedingt an solchen zukünftigen Lösungen mitarbeiten wollten.

Literaturverzeichnis

- [El23] Elfgen, M. u.a.: Abschlussbericht zum ZIM-Kooperationsprojekt „Entwicklung eines Drohnenschwarms zum Einsatz in Katastrophenregionen“, April 2023
- [RN22] RND, Redaktionsnetzwerk Deutschland; Waldbrand im Nationalpark Sächsische Schweiz: Lage unverändert ernst, in: RND Panorama, 03.08.2022, <https://www.rnd.de/panorama/waldbrand-in-saechsischer-schweiz-lage-in-nationalpark-unveraendert-ernst-Y3FE2F7S553XSEE5K5YZ2PRWEI.html>, Stand: 24.05.2023
- [Sp22] Spiegel, Großer Waldbrand in Brandenburg – auf Fläche werden Kampfmittel vermutet, Spiegel Panorama, 18.06.2022 , <https://www.spiegel.de/panorama/treuenbrietzen-waldbrand-in-brandenburg-auf-flaeche-mit-kampfmittelverdacht-a-59ac30e5-a2a0-4eec-9830-1b94956625aa> Stand: 24.05.2023
- [El22] Elfgen, M. u.a.: Comprehensive Drone System for Deployment in Disaster Scenarios with Focus on Forest Fire Fighting, In (Wohlgemuth, V. u.a. Hrsg.): Advances and New Trends in Environmental Informatics. Springer Verlag, Hamburg 2022, S. 149-164, 2022

RoboTwins

Kommunizierende Knickarm-Roboter mit haptisch sensitivem Feedback

Kai Schauer¹ 

Abstract: Die RoboTwins wurden für Demonstrationszwecke in der Lehre entwickelt. Sie bestehen aus verschiedenen Komponenten, die sich wie Module eines Baukastens unter den Restriktionen von Anwendungen individuell anpassen, austauschen und durch Datenkommunikation verbinden lassen. Die zwei miteinander kommunizierenden Knickarm-Roboter veranschaulichen das haptische Feedback in ferngesteuerten Systemen. Sämtliche Systemkomponenten sind durch andere mit gleicher Funktion oder andersartigen Wirkprinzipien auswechselbar. Einer der RoboTwins fungiert als Remote Control, der andere wird gesteuert und sendet drahtlos dem steuernden ein sensitives Feedback zurück, wenn z.B. Hindernisse die freie Bewegung stören. Es können auch beliebig andere Informationen, wie Temperaturen, Geräuschpegel usw. übermittelt werden. Studierende der Ingenieurinformatik erfahren, wie Sensoren und Aktoren bekannte physikalische Effekte umsetzen. Sie lernen die Wirkungsweise und das Zusammenspiel der Komponenten sowie den Systementwurf kennen und die mechanischen Einflüsse in der Programmierung zu berücksichtigen.

Keywords: Mechatronik, Robotik, Mikrocontroller, Sensitives Feedback, IoT

1 Haptische Systeme

1.1 Begriffsklärung

"Haptik" ist von dem griechischen Wort "haptesthai" abgeleitet und bedeutet „ergreifen“, „anfassen“ oder „berühren". Haptische Systeme ermöglichen die physische Interaktion mit einer realen, virtuellen oder örtlich verschiedenen Umgebung. Sie umfassen auch thermische und schmerz-sensitive Wahrnehmungen [BIR03]. Vergleichsweise sind im allgemeinen Umfeld "Optik" oder "Akustik" bekannt.

Die Vereinigung von Sensor-, Aktor- und Signalverarbeitungsfunktionen ermöglicht Anwendern haptischer Systeme bzw. Geräte den physischen Kontakt, ohne ihn mit eigenen Händen auszuführen oder vor Ort zu sein. Die Anwender agieren und erhalten vom System ein Feedback, indem die Sinne (Temperatur, Kraft, Vibration u.v.m.) an bestimmten Stellen des Körpers angesprochen werden [KLJ17]. „Haptische Schnittstelle“ ist dafür ein alternativer Begriff.

¹ Hochschule für Technik und Wirtschaft Berlin, Fachbereich 2, SG Ingenieurinformatik,
Wilhelminenhofstr. 75A, 12459 Berlin, kai.schauer@htw-berlin.de

Aus dem Umgang mit Smartphones ist das "tactile feedback" gut bekannt. Das Gerät reagiert auf eine Bewegung an der berührten Oberfläche. Dies ist vergleichbar mit dem Gefühl, einen Schalter oder eine Taste zu betätigen. Die Wirkprinzipien unterscheiden sich in jene, die an einem einzelnen Berührungspunkt einen Impuls im Mikrometerbereich auslösen, gerade groß genug, um ihn mit den Fingerspitzen zu spüren und in jene, bei denen die gesamte Oberfläche/das Gerät z.B. mit einem Vibrationsmotor bewegt wird [KLJ17].

Ein Bereich der Haptik sind Force-Feedback-Systeme [LIE16, PRA12]. Hier vermitteln Muskeln und Sehnen das Gefühl, dass Kräfte wirken. Ein breites Betätigungsfeld bietet die Medizintechnik. So finden diese Systeme Anwendung in der minimalinvasiven Chirurgie [CHA23], in der Telemanipulation [CHA23, KAS13] oder in der Bein- oder Armprothetik [SOR14] und vielem mehr.

1.2 Motivation für die Entwicklung der RoboTwins

Die meisten kommerziellen haptischen Systeme und Force-Feedback-Systeme bilden nur Bewegungen ab. Reale Belastungen, Hindernisse und Umwelteinflüsse werden nicht nachempfunden oder übertragen. Die Sensorik ist in spielzeugähnlichen Tools, Lenkrädern oder Joysticks aus Plastik verbaut. Anwender sehen durch Virtual Reality (VR) zwar was sie tun, spüren dabei aber nicht, dass Gegenstände ein Gewicht oder eine Temperatur haben, dass sie schwingen oder auf harte oder weiche Hindernisse stoßen, dass sie Schwerpunkte haben, die ausbalanciert werden müssen. Diese im Realen „gefühlten“ Informationen bleiben in der VR meist verschlossen. Es fehlt das wirklichkeitsnahe „Feedback“ des Handelns. Force-Feedback beschränkt sich auf Kräfte, auch wenn das allein schon eine Herausforderung ist. Ließe sich den Anwendern ein umfassendes sensitives Feedback vermitteln, könnten Personen unvorhersehbare Ereignisse (z.B. Brandbekämpfung) in einem virtuellen Szenario oder in der Realität außerhalb von Gefahrenbereichen trainieren, so dass sie dann intuitiv richtig handeln lernen.

Hierzu ein Demonstrationsbeispiel: Wie könnte ein Feuerwehrmann den professionellen Umgang mit einem Feuerwehrschauch zur Brandbekämpfung trainieren? Die Situation ließe sich in VR (virtual reality) schon abbilden, ähnlich einem Computerspiel. Doch wäre es nicht realistisch, wenn der Feuerwehrmann nur ein leichtes Sensortool aus Plastik in der Hand hielte und im Übungsraum Raumtemperatur herrscht.

Szenenwechsel: Der Feuerwehrmann trägt seine vollständige Ausrüstung, eine AR-Brille (augmented reality) und nimmt ein Modell des Feuerwehrschauchs in die Hand, das mit der Master-Mechanik verbunden ist. Er befindet sich in einem Raum jenseits der Gefahrenzone. Zur selben Zeit ist ein Roboter, die Slave-Mechanik, vor Ort im Gefahrenbereich und führt synchron die Handlungen des Feuerwehrmanns aus. In der Datenbrille wird der Video-Stream aus der Gefahrenzone, gerne überlagert mit Infrarotaufnahmen, eingespielt. Die Slave-Mechanik sendet ein sensitives Feedback

zurück zur Master-Mechanik. Diese erzeugt mechanische Widerstände proportional zum Szenario vor Ort. Der Feuerwehrmann sieht sein Handeln, spürt und beobachtet seine Aktionen außerhalb der Gefahrenzone!

Es geht also darum, die Aktivitäten der Anwender von VR bei Nutzung von Tools auf die reale Situation, ähnlich einer Fernsteuerung zu transformieren. Gleichzeitig erhalten sie ein direktes Feedback durch Ansprechen der Sinne oder Muskeln. In den meisten Applikationen von VR liegt der Schwerpunkt heute noch auf „Virtual“. Das hier vorgestellte Konzept verleiht dem Anspruch von „Reality“ eine neue Dimension.

Dieselbe technische Ausstattung kann auch für andere Szenarien durch Einspielen eines neuen Szenarios in der Software und der Datenbrille genutzt werden. Branchen wie Luft- und Raumfahrt, Medizin, Munitionsräumung oder Arbeitsschutz, Notfallmedizin und Rehasport bieten dafür vielfältigste Anwendungsgebiete.

1.3 Bezug zum Studium der Ingenieurinformatik

Studierende nehmen Module des Curriculums oft als zusammenhangslose Einzelleistungen wahr. Eine zentrale Frage stellt sich für sie dann immer wieder: Wozu wird das gebraucht? Der modulare Aufbau der RoboTwins und die Systemzusammenhänge beantworten diese Frage durch Demonstration. Darüber hinaus können andere Wirkprinzipien für Sensoren und Aktoren vorgestellt und verglichen sowie typische industrielle Anwendungen aufgezeigt werden.

Semester	Modul	Inhalte
1; 2	Technische Mechanik	Kräfte und Momente, statische Bestimmtheit
1; 2; 3	Technisches Zeichnen und Konstruktion	3D-CAD, Wellenlagerungen, kinematische Berechnungen und Dimensionierung, Maschinenelemente, montagegerechte Konstruktion
2; 3; 4	Mechatronik 1...3	Elektronische Bauelemente, Signalverarbeitung, elektr. Antriebstechnik, Sensorik, Steuern und Regeln, μ C-Steuerungen programmieren, Entwurf von Mikroprozessorsteuerung und deren Programmierung
5	Fachübergreifendes Projekt Ingenieurinformatik	Eigenständige Lösung von Projektaufgaben, Projektmanagement, Arbeiten im Team, Projekt-Lifecycle, Qualitätssicherung

Tab. 1: Thematischer Bezug der RoboTwins zu Lehrinhalten des Bachelor Ingenieurinformatik

Das erlernte Wissen zu den Komponenten wird in Laborübungen vertieft und variiert. Die RoboTwins kehren wie ein roter Faden als Demonstratoren über mehrere Semester und in unterschiedlichen Modulen [Tab. 1] wieder. Das System lässt sich weiterentwickeln und jederzeit mit modernen Komponenten ausstatten und ergänzen.

2 Systembeschreibung der RoboTwins

2.1 Kinematisches Konzept

Die hier vorgestellten Demonstratoren veranschaulichen das Prinzip der kommunizierenden Kinematiken. Eine Randbedingung war es, den Aufbau mit Kaufteilen und in Ausnahmen mit Teilen aus dem 3D-Drucker zu realisieren. Das Sortiment von Makeblock [MB23] erfüllt diese Forderung. Es gibt zahlreiche mechanische Komponenten, Profile, Wellen und Wellenlagerungen und eine überschaubare Auswahl an Getriebemotoren. Alle mechanischen Komponenten sind als 3D-CAD-Modelle verfügbar. Das macht eine detaillierte Neu- oder Änderungskonstruktion möglich.

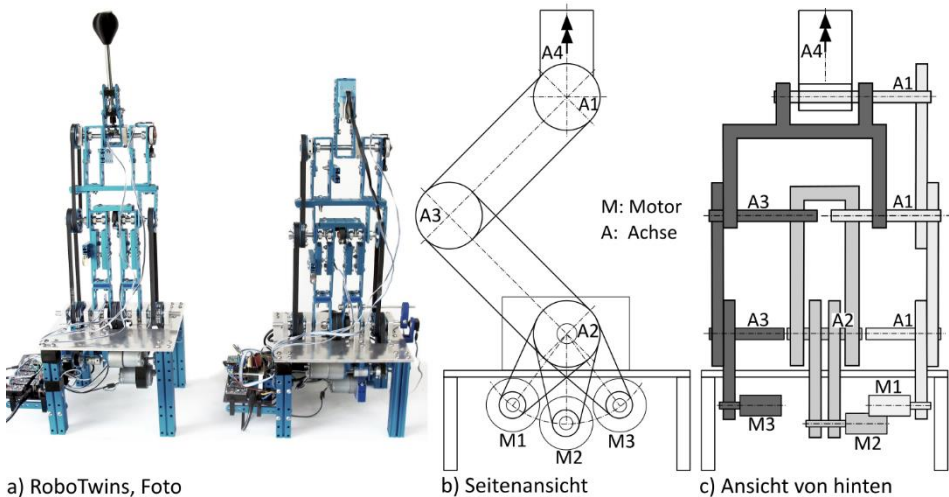


Abb. 1: Prototypen der RoboTwins und mechanische Prinzipskizze

Nach Vergleich verschiedener Bewegungskonzepte fiel die Entscheidung zugunsten der Knickarm-Kinematik. Im senkrechten Stand müssen alle Achsen harmonieren. Der 90° -Winkel des gesamten Arms ist für die Antriebsmotoren wegen der zu leistenden Drehmomente anspruchsvoll. Damit die passive Last der Getriebemotoren die Nutzlast der Kinematik nicht reduziert, sind diese unterhalb der Plattform montiert (M1 ... M3). Die Antriebsmomente (M1 ... M3) werden durch Zahnriemen bis zur jeweiligen Antriebswelle mittels Achsen übertragen. Die Achse A4 ist für eine Kamera reserviert.

Das kinematische Konzept sieht sieben Riemen mit drei verschiedenen Längen vor. Die Riemen müssen ohne Verbindungselement umlaufend sein. Im Katalog zum Makeblock-Baukasten reichen die Informationen zu den Komponenten der Zahnriementriebe nicht für die Konstruktion aus. So werden die Zahnriemenräder, wie auch die Zahnriemen lediglich mit der Anzahl der Zähne unterschieden. Die Verzahnung hat ein MXL-Profil. Das ist genormt und entspricht einem Pulley-Pitch PP von 2,032mm. Das Lochraster der mechanischen Komponenten, mit denen die Achs-/Wellenlagerungen gebaut werden, definiert die Sprünge der möglichen Achsabstände A . Bei einer Übersetzung zwischen großem ($Z1$) und kleinem ($Z2$) Riemenrad oder bei gleich großen Riemenrädern ($Z1=Z2$) sind nun alle Informationen vorhanden, um einen passenden Zahnriemen auszuwählen. In Abb. 2 sind die geometrischen Zusammenhänge für die nachfolgend beschriebenen Berechnungen skizziert.

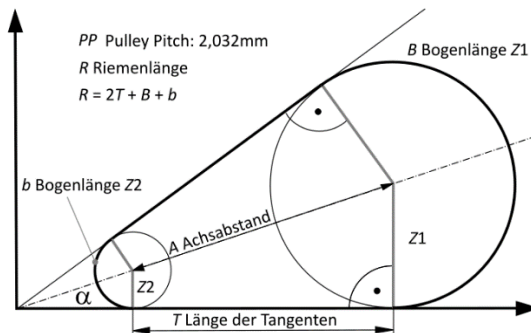


Abb. 2: Skizze zur Berechnung der Riemenlänge R

Weisen die Zahnriemenräder unterschiedliche Zähnezahlen auf, stellt sich ein Winkel α der verbindenden Tangenten in Abhängigkeit vom Achsabstand A ein. Mit Gleichung (1) ist der Winkel α berechenbar.

$$\alpha = \arctan\left(\frac{PP \cdot (Z1 - Z2)}{2\pi \cdot A}\right) \quad (1)$$

Die Riemenlänge R addiert sich nach Gleichung (2) aus dem doppelten der Tangente T und den Bogenlängen b und B entsprechend den Umschlingungswinkeln. Weil die Auswahl der Zahnriemen auch nach der Zähnezahl Z_{Riemen} erfolgt, muss die Riemenlänge R , die sich in Millimetern ergibt, noch durch den Pulley Pitch PP dividiert werden. Gleichung (2) ist nun nur noch von den Zähnezahlen der Zahnriemenräder $Z1$ und $Z2$ sowie dem Achsabstand A abhängig.

$$Z_{\text{Riemen}} = 2 \cdot \frac{A \cdot \cos \alpha}{PP} + Z1 \cdot \frac{180^\circ + 2 \cdot \alpha}{360^\circ} + Z2 \cdot \frac{180^\circ - 2 \cdot \alpha}{360^\circ} \quad (2)$$

Für die Auswahl der passenden Zahnriemen werden die drei Parameter nun so lange variiert, bis ein nach Zähnezahl Z_{Riemen} passender Zahnriemen im Katalog zu finden ist.

Gleichstromgetriebemotoren treiben die beiden Knickarmroboter an. Die Nennspannung der Motoren beträgt 12V. Bei einer Stromaufnahme von 450mA unter Last liefern sie ein Nenn Drehmoment von 50mNm bei ca. 480U/min. Die Entscheidung fiel zugunsten der Gleichstromgetriebemotoren wegen zwei sehr wichtiger Merkmale aus. Zum einen lassen sie sich mittels H-Brücke und PWM sehr praktisch regeln. Zum anderen dämpft das Getriebe die PID-Regelung sehr robust [MOR09, SCH08]. Bei der Positions-Regelung wird ein DC-Motor atypisch genutzt. Häufige Drehrichtungswechsel und Lastmomente bewirken Drehzahlen, die sehr viel niedriger sind als im idealen Motorbetrieb. Das führt letztendlich zu sehr großen Strömen von bis zu 2400mA, wenn die Motoren anlaufen oder blockieren. Auch dieser Form der Beanspruchung müssen die Motoren standhalten.

2.2 Elektronik

Die Gleichstromgetriebemotoren sind in der Schaltung die Verbraucher mit einer maximalen Spannung von 12V. Hierfür gibt es im Handel geeignete Netzteile. Es ist auch möglich, gleich 12V, z.B. aus der Spannungsverteilung eines PKW's zu verwenden. Damit wird der Forderung für die mobile Nutzung der RoboTwins sehr gut entsprochen.

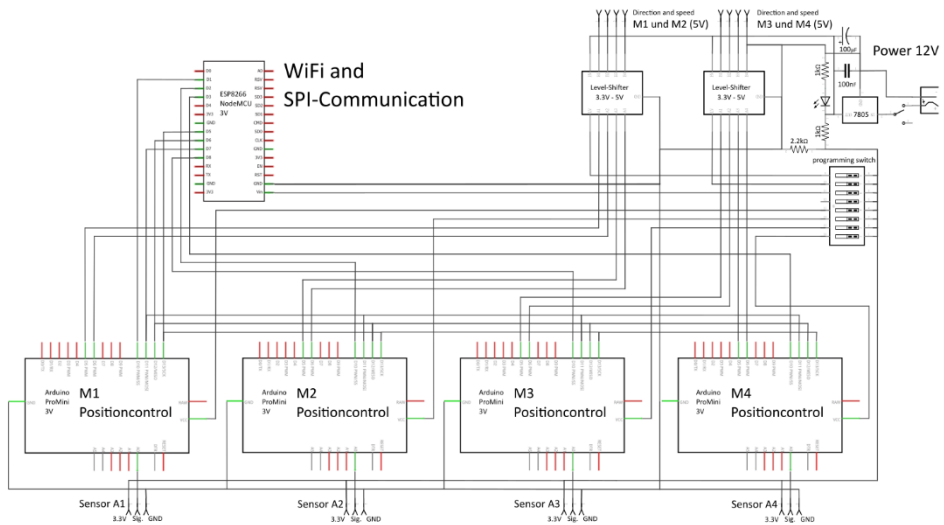


Abb. 3: Schaltplan der Steuerungen, für Master und Slave identisch

Das NodeMCU arbeitet mit 3,3V. Die ProMini-Arduino-Shields können wahlweise 3,3V oder 5V nutzen. Sie kommunizieren mit dem NodeMCU via SPI-Schnittstelle. D.h. die Spannung für die Mikrocontroller muss für 3,3V festgelegt werden, um mit den ProMinis das NodeMCU nicht zu beschädigen. Die H-Bridge-Motortreiber sind vom Typ LN298N. Diese benötigen einen 5V-Spannungspegel für die Signalleitungen. Das wiederum macht Levelshifter erforderlich. Diese heben die von den PID-Reglern der ProMinis kommenden

PWM-Signale vom 3,3V- auf den 5V-Pegel. Bei der Auswahl der Levelshifter ist darauf zu achten, dass die Transistoren der Shields schnell genug schalten, um die mit minimal 31KHz getakteten PWM-Signale störungsfrei zu übertragen.

Somit sind drei Spannungsniveaus zu bedienen, 12V, 5V und 3,3V. Wie bereits erwähnt liefern Netzteile die 12 V. Mit einem Spannungsregler 7805 werden die 5V umgesetzt. Von dort geht es über einen Spannungsteiler noch weiter runter auf 3,3V.

Zuletzt sei hier der Programming-Switch erwähnt. Die Mikrocontroller müssen einzeln geflasht werden. Der Programming-Switch unterbricht die Spannungsversorgungen. Die anderen Mikrocontroller sind in diesem Moment passiv.

2.3 Drahtlose Kommunikation

Haptische Systeme, bei denen die Anwender örtlich vom ausführenden Gerät getrennt sind, benötigen sichere Verbindungen, im Idealfall drahtlos. Je nach Anforderung der jeweiligen Applikation definieren sich die zur Verfügung stehenden Funktechnologien. Tab. 2 vergleicht diese anhand der für diese Anwendung wichtigen Merkmale.

	WLAN	Bluetooth	2,4GHz	433kHz
Produkt Beispiel	ESP8266	HC-05	NRF24L01	RF433
Frequenz	2,4GHz	2,4GHz	2,4GHz	433GHz
Betriebsspannung	3,3V	3,3 ... 6V	3,3V	3,3 ... 12V
Kommunikation	IP	Serial	SPI	Serial
Komm.-Richtung	bidirektional	bidirektional	bidirektional	unidirektional
Stromverbrauch (ca.)	0,3A	0,04A	0,0126A	0,004A
CPU on board	ja	nein	nein	nein
Reichweite	100 ... 250m	10m	1000m	100m
Transfer-Variablen	char, String	int, char	byte, int, String	uint8_t
Preise (ca.)	2€	6€	2€	1€

Tab. 2: Vergleich gängiger Funktechnologien

WLAN-Module von Espressif oder vergleichbar, haben eine eigene CPU on board. Sie benötigen somit keinen separaten Mikrocontroller. NRF24L01-Module erreichen sehr hohe Reichweiten. Weitere Merkmale sind die Preise der μC 's und das Format der zu versendenden Nachricht. Einzelne Zahlen in unidirektionaler Richtung können einfach mittels 433kHz-Modulen versendet werden.

Die RoboTwins sind mit WLAN-Modulen ausgestattet. Die Sensorwerte werden in Strings oder Charakters verpackt und als ein Paket versendet. Die jeweils andere Seite zerlegt die empfangene Nachricht in einzelne Werte und prüft ihre Richtigkeit.

2.4 Programmierung

Die Bezeichnungen Master und Slave wurden bewusst gewählt. Wenn die beiden Systeme miteinander kommunizieren, muss ein System die Priorität haben. Das sei hier der Master. Der Slave führt aus, kann aber möglicherweise die Zielwerte nicht erreichen. Weil beide Systeme permanent miteinander kommunizieren, findet eine Rückkopplung in Gestalt des hier praktizierten Feedbacksystems statt. In dem Moment hätten die Werte des Slaves Priorität. Um diesen Konflikt zu vermeiden, werden die Werte des Slaves als informativ ohne Priorität weiterverarbeitet. Das Master-System behält die Priorität.

Damit die RoboTwins ihre Funktionen korrekt ausführen und die dafür notwendigen Informationen pünktlich bereitstehen, ist der Datenfluss in Teilaufgaben unterteilt und durchstrukturiert, vergleichbar mit einer Choreografie (Abb. 4). Alles beginnt mit der Winkelmessung aller Gelenke des Masters, derzeit mittels Potentiometern an den AD-IO's der ProMinis.

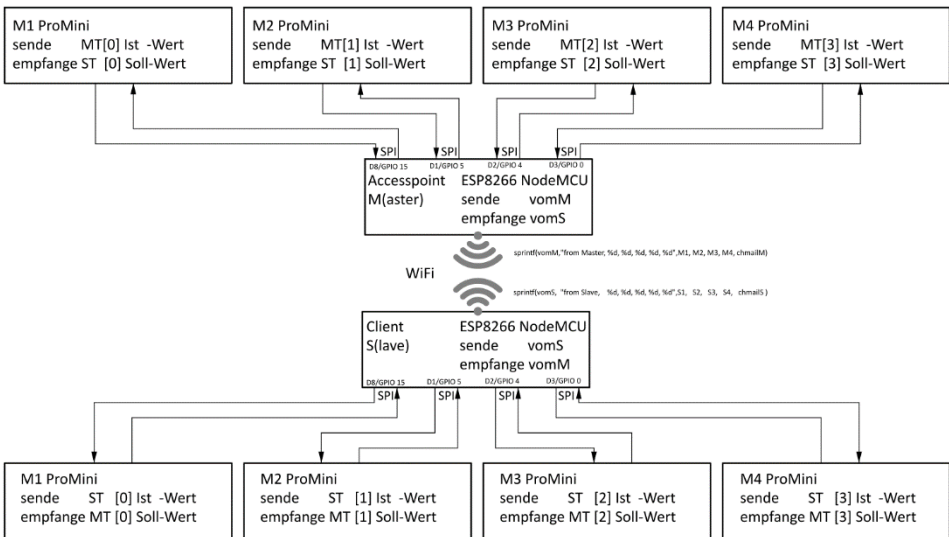


Abb. 4: Diagramm des Datenflusses der RoboTwins

Die SPI-Schnittstelle ist die schnellste, leider aber auch empfindlichste Kommunikation. Das NodeMCU des Masters sammelt förmlich die Werte über den SPI-Transfer ein. Dafür wird eine Bibliothek SPI.h genutzt [GIH23-1]. Der nachfolgende Code-Ausschnitt vom Mastersystem erledigt diesen Transfer. Der Prozess dahinter kann wie folgt beschrieben

werden. Sämtliche ProMinis des Masters erhalten alle Werte der ProMinis des Slaves. Das NodeMCU schaltet die Chip-Select-IO's (CS) zu den ProMinis der Reihe nach auf LOW. Ist CS zu einem ProMini LOW, kann die Kommunikation beginnen und das NodeMCU sendet die Slave-Werte an diesen. Das funktioniert bei bidirektionaler Kommunikation mit einem Schieberegister, vergleichbar mit einem Postfach. Die vier Slave-Werte werden nacheinander gesendet, also in das Postfach gelegt und versandt. Sind die vier Werte durchgestellt, folgt ein 'a'. Für den ProMini ist das 'a' das Startsignal, nun seine vier Werte als Antwort auf die vom NodeMCU empfangenen Buchstaben zu diesem zurückzusenden. Das NodeMCU empfängt also die vier Nachrichten und weist ihnen die Werte MT[0 ... 3] zu. Das Procedere mit den Buchstaben hat auch den Zweck, das Register für die erwartete neue Nachricht zu leeren. Ohne diesen Trick würden die Werte wie Fließtext ohne Unterscheidung oder Kommas durch das Register geschoben werden. Sind alle, zwei mal vier Werte ausgetauscht, schaltet das NodeMCU den CS-IO zum aktuellen ProMini wieder auf HIGH und wendet sich dem nächsten ProMini zu.

Für die WiFi-Verbindung wird die ESP8266WiFi.h-Bibliothek verwendet [GIH23-2]. Damit ist die Kommunikation des Variablentyps 'String' möglich. Allerdings ist es nicht so einfach, die vier als 'int' deklarierten Werte der Potentiometer direkt in einem String zu verpacken. Darum wird der Umweg über eine Variable vom Typ 'char' genommen. Der Zusammenbau beginnt mit dem Text 'from Master' und endet mit der definierten Zahl 'chmailM'. Der Versand erfolgt via UDP, um eine möglichst hohe Geschwindigkeit zu erreichen und nicht auf die Antwort, wie bei TCP, warten zu müssen. War die SPI- oder WLAN-Kommunikation fehlerhaft, kann das empfangende NodeMCU anhand des Anfangstextes und der Endzahl der Nachricht prüfen, ob die Nachricht überhaupt korrekt sein kann. Ist dem so, werden der Text anschließend zerlegt und die vier Integer ausgelesen und weiterverarbeitet.

Für die PID-Regelung der Winkelposition der Hauptachsen mit den ProMinis wird die Bibliothek PID_v1.h verwendet [GIH23-3]. Das reduziert den notwendigen Aufwand auf das Suchen der für den Antrieb passenden PID-Koeffizienten. Diese werden im Programm hinterlegt und dann bei der Kalkulation verwendet. An dieser Stelle sei darauf hingewiesen, dass die Taktung der PWM-Frequenz möglichst hoch, zumindest größer als 31kHz sein sollte. Es wäre auch möglich, mehrere PID-Regler mit verschiedenen Koeffizienten zu nutzen.

Abb. 5 zeigt die stark vereinfachte Grundschaltung der ProMinis für die PID-Regelung der Hauptwellen der RoboTwins. Diese Schaltung ist auch für do-it-yourself-Servos nutzbar. Ein Potentiometer ist koaxial zur Achse der jeweiligen Antriebswelle montiert und generiert den Istwert für den PID-Regler. In dem Fall wurden 340°-Endlospotentiometer ohne mechanische Limits verwendet, um einer Beschädigung vorzubeugen. Mit einem zweiten Potentiometer kann der User exemplarisch einen Sollwert erstellen. Dieser Wert käme bei den RoboTwins vom NodeMCU.

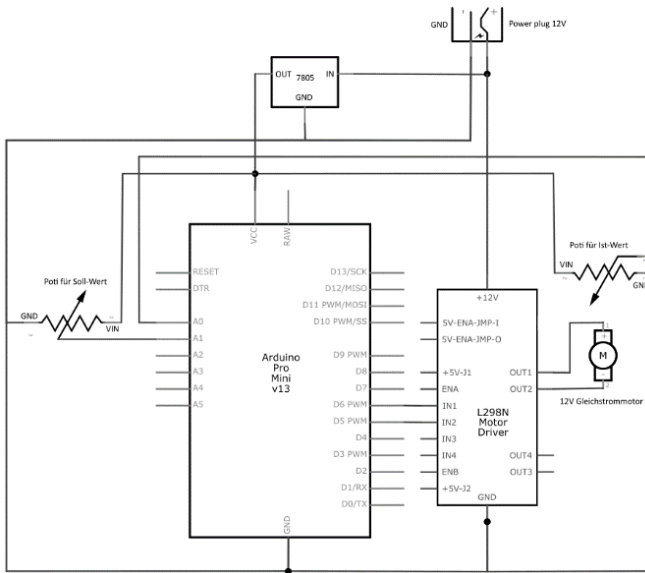


Abb. 5: Schaltplan für eine Achse mit Positionsregelung

3 Zusammenfassung

Die RoboTwins sind Prototypen kommunizierender Mechaniken. Die bidirektionale Kommunikation zwischen Master- und Slave-Aufbau macht Rückkopplungen im Handeln bzw. ein Feedback möglich. Durch ihren modularen Aufbau sind sie für verschiedenste Inhalte in der Lehre der Ingenieurinformatik anschaulich als Demonstratoren geeignet. Die Anwendungsmöglichkeiten bedienen aktuelle Themen in diversen Branchen. Besonders hervorzuheben sind Trainingsszenarien, die dem Anwender ein intuitives Handeln in anspruchsvollen haptischen Aufgaben vermitteln.

Die RoboTwins verfügen über WiFi-Anbindungen. Damit könnten Sie, sofern WiFi verfügbar ist, über weite Distanzen, örtlich voneinander getrennt, über das Internet kommunizieren. Der Master weiß immer, welche Winkel die Gelenke des Slaves einnehmen und andersherum. Die übertragenen Positionsdaten können beliebig große Zahlen annehmen. Somit können auch Lagesensoren, hochauflösende Encoder oder andere Sensoren verwendet werden. Der exemplarische Aufbau der RoboTwins verfügt über vier Freiheitsgrade. Die Zahl der Freiheitsgrade ist beliebig erweiterbar. Lediglich die Verarbeitungsgeschwindigkeit und die Zahl der Ein- und Ausgänge der verwendeten Mikrocontroller limitieren diese. Die als Studie vorgestellten Demonstratoren lassen es zu, sämtliche Komponenten, ähnlich einem Baukasten, durch andersarbeitende Komponenten auszutauschen. Statt der rotatorischen Gelenke könnten translatorische Antriebe verwendet werden. Die DC-Motoren lassen sich durch Schrittmotoren, Servos

oder bürstenlose Antriebe ersetzen. Die Leistung der Motoren ist beliebig skalierbar. Es ließen sich Kraftstrommotoren oder Miniaturantriebe steuern. Selbst anders arbeitende Antriebsprinzipien wie zum Beispiel Hydraulik sind möglich. Die Lagesensoren sind beliebig wählbar. Als Schnittstellen für die Sensoren sind I2C, SPI, analoge Eingänge der digitale IO's möglich.

4 Literaturverzeichnis

- [BIR03] Birmanns, S.: Haptisches Rendern zum Einpassen von hochaufgelösten Molekülstrukturdaten in niedrigaufgelöster Elektronenmikroskopie Dichteverteilungen. Jülich: NIC, 2003.
- [CHA23] Charite [Online]. https://chi.charite.de/leistungen/allgemeine_bauchchirurgie/minimal_invasive_chirurgie/, Stand 6.5.2023.
- [GIH23-1] GitHub, Inc., <https://github.com/arduino/ArduinoCore-avr/tree/master/libraries/SPI>, Stand 6.5.2023.
- [GIH23-2] GitHub, Inc., <https://github.com/arduino/esp8266/tree/master/libraries/ESP8266WiFi>, Stand 6.5.2023.
- [GIH23-3] GitHub, Inc., https://github.com/br3ttb/Arduino-PID-Library/blob/master/PID_v1.h, Stand 6.5.2023.
- [KAS13] Kassner, S.: Haptische Mensch-Maschine-Schnittstelle für ein laparoskopisches Chirurgie-System. Dissertation, TU-Darmstadt, 2013.
- [KER08] Kern, Thorsten A.: Entwicklung Haptischer Geräte: Ein Einstieg für Ingenieure. Berlin, Heidelberg: Springer Science & Business Media, 2008.
- [KLJ17] Kljajic, Z; Schill, E.: Haptische Feedback-Systeme für Displays - Den Tastsinn wieder nutzen. <https://www.elektroniknet.de/optoelektronik/displays/den-tastsinn-wieder-nutzen.147279.html>, Stand 6.5.2023.
- [LIE16] Liedecke, Christoph: Haptische Signale am Fahrerfuß für Aufgaben der Fahrzeugsteuerung. Berlin, Heidelberg, New York: Springer, 2016.
- [MB23] Makeblock Co. Ltd, www.makeblock.com, Stand 6.5.2023.
- [MOR09] Morper, Thomas: Hybride Kraft- und Positionsregelung. 1. Aufl., München: GRIN Verlag, 2009.
- [PRA12] Prats, Mario; Pobil, Ángel P. del; Sanz, Pedro J.: Robot Physical Interaction through the combination of Vision, Tactile and Force Feedback: Applications to Assistive Robotics. Berlin, Heidelberg: Springer, 2012.
- [SCH08] Schiebl, Frank: Force-Feedback unter besonderer Berücksichtigung interner Modelle. 1. Aufl., Pieterlen: Peter Lang, 2008.
- [SOR14] Sørensen, D., A., heise online: <https://www.heise.de/news/Bionische-Handprothese-ermoglicht-Tasten-und-Fuehlen-2106370.html>, Stand 6.5.2023.

**Bildung - Serious XR -
Wissenstransfer mit Augmented und
Virtual Reality (SXR)**

Teaching Optical Principles in XR

Tobias Zengerle¹ David A. Plecher² Salome Flegr³ Jochen Kuhn⁴ and Martin R. Fischer⁵

Abstract: Extended Reality (XR) is seeing a massive rise in interest for the use in educational contexts, with a number of studies pointing out the beneficial effects on learning outcome when using immersive XR platforms, compared to classic educational approaches. However, since the necessary hardware has only recently become more accessible and consumer-level XR is just now beginning to push towards more widespread use, these technologies have not yet fully arrived in everyday university teaching. In this work, we will explore how modern XR technologies can be utilized to create new types of gaining knowledge and investigate how they can enhance the learning experience in actual everyday university practice. To this end, we entered into a cooperation with a chair of physics education (LMU) and developed an immersive Virtual Reality (VR) application for their optical physics lab course for students of human medicine. A usability study with a small number of participants yielded a very strong System Usability Score of 82.52, indicating that the resulting application is easy to use and highly accessible, even to users without prior VR experience and putting it in a great spot to be used in future university teaching.

Keywords: Extended Reality; Virtual Reality; Science Lab Courses; Optical Physics

1 Introduction

Extended Reality (XR) technologies offer versatile visualisation possibilities for concepts that are not visible to the human eye. At the same time, they enable a natural kind of interaction and practical engagement with the presented content. This combination makes the two sub-branches, Augmented Reality (AR) and Virtual Reality (VR) a perfect solution for science laboratory scenarios. In the area of AR, this might entail augmenting a real-world experiment and overlaying additional visualisations to supplement the observations in the experiment. In the case of VR there are two main possibilities. On the one hand you can either faithfully re-create an existing experiment inside the virtual world, or on the other hand you can create novel experimental setups from scratch.

¹ Technische Universität München, School of Computation, Information and Technology, Arcisstraße 21, 80333 München, tobias.zengerle@tum.de

² Technische Universität München, Research Group Augmented Reality, Boltzmannstraße 3, 85748 Garching b. München, plecher@in.tum.de

³ Ludwig-Maximilians-Universität München, Lehrstuhl für Didaktik der Physik, Edmund-Rumpler-Straße 13, 80939 München, Salome.Flegr@physik.uni-muenchen.de

⁴ Ludwig-Maximilians-Universität München, Lehrstuhl für Didaktik der Physik, Edmund-Rumpler-Straße 13, 80939 München, Jochen.Kuhn@physik.uni-muenchen.de

⁵ LMU Klinikum, Institut für Didaktik und Ausbildungsforschung in der Medizin, Pettenkoferstraße 8a, 80336 München, Martin.Fischer@med.uni-muenchen.de

This work will be dedicated to this second approach and the potential benefits it can yield for university students in science lab course scenarios. More specifically, we will explore how established experimental setups can be translated into an immersive VR environment in a way that exploits the technology's strengths to provide new ways of gaining knowledge for students in higher education.

To this end, we entered a cooperation with the chair of physics education (LMU) to create a VR application for the use in their lab course module on optical physics for students of human medicine. The current LIN module (abbreviated from "Linse", German for "lens") features a classic optical rail experiment, through which the students learn about the underlying optical principles of the human visual system. For the LIN-VR application, we aim to take the same underlying abstract principles and re-wrap them with a visual representation that is closer to the actual medical context. This way, we will introduce an interactive cross-sectional model of the human eye with simulated behavior for the basic optical image formation process, that the students can engage with in the virtual environment. For the resulting VR application, two aspects are most important for us: a high level of usability of the system and its interactions, and clear, descriptive visualisations of the optical principles at work in the human eye.

To assess these factors and to identify any existing usage hurdles, we will conduct a usability study with a small batch of participants. A future second evaluation stage will then see the application being used in real teaching practice of the physics lab course. The accompanying study will investigate differences in learning outcome for the use of our application, versus using the analogous optical rail setup.

2 Related Work

The potential application of VR in different areas of education and training has been of great interest in recent years. As VR headsets become more affordable for the end user, schools, universities and private companies are looking for ways to enhance traditional learning methods with these new technologies.

Many previous works have shown the effectiveness and versatility of using VR in a wide range of educational fields [EPK22; Mu23]. Authors have employed VR in the domains of public speaking training [Pa21], leadership development [Ci19], engagement with history and cultural heritage [PWK19] or as a training tool for medical examinations [Wi17; Wi22].

In the field of physics lab course settings, a lot of research has been done for both AR and VR scenarios. Regarding AR, previous work has focused on the effects augmented physics experiments have on knowledge acquisition [Al20; Th22] or on students' cognitive load while using such AR solutions [Th20].

Looking into VR, Sidanin et al. [Ši20] recreated a nuclear physics laboratory experiment in a virtual environment. This work is designed as a pilot project to show the capabilities of VR

to be used as a replacement tool for the highly specialized equipment usually required for such an experiment, which only a few universities have access to. Based on real spectroscopy data they have created a VR simulation for an experiment to determine the mass of the deuterium. In the resulting application, students can swap between interacting with the virtually recreated equipment in the way they would for the real experimental setup, and “leaving” their virtual body to take a look inside the experiment apparatus, allowing them to zoom all the way down to see what happens on the atomic scale. After the interaction, the students were tasked with analyzing the gathered data, like they would for the regular experiment.

A similar “proof-of-concept” approach was applied to the field of optical physics by Engeln and Gómez Puente [EG19]. Here, they set out to develop a virtual lab for university students to safely experiment with lasers and optical measurement techniques. Students were taught concepts like emission and absorption spectroscopy or laser-induced fluorescence, all without the need for expensive and potentially dangerous equipment.

Much of the existing works in the area of VR in physics lab course settings focus on a more or less faithful re-creation of existing experiments in a virtual environment, some with the addition of further visualisations to better convey the taught concepts. This work however aims to go another route and to create an altogether novel experimental setup, using the classic experiment as an inspiration but transferring the underlying concepts into a new representation closer to the specific domain of the lab course content.

3 LIN-VR Application

Currently in the lab course module, the medical students work with an optical rail. This is a classic experiment where different optical elements, including a light source, a projection slide, a screen and various lenses, can be slid along a metal rail to change their distances from each other. Throughout a number of sub-experiments with different lens setups, the students were asked to move the elements on the rail to several given positions. After every step, they took distance measurements between certain elements in order to calculate the optical properties of the system.

While this experiment is well suited to consolidate the students’ knowledge of optical physics through a hands-on approach, it is also highly abstract and far from the actual subject matter of the human eye. However, we intend to present the same content in a way that is much closer to the medical domain and thus easier accessible to students of human medicine. Therefore, in our application, the students will perform the optical experiments on an enlarged cross-sectional model of the eye, which can be seen in Figure 1. This model covers all the basic optical principles that can be investigated through the optical rail and provides additional visualizations for them, that are only possible through the use of XR technologies. Going beyond the optical basics, the eye model also incorporates simulated

physiological behavior to convey how these principles are practically applied in the human visual organ.

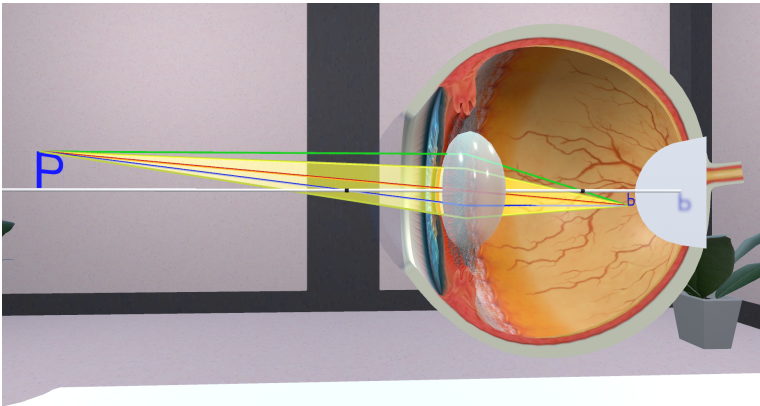


Fig. 1: Closer view of the eye simulation and its individual components

The presented application was created using the game engine Unity⁶ and during development, we worked with the HP Reverb G2 VR Headset⁷. The same device will also be used during evaluation and, ultimately, in the actual lab course practice. We did not build the entire environment completely from scratch, but were kindly given access to the project codebase of a VR demo room on optical physics by the RPTU Kaiserslautern-Landau, which served as the basis for our own application. The respective demo room was created as part of a larger cooperation between the RPTU and the LMU under the umbrella of a government-subsidized project of the name “Au.Ge”⁸.

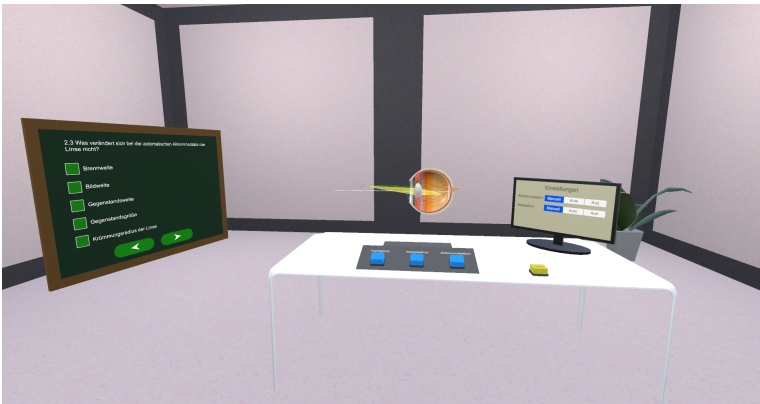


Fig. 2: Overview of the application’s working space and its elements

⁶ <https://unity.com/> (visited on 05/13/2023)

⁷ <https://t1p.de/k79nc> (visited on 04/21/2023)

⁸ <https://rptu.de/uedu/arbeitsfelder/ausbildungskonzepte-af2/auge> (visited on 05/04/2023)

Working through the learning content of our application will be divided into a series of exercises. Through these, the students will gradually build an understanding of the physical workings of the dioptric apparatus and the associated advanced physiological behavior of the eye. Three initial exercises have been implemented on the topics of Basic Image Formation, Accommodation (deformation of the eye lens to put objects at varying distances into focus), and Adaptation (pupillary response to light).

In each exercise, the students will engage with the cross-sectional eye model, make changes to the simulation state and observe what happens as a result. Here, advanced rendering techniques are used to, among other things, illustrate the effects that different lighting conditions or an ill-adjusted eye lens have on the final seen image. Some of the interactive elements of the simulation can be directly grabbed and moved along a rail-like “optical axis”, while other parameters can be changed via a control panel on the table in front of the eye model.

Instructions on what to do in the simulation and what to look for are displayed on the exercise blackboard to the side of the working table, shown in Figure 3. On the same blackboard, the user will also be presented with a set of multiple-choice questions regarding their observations in the simulation or as follow-up for other informational content previously shown on the blackboard. The UI elements on the blackboard can be manipulated with via a basic point-and-click interaction.



Fig. 3: User interaction with one of the multiple choice questions of the application exercises

4 Evaluation

The evaluation of our application will be split into two stages. A preliminary first study, that is included here, deals with the topics of usability and general acceptance of the use of VR in university teaching. A second stage, planned for the near future, will then see the application being utilized during real teaching practice of the lab course module, in order to compare the learning outcome for students using our VR application with that of students working with the classic optical rail experiment. Against this background, the main objective of the first stage is to identify any glaring issues in the application or any poorly usable interactions, that need to be addressed ahead of the upcoming large-scale learning-benefit study and the field usage beyond.

Subject to the usability study of this first stage were 10 university students, ages 19-26, from a variety of academic backgrounds. Half of the participants are studying for a teaching degree in physics, while other subject areas include computer science, history and sociology. The gender distribution of the study was 8 male and 2 female. 3 of the 10 participants indicated, that this study was the first time, they ever came in contact with VR Technology of any kind.

After a 15-20 minute trial run through the application, the participants were presented with a questionnaire, containing the 10 items of the System Usability Scale (SUS) [Br95] and a set of three qualitative open questions on their usage experience. To gather more contextual information on the participants, we also included the question sets for two auxiliary rating scales, which we adopted from the scales manual of the 2015 PISA program [Ma19]. The first one captures the participants' self-reported level of competence using modern digital technology, while the second one is concerned with their general enjoyment of engaging with medical or physiological topics.

The assessment of the 10 filled-out questionnaires yielded an average SUS score of 82.53. Using the qualitative benchmarking scale proposed by Sauro [Sa; Sa11] to put this result into perspective, our application receives the grade "A" with the adjective "Excellent". This is an extremely positive result and it puts our system in the 90-95 percentile range of applications in terms of usability [Sa]. This highly suggests that we have succeeded in our goal of creating an application that is accessible and sufficiently easy to use for employment in university teaching.

Considering the relationship between the SUS scores and the collected contextual information of the individual participants, we compare the average scores for participants that have already engaged with VR systems before with those for participants that had no VR experience, prior to the study. The respective breakdown can be seen in Table 1. First-time users generally rated the system's usability a little higher than non-first-time users. Perhaps this difference can be attributed to the experienced users having a more solid frame of reference and something to compare the system to, in terms of user-friendliness. Or perhaps

the excitement and the novelty factor played a part in a generally positive experience with the system that impacted the first-time users' "generosity" when assessing the system.

Furthermore, in Table 2 we have broken down the SUS scores by the participant's self-reported level of competence in dealing with digital technology. Here we have divided the participants into three categories (high, medium and low) based on their score on the respective auxiliary scale. In this particular study, no participant reported a "low" level of technological competence, as is somewhat expected since we are dealing exclusively with a very young age group. The other participants are distributed among the "high" and "medium" categories in a fashion that reminds of the "First Time vs. Not First Time" table, with a 7 to 3 split between the "high" and the "medium" level. The assembled SUS scores show that subjects who report a high level of technological competence rate the usability slightly lower on average. This is again reminiscent of the situation between first-time VR users and more experienced users. That might suggest, that a higher overall skill level when dealing with technology has a similar effect on the assessment of the application's usability, as the more specialized experience with VR technology in particular.

Last, we take a look if there is a connection between the user's ratings and their interest level in medical topics. Here, we again applied the same categorization into "high", "medium" and "low" based on the score of the scale for the enjoyment of the topics medicine and physiology. The broken-down results in Table 3 indicate an inverse correlation between the SUS score and the participants' level of medical interest. A deeper dive into the responses to the open questions, to search for a potential underlying reason for this connection, came back with no results. The positive and negative feedback the participants shared with us did not display any particular differences based on the sub-category of medical interest they fall into.

5 Discussion

The data collected during the study indicates a highly satisfactory level of usability of the LIN-VR application. While users who are more involved with either the medical field or with digital technology, rated the application slightly more critical, the measured usability scores for these groups still remain at a high level, well within our expectations of the system.

Based on the user's answers to the three open questions, we were able to gather some more valuable insights and received a number of interesting suggestions. For one we were able to identify two cases where the interactions could be further improved. This includes a too small collider for one of the elements that can be grabbed by the user, leading to unintentional manipulation of surrounding elements, and an issue regarding the readability of the labels of the buttons on the control panel.

The participants were also asked whether they could imagine using such a VR application in their own studies and what they would consider to be the most important aspect that such

First VR Usage	N	SUS
no	7	80.75
yes	3	86.67

Tab. 1: Average scores for users with and without prior VR experience

Technology Competence	N	SUS
high	7	80.40
medium	3	87.50
low	0	-

Tab. 2: Scores broken down based on the participants' self-reported level of technology competence

Medical Interest	N	SUS
high	4	75.70
medium	3	85.00
low	3	89.17

Tab. 3: Score breakdown for the participants' level of interest in medical topics

an application would have to address. Here we observed an overwhelming level of openness to the subject matter, as every single student answered that they can indeed imagine using VR-supported learning scenarios in one way or another at university. The largest cluster of answers, regarding their demands of such applications, centered around a good introduction to the system and its basic functionalities, as well as the availability of technical guidance in case they run into a problem with software or hardware during the lab course session.

6 Future Work

The first step going forward should be to implement the takeaways of the usability study and the test users' feedback into the application. This includes cleaning up the two interactions that some users struggled with and adding sound, as well as haptic feedback, to increase the level of immersion for the users.

After that, there are three other planned exercises that need to be implemented to cover the rest of the intended learning content. However, with all the basic systems in place, this shouldn't take an excessive amount of work, especially since the systems were specifically designed with the exercises' planned functionality already in mind. Following this, the application should be ready to be used in the upcoming learning benefit study.

Going beyond the scope of the study, it could be interesting to explore the concept of cooperative learning for the lab course application. As a reference, the old analogous experimental setup of the optics lab course had the student working in pairs. The LIN-VR application technically already provides the ability for multiple people to be in the same virtual room, inherited from the VR demo room it was based upon. However, the respective feature has not seen further pursuit, up to this point, as the learning benefit study is planned to be conducted for a single-person working style. When extending the application in that direction, there should be a special emphasis to create a truly cooperative experience with all involved students actively working together, instead of one student conducting the experiment, while the others are limited to an observer role.

Going further, it would be imaginable to perform another subsequent field study to compare how such cooperative approaches affect the learning experience and its success, compared to the single-person working setup.

7 Conclusion

As part of this work, a VR learning application has been created that can be used in teaching optical physics to students of human medicine, as a replacement for the classic optical rail experiment. The application enables students to engage hands-on with the basic concepts of geometrical optics, utilizing an interactive eye model with simulated physical and physiological behavior.

We initially set out to create a VR learning environment that provides students with clear and descriptive visualizations of the optical principles involved in human vision. At the same time, the application should exhibit a strong degree of usability, to ensure that students will not be held up by poorly usable interactions in the upcoming learning study or the lab course use beyond that.

The conducted usability study indicated that we were successful in both regards. The evaluated data shows that our application is highly accessible and easy to use, even for first-time VR users. Additionally, the handed-out questionnaire revealed a large level of satisfaction with the visualization techniques used to illustrate the various taught concepts and with the hands-on nature of the engagement. More generally, the responses provided by the participants also hinted at a great deal of openness to the concept of using VR applications in university lab course settings.

Further, we have analyzed the study results and the provided feedback and have devised a plan to alleviate the last remaining points of friction in terms of usability and to further enhance the immersion of the engagement. Building on the foundation we laid out, future work will aim at implementing these improvements and finalizing the LIN-VR application to be ready for the use in the optics lab course.

Following this, a large-scale study will be conducted in university use to investigate the potential learning benefits of using our VR application in comparison with the traditional optical rail experiment.

Acknowledgements

We want to express our gratitude to the LMU chair of physics education and thank Dr. Christoph Hoyer, Bernhard Emmer and Vanessa Weber for their invaluable input.

We also want to thank the RPTU Kaiserslautern-Landau department of sport science for providing us access to the project codebase that served as the foundation for our own application and we thank Sergey Mukhametov and Steffen Menne for their support throughout the development process.

Lastly, our thanks go out to all participants of the usability study.

References

- [Al20] Altmeyer, K.; Kapp, S.; Thees, M.; Malone, S.; Kuhn, J.; Brünken, R.: The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. *British Journal of Educational Technology* 51/3, Publisher: John Wiley & Sons, Ltd, pp. 611–628, May 2020, ISSN: 0007-1013, URL: <https://doi.org/10.1111/bjet.12900>, visited on: 06/13/2023.
- [Br95] Brooke, J.: SUS: A quick and dirty usability scale. *Usability Evaluation In Industry* 189/, Nov. 1995.
- [Ci19] Cichor, J.; Egorov, M.; Plecher, D.; Schmid, E.; Peus, C.: Everything Starts with a Handshake: Effects of Character Design and Character Interactions on Leadership Development in Virtual Reality. In: *5th International Augmented Reality & Virtual Reality Conference (IAVR)*. 2019.
- [EG19] Engeln, R. A.; Gómez Puente, S. M.: Does virtual reality help students learn to use optical measurement techniques? *10th International Conference on Physics Teaching in Engineering Education PTEE 2019/*, May 2019, visited on: 05/23/2019.
- [EPK22] Eichhorn, C.; Plecher, D.; Klinker, G.: VR enabling knowledge gain for the user (VENUS), tech. rep., 2022.
- [Ma19] Mang, J.; Ustjanzew, N.; Leßke, I.; Schiepe-Tiska, A.; Reiss, K.: *PISA 2015 Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Waxmann Verlag, 2019, ISBN: 3-8309-9032-4.

- [Mu23] Mukhametov, S.; Wörner, S.; Hoyer, C.; Becker, S.; Kuhn, J.: Unterstützung von Experimenten zu Linsensystemen mit Simulationen, Augmented und Virtual Reality: Ein Praxisbericht. In: Die Zukunft des MINT-Lernens–Band 2: Digitale Tools und Methoden für das Lehren und Lernen. Springer Berlin Heidelberg Berlin, Heidelberg, pp. 63–76, 2023.
- [Pa21] Palmas, F.; Reinelt, R.; Cichor, J. E.; Plecher, D. A.; Klinker, G.: Virtual reality public speaking training: Experimental evaluation of direct feedback technology acceptance. In: 2021 IEEE Virtual Reality and 3D User Interfaces (VR). IEEE, pp. 463–472, 2021.
- [PWK19] Plecher, D. A.; Wandinger, M.; Klinker, G.: Mixed reality for cultural heritage. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, pp. 1618–1622, 2019.
- [Sa] Sauro, J.: 5 Ways to Interpret a SUS Score – MeasuringU, URL: <https://measuringu.com/interpret-sus-score/>, visited on: 06/10/2023.
- [Sa11] Sauro, J.: A practical guide to the system usability scale: Background, benchmarks & best practices. Measuring Usability LLC, Denver, 2011, ISBN: 1-4610-6270-5.
- [Ši20] Šiđanin, P.; Plavšić, J.; Arsenić, I.; Krmar, M.: Virtual reality (VR) simulation of a nuclear physics laboratory exercise. *European Journal of Physics* 41/6, Publisher: IOP Publishing, p. 065802, Oct. 2020, ISSN: 0143-0807, URL: <https://dx.doi.org/10.1088/1361-6404/ab9c90>.
- [Th20] Thees, M.; Kapp, S.; Strzys, M. P.; Beil, F.; Lukowicz, P.; Kuhn, J.: Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior* 108/, p. 106316, July 2020, ISSN: 0747-5632, URL: <https://www.sciencedirect.com/science/article/pii/S0747563220300704>.
- [Th22] Thees, M.; Altmeyer, K.; Kapp, S.; Rexigel, E.; Beil, F.; Klein, P.; Malone, S.; Brünken, R.; Kuhn, J.: Augmented reality for presenting real-time data during students' laboratory work: comparing a head-mounted display with a separate display. *Frontiers in Psychology* 13/, ISBN: 1664-1078, p. 804742, 2022.
- [Wi17] Wilson, A. S.; O'Connor, J.; Taylor, L.; Carruthers, D.: A 3D virtual reality ophthalmoscopy trainer. *The Clinical Teacher* 14/6, Publisher: John Wiley & Sons, Ltd, pp. 427–431, Dec. 2017, ISSN: 1743-4971, URL: <https://doi.org/10.1111/tct.12646>, visited on: 01/09/2023.
- [Wi22] Williams, M.: Virtual reality in ophthalmology education: simulating pupil examination. *Eye* 36/11, pp. 2084–2085, Nov. 2022, ISSN: 1476-5454, URL: <https://doi.org/10.1038/s41433-022-02078-3>.

Generating an Environment for Socializing Between Older Adults in a VR Supermarket

Sebastian Walchshäusl¹, Christian Eichhorn², David A. Plecher³, Tim Simecek⁴,
Gudrun Klinker⁵, Atsushi Hiyama⁶ and Masahiko Inami⁷

Abstract: Socialization is crucial for the well-being of individuals, but elderly people often face challenges due to physical and psychological issues, as well as the COVID-19 pandemic. To address this issue, a novel approach has been developed using a Japanese Virtual Reality (VR) supermarket, where older adults can purchase products that could be delivered to their homes while socializing with others in a familiar environment. The VR supermarket is based on Japanese supermarket shelf layouts, with hand-tracking and gestures used for interaction. Avatars, attached with images of the participants and their synchronized voice input are used to represent them. In a user study involving 14 older adults, the application was found to be effective for social communication, and potential negative effects such as cybersickness were successfully mitigated. The VR supermarket thus provides a valuable tool for older adults to socialize and meet their need to interact with people from the same age group, even in isolation (pandemic scenario).

Keywords: older adults; Virtual Reality; supermarket; cybersickness; hand-tracking; optimization; multiplayer; socializing; Mixed Reality

1 Introduction

The COVID-19 pandemic has posed a significant challenge for older adults (aged 60+) by impairing their ability to socialize, leading to feelings of loneliness [Ha22]. Consequently, there is an increased risk of depression among this demographic [Ca06]. Several factors contribute to the decline in socialization among older adults, including the rising number of individuals classified as older adults due to longer life expectancy [MEH19], the prevalence of health issues, and the difficulty of forming new friendships [BA98].

To address the issue of loneliness and reduced socialization, Virtual Reality (VR) technology can be harnessed. VR offers natural interactions, immersive experiences, and the potential for

¹ Technical University Munich, FAR, Boltzmannstraße 3, 85748 Garching, Germany; s.walchshaeusl@tum.de

² Technical University Munich, FAR, Boltzmannstraße 3, 85748 Garching, Germany; christian.eichhorn@tum.de

³ Technical University Munich, FAR, Boltzmannstraße 3, 85748 Garching, Germany; plecher@in.tum.de

⁴ Technical University Munich, FAR, Boltzmannstraße 3, 85748 Garching, Germany; tim.simecek@tum.de

⁵ Technical University Munich, FAR, Boltzmannstraße 3, 85748 Garching, Germany; klinker@in.tum.de

⁶ University of Tokyo, Information Somatics Lab, 4-chōme-6 Komaba Meguro, 153-0041 Tokyo, Japan; hiyama@star.rcast.u-tokyo.ac.jp

⁷ University of Tokyo, Information Somatics Lab, 4-chōme-6 Komaba Meguro, 153-0041 Tokyo, Japan; drinami@star.rcast.u-tokyo.ac.jp

increased social engagement through hand and head gestures. A VR supermarket, inspired by a Japanese supermarket, has been developed as a means of promoting socialization among older adults. By creating a familiar environment, the VR supermarket provides comfort and allows users to purchase items for later delivery, offering additional convenience. By incorporating multiplayer functionality, individuals can interact with both familiar and unfamiliar users, providing a valuable socialization opportunity, particularly for older adults who may be isolating themselves due to the pandemic or other health conditions.

To evaluate the feasibility of the VR supermarket as a socialization tool for older adults, a user study was conducted involving 14 participants. Each user had the opportunity to interact with their communication partner and explore the supermarket aisles filled with Japanese food products together. The supermarket was specifically designed to minimize the risk of cybersickness while ensuring high usability.

2 Related Work

Moreover, VR applications exist, which can have positive effects on physical health. Exercise-focused games such as BOXVR [fi19], Beatsaber [Be19], and KnockoutLeague [Gr18] have been developed to engage users in physical activity and consequently improve fitness. These games provide an interactive and immersive experience, making workouts more enjoyable and motivating. In addition to physical health, social isolation has been recognized as a significant issue affecting mental well-being, especially among older adults [EF21]. Socialization plays a crucial role in mitigating the harmful effects of social isolation and improving mental health. VR technology has been found to be effective in enhancing physical health and promoting mental well-being [Je18]. By creating virtual environments that facilitate social interactions, VR can provide opportunities for older adults to connect with others, engage in meaningful conversations, and combat feelings of loneliness. Several studies have explored the use of VR applications to address the specific needs of older adults in terms of rehabilitation and mental well-being. For instance, Lin et al. [Li18] focused on socializing among older adults in a VR environment. While this approach proved to be more beneficial to stabilize their mental health compared to passive activities like watching TV, the lack of interactivity limited participants to a static environment without the ability to interact with their virtual surroundings. VR has also been employed to simulate supermarket environments for the purpose of improving the health and cognitive abilities of older adults. However, previous studies have primarily focused on predetermined paths and limited interactions [LU11]. Lee et al. [Le03] aimed to rehabilitate activities of daily living (ADL) through a VR supermarket, showing positive improvements in participants' ADL capabilities. Nevertheless, this study was confined to a single-player scenario, limiting the social aspect of the experience. Mondellini et al. [Mo18] designed a VR supermarket specifically for older adults undergoing cognitive rehabilitation. Participants in this study were able to interact with virtual products and perform tasks, but the verification of the system's effectiveness was limited to younger adults (age < 40). In the work by Baler et

al. [Ba21], older adults were provided with a virtual classroom setting where they could communicate with each other through avatars. Although this application focused solely on socializing and did not extend to other functionalities, it highlighted the importance of avatar representation and communication in creating a more immersive and engaging experience.

Previous work [P119] focused on developing an interactive drinking gadget for older adults and specifically targeted Alzheimer's patients. It aimed to automate the drinking protocol, considering design strategies, user requirements, and convenience. In a study with dementia patients [Ei19], the drinking gadget combined with Serious Games was explored to encourage water consumption. Additionally, "Tangible Chess" was envisioned as a dementia-specific Serious Game [Ei20; Ei22; EPK23; PEK22] which utilized conductive 3D printed chess figures on the touchscreen. It featured a reduced complexity, customizable visual appearance, and a virtual opponent that adjusts to the player's cognitive skill level. The game aimed to enhance cognitive capabilities and support the performance of ADL. In addition to these studies, prior research has investigated the use of Augmented Virtuality to incorporate smartphones in VR supermarket simulations [Ei21]. This approach provided insights into customers' buying decisions when utilizing health apps within the VR supermarket context. Furthermore, another study [EPK22] examined various VR projects, including a travel application that laid the groundwork for multiplayer functionality and emphasized the significance of avatar representation. By incorporating and expanding upon these insights gained from previous research, this work aims to develop a comprehensive and engaging VR supermarket experience tailored to the socialization and well-being of older adults. The utilization of realistic avatars, interactive product interactions, and multiplayer functionality will contribute to creating this immersive environment.

3 Solution Approach



(a) Self-developed layout for the store



(b) Example of a generated product through the product creator



(c) Example of shelf used in the market

Fig. 1: Supermarket interior

One of the objectives encompasses the meticulous design of a Japanese supermarket environment that would offer a comfortable and familiar atmosphere for elderly individuals, accentuated by the inclusion of authentic Japanese products on its shelves. In pursuit of this goal, an in-depth analysis of various Japanese supermarkets was conducted to extract

valuable insights and establish comprehensive design guidelines for the arrangement of shelves within the virtual supermarket. Although Tokyo's supermarkets exhibit diverse layouts, certain discernible patterns were identified, serving as useful references during the creation of a general store layout, as exemplified in Fig. 1a. Notably, strategic considerations were made to encourage social interactions among users, leading to the deliberate expansion of certain spaces beyond the typical dimensions observed in Japanese supermarkets [Wi21]. By implementing these design principles, the project aimed to cultivate an immersive environment that fosters a sense of belonging and facilitates meaningful social engagements among elderly users.

To populate the virtual supermarket, two distinct methods were employed. The first method involved leveraging existing assets from a comprehensive convenience store asset package [Ni21]. This package provided a range of pre-designed elements commonly found in Japanese stores, such as shelves and air conditioners. These assets were carefully selected and integrated into the supermarket application after undergoing modifications to meet the desired requirements. Notably, the shelves were adapted (see Fig. 1c) to align with the specific design specifications of the virtual environment. Furthermore, all the essential features typically present in Japanese supermarkets, including the floor, walls, ceiling, cash registers, entrance, and even the external view visible through a window wall, were incorporated. The second method employed the product creation tool available within the editor framework provided through previous research [Ei21]. This versatile tool facilitated the sculpting of custom product forms according to the desired specifications. Moreover, it provided the flexibility to include product images sourced from a database or produced using manual creation. These images underwent preprocessing to ensure optimal visual fidelity and were automatically positioned on the front-facing surface of the respective virtual products. This placement aimed to closely resemble the appearance of real-life products, hence enhancing the authenticity and immersion within the virtual supermarket environment (see Fig. 1b).

A generic 3D avatar served as a representation of the player, enabling interactions and communication within the virtual environment. To enhance user identification, the face of the avatar was replaced with a processed image of the user. To achieve this, the OpenCV library was employed to remove the background using a mask (see Fig. 2a). The processed image was then utilized for avatar customization, making it easier for older adults to recognize other users in the virtual space.

For user interaction, hand-tracking with the internal Oculus Quest 2 system was incorporated. This technology allowed for the direct mapping of the users' physical hands to the virtual avatar's hands. By leveraging hand-tracking, the technical complexity for users was reduced, thereby promoting a more comfortable experience for older adults engaging with the new technology. To provide artificial input in the absence of traditional controllers that are incompatible with hand-tracking, gestures were employed as an alternative. These gestures served as the means of navigation within the virtual environment. To facilitate teleportation, a specific two-handed gesture was utilized. This gesture involved a palm-based pointing

gesture of the right hand [SRS21], which indicated the desired teleportation location and closing the left hand to accept the transportation (see Fig. 2b). Additionally, a one-handed thumbs-up gesture (refer to Fig. 2c) for both hands was employed to enable snap turning of 45 degrees. Activating this gesture involved holding it for three seconds to choose the desired direction for turning.

Given the susceptibility of older adults to cybersickness, measures were taken to prevent discomfort and enhance the overall experience. Strategies included avoiding low frame rates as the top priority, implementing artificial locomotion instead of involuntary camera movement, and incorporating techniques outlined in [PTC20]. To further mitigate cybersickness, a virtual nose [Wi18] (see Fig. 3b) was included as a reference frame for users, aiding in reducing the potential for discomfort. To further optimize performance and minimize cybersickness, various techniques were implemented: The Unity editor provided essential built-in tools, such as the FPS counter and profiler, for identifying areas requiring improvement. Scene adjustments were made by deactivating or removing performance-intensive objects. Furthermore, memory optimization techniques were employed, including culling Level of Detail (LOD), draw call batching and occlusion culling. These measures significantly improved rendering efficiency and reduced unnecessary draw calls based on object visibility and distance.

The integration of the multiplayer functionality was executed to facilitate interactive experiences among players by employing the Mirror Framework [vi17]. It served as the foundation for enabling seamless interactions within the virtual environment. To achieve this, an array of components were synchronized over the network to ensure smooth collaboration. Notably, hand-tracking, head-tracking, and area-tracking were incorporated into the networked player, guaranteeing precise tracking and recognition capabilities for other users over the air. For effective communication between players, the Dissonance plugin [P17] was utilized to facilitate voice transference. Furthermore, the synchronization of objects within the store, including shelves and asset products, was accomplished through the utilization of network transforms and identities. By employing these techniques, interaction with products could be synchronized across users, fostering a cohesive and shared experience. Consequently, players were able to perceive each other's avatars, synchronize head and hand positions, and engage with objects in the virtual environment. A comprehensive overview of the final Japanese VR supermarket is presented in Fig. 3a. By constructing a prototype of the supermarket and ensuring its explorability, we sought to evaluate its potential for fostering socialization among older adults. By enhancing immersion and promoting social interactions, this initiative aimed to create an environment conducive to the comfortability and engagement of its users.

4 Evaluation

The primary objective of this study was to determine the efficacy of the VR supermarket as a means of facilitating social interactions among older adults, particularly in situations

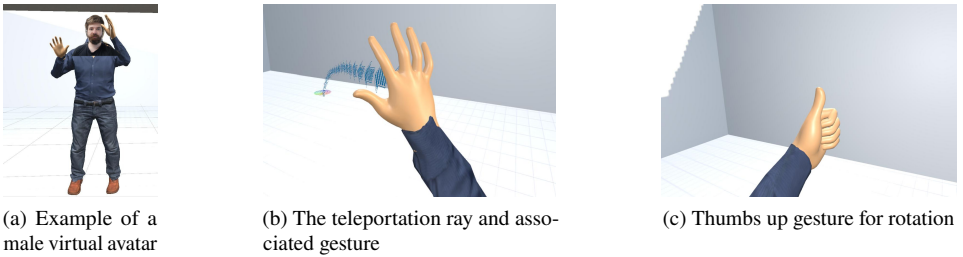


Fig. 2: Player representation and control

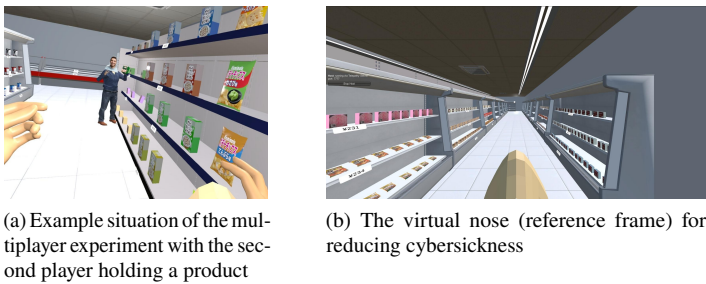


Fig. 3: Multiplayer scenario and cybersickness prevention

where face-to-face communication is not feasible, such as during the COVID-19 pandemic. To achieve this goal, the participants were physically separated by being placed in different rooms, ensuring that sound transmission between them was minimal. For the experiment, each participant utilized an Oculus Quest 2 VR headset and remained seated on a chair for safety purposes. To foster communication and interactions between the participants, an icebreaker activity was integrated into the user study. As part of this task, objects were deliberately misplaced on the shelves of the virtual supermarket. The participants were then challenged to work together and accurately identify these misplaced items. Each participant had access to only partial information and this underscored the significance of communicating with each other to accomplish the task's objective within the virtual environment (refer to Fig. 4).

Upon completing the experiment, the participants were invited to provide feedback through a comprehensive questionnaire. The questionnaire encompassed various aspects, including the usability of interactions, comfort levels during the interaction process, and the likelihood of utilizing the VR supermarket in the future. Participants above the age of 60 were exclusively recruited, with an average age of 70.7, and individuals with dementia, deafness, muteness, blindness, or severe motion sickness were excluded from the study. The recruitment process involved pairing participants within the available time slots. While some participants had prior experience with VR technology, their exposure primarily revolved around non-

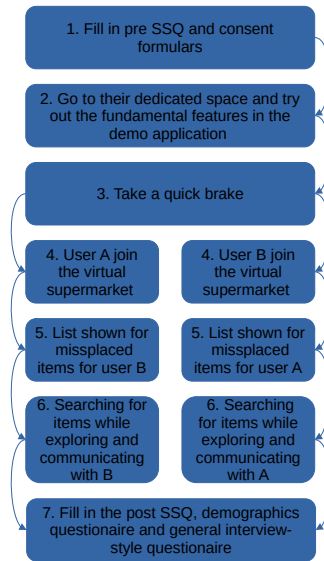


Fig. 4: Step-by-step process for the user study

social applications. Thus, the study aimed to evaluate cybersickness by employing two Simulator Sickness Questionnaires (SSQ): one administered before the experiment and another afterward, thereby assessing any variation in cybersickness levels resulting from the VR experience. The questionnaire followed a neutral interview-style format, commencing with general demographic information, queries regarding participants' familiarity with VR and technology for communication, their engagement in social activities, and their connection with the other user. Subsequent questions focused on usability, immersion, and social presence, with participants providing ratings for each aspect and the opportunity to explain their decisions. These inquiries encompassed aspects such as the resemblance to face-to-face meetings and the ease of gesture-based teleportation. The experimental procedure commenced by capturing a photograph of each participant to be utilized for the avatar representation. Additionally, the initial SSQ was administered. To enhance the feeling of embodiment, a scene featuring a large mirror and a shelf with products was introduced as a test scenario (refer to Fig. 2a). Following the adjustment of headsets for the participants, they were allotted time to familiarize themselves with the VR environment, including testing the rotation feature, teleportation, and interacting with the virtual products. Once the participants felt comfortable, the supermarket simulation commenced, with the objective of maintaining their presence inside the VR environment for 20 minutes, allowing

them to interact with each other while completing the assigned task. Upon completion of the allocated time, the experience concluded, and the participants filled out the second SSQ and the comprehensive questionnaire.

Upon conducting a thorough analysis of the SSQ results, which can be referenced in Fig. (5a), (5b), and (5c), it was evident that the perceived levels of cybersickness among the participants remained generally low throughout the study. The highest observed level of change in cybersickness was 7.48. Interestingly, only three individuals reported a value of 3.47, while all other participants exhibited a value of 0 or even negative values, indicating a decrease in cybersickness symptoms post-experiment. It is noteworthy that the maximum SSQ value is 179.52, thus highlighting the relatively low values obtained in this study. It is important to emphasize that the average SSQ value before the experiment was 8.82, which subsequently decreased to 5.34 after the VR experience. This suggests a decline in cybersickness symptoms, with an average difference of -3.47. The negative average difference indicates that, for the majority of participants, their cybersickness levels reduced following the experiment. Several factors may have contributed to this outcome, such as increased fatigue resulting from physical exhaustion before arriving at the laboratory or a slight overestimation of symptoms prior to the commencement of the experiment. It can be noted that the objective of minimizing cybersickness during the experiment was successfully achieved. This positive outcome indicates the potential for enhancing user comfort and satisfaction during VR experiences, particularly for the elderly population.

The icebreaker activity successfully facilitated socialization and communication among older adults within the VR supermarket. It provided an engaging and interactive experience that encouraged participants to interact, collaborate, and engage in conversations, ultimately achieving the primary objective of its intended purpose.

In terms of interpersonal communication, participants effectively utilized body language and voice transference to facilitate interaction within the VR environment. The voice transference feature was well-received by all users and garnered positive feedback in the comments. The avatars, which accurately represented the participants, were also highly regarded, as they contributed to a sense of co-presence, akin to being physically present in the same room as their conversation partner. This sense of co-presence is crucial for fostering effective communication among users. The visual presence of other avatars within the environment played a significant role in creating a feeling of connection among participants. Moreover, participants expressed satisfaction with the avatars' ability to react appropriately to their verbal instructions and gestures, such as pointing. This responsiveness further enhanced the conversational experience. Analyzing the participants' spatial behavior revealed that they tended to spend a considerable amount of time together when occupying the same location, particularly in areas designed for socializing that offered ample space. However, other locations that required participants to search for misplaced products also served as hot spots where users spent an extended period.

An examination of conversation duration indicated that most participants engaged in consecutive dialogues for approximately 11 minutes and 20 seconds, with an equal distribution of speaking time between both parties. This balanced participation in conversations contributes to a sense of equality and active engagement among users. Overall, the communication aspect of the VR application garnered positive feedback from the majority of users. Although some critiques were voiced, most participants expressed a favorable opinion regarding their ability to perceive the avatars of other users and listen to their voices. Many individuals thoroughly enjoyed the social experiment and envisioned the significant potential for the application in the future. Some participants suggested that additional exposure to VR would enhance their comfort and familiarity with the technology. Furthermore, users proposed various potential use cases for the application, such as virtual window shopping to try on clothing or read books with friends. Notably, there were mentions of the application's potential benefits for physically impaired users. Additionally, participants frequently expressed a desire for more customization options for their avatars and commented on the weight of the VR headset, which some found cumbersome. These findings affirm that the VR application successfully facilitated socialization among users, ultimately contributing to improved mental well-being, as previously mentioned. The scientific exploration and analysis of user experiences shed light on the efficacy of the communication features, users' preferences, and potential areas for further improvement in the application.

Usability plays a crucial role in the success of VR experiences, as poor usability can lead to user frustration and limited engagement. To address this, hand-tracking technology was implemented instead of controllers, as the presence of numerous buttons on controllers could introduce complexity and technological barriers. A natural user interface was a key consideration, allowing users to freely move their hands and navigate the virtual environment through intuitive gestures. However, achieving usability for all older adults using hand-tracking with the Oculus Quest 2 headset remains challenging. Inconsistent hand recognition was observed and mentioned in the questionnaires due to factors such as the recognition window and user movement habits, impeding reliable gesture recognition. To enhance the reliability of gesture recognition, it is necessary to develop methods that strike a balance between ease of recognition and low erroneous input rates. Furthermore, comprehensive testing is essential, encompassing various hand shapes and physical conditions to ensure compatibility across diverse user profiles.

Immersion, as an integral aspect of the VR experience, has the potential to enhance usability, social presence, embodiment, and overall enjoyment. Participants perceived the VR supermarket as a real place they visited, rather than a static scene, demonstrating a high level of immersion. This sense of presence was attributed to the participants' ability to observe and navigate the virtual store's surroundings. Notably, the inclusion of a believable environment, such as providing views of the outside world through virtual store windows, contributed to the participants' feeling of being physically present within an authentic supermarket setting. Consequently, the supermarket environment felt more immersive and engendered a stronger sense of realism among the users.

Overall, the study demonstrated the potential of VR technology to facilitate social interactions among older adults. The findings emphasize the importance of considering usability, immersion, and social presence in designing VR experiences for this population.

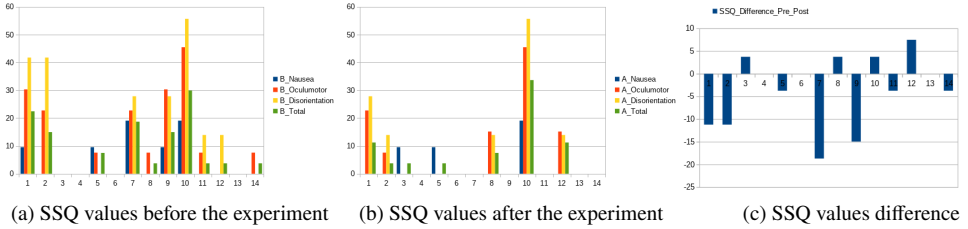


Fig. 5: SSQ graphs

5 Future Work

To improve the current state of the project, the vision of utilizing the application as a virtual shopping solution for older adults would be a desired goal. A collaboration with a Japanese supermarket brand can be established and reworking the layout and products is efficiently possible through the underlying framework. This vision combines socializing with an additional benefit and technology is being mediated as an opportunity to solve daily challenges for older adults. Additionally, more control over the virtual avatar would be useful for the target group, as described in the results of the user study. Users also suggested that there should be a way to be able to invite friends globally. A form of window-shopping may be possible even over distance, where users could meet and exchange and even share clothes or other accessories on their avatars by shopping together with others. Finally, additional consideration could be given to the interaction scheme and gesture recognition.

6 Conclusion

The aim of this project was to develop a Japanese VR supermarket that promotes socialization among older adults by providing a familiar and comfortable environment for social interactions through technology. The primary objectives were to ensure high usability, minimize cybersickness, and maximize immersion. In the user study, older adults had the opportunity to explore the VR supermarket with their conversation partners and interact with supermarket products. Afterward, participants were given interview-style questionnaires to gather feedback on their experiences and evaluate their level of cybersickness. The results showed that the efforts to reduce cybersickness were successful, as no complaints were reported and as seen in the recorded SSQ values. Moreover, the main goal of facilitating communication among users was accomplished, as they were able to interact naturally and comfortably. However, the effectiveness of hand-tracking requires further investigation, as

some users encountered challenges with recognition and accuracy. Immersion was generally perceived positively, and by enhancing the environment through the incorporation of an existing Japanese supermarket layout, the sense of presence can be further increased.

References

- [Ba21] Baker, S.; Waycott, J.; Carrasco, R.; Kelly, R.M.; Jones, A.J.; Lilley, J.; Dow, B.; Batchelor, F.; Hoang, T.; Vetere, F.: Avatar-mediated communication in social VR: an in-depth exploration of older adult interaction in an emerging communication platform. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Pp. 1–13, 2021.
- [BA98] Blieszner, R.; Adams, R. G.: Problems with friends in old age. *Journal of Aging Studies* 12/3, pp. 223–238, 1998.
- [Be19] *Beat-Games: Beatsaber*, Steam, 2019.
- [Ca06] Cacioppo, J. T.; Hughes, M. E.; Waite, L. J.; Hawkley, L. C.; Thisted, R. A.: Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychology and aging* 21/1, p. 140, 2006.
- [EF21] Evans, M.; Fisher, E. B.: Social isolation and mental health: The role of nondirective and directive social support. *Community mental health journal*/, pp. 1–21, 2021.
- [Ei19] Eichhorn, C.; Plecher, D. A.; Lurz, M.; Leipold, N.; Böhm, M.; Krcmar, H.; Ott, A.; Volkert, D.; Hiyama, A.; Klinker, G.: THE innovative reminder in senior-focused technology (THIRST)—evaluation of serious games and gadgets for Alzheimer patients. In: *Human Aspects of IT for the Aged Population. Social Media, Games and Assistive Environments: 5th International Conference, ITAP 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part II* 21. Springer, pp. 135–154, 2019.
- [Ei20] Eichhorn, C.; Plecher, D. A.; Lurz, M.; Leipold, N.; Böhm, M.; Krcmar, H.; Ott, A.; Volkert, D.; Hiyama, A.; Klinker, G.: Combining motivating strategies with design concepts for mobile apps to increase usability for the elderly and Alzheimer patients. In: *Human Aspects of IT for the Aged Population. Healthy and Active Aging: 6th International Conference, ITAP 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II* 22. Springer, pp. 47–66, 2020.

- [Ei21] Eichhorn, C.; Lurz, M.; Plecher, D. A.; Weber, S.; Wintergerst, M.; Kaiser, B.; Holzmann, S. L.; Holzapfel, C.; Hauner, H.; Gedrich, K., et al.: Inspiring healthy Food Choices in a Virtual Reality Supermarket by adding a tangible Dimension in the Form of an Augmented Virtuality Smartphone. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). IEEE, pp. 548–549, 2021.
- [Ei22] Eichhorn, C.; Plecher, D. A.; Trilk, A.; Hiyama, A.; Klinker, G.: GuessingCarbs- a serious game about healthy nutrition in old age combining virtual and tangible components. In: International Conference on Human-Computer Interaction. Springer, pp. 407–415, 2022.
- [EPK22] Eichhorn, C.; Plecher, D.; Klinker, G.: VR enabling knowledge gain for the user (VENUS), tech. rep., TUM, 2022.
- [EPK23] Eichhorn, C.; Plecher, D. A.; Klinker, G.: IEEE VR 2023 Workshops: Workshop: ARES-Augmented Reality Enabling Superhuman Sports & Serious Games (2nd Annual Workshop). In: 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). IEEE, pp. 237–238, 2023.
- [fi19] fitXR: BOXVR, Steam, 2019.
- [Gr18] Grab-Games: Knockout League - Arcade VR Boxing, Steam, 2018.
- [Ha22] Hawkey, L. C.; Buecker, S.; Kaiser, T.; Luhmann, M.: Loneliness from young adulthood to old age: Explaining age differences in loneliness. *International journal of behavioral development* 46/1, pp. 39–49, 2022.
- [Je18] Jerdan, S. W.; Grindle, M.; Van Woerden, H. C.; Boulos, M. N. K.: Head-mounted virtual reality and mental health: critical review of current research. *JMIR serious games* 6/3, e9226, 2018.
- [Le03] Lee, J. H.; Ku, J.; Cho, W.; Hahn, W. Y.; Kim, I. Y.; Lee, S.-M.; Kang, Y.; Kim, D. Y.; Yu, T.; Wiederhold, B. K., et al.: A virtual reality system for the assessment and rehabilitation of the activities of daily living. *CyberPsychology & Behavior* 6/4, pp. 383–388, 2003.
- [Li18] Lin, C. X.; Lee, C.; Lally, D.; Coughlin, J. F.: Impact of virtual reality (VR) experience on older adults' well-being. In: International Conference on Human Aspects of IT for the Aged Population. Springer, pp. 89–100, 2018.
- [LU11] Liu, C.-L.; Uang, S.-T.: Effects of presence on causing cybersickness in the elderly within a 3D virtual store. In: International conference on human-computer interaction. Springer, pp. 490–499, 2011.
- [MEH19] Max, R.; Esteban, O.-O.; Hannah, R.: Life Expectancy, <https://ourworldindata.org/life-expectancy>, (Accessed on 10/10/2022), 2019.

- [Mo18] Mondellini, M.; Arlati, S.; Pizzagalli, S.; Greci, L.; Sacco, M.; Ferrigno, G.: Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: A pilot study on young adults. In: 2018 IEEE 6th international conference on serious games and applications for health (SeGAH). IEEE, pp. 1–8, 2018.
- [Ni21] Nimikko: Japanese Convenience Store, <https://assetstore.unity.com/packages/3d/environments/japanese-convenience-store-195956>, (Accessed on 10/10/2022), 2021.
- [PEK22] Plecher, D.; Eichhorn, C.; Klinker, G.: Roar-role of augmented reality in serious games and superhuman sports./, 2022.
- [PI17] Placeholder-Software: Dissonance Voice Chat, <https://assetstore.unity.com/packages/tools/audio/dissonance-voice-chat-70078>, (Accessed on 10/10/2022), 2017.
- [PI19] Plecher, D. A.; Eichhorn, C.; Lurz, M.; Leipold, N.; Böhm, M.; Krcmar, H.; Ott, A.; Volkert, D.; Klinker, G.: Interactive drinking gadget for the elderly and Alzheimer patients. In: Human Aspects of IT for the Aged Population. Social Media, Games and Assistive Environments: 5th International Conference, ITAP 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part II 21. Springer, pp. 444–463, 2019.
- [PTC20] Porcino, T.; Trevisan, D.; Clua, E.: Minimizing cybersickness in head-mounted display systems: causes and strategies review. In: 2020 22nd Symposium on Virtual and Augmented Reality (SVR). IEEE, pp. 154–163, 2020.
- [SRS21] Schäfer, A.; Reis, G.; Stricker, D.: Controlling teleportation-based locomotion in virtual reality with hand gestures: a comparative evaluation of two-handed and one-handed techniques. *Electronics* 10/6, p. 715, 2021.
- [vi17] vis2k: Mirror, <https://github.com/vis2k/Mirror>, (Accessed on 10/10/2022), 2017.
- [Wi18] Wienrich, C.; Weidner, C. K.; Schatto, C.; Obremski, D.; Israel, J. H.: A virtual nose as a rest-frame-the impact on simulator sickness and game experience. In: 2018 10th international conference on virtual worlds and games for serious applications (VS-Games). IEEE, pp. 1–8, 2018.
- [Wi21] Williamson, J.; Li, J.; Vinayagamoorthy, V.; Shamma, D. A.; Cesar, P.: Proxemics and social interactions in an instrumented virtual reality workshop. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Pp. 1–13, 2021.

Bildung - KI-Bildung - Ein
Workshop zu Aus- und
Weiterbildung über Künstliche
Intelligenz

Assessment of AI literacy – Development and testing of a customizable set of items

Anna Faust¹, Martin Dröge² and Carolin Odebrecht³

Abstract: In this paper, we argue that there are two types of learners who seek to build AI literacy. However, recent studies that provide questionnaires to assess AI literacy seem suitable for one type of learners only. Based on previous literature, we generated items that also addresses the second type of learners. The items can be adapted to the needs of diverse fields. Further, in the following pages, we give insights in the preliminary result of evaluating AI literacy based on these items in the beginning of a course in summer term 2023. Results suggest, that students have very limited AI knowledge in the beginning of the course and lack a linkage of their knowledge.

Keywords: AI literacy, Assessment, Scale Questionnaire, Higher Education

1 Introduction

AI literacy has become a central issue in higher education [Ka2023]. Open access to technologies such as large language models provided by Chat GPT and Bard are key drivers in this context [We23]. These text generating AI technologies have a disruptive impact on higher education institutions: „[...] some academics are calling the bot the death knell for conventional forms of educational assessment.“ [St22]. Moreover, the environment of AI applications is very dynamic resulting in an ever-increasing range of AI-tools and applications (e.g., see [V23] for a comprehensive overview provided by the Virtual Competence Center on Writing and Learning with AI).

In this context, higher education institutions are challenged to integrate knowledge, skills, and attitudes into their curricula that will help students engage with AI technologies in a confident, critical, and safe manner. In recent years, several studies have focused on AI in education. A systematic review of the literature on AI applications in higher education is provided by [Za19]. More recent studies focus on the development of a competency framework for AI literacy [LM20] and the assessment of these competencies [LAR23], [Ca23].

¹ Humboldt-University Berlin, Research Group Digital Knowledge Management in Teaching and Learning, Unter den Linden 6, 10099 Berlin, anna.f Faust@hu-berlin.de, <https://orcid.org/0009-0006-5500-4860>

² Humboldt-University Berlin, Digital History, Friedrichstraße 191-193, 10117 Berlin, martin.droege@hu-berlin.de

³ Humboldt-University Berlin, Corpus Linguistics and Morphology, Dorotheenstr. 24, 10117 Berlin, carolin.odebrecht@hu-berlin.de, <https://orcid.org/0000-0003-4887-7701>

In this paper, we argue that there are two types of students when it comes to AI literacy, but current questionnaires are only suitable for assessing one of the two groups. We aim to add to the existing literature on AI literacy assessment by introducing adaptive items to assess AI literacy. In addition, we provide insights into the preliminary results of using this questionnaire at the beginning of a course on AI methods for text generation. The study is part of the AI-SKILLS project funded by the BMBF and the Berlin Senate Department for Science, Health and Care.

For the remainder of this paper, we will refer to AI literacy as proposed by Long & Magerko, who define AI literacy as "a set of competencies that enable individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool online, at home, and in the workplace." [LM20, p.2].

2 Assessing AI literacy across diverse fields of study

In the AI-SKILLS project, we aim to increase the opportunities for students to acquire AI skills. In this context, we also evaluate AI-related courses from different disciplines (ranging from natural sciences to social sciences to humanities). Some of these courses are being offered for the first time, while others are well established. Moreover, the context in which AI is embedded and used varies considerably from course to course, also depending on the field of study and the specific technology. The general AI competencies remain the same (e.g., recognition of AI, ethical aspects, strengths and weaknesses of AI). However, depending on the technology taught (e.g., cognitive systems, computer vision, robotics, natural language processing), the learning outcomes vary considerably.

Our motivation for evaluating the courses comes from three directions. First, we want to evaluate how they contribute to students' AI competency development based on a pre-post design. Second, based on the students' competence development, we want to identify course sections that might need didactic revision. Third, based on the selection and adaptation of items by course instructors, we aim to build a cross-disciplinary overview of AI-related courses and their competency profile.

The challenge of choosing an appropriate questionnaire to assess AI literacy is rooted in the diversity of fields of study; e.g., teaching AI in philosophy is different from teaching AI in digital history or biology. Recent studies have found different ways to address this challenge. They either define general AI competencies in terms of a technology-independent higher level [VKP22]; or they refer to selected, specific technologies. For example, Long & Magerko define competencies for machine learning and robotics [LM20, p. 6], and Laupichler et al. specify an item on natural language processing [LAR23, p. 7].

However, we observe that there are two types of learners that can be distinguished in terms of AI literacy. The first type of student is looking for a broad overview of AI-related aspects and topics without going into technical details. These students want to understand the basic concepts of AI, the capabilities of AI, and the pros and cons associated with

them. They will be users of AI in the future and therefore focus on interacting with AI, but they are not interested in understanding the technical aspects of a large language model in depth; e.g., a future elementary school teacher who wants to understand ChatGPT and how to handle/teach it in class.

The second type aims at subject-specific in-depth knowledge, e.g., history students who want to learn computer vision to evaluate ancient seals. In the future, these students will not only be users of AI, but will also be able to discuss AI design issues and design AI themselves, based on their mix of domain and technological knowledge.

Considering the first type of students, recently developed questionnaires to assess AI literacy seem to be well suited. However, these questionnaires seem to be less suitable for assessing the AI literacy of the second type of students.

3 Development of questionnaire

To fill this gap in the literature, we created additional items to assess AI literacy with a focus on AI technology. For the assessment tool, we kept items on general AI competencies suggested by previous studies. We then added technology-focused items to the questionnaire. The technology items are formulated in a generic way so that they can be adapted to all relevant technologies. Thus, they can be used in different fields of study. More importantly, these items allow us to assess the competencies of the second type of students who acquire AI literacy using a specific technology.

To include in-depth knowledge of specific technologies, we collected existing items that indicate some type of technology focus and then generalized them. (e.g., "I understand the steps of [AI TECHNOLOGY] and the practices and challenges associated with each step. and "I can describe how [AI TECHNOLOGY] is trained, validated, and tested"). In this way, we have designed items that can be easily adapted by course instructors by adding the appropriate AI technology in square brackets. The questionnaire also includes general items to assess awareness of AI, its history, strengths and weaknesses, and ethical aspects, as proposed by Laupichler et al. [LM20], Long & Magerko [LM20] and Vuorikari et al. [VKP22]. In total, our AI literacy questionnaire consists of 45 items. For the use of the questionnaire, course instructors are asked to select a maximum of 20 items with a focus on the learning objectives of their course and to adapt these items accordingly, if necessary.

With this type of questionnaire design, we have several goals. First, we want to assess students' perceived competence growth. The adaptable items allow the questionnaire to be used for different AI technologies and learning goals. This makes the questionnaire applicable to a wide range of disciplines and suitable for both type one and type two students seeking to acquire AI literacy.

Second, using a pre-post design, we will be able to evaluate the skill development in a course and identify areas with a low or high learning curve. Based on the skill development, we will be able to identify course sections that can be classified as best practice and sections that may need didactic revision.

Item AI Competencies Items - German

- 1 Ich kann erkennen, ob die Dinge, die ich häufig benutze, von künstlicher Intelligenz unterstützt werden.
- 2 Ich kann Beispiele für technische Anwendungen nennen, die durch künstliche Intelligenz unterstützt werden.
- 3 Ich kann Anwendungen nennen, in denen [KI-unterstütztes Natural Language Processing/Understanding] zum Einsatz kommt.
- 4 Ich kann den Unterschied zwischen allgemeiner (oder starker) und enger (oder schwacher) künstlicher Intelligenz erklären.
- 5 Ich kann verschiedene Schwächen künstlicher Intelligenz benennen und bewerten.
- 6 Ich kann verschiedene Stärken künstlicher Intelligenz benennen und bewerten.
- 7 Ich kann Risiken beschreiben, die beim Einsatz von Systemen der künstlichen Intelligenz auftreten können.
- 8 Ich kann die wichtigsten ethischen Fragen im Zusammenhang mit künstlicher Intelligenz benennen und beschreiben.
(z. B. Datenschutz, Beschäftigung, Fehlinformationen, die Singularität, ethische Entscheidungsfindung, Diversität, Voreingenommenheit)
- 9 Ich kann erklären, wie [*Large Language Modelle*] funktionieren.
- 10 Ich kann beschreiben, wie [*Large Language Modelle*]-Modelle trainiert, validiert und getestet werden.
- 11 Ich kann den Unterschied zwischen „supervised learning“ und „unsupervised learning“ im Zusammenhang mit [*Large Language Modellen*] erklären.
- 12 Ich kann erklären, wie Deep Learning mit [*Large Language Modellen*] zusammenhängt.
- 13 Ich kann erklären, was der Begriff "künstliches neuronales Netz" bedeutet.
- 14 Ich kann erklären, warum Menschen eine wichtige Rolle bei der Programmierung, der Auswahl von Modellen und der Feinabstimmung von Systemen der künstlichen Intelligenz spielen.
- 15 Ich verstehe grundlegende Konzepte der Datenkompetenz.
- 16 Ich kann erklären, warum Daten bei der Entwicklung und Anwendung künstlicher Intelligenz eine wichtige Rolle spielen.
- 17 Ich erkenne, dass Computer oft aus Daten (einschließlich meiner Daten) lernen.

Tab. 1: Items selected and adapted by course instructor

Third, we plan to create an overview of AI-related courses and their learning content. For this purpose, the questionnaire will be distributed to all AI-related courses at the university at the beginning of the winter semester. Based on the selection of items by the instructors,

we will be able to develop a general map of AI competencies, which can later be used by students to search for specific competencies and related courses. In addition, we can identify gaps in the curriculum that need to be filled by developing new courses.

4 Testing of questionnaire

The adaptive questionnaire was submitted to the instructor of the course "AI methods for text generation. Evaluating chat GPT in the humanities' questions about text generation and digitality". The course was offered for the first time in the summer semester 2023. The students were partly undergraduates and graduate students. The instructor was asked to select appropriate items and adapt the questionnaire according to the learning objectives of the course. Table 1 shows the items selected and how the instructor adapted the items to fit the course. In doing so, the instructor replaced the [AI technology] in the square brackets with "Large Language Model(s)". The selected 17 items were made available as a questionnaire on the learning platform (Moodle). At the beginning of the first session, 16 students were asked to express their level of agreement with the provided items on a seven-point Likert scale (ranging from 1 strongly disagree to 7 strongly agree). Table 1 provides information on the items selected and adapted by the course instructor.

5 Results

Due to the limited number of pages, we will focus on selected results. The selection of items shows that the course instructor also selected items with technical depth. This indicates that these items were recognized as necessary to assess AI competencies in the considered course and thus supports our argumentation about the two types of students. The alignment of the items with the learning objectives was perceived as easy.

As expected, the perceived AI competence of the students was low at the beginning of the course. This is exemplified by the item " Ich kann Anwendungen nennen, in denen KI-unterstütztes Natural Language Processing/ Understanding zum Einsatz kommt ", with which 12 out of 16 students disagreed and another two students neither disagreed nor agreed. We also found that perceived competence in naming advantages and disadvantages of AI was very low. Results show that students lack the ability to name weaknesses of artificial intelligence. This highlights the need to address the pros and cons of AI, including ethical aspects, in the classroom.

Another observation is that 4 students perceive themselves as competent regarding the role of humans in the development of artificial intelligence systems. In addition, eight students are aware that computers learn from data (including their data). These results might suggest that knowledge is not linked in a meaningful way at the beginning of the course. Some students are aware that systems learn from their data, but are not able to derive ethical considerations / weaknesses and strengths of AI.


6 Discussion

Finally, we would like to point out the methodological limitations of this short paper. Assessing AI literacy through perceived competencies runs the risk of students over- or underestimating their knowledge. This could have the following implications: For example, if students have misconceptions about how large language models work, they may mistakenly overestimate their knowledge level. If students have misconceptions at the end of the course, they may misperceive themselves as competent. For the planned pre-post design, there is also the risk that misconceptions that are corrected during the course will result in the same level of perceived competence in the pre- and post-test. This limits the interpretability of the data in terms of competence growth. Furthermore, the chosen design does not allow us to control whether students might miss a lecture.

References

- [Ca23] Carolus, A. et al.: MAI LS -- Meta AI Literacy Scale: Development and Testing of an AI Literacy Questionnaire Based on Well-Founded Competency Models and Psychological Change- and Meta-Competencies. <https://arxiv.org/abs/2302.09319>, 2023.
- [Ka2023] Kasneci, E., et al.: ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. <https://doi.org/10.35542/osf.io/5er8f>, 2023.
- [LAR23] Laupichler, M. C., Aster, A. and Raupach, T.: Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy. *Computers and Education: Artificial Intelligence* 4, 100126. <https://doi.org/10.1016/j.caeai.2023.100126>, 2023.
- [LM20] Long, D., & Magerko, B.: What is AI literacy? Competencies and design considerations. In *Conference on human factors in computing systems - proceedings*. <https://doi.org/10.1145/3313831.3376727>, 2020.
- [St22] Stokel-Walker, C.: AI bot ChatGPT writes smart essays — should professors worry? The bot is free for now and can produce uncannily natural, well-referenced writing in response to homework questions., <https://www.nature.com/articles/d41586-022-04397-7>, 2022, last accessed 15.05.2023.
- [Vi23] Virtuelles Kompetenzzentrum – KI Ressourcen, <https://www.vkkiwa.de/ki-ressourcen/>, last accessed 15.05.2023.
- [VKP22] Vuorikari, R., Kluzer, S. and Punie, Y., DigComp 2.2: The Digital Competence Framework for Citizens, EUR 31006 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-48882-8, doi:10.2760/115376, JRC128415, 2022.
- [We23] Weßels, D.: ChatGPT – ein Meilenstein der KI-Entwicklung. *Mitteilungen der Deutschen Mathematiker-Vereinigung*. <https://doi.org/10.1515/dmvm-2023-0007>, 2023.
- [Za19] Zawacki-Richter et al.: – Systematic review of research on artificial intelligence applications in higher education – where are the educators? 2019.

Integrating Artificial Intelligence and Robotics into Art Curriculum


Ivan Iovine  ¹

Abstract: Artificial intelligence (AI) and robotics are taking an increasingly central role in the creative industries. In recent years a growing number of artistic works have been produced and embedded with AI or robotic driven technologies. While these technologies are introduced and taught at the university level in creative domains such as architecture and design, their integration into multimedia and electronic arts is often limited or partially addressed. This study presents a practical experience of integrating machine learning and robotics into an art curriculum, specifically targeting students with little programming experience. To teach these disciplines, a constructivism approach is explored, allowing students to learn AI and robotics pipelines through the familiar subject of physical computing. From this starting point, through code samples and experimentations students were able to explore increasingly technical realms such as computer vision and human-robot-interaction.

Keywords: Art Education, Artificial Intelligence, Robotics

1 Coding in Art Education: An Overview of Current Practices

The advancement of digitization and automation, specifically the integration of artificial intelligence (AI) in robotic agents, brings new opportunities not only in industrial contexts but also in creative and educational domains. Extensive research has been conducted over several years to explore the incorporation of AI in education, including its intersection with related fields like computer vision and robotics [WRP20][WB21a]. These studies have focused on the practical implementation and potential advantages of these technologies in various academic disciplines, particularly in technical faculties [WB21b]. In artistic contexts, over the past two decades, first robotics and then AI have led to visual aesthetic explorations [J23][C16] and the development of new classes of artistic works [K01]. However, despite their recent significant role in the experimentation and creation of multimedia artworks, these subjects are not adequately integrated into art curricula. The embedding of technical and programming aspects into design and art programs is currently limited to the discipline of creative coding, whereby the development of generative art and interactive art work is taught. The term creative coding refers to a type of computer programming in which the goal is to create something expressive rather than functional [M04]. To this end, several tools and

¹ University of Art and Design Offenbach am Main, Department of Electronic Media, Schloßstraße 31, Offenbach am Main, 63065, iovine@hfg-offenbach.de,  <https://orcid.org/0009-0006-4421-4483>

platforms have been developed that facilitate learning object-oriented programming and computational thinking [KP15]. These have enabled art education lecturers to teach the fundamentals of computer programming by allowing students to develop art projects in visual and interaction design. In this sense, creative coding has given study access to technical-art topics such as generative art, interactive art and physical computing [S08]. While these subjects serve an important purpose in teaching students how to use hardware and software for creative projects, they no longer fully reflect the contemporary technology stack that encompasses AI and robotics systems. As a result of this, interactive machine learning and explainable AI projects and libraries have been developed for educational purposes in recent years that empower students in non-technical faculties to understand and develop machine learning-based systems in the fields of sound, generative visual art, and interaction [FC10][RBLEO21][PSF11]. Although such tools can cover a wide range of use cases, there are currently no educational tools for teaching AI and robotics in artistic contexts.

2 From Physical Computing to AI and Robotics: A Constructivist Approach

To fill existing gaps in the integration of AI and robotics in arts education, the Robotics Lab at University of Art and Design in Offenbach embarked on an experimental journey. Recognizing the importance of maintaining a cohesive educational framework and avoiding student disorientation, the lab sought to introduce a pipeline for AI and robotics by building on the already well-known discipline of physical computing. This aimed to give continuity and initial familiarity to the topic, which are essential for laying the foundation for learning complex and unfamiliar subjects [GKBE22]. The course "Robotik und Computer Vision für KünstlerInnen" adopted a constructivist approach, empowering students to explore AI and robotics and develop their own artistic projects, aligning with the principle of open-ended learning [H95]. Emphasizing the active role of learners, the course provided a platform for creative expression and enabled students to gain insights into these technologies. By getting the basics of the physical computing platform Arduino [A23] and the Python programming language [P23] from the lecturer, students were able to explore the vast possibilities of AI pipeline technologies and robotics, including computer vision, sensors and actuators. This approach fostered students' sense of ownership as they developed and implemented their own projects, leveraging their new skills to experiment with new art forms, or integrate such pipeline to existing art skill forms (see Fig 1). The course took place during both the summer semester (SS22) and winter semester (WS22/23) and was structured as follows:

- To ensure optimal student support, the course limited the number of participants to a maximum of 8 for the summer semester and 10 for the winter semester.

- The course consisted of 14 teaching units, divided into seven units focused on physical computing and seven units addressed to AI-robotic pipeline.
- The last part of the course introduced students to the fundamentals of Python programming and provided theoretical knowledge about computer vision models. For the practical component, Python code examples were provided, using the Mediapipe machine learning framework [M23a]. The goal was to explore with machine learning models in areas such as pose estimation, hand pose estimation, and facial recognition. In addition, students learned how to implement serial communication to establish a connection between Python and Arduino code. This allowed the aforementioned machine learning models to be used to control actuators, thus demonstrating to students the integration of a whole AI-robotic pipeline.
- Students were given the opportunity, both to develop physical computing projects that combined their current artistic crafts abilities, as well as to experiment and create projects that employed the AI-robotic pipeline.

3 Course outcomes

The results of the course are as follows:

- Out of the total of 18 available slots, 16 students enrolled in the course, with 6 students during the summer semester and 10 students during the winter semester. Most students had initial experience in Arduino programming, while their knowledge of Python was limited to basic usage for terminal-based script handling and execution.
- The dropout rate for the summer semester was 0%, indicating a high level of engagement and commitment from the students. However, during the winter semester, there was a 10% dropout rate, with 1 out of 10 students discontinuing their participation.
- In total, 15 students completed the course. Among these, 13 artworks were showcased, with 7 artworks that combine artistic knowledge already owned with physical computing (Fig. 1), while 6 artworks were developed using the AI-robotic pipeline (Fig. 2, Fig. 3). With the help of code examples and teacher advises on which block of code was editable for experimentation and achieving results of different aesthetics, the students demonstrated the ability to have grasped the basics of Python code and the workings of computer vision based libraries, exploiting these to develop their own projects.



Fig. 1: Rahel Pabst's "Femme Fontaine" work combines student's sculpture abilities with physical computing. ©Rahel Pabst



Fig. 2: Malte Niels Möllers' work, "Handling with Care", involves an interface that uses Mediapipe's Handpose model to track the acceleration of the wrist landmark on the x-axis. This parameter is used to control a stepper motor, which flies a small kite. ©Ivan Iovine



Fig. 3: Marina Köstel's project combines semisolid materiality, physical computing, and machine learning. Using Mediapipe's Handpose model, an interface tracks the y-axis position of the hand. The servo motor that controls the plate is adjusted according to the angle of the user's hand to ensure that the gelatinous material remains in balance. ©Ivan Iovine

4 Conclusion and Future Perspectives

This contribution aimed to explore the teaching of AI and robotics in an artistic educational context, introducing an AI-robotic pipeline based on the Arduino physical computing platform and the Python language, along with the Mediapipe machine learning models. Adopting a constructivist approach, students were taught the fundamentals of these disciplines and given the freedom to develop their own artistic projects using these technologies. Despite the experimental nature of this educational approach and the relatively unfamiliar topics covered in the course, taking into account the outcomes achieved by the students, the experience gained from the two courses titled "Robotik und Computer Vision für KünstlerInnen" offered in the summer semester of 2022 and the winter semester of 2022/23 can be considered positive. The perspective moving forward is to establish a culture of robotics-driven and AI-driven art through the courses and laboratory activities, and to disseminate these knowledge and skills to a larger number of students in the upcoming semesters. By further refining the curriculum, providing ongoing support, and expanding access to these educational opportunities, it is expected that more students will be able to engage with and benefit from the integration of AI and robotics in the arts.

Bibliography

[A23] Arduino. <https://www.arduino.cc/>, Last Access: 30.05.23

- [C16] Champanand, A. J. (2016). "Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artworks". CoRR abs/1603.01768 (2016). arXiv:1603.01768 <http://arxiv.org/abs/1603.01768>
- [DRP20] de Witt, C., Rampelt, F., & Pinkwart, N. (2020). "Künstliche Intelligenz in der Hochschulbildung". <https://doi.org/10.5281/zenodo.4063722>
- [FC10] Fiebrink, R., Cook, P. R. (2010). "The Wekinator: A System for Real-time, Interactive Machine Learning in Music". Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)(Utrecht).
- [GKBE22] Grace, K., Klaassens, B., Bray, L., Elton-Pym, A. (2022). "An Open-Ended Blended Approach to Teaching Interaction Designers to Code". *Front. Comput. Sci.* 4:813889. doi: 10.3389/fcomp.2022.813889
- [H95] Hannafin, M. J. (1995). "Open-Ended learning environments: Foundations, assumptions, and implications for automated design," in *Automating Instructional Design: Computer-Based Development and Delivery Tools* (Berlin; Heidelberg: Springer Berlin Heidelberg), 101–129.
- [J23] Jørgensen, J. (2023). "Towards a Soft Science of Soft Robots. A Call for a Place for Aesthetics in Soft Robotics Research". *J. Hum.-Robot Interact.* 12, 2, Article 15 (June 2023), 11 pages. <https://doi.org/10.1145/3533681>
- [K01] Kac, E. (2001). The Origin and Development of Robotic Art. *Convergence*, 7(1), 76–86. <https://doi.org/10.1177/135485650100700108>
- [M04] Maeda, J. (2004). "Creative code. Aesthetics + Computation". New York, NY: Thames & Hudson, Inc. ISBN: 0500285179
- [M23] Mediapipe. <https://developers.google.com/mediapipe/framework>, Last Access: 30.05.23
- [P23] Python. <https://www.python.org/>, Last Access: 30.05.23
- [PSF11] Perry, P., Schedel, M., Fiebrink, R. (2011). "Wekinating 000000Swan: Using Machine Learning to Create and Control Complex Artistic Systems", *Proc. Int. Conf. New Interfaces Music. Expr.* - 30 May - 1 June 2011, p. 4, 2011.
- [RBLEO21] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B. (2021). "High-Resolution Image Synthesis with Latent Diffusion Models". <https://arxiv.org/abs/2112.10752>
- [S08] Shiffman, D. (2008). "Learning Processing: A Beginner's Guide to Programming Images, Animation, and Interaction". (1st ed.), Morgan Kaufmann, p. 450, ISBN 978-0-12-373602-4
- [WB21a] Wannemacher, K., & Bodmann, L. (2021). "Künstliche Intelligenz an den Hochschulen. Potenziale und Herausforderungen in Forschung, Studium und Lehre sowie Curriculumentwicklung". No. 59.
- [WB21b] Wannemacher, K., & Bodmann, L. (2021). "Künstliche Intelligenz an den Hochschulen. Potenziale und Herausforderungen in Forschung, Studium und Lehre sowie Curriculumentwicklung". No. 59.

Foundations for applied artificial intelligence: enabling and supporting AI teaching and research

Ricardo Knauer¹, Raphael Wallsberger², Marvin Grimm³, Marian Bookhahn⁴, Frank Neumann⁵, Stephan Matzka⁶ and Erik Rodner⁷

Abstract: The field of artificial intelligence (AI) has experienced tremendous advances and will undoubtedly shape a range of vertical industries over the next years. Understanding and applying AI technologies are key skills for the future and of utmost importance for virtually all disciplines. At the University of Applied Sciences Berlin, we are designing an infrastructure and community called KI-Werkstatt that allows us to tackle the challenges of applied research and teaching across vertical domains. The following paper outlines our current efforts and principles.

Keywords: AI education, applied AI research, hardware infrastructure

1 Introduction

Data is sometimes regarded as the oil of the 21st century. Indeed, algorithms with the ability to learn from data are ubiquitous in our daily lives - from voice assistants and chatbots to new medical products [DD22]. The range of services that derive insights from data is vast, and therefore it is increasingly important that students already understand at an early stage how machines can leverage data to do things “intelligently” and apply algorithms that can learn from data. Undergraduate courses do not typically cover core concepts about artificial intelligence (AI), especially in non-technical fields. Furthermore, the computational resources for tackling complex economic and societal problems with AI tools are often scarce for students and even researchers. In the KI-Werkstatt project at the University of Applied Sciences Berlin (HTW Berlin), we aim to tackle each of these issues. In this paper, we first describe how we set up a state-of-the-art infrastructure that acts as an enabler for AI applications in teaching and research, then illustrate our approach to anchor AI literacy as a multidisciplinary core competency set and finally report on the use of our infrastructure for an academic research project in the field of 3D feature extraction. Each of these topics is essential to enable students, lecturers, and researchers.

¹ University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, ricardo.knauer@htw-berlin.de

² University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, raphael.wallsberger@htw-berlin.de

³ University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, marvin.grimm@htw-berlin.de

⁴ University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, marian.bookhahn@htw-berlin.de

⁵ University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, frank.neumann@htw-berlin.de

⁶ University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, stephan.matzka@htw-berlin.de

⁷ University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, erik.rodner@htw-berlin.de

2 Technical prerequisites for AI teaching and research

In machine learning, finding complex associations between input and output and using current state-of-the-art models is extremely resource-intensive. It is therefore crucial to provide sufficient computational resources for applying AI in teaching and research. In the KI-Werkstatt project, we logically split our computing infrastructure into 3 tiers, solving 3 different problems: (1) prototyping, (2) exploration and training as well as (3) inference. This ensures that our AI infrastructure is both scalable, allowing educators and researchers to scale their computational resources as needed, and consistently utilized.

We facilitate rapid prototyping with individualized GPU-powered virtual machines (VMs) and mobile workstations. For exploration and training, we set up a high-performance computing (HPC) cluster with a head node running SLURM, scheduling jobs across multiple Nvidia A100 GPUs. SLURM allows for full flexibility on the server side without being tied to specific machine learning frameworks. However, this also requires our users to use external or self-hosted systems for experiments and dataset tracking. For inference, we rely on a multi-node Kubernetes cluster with A30 GPUs to allow for a more granular control of GPU compute time scheduling. Finally, we deploy a large quantity of small, standardized GPU VMs for students to gain hands-on experience in applying AI while keeping the barrier of entry as low as possible. As can be seen, our infrastructure was designed with flexibility and diversity in mind without the need to achieve maximum compute performance.

3 Enabling and supporting students and lecturers

The ubiquity of AI applications in our personal and professional lives, combined with common misunderstandings about the underlying technology, makes AI literacy a core competency set of the 21st century [LM20]. We believe that the groundwork for AI literacy should be laid at an early stage, and that students of all backgrounds should become familiar with basic AI concepts. Based on the literature [LAP23, LM20, MRS23] and 25 semi-structured interviews with teaching staff and students at HTW Berlin as well as enterprises, we have therefore developed a range of competency-oriented, one- to two-hour-long teaching bricks that can be integrated into and adapted for existing lectures:

- **Basics of data understanding:** In this module, we address how data can be used for decision-making, why the quality of data is important (algorithmic bias) and how it can be evaluated with minimal coding. For example, students interactively explore data using pandas' Plotly backend.
- **Basics of AI:** In "Basics of AI", we discuss domain-specific applications of AI, buzzwords around the technology as well as the business and data understanding

phases of a typical AI lifecycle. Students cooperatively build an exemplary AI algorithm in a no-code setting to classify real-world or virtual playing cards [Mi23].

- **AI and ethics:** This teaching brick tackles the question how ethics relates to data and AI. We zoom in on the ethical, legal and societal implications of trending AI technologies such as large language models, and discuss opportunities and challenges. Students actively apply web-based AI tools⁸ to classify animal images and understand how low-quality data can lead to biased decisions.
- **Trending topics in AI:** In this teaching brick, we sketch the inner workings of selected AI models. State-of-the-art AI models can require dozens of gigabytes of memory for low-latency inferences, so students must commonly resort to using closed-source models that ought not be fed with sensitive data. Our AI infrastructure enables them to apply and evaluate open-source models for sensitive and insensitive data alike. For *the* major trending topic at present, large language models, this flexibility also allows them to experience how prompt engineering strategies differ between larger closed-source and smaller open-source models [e.g., Ko22], and how iterative prompt development is key to avoid hallucinations and biases across model architectures.

To lower the barriers for lecturers to implement these teaching bricks, they are individually supported along the way in integrating these bricks into their own teaching. Additionally, we offer workshops both at the HTW Berlin and the Berliner Zentrum für Hochschullehre (BZHL) that address teaching about and with AI⁹. By providing lecturers with multiple means to incorporate AI into their own teaching context, including an infrastructure for resource-intensive computations, we take important steps towards anchoring AI as a multidisciplinary core competency set in technical and nontechnical domains – both at the HTW Berlin and beyond. We are currently in the process of releasing our materials as open educational resources under the CC BY-SA 4.0 license¹⁰.

4 Enabling AI research at scale

Apart from AI applications in teaching, we also support a range of AI projects in research¹¹. In the following, we focus on the project SparePartAssist as a concrete example. SparePartAssist aims at finding spare parts in a huge (> 1 million) and ever-changing collection of computer-aided design (CAD) models, based on 3D scans. With a specifically developed app, service technicians can identify urgently needed components, when documentation is not available or up-to-date. To train and evaluate robust feature

⁸ <https://teachablemachine.withgoogle.com/>

⁹ <https://bzh1programm.tu-berlin.de/e/FACHBEZOGENE-KI-INHALTE>

¹⁰ e.g., <https://mediathek.htw-berlin.de/video/Introduction-to-kernel-density-estimation-2f3c144721aadae43b8564b52519ca07>

¹¹ <https://kiwerkstatt.htw-berlin.de/alle-projekte>

extractors for 3D scans, a large synthetically generated dataset was created utilizing different rendering techniques including computationally expensive raytracing algorithms. For the first prototyping, comparatively small GPU VMs were used. Having a tested container setup, it was easy to transfer the different pipelines to our HPC cluster for training. Without the HPC cluster, we performed a training run in roughly 3 days on an Nvidia RTX 3090 GPU; with the cluster, we were able to start 3 parallel jobs which finished in less than 1 day without even utilizing the entire capacity of the cluster. It is important to note that our infrastructure also allows for experiments with state-of-the-art models on a much larger scale. Next to high-performance computing, our infrastructure, in particular the mobile workstations, also enabled students to easily test the interplay of the app with external equipment like cameras or depth sensors.

5 Conclusion

The KI-Werkstatt project lays the groundwork for applied AI teaching and research at the HTW Berlin and beyond, by providing a hardware infrastructure for resource-intensive computations, designing educational resources for AI literacy across technical and nontechnical fields, supporting lecturers in implementing core AI concepts into their own teaching as well as facilitating state-of-the-art applied AI research. Open challenges remain how to best trade-off infrastructure performance against maintenance effort, as well as how to foster teaching brick integration with minimal external support.

Bibliography

- [DD22] DIN, DKE: Deutsche Normungsroadmap Künstliche Intelligenz (Ausgabe 2), <https://www.din.de/go/normungsroadmapki/>, accessed: 31/05/2023.
- [LAP23] Laupichler, M.C. et al.: Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy. *Computers and Education: Artificial Intelligence* 4/23, 2023.
- [LM20] Long, D.; Magerko, B.: What is AI literacy? Competencies and design considerations. In (Bernhaupt, R. et al.): *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-16, 2020.
- [Ko22] Kojima, T. et al. Large language models are zero-shot reasoners. In (Oh, A.H. et al.): *Advances in Neural Information Processing Systems*, pp. 22199-22213, 2022.
- [Mi23] Michaeli, T. et al.: Data, trees, and forests – decision tree learning in K-12 education. *arXiv preprint arXiv:2305.06442*, 2023.
- [MRS23] Michaeli, T. et al.: What students can learn about artificial intelligence – recommendations for K-12 computing education. *arXiv preprint arXiv:2305.06450*, 2023.

Unterstützung des juristischen Schreibens durch ein KI-basiertes Schreibsystem

Einblicke aus der KI-unterstützten Lehre im Studium der Rechtswissenschaft

Florian Weber¹, Thiemo Wambsganss² und Matthias Söllner³

Abstract: In einer zunehmend vernetzten Welt gewinnt die Fähigkeit klar strukturiert und überzeugend zu schreiben, insbesondere im Bereich des Rechts, an Bedeutung, da sie eine grundlegende Komponente effektiver juristischer Kommunikation darstellt. Jedoch zeigen Studien, dass die Schreibfähigkeiten von Studierenden in diesem Bereich abnehmen. Um diese Probleme zu überwinden, präsentieren wir ein innovatives KI-basiertes Schreibunterstützungssystem, das Studierenden beim Erlernen des juristischen Schreibens hilft. Das System wurde in mehreren Sitzungen eines Tutoriums an einer deutschen Universität eingesetzt und evaluiert. Die Ergebnisse der Evaluation zeigen die Nützlichkeit unseres KI-basierten Schreibsystems. Unsere Forschung markiert einen wichtigen Meilenstein in der Entwicklung von KI-basierten Lernsystemen für das juristische Schreiben. Sie bildet die Grundlage für zukünftige Fortschritte auf diesem Gebiet und eröffnet neue Möglichkeiten zur Erkundung des Potenzials von KI im Bereich des juristischen Schreibens.

Keywords: Juristische Ausbildung, KI-basiertes Lernsystem, Schreibunterstützung

1 Einleitung

Die Fähigkeit von Studierenden, strukturiert und überzeugend zu schreiben, hat in den letzten Jahrzehnten abgenommen [CH13]. Die mangelnde Fähigkeit, strukturierte Texte zu schreiben, führt zu einer geringeren Überzeugungskraft der Texte und einer schwierigeren Verständlichkeit der Inhalte [KV07]. Insbesondere Studierende im juristischen Bereich stehen vor der Herausforderung, strukturiert und überzeugend zu schreiben, um die komplexen Anforderungen der Rechtslehre und spezifische Probleme darzustellen. Die Herausforderungen beim Verfassen juristischer Texte ergeben sich vor allem aus den strengen Formalismen des juristischen Schreibens und der Schwierigkeit, ein klares und präzises Argument auf der Grundlage von Fakten und Gesetzen zu formulieren. Jurastudierende in Deutschland müssen den sogenannten Gutachtenstil

¹ Universität Kassel, Fachgebiet Wirtschaftsinformatik und Systementwicklung, Pfannkuchstraße 1, 34121 Kassel, weber@uni-kassel.de

² Berner Fachhochschule, Institut Digital Technology Management, Brückenstrasse 73, 3005 Bern, thiemo.wambsganss@bfh.ch

³ Universität Kassel, Fachgebiet Wirtschaftsinformatik und Systementwicklung, Pfannkuchstraße 1, 34121 Kassel, soellner@uni-kassel.de

beherrschen, welcher die Form und den Schreibstil eines Rechtsgutachtens beschreibt (vgl. Tab. 1). Durch die Einhaltung des Gutachtenstils können Studierende ihre schriftlichen juristischen Arbeiten strukturiert, präzise und überzeugend gestalten.

Komponenten des Gutachtenstils	Erläuterung
Obersatz	Der Obersatz bildet die Fragestellung des jeweiligen Prüfungsgegenstandes, der im Folgenden untersucht wird.
Definition	Die Definition benennt die abstrakten Voraussetzungen des im Obersatz aufgeworfenen Prüfungsgegenstands.
Subsumtion	In der Subsumtion wird geprüft, inwieweit die im Sachverhalt geschilderten konkrete Tatsachen mit den abstrakten Voraussetzungen der Definition übereinstimmen.
Konklusion	Die Konklusion ist die Antwort auf den gestellten Obersatz. Somit kommt die Falllösung hier zu einem Endergebnis.

Tab. 1: Komponenten des Gutachtenstils.

Forscher haben, insbesondere im Bereich der Bildungstechnologie, haben Systeme entwickelt, die Studierende beim überzeugenden Schreiben unterstützen [Wa21]. Dennoch sind diese Systeme für Jurastudierende nur von begrenztem Interesse, da sich der Schreibstil in der Rechtswissenschaft aufgrund der strengen Formalismen vom allgemeinen Schreibstil unterscheidet. Forscher und Pädagogen behaupten, dass der gezielte Einsatz von IT-Lösungen in der juristischen Ausbildung hinter den Erwartungen zurückbleibt, da es nur einige wenige Systeme gibt, die Jurastudierende beim Schreiben unterstützen [Al03, Pi08]. Diese Systeme können aufgrund von technischen Einschränkungen Studierenden kein individuelles Feedback auf die eigens geschriebenen Rechtstexte geben. Eine Möglichkeit, Studierende beim Verfassen strukturierter und überzeugender Rechtstexte zu unterstützen, ist die Entwicklung eines intelligenten Schreibunterstützungssystems auf der Grundlage von Modellen des maschinellen Lernens (ML). Intelligente Schreibunterstützungssysteme werden bereits erfolgreich für verschiedene Aufgaben eingesetzt, wie z. B. zur Verbesserung der Grammatik [Ka18]. Die aktuelle Literatur bietet jedoch keine Beweise für die Effektivität eines intelligenten Schreibunterstützungssystems, das Studierenden dabei hilft, Rechtstexte zu verfassen. Daher wollen wir die folgende Forschungsfrage beantworten:

Inwieweit hilft ein KI-basiertes Schreibunterstützungssystem Jurastudierenden bei der Verbesserung ihrer strukturellen und überzeugenden Schreibfähigkeiten?

Um die Forschungsfrage zu beantworten, haben wir ein KI-basiertes Schreibunterstützungssystem entwickelt und in einem Juratorium an einer deutschen Universität mit 42 Jurastudierenden in einem Feldexperiment evaluiert. Zunächst geben wir einen Überblick über den aktuellen Stand der Forschung. Anschließend präsentieren wir *LegalWriter*, unser KI-basiertes Schreibunterstützungssystem. Danach beschreiben wir die Evaluierung des Systems, fassen die Ergebnisse zusammen und geben einen Ausblick auf zukünftige Forschungsbemühungen.

2 Stand der Forschung

Forscher und Pädagogen fordern Bildungseinrichtungen auf, der Entwicklung des überzeugenden und strukturierten Schreibens einen größeren Stellenwert einzuräumen [DNO00]. Folglich haben Forschungsgruppen Systeme entwickelt, die Studierenden beim überzeugenden und strukturierten Schreiben unterstützen sollen. Diese Systeme wurden in verschiedenen Bereichen eingesetzt, z. B. in der Wissenschaft und in Online-Debatten [OH16]. Forscher und Rechtspädagogen stellen jedoch fest, dass der Einsatz von IT-Systemen in der juristischen Ausbildung hinter den Erwartungen zurückbleibt [Be16]. Nichtsdestotrotz gibt es einige Systeme, die Studierenden helfen sollen, überzeugendes juristisches Schreiben und strukturiertes Argumentieren zu lernen. Die meisten dieser Systeme verwenden Ansätze der visuellen Unterstützung [Pi08]. Pionierarbeiten im juristischen Bereich haben gezeigt, dass die Erstellung von Argumentationsdiagrammen die Fähigkeit von Studierenden verbessern kann, qualitativ hochwertige Argumente zu formulieren [Ca03]. Pinkwart et al. 2008 haben das LARGO-System (Legal ArgumentGraph Observer) entwickelt, mit dem Jurastudierende Beispiele für Rechtsauslegungen mit hypothetischen Argumenten grafisch darstellen können [Pi08]. Neben den Ansätzen der visuellen Unterstützung gibt es noch einige andere Systeme, wie z. B. CATO [Al03]. Dieses System unterstützt Studierende bei der Argumentation, indem es Argumente von bereits bestehenden Fällen anbietet. Insgesamt lässt sich festhalten, dass bisher keine geeigneten Systeme verfügbar sind, die Studierende adaptiv unterstützen, strukturierte und überzeugende Falllösungen zu schreiben.

3 Vorstellung des Systems *LegalWriter*

Das Schreibunterstützungssystem *LegalWriter* wird seit 2020 in mehreren iterativen Schritten entwickelt. Um einen Feedback-Algorithmus implementieren zu können, haben wir im ersten Schritt einen Korpus entwickelt, der auf 413 Falllösungen von Jurastudierenden basiert. Basierend auf dem Corpus haben wir drei BERT Transformer Modelle trainiert, die die Komponenten des Gutachtenstils und die Verbindungen der Komponenten klassifizieren. In einem Vergleich haben wir festgestellt, dass die BERT-Modelle im Vergleich zu anderen Transformer-Modellen am besten für unsere Klassifizierungsaufgaben geeignet sind [We23]. Das Systemdesign wurde nutzerzentriert entwickelt und folgt im Wesentlichen der Theorie des Lernens aus Fehlern [Me17]. Das Interface von *LegalWriter* besteht aus drei Hauptkomponenten: einem Texteditor, einer Checkliste und einem Dashboard für juristische Argumentation (vgl. Abb. 1). Lernende erhalten im Texteditor Feedback in Form von farbigen Hervorhebungen im Text. Die Textmarkierungen kennzeichnen die richtige Verwendung der Komponenten des Gutachtenstils (vgl. Abb. 1). Sätze, die nicht in den Gutachtenstil passen, bleiben unmarkiert. Weiterhin erhalten Studierende in der Checkliste Empfehlungen, wie sie den Gutachtenstil besser einhalten können. Die Empfehlungen ergeben sich aus einem Matching der durch die ML-Modelle markierten Komponenten und der Anwendung

bestimmter Heuristiken. Aus diesem Matching ergibt sich auch die Anzahl der Fehler. Zusätzlich erhalten Lernende im Dashboard einen Gesamtüberblick über die verwendeten Komponenten im Text und weitere Hinweise.

Checklist
Die Checkliste wird dir helfen, die juristische Argumentationsstruktur bzw. dem Gutachtenstil besser einzuhalten. [Mehr](#)

127 Wörter

Obersatz ↓
Count: 4 | Errors: 0

Definition ↓
Count: 2 | Errors: 1

Subsumtion [Mehr](#)

Prämisse ↓
Count: 1 | Errors: 0

Schlussfolgerung (Claim) ↓
Count: 1 | Errors: 0

Konklusion ↓
Count: 1 | Errors: 3

Fallstudie ↓ **Nützliche Paragraphen** ↓

Texteditor - Schreiben Sie hier Ihre Falllösung

1. Obersatz:
H könnte einen Anspruch auf Zahlung des Luftreinigers haben, wenn ein Kaufvertrag zwischen H und R wirksam geschlossen wurde.

2. Vertragsprüfung:
Ein Vertrag kommt nach § 145 ffdurch zwei aufeinander bezogene Willenserklärungen zustanden.
Ein Angebot könnten in der Mail von R an H liegen.
Ein Angebot ist eine empfangsbedürftige Willenserklärung, die alle grundlegenden Bestandteile eines Vertrage erhält.
R hat die Mail an den Elektronikhändler geschickt und diese ist auch zugegangen. **Den Kaufpreis nannte R nicht, gab aber an den Ordner für die üblichen Konditionen kaufen zu wollen.**
Insofern kann vom objektiven Empfängerhorizont aus, davon ausgegangen werden, dass R den Gegenstand für 125 Euro kaufen möchte.
Somit hat R gegenüber H ein Angebot von 125€ abgegeben.
Prüfung des Angebots: H müsste das Angebot der R angenommen haben.

Feedback

Wenn du auf den Button "Feedback" klickst, zeigt dir das System an, welche deiner Sätze der rechtlichen Argumentationsstruktur bzw. dem Gutachtenstil zugeordnet werden können. [Mehr](#)

Dashboard für juristische Argumentation
Zusätzlich zur Checkliste bietet dir das Dashboard ein allgemeines Feedback.

Verteilung der Komponenten
In dieser Grafik siehst du, wie viel Prozent der Sätze auf die einzelnen Komponenten entfallen. [Mehr](#)

Konklusion (11%) Obersatz (33%)
Definition (22%) Subsumtion (33%)

Bewertung der Überzeugungskraft
Dieses Diagramm zeigt dir, wie viel von deinem Text der Argumentationsstruktur bzw. dem Gutachtenstil folgt. [Mehr](#)

100%

Abb. 1: Screenshot des KI-basierten Schreibunterstützungssystems *LegalWriter*.

4 Evaluation und Ergebnisse

Um die Effektivität von *LegalWriter* nachzuweisen, haben wir das System in einem Feldexperiment evaluiert. Das System wurde in ein deutschsprachiges juristisches Tutorium implementiert. Die Studierenden sollten im Zuge des Tutoriums im Abstand von einer Woche drei juristische Falllösungen verfassen. Die Studierenden, die teilnehmen wollten, wurden zufällig in zwei Gruppen aufgeteilt. Die Kontrollgruppe löste das juristische Problem mit Stift und Papier und erhielt anschließend eine Erklärung des zu lösenden Falles von einem Tutor in Form einer Musterlösung. Die Behandlungsgruppe benutzte das System *LegalWriter* und erhielt automatisiertes intelligentes Feedback. Die Teilnehmenden in der Behandlungsgruppe hatten ein Durchschnittsalter von 19,57 Jahren ($SD = 5,20$), 9 waren männlich, 14 waren weiblich. In der Kontrollgruppe lag das Durchschnittsalter der Teilnehmer bei 18,32 ($SD = 1,47$), 6 waren männlich, 12 weiblich. Die Versuchsphase dauerte fünf Wochen. Sie bestand aus drei Hauptphasen: einer Vorerhebung, einer individuellen Schreibphase und einer Nacherhebung. In der Vorerhebung erhielten die Studierenden einen Fragebogen mit 12 Fragen. In der Umfrage wurden die demografischen Daten und die Erfahrungen der Studierenden beim Schreiben von juristischen Falllösungen erfasst. Um die Erfahrungen mit dem Verfassen von juristischen Falllösungen zu messen, wurden die Studierenden gebeten, ein juristisches

Problem zu lösen. In der individuellen Schreibphase wurden die Studierenden gebeten, in drei Tutorien jeweils ein juristisches Problem zu lösen. In jedem Tutorium hatten die Studierenden dafür 60 Minuten Zeit. Die Behandlungsgruppe erhielt Feedback von *LegalWriter*, um ihre Falllösungen zu verbessern, und die Kontrollgruppe erhielt ein allgemeines Feedback von einem Dozierenden. In der Nacherhebung haben die Studierenden eine Probeklausur ohne die Hilfe von *LegalWriter* oder des Dozierenden verfasst. Um die Daten aus dem Feldversuch auszuwerten, analysierten wir die Qualität des juristischen Schreibens in den Probeklausuren. Die Qualität des juristischen Schreibens wird durch die Einhaltung des Gutachtenstils bestimmt. Studierende sollten in der Lage sein, logische Schlussfolgerungen basierend auf Definitionen und spezifischen Fakten der Falllösung zu ziehen, um überzeugend zu schreiben. Die Bewertung erfolgte durch den Dozierenden des Tutoriums auf einer Skala von 1 bis 15 und orientiert sich an den juristischen Klausurstandards. Um die Klausurergebnisse auszuwerten, führten wir einen t-Test durch. Die Daten wurden vorher auf Normalverteilung und Homogenität der Varianz geprüft. Um die möglichen Auswirkungen von Störvariablen zu kontrollieren, haben wir die Noten der Falllösungen aus der Vorerhebung zwischen den beiden Gruppen verglichen. Wir erhielten einen p-Wert größer als 0,05 zwischen der Behandlungs- und der Kontrollgruppe. Daher gehen wir auch davon aus, dass es keinen signifikanten Unterschied in der Qualität des juristischen Schreibens zwischen der Kontrollgruppe und der Behandlungsgruppe in der Vorerhebung gibt. Die Analyse der Prüfungstexte zeigt, dass die Behandlungsgruppe bei der Qualität des juristischen Schreibens signifikant besser abschneidet als die Kontrollgruppe (vgl. Tab. 2).

Gruppe	n	Qualität des juristischen Schreibens			
		Mittelwert (Prä-Test)	SD (Prä-Test)	Mittelwert (Post-Test)	SD (Post-Test)
Kontrollgruppe	18	5,42	2,99	8,84	2,59
Behandlungsgruppe	24	6,83	2,43	11,08**	2,19
Total	42	p = 0,095 t-Wert = 1,712		p = 0,004 t-Wert = 3,079	
**p < 0,01					

Tab. 2: Mittelwert und Standardabweichung auf einer Skala von 1-15 (1: mangelhaft, 15: sehr gut).

5 Schlussfolgerung und Ausblick

Mit unserem Beitrag zeigen wir, inwieweit KI-basierte Schreibunterstützungssysteme die juristische Bildung verbessern können. Unser Feldexperiment in einem juristischen Tutorium hat gezeigt, dass Studierende, die begleitend mit dem System *LegalWriter* arbeiten, eine bessere Qualität des juristischen Schreibens aufweisen als Studierende, die in einem klassischen Juratutorium lernen. Zukünftig möchten wir das System in größeren Juravorlesungen einsetzen, um die Wirksamkeit des Systems genauer untersuchen zu können. Geplant sind Experimente mit unterschiedlichen Design-Ausprägungen des

Systems, um nachweisen zu können, welche Ausprägungen des Systems am effektivsten sind. Das System soll mit Large Language Modellen verglichen werden, um die Leistung und Qualität des intelligenten Schreibunterstützungssystems im Vergleich zu diesen Modellen zu beurteilen.

Literaturverzeichnis

- [Al03] Alevan, V.: Using background knowledge in case-based legal reasoning: a computational model and an intelligent learning environment. *Artificial Intelligence* 150/1–2, S. 183–237, 2003.
- [Be16] Beurskens, M.: Neue Spielräume durch Digitalisierung? E-Learning in der deutschen Rechtslehre. In: *ZDRW Zeitschrift für Didaktik der Rechtswissenschaft* 3, Nomos, Nr. 1, S. 1–17, 2016.
- [CH13] Carter, M. J.; Harper, H.: Student writing: Strategies to reverse ongoing decline. *Academic Questions* 26/3, S. 285–295, 2013.
- [Ca03] Carr, C.: Using computer supported argument visualization to teach legal argumentation. In: *Visualizing argumentation: Software tools for collaborative and educational sense-making*, Springer, S. 75–96, 2003.
- [DNO00] Driver, R.; Newton, P.; Osborne, J.: Establishing the norms of scientific argumentation in classrooms. *Science education* 84/3, Wiley Online Library (2000), S. 287–312, 2000.
- [Ka18] Karyuatry, L.: Grammarly as a tool to improve students' writing quality: Free online-proofreader across the boundaries. *JSSH* 2/1, S. 83–89, 2018.
- [KV07] Kendeou, P.; Van Den Broek, P.: The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & cognition* 35/7, S. 1567–1577, 2007.
- [Me17] Metcalfe, J.: Learning from errors. *Annual review of psychology* 68, Annual Reviews, S. 465–489, 2017.
- [OH16] Osborne, J. F.; Henderson, J. B.: The development and validation of a learning progression for argumentation in science. *Journal of research in science teaching*, 53/6, Wiley Online Library, S. 821–846, 2016.
- [Pi08] Pinkwart, N. et al.: Graph Grammars: An ITS Technology for Diagram Representations. In: *FLAIRS Conference*, 2008, S. 433–438
- [Wa21] Wambsganss, T.: Designing Adaptive Argumentation Learning Systems Based on Artificial Intelligence. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, S. 1–5, 2021.
- [We23] Weber, F. et al.: Modeling Structured Persuasive Writing of Case Solutions in German Law Courses to Support Students in Legal Education. *Findings of ACL 2023*. Toronto, Canada, S. 1–15, 2023.

Interdisciplinary Perspectives on the Use of Competency-Based Learning and AI in Personalized Learning Environments

André Selmanagić ¹, Juliane Ahlborn ²

Abstract: Contemporary personalized learning environments emphasize the importance of competency-based learning, while leveraging the growing potential of AI-based techniques. The authors present a joint perspective on personalized learning environments integrating both computer and educational science, considering the influence of educational policies. The publication emphasizes the need for critical reflection of competency-based approaches in educational settings. Furthermore, it discusses the challenges of designing educational settings and the promotion of AI literacy among educators and learners. The paper highlights the importance of interdisciplinary collaboration to ensure a comprehensive understanding and responsible use of AI in education.

Keywords: Personalized Learning Environments, Competency-based Learning, AI Literacy

1 Introduction

The modelling and acquisition of competencies take on a heightened level of importance in light of digital technologies in general, and AI specifically. The data collected through competency-oriented measurement procedures, potentially utilizing AI methods, can offer insights into learners' support needs [WA21]. Additionally, it can be employed to evaluate competency-oriented teaching and learning concepts, or facilitate automated instruction, i.e. the delegation of “educational tasks” to machines [WA21]. This should be critically reflected with regard to the implications of competency-based approaches in educational settings. In this interdisciplinary article we present a joint perspective on the subject of competency-based digital learning environments, integrating both perspectives from computer and educational science. Moreover, it considers the influence of educational policy agenda-setting, adding a comprehensive dimension to the analysis. While the computer science perspective provides insights into the limitations and feasibility of personalized learning environments, the educational science perspective offers a deeper understanding of learning, the educational promotion of competency development and the challenges involved in designing educational settings. During this discussion, the needs and approaches for promoting AI literacy among educators and learners are being explored.

¹ Hochschule für Technik und Wirtschaft Berlin, <https://orcid.org/0000-0002-6457-2791>

² Universität Bielefeld, <https://orcid.org/0000-0002-7960-192X>

2 Politics of Education and Competencies

Given the current pace at which new digital technologies are being introduced and the accompanying mix of enthusiasm and concerns, there is a broad consensus that a solely instrumental approach to subjects such as AI, automation, and datafication inadequately captures the complexity of pedagogical relationships [WMP23]. Educational institutions, schools in particular, are thus confronted and challenged by this dynamic of change. In this respect, educational policy efforts can be observed at the (inter-)national level, exemplified by recommendations of the German KMK [Se16] or the DigComp Framework [EU17]. These initiatives address the emerging needs by developing competency frameworks for cultivating digital competencies among both teachers and learners, and by defining what is essential for the future and what needs to be learned [Ke20]. Macgilchrist's discourse analysis reveals a lack of differentiation among digital technologies in the ongoing debates within educational policy-making regarding the subject of digitalization [Ma23]. The education policy documents predominantly address digitization, technology, and digital media in a very general sense [Ma23]. Moreover, Macgilchrist points out the lack of critical reflection on which types of digital media should be used for what purpose, and with whom, as well as the absence of critical examination of the various forms of learning embedded within diverse educational media [Ma23]. However, due to their agenda-setting nature, these policy documents also have a significant influence on “hopes, visions, and future decisions” [Ma23]. The discourse they shape becomes ingrained in software and, in turn, exerts a substantial impact on users, including learners and educators: “A discourse inscribed in software can – precisely through its seemingly banal, casual, subtle mechanics – have a powerful effect on everyday practices, standardizations and normalizations” (translated from German) [Ma23]. In education policy documents, the development of competencies is often presented as *the* solution to the respective problem. While there is a significant overlap between the prevailing understandings of education, as envisioned by the German term *Bildung*, and competencies, there is a growing tendency to associate the concept of competencies with an understanding of education that is no longer holistic-critical-humanistic, but rather narrowly focused on economic functionality [Le14]. In our view, there is no doubt that competencies are important, but in a holistic-critical-humanistic perspective they are insufficient for the description of learning and educational processes.

3 Learning Environments: Between Competencies and AI

In this context, it becomes evident to what extent this economic-functional understanding of competencies is inscribed in the research and development of software for teaching and learning contexts, consequently exerting an influence on the design of educational settings. “More notably, personalized learning systems will choose and implement these values, even as they are still being debated. This raises a fundamental question: who gets to decide what education should be optimizing for?” [Bu16]. According to Bulger, competency

acquisition has become a primary optimization goal in personalized learning environments [Bu16], which can be attributed to public discourse. Pioneers of personalized learning environments envisioned “that computing could allow [...] a personalized, one-teacher-one-student experience to be available to the masses, replacing or supplementing the one-teacher-many-students experience of most classrooms” in order “to catalyse a revolution in learning” [Bu16]. The common usage of terms like “intelligent tutor” suggests that personalized learning environments aim to emulate the roles of human tutors, educators, and mentors in their functionality. However, five decades of research on these environments have demonstrated the limitations of rule-based algorithms in emulating the educational expertise of human educators [Wa21]. It follows that in the field of education, similar to other domains, statistical methods and AI techniques are regarded as promising approaches, e.g. for the creation of intelligent tutors [AH21]. But in education, the use of AI is currently conveyed “as if it must be right, as if the (anticipated) potential justifies its use - even if teachers who employ AI do not (fully) understand the actual mechanisms and models, and the question of legitimacy and accountability remains entirely unclear” (translated from German) [AH21]. On the one hand, the use of AI highlights the importance of AI literacy for educators and students in their interaction with these systems. On the other hand, it could potentially lead to a new dependency on (large) companies and suppliers that provide these services, raising concerns about data security [MPW21]. In a broader scope, the underlying functionality of a computerized educational agent is similar to that of a human instructor. Personalized learning environments, whether human or artificial, make decisions regarding the adaptation of the learner's environment and instruction, specifically tailored to an individual or a small group of learners [VDC11]. These targets of adaptation are manifold, including the instructional sequence, media types, supportive feedback within learning objects, and adaptive task difficulty [VDC11]. In order to make these decisions, both human and computational agents rely on information about the learner (sources of adaptation) [VDC11]. Such information can encompass various cognitive, affective, and behavioural learner characteristics, such as learning styles, frustration levels, personal goals, and support needs [VDC11]. This is often referred to as the learner model. While learner models can be quite complex in order to capture various facets of a learner [VDC11], competency-based learning strongly emphasizes the importance of competencies as targets of learning and sources of adaptation [Bu16].

The desire to digitally measure competencies and utilize this data to inform educational decisions, necessitates a theoretical modelling of these competencies. “The essential preliminary activity that must be deployed in order to use a computer [...], is that of modelling: only by building an explicit model of a given object can that object become computable” [Pi18]. The unpredictability and contingencies inherent in human learning processes pose a significant challenge when it comes to the modelling and automation of human learning and educational processes. These processes are highly complex and dynamic, making them only vaguely measurable and poorly formalizable or datafiable [AH21]. Consequently, modelling is always a simplification of reality, inherently subjective [Pi18], and potentially ambiguous due to its formulation in human language. Human competencies are latent – they cannot be measured directly. Instead, it is human behaviour, such as

performance in specific activities, that can be quantified [Be10]. As a result, behavioural patterns serve as indicators for approximating competencies [Be10]. The digital measurement of competencies faces various challenges. Certain competencies, such as creativity and problem-solving skills, are too abstract and thus practically unmeasurable. Moreover, there are competencies that can be reasonably approximated, yet their essential indicators cannot be captured due to ethical or data privacy considerations (e.g., when measurement is only possible using personal recordings in image or sound). In the light of technological advancements, certain competencies have only recently become susceptible to digital approximation. Previously, machines were unable to evaluate various indicators that humans could discern, such as identifying relevant keywords in free-text passages. However, advancements in machine learning, particularly in the domain of Natural Language Processing, have now facilitated the extraction of such indicators [Be10]. As a result, there is a growing need for AI literacy among learners and educators to cultivate critical reflection regarding the approximation of competencies.

4 Competency-Based Learning Objects and Education

While the automation of teaching and learning has long been a subject of controversy [Wa21] and has not yet fulfilled the anticipated success, the behaviourist approaches, which were already inscribed in these early forms of automated learning, seem to be perpetuated today [Ma23]. The educational policy's emphasis on competencies is also projected on the individual learning objects that are offered within personalized learning environments and the educational principles these encompass. Macgilchrist summarizes this the following way (translated from German): “It is about solving tasks correctly with pre-defined answers. Success is reported to the learner by the system. Learning content is prepared in a way that makes the learning goal clear to the learner. Learning steps lead in a logical sequence towards the learning goal. With adaptive software, repetition or iterative practice is employed until a significant majority of learning units have been successfully mastered” [Ma23]. While the focus on such learning formats is by no means dictated by the characteristics of personalized learning environments, it is common practice [Ma23]. Generally, it is possible to provide educationally meaningful learning materials in various kinds of digital learning environments that allow for unrestricted exploration and reflection on media, going beyond mere instrumental or target-oriented methods. Examples include exploratory simulations or game-based learning. However, the development of such materials requires significant effort and thus tends to be the exception rather than the norm [LS115]. Furthermore, within personalized, competency-oriented learning environments, there is a tendency to optimize learning objects for measurability [Ro19], thereby ensuring they provide valuable data to the system. Consequently, learning objects that do not provide this information, are considered of lesser significance to the system. Regardless of the learning theory applied, the rise of generative AI will significantly impact the production of learning content. While context such as examples and analogies were previously firmly inscribed in digital learning objects, generative AI for texts and images will soon allow the instantaneous creation of personalized examples tailored to a learner's real-world

context. However, this necessitates learners and educators to engage in critical reflection regarding the implications of such AI technologies.

5 Conclusion

This interdisciplinary contribution aims to provide a unified perspective on the topic of competency-based, digital learning environments, both from the viewpoints of computer science and educational science. In conclusion, the development of personalized learning environments within the teaching/learning context necessitates an interdisciplinary approach to ensure that neither educational nor computer science falls into the trap of developing a reductionist understanding of the other discipline. In this regard, we have identified four different levels that exhibit disciplinary overlaps, but where a need for discussion has been identified: 1) Public discourse plays a significant role in shaping the attitudes and actions of all parties involved. For instance, the understanding of *Bildung* as a combination of competencies, as portrayed in educational policy documents, molds the mindset of those working on digital (personalized) learning environments. This imprint is ingrained in the developed software and influences its users, in turn exerting an influential force on its users. The same discourse also has a formative impact on actors in the field of AI education, including the modelling of this area and the design of learning resources. A central objective of this article is to raise awareness of this type of shaping and its potential consequences. 2) At the level of developing personalized learning environments, it is inevitable to model and effectively simplify learners and learning processes, followed by their approximation using diverse measurement methods. *Bildung* thus undergoes further reduction, this time driven by pragmatic technological considerations. AI techniques serve as indispensable tools for these learning analytics. On one hand, this requires a growing awareness of the inherent complexity associated with such modelling and measurement of learning, along with the responsible utilization of the employed technologies. On the other hand, it is crucial to include both learning theoretical considerations and the integration into educational contexts as essential subjects of discussion. 3) At the learner level, a demand for AI literacy becomes evident, as learners progressively immerse themselves in AI-based learning environments. Moreover, these types of learning environments have faced criticism from the field of educational science for encapsulating a simplified understanding of human learning processes, where learners are directed along specific learning paths that offer limited room for deviation, thereby constraining the exploration of content beyond the predetermined curriculum. Therefore, it is crucial to explore personalized learning environments beyond their mere role as mediators and, as suggested by Macgilchrist [Ma23], and critically reflect on which types of personalized learning environment should be used for what purpose, and with whom. 4) At the educator level, the insights derived from personalized learning environments, specifically learner data, can serve as valuable guidance for (re)designing educational settings and the identification of individual support needs. However, this can only be achieved through a thoughtful approach to these systems, underscoring the importance of AI literacy among educators.

Bibliography

- [AH21] Allert, H.; Hartong, S.: Alternativlose Zukunft? Wie wichtig es ist, sich kritisch mit den Lösungsversprechen der KI auseinanderzusetzen. *On. Lernen in der digitalen Welt* 5/2021, pp.8–9, 2021.
- [Be10] Behrens, J. T. et al.: *An evidence centered design for learning and assessment in the digital world*, 2010.
- [Bu16] Bulger, M.: Personalized learning: The conversations we're not having. *Data and Society* 1/22, pp.1–29, 2016.
- [Eu17] European Commission. Joint Research Centre.: *European framework for the digital competence of educators: DigCompEdu*. Publications Office, LU, 2017.
- [Ke20] Kerres, M.: Bildung in der digitalen Welt: Über Wirkungsannahmen und die soziale Konstruktion des Digitalen. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, pp.1–32, 2020.
- [Le14] Lederer, B.: *Kompetenz oder Bildung. Eine Analyse jüngerer Konnotationsverschiebungen des Bildungsbegriffs und Plädoyer für eine Rück- und Neubesinnung auf ein trans-instrumentelles Bildungsverständnis*. Innsbruck University Press, Innsbruck, 2014.
- [LSI15] Loh, C. S.; Sheng, Y.; Ifenthaler, D.: *Serious Games Analytics: Theoretical Framework*. In (Loh, C. S.; Sheng, Y.; Ifenthaler, D. Hrsg.): *Serious Games Analytics*. Springer International Publishing, Cham, pp.3–29, 2015.
- [Ma23] Macgilchrist, F.: *Diskurs der Digitalität und Pädagogik*. In (Aßmann, S.; Ricken, N. Hrsg.): *Bildung und Digitalität*. Springer Fachmedien Wiesbaden, Wiesbaden, pp.47–71, 2023.
- [MPW21] Macgilchrist, F.; Potter, J.; Williamson, B.: *Shifting scales of research on learning, media and technology*. *Learning, Media and Technology* 4/46, pp.369–376, 2021.
- [Pi18] Pierazzo, E.: *How subjective is your model?* In (Flanders, J.; Jannidis, F. Hrsg.): *The Shape of Data in Digital Humanities. Modeling Texts and Text-Based Resources*. Routledge, London, pp.117–132, 2018.
- [Ro19] Robby Robson: *Assessing competency instead of learning outcomes: AI in Adaptive Education*, 2019.
- [Se16] Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik (KMK): *Strategie der Kultusministerkonferenz "Bildung in der digitalen Welt", Version 1.0 (Entwurf), Stand 27. April 2016*, 2016.
- [VDC11] Vandewaetere, M.; Desmet, P.; Clarebout, G.: *The contribution of learner characteristics in the development of computer-based adaptive learning environments*. *Computers in Human Behavior* 1/27, pp.118–130, 2011.
- [Wa21] Watters, A.: *Teaching machines*. The Mit Press, Cambridge, Massachusetts, 2021.
- [WMP23] Williamson, B.; Macgilchrist, F.; Potter, J.: *Re-examining AI, automation and datafication in education*. *Learning, Media and Technology* 1/48, pp.1–5, 2023.

Entscheidungsbäume in der Mittelstufe

Franz Walgenbach,¹ Lukas Wachter²

Abstract: In einigen kürzlich überarbeiteten Lehrplänen zur Informatik in der Sekundarstufe I taucht das Thema Künstliche Intelligenz insbesondere in Form des Maschinellen Lernens auf. Dieses Abstract gibt einen Vorschlag zur Durchführung einer Unterrichtsreihe zum Thema Entscheidungsbäume, in der neben der Grundidee des Maschinellen Lernens auch Gütekriterien von Modellen sowie der Einfluss der Datenauswahl zum Trainieren des Modells anhand konkreter alltagsnaher Beispiele erarbeitet werden. Die Unterrichtsreihe befindet sich derzeit noch in der Entwicklungsphase und soll im Weiteren Verlauf erprobt werden.

Keywords: Maschinelles Lernen; Entscheidungsbäume; Unterrichtsvorschlag

1 Lerninhalte und -ziele der Unterrichtsreihe

Maschinelles Lernen besitzt eine große Bedeutung für den Alltag. Immer öfter taucht es daher in den Informatik-Curricula der Bundesländer auf, z. B. in NRW [Mi23] oder Bayern [St23] – auch in der Sekundarstufe I. Die Schülerinnen und Schüler lernen die Funktionsweise verschiedener Algorithmen zu den unterschiedlichen Bereichen Maschinellen Lernens (ML), deren Chancen und Risiken sowie gesellschaftliche Auswirkungen kennen. Nachfolgend wird ein Vorschlag vorgestellt, Maschinelles Lernen anhand des inhaltlichen Schwerpunkts *Entscheidungsbäume* am Ende der Sekundarstufe I zu unterrichten.

Die Lerninhalte untergliedern sich in die Vermittlung der Grundidee von ML, Muster in Daten zu erkennen und so ein Modell zur Vorhersage von Eigenschaften unbekannter Datenpunkte zu entwickeln. Die Idee wird anhand von Entscheidungsbäumen konkretisiert. Im Anschluss werden Kriterien zur Bewertung des Modells diskutiert, wie Genauigkeit, Größe des Baums und der Einfluss der Trainingsdaten sowie eine Diskussion des *Bias-Variance-Tradeoffs*. Das führt zu folgenden Lernzielen:

- Die Schülerinnen und Schüler wenden einen Entscheidungsbaum zur manuellen Vorhersage an.
- ... entwerfen einen eigenen Entscheidungsbaum zu vorgegebenen Daten.
- ... erklären die Grundidee des Maschinellen Lernens mit eigenen Worten.

¹ Universität des Saarlandes, Lehrstuhl Modeling and Simulation, Campus E1 3, 66123 Saarbrücken, Deutschland
walgenbach@cs.uni-saarland.de

² Universität des Saarlandes, Lehrstuhl Didaktik der Primarstufe – Schwerpunkt Mathematik, Campus E2 4,
66123 Saarbrücken, Deutschland lukas.wachter@uni-saarland.de

- ... vergleichen verschiedene Entscheidungsbäume zu gleichen Daten miteinander.
- ... erklären den Bias-Variance-Tradeoff in eigenen Worten.

2 Ablauf der Reihe

Die Schülerinnen und Schüler (i. F. SuS) erhalten zunächst einen vorgefertigten Entscheidungsbaum anhand dessen sie Daten (beispielsweise Obst und Gemüse anhand von Größe, Form und Farbe) klassifizieren und sich mittels Think-Pair-Share den Aufbau und die Verwendung eines Entscheidungsbaums selbstständig erarbeiten. Zur Sicherung des Gelernten erhalten die Lernenden anschließend weitere neue Daten, die in ihren Tandems klassifiziert werden sollen. Anhand eines im Vorfeld konstruierten fehlerhaft eingeordneten Beispiels wird ein kognitiver Konflikt ausgelöst (Bsp.: Eigenschaften rot & rund → Entscheidung: Tomate; Tatsächlich: Kirsche), der Anlass zur Diskussion im Plenum über die Güte eines Modells schafft.

Aus den einführenden Erfahrungen der SuS heraus, resultiert die Frage, was einen Baum gegenüber einem anderen besser macht. Die SuS kommen erfahrungsgemäß schnell auf die Idee, den Anteil von korrekt dem Anteil von falsch klassifizierten Datenpunkten gegenüberzustellen und entdecken somit das Gütekriterium der *Accuracy*. Als Lernerfolgskontrolle bietet sich an, dass die SuS die Genauigkeit der in der letzten Einheit festgehaltenen Modelle bestimmen.


Zur Erarbeitung des Bias-Variance-Tradeoffs erhalten die SuS verschiedene Entscheidungsbäume im Kontext Phishing-Mails (Features: u. a. Domain, Anhang, Link, ...) als Experten bzw. Hypothesen mit unterschiedlicher Komplexität hinsichtlich der Größe der Bäume. Hierbei entdecken die SuS, aufgeteilt in zwei Großgruppen, dass sowohl ein sehr kleiner, als auch ein sehr großer Entscheidungsbaum – trotz jeweils guter Ergebnisse auf den Trainingsdaten – bei unbekanntem Daten nur schlechte Vorhersagen trifft. Dabei stoßen die SuS zwangsläufig auch auf die Problematik des Overfittings, also des Memorierens der Trainingsdaten. Darauf basierend wird die Unterscheidung von Trainings- und Testdaten zur Bewertung von Modellen hinsichtlich der bereits bekannten Gütekriterien motiviert.

Literaturverzeichnis

[Mi23] Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen: , Kernlehrplan für die Sekundarstufe I Gymnasium in Nordrhein-Westfalen. https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SI/g9_wpi_f_klpentwurf_vb_2023_01_23.pdf, 2023. Letzter Aufruf: 2023-05-12.

[St23] Staatsinstitut für Schulqualität und Bildungsforschung: , Lehrplan Informatik 11 (NTG). <https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/11/informatik/ntg>, 2023. Letzter Aufruf: 2023-05-12.

Der KI-Campus 2.0 an der FernUniversität in Hagen – lernen, vernetzen, entdecken!

Caroline Berger-Konen ¹, Jessica Felgentreu² und Bettina Ötvös³

Abstract: Im Hochschulbereich wird verstärkt über das Lehren und Lernen mit und über KI diskutiert. Im Fokus steht dabei das Erfordernis, KI- und Datenkompetenzen zu stärken und weiterzuentwickeln. Das Verbundprojekt KI-Campus 2.0 greift genau diese Notwendigkeit mit zahlreichen Online-Kursen, Videos, Podcasts und Vernetzungsmöglichkeiten auf. Der Beitrag vermittelt praxisnah, wie die FernUniversität in Hagen im Rahmen ihrer Arbeitspakete KI-Campus-Lernangebote samt Micro-Degrees strukturiert in die Hochschullehre integriert und welche Mehrwerte im Rahmen von Open Think Tanks am KI-ExpertLab Hochschullehre durch Wissenstransfer und Vernetzung entstehen.


Keywords: KI in der Hochschullehre, KI-Lernangebote und Micro-Degrees, FernUniversität in Hagen, KI-Campus 2.0, Communities of Practice, Open Think Tanks, KI-HUB NRW, KI und Didaktik

1 Einleitung

Mit dem Hype um KI-Anwendungen - wie zuletzt ChatGPT - treten nicht nur vermehrt Fragen zur technologischen Souveränität und zur Nutzung von Potentialen KI-basierter Anwendungen für Bildungsprozesse in der Hochschullehre in den Vordergrund. Die Forschungsarbeiten und Diskussionen rücken u. a. eine zentrale Notwendigkeit in den Vordergrund: Die Stärkung von Kompetenzen (“AI Literacy”). Genau dies ist das Hauptziel im Verbundprojekt KI-Campus 2.0 (KI-C). Der KI-Campus (www.ki-campus.org) ist die Lernplattform für Künstliche Intelligenz mit kostenlosen Online-Kursen, Videos, Podcasts und zahlreichen Vernetzungsmöglichkeiten zur Stärkung von KI- und Datenkompetenzen. Als Forschungs- und Entwicklungsprojekt wird der KI-Campus vom Bundesministerium für Bildung und Forschung (BMBF) gefördert.

Die FernUniversität in Hagen (FeU) ist im Projekt KI-Campus 2.0 mit dem Lehrgebiet Bildungstheorie und Medienpädagogik, dem Zentrum für Lernen und Innovation und dem Zentrum für Digitalisierung und IT, in insgesamt sechs Arbeitspaketen vertreten. Im

¹ FernUniversität in Hagen, Lehrgebiet Bildungstheorie und Medienpädagogik, Universitätsstr. 33, 58097

Hagen caroline.berger-konen@fernuni-hagen.de, 

² FernUniversität in Hagen, Lehrgebiet Bildungstheorie und Medienpädagogik, Universitätsstr. 33, 58097
Hagen Einrichtung/Universität, jessica.felgentreu@fernuni-hagen.de

³ FernUniversität in Hagen, Lehrgebiet Bildungstheorie und Medienpädagogik, Universitätsstr. 33, 58097
Hagen Einrichtung/Universität, bettina.oetvoes@fernuni-hagen.de

Rahmen der Informatiktagung berichten die Autorinnen über ihre Praxiserfahrungen im Hinblick auf die Aktivitäten und Formate, die die Integration von KI-C-Lernangeboten und Micro-Degrees, die Initiative der Open Think Tanks am KI-ExpertLab Hochschullehre sowie den Aufbau eines Didaktik Labs betreffen. Im Folgenden werden die thematische Ausrichtung und Ziele der Teilprojekte sowie die Bedeutung dieser für andere Standorte in Kürze erläutert.

2 Teilprojekte im KI-Campus 2.0

2.1 Teilprojekt „KI-Campus-Lernangebote für die Hochschullehre“

Der KI-Campus möchte nicht nur rein selbstgesteuerte Lernangebote zum Thema KI anbieten, sondern auch Lehr-/ Lernszenarien entwickeln, pilotieren und umsetzen, in denen die digitalen Inhalte und Formate strukturiert in die Hochschullehre integriert werden. Das Lehrgebiet Bildungstheorie und Medienpädagogik der FeU plant in diesem Teilprojekt die Entwicklung und Erprobung innovativer Blended Learning- und Selbstlernformate u. a. mit einem überfachlichen Bildungsangebot zu KI in der Hochschulbildung. KI-C-Lernangebote sind curricular und auf bildungswissenschaftlicher und interdisziplinärer Studiengangs- und Weiterbildungsebene zu verankern. Die KI-C-Lernangebote werden genutzt, adaptiert und in spezifischen mediendidaktischen Formaten strukturiert integriert. Für den Einsatz von KI-C-Lernangeboten, digitalen Formaten und Inhalten in der Hochschullehre wurden bereits Zielgruppen und thematische Schwerpunkte identifiziert und didaktische Konzepte für die Hochschulpraxis entwickelt, Kurse und/oder Mikro-Lernformate zielgruppenadäquat ausgewählt, zusammengestellt und ihm Rahmen von Integrationsgesprächen mit relevanten Akteuren in Studiengängen und der Weiterbildung der FeU vorgestellt. Die Kurse werden darüber hinaus auf ihre Anerkennung bzw. Anrechenbarkeit hin geprüft. Geplant sind mithin ein eigenes KI-C-Lernangebot einzusetzen, die Auswahl von KI-C-Kursen, die bereits auf dem KI-Campus als Micro-Degree-Programm zu absolvieren sind sowie die Kuratierung von KI-C-Lernangeboten und die Pilotierung von Blended Learning- und Qualifizierungsformaten für die wissenschaftliche Weiterbildung. Die FeU stellt im Rahmen des Vortrags erste Erfahrungen, Möglichkeiten und Herausforderungen vor.

2.2 Teilprojekt „Anwendungsorientierte Erprobung von Micro-Degrees als neue Qualifizierungsformate“

Die Bundesregierung will in ihrem Koalitionsvertrag die Einführung von Micro-Degrees prüfen. Auch die Hochschulrektorenkonferenz (HRK) empfiehlt eine proaktive Auseinandersetzung mit den digitalen Bildungsnachweisen [vgl. Ho20], denen aufgrund der rasanten technologischen und der daraus resultierenden gesellschaftlichen, privaten und beruflichen Entwicklungen eine wachsende Bedeutung zukommt. Auch auf

internationaler Ebene hat sich die Europäische Kommission bis 2025 das Ziel gesetzt, „an der Entwicklung eines europäischen Ansatzes für Micro-Credentials“ [Eu20, S. 21] zu arbeiten. Micro-Credentials als „digitale Zertifikate, die einen bestimmten Kompetenzerwerb dokumentieren“ [Fl22, S. 11] und als inhaltlich aufeinander abgestimmtes (Lern-)Set aus mindestens drei Online-Kursen einen Micro-Degree auf dem KI-Campus ergeben [vgl. ebd.], bieten vielfältige Potenziale zur neuen Gestaltung von Hochschullehre an der FernUniversität in Hagen. Die FeU zielt innerhalb des Forschungsprojektes u. a. darauf ab, Micro-Degrees als neue Qualifizierungsformate anwendungsorientiert zu erproben und strukturiert in die eigene Hochschullehre und wissenschaftliche Weiterbildung zu verankern. Hierbei werden KI-Campus-Lernangebote kuratiert, die maßgeblich für die Anerkennungs- und Anrechnungsverfahren sind.

2.3 Teilprojekt „Communities of Practice am KI-Expert-Lab Hochschullehre“

Am KI-Campus sind drei KI-ExpertLabs zu folgenden Schwerpunkten angesiedelt: Hochschullehre, Schule und Medizin. Die ExpertLabs verstehen sich als Open Think Tanks und haben das Ziel, den Transfer von Expertise und Methodenwissen rund um das Thema KI im Kontext der Schwerpunkte zu stärken. Das Lehrgebiet Bildungstheorie und Medienpädagogik der FeU ist für die Organisation und Durchführung von Community-Aktivitäten am KI-ExpertLab Hochschullehre verantwortlich. Die Mitwirkenden beschäftigen sich mit Herausforderungen an Hochschulen, die durch das Thema KI entstehen, beispielsweise die KI- Kompetenzvermittlung, die curriculare Einbindung des Themas KI, die sinnvolle Nutzung von KI-Technologien in der Lehre, ethische Fragestellungen und vieles mehr. Diese Herausforderungen sind zugleich die Themenschwerpunkte, die in den Open Think Tanks – wie in einer Zukunftswerkstatt - diskutiert werden und zu denen neue Herangehensweisen oder auch Handlungsempfehlungen entstehen. Im Rahmen des Vortrags werden erste Erkenntnisse, Diskussionspunkte und Fragestellungen aus den Open Think Tanks vorgestellt. Darüber hinaus wird die thematische Ausrichtung des Didaktik Labs erläutert und aufgezeigt, inwiefern interessierte Lehrende ihre Didaktik mit und über KI ausrichten und dabei in Kommunikation mit anderen Lehrenden treten können.

2.4 Bedeutung der Teilprojekte für andere Standorte und erste Ergebnisse

Die drei vorgestellten Teilprojekte tragen zu einer Stärkung der KI- und Datenkompetenzen durch innovative, digitale Lernangebote und deren Verfestigung in der Hochschule bei. Mit der anvisierten Verankerung der Lernangebote auf bildungswissenschaftlicher und interdisziplinärer Studiengangsbildungsebene der FeU werden Möglichkeiten für die Hochschulpraxis sowie für andere Hochschulen aufgezeigt. Darüber hinaus erfolgt ein Wissenstransfer in regionalen und überregionalen themenspezifischen Netzwerken. In den aufgebauten und bereits durchgeführten Communities of Practice kommen Hochschulangehörige,

Bildungsakteure, Unternehmen, regionale Akteure und alle Interessierten zusammen, um sich über KI in der Hochschule auszutauschen, voneinander zu lernen und Vorhaben umzusetzen. Es haben sich bereits fünf Open Think Tanks zu folgenden Themenschwerpunkten gebildet: *KI & Didaktik I und II*, *KI, Curriculum & Micro-Degrees*, *KI & Prüfung* sowie *KI und Ethik*. In den monatlich stattfindenden offenen Austauschformaten werden die (Praxis-)Erfahrungen der FeU eingebracht mit dem Ziel, den kreativen Austausch zu KI in der Hochschulbildung zu führen und das Thema praktisch voranzubringen.

3 Praxisbeitrag

Der Praxisbeitrag setzt damit zentrale Impulse aus der aktuellen Projektpraxis entlang folgender Fragestellungen: Inwiefern lassen sich nachhaltige Umsetzungskonzepte entwickeln, erproben und in mediendidaktische Formate mit offenen KI-C-Lernangeboten curricular an der FernUniversität integrieren? Wie lassen sich Micro-Degrees als neue Qualifizierungsformate anhand von KI- C-Lernangeboten in die Hochschulbildung der FernUniversität kriterienorientiert integrieren sowie anerkennen und anrechnen? Inwiefern kann der Wissenstransfer mit und über KI in der Hochschullehre nachhaltig gestaltet werden und welche Erkenntnisse sind bereits im Rahmen der Initiative der Open Think Tanks entstanden? Insgesamt ist der Beitrag darauf ausgerichtet, Erfahrungen und Einsichten in die aktuelle Projektpraxis zu gewähren. Dabei sind alle Teilnehmenden herzlich eingeladen, den Community-Aktivitäten beizutreten und Feedback zu den verschiedenen Schwerpunkten einzubringen.

Literaturverzeichnis

- [EU20] Europäische Kommission, Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen über die Vollendung des europäischen Bildungsraums bis 2025. Brüssel, 2020, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0625>, Stand: 30.08.2023.
- [Fl22] Flasdick, J.; Mah, D.-K.; Bernd, M.; Rampelt, F.: Micro-Credentials und Micro-Degrees. Aktuelle Entwicklungen und Perspektiven aus der Praxis des KI-Campus. Berlin: KI- Campus, 2022.
- [Ho20] Hochschulrektorenkonferenz (HRK), Micro-Degrees und Badges als Formate digitaler Zusatzqualifikation. Empfehlung der 29. HRK-Mitgliederversammlung vom 24.11.2020, <https://www.hrk.de/positionen/beschluss/detail/micro-degrees-und-badges-als-formate-digitaler-zusatzqualifikation/>, Stand: 30.08.2023.
- [If20] Ifenthaler, D.: Learning Analytics im Hochschulkontext – Potenziale aus Sicht von Stakeholdern, Datenschutz und Handlungsempfehlungen. In (Fürst, R. A., Hrsg.): Digitale Bildung und Künstliche Intelligenz in Deutschland. Nachhaltige

Wettbewerbsfähigkeit und Zukunftsagenda. Springer Fachmedien, Wiesbaden, S. 520-531, 2020.

Die Chatbot-Challenge – Spielend mit KI von der Idee zum Dialogsystem

Ingo Siegert¹⁷, Stefan Hillmann²⁷, Philline Thalia Görzig³⁷, Matthias Busch⁴⁷, Jan Nehring⁵⁷ und Xenia Klinge⁶⁷

Abstract: In der Chatbot-Challenge sollen Gruppen von vier bis fünf Studierenden ein spannendes Escape-Spiel erstellen, das durch einen Dialog mit einem Conversational Interface erlebt wird. Aufgabe ist es, den Spielverlauf zu planen, mittels des Rasa-Frameworks ansprechend umzusetzen und die Ergebnisse in einer Studie zu prüfen. Hierzu muss nicht nur das KI-gestützte Framework erlernt und angewandt werden, sodass der Chatbot Eingaben verstehen und passend beantworten kann, um das Spiel voranzubringen, sondern auch das Design des Spiels und der Interaktionen iterativ mithilfe von Probanden evaluiert werden. Zum Abschluss der Challenge wird aus allen erstellten Bots der Gewinner in einer Nutzungsstudie ermittelt. Die Chatbot-Challenge wird als hybrides, paralleles Master-Modul an der OvG-Universität Magdeburg und der TU Berlin durchgeführt.

Keywords: AI-Literacy, Chatbot, Escapespiel, Hochschule

1 Einleitung

Der zunehmende Einsatz von KI in Freizeit und Beruf erfordert eine stärkere Förderung im Umgang mit angewandter KI in der Bildung. Es ist wichtig, dass Studierende in der Lage sind, KI-Methoden und -Technologien effektiv zu nutzen und zu verbessern. Dazu gehört auch, KI-Systeme zu entwickeln und zu evaluieren, sowie das Verständnis von deren Grenzen und Möglichkeiten. Wichtig ist hierbei die Entwicklung von Lehrmaterialien, die sich nicht nur auf Grundlagen und technische Hintergründe konzentrieren, sondern insbesondere auch die sichere Verwendung von KI als Werkzeug zum Erreichen bestimmter Ziele vermitteln, sodass Studierende nicht allein die abstrakte Funktionsweise der Algorithmen verstehen, sondern auch, wie und wozu diese in der Praxis verwendbar sind.

Das hier vorgestellte Studienprojekt ist eine auf ein Semester ausgelegte Projektarbeit, in der die Anwendung von KI-Methoden praktisch geübt und in Kombination mit Fähigkeiten des Projektmanagements und dem wissenschaftlichen Evaluieren und Präsentieren gebracht

¹ Otto-von-Guericke-Universität Magdeburg, Fachgebiet Mobile Dialogsysteme, Universitätsplatz 2, 39106 Magdeburg, Deutschland siegert@ovgu.de

² TU Berlin, Q&U Lab, Ernst-Reuter-Platz 7, 10587 Berlin, Deutschland stefan.hillmann@tu-berlin.de

³ TU Berlin, Q&U Lab, Ernst-Reuter-Platz 7, 10587 Berlin, Deutschland p.goerzig@tu-berlin.de

⁴ Otto-von-Guericke-Universität Magdeburg, Fachgebiet Mobile Dialogsysteme, Universitätsplatz 2, 39106 Magdeburg, Deutschland matthias.busch@ovgu.de

⁵ DFKI, Speech and Language Technologies Lab, Alt-Moabit 91c, 10559 Berlin, Deutschland jan.nehring@dfki.de

⁶ DFKI, COS, Alt-Moabit 91c, 10559 Berlin, Deutschland xenia.klinge@dfki.de

⁷ Alle Autoren haben gleichermaßen beigetragen.

wird. Dies geschieht anhand einer klar umrissenen Zielsetzung: Ein Escape-Spiel mit einem Chat-Interface zu erzeugen, das sich in einer abschließenden Challenge gegen andere Spiele in Bezug auf Spielspaß und Qualität der technischen Realisierung durchsetzen kann. Die Challenge wird als hybrides, paralleles Lehr-Modul an der Otto-von-Guericke-Universität Magdeburg und der Technischen Universität Berlin angeboten.

2 Die Chatbot-Challenge

Die Aufgabe ist, mithilfe des KI-gestützten Chatbot-Frameworks Rasa ein Escape-Room-Spiel zu entwickeln. In einem Escape-Room müssen die Spielenden eine Reihe von Rätseln lösen, um sich aus einem Raum zu befreien. In der Chatbot-Challenge interagieren die Spielenden mit dem Spiel lediglich über Textnachrichten. Sie schreiben beispielsweise “Öffne die Tür”, worauf das Spiel antwortet “Die Tür ist verschlossen”. Das Setting eines Escape-Spiels ist allseits bekannt, bietet damit auch internationalen Studierenden einen Zugang, erlaubt auch kurze Spiele ohne limitiert zu wirken und erlaubt trotzdem viel Freiheit in der Entwicklung. Ergebnis des Moduls sollte ein bedienbarer Chatbot mit spielbarer und konsistenter Story sein. Im Fokus steht hier aber vor allem die ganzheitliche Erstellung eines KI-Projekts von der Idee über die Umsetzung bis zur Evaluation im “produktiven Einsatz”.

2.1 Inhalte

Die größte Herausforderung bei der Erstellung des Chatbots ist es, dass die Menge der möglichen Texteingaben nicht im Voraus bekannt ist. Es lassen sich nicht alle Intents (Handlungsabsichten⁸) und Entities (Elemente der “Spiel-”Welt⁹), die in den Eingaben vorkommen können, vorhersehen. Daher kommen KI-Methoden zum Einsatz, um die Interaktion mit dem Chatbot mittels NLU (Natural Language Understanding) zu verbessern. Als Plattform für den Chatbot wird Rasa¹⁰ genutzt. Dessen NLU-Komponente basiert auf einem zweistufigen Ansatz. Zunächst werden mit einem Transformer-basierten [Va17, De19] vortrainiertem Sprachmodell (bspw. [Fe22]) die Eingaben der Nutzenden in sog. Embeddings (eine numerische Repräsentation, die es z.B. erlaubt semantische Ähnlichkeit zu messen) überführt. Diese werden dann von einem zweiten Klassifikator verwendet, dem Dual Intent und Entity Classifier (DIET) [Bu20], der für die eigentliche Erkennung der Intents und der Entities trainiert bzw. verwendet wird.

Die Teilnehmenden lernen ein KI-Modell zu trainieren, zu evaluieren und direkt anzuwenden. Sie üben am praktischen Beispiel Escape-Room, wie Daten für das Training eines Klassifikators gesammelt und aufbereitet werden müssen (benötigte Datenmenge, Datenbalance in verschiedenen Klassen) und wie die trainierten Modelle hinsichtlich ihrer Klassifikationsleistung evaluiert werden können (geeignete Maße, Kreuzvalidierung).

Neben der reinen Erkennungsrate sind in der Challenge auch User Experience (UX) und Spielspaß (Game Experience) wichtige Metriken der Evaluation des Chatbots. Nicht nur die Erzeugung des KI-Modells wird erlernt, das den Chatbot befähigt, Eingaben in natürlicher

⁸Ein Intent beschreibt das Ziel einer Interaktion, z.B. “Tür öffnen”, “Licht anschalten”.

⁹Eine Entity bezieht sich auf Personen, Orte oder Objekte, die der Chatbot “kennen” soll, z.B. Tür, Schlüssel.

¹⁰Ein kostenloses Open-Source-Chatbot-Framework basierend auf maschinellem Lernen.

Sprache semantisch korrekt zu interpretieren, sondern auch, wie die UX eines interaktives Dialogsystem evaluiert werden kann. In dem Studienprojekt wird also die äußerst relevante Verbindung zwischen dem Erzeugen eines KI-gestützten Systems und dessen Nutzung durch und Wirkung auf die Nutzenden hergestellt.

2.2 Ablauf und Regeln

Die Veranstaltung ist in mehrere Praxisblöcke, Ergebnis-Vorstellungen sowie Selbstlernphasen aufgeteilt. Die Praxisblöcke umfassen die Vorstellung der Chatbot-Challenge, der Lehrenden und der Regularien. Weiterhin gibt es Blöcke zur Einführung Rasa, zum Trainieren der NLU, zum Planen, Durchführen und Auswerten von Nutzungstudien und zum Deployment im Frontend. Die Ergebnisvorstellungen finden immer als parallele hybride Veranstaltung an der Uni Magdeburg und der TU Berlin statt. Die Termine und erwarteten Ergebnisse sind als entsprechende Meilensteine (MS) vordefiniert und umfassen: (MS1) Vorstellung Gesamtidee, 1. Prototyp, (MS2) 1. Version vollständiger Escape-Room, Konzept Nutzerstudie, (MS3) Escape-Room nach Nutzerstudie, Auswertung Nutzerstudie, (MS4) Auswertung finale Nutzerstudie, Lessons-Learnt. Am Ende schreiben alle Gruppen einen Bericht, der die Ergebnisse der Projektarbeit zusammenfasst.

Damit die Ausgangsbedingungen zwischen den Teams vergleichbar sind und die Aufgabe insgesamt nicht zu groß wird, gibt es klare Regeln, die am Beginn der Veranstaltung kommuniziert werden. Hierzu gehört, dass alle Gruppen Rasa verwenden müssen, dass in den Antworten keine Bilder erlaubt sind und keine externen LLMs (Large Language Models) benutzt werden dürfen. Weiterhin ist der Fragebogen für die UX-Evaluation festgelegt (UEQ+[Sc21]), für die Evaluierung des Spielspaßes sollen die Teams Vorschläge geben.

3 Lernziele

Künstliche Intelligenz	Benutzung des Rasa Frameworks Generierung von Trainingsdaten für maschinelles Lernen Evaluation von KI-basierten Klassifikatoren
UX-Design	Design eines textbasierten Spiels Iterative Durchführung von Nutzungstudien
Projektmanagement	Selbstorganisation in Gruppen, Management Tools Selbstständige Erarbeitung von Teilzielen
Evaluierung der Interaktion	User Experience und Game Experience Fragebögen

Tab. 1: Lernziele der Lehrveranstaltung

In der Challenge haben die Studierenden Gelegenheit, sich mit dem Thema KI in einem anwendungsorientierten Rahmen auseinanderzusetzen, der von ihnen einerseits ein Verständnis der zugrunde liegenden Prinzipien verlangt (z.B. beim Trainieren eines NLU-Modells), andererseits aber die Aufmerksamkeit direkt auf das Ergebnis lenkt, das sich in Gesprächen mit relativ schnell iterierbaren Versionen des Chatbots einfach überprüfen lässt. Somit liegt der Fokus auf der bewussten und souveränen Anwendung des KI-gestützten Dialogsystem-frameworks, wie sie in der heutigen Softwarelandschaft mehr und mehr erforderlich wird.

Darüber hinaus werden im Gesamtprojektrahmen weitere wichtige Fähigkeiten und Kenntnisse in UX-Design, Projektmanagement und Evaluierung vermittelt (Tab. 1). Durch die Wettbewerbsrahmung setzten sich die Teilnehmenden mit den anderen Lösungen auseinander und identifizierten die Vor- und Nachteile, um diese im eigenen Chatbot zu berücksichtigen.

4 Ausblick

Die Challenge hat im April 2023 als universitätsübergreifende Lehrveranstaltung begonnen. Dies ist das erste Mal, dass sie durchgeführt wird, in kommenden Jahren soll sie jedoch in dieser oder angepasster Form weitergeführt werden.

Während die Studierenden in ihrer Gruppe jeweils einen Chatbot entwickeln, wurde durch uns ein Portal entwickelt, in das die Chatbots integriert werden können. Dies erlaubt es, dass alle Teilnehmenden alle Chatbots unter gleichen Bedingungen testen können – und die interessierte Öffentlichkeit sowie Teilnehmende auch zukünftig die Spiele spielen können. Die Challenge ist in vielen Punkten eingeschränkt, damit sich die Studierenden auf die zu bewertenden Aspekte konzentrieren und die abschließende Bewertung der Ergebnisse fair bleibt, da alle dieselben Mittel nutzen. Dennoch bieten sich Möglichkeiten, die Materialien in zukünftigen Challenges anzupassen, insbesondere im Bereich von Textgenerierung. So hat sich bereits in diesem Jahr in den Fragen Studierender ein großes Interesse am Einsatz von ChatGPT gezeigt und auch für Rasa gibt es erste Arbeiten LLMs zu integrieren¹¹.

Die bisherigen Erfahrungen mit der diesjährigen Challenge zeigen, dass alle Gruppen in der Lage waren, mit Rasa erste Prototypen zu entwerfen, und NLU-Modelle zu trainieren. Unterschiede zeigen sich, wie die Gruppen den Chatbot im Spiel interpretieren (Erzähler, Mit-/Gegenspieler). Auch zeigen sich hier aber bereits erste Herausforderungen bei der Entwicklung. Der Ansatz von Versuch und Irrtum, der sonst für kleine Demonstratoren ein gangbarer Weg ist, stößt hier schnell an zeitliche Grenzen.

Literaturverzeichnis

- [Bu20] Bunk, T.; Varshneya, D.; Vlasov, V.; Nichol, A.: DIET: Lightweight Language Understanding for Dialogue Systems, 2020.
- [De19] Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proc. 2019 Conf. of the North American Chap. ACL: Human Lang. Tech. ACL, Minneapolis, Minnesota, S. 4171–4186, Juni 2019.
- [Fe22] Feng, F.; Yang, Y.; Cer, D.; Arivazhagan, N; Wang, W.: Language-agnostic BERT Sentence Embedding. In: Proc. of the 60th Annual Meeting of ACL. Dublin, Ireland, S. 878–91, 2022.
- [Sc21] Schrepp, M.: Measuring User Experience with Modular Questionnaires. In: 2021 Int. Conf. on Advanced Computer Science and Information Systems (ICACSIS). 2021.
- [Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I.: Attention is All you Need. In: Advances in Neural Information Processing Systems. Jgg. 30. Curran Associates, Inc., 2017.

¹¹etwa in Projekten wie RasaGPT (<https://github.com/paulpierre/RasaGPT>).

AI Engineering als interdisziplinäres Einführungsmodul zwischen Künstlicher Intelligenz und Ingenieurwesen

Sebastian Lang¹, Ingo Siegert², Viktor Artiushenko³ und Johannes Schleiss⁴

Abstract: Im Wintersemester 2023/2024 wird in Sachsen-Anhalt der Kooperationsstudiengang »AI Engineering — Künstliche Intelligenz in den Ingenieurwissenschaften« eingeführt. Ziel des Studiengangs ist es, Fachkräfte in der Entwicklung und Implementierung industrieller KI-Lösungen auszubilden. Im ersten Fachsemester absolvieren die Studierenden das Modul »Einführung ins AI Engineering«, das einen übergeordneten Rahmen für die im weiteren Studienverlauf vermittelten KI- und Ingenieurmethoden schafft. Dieser Beitrag diskutiert die Herausforderungen in der Entwicklung eines solchen interdisziplinären Einführungsmoduls. Anschließend wird ein Konzept zur Ausgestaltung und Didaktik des Einführungsmoduls präsentiert, das diesen Herausforderungen begegnet.

Keywords: AI Engineering, Interdisziplinäre Hochschullehre, Kursdesign, Studieneingangsphase

1 Hintergrund

Der Begriff »AI Engineering« beschreibt die systematische Konzeption, Entwicklung, Integration und den Betrieb von KI-basierten Lösungen für industrielle Problemstellungen nach Vorbild ingenieurwissenschaftlicher Methoden. Gleichzeitig schlägt »AI Engineering« eine Brücke zwischen der KI-Grundlagenforschung und den anwendungsorientierten Ingenieurwissenschaften, indem es den Einsatz von KI-Methoden und -Technologien für ingenieurwissenschaftliche Problemstellungen als Forschungsschwerpunkt setzt. Die grundlegende Hypothese des gleichnamigen Studiengangs ist, dass die Entwicklung von KI-Lösungen für industrielle Problemstellungen zwei wesentliche Voraussetzungen bedingen, nämlich (1) das Vorhandensein eines interdisziplinären Wissens über KI- und ingenieurwissenschaftliche Methoden sowie (2) die Fähigkeit zur integrativen Anwendung beider Methodenwelten [Sc22].

Hinsichtlich (1) umfasst das 7-semesterige Curriculum Veranstaltungen, die vornehmlich KI- oder ingenieurwissenschaftliche Grundlagen lehren. Im 5. und 6. Fachsemester

¹ Fraunhofer-Institut für Fabrikbetrieb und -automatisierung IFF, Geschäftsfeld Robotersysteme, Sandtorstr. 22, 39106 Magdeburg, Deutschland, sebastian.lang@iff.fraunhofer.de

² Otto-von-Guericke-Universität Magdeburg, Fakultät für Elektro- und Informationstechnik, Universitätsplatz 2, 39106 Magdeburg, Deutschland, siegert@ovgu.de

³ Otto-von-Guericke-Universität Magdeburg, Fakultät für Maschinenbau, Universitätsplatz 2, 39106 Magdeburg, Deutschland, viktor.artiushenko@ovgu.de

⁴ Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, Universitätsplatz 2, 39106 Magdeburg, Deutschland, johannes.schleiss@ovgu.de

werden zudem vertiefende Fachkenntnisse in einer aus fünf wählbaren ingenieurwissenschaftlichen Vertiefungen vermittelt. Hingegen stellt (2) eine Herausforderung dar, weil bisher kein übergeordneter methodischer, beide Disziplinen vereinender Rahmen existiert. Infolgedessen mangelt es gegenwärtig an etablierten Lehrkonzepten für eine integrierende Anwendung von KI- und ingenieurwissenschaftlichen Methoden.

2 Intention und Herausforderungen eines Einführungsmoduls

Zu diesem Zweck wird im ersten Semester des Studiengangs das Modul »Einführung ins AI Engineering« angeboten. Dieses dient als zentrale Einführungsveranstaltung mit je 2 SWS für Vorlesung und Übung. In dieser Funktion soll das Modul über den Aufbau und das Konzept des Studiengangs informieren, bei dem ein über alle Semester integrierender Ansatz zur Ausbildung von KI- und Ingenieurwissenschaften angestrebt wird. In der Einführungsveranstaltung wird zu diesem Zweck der notwendige methodische Rahmen vermittelt. Dieser dient als Grundlage zur integrativen Anwendung von KI- und ingenieurwissenschaftlichen Methoden, um KI-Lösungen für industrielle Problemstellungen zu entwickeln. Dreh- und Angelpunkt des methodischen Rahmens ist ein Vorgehensmodell, welches den vollständigen Entwicklungsprozess von industriellen KI-Lösungen - von der Problemanalyse und Zieldefinition bis hin zur Inbetriebnahme, Instandhaltung und Wartung der KI-basierten Anwendung - beschreibt und in allen Phasen methodisch unterstützt.

Hierbei stellen sich zunächst die für Einführungsveranstaltungen üblichen Herausforderungen [BT14]. Beispielsweise muss auf die unterschiedlichen Vorkenntnisse der Studierenden in der Studieneingangsphase eingegangen werden, insbesondere da der Studiengang als Kooperationsstudiengang zwischen Universität und Hochschulen für angewandte Wissenschaften eine größere Variabilität der Studierendenschaft erwarten lässt. Des Weiteren muss berücksichtigt werden, dass die Einführungsveranstaltung der erste prägende Kontakt der Studierenden mit den Hochschulen im Allgemeinen und dem Studiengang im Speziellen ist. Darüber hinaus entstehen jedoch auch neue Probleme, die aus der Interdisziplinarität zwischen KI- und Ingenieurwissenschaften resultieren. Etablierte ingenieurwissenschaftliche Bachelorstudiengänge vermitteln in den ersten Semestern zunächst notwendige mathematische und technische Grundlagen sowie erste Kenntnisse in der Programmierung von Software. Hingegen fokussieren informatikorientierte Bachelorstudiengänge meist die Ausbildung von Mathematik- und Informatikgrundlagen, ohne die Theorie von ingenieurwissenschaftlichen Anwendungsdomänen zu beleuchten. Die integrierte Anwendung von KI- und Ingenieurmethoden wird oftmals erst in Masterstudiengängen in Form von praxisnahen Projekten vermittelt. Die Herausforderung besteht somit darin, die Ausbildungsinhalte der Einführungsveranstaltung entsprechend aufzubereiten und zu vermitteln, sodass diese von den Studierenden aufgenommen und verstanden werden, ohne dass diese zuvor eine Grundlagenausbildung in den KI- und Ingenieurwissenschaften erfahren haben.

3 Lernziele und Ausgestaltung der Einführungsveranstaltung

Aus diesen Überlegungen lässt sich das folgende Hauptlernziel formulieren, nämlich dass Studierende nach erfolgreichem Abschluss des Einführungsmoduls in der Lage sein sollen, Forschungs- und Entwicklungsprozesse für industrielle KI-Lösungen zu verstehen, zu planen und koordinativ zu begleiten. Das Hauptlernziel kann wiederum auf die im Folgenden gelisteten Teillernziele heruntergebrochen werden.

1. *Definition, Konzepte und Methoden der KI:* Zunächst muss den Studierenden ein grundlegendes KI-Verständnis vermittelt werden. Neben einer griffigen Definition gilt es herauszustellen, welche Methoden unter dem KI-Begriff verstanden werden und wie diese sich konzeptionell unterscheiden.
2. *Definition und Merkmale von AI Engineering:* Aufbauend auf Teillernziel (1) werden die Studierenden an den Begriff und die Wesenszüge des AI Engineering herangeführt, u. a. durch Formulierung einer prägnanten Definition und anhand von Praxisbeispielen.
3. *Charakteristiken und Differenzierung von KI- und ingenieurtechnischen Entwicklungsprozessen:* Um das Verständnis für das Wesen des AI Engineering zu vertiefen, werden vorbereitend und zur besseren Einordnung typische KI- und ingenieurtechnische Entwicklungsprozesse analysiert und beschrieben.
4. *AI-Engineering-Prozess:* In Abgrenzung zu den konventionellen KI- und ingenieurtechnischen Entwicklungsprozessen (Teillernziel 3) wird der AI-Engineering-Prozess hergeleitet. Hierbei handelt es sich um ein Vorgehensmodell zur Projektierung und Umsetzung industrieller KI-Systeme.
5. *Begriffe, Methoden, Modelle und Gütemaße für AI-Engineering-Prozessphasen:* Die Studierenden erlangen ein tiefgreifendes Verständnis über den AI-Engineering-Prozess und dessen Anwendung. Für jede Phase des Prozesses werden den Studierenden notwendige Begriffsdefinitionen, Methoden und Modelle zur erfolgreichen Absolvierung der Prozessphase sowie Gütemaße zur Evaluation der Phasenergebnisse an die Hand gegeben.

Die Nummerierung der Teillernziele impliziert die Reihenfolge der zu vermittelnden Inhalte und gibt somit Rückschluss auf die Ausgestaltung des Einführungsmoduls. Hierbei wird ein Großteil des Moduls anhand des eingeführten AI-Engineering-Prozesses strukturiert. Zunächst ist jedoch in den ersten Wochen des Moduls eine allgemeine Einführung in KI geplant, um notwendige Begriffe und Definitionen zum Verständnis der nachfolgenden Kerninhalte zu legen. Gleichmaßen soll in den ersten Wochen ein Verständnis für prozessorientiertes Denken und Handeln vermittelt und anhand von unterschiedlichen KI- und ingenieurtechnischen Entwicklungsprozessen demonstriert werden. Darauf folgend wird der AI-Engineering-Prozess als Kerninhalt des Moduls eingeführt. In den folgenden Wochen orientiert sich der Lehrplan an den einzelnen Phasen des AI-Engineering-Prozesses.

4 Didaktisches Konzept

Das didaktische Konzept soll sich an die Struktur und Inhalte des angestrebten Lehrplans anpassen. Dafür verknüpft das Modul verschiedene Lehrformen. Notwendige KI-Grundlagen sollen über ein integriertes Angebot von Vorlesungen und Übungen gelehrt werden, d. h. die Wissensvermittlung erfolgt über Vorlesungen, wobei das Wissen der Studierenden innerhalb der Vorlesung regelmäßig und interaktiv abgefragt wird. Der AI-Engineering-Prozess mit seinen einzelnen Phasen soll den Studierenden durch eine Kombination aus klassischer Vorlesung, Fallstudien und umgedrehter Unterricht (Flipped Classroom) nähergebracht werden. Konkret hält der Dozierende zu jeder Prozessphase einen einführenden Impulsvortrag, um wesentlichen Inhalte zusammenzufassen. Anschließend diskutieren die Studierenden gemeinsam mit dem Dozierenden die Herausforderungen der jeweiligen Prozessphase. Methoden, Modelle und Gütemaße, die zur Bearbeitung der Prozessphase notwendig sind, werden von den Studierenden für die kommende Woche recherchiert, in Form von kurzen Vorträgen präsentiert und im Anschluss diskutiert, bevor zur nächsten Prozessphase übergeleitet wird. Parallel zur Behandlung der Prozessphasen sollen die Studierenden in Gruppenarbeit einen AI-Engineering-Prozess für eine jeweils individuell vorgegebene Problemstellung entwickeln. Hierbei dient das Ergebnis der Gruppenarbeit als Prüfungsleistung des Moduls.

5 Diskussion und Ausblick


Der vorliegende Beitrag diskutiert die Herausforderungen eines Einführungsmoduls zum Thema »AI Engineering« und präsentiert ein erstes Konzept zur Ausgestaltung und Didaktik des Einführungsmoduls, um diesen zu begegnen. Für die konkrete Ausgestaltung des Einführungsmoduls verbleibt weiterhin als Herausforderung, einen AI-Engineering-Prozess zu definieren, der angrenzenden Prozessmodellen nicht widerspricht und darüber hinaus eine breite Anerkennung in Wissenschaft und Praxis erzielt. Das Modul wird erstmalig im Wintersemester 2023/2024 angeboten. Es ist festzustellen, dass die konkrete Umsetzung des Konzepts eine ambitionierte Aufgabe bleibt, und daher nach erster Absolvierung des Einführungsmoduls eine Evaluation erfolgen muss. Dabei gilt es insbesondere zu eruieren, ob die ex ante angedachten KI- und ingenieurstechnischen Grundlagen geeignet und in sich vollständig sind, um eine verständliche und umfängliche Einführung in beide Themengebiete zu leisten.

Literaturverzeichnis

- [BT14] Bosse, E.; Trautwein, C.: Individuelle und institutionelle Herausforderungen der Studieneingangsphase. *Zeitschrift für Hochschulentwicklung*, 9(5), S. 41-62, 2014.
- [Sc22] Schleiss, J.; Bieber, M. I.; Manukjan, A.; Kellner, L.; Stober, S.: An Interdisciplinary Competence Profile for AI in Engineering. In: *Proceedings of 2022 Annual Conference of the European Society for Engineering Education (SEFI)*, S. 1601–1608, 2022.

Harnessing Conversational AI in Higher Education: The Education Buddy - An Innovative Intervention for International Students Facing Study Difficulties

AI-Assisted Learning for International Students


Martin Hänel ¹ and Matthias Söllner²

Abstract: As Germany anticipates a demographic shift necessitating reliance on foreign professionals, the integration of international students becomes critical. Despite being a preferred non-English speaking study destination, the country's academic environment – encompassing course materials, lectures, and organizational aspects – is predominantly German. This poses a significant language barrier for the 440,000 foreign students currently studying in Germany, potentially impeding their academic success and future professional integration. To mitigate this, we introduce the Education Buddy, a Conversational Agent (CA) designed to assist international students in navigating their studies. The Education Buddy emerges as a potentially effective tool for facilitating the successful participation of international students within the German academic system.

Keywords: Education Buddy, Conversational Agent, Digital Learning, Coach, Integration

1 Introduction

Undertaking higher education in a foreign country can be a daunting challenge, especially for international students. In Germany, there are currently 440,000 international students [ER22], and at the same time, there is a shortage of skilled workers [HA23]. In response to the shortage of skilled labor, the German Council of Economic Experts proposes that an annual influx of 1.5 million immigrants is necessary [TA23]. The country actively seeks to attract international students and professionals to contribute to its workforce and combat the skills gap [HA23]. By encouraging the integration of these students into the German job market after graduation, Germany aims to harness their knowledge and skills to bolster its industries and economic growth. However, international students, who are often not native German speakers, face several challenges in their studies [Pi18]. The German academic exchange service clusters the study difficulties of international students into nine parts: Discrepancy between

¹ University of Kassel, Chair of Information Systems and Systems Engineering, Pfannkuchstraße 1, Kassel, 34121, haenel@uni-kassel.de,  <https://orcid.org/0009-0009-5035-0782>

² University of Kassel, Chair of Information Systems and Systems Engineering, Pfannkuchstraße 1, Kassel, 34121, soellner@uni-kassel.de.

expectations and reality, institutional heterogeneity, language problems, lack of preparation for studies and daily life, study financing, career prospects for students, administrative processes like residence permit, stress, and other emotional strains.

The Education Buddy primarily focuses on the scope of resolving the problem of students' inadequate preparation for their academic studies. The current concept for providing educational coaching to students with study difficulties in Germany relies solely on lecturers. However, this approach is not effective, particularly in large departments where there are capacity issues. This shortage of coaching can lead to lower success rates and dropouts which is a permanent issue for the society [GR2023]. Uniquely, the Education Buddy marries the human coaching approach with an Conversational Agent (CA) assistance, such a hybrid coaching approach has not been previously realized in higher education and is a scalable education technology solution. The novel approach of this idea is the hybrid collaborative coaching, in which support is provided by either a human or a CA, depending on the issue of the international student. A continuous handover takes place between the two entities, depending on which type of coaching is currently beneficial.

Our research proposes the creation and implementation of a hybrid CA system. The core objective of the project is to enrich educational coaching for international students by integrating digital elements. It is essential to identify which aspects of educational coaching can be effectively supplemented by digital offerings. Furthermore, the role of the Education Buddy in relation to lecturers needs to be defined.

For this purpose, we pose the research question:

In the realm of higher education for international students facing study difficulties because of inadequate preparation and insufficient prior knowledge, what design principles should be employed to optimize the support provided by conversational agents, in which specific tasks can they offer assistance to support traditional lecturers and how the CA will be integrated into the existing support services?

2 Theoretical Background

The latest developments in AI have opened new interactive opportunities for the support and interaction with the user [Wi22]. The advanced capabilities of AI have been instrumental in extending user support across a progressively broadening range of applications [Wi22]. Until nowadays the capabilities of such a CA as coach were limited, however due to the newest developments CA coaches can match human coaches in adult education [Te22]. CA are AI-powered software programs designed to simulate human-like conversations and interactions [GCM13]. These digital assistants provide personalized advice, support, and guidance to users in the education context [Wi20]. Their purpose extends beyond answering queries to facilitating a learning process, tailoring their interactions based on users' individual needs, preferences, and progress

over time [Wi20].

The Education Buddy is a CA designed to provide personalized assistance to international students with a lack of preparation for studies. The CA serves as the first point of contact allowing lecturers to focus on complex knowledge issues necessitating their attention. The technical design of the Education Buddy is anchored in the use of Large Language Models, particularly Natural Language Processing (NLP) techniques and transformer-based architectures like GPT.

Interaction with stakeholders is crucial in the successful implementation. The collaboration begins by engaging these stakeholders in discussions to understand their perspective and insights, thus ensuring the system is designed to meet the needs of all parties involved. The International Office would play a vital role in the integration and promotion of the Education Buddy. Student advisors would be integral in providing content and context for the Education Buddy. Exam offices would be crucial partners in supplying the CA with up-to-date information which students may struggle and need support.

3 Implications and Conclusions

In our quest for designing the Education Buddy, we endeavour to employ the Design Science Research approach [He04]. Our primary objective is to glean practical implications related to the requirements of international students and the utilization of the Buddy at German universities. In terms of practice, the Education Buddy could change the way support is provided to international students. It could significantly improve student outcomes and wellbeing, contributing to a more inclusive and supportive educational environment. On a theoretical level, our research aims to delve deeper into how conversational agents can contribute to the elevation of accessibility and inclusivity of the processes and procedures in the higher education domain. The project could contribute to our understanding of AI in education, particularly in the context of international student support. In conclusion, the development and research of the Education Buddy represents an exciting, promising venture in the field of educational technology.

Bibliography

- [ER22] ER, Erudera, <https://erudera.com/statistics/germany/germany-international-student-statistics>, accessed: 10/05/2023.
- [GCM13] Griol, D.; Carbo, J.; Molina, J.: An automatic dialog simulation technique to develop and evaluate interactive conversational agents. *Applied Artificial Intelligence* 10/13, pp. 759-780, 2013.
- [GR23] Göldi, A.; Rietsche, R.: Where to for automated coaching conversation: structured intervention or adaptive generation?. In: *Proceedings Research-In-Progress Papers of the European Conference on Information Systems. ECIS 2023*, Kristiansand, pp. 1-13, 2023.
- [HA23] HA, Handelsblatt, <https://www.handelsblatt.com/politik/deutschland/arbeitsmarkt-ingenieurmangel-spitzt-sich-zu-mehr-als-170-000-stellen-offen/29140000.html>, accessed: 27/05/2023.
- [He04] Hevner, A.; March, S.; Park, J.; Ram, S.: Design Science in Information Systems Research. *MIS Quarterly* 03/04, pp. 75-106, 2004.
- [Pi18] Pineda, J.: *Problemlagen und Herausforderungen internationaler Studierender in Deutschland*. DAAD Deutscher Akademischer Austauschdienst, Bonn, 2018.
- [TA23] TA, Tagesschau, <https://www.tagesschau.de/wirtschaft/wirtschaftsweise-schnitzer-zuwanderung-fachkraefte-100.html>, accessed: 05/07/2023.
- [Te22] Terblanche, N.; Molyn, J.; de Haan, E.; Nilsson, V.: Comparing artificial intelligence and human coaching goal attainment efficacy. *Plos one* 06/22, pp. 1-17, 2022.
- [Wi20] Winkler, R.; Hobert, S.; Salovaara, A.; Söllner, M.; Leimeister, Jan Marco.: Sara, the Lecturer: Improving Learning in Online Education with a Scaffolding-Based Conversational Agent. In: *Proceedings of the Conference on Human Factors in Computing Systems. CHI 2020*, Honolulu, pp. 1-14, 2020.
- [Wi22] Wintersberger, P.; van Berkel, N.; Fereydooni, N.; Tag, B.; Glassman, E.: Designing for Continuous Interaction with Artificial Intelligence Systems. In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. CHI '22*, New Orleans, pp. 1-4, 2022.

Entscheidungsbäume erstellen ab Klasse 6

Susanne Podworny ¹, Yannik Fleischer² und Rolf Biehler ³

Abstract: Entscheidungsbäume können eine Methode sein, die den steigenden Bedarf an schulischen Lerninhalten zu KI bedienen. Im ProDaBi-Projekt wird Unterrichtsmaterial entwickelt, erprobt und beforscht, um datenbasiert Entscheidungsbäume mit Datenkarten ab Klasse 6 zu lehren. Evaluationen zeigen positive Ergebnisse zum Verstehen und Anwenden von Entscheidungsbäumen von Lernenden. Dies kann einen ersten Schritt für KI-Verständnis im Schulcurriculum legen.


Keywords: machine learning, Unterricht, Unplugged Aktivitäten, KI in der Schule

1 Einleitung und Hintergrund


Automatisierte Entscheidungsmodelle finden in vielen Bereichen wie Medizin, Kriminalistik oder Empfehlungsdiensten Anwendung. Aufgrund der Relevanz dieser Modelle besteht eine steigende Nachfrage nach schulischer Ausbildung in Datenwissenschaft und KI [Bi07]. Für den Unterricht geeignetes ist das Entscheidungsbaummodell, das Entscheidungsregeln hierarchisch darstellt, und das Grundkonzept von maschinellem Lernen als Generalisieren über Daten verdeutlicht [Be10]. Entscheidungsbäume eignen sich für den Einstieg in das Thema KI für Lernende, weil sie als Modelle transparent sind und kaum mathematisches oder informatisches Hintergrundwissen benötigt wird [ML19].

2 Entscheidungsbäume unterrichten

Im Projekt Data Science und Big Data in der Schule⁴ werden innovative Unterrichtseinheiten zur Einführung von datenbasierten Entscheidungsbäumen entwickelt, um das Verständnis für ML-Methoden ab Klasse 6 zu fördern⁵. Diese Reihe zielt auf die Beantwortung der Frage „Wie kann KI dabei helfen, datenbasiert zu entscheiden, ob ein Lebensmittel eher empfehlenswert oder eher nicht ist?“ Dazu gibt es weitere „plugged“ Unterrichtsreihen ab Klasse 9, um Entscheidungsbäume vertieft zu behandeln [BF21]. Dort werden weiterführende Themen wie z.B. Überanpassung

¹ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, podworny@math.upb.de, 
<https://orcid.org/0000-0002-6313-5987>

² Universität Paderborn, Warburger Str. 100, 33098 Paderborn, yanflei@math.upb.de

³ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, biehler@math.upb.de, 
<https://orcid.org/0000-0002-9815-1282>

⁴ ProDaBi; <http://www.prodabi.de>, 31.08.2023.

⁵ Apfel oder Popcorn?, <https://go.upb.de/Entscheidungsbaeume-KI5u6>, 31.08.2023.

In den Unterrichtsreihen erkunden Schülerinnen und Schüler, wie Modelle auf der Grundlage von Daten trainiert werden. Dazu erstellen sie manuell Entscheidungs­bäume und verstehen, dass ein Entscheidungsbaum ein Regelsystem ist, das aus Daten abgeleitet werden kann. Mit dem innovativen Ansatz der Verwendung von in der Statistikdidaktik bekannten Datenkarten (HK06) zur Repräsentation von Objekten können Lernende KI-Konzepte erforschen, indem sie mit den Datenkarten handeln und ihre eigenen Repräsentationen erstellen. Mit dem Übergang zu automatisch erzeugten Entscheidungs­bäumen lernen sie im Sinne eines Spiralcurriculums eine Methode der künstlichen Intelligenz kennen, sowie die Rolle von Daten und Menschen in diesem Prozess und welche Fehler auftreten können.

3 Diskussion

Die Unterrichtsreihe wurde in fünf Klassen der Jahrgangsstufe 6 erprobt. Erste Ergebnisse der Begleitforschung zeigen, dass das Thema Entscheidungs­bäume großes Interesse hervorruft. Von den Lernenden konnten 95% nach der Unterrichtsreihe einen Entscheidungsbaum anwenden, um die Klassifizierung eines neuen Lebensmittels korrekt vorherzusagen und 66% lieferten eine ideale Verbalisierung dazu [PFH22]. Mit dem Ansatz von Datenkarten kann eine erste Grundlage im Sinne des Spiralcurriculums gelegt werden, um Schülerinnen und Schüler in KI-Konzepte einzuführen, welche später mit anderen Themen, z.B. dem Programmieren von Bäumen oder zur Statistik, verbunden wird.

Literaturverzeichnis

- [Be10] Berthold, M. R., Borgelt, C., Höppner, F., Klawonn, F.: Guide to intelligent data analysis. Springer, 2010.
- [BF21] Biehler, R., Fleischer, Y.: Introducing students to machine learning with decision trees using CODAP and Jupyter Notebooks. *Teaching Statistics* 43/S1, S. 133-142.
- [Bi10] R. Biehler et al., Paderborn Symposium on Data Science Education at School Level 2017: The Collected Extended Abstracts. Universitätsbibliothek Paderborn, 2018.
- [HK06] Harradine, A., Konold, C.: How representational medium affects the data displays students make. In (Rossman, A., Chance, B. Hrsg.): *Working cooperatively in statistics education. Proceedings of the Seventh International Conference on Teaching Statistics*, 2006.
- [ML19] Martignon, L., Laskey, K.: Statistical literacy for classification under risk: an educational perspective. *AStA Wirtschafts- und Sozialstatistisches Archiv* 13/3, S. 269-278, 2019.
- [PFH22] Podworny, S., Fleischer, Y., Hüsing, S.: Grade 6 students' perception and use of data-based decision trees. In (Zapata-Cardona, L., Bonafini, F., Fan, A. Hrsg.): *Bridging the Gap: Empowering and Educating Today's Learners in Statistics. Proceedings of the Eleventh International Conference on Teaching Statistics*, 2022.

⁶ Entscheidungs­bäume Do It Yourself, <https://learn.ki-campus.org/courses/baecumedi-y-upb2021/>, 31.08.2023.

D^{AI} dalos: Forschen und Lernen zugleich?

Data Literacy als Lernaufgabe für die Klassisch-philologische Forschung


Andrea Beyer ¹, Konstantin Schulz ²


Abstract: Die Daidalos-Infrastruktur soll es Forschenden der Klassischen Philologie und verwandter Disziplinen ermöglichen, verschiedene Methoden des Natural Language Processing an selbst zusammengestellten Forschungskorpora anzuwenden. Dabei ist Daidalos als interaktive Lern- und Forschungsinfrastruktur konzipiert, die den Ausbau wesentlicher Teilfähigkeiten von Data Literacy, z. B. die Zusammenstellung und Analyse von Korpora oder den Umgang mit Annotationen, TEI-XML und graphischen Auswertungen, unterstützt.

Keywords: Data Literacy, Forschendes Lernen, KI-Bildung, NLP-Forschungsinfrastruktur

1 Data Literacy in der Klassischen Philologie

In der Klassischen Philologie besteht zunehmend das Bedürfnis, Methoden aus den Digital Humanities (DH) und insbesondere aus dem Bereich Natural Language Processing (NLP) in die Forschung zu integrieren. Im Gegensatz zur deutschsprachigen Fachcommunity sind diese auf ein anderes Verständnis von Text zurückgreifenden Methoden international bereits deutlich stärker etabliert, wenn u. a. literaturwissenschaftliche (z. B. zu Autorschaft [OE19]) oder philologische Fragen (z. B. zu Textrekonstruktionen [As22]) mit DH-Methoden untersucht werden. Als wir zur Vorbereitung unseres Daidalos-Projektes stichprobenartig deutschsprachige Klassische Philologen und Philologinnen nach den Ursachen für ihr vermeintliches Desinteresse an den DH befragten, wurde durchgehend auf fehlende eigene Kompetenzen und den Mangel an informatisch gebildeten Personen hingewiesen. Eine genauere Analyse des Umfeldes zeigte uns, dass es kaum Tools und Lernmaterialien gibt, die sich sowohl an authentischen Forschungsfragen und dem Gesamtworkflow der Klassischen Philologie orientieren als auch didaktisch für den Erwerb von Data Literacy [Sc19] aufbereitet sind. Ebenso fehlt ein gemeinsames Verständnis davon, was eine fachspezifische Data Literacy charakterisiert und wie ihre Teilfähigkeiten ausgebaut werden können. Daher entschieden wir uns, die geplante Forschungsinfrastruktur Daidalos von Beginn an auch als eine an authentischen *user stories* ausgerichtete Lerninfrastruktur zu konzipieren (Vernetzung mit Text+ und KI-Campus), um zusammen mit einer Community of Practice interessierte Forschende als

¹ Humboldt-Universität zu Berlin, Institut für Klassische Philologie, Unter den Linden 6, 10099 Berlin, beyeranz@hu-berlin.de,  <https://orcid.org/0000-0002-8468-6411>

² Deutsches Forschungszentrum für Künstliche Intelligenz, Alt-Moabit 91c, 10559 Berlin, Konstantin.Schulz@dfki.de,  <https://orcid.org/0000-0002-3261-9735>

Lernende möglichst niedrigschwellig an NLP-Methoden heranzuführen und mit ihnen über den Einsatz von KI-Methoden in der Community zu reflektieren.

2 Das Daidalos-Projekt und KI-Bildung

Das interdisziplinäre DFG-Projekt Daidalos³ entwickelt eine Forschungsinfrastruktur, die dem methodischen Ansatz des Forschenden Lernens [HR19] folgt und vorhandene Tools, z. B. LatinCy [Bu23], nachnutzt. Sie soll Forschende der Klassischen Philologie sowie angrenzender Disziplinen befähigen,

1. ein spezifisches Textkorpus zusammenzustellen, zu analysieren und zu adaptieren,
2. die Analyseergebnisse zu visualisieren, zu speichern, zu teilen und zu exportieren,
3. über Ergebnisse und Methodik zu reflektieren, Forschungsfragen zu verfeinern oder aus neuen Perspektiven zu betrachten sowie
4. die eigene Data Literacy auszubauen und über den Einsatz von KI zu reflektieren.

2.1 Verbesserung der Data Literacy als Teil von KI-Bildung

Um unterschiedliche Voraussetzungen und Lernausgangslagen der User berücksichtigen zu können, bietet die Software verschiedene Niveaustufen und Lernangebote:

- Level 1: starke Lenkung durch Demo-Workflows anhand von *user stories* mit konfigurierbaren Elementen, z. B. Textkorpusauswahl; kleinschrittige, detaillierte Erklärungen zu und angeleitete Reflexion über die gewählte NLP-Methode für die *user story* begleitet von Kurztutorials und Übungen (Kooperation mit KI-Campus).
- Level 2: domänenspezifische Oberfläche mit Interaktionen zu Forschungsinteresse, Vorwissen, Korpuszusammenstellung und Methodenwahl; Speicherung von personalisierten Einstellungen; ergänzt durch komplexere Übungen und Code-Snippets in Jupyter Notebooks sowie Workshops zu Datenqualität und Bias (Text+).
- Level 3: freie Konfiguration, Methoden-Triangulation, Einbindung von eigenem Quellcode; weiterführende Literatur und Vernetzung zu anderen DH-Projekten.







Literaturverzeichnis

- [OE19] Ochab, J. K.; Essler, H.: Stylometry of literary papyri. In Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage, 139–142, 2019.

³ <https://hu.berlin/daidalos>

- [As22] Assael, Y. et al.: Restoring and attributing ancient texts using deep neural networks. In *Nature* 603, 280–283, 2022.
- [Sc19] Schüller, K: Ein Framework für Data Literacy. In *AStA Wirtsch Sozialstat Arch* 13, 297–317, 2019.
- [HR19] Huber, L.; Reinmann, G.: *Vom forschungsnahen zum forschenden Lernen an Hochschulen: Wege der Bildung durch Wissenschaft*. Springer, Wiesbaden u. a. 2019.
- [Bu23] Burns, P. J.: *LatinCy: Synthetic Trained Pipelines for Latin NLP*, <https://arxiv.org/abs/2305.04365v1>, 2023.

Learning with ALEA: Tailored experiences through annotated course material

Theresa Kruse ¹, Marc Berges ¹, Jonas Betzendahl ¹, Michael Kohlhasse ¹,
Dominic Lohr ¹ and Dennis Müller ¹

Abstract: We present ALEA, an adaptive learning assistant for university courses. The intelligent tutoring system (ITS) provides services for both learners and instructors, based on semantically annotated course material. The system can generate learning tools like flashcards, guided tours, and quizzes, which can be tailored to individual learning progress. Currently, the system is only available for courses from the field of computer science, like artificial intelligence. We plan to expand it for further courses from different fields.

Keywords: learning assistant, self-learning, learning analytics, adaptive learning

1 Introduction

We present ALEA, an Aaptive Learning Assistant and a university course material platform, currently deployed for ≥ 1000 students in six computer science courses. ALEA adapts the basic teacher knowledge facets of Shulman [Sh86] and encompasses the components of intelligent tutoring systems (ITSs), namely a student module, an expert module and a pedagogical module as well as the user interface [An90]. A *Content Object Server* provides an ontology of abstract concepts to be learned. These concepts are shareable across courses, subjects, and universities. A *Learning Object Server* provides course material as HTML, ranging from individual definitions, exercises, examples, or explanatory snippets, up to full lecture notes. This material is annotated with references to the ontology from the Content Object Server. A *Learner Model Server* (corresponding to the student model of the ITS), associates learners with their estimated mastery of a given concept; represented as a set of triples $\langle C, D, p \rangle$, where C is a concept, D a cognitive dimension according to Anderson & Krathwohl's revision of Bloom's taxonomy for learning objectives [AK09], and $p \in [0, 1]$ a probability representing the learner's assumed competency.

2 ALEA for learners

ALEA's main features for learners all use the *learner model* to tailor content to the individual student. Every learner model starts out empty. Certain actions in ALEA serve as evidence

¹ FAU Erlangen-Nürnberg, Dep. Informatik, Martensstr. 3, 91058 Erlangen, <firstname>.<lastname>@fau.de

of a student's competence level. They are logged and the learner model is updated for all relevant concepts. For example, incorrectly answering a question about the binomial theorem might also lower a student's score for the related concept of polynomials.

2.1 Dashboard and Course material

The core of ALEA is the course material, in particular lecture notes, slides, and videos. For learner's, a main feature is to access definitions of terms by hovering over them in the text. This way the workflow is not interrupted and students are sure to get the correct definition for their context and do not have to navigate homonyms (e.g. *graph* meaning different things in graph theory or as a representation of a function) or instructor-specific re-formulations.

ALEA can also generate a *learner's dashboard*, which presents the learner's progress in a course based on the learner model. Currently, the dashboard is still under development but all necessary information is available in the system. This dashboard also gives hints on which topics need further study. Based on earlier cohorts, ALEA could even estimate of the learner's preparedness for exams.

2.2 Guided Tours

Guided tours allow learners to generate an individualised mini-course for concepts they do not understand starting with their current level of understanding. Currently, this takes the form of a list of definitions the system assumes the student needs to understand to grasp the final concept (all prerequisites minus those the learner model already knows). We plan to make the guided tours more interactive, presenting one learning object at a time and assessing the learner's progress in the style of a dialogue.

2.3 Flashcards and Quizzes

Quizzes and flashcards are two further self-study tools that are included in ALEA, meant to enhance self-regulated learning [Ba77]. Quizzes are built on rapid feedback enhancing students' engagement and providing immediate feedback [Co08] and flashcards have a direct impact on self-regulated learning [WRP12; ZIP22].

The flashcards are created from the annotated course material and thus provide the content exactly as in the course, at no additional investment for the instructor: For each definition in the material, the system knows the terms defined and the defining text fragments. When learners review a flashcard, they can indicate on a five-point scale to which extent they understand and remember each concept. This self-assessment updates the learner model and determines when to show a card to a student again.

The generation of quizzes is not automatic but requires semantically annotated problems. Each of them carries information about its *prerequisites* and its *learning goals* (a question might be about applying the Pythagorean Theorem and require remembering a theorem about angle sums in a triangle). This allows the system to present students appropriate questions and to track their progress. Currently, we support single-choice and multiple-choice questions and clozes.

3 ALeA for Instructors

The role of the instructor is a dual one in ALeA: they annotate the material to create the domain model and learning objects which induce the learning support services, and they profit from teaching support services resulting from the learner data.

3.1 Domain and Learning Object Modeling

We use the $\mathcal{S}\mathcal{T}\mathcal{E}\mathcal{X}$ system [MK22] for the annotation of course materials. $\mathcal{S}\mathcal{T}\mathcal{E}\mathcal{X}$ is a $\mathcal{L}\mathcal{T}\mathcal{E}\mathcal{X}$ package that provides macros and environments for embedding markup for knowledge items, learning objects, and their relations. The union of all concepts introduced in these annotations and their specified relations to each other forms the ontology that the rest of the system operates on. This means there is no necessity to maintain (and ensure compliance with) a separate ontology. This allows easy re-use of concepts but also specification of new concepts or alternative perspectives.

$\mathcal{S}\mathcal{T}\mathcal{E}\mathcal{X}$ source files are convertible to PDF and the HTML-based OMDoc/MMT format for learning/teaching support services. The MathHub content commons (see <https://gl.mathhub.info>) already provide annotated material for more than a dozen courses as well as domain modules for many topics in computer science and mathematics.

3.2 Course dashboard

Analogously to the learner side, ALeA will provide an *instructor dashboard* for each course. It presents activities and progress of students in a course. This includes the mastery of material covered so far (aggregated from the learner model) now and over time. This constitutes a valuable tool that empowers instructors in evaluating their teaching success. As with the learners, the instructor dashboard indicates exam readiness and which topics are best and least understood.

4 Conclusion and future work

We have presented the main features of ALeA from the learner's and instructor's perspectives. The system features dashboards with an overview of learner progress in the course as well

as study features for individuals. The dashboards as well as the annotated course data are special features which we do not see in other ITSs [AF21; Wa23]. Our next steps are to further improve learner models, to enhance guided tours as educational dialogues and to realize the features we designed prototypes for.

4.1 Acknowledgements

The presented research is part of the VoLL-KI project, supported by the Bundesministerium für Bildung und Forschung (BMBF) under grant 16DHBKI089.

References

- [AF21] Alabdulhadi, A.; Faisal, M.: Systematic literature review of STEM self-study related ITSs. *Education and Information Technologies* 26/2, pp. 1549–1588, Mar. 2021.
- [AK09] Anderson, L. W.; Krathwohl, D. R.: *A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives*. Longman, New York, 2009.
- [An90] Anderson, J. R.; Boyle, C. F.; Corbett, A. T.; Lewis, M. W.: Cognitive Modeling and Intelligent Tutoring. *Artif. Intell.* 42/1, pp. 7–49, Feb. 1990.
- [Ba77] Bandura, A.: *Social learning theory*. Prentice Hall, Englewood Cliffs, N.J, 1977.
- [Co08] Cotner, S. H.; Fall, B. A.; Wick, S. M.; Walker, J. D.; Baepler, P. M.: Rapid Feedback Assessment Methods: Can We Improve Engagement and Preparation for Exams in Large-enrollment Courses? *Journal of Science Education and Technology* 17/5, pp. 437–443, 2008.
- [MK22] Müller, D.; Kohlhase, M.: sTeX3 – A L^AT_EX-based Ecosystem for Semantic/Active Mathematical Documents. *TUGboat* 43/2, ed. by Berry, K., pp. 197–201, 2022.
- [Sh86] Shulman, L. S.: Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher* 15/2, pp. 4–14, 1986.
- [Wa23] Wang, H.; Tlili, A.; Huang, R.; Cai, Z.; Li, M.; Cheng, Z.; Yang, D.; Li, M.; Zhu, X.; Fei, C.: Examining the applications of intelligent tutoring systems in real educational contexts. *Education and Information Technologies* 28/7, pp. 9113–9148, July 2023.
- [WRP12] Wissman, K. T.; Rawson, K. A.; Pyc, M. A.: How and when do students use flashcards? *Memory (Hove, England)* 20/6, pp. 568–579, 2012.
- [ZIP22] Zung, I.; Imundo, M. N.; Pan, S. C.: How do college students use digital flashcards during self-regulated learning? *Memory* 30/8, pp. 923–941, 2022.

Künstliche Intelligenz im verpflichtenden Informatikunterricht: Ein skalierbares Fortbildungsangebot für Informatik Lehrkräfte

Sven Baumer, Franz Jetzinger und Tilman Michaeli¹

Abstract: Im Informatikunterricht an bayerischen Gymnasien wird ab dem Schuljahr 23/24 Künstliche Intelligenz unterrichtet. Um Lehrkräfte für dieses neue Thema im Unterricht zu qualifizieren, wurde an der TU München eine Fortbildung entwickelt. Im Beitrag wird die Konzeption des skalierbaren Angebots beschrieben, die insbesondere auch die heterogenen Vorkenntnisse der Lehrkräfte adressiert. Weiterhin werden Erfahrungen und Ergebnisse der ersten Durchführungen berichtet.

Keywords: Künstliche Intelligenz; Fortbildung; Blended Learning; Informatikunterricht

1 Motivation und Hintergrund

Im Rahmen der Umstellung des bayerischen Gymnasiums vom G8 zurück zum G9 wurde in der 11. Jahrgangsstufe verpflichtender Informatikunterricht (zweistündig) für alle Schüler:innen verankert. Dabei wurde ebenfalls ein neuer Inhalt in den Lehrplänen eingeführt: So nimmt das Thema Künstliche Intelligenz (KI) in Klasse 11 etwa ein Viertel des vorgesehenen Informatikunterrichts ein, und zwar sowohl für Schüler:innen, die bereits in den zwei Jahren zuvor Informatikunterricht erhalten haben, als auch in einer „spätbeginnenden“ Variante. Inhalte des Lehrplans sind die Erörterung von möglichen Definitionen, Grundideen sowie Chancen und Risiken von KI. Weiterhin soll ein konkreter Algorithmus des maschinellen Lernens (wahlweise k-nächster-Nachbar oder Lernen von Entscheidungsbäumen) erläutert und angewandt sowie ein einzelnes künstliches Neuron implementiert (spätbeginnend: simuliert) werden. Darüber hinaus wird das Thema KI auch in Klasse 13 erneut intensiv aufgegriffen².

Allerdings stellt KI ein für die informatische Bildung weitgehend neues Themengebiet dar, das in der Didaktik der Informatik gerade erst erschlossen wird [MRS23]. Auch den Lehrkräften fehlt es sowohl an fachlichen sowie unterrichtspraktischen Grundlagen. Um Informatik Lehrkräfte also für diesen neuen Inhaltsbereichs vorzubereiten und bei der Ausgestaltung eines handlungsorientierten und abwechslungsreichen Unterrichts zu unterstützen, haben wir an der TUM in Kooperation mit dem Bayerischen Staatsministerium für Unterricht und Kultus ein Fortbildungsangebot entwickelt.

¹ TU München, Didaktik der Informatik, Arcisstr. 21, 80333 München, sven.baumer@tum.de

² <https://www.lehrplanplus.bayern.de/>

2 Rahmenbedingungen und Herausforderungen

Zur Gestaltung von Fortbildungsangeboten für Lehrkräfte liegen umfangreiche Forschungsergebnisse vor (siehe etwa [LR12]). Allerdings fehlt es bisher an Untersuchungen spezifisch für das Themengebiet KI, die etwa Hinweise liefern könnten, wie die besonderen Herausforderungen eines solch neuen Themas informatischer Bildung bei einer so umfassenden Integration in die Curricula adressiert werden können.

Unterricht zu KI findet im bayerischen Gymnasium in Klassenstufe 11 (verpflichtend, ab 23/24) sowie Klassenstufe 13 (Wahlbereich, ab 25/26) statt. Aufgrund der umfangreichen zu vermittelnden Inhalte und um *Wissen auf Vorrat* zu vermeiden, haben wir uns für eine spiralcurriculare Aufteilung in eine Basisfortbildung (Grundlagen, 11. Klasse) und Aufbaufortbildung (Vertiefung, 13. Klasse) entschieden. Letztere beginnt dabei erst ab dem Schuljahr 24/25, ist für einen deutlich kleineren Kreis der Lehrkräfte relevant und kann auf deren erste Unterrichtserfahrung zum Thema aufbauen.

Zielsetzung der Basisfortbildung ist sowohl die grundlegende fachliche als auch fachdidaktische Qualifizierung der Lehrkräfte, die auch die konkreten Anforderungen des Lehrplans berücksichtigt. KI stellt dabei bisher keinen verpflichtenden Inhalt des Lehramtsstudiums in Bayern dar. Darüber hinaus hat der mit Abstand größte Teil der bayerischen Informatiklehrkräfte kein grundständiges Informatik-Lehramtsstudium absolviert, sondern lediglich ein stark verkürztes *Erweiterungsstudium* [Mi23]. Damit ist einerseits von *geringen Vorerfahrungen* zum Thema KI auszugehen, wobei vereinzelt Lehrkräfte sich bereits im Selbststudium in das Thema eingearbeitet haben dürften. Andererseits ist insgesamt eine *große Heterogenität* auch bezüglich weiterer informatischer Grundlagen und Unterrichtserfahrung im Informatikunterricht zu erwarten.

Im regional weit gestreuten Einzugsgebiet der TU München sind etwa 450 Informatiklehrkräfte an Gymnasien beschäftigt. Mit einem Vorlauf von lediglich knapp über einem Jahr und begrenzten personellen Ressourcen erscheint daher ein *skalierbares* Format zentral (etwa mit umfangreichen Selbstlernphasen), um für einen Großteil der Lehrkräfte ein Fortbildungsangebot bereitzustellen. Gleichzeitig ist die persönliche Vernetzung der Lehrkräfte im Sinne einer **Community of Practice** für dieses neue Thema wichtig.

3 Aufbau des Fortbildungsangebots

Zunächst haben wir anhand von grundlegenden Strukturierungen des Themengebiets aus fachdidaktischer Perspektive [MRS23, To19] sowie dem Lehrplan die intendierten fachlichen Lernergebnisse des Angebots abgeleitet. Um die persönliche Vernetzung der Lehrkräfte bei gleichzeitiger zeit- und ressourcenschonender Skalierbarkeit sicherzustellen, haben wir uns für ein Blended-Learning-Format entschieden: Zu Beginn und Ende einer Fortbildungssequenz findet jeweils ein Präsenztage (Dauer 3 bzw. 6 Stunden) im Abstand von ca. 6 Wochen statt. Dazwischen wird individuell der MOOC (Massive Open Online Course)

"Die Welt der KI entdecken"³ absolviert, der bereits 2021 von einem Teil der Autor:innen zur Vermittlung der Grundlagen von KI für die interessierte Allgemeinheit entwickelt wurde. Der MOOC besteht aus drei Wochen-Einheiten (3 Stunden Lernaufwand pro Woche) mit kurzen Videos, Aufgaben und Selbsttests. Neben der Skalierbarkeit ermöglicht dieser Ansatz, selbstdifferenziert die unterschiedlichen Bedürfnisse und den individuellen Kenntnisstand der Teilnehmer:innen zu adressieren.

Die flankierenden Präsenztage vor und nach dem MOOC sind auf etwa dreißig Personen begrenzt. Zur Vermittlung der Inhalte wurde, soweit möglich, im Sinne des didaktischen Doppeldeckers auf erprobte Materialien für den Informatikunterricht zurückgegriffen, um die direkte Anwendung im Unterricht zu ermöglichen. Der erste Präsenztage dient der Vernetzung der Teilnehmer:innen und um „Lust auf mehr“ zu machen. Dazu wird ein Gesamtüberblick zu KI und den Anforderungen des Lehrplans gegeben und mithilfe vieler spielerischer Ansätze werden [SLR19] grundlegende Ideen vermittelt. Am abschließenden Präsenztage werden einerseits bestimmte lehrplanspezifische fachliche Inhalte vertieft, die im MOOC nur angeschnitten werden. Andererseits werden vielfältige Möglichkeiten zur Umsetzung anhand erprobter Materialien im Unterricht exploriert. Diese Sequenz wird seit Januar 2023 bis April 2024 je nach Bedarf ca. zwölfmal angeboten.

4 Auswertung und Erfahrungen

Zum aktuellen Zeitpunkt wurde die Fortbildung in zwei Kohorten mit knapp über 60 Teilnehmer:innen vollständig durchgeführt, die dritte Kohorte hat bisher lediglich den ersten Präsenztage absolviert. Um sowohl die Vorerfahrungen der Lehrkräfte zu erheben und das Fortbildungsangebot zu evaluieren, als auch den Lehrkräften die Breite des Themengebiets zu verdeutlichen, haben wir einen Progress-Test eingesetzt. Dieser besteht aus 15 Fragen in einem geschlossenen Antwortformat, die einen Großteil der für die gesamte Fortbildung intendierten Lernziele abdecken. Die Lehrkräfte werden gebeten, bereits vor dem ersten Termin in ca. 15 Minuten den Test online durchzuführen, und wiederholen ihn am Ende der Sequenz (also fast zwei Monate später) mit identischen Fragen. Die Teilnehmer:innen erhalten dabei keine Rückmeldung zu den einzelnen Antworten, sondern nur das prozentuale Gesamtergebnis und können bei jeder Frage „weiß ich nicht“ angeben, um Raten vorzubeugen.

Betrachtet man die Vorerfahrungen der Lehrkräfte, zeigt sich die erwartete Heterogenität (vgl. Abb. 1). Allerdings hat uns der große Anteil an Teilnehmer:innen überrascht, der bereits im Pretest eine sehr hohe Punktzahl erreicht hat. Im Vergleich der einzelnen Kohorten zeigt sich dabei ein abnehmender Trend. Unsere Vermutung ist daher, dass es sich bei vielen der Teilnehmer:innen der ersten Fortbildungen um besonderes motivierte Lehrkräfte handelt, die sich selbst bereits umfangreich in das Thema eingearbeitet haben – so dauerte es zu diesem Zeitpunkt noch über ein Jahr bis zum ersten Unterricht zu KI.

³ <https://open.sap.com/courses/ai1-de>

Die anekdotischen Eindrücke aus den Fortbildungen und Gesprächen mit den Lehrkräften bestätigen diese Hypothese. Wir gehen daher auch für die weiteren Durchführungen von im Schnitt geringeren Vorerfahrungen aus. Gebiete, in denen sich unterdurchschnittliches Vorwissen zeigte, waren insbesondere zu wissensbasierten Ansätzen der KI, der Geschichte der KI, dem Datenlebenszyklus sowie der Funktionsweise eines Perzeptrons.

Für die bereits vollständig durchgeführten ersten beiden Kohorten zeigt die Auswertung der Ergebnisse (vgl. Abbildung 2, nur Personen, für die wir ein Matching von Pre- zu Posttest herstellen konnten) ein signifikant besseres Abschneiden. Die positiven Rückmeldungen der Lehrkräfte bestätigen den Erfolg des Fortbildungsangebots. So heben die Lehrkräfte v.a. den flexiblen Aufbau mit vielen unterschiedlichen Zugängen hervor. Besonders betont wird der zur Verfügung gestellte Advance Organizer, welcher von Beginn an Überblick und Hilfestellung bietet. Außerdem werden die vielen direkt im eigenen Unterricht einsetzbaren Materialien von den Lehrkräften geschätzt.

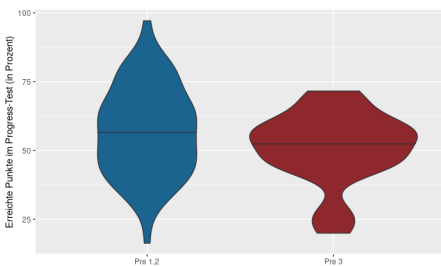


Abb. 1: Vorerfahrungen der Kohorten 1 und 2 (n=63) vs. 3 (n=25)

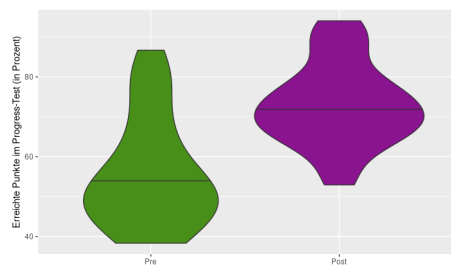


Abb. 2: Pre-Post Vergleich der Kohorten 1 und 2, falls Matching möglich (n=24)

Literaturverzeichnis

- [LR12] Lipowsky, F.; Rzejak, D.: Lehrerinnen und Lehrer als Lerner – Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen. Reform der Lehrerbildung in Deutschland, Österreich und der Schweiz. Teil, S. 235–253, 2012.
- [Mi23] Michaeli, T.: Die Henne und das Ei – Absolvent:innen im Lehramt Informatik 20 Jahre nach Einführung des Schulfaches in Bayern. In: INFOS 2023. Gesellschaft für Informatik, Bonn, 2023.
- [MRS23] Michaeli, T.; Romeike, R.; Seegerer, S.: What students can learn about artificial intelligence – recommendations for K12 computing education. In: Proceedings of WCCE 2022. Springer, 2023.
- [SLR19] Seegerer, S.; Lindner, A.; Romeike, R.: AI Unplugged – Wir ziehen Künstlicher Intelligenz den Stecker. In: INFOS 2019. Gesellschaft für Informatik, 2019.
- [To19] Touretzky, D.; Gardner-McCune, C.; Martin, F.; Seehorn, D.: Envisioning AI for K-12: What should every child know about AI? In: Proceedings of the AAAI conference on artificial intelligence. Jgg. 33, S. 9795–9799, 2019.

Künstliche Intelligenz in der universitären Lehrerinnen- und Lehrerbildungsphase

Gia Minh Vo und Nils Pancratz¹

Abstract: Die zunehmende Relevanz von Künstlicher Intelligenz (KI) im täglichen Leben führt zu einer steigenden Nachfrage nach *KI in der Bildung*. Es ist daher erforderlich, dass angehende Lehrerinnen und Lehrer im Rahmen ihrer universitären Lehramtsausbildung angemessenes Professionswissen zu diesem Thema vermittelt bekommen. Im Rahmen eines Forschungsprojekts wurden für den vorliegenden Work-In-Progress Beitrag 14 Lehramtsstudierende mittels leitfadengestützter Einzelinterviews untersucht, um ihre Vorstellungen zur Integration von KI in die universitäre Lehramtsausbildung zu ermitteln. Die Lehramtsstudierenden setzen sich dabei aus zwei Gruppen zusammen: einer ersten Gruppe von Lehramtsstudierenden, die das Fach Informatik studieren, und einer zweiten Gruppe von Lehramtsstudierenden, die das Fach Informatik nicht studieren. Die Ergebnisse zeigen, dass sich die Meinungen und Bedürfnisse beider Gruppen hinsichtlich der Integration von KI in die universitäre Lehramtsausbildung unterscheiden. Darüber hinaus äußerten die Lehramtsstudierenden erste Ideen und Inhalte zur konkreten Umsetzung von KI-Bildung an der Hochschule. Basierend auf den vorläufigen Ergebnissen werden mögliche thematische Inhalte und didaktische Implikationen für KI-Bildung in der universitären Lehrerinnen- und Lehrerbildungsphase für alle diskutiert.

Keywords: Künstliche Intelligenz, Lehramtsstudierende, Leitfadeninterviews, Hochschuldidaktik

1 Einleitung und Motivation

Das Gebiet der Künstlichen Intelligenz (KI) ist derzeit eines der am schnellsten wachsenden Forschungsfelder in der Informatik [RN21]. Die jüngste Integration von KI-Themen in die Lehrpläne einiger Bundesländer wie bspw. Nordrhein-Westfalen [Mi21] zeigt bereits, dass das Thema zunehmend eine Rolle im Informatikunterricht einnimmt. Die aktuell kontrovers geführte Debatte über den Einsatz von KI-Anwendungen im Unterricht, wie bspw. die Verwendung von Chatbots wie ChatGPT, verdeutlicht außerdem die zunehmende Relevanz des Themas auch in allen anderen Schulfächern. Dies stellt sowohl Informatik-Lehramtsstudierende (INF-Lehramtsstudierende) als auch Lehramtsstudierende, die nicht das Fach Informatik studieren (NINF-Lehramtsstudierende), vor die Herausforderung, sich mit einer Vielzahl neuer und anspruchsvoller Themen für die Schule auseinanderzusetzen zu müssen. Eine notwendige Aufgabe der universitären Lehrerinnen- und Lehrerbildung (erste Bildungsphase) besteht darin, die aktuellen Entwicklungen zu berücksichtigen und sie angemessen als Teil des *Professionswissens* [Bo10] in die Lehramtsausbildung zu integrieren. Professionswissen umfasst sämtliche theoretisch fundierten Wissensbestandteile, die im

¹ Universität Hildesheim, Universitätsplatz 1, Didaktik der Informatik, 31141 Hildesheim, {vo,pancratz}@imai.uni-hildesheim.de

Zuge der Lehrerinnen- und Lehrerausbildung sowie während der Unterrichtspraxis erlangt werden können [Bo10]. Shulman [Sh86] identifiziert in seiner Taxonomie insgesamt sieben Bereiche des professionellen Lehrerinnen- und Lehrerwissens, wobei insbesondere die drei Bereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen als besonders relevant für den Unterricht erachtet und dementsprechend häufig untersucht werden [Bo10]. So untersucht auch Lindner [Li21] das Professionswissen von Lehrpersonen zu KI und stellt erste KI-Kompetenzen für das Fachwissen (content knowledge) und das fachdidaktische Wissen (pedagogical content knowledge) auf. In der deutschen Informatikdidaktik-Community herrscht Konsens darüber, dass *alle* Lehrpersonen, unabhängig von ihren konkreten Fächern, angesichts einer von Informatik durchdrungenen Welt einen Zugang zu allgemeinbildenden Elementen der Informatik benötigen [LH19]. Ein solches Element ist das Thema KI, das aufgrund seiner steigenden interdisziplinären Bedeutung und Relevanz für die Allgemeinbildung für alle Lehramtsstudierenden von großer Bedeutung ist. In einer vorangegangenen Arbeit wurden bereits die Vorstellungen von Lehramtsstudierenden zu KI untersucht [VP23]. In diesem Beitrag wird die Perspektive von Lehramtsstudierenden hinsichtlich der Integration von KI-Bildung in ihre eigene Lehramtsausbildung untersucht. Brooman et al. [BDP15] betonen in ihrer Studie die Bedeutung, die Sichtweise der Studierenden bei der Entwicklung des Lehrplans zu berücksichtigen, da dieser häufig ausschließlich aus der Perspektive der Lehrenden entsteht und nicht immer den Bedürfnissen der Studierenden entspricht. Der Einbezug der Studierenden kann dazu beitragen, die Unterrichtsqualität zu verbessern und die Akzeptanz der vermittelten Inhalte zu erhöhen. Die vorliegende Studie adressiert die beschriebene Ausgangslage anhand der folgenden Forschungsfragen:

1. Welche Meinungen und Positionen vertreten die Informatik-Lehramtsstudierenden und Nicht-Informatik-Lehramtsstudierenden hinsichtlich einer Einführung von Künstlicher Intelligenz in die universitäre Lehramtsausbildung?
2. Welche Themen und Kompetenzen sollten aus Sicht der Lehramtsstudierenden in einer Veranstaltung zur Künstlichen Intelligenz in der universitären Lehramtsausbildung behandelt werden?

2 Darstellung des methodischen Vorgehens

Im Rahmen der vorliegenden Studie fiel die Wahl des Datenerhebungsinstruments auf ein leitfadengestütztes Einzelinterview. Die Einzelinterviews wurden aufgezeichnet, transkribiert und pseudonymisiert. Die vorliegenden Ergebnisse basieren auf der Durchführung und Auswertung von 14 Interviews, davon sechs mit NINF-Lehramtsstudierenden und acht mit INF-Lehramtsstudierenden. Die Auswahl der Lehramtsstudierenden erfolgte nach dem Prinzip der minimalen und maximalen Kontrastierung (Stichprobenziehung), um eine möglichst große Variation in Bezug auf Geschlecht, Fächerkombination und Studienerfahrung (Bachelor oder Master) zu gewährleisten. Zur Auswertung wurden die Einzelinterviews (n=14) einer qualitativen Inhaltsanalyse nach Mayring [Ma02] mit induktiver Kategorienbildung unterzogen und mithilfe der Software MAXQDA 2022 analysiert.

3 Erste Ergebnisse

Die Frage nach dem Ort bzw. der Form der Integration von KI in die *eigene* universitäre Lehrerinnen- und Lehrerbildung (Forschungsfrage 1) gestaltet sich äußerst heterogen bei den INF- (K1) und NINF-Lehramtsstudierenden (K2). Abbildung 1 zeigt die unterschiedlichen Kategorien der Lehramtsstudierenden.

Kategorie	Definition	Ankerbeispiel
K1.1: <i>Schwerpunkt Fachwissenschaft mit Integration in die Fachdidaktik</i>	Die Lehramtsstudierenden wünschen sich eine verpflichtende fachwissenschaftliche Veranstaltung, KI-Inhalte können zusätzlich in eine fachdidaktischen Veranstaltung integriert werden.	„[...] die Vorlesung "Machine Learning" sollten wir auch gemeinsam belegen wie die Informatiker [...]. Außerdem könnte man dann noch KI-Inhalte in dem Didaktik-Seminar [...] von dir oder [...] integrieren.“
K1.2: <i>Eigenständige fachdidaktische Veranstaltung</i>	Die Lehramtsstudierenden wünschen sich eine eigenständige fachdidaktische Veranstaltung, in die auch fachwissenschaftliche Inhalte integriert werden.	„[...] man kann daraus natürlich auch ein eigenes Seminar machen wie in Mathematik [...]. Die technischen Hintergründe könnte man dann auch in das Seminar einbauen.“
K1.3: <i>Ausschließliche Ein- bindung von KI in die Fachdidaktik</i>	Die Lehramtsstudierenden wünschen sich eine Integration von KI ausschließlich in die fachdidaktischen Veranstaltungen.	„[...] ich denke, dass man KI gut in der Didaktik-Vorlesung von [...] oder in einem Seminar von dir integrieren kann. In die Tiefe der Mathematik [...] muss man auch nicht gehen, die Grundlagen würden mir ausreichen [...].“
K2.1: <i>Teil einer Ring- vorlesung</i>	Die Lehramtsstudierenden wünschen sich eine Integration von KI als Teil einer Ringvorlesung für alle Lehramtsstudierenden.	„[...] ich fänds interessant, wenn wir vielleicht so 1-2 Vorlesungen in der Ringvorlesung hören könnten. Da gibt es ja im Master die Ringvorlesung [...]. Die mussten wir auch alle belegen.“
K2.2: <i>Integration in die eigenen Fächer</i>	Die Lehramtsstudierenden wünschen sich eine direkte Verknüpfung der KI-Inhalte mit ihren eigenen studierten Fächern.	„[...] schön wäre es auch [...] wenn ich [...] z.B. auch in Religion im Seminar über Ethik und KI diskutieren könnte [...]. Also für meine eigenen Fächer [...] in der Schule.“
K2.3: <i>Keine verpflicht- ende Einbindung</i>	Die Lehramtsstudierenden möchten keine verpflichtende KI-Veranstaltung im Studium, kann jedoch als Wahlmöglichkeit im Optionalbereich angeboten werden.	„[...] mein Stundenplan ist wirklich sehr voll und man muss dann ja eine Veranstaltung streichen oder so. Vielleicht als Wahlveranstaltung wäre das in Ordnung [...]“.

Hinweis. Die direkten Aussagen der Studierenden wurden geringfügig sprachlich und grammatikalisch angepasst.

Abb. 1: Bei der Inhaltsanalyse verwendeter Kategorienleitfaden

Einige INF-Lehramtsstudierende betonen die Notwendigkeit, *alle* Lehramtsstudierende für KI und weitere Inhalte der Informatik zu sensibilisieren. Die Forderung dieser Studierenden basiert auf den Erfahrungen mit den betreuenden Informatik-Mentoren während der Schulpraktika: Sie wurden als fachfremde Lehrkräfte beschrieben, die aufgrund des Lehrkräftemangels das Fach Informatik unterrichten (müssen). Ein INF-Lehramtsstudent berichtet von teilweise naiven bzw. fehlerhaften Vorstellungen eines betreuenden fachfremden Informatiklehrers während des Schulpraktikums und appelliert daher auch an die Bedeutung eines angemessenen fachwissenschaftlichen Hintergrunds. Es wurde auch der Vorschlag geäußert, fachdidaktische Veranstaltungen mit KI-Inhalten als optionale Wahlveranstaltung für *alle* informatiknahen Studiengänge zu öffnen.

Die bisherigen Antworten der Lehramtsstudierenden (n=14) zu konkreten Inhalten von KI-Bildung in der Hochschule (Forschungsfrage 2) lassen sich im laufenden Forschungsprojekt in vier Dimensionen (Dim.) einordnen.

(Dim. 1) Perspektiven des Dagstuhl-Dreiecks. Bildung in der digital vernetzten Welt muss nach dem Dagstuhl-Dreieck aus technologischer, gesellschaftlich-kultureller und anwendungsbezogener Perspektive in den Blick genommen werden [Ge16]. Die Aussagen der Lehramtsstudierenden lassen sich in diese drei Dagstuhl-Perspektiven einordnen. In der gesellschaftlichen Perspektive äußern sie den Wunsch nach einer Auseinandersetzung mit ethischen Aspekten der KI sowie den Auswirkungen von KI-Systemen auf die Gesellschaft. Auch historische Entwicklungen und Meilensteine in der KI-Entwicklung wurden genannt. In der technologischen Perspektive äußern die INF-Lehramtsstudierenden den Wunsch nach einer Einführung in das „Maschinelle Lernen“ (K1.1). Einige INF-Lehramtsstudierende weisen jedoch darauf hin, dass sich die technologische Perspektive auf die Vermittlung von „Grundlagen“ (K1.3) beschränken sollte und weiterführende KI-Inhalte optional in fachwissenschaftlichen Veranstaltungen angeboten werden könnten. Für den Großteil der NINF-Lehramtsstudierenden wird die technologische Perspektive hingegen abgelehnt, da sie nach eigener Aussage bewusst kein MINT-Fach gewählt haben und daher die Verantwortung bei den MINT-Lehramtsstudierenden sehen. Die INF-Lehramtsstudierenden verknüpfen die anwendungsbezogene mit der technologischen Perspektive. Im Gegensatz dazu haben die NINF-Lehramtsstudierenden eine abweichende Auffassung dieser Perspektive und sehen keine Verbindung zur technologischen Perspektive. Für sie liegt der Fokus ausschließlich auf der Nutzung und sie wünschen sich bspw. in einer Ringvorlesung (K2.1) eine „KI-Werkzeugliste“ mit verschiedenen KI-Systemen. Häufig wurden KI-Systeme erwähnt, die den individuellen Lernfortschritt der Schülerinnen und Schüler (SuS) überprüfen und korrigieren können sowie KI-Tools zur Erkennung von Plagiaten in schriftlichen Arbeiten.

(Dim. 2) Didaktische Reduktion: Förderung der Entwicklung von Fachwissen und fachdidaktischem Wissen. Die INF-Lehramtsstudierenden wünschen sich in der KI-Veranstaltung einen stärkeren Praxisbezug. Sie gaben an, Schwierigkeiten zu haben, die „Komplexität“ des Themas für die Sekundarstufe I angemessen zu reduzieren. Ihr Ziel ist es, konkrete Unterrichtsmaterialien zu KI zu erstellen, die auch als Open Educational Resources (OER) entwickelt und gegebenenfalls erprobt werden können. Durch die Entwicklung und Veröffentlichung dieser OER-Unterrichtsmaterialien könnten nach eigenen Aussagen auch praktizierende Informatiklehrpersonen von den Materialien profitieren.

(Dim. 3) Analyse von Lehrplänen zu KI. Für die INF-Lehramtsstudierenden gestaltet sich die Frage, welche konkreten Kompetenzen die SuS für eine KI-Bildung in der Schule erwerben sollten, als äußerst schwierig. Die INF-Lehramtsstudierenden äußern Kritik am Kerncurriculum (KC) in ihrem Bundesland (Niedersachsen), da es keine explizite Behandlung von Themen wie KI oder maschinellem Lernen (ML) vorsieht. Sie appellieren daher an eine Aktualisierung des KCs, um die Themen KI und ML einzubeziehen. Bezüglich der KI-Veranstaltung gaben die INF-Lehramtsstudierenden an, dass sie auch Lehrpläne aus anderen Bundesländern heranziehen könnten, in denen KI-Inhalte bereits explizit in die

Sekundarstufe I integriert wurden. Nach eigenen Aussagen könnten sie diese Lehrpläne analysieren und für ihr eigenes Bundesland reflektieren.

(Dim. 4) *Möglichkeiten des fachübergreifenden Unterrichts*. Angesichts der zunehmenden Bedeutung und Anwendung von KI-Systemen, wie bspw. von Chatbots wie ChatGPT im Unterricht, denken Lehramtsstudierende darüber nach, ob und wie sie dieses Thema in ihren eigenen Lehrplan und Unterricht integrieren können. Die NINF-Lehramtsstudierenden wünschen sich bspw. in einer Ringvorlesung (K2.1) Anregungs- und Vernetzungsmöglichkeiten. Ziel ist es, KI-Inhalte mit ihrem eigenen Lehrplan in den entsprechenden Fächern zu verknüpfen und potenzielle Diskussionspunkte für den Unterricht zu identifizieren.

4 Diskussion und Ausblick



Um Lehramtsstudierende angemessen auf den Unterricht über KI vorzubereiten, ist es erforderlich, dass sie das notwendige Professionswissen im Bereich KI erwerben. Dieses Lehrerinnen- und Lehrerwissen sollte die gesellschaftlich-kulturellen, anwendungsbezogenen und technologischen Perspektiven im Sinne des Dagstuhl-Dreiecks [Ge16; MRS22] umfassen. Ein fehlender technologischer Hintergrund, wie von einigen Lehramtsstudierenden (K1.3; Dim. 1) gewünscht wird, stellt ein potenzielles Problem dar, da sie möglicherweise ihre eigenen vorhandenen naiven bzw. fehlerhaften Vorstellungen über KI an ihre SuS weitergeben könnten. Nur eine *ganzheitliche* Betrachtung des Dagstuhl-Dreiecks [Ge16] kann einen adäquaten Unterricht über KI gewährleisten und bestehende naive Vorstellungen adressieren. Es liegt auch in der Verantwortung der einzelnen Hochschulen, eine verpflichtende Einführung in KI im Rahmen einer informatischen Bildung für *alle* in die Studienpläne der Lehramtsstudierenden zu integrieren. Um die KI-Unterrichtsmaterialien der Lehramtsstudierenden (Dim. 2) zu erproben und zu evaluieren, könnte eine Verbindung zu außerschulischen Lernorten der SuS [PS17] hergestellt werden. Durch diese Zusammenarbeit ließen sich auch vorhandene Schülerinnen- und Schülervorstellungen zu KI [LBL21] unter Berücksichtigung der Didaktischen Rekonstruktion für den Informatikunterricht [Di11] nutzen und didaktisch aufbereiten. Im Sinne des forschenden Lernens (Dim. 3) wäre es auch möglich, im Rahmen der KI-Veranstaltung internationale ML-Frameworks für den Unterricht wie die Arbeit von Lao [La20] sowie wissenschaftliche Arbeiten zur Entwicklung von KI-Lehrplänen wie die Arbeit von Michaeli et al. [MRS22] zu untersuchen und in die Lehrpläne des eigenen Bundeslandes für die Sekundarstufe zu integrieren.

Um eine holistische Perspektive sicherzustellen, werden im nächsten Schritt Lehrpersonen und Experten aus den Bereichen Informatik, Informatikdidaktik, Erziehungs- und Bildungswissenschaften mittels einer *Delphi-Studie* untersucht. Das Ziel dieser Untersuchung besteht darin, einerseits die zentralen Inhalte und Kompetenzen für eine KI-Bildung bei INF-Lehramtsstudierenden zu ermitteln, die im Rahmen einer universitären Lehrveranstaltung vermittelt werden sollten. Andererseits soll auch – im Sinne einer informatischen Bildung für *alle* [LH19; VP23] – eine ausreichende KI-Bildung für NINF-Lehramtsstudierende abgeleitet werden.

Literatur

- [BDP15] Brooman, S.; Darwent, S.; Pimor, A.: The student voice in higher education curriculum design: is there value in listening? *Innovations in Education and Teaching International* 52/6, S. 663–674, 2015.
- [Bo10] Borowski, A.; Neuhaus, B.; Tepner, O.; Wirth, J.; Fischer, H.; Leutner, D.; Sandmann, A.; Sumfleth, E.: Professionswissen von Lehrkräften in den Naturwissenschaften – Kurzdarstellung des BMBF-Projekts. In: *ZfDN*. 2010.
- [Di11] Diethelm, I.; Dörge, C.; Mesaros, A.-M.; Dünnebier, M.: Die Didaktische Rekonstruktion für den Informatikunterricht. In: *Informatik in Bildung und Beruf*. Gesellschaft für Informatik e.V., Bonn, S. 77–86, 2011.
- [Ge16] Gesellschaft für Informatik e.V.: Dagstuhl-Erklärung. *Bildung in der digitalen vernetzten Welt*, 2016, URL: <https://dagstuhl.gi.de/dagstuhl-erklaerung>, Stand: 28.08.2023.
- [La20] Lao, N.: Reorienting machine learning education towards tinkerers and ML-engaged citizens. PhD thesis. Massachusetts Institute of Technology, Cambridge, 2020.
- [LBL21] Lindner, A.; Berges, M.-P.; Lechner, M.: KI im Toaster? Schüler:innenvorstellungen zu künstlicher Intelligenz. In: *INFOS 2021*. Gesellschaft für Informatik, Bonn, S. 133–142, 2021.
- [LH19] Losch, D.; Humbert, L.: Informatische Bildung für alle Lehramtsstudierenden. In: *INFOS 2019*. Gesellschaft für Informatik, Bonn, S. 119–128, 2019.
- [Li21] Lindner, A.: Designing a Teacher PD Programme for AI – First Steps. In: *WiPSCE 21*. ACM, New York, USA, S. 1–2, 2021.
- [Ma02] Mayring, P.: *Einführung in die qualitative Sozialforschung*. Beltz Verlag, Weinheim und Basel, 2002.
- [Mi21] Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen: Kernlehrplan für die Sekundarstufe I - Klasse 5 und 6 in NRW, Informatik. 2021.
- [MRS22] Michaeli, T.; Romeike, R.; Seegerer, S.: What students can learn about artificial intelligence - recommendations for K-12 computing education. In: *Proc. of IFIP WCCE 2022: World Conference on Computers in Education*. Hiroshima, 2022.
- [PS17] Prädell, L.; Schreiber, G.: Hochschule als außerschulischer Lernort für Schülerinnen und Schüler. In: *INFOS 2017*. GI, Bonn, S. 221–226, 2017.
- [RN21] Russell, S. J.; Norvig, P.: *Artificial Intelligence: A Modern Approach*. Pearson, Harlow, 2021.
- [Sh86] Shulman, L. S.: Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher* 15/2, S. 4–14, 1986.
- [VP23] Vo, G. M.; Pancratz, N.: Vorstellungen von Lehramtsstudierenden zu künstlicher Intelligenz. In: *INFOS 2023*. Gesellschaft für Informatik, Bonn, 2023.

Design Review von offenen Online-Kursen zum Thema Künstliche Intelligenz

Marc Egloffstein ¹ und Kristina Kögler ²


Abstract: Der Beitrag berichtet die ersten Ergebnisse eines explorativen Design Reviews von offenen Online-Kursen zum Thema Künstliche Intelligenz (KI). In einem Rating-Verfahren wurden N=41 zufällig ausgewählte Kurse des KI-Campus auf 15 theoriebasiert entwickelte didaktische Gütekriterien hin überprüft. Die Ergebnisse zeigen, dass die untersuchten Kurse gut strukturierte und systematisch aufbereitete Lernangebote darstellen. Gleichzeitig deuten sich Unterstützungspotenziale durch KI-Technologien an. Für die Weiterentwicklung von digitalen Lernangeboten zum Thema KI erscheint die Kopplung der Perspektiven „KI als Lerninhalt“ und „KI als Lernwerkzeug“ Erfolg versprechend.

Keywords: MOOCs, Instructional Design, KI-Bildung, Design Review

1 Hintergrund und Problemstellung

Die gegenwärtigen und zukünftigen Umwälzungen durch Künstliche Intelligenz (KI) betreffen alle Gesellschaftsbereiche. Um den daraus erwachsenden enormen Aus- und Weiterbildungsbedarfen zeitnah begegnen zu können, sind Online-Lernangebote auf digitalen Bildungsplattformen von zentraler Bedeutung.

Skalierbarkeit, Flexibilität, Time-to-market, die mögliche Passung von Inhalt und Methode sowie die Verfügbarkeit von etablierten Plattformen sprachen dafür, die initialen Angebote des KI-Campus als Massive Open Online Courses (MOOCs) zu realisieren. Im aktuellen Verständnis sind damit in der Regel Selbstlernumgebungen gemeint, die videobasierte Inhalte mit Selbsttests zur Lernerfolgskontrolle kombinieren [Eg18]. Aus der akademischen Perspektive (MOOCs als Lehrveranstaltungen oder Kurse) werden auch nach der inzwischen erfolgten Etablierung und Konsolidierung des Lernformats didaktische Herausforderungen wie hohe Drop-Out-Raten [BCP22] und offene Fragen hinsichtlich der Instruktionsqualität [MBL15] berichtet, die auf konzeptionelle und praktische Optimierungspotenziale hindeuten. Ob dies in gleicher Weise für MOOCs im Bereich Aus- und Weiterbildung zutrifft, steht indes zur Debatte.

¹ Universität Mannheim, Wirtschaftspädagogik – Technologiebasiertes Instruktionsdesign, L4,1,
68161 Mannheim, egloffstein@uni-mannheim.de,  <https://orcid.org/0000-0002-2377-7889>

² Universität Stuttgart, Berufs-, Wirtschafts- und Technikpädagogik, Geschwister-Scholl-Str. 24 D,
70174 Stuttgart, koegler@bwt.uni-stuttgart.de,  <https://orcid.org/0000-0002-6766-1685>

Anknüpfend an bestehende Forschungsansätze [MBL15; EKI19; OCP20; KVK21] haben wir vor diesem Hintergrund verschiedene Online-Angebote aus dem Portfolio des KI-Campus einem explorativen Rating unterzogen. Die zentralen Fragestellungen lauteten dabei, wie sich diese Kurse hinsichtlich wesentlicher Instructional Design-Kriterien darstellen und welche Implikationen sich daraus für die Optimierung und Implementierung der Kurse in verschiedenen Lernsettings ergeben könnten.

2 Design Review

2.1 Instrument und Vorgehen

Die Analyse des didaktischen Designs der Kurse zur KI-Bildung wurde ausgehend von den „First Principles of Instruction“ [Me13], deren Operationalisierung in Referenzstudien [MBL15] sowie einschlägigen Voranalysen [EKI19] mit einem Ratinginstrument durchgeführt, das über drei Bereiche hinweg insgesamt 20 Items enthält. Bereich A umfasst fünf Items in fünf Kategorien, die sich auf die Strukturiertheit und Klarheit von Kursen beziehen. Eine hohe Bewertung impliziert eine klare und umfassende Beschreibung des Kurses, der Struktur, des Inhalts, des erwarteten Aufwands, der Zielgruppe und der entsprechenden Lernziele. Bereich B enthält, anknüpfend an [Me13], die fünf zentralen didaktischen Kategorien Problemzentriertheit, Aktivierung, Demonstration, Anwendung und Integration mit jeweils zwei Rating-Items. Bereich C besteht aus fünf Items in fünf weiteren didaktischen Kategorien wie Feedback, Zusammenarbeit und Kooperation, Authentizität von Lernmaterialien sowie Individualisierung und Differenzierung. Jedes Item wird auf einer Skala von 0 (trifft überhaupt nicht zu, d. h. nicht vorhanden) bis 3 (trifft sehr zu, d. h. in hohem Maße vorhanden) bewertet. Zur Berechnung eines Gesamtscores werden die Bereiche ausgehend von inhaltlichen Erwägungen im Verhältnis 1:2:2 für die zu erreichenden Punkte gewichtet. Alles in allem ergibt sich eine gewichtete Gesamtpunktzahl von maximal 75 Punkten über die drei Bereiche als theoriebasiertes Maß für die didaktische Qualität eines MOOCs.

Mittels des Ratinginstruments wurden im Frühjahr 2023 N=41 zufällig ausgewählte Kurse des KI-Campus einem Design-Review unterzogen. Vier Studierende im Masterstudiengang Berufspädagogik und Personalentwicklung der Universität Stuttgart (m=0, w=4; M_{Alter}=24) führten die Analyse im Rahmen des Hauptseminars Instruktionsdesign nach intensiver Schulung durch. Vier der Kurse wurden dabei von allen Raterinnen analysiert, so dass im Rahmen einer konsensuellen Validierung ein gemeinsames Grundverständnis der Kategorien herbeigeführt werden konnte.

2.2 Ergebnisse

Tabelle 1 zeigt die deskriptiven Ergebnisse des explorativen Reviews. Es zeigt sich für den Bereich A, der die Klarheit und Strukturiertheit der Kurse in den Blick nimmt, über alle Kategorien gemittelt eine insgesamt überdurchschnittliche Ausprägung - in der Regel werden sowohl die Lernziele als auch die Anforderungen sowie die Kursinhalte und Kursstruktur gut bis sehr gut beschrieben. In Bereich B liegen die Ratings einiger Kategorien wie etwa Wissensaktivierung und Wissensintegration deutlicher unter dem theoretischen Skalenmittelwert, hier zeichnen sich Optimierungspotenziale ab. Der Bereich C liegt mit Ausnahme der Kategorie Authentizität der Materialien durchgängig unter dem theoretischen Mittelwert - insbesondere für die Kategorien Feedback und Differenzierung/Individualisierung sowie Kooperation und Kollaboration werden geringe Ausprägungen festgestellt.

	M	SD	Min	Max
Bereich A – Klarheit und Strukturiertheit				
1. Lernziele	1,93	,76	1	3
2. Zielgruppe	1,34	,99	0	3
3. Anforderungen	1,34	,79	0	3
4. Kursinhalte	2,00	,84	1	3
5. Kursstruktur	2,27	,63	1	3
Bereich B – Zentrale didaktische Kriterien				
6. Problemzentrierung	1,57	,84	0	3
7. Wissensaktivierung	0,99	,81	0	2
8. Demonstration	2,15	,65	0	3
9. Anwendung	1,44	,68	1	2
10. Wissensintegration	1,15	,50	0	3
Bereich C – Zusätzliche didaktische Kriterien				
11. Feedback	0,95	,44	0	2
12. Authentische Materialien	2,15	,62	1	3
13. Differenzierung / Individualisierung	0,68	,57	0	2
14. Kooperation / Kollaboration	0,80	,46	0	2
15. Lernendenaktivität	1,07	,61	0	2

Tab. 1: Kategorien und Deskriptiva im Überblick (ungewichtet)

Die zehn Kurse mit den höchsten Gesamtscores (nach Gewichtung) werden in Tabelle 2 dargestellt. Es handelt sich dabei um Kursangebote, die durch wissenschaftliche oder wissenschaftsnahe Institutionen verantwortet werden. Inhaltlich ist ein breites Spektrum vertreten, das von grundlegenden Einführungen in die KI über algorithmennahe Inhalte und maschinelles Lernen auch und vor allem konkrete Anwendungskontexte wie Schule oder Berufsbildung enthält. Auch hier zeigen sich in der Gesamtschau mittlere bis

überdurchschnittliche Ausprägungen der Kategorien A und B sowie eher unterdurchschnittliche Ausprägungen für den Bereich C.

Kurstitel	Institution	A	B	C	gesamt
Erklärbares Maschinelles Lernen für Ingenieurwissenschaften	Fraunhofer IIS	14	22	16	52
Data Literacy-Basiskurs Data Lifecycle	TH Köln	13	18	16	47
Building AI	U Helsinki	9	20	14	43
AI-VET 4: KI als Werkzeug in der beruflichen Bildung	U Stuttgart, U Mannheim	13	18	12	43
AI_VET 1: Einführung in die KI	U Mannheim, U Stuttgart	7	21	14	42
AutoML – Automated Machine Learning	U Hannover U Freiburg LMU Münch.	8	19	14	41
Daten- und Algorithmenethik	HWR Berlin	10	18	12	40
Anwendung von KI in der Produktion	U Stuttgart	11	17	12	40
Data2Teach - Wie Daten Schule u. Schulen Daten machen	PH Heidelberg TU Kaiserslautern	11	16	12	39
Building Visual Machine Learning Models	DHBW Stuttgart	10	17	12	39

Tab. 2: Kurse mit den höchsten Ratings

3 Diskussion und Ausblick

Die Ergebnisse des explorativen Design Reviews von offenen Online-Kursen zum Thema KI sind anschlussfähig an die bisherige Forschung zum Instruktionsdesign von MOOCs. Bei den „Top Ten“ der Kurse mit den höchsten Ratings zeigt sich, dass die Beurteilung

hinsichtlich didaktischer Gütekriterien weder von der Art der anbietenden Institution (Universitäten, HAWs, Duale Hochschulen sowie außeruniversitäre Forschungsinstitutionen sind einzeln oder in Verbänden vertreten), noch von der inhaltlichen Ausrichtung der Kurse (eher technisch oder eher anwendungsorientiert) abhängig zu sein scheint. Aus methodischer Sicht sind diese ersten Ergebnisse allerdings noch durch differenzierte Analysen zu vertiefen.

Mit Blick auf die Ausprägung der verschiedenen Design-Kriterien über alle Kurse zeigt sich zunächst ein erwartungskonformes Bild. Die überdurchschnittlichen Ausprägungen im Bereich A unterstreichen einmal mehr, dass MOOCs gut strukturierte, systematisch aufbereitete Lernangebote sind. Das trifft auch für die untersuchten Kurse des KI-Campus zu. Die teils unterdurchschnittlichen Ausprägungen in den Bereichen B und C könnten dagegen auf vorhandene Optimierungspotenziale hindeuten. Dabei ist allerdings zunächst einschränkend zu erwähnen, dass Kriterien wie Feedback, Differenzierung, Kooperation und Kollaboration oder auch Wissensaktivierung und –integration in vielen Fällen nicht in den Online-Angeboten selbst, sondern über deren Einbindung in hybride Lernsettings, bspw. innerhalb von akademischen Lehrveranstaltungen, angesprochen werden, die nicht Gegenstand dieses Design Reviews waren. Wenn die MOOCs allerdings als eigenständige Lehr-Lern-Arrangements betrachtet werden sollen, so zeigen sich gerade hier die Unterstützungspotenziale von KI-Technologie: Kooperation und Kollaboration könnte durch virtuelle Assistenten oder Chatbots unterstützt, Feedback automatisiert und adaptiv gestaltet werden. Differenzierung und Individualisierung ließe sich über Learning Analytics-Verfahren oder über adaptive, individuelle Lernpfade realisieren. Die Lernendenaktivität ließe sich – soweit inhaltlich passend – durch die Einbindung und aktive Nutzung von KI-Werkzeugen steigern. Genauere, systematische Analysen der Unterstützungspotenziale erscheinen daher notwendig. Die (weitere, tiefere, systematische) Einbindung von KI-basierten Lernwerkzeugen [CCL20] könnte folglich einen wichtigen Meilenstein auf dem Entwicklungspfad der Bildungsplattform darstellen. Für die Weiterentwicklung von digitalen Lernangeboten zum Thema KI erscheint eine Kopplung der Perspektiven „KI als Lerninhalt“ und „KI als Lernwerkzeug“ Erfolg versprechend.

Aus methodischer Sicht bleibt festzuhalten, dass die designorientierte Betrachtung nur einen – wenn auch wichtigen – Teilaspekt einer Qualitätsbetrachtung von MOOCs ausmacht [ST21]. Dabei erscheint es zudem notwendig, die wissenschaftliche, theorieorientierte Sicht des Ratings durch weitere Perspektiven zu erweitern. Die Sicht der Stakeholder*innen, und dabei insbesondere die der Lernenden, sollte ebenfalls in die Betrachtung einfließen, um ein differenziertes Urteil zur didaktischen Passung von MOOCs zu erlangen [OCC23]. Von daher erscheint es sinnvoll, das Design Review mit Daten aus der Teilnehmendenevaluation oder zusätzlich erhobenen Lernendendaten zu koppeln. Über eine systematische Befragung der Kursanbietenden könnten zudem die Spezifika der curricularen und didaktischen Einbindung der MOOCs in Aus- und Weiterbildungskontexte erhoben und systematisiert werden. Vor dem Hintergrund des Bedeutungswandels von MOOCs weg von lehrveranstaltungsäquivalenten akademischen

Kursen hin zu Weiterbildungsbausteinen erscheint eine solche Betrachtung für die Beurteilung der didaktischen Passung notwendig.

Literaturverzeichnis

- [BCP22] Borrella, I.; Caballero-Caballero, S.; Ponce-Cueto, E.: Taking action to reduce dropout in MOOCs: Tested interventions. *Computers & Education* 179, 104412, 2022.
- [CCL20] Chen, L.; Chen, P.; Lin, Z.: Artificial Intelligence in Education: A Review. *IEEE Access* 8, pp. 75264–75278, 2020.
- [Eg18] Egloffstein, M.: Massive open online courses in digital workplace learning: current state and future perspectives. In (Ifenthaler, D., ed.): *Digital workplace learning: bridging formal and informal learning with digital technologies*. Springer, Cham, pp. 149–166, 2018.
- [EKI19] Egloffstein, M.; Kögler, K.; Ifenthaler, D.: Instructional quality of business MOOCs: Indicators and initial findings. *Online Learning Journal* 23, pp. 85–105, 2019.
- [KI19] Kloos, C. D., et al.: Taxonomy of MOOC-Based Hybrid Educational Models in Higher Education. *IEEE Learning with MOOCs (LWMOOCs)*, Milwaukee, WI, pp. 128–132, 2019.
- [KVK21] Kasch, J.; Van Rosmalen, P.; Kalz, M.: Educational scalability in MOOCs: Analysing instructional designs to find best practices. *Computers & Education* 161, 104054, pp. 1–12, 2021.
- [MBL15] Margaryan, A.; Bianco, M.; & Littlejohn, A.: Instructional quality of massive open online courses (MOOCs). *Computers & Education* 80, pp. 77–83, 2015.
- [Me13] Merrill, M. D.: *First principles of instruction: Identifying and designing effective, efficient and engaging instruction*. Pfeiffer, Hoboken, 2013.
- [OCC23] Oh, E. G.; Cho, M. H.; Chang, Y.: Learners’ perspectives on MOOC design. *Distance Education*, 2023.
- [OCP20] Oh, E. G.; Chang, Y.; Park, S. W.: Design review of MOOCs: application of e-learning design principles. *Journal of Computing in Higher Education* 32, pp. 455–475, 2020.
- [ST21] Stracke, C. M.; Trisolini, G.: A Systematic Literature Review on the Quality of MOOCs. *Sustainability* 13, 5817, 2021.

Enaktive Bestimmung der Hyperparameter beim Entscheidungsbaum- und k-nächste-Nachbarn-Algorithmus

Vorstellung von Materialien eines KI-Experimentierkastens

Silvia Joachim ¹ und Martin Hennecke ²

Abstract: Für den Bereich der Künstlichen Intelligenz werden momentan zahlreiche Unterrichtsmaterialien veröffentlicht. Viele davon konzentrieren sich dabei auf den Teilbereich des Maschinellen Lernens. Die hier vorgestellten Unterrichtsmaterialien gehören ebenfalls in diese Kategorie – sie sprechen allerdings gezielt verschiedene Repräsentationsebenen an und sollen durch die haptische Zugänglichkeit auch Menschen mit einer Sehbeeinträchtigung den Zugang erleichtern. Bei den in diesem Beitrag vorgestellten Materialien handelt es sich um aus Holz gefertigte Pilze, die sich u.a. für die Erarbeitung des k-nächste-Nachbarn- und Entscheidungsbaum-Algorithmus eignen. Die Pilze verfügen dazu über binäre und metrische Merkmale. Der zugrunde liegende Datensatz ist bewusst klein und wurde so gewählt, dass der Einfluss von Hyperparametern gezeigt werden kann. Er enthält Validierungsdaten, die die Optimierung der Baumtiefe und der Anzahl der Nachbarn ermöglichen. Der Datensatz kann für weitere Algorithmen verwendet werden und ermöglicht somit eine effiziente Unterrichtsgestaltung.


Keywords: Unterrichtsmaterial, Experimentierkasten, Künstliche Intelligenz, k-nächste-Nachbarn-Algorithmus, Entscheidungsbaum-Algorithmus, Validierungsdaten, Hyperparameter, enaktiv

1 Einleitung

In der Öffentlichkeit wird das Thema Künstliche Intelligenz immer präsenter. Dies spiegelt sich auch in ersten Lehrplänen an allgemeinbildenden Schulen wider. So sieht beispielsweise der bayerische LehrplanPLUS für die Gymnasien in der Jahrgangsstufe 11 den Einsatz des k-nächste-Nachbarn- oder des Entscheidungsbaum-Algorithmus vor [ISB23a] [ISB23b]. Da das Fach Informatik in Bayern Pflichtfach ist muss daraus der Anspruch für eine inklusive Medienbildung abgeleitet werden, d.h. Unterrichtsmaterialien müssen z.B. auch für Schülerinnen und Schüler mit Sehbeeinträchtigung zugänglich sein.

Die Arbeit mit haptischen Unterrichtsmaterialien hat sich als Einstieg in neue Unterrichtsthemen in der Schulpraxis bewährt. Dieser Idee folgend stellt dieser Beitrag, real erfährt, begreif- und fühlbare Unterrichtsmaterialien in Form aus Holz gefertigter Pilze vor. Sie

¹ Universität Würzburg, Didaktik der Informatik, Emil-Fischer-Str. 30, 97074 Würzburg, silvia.joachim@uni-wuerzburg.de,  <https://orcid.org/0009-0004-8095-0133>

² Universität Würzburg, Didaktik d. Informatik, Emil-Fischer-Str. 30, 97074 Würzburg, martin.hennecke@uni-wuerzburg.de,  <https://orcid.org/0000-0001-5586-0647>

verfügen über binäre und metrische Merkmale und sind daher für verschiedene Algorithmen des Maschinellen Lernens nutzbar, insbesondere für den k-nächste-Nachbarn-Algorithmus oder das Maschinelle Lernen von Entscheidungsbäumen. Die Arbeit mit den Pilzen konfrontiert die Schülerinnen und Schüler mit dem Problem, Pilze in „giftig“ (X) oder „essbar“ (O) zu klassifizieren. Die Bedeutung der Fragestellung wird dabei durch die Bedeutung der Entscheidung betont und eröffnet für eine anschließende kritische Reflexion der Vorgehensweise zahlreiche Ansätze für die Diskussion.



Abb. 1: Pilze mit sicht- und fühlbaren Merkmalen (Bildquelle: MEKRUPHY GMBH)

Als binäre Merkmale der Pilze sind Muster (Punkte, Streifen), Farbe (blau, rot), Randlinie (gerade, gewellt) und Kragen (ja, nein) ausgeprägt. Die metrischen Merkmale x (0 bis 12) und y (0 bis 8) geben den Standort der Pilze auf einem Koordinatensystem an. 14 dieser Datensätze sind als „giftig“ oder „essbar“ gelabelt und zwei stehen zur Klassifizierung zur Verfügung. Die Symbole oben im Pilzkopf geben zudem eine Unterteilung in Trainings-, Validierungs- und Testdaten vor. Der durch die Pilze kodierte Datensatz ermöglicht eine Optimierung der jeweiligen Modelle, ist aber dennoch klein genug um die Algorithmen auch enaktiv durchführen und manuell berechnen zu können.

2 Entscheidungsbaum-Algorithmus

Für den Entscheidungsbaum-Algorithmus [Er21] gibt es bereits ansprechende Materialien. So können etwa die „beißenden“ oder „nicht beißenden“ Äffchen von Seegerer, Lindner und Romeike [SLR19] genutzt werden. Die Arbeit mit den Pilzen ist grundsätzlich ähnlich gestaltbar. Die binären Merkmale Muster, Farbe (bzw. Form), Randlinie und Kragen erleichtern jedoch die direkte Nutzbarkeit der Daten beim Erlernen von binären Entscheidungsbäumen. Die gegebene Unterteilung in Trainings-, Validierungs- und Testdaten ist so gewählt, dass im Anschluss eine Optimierung des Hyperparameters Baumtiefe sinnvoll möglich ist. Die Pilze sind dabei nicht (nur) alternativ zu den Äffchen nutzbar, sondern können nach einer stärker spielerischen Begegnung mit den Äffchen auch für eine weitere Formalisierung der algorithmischen Idee des Lernalgorithmus eingesetzt werden.

Die fühlbaren binären Merkmale der Pilze sollen auch sehbeeinträchtigten Menschen einen Zugang zum Entscheidungsbaum-Algorithmus ermöglichen. Eine Platte mit Aussparungen für Wurzel, Knoten und Kantenbeschriftungen des Entscheidungsbaums ermöglicht eine graphische Repräsentation des Entscheidungsbaums, die für Schülerinnen und Schüler mit Sehbeeinträchtigung zugänglich ist. Die Pilze können immer wieder neu nach den Merkmalen Muster, Farbe (bzw. Form), Randlinie und Kragen sortiert werden, um das Merkmal mit dem größten Informationsgewinn zu ermitteln. Das anschauliche Ziel dabei ist, möglichst gut zwischen „essbaren“ und „giftigen“ Pilzen unterscheiden zu können.

Durch das im Folgenden beschriebene Zählen von Pilzen findet eine didaktische Reduktion beim Maschinellen Lernen eines Entscheidungsbaums statt. Die Schülerinnen und Schüler können anhand der acht Trainingsdaten (Pilze mit Dreieck) zunächst feststellen, dass es unter ihnen genauso viele „giftige“ wie „essbare“ Pilze gibt. Werden die Pilze nach dem Merkmal Muster sortiert, so sind sowohl unter den Pilzen mit Punkten als auch unter denen mit Streifen wieder genauso viele „essbare“ wie „giftige“ Pilze. Gleiches gilt für eine Sortierung nach den Merkmalen Farbe und Kragen. Bei einer Sortierung der Trainingsdaten nach dem Merkmal Randlinie können die Schülerinnen und Schüler feststellen, dass in der Gruppe mit gewellter Randlinie drei Pilze „essbar“ und zwei „giftig“ und in der Gruppe mit gerader Randlinie einer „essbar“ und zwei „giftig“ sind. Als Wurzel des Entscheidungsbaumes erhalten sie so das Merkmal Randlinie. Pilze mit gewellter Randlinie würden per Mehrheitsentscheidung als „essbar“ und Pilze mit gerader Randlinie als „giftig“ klassifiziert werden und ergäben so einen Entscheidungsbaum der Tiefe 1. Nach dem Divide-and-Conquer-Ansatz werden die Pilze danach entsprechend der Randlinie getrennt betrachtet (usw.).

Ein beim Entscheidungsbaum interessanter Hyperparameter ist dessen Tiefe. Bei großen Datensätzen kann Pruning vorteilhaft sein. Mit Hilfe der Validierungsdaten kann, wie bei Russel/Norvig [RN22] für einen großen Datensatz beschrieben, die optimale Baumtiefe gewählt werden, indem die minimale Fehlerrate für die Validierungsmenge bestimmt wird. Um dies nachvollziehbar zu machen, wurde der zugrunde liegende Datensatz so gestaltet, dass die Vorteile von Pruning nachvollzogen werden können.

3 k-nächste-Nachbarn-Algorithmus

Der k-nächste-Nachbarn-Algorithmus lässt sich mit den metrischen Merkmalen umsetzen. Diese geben die Koordinaten der Fundorte der Pilze an und können in ein Koordinatensystem eingetragen werden. Die Klassifikation anhand der Koordinaten folgt der Idee, dass Pilze einer Art häufig in kleinen Gruppen beieinanderstehen. Beim Eintragen der Pilze lassen sich die giftigen und die essbaren Pilze z. B. durch die Symbole X und O visualisieren. Schülerinnen und Schüler mit einer Sehbeeinträchtigung können die Fundorte der Pilze räumlich auf einem Koordinatensystem ertasten. Entfernungen sind mit der Manhattan-Metrik erfassbar.

Anschließend kann nacheinander der Fundort der drei Validierungspilze betrachtet und der k-nächste-Nachbarn-Algorithmus (k ungerade) mit euklidischer Norm angewendet werden. D.h. es werden die Abstände von jeweils einem Validierungspunkt zu allen Trainingsdatenpunkten mit dem Lineal gemessen und die k-nächsten Nachbarn bestimmt. Dabei wird der Parameter k bewusst ungerade gewählt, so dass hier immer eine Mehrheitsentscheidung möglich ist. Da es 8 Trainingsdaten gibt, wird $k=1$, $k=3$, $k=5$ und $k=7$ betrachtet. Jeder Validierungspilz wird dann, entsprechend der Mehrheit der k-nächsten Nachbarn, klassifiziert und das Ergebnis mit seinem Label verglichen. Für $k=5$ werden alle Validierungs- und Testdaten richtig klassifiziert. Von den beiden unbekannt Pilzen wird einer als „essbar“ und einer als „giftig“ klassifiziert. Hier liefert der k-nächste-Nachbar-Algorithmus didaktisch sinnvoll ein anderes Ergebnis als der Entscheidungsbaum-Algorithmus mit binären Merkmalen; schließlich wurden ja sogar andere Merkmale betrachtet. Die Trefferquote ist maximal und bietet eine Möglichkeit zur Diskussion über das Vertrauen in KI-Systeme.

4 Fazit

Der durch Speise- bzw. Giftpilze gegebene Realitätsbezug macht unmittelbar klar, wie wichtig eine korrekte Klassifizierung ist. Die sich durch die unterschiedlichen Algorithmen ergebenden unterschiedlichen Klassifizierungen der unbekannt Pilzsorten bieten eine Möglichkeit zur Diskussion über die Qualität von Daten und die Vertrauenswürdigkeit von KI-Systemen, wie etwa einer Smartphone-App zur Pilzbestimmung. Die hier vorgestellten Pilze sind Teil eines Experimentierkastens, der in zahlreichen Lehrerfortbildungen getestet wurde und dort auf sehr positive Resonanz gestoßen ist.

Literaturverzeichnis

- [Er21] Ertel, W.: Grundkurs Künstliche Intelligenz. Eine praxisorientierte Einführung, 5. Auflage, Springer Fachmedien Wiesbaden, 2021.
- [ISB23a] Staatsministerium für Schulqualität und Bildungsforschung München, LehrplanPLUS Gymnasium Bayern, Informatik 11 (NTG), <https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/11/informatik/ntg>, Stand: 14.05.2023.
- [ISB23b] Staatsministerium für Schulqualität und Bildungsforschung München, LehrplanPLUS Gymnasium Bayern, spätbeginnende Informatik 11, https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/11/informatik/mug_swg_sg, Stand: 14.05.2023.
- [RN22] Russel, S.; Norvig, P.: Artificial Intelligence. A Modern Approach, 4. Auflage, Pearson, United Kingdom, 2022.
- [SLR19] Seegerer, S.; Lindner, A; Romeike, R.: AI Unplugged – Wir ziehen Künstlicher Intelligenz den Stecker. In (Pasternak, A., Hrsg.): Informatik für alle. Bonn: Gesellschaft für Informatik. S. 325-334. DOI: 10.18420/infos2019-c18

Einstieg in den ML-Workflow durch Entwicklung einer intelligenten Museumsapp

Erik Marx^{1,2} und Nadine Bergner³

Abstract: In diesem Praxisbeitrag wird der Workshop „Die intelligente Museumsapp“ vorgestellt, mit dem die zentralen Konzepte des Maschinellen Lernens (ML) für die Sekundarstufe I vermittelt werden können. Ausgehend vom ML-Workflow werden die Ziele des Workshops sowie der Aufbau und die verwendeten Lehr-Lern-Materialien vorgestellt und beschrieben. Im Weiteren werden erste Pilotierungen reflektiert.

Keywords: Künstliche Intelligenz, Maschinelles Lernen, Lehr-Lern-Material, Workshop

1 Einleitung - Maschinelles Lernen in der Sekundarstufe I

Schüler:innen interagieren im Alltag, teils unbewusst, mit künstlicher Intelligenz (KI), welche insbesondere auf der Technologie des *maschinellen Lernens (ML)* beruht. Um die Möglichkeiten und Herausforderungen dieser neuen Technologie einschätzen zu können, müssen Schüler:innen KI nicht nur anwenden können, sondern auch deren Funktions- und Entstehungsweise verstehen. Entsprechend wird ML auch in diversen KI-Kompetenzkatalogen [LM20; TGS22] thematisiert und auch in den Informatiklehrplänen mehrerer Bundesländer explizit genannt: u. a. in Sachsen in Klasse 9 [Sä22] und in Nordrhein-Westfalen in Klasse 6 [Mi21]. Im Folgenden wird eine Möglichkeit vorgestellt, wie der Themenkomplex ML in der Sekundarstufe I schüler:innenzentriert behandelt werden kann.

2 Das Workshopkonzept der intelligenten Museumsapp

Kontext des Workshops ist die Entwicklung einer „intelligenten Museumsapp“, mit der Bilder von Exponaten gemacht werden können, um Erklärungen zu erhalten. Dabei werden die Kernschritte (1) *Daten aggregieren*, (2) *Modell testen* und (3) *Modell anwenden* des *ML-Workflows* (siehe Abb. 1) des ML-Prozesses vermittelt und von den Schüler:innen angewendet, wodurch die Rolle der Trainingsdaten und der Einfluss des Menschen auf den ML-Prozess verdeutlicht werden [LM20; TGS22].

¹ TU Dresden, Professur für Didaktik der Informatik, Nöthnitzer Str. 46, 01187 Dresden, erik.marx@tu-dresden.de, <https://orcid.org/0000-0002-5918-804X>

² Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) Dresden/Leipzig, Germany

³ RWTH Aachen University, Professur für Didaktik der Informatik, Tempelgraben 55, 52062 Aachen, bergner@informatik.rwth-aachen.de, <https://orcid.org/0000-0003-3527-3204>

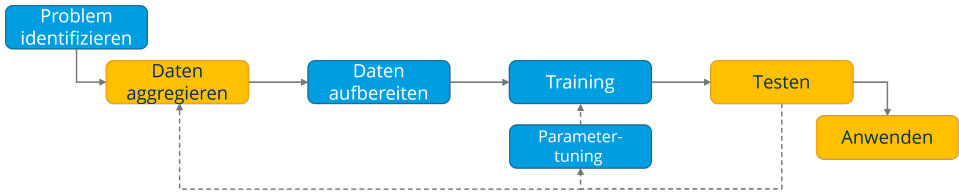


Abb. 1: Prototypischer ML-Workflow nach [Zi18] mit drei Kernschritten (gelb markiert)

Die Lernziele für den Workshop rund um die Museumsapp lauten: Die Schüler:innen...

1. ...vollziehen den reduzierten prototypischen ML-Workflow nach, indem sie ihr eigenes ML-Modell erstellen und iterativ verbessern.
2. ...sind sich dem Einfluss des Menschen im ML-Prozess bewusst, indem sie geeignete Trainingsdaten auswählen und das Modell testen.
3. ...differenzieren zwischen dem Erstellen/Trainieren und Nutzen/Anwenden eines ML-Modells, indem sie die jeweiligen Schritte beschreiben.

Um diese Lernziele zu erreichen, werden die von Touretzky et al. in [TGS22] empfohlenen Werkzeuge „Teachable Machine (TM)“ [Te23] und „App Inventor (AI)“ [Ap23] genutzt. Durch die Verzahnung dieser Tools können die Problemlösestrategien des ML, die sich von denen traditioneller Softwareentwicklung unterscheiden [TDT21], handlungsorientiert nachvollzogen werden. Dadurch wird TM neben dem iterativen Experimentieren [TGS22] auch zur systematischen Problemlösung eingesetzt. Ein weiterer Vorteil des Konzepts ist, dass es modular ist und daher auch im Informatikunterricht eingesetzt werden kann. Durch die leichte Einbindung in den AI können Anwendungsbeispiele geändert werden, z. B. von der Museums-App zu einem Yoga-Trainer mithilfe der Posenerkennung von TM. Auch kann das Konzept erweitert werden, indem z. B. mit vorbereiteten Datensätzen spezifische Konzepte wie Bias thematisiert werden. Zusätzlich kann der ML-Workflow im AI auch mit klassischer Programmierung verknüpft werden. Im Folgenden werden die drei Phasen des Workshops detaillierter vorgestellt. Alle Lehr-Lern-Materialien wurden zusätzlich als Open Educational Resources veröffentlicht [OE23].

2.1 Einführungsaufgabe

Um das Tool TM und den allgemeinen ML-Workflow kennenzulernen, trainieren die Schüler:innen ein Modell, welches ihre Gesichter unterscheiden kann⁴. Die Schüler:innen machen sich mit der Oberfläche von TM vertraut und erstellen Kategorien. Dann nehmen sie Bilder auf, trainieren das Modell und testen es. Anschließend wird der ML-Workflow in

⁴ Alle Bilder werden dabei lokal im Browser verarbeitet und nicht an Google-Server geschickt.

einem Unterrichtsgespräch gesichert, indem die drei Kernschritte direkt auf die Anwendung TM übertragen werden.

2.2 Erstellen des Modells für die Museumsapp

Zuerst wird diskutiert, worauf zu achten ist, damit das Modell möglichst gut funktioniert. Dazu wurde in der Einführungsaufgabe die Bedeutung der Datenauswahl thematisiert. Das Projekt selbst bietet hier die Möglichkeit, eine Variation der Daten zu motivieren. Anschließend wird das Modell zur Erkennung der Exponate trainiert. Es können beliebige Gegenstände als Exponate genutzt werden. Damit den Schüler:innen die Bedeutsamkeit der Datenaggregation und der Evaluation des Modells bewusst wird, sollten die Exponate so gewählt werden, dass eine grobe Unterscheidung der Exponate mittels TM leicht möglich ist, jedoch ein robusteres Modell, welches die Variation von Winkel, Hintergrund etc. einschließt, fehleranfällig bleibt. Je nach Lerngruppe können die Schüler:innen freier bzw. angeleiteter am Projekt arbeiten. Durch die Arbeit am Modell wird der ML-Workflow nicht nur nachvollzogen, sondern selbst (mehrfach) durchlaufen. Als Strukturierungshilfe kann die Lehrperson vorgeben, dass das Modell zuerst zwei Exponate unterscheiden soll und anschließend erweitert wird. Je nach verfügbarer Zeit kann diese Phase durch Zwischendiskussionen intensiviert werden.

2.3 Integration des Modells in die App & Diskussion

Im dritten Schritt integrieren die Lernenden ihr Modell in eine (vorgefertigte oder selbstentwickelte) App. Dafür stehen Lehr-Lern-Materialien in mehreren Schwierigkeitsstufen zur Verfügung. Durch eine Erweiterung des AI [TM23] kann das Modell mittels blockbasierter Programmierung weiterverarbeitet werden. Anschließend können die Schüler:innen die App mit mobilen Geräten testen. Hier steht die Unterscheidung zwischen dem Erstellen und Nutzen eines ML-Modells im Fokus. Anhand von Beispielen wie dem autonomen Fahren oder Empfehlungssystemen kann eine Diskussion über die gesellschaftlichen Auswirkungen von KI motiviert werden. Ein wichtiger Aspekt ist dabei, dass bei der Nutzung von KI-Systemen häufig die Unsicherheiten, mit der ein ML-Modell eine Entscheidung trifft, nicht sichtbar werden.

Am Ende des vorgestellten Workshops haben die Schüler:innen handlungsorientiert einen reduzierten Ablauf des ML-Workflow kennengelernt, indem sie ein eigenes ML-Modell trainiert und iterativ verbessert haben. Damit soll den Lernenden die Bedeutung der Trainingsdaten und somit der Einfluss des Menschen sowie die Relevanz des Testens und iterativen Verbesserns bewusst werden.

3 Erfahrung aus Pilotierungen & Ausblick

Der Workshop „Die intelligente Museumsapp“ wurde in verschiedenen außerschulischen Workshops über 1,5 - 3 Stunden mit Schüler:innen der Klassenstufen 5 bis 9 erprobt, die das Thema ML noch nicht im Unterricht behandelt hatten. Bereits in der Einführungsaufgabe diskutierten die Schüler:innen selbstständig mögliche Einflusseffekte der Trainingsdaten auf das Ergebnis des ML-Modells und konnten im anschließenden Unterrichtsgespräch verschiedene Gründe für Probleme der Bildererkennung (wie Winkel, Farbe etc.) nennen. Jüngere Schüler:innen tendierten dazu, länger mit dem Tool zu experimentieren, wohingegen Ältere häufig direkt versuchten, die Grenzen des Modells auszuloten. Älteren Schüler:innen war schnell bewusst, worin die Schwierigkeit bei der Datenauswahl liegt. Sie gingen systematisch vor und versuchten die Daten entsprechend den vermuteten Einflussfaktoren zu variieren. Die weiteren Durchführungen des Workshops mit zwei 11. Klassen in 90 Minuten zeigten, dass sich die Museumsapp auch anbietet, um in das Thema ML in der Sekundarstufe II einzusteigen. Die Schüler:innen zeigten insbesondere in den Diskussionen, dass sie bereits nach kurzer Zeit die zentralen Konzepte des ML-Workflow einordnen konnten. In Zukunft wird der Workshop auch im schulischen Kontext erprobt und mittels Lehrkräftefortbildungen weiterverbreitet. Auch eine formale Evaluation hinsichtlich der angestrebten Lernziele ist geplant.

Literatur

- [Ap23] App Inventor, 2023, URL: <https://appinventor.mit.edu/>, Stand: 27.08.2023.
- [LM20] Long, D.; Magerko, B.: What Is AI Literacy? Competencies and Design Considerations. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM, Honolulu HI USA, S. 1–16, 2020.
- [Mi21] Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, Hrsg.: Kernlehrplan Für Die Sekundarstufe I - Klasse 5 Und 6 in Nordrhein-Westfalen, 2021.
- [OE23] OER-Materialien, 2023, URL: <https://tu-dresden.de/inf/eduinf/ki-materialien>, Stand: 27.08.2023.
- [Sä22] Sächsisches Staatsministerium für Kultus, Hrsg.: Lehrplan Gymnasium Informatik, 2022.
- [TDT21] Tedre, M.; Denning, P.; Toivonen, T.: CT 2.0. In: 21st Koli Calling International Conference on Computing Education Research. ACM, Joensuu Finland, S. 1–8, 2021.
- [Te23] Teachable Machine, 2023, URL: <https://teachablemachine.withgoogle.com/>, Stand: 27.08.2023.
- [TGS22] Touretzky, D.; Gardner-McCune, C.; Seehorn, D.: Machine Learning and the Five Big Ideas in AI. International Journal of Artificial Intelligence in Education 33/22, S. 233–266, 2022, ISSN: 1560-4306.

- [TM23] TMIC, 2023, URL: <https://community.appinventor.mit.edu/t/tmic-app-inventor-extension-for-the-deployment-of-image-classification-models-exported-from-teachable-machine/64411>, Stand: 27.08.2023.
- [Zi18] Zimmerman, M.: Teaching AI: Exploring New Frontiers for Learning. International Society for Technology in Education, Portland, Oregon, 2018, ISBN: 978-1-56484-705-8.

Better ready than just aware: Data and AI Literacy as an enabler for informed decision making in the data age

Katharina Schüller ¹, Florian Rampelt ², Henning Koch ³, and
Johannes Schleiss ⁴


Abstract: Data and AI literacy is an important enabler for informed decision making in the data age. To inform educational programs and policies, it is important to create a common understanding about the required knowledge and skills. In this paper, we propose a novel taxonomy to data and AI literacy based on qualitative literature analysis and expert group discussions. We introduce three key roles related to Data and AI Literacy: the informed prosumer, the skilled user, and the expert creator. Moreover, we argue that Data and AI Awareness as the lowest level of understanding and recognizing is a necessary prerequisite but not a sufficient condition to Data and AI Literacy. We rather equate Data and AI Literacy with Data and AI Readiness. Further work will focus on defining the core knowledge, skills and competences of the taxonomy.


Keywords: Data Literacy, AI Literacy, AI Skills, AI Readiness, Competences, Taxonomy


1 Motivation


Data has long determined our everyday lives. Whether it's a search on Google, navigation in the car, suggestions in the online shop or the credit rating at the bank. Far too often, we are unable to make informed decisions for or against data, data sharing and data use. In some cases, we are neither aware nor prepared to deal with data and digital technologies in everyday life and at work. The current developments around generative AI increase the challenges (but also opportunities) that go hand in hand with this.

This realisation emphasises the great need for data literacy and AI literacy for everyone. But what does that mean? Data and AI literacy is needed on various levels, depending on specific roles of a society's members, to weigh up the opportunities and risks of data and data analyses in a sensible way and to be able to make informed decisions in private and in professional situations. One needs to be aware and, even more, one needs to be ready for the data age.

¹ STAT-Up, Germany, katharina.schueller@stat-up.com,  <https://orcid.org/0000-0003-4957-5115>

² KI Campus / Stifterverband, Germany, florian.rampelt@stifterverband.de,  <https://orcid.org/0000-0003-3604-813X>

³ Stifterverband, Germany henning.koch@stifterverband.de,  <https://orcid.org/0000-0003-2898-5692>

⁴ Otto-von-Guericke-Universität Magdeburg, Germany, johannes.schleiss@ovgu.de,  <https://orcid.org/0009-0006-3967-0492>

Each and every one of us, whether the highest decision-maker at the national or even international level or simply as a private individual, is continuously exposed to a flood of data. Data-informed and algorithm-supported decisions have to be made in everyday life.

In the 1992 United Nations Agenda 21, one reads, "In sustainable development, each individual is a user and provider of information in the broadest possible sense." Data and AI literacy is to be understood in an equally broad sense. It includes not only a tool set or skill set, but above all a mindset: to critically question the quality of data and information sources; to be able to separate facts from opinions; to recognise the limits of data, including the capabilities and limits of AI applications and algorithms. And to make the right decision based on this recognition.

Our contribution intends to create a common understanding of data and AI literacy which aims to support appropriate and practical education programs along the entire education chain. Education that enables informed action and decision-making in dealing with data and artificial intelligence.

Given that, our data and AI literacy concept is the basis for eventually developing a convergent, coherent literacy assessment, measurement and impact evaluation framework for the national, regional and global levels that can help to fill the empirical evidence gap on outcomes from data and AI literacy interventions and programs.

1.1 Conceptual and theoretical foundations

The work of our research group is based on years of engagement with data literacy and increasingly also with AI literacy and AI skills. A central result of an initial, predominantly German-speaking community on data literacy was a Data Literacy Charter, which was published in 2021 [Sc21]. In this charter, data literacy is defined, following [Ri15], as "the ability to collect, manage, evaluate and apply data in a critical manner" [Sc21, Ri15]. A definition for AI literacy that has been widely cited in recent years has been developed by [LM20]: "AI literacy is a set of competencies [*sic*] that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace."

Combining these conceptional considerations we argue that data literacy and AI literacy cannot be separated from each other as data serves as the fuel for AI. Developing data and AI literacy is therefore crucial in today's technological landscape to harness the power of data and AI responsibly, enabling individuals to ask relevant questions, navigate through complex algorithms, interpret data- and AI-driven outputs, and critically evaluate their implications; generally speaking, to be able to make informed decisions in the data age.

To further specify this idea, amongst others the IEEE Standards Association has attempted a comprehensive definition in one of its standards, which is a central basis for many of the aspects to be discussed later: "Data and AI Literacy is the ability to generate, process, analyze, present meaningful information from data and develop, use, and apply artificial

intelligence (AI) and related algorithmic tools and strategies in order to guide informed, optimized, and contextually relevant decision-making processes.” (IEEE 3527.1™ Standard).

1.2 Related Work

Related competence frameworks are partly more specific, partly more general. Examples include the European Unions’ Digital Competence Framework for Citizen (DigComp) [Vu22], the HFD Data Literacy Framework [Sc19], the IEEE 3527.1™ Standard on Digital Literacy, Digital Skills, and Digital Readiness, the basic data literacy framework [Ai21] as well as work focusing on competences for AI literacy [LM20, Ng21]. Laupichler et al. [La22] recently scoped the literature specifically for AI literacy and found the topic is still in an early stage in higher and adult education. Multiple works build on defining roles or groups of stakeholders [Ai21, Wi18, Fa21] or using steps of action through a process in order to define competences [Sc19, Ai21]. At the same time, there is currently no clearly defined set of skills and competences over multiple frameworks and we are only beginning to understand the impact of data and AI literacy to society [La22].

2 Methodology

Methodology Literature research, qualitative content analysis [Ma04] and group discussions [Bo10] were the main methods used for this study. The methodology connects this study to the authors' involvement in the IEEE Working Group on “Data and AI Literacy, Skills and Readiness”. In this working group international data and AI literacy experts meet regularly to develop a new IEEE standard. Thus, the results of the literature research and the qualitative content analysis could be presented for discussion in this working group. This mix of methods is suitable for gaining new insights into the meaning of definitions, roles and competences in the area of data and AI literacy.

Sample and Data Collection In the course of the literature review, more than 90 competency frameworks and curricula related to data literacy and AI literacy were identified worldwide. This semi-systematic review leads to an overview of the global activities in this field and is the starting point for the qualitative content analysis. The competency frameworks were then subjected to a qualitative content analysis. Particular attention was paid to the description of processes, roles and the specific descriptions of the data and AI competences. In addition, the described competences were assigned to different competence levels. The results of these qualitative content analyses were subsequently presented for discussion in the focus group of IEEE experts and thus assessed and validated. The data collection thus followed a mixed methods approach and is characterised by different perspectives on the object of study.

Validity and Reliability To ensure validity of this study's results, various scientific techniques were used. For example, qualitative content analysis was conducted by

independent coders to verify classifications between individuals. Additionally, feedback was sought from individual IEEE Working Group members to accurately represent their expertise and opinions.

3 Data and AI Literacy: From Awareness to Readiness

The discussions with education experts have shown that there are indeed some general bases for addressing knowledge, skills and competences. For example, Bloom's Taxonomy or a version supplemented by Anderson and Krathwohl served many experts as a central reference for the question of how knowledge, skills and competences can be transferred into learning outcomes. A simplified representation according to Rampelt et al. [Ra22] defines the descriptors "reflect" and "understand" at the knowledge level and the descriptors "apply", "analyse", "assess" and "create" at the skills level (or application and problem-solving) (see Tab. 1). These form a good basis for the question of what is meant by AI literacy in relation to the concept of AI Readiness.

Dimension	Level	Description
Knowledge	1	Remember
	2	Understand
Application and Problem Solving	3	Apply
	4	Analyse
	5	Assess / Evaluate
	6	Create

Tab. 1: A taxonomy of learning [Ra22] based on Bloom's Taxonomy

We further developed the basic taxonomy of [Ra22] and applied it to the principles and definitions from the Data Literacy Charta [Sc21]. In this context, we derived a new taxonomy to describe the concepts of data and AI Awareness and Readiness (see Fig. 1). It proposes seven literacy levels in the field of data and AI, namely (1) Understand and Recognise, (2) Consume, (3) Provide, (4) Collect, (5) Manage and Protect, (6) Evaluate and Analyse, and (7) Create and Develop.

We argue that the first level, "understand and recognise", corresponds to the characteristic of general awareness. Data and AI awareness refers to having a basic understanding and knowledge about data and AI, its characteristics, and its potential applications. Levels 2-7, according to our working hypothesis, are characterized by Data and AI Readiness. This is what we understand as the core of Data and AI Literacy. Data readiness describes the ability to apply such (basic) understanding in an informed, responsible, and outcome-oriented manner in private and professional practice. AI Awareness is, therefore, a necessary prerequisite for AI Literacy, but not a sufficient condition.

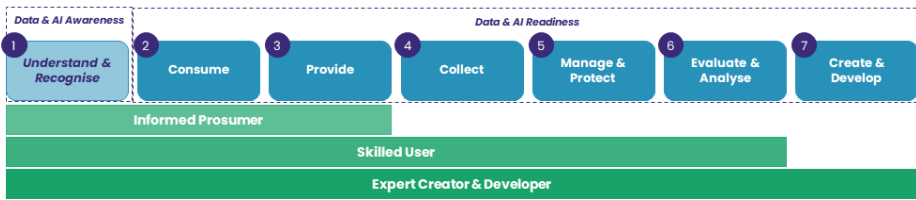


Fig. 1: Data and AI literacy levels and corresponding roles.

To describe data and AI readiness and the respective competences, it has shown useful to define exemplary roles as well as to match competences to these roles (see Fig. 1). The underlying idea is that competences are individual enablers and their respective levels depend on specific roles that are expected to fulfil certain tasks. In this context, each level might be present in different complexity depending on the role.

We identified three main roles that we describe as follows:

Informed Prosumer: A person who produces and consumes data and AI in an informed way. Thus, the informed prosumer is aware of and recognises the importance of data and AI in the world and takes a conscious decision to provide and consume data and AI tools.

Skilled User: A person who uses data and AI in a skilled and responsible way, including application-relevant skills that go beyond consumption and sharing. These include collecting, managing and protecting data, as well as the ability to evaluate and analyse data.

Expert Creator: A person who creates new insights, solutions, and tools with and based on data and AI. In this context, the creation and development go beyond sole application and usage of data and AI.

4 Discussion and Outlook

In conclusion, we have presented a novel taxonomy for data and AI literacy that is based on a combination of qualitative literature analysis and expert group discussions. This taxonomy introduces seven literacy levels that contribute to informed decision-making processes. In this context, literacy includes not only knowledge of data and AI but also attitudes and values towards it. Moreover, the taxonomy defines three major roles or groups, namely informed prosumers, skilled users, and expert creators, which help to distinguish and describe levels of proficiency in data and AI literacy and to further elaborate on its application in practice. We have also highlighted the importance of moving beyond data and AI awareness as the foundation of understanding. While data and AI awareness is necessary, it is not sufficient to achieve true literacy and data and AI readiness.

Overall, this work creates a common understanding of data and AI readiness and can support work on defining policies, competence frameworks, education curricula and the preparation of the digital and data-driven society. It can support policy makers, curricula designers in institutions and companies, educators, and individuals. Specifically, this work will contribute to the development of an official IEEE standard on Data and AI Literacy, Skills, and Readiness to be finalised this year. Further research will focus on providing competence indicators per role, to further elaborate on the actual roles in regards to their application in practice and to provide a basis to assessment and impact evaluation of interventions and programs.

5 References

- [Ai21] Aiken, P.; Harbour, T.: *Data Literacy: Achieving Higher Productivity for Citizens, Knowledge Workers, and Organizations*, Technics Publications, 2021.
- [Bo10] Bohnsack, R.: *Documentary method and group discussions*. 2010.
- [Fa21] Faruqe, F.; Watkins, R.; Medsker, L.: *Competency Model Approach to AI Literacy: Research-based Path from Initial Framework to Model*. arXiv:2108.05809. 2021.
- [La22] Laupichler, M. C., et al.: *Artificial intelligence literacy in higher and adult education: A scoping literature review*. In *Computers and Education: Artificial Intelligence*, 100101, 2022.
- [LM20] Long, D.; Magerko, B.: *What is AI literacy? competences and design considerations*. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, 2020.
- [Ma04] Mayring, P.: *Qualitative content analysis. A companion to qualitative research 1.2*: 159-176, 2004.
- [Ng21] Ng, D. T. K., et al.: *Conceptualizing AI literacy: An exploratory review*. *Computers and Education: Artificial Intelligence 2*: 10004, 2021.
- [Ra22] Rampelt, F.; Bernd, M.; & Mah, D.-K.: *Wissen, Kompetenzen und Qualifikationen zu Künstlicher Intelligenz. Eine Systematisierung von digitalen Formaten am Beispiel des KI-Campus und seiner Partner*. Berlin: KI-Campus. 2022.
- [Ri15] Ridsdale, C., et al.: *Strategies and best practices for data literacy education: Knowledge synthesis report*, 2015.
- [Sc19] Schüller, K.: *Future skills: a framework for data literacy*. In: *Hochschulforum Digitalisierung 46*, 1-128, 2021.
- [Sc21] Schüller, K., Koch, H.; Rampelt, F.: *Data Literacy Charta*, Stifterverband, 2021.
- [Vu22] Vuorikari, R., Kluzer, S. & Punie, Y., *DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes*, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022.
- [Wi18] Wienrich, C., et al.: *AI literacy: Kompetenzdimensionen und Einflussfaktoren im Kontext von Arbeit*. *Economics*, 12(1), 2018.

KI-Kompetenzen in ingenieurwissenschaftlichen Studiengängen entwickeln und fördern

Ein Praxisbeispiel für aktivierende Lehr-Lern-Angebote mittels Lerntagebuch

David Weigert¹, Ulrike Scorna² und Fabian Behrendt³

Abstract: Alle Akteure in der Hochschullandschaft auf die Herausforderungen und Chancen der KI vorzubereiten und zu begleiten ist ein wesentlicher Beitrag für zukünftiges Lehren, Lernen und Forschen. Die Autoren der Hochschule Magdeburg-Stendal zeigen in diesem Artikel auf, wie KI-Kompetenzen in ingenieurwissenschaftlichen Studiengängen durch den Einsatz eines Lerntagebuches entwickelt und gefördert werden können. Durch das Modul „Einführung in die KI – Grundlagen und Anwendungsfelder“ ist ein spezielles Kursdesign geschaffen worden, welches als wertvolles Lehr-Lernangebot der Studentinnen und Studenten evaluiert wurde.

Keywords: Lehr-Lernangebote KI, Lerntagebuch, ingenieurwissenschaftliche Studiengänge, ZAKKI

1 KI in der Hochschulbildung - Motivation und Problemstellung

Die Bedeutung von Künstlicher Intelligenz (KI) im Allgemeinen und als Ausbildungsinhalt an Hochschulen hat in den letzten Jahren erheblich zugenommen. Mit der wachsenden Verfügbarkeit von Daten, Rechenleistung und Algorithmen hat KI das Potenzial viele Aspekte unseres täglichen Lebens und unserer Gesellschaft zu verändern. Studentinnen und Studenten denen fundierte Kenntnisse über KI zur Verfügung stehen, sollten somit in der Lage sein, zukünftigen Anforderungen mit innovativen Lösungen zu begegnen. Dass es ein hohes Interesse von Studentinnen und Studenten an dem Thema Künstliche Intelligenz gibt, belegt eine hochschulweite Befragung zu den zukünftigen Bedarfen an der Hochschule Magdeburg-Stendal aus dem Jahr 2022. Allerdings wurde in der Befragung auch deutlich, dass es derzeit noch an KI-Lehr-Lern-Angeboten mangelt [Ho22]. Das belegt auch der VDI-Statusreport 2022, wonach von den 300 Befragten 32% angaben, keine KI-Methoden zu kennen und nur 11% angaben, KI-Methodenwissen im Grundstudium vermittelt bekommen zu haben[Ve22].

¹ Institut für Technische BWL/Hochschule Magdeburg-Stendal, Osterburger Str. 25, 39576 Hansestadt Stendal, david.weigert@h2.de

² Gesundheits- und Sozialwissenschaften/Hochschule Magdeburg-Stendal, Breitscheidstr. 2, 39114 Magdeburg, ulrike.scorna@h2.de

³ Institut für Technische BWL/Hochschule Magdeburg-Stendal, Osterburger Str. 25, 39576 Hansestadt Stendal, fabian.behrendt@h2.de

Ein weiteres Hemmnis für die Entwicklung von KI-Lehr-Lernangeboten ist das Vorhandensein von isoliertem Expertenwissen zu themenspezifischen KI-Inhalten. Diese „Wissensilos“ fragmentieren das KI-Wissen innerhalb der Hochschule Magdeburg-Stendal und führen darüber hinaus zu unsystematischer Wissensvermittlung. Um diesen Mangel entgegenzuwirken, wurde an der Hochschule Magdeburg-Stendal das Projekt „ZAKKI – zentrale Anlaufstelle für innovatives Lehren und Lernen interdisziplinärer Kompetenzen der KI“ initiiert, um didaktisch-innovative KI-Lehr-Lern-Angebote zu entwickeln und KI-Kompetenzen, wie Datenkompetenz, stochastisches und statistisches Denken, algorithmisches Denken und Programmieren, kritisches Denken und Menschsein im Kontext von KI [De19], bereits im Studium zu vermitteln.

2 Kursdesign, Aufbau und Zielstellung

Die Lehrveranstaltung „Einführung in die KI – Grundlagen und Anwendungsfelder“ ist als Wahlpflichtfach des Fachbereichs Ingenieurwissenschaften und Industriedesign (IWID) für die Bachelorstudiengänge Wirtschaftsingenieurwesen, Mensch-Technik-Interaktion, Maschinenbau mit den Spezialisierung Produktion und Elektrotechnik des 5. und 7. Fachsemesters der Hochschule Magdeburg-Stendal konzipiert.

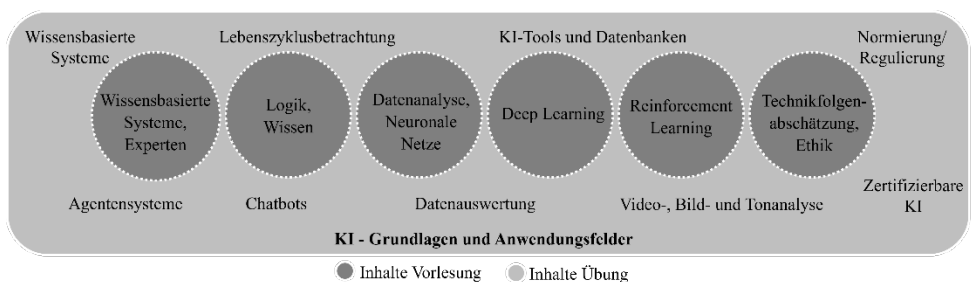


Abb. 1: Inhalte von Vorlesung und Übung

Mit 3 SWS und 5 CPs ist das Fach als seminaristische Vorlesung und Workshops mit hohem Praxisanteil ausgewiesen. Das vorliegende Paper adressiert primär das Lerntagebuch als didaktische Form der Wissensvermittlung. Durch die bewussten Entscheidungen als Grundlagenkurs zu firmieren sowie themenoffene Schwerpunkte zu setzen sind die Kursinhalte auf die Teilnehmerinnen und Teilnehmer und deren Fachrichtung flexibel adaptierbar. Die aufgegriffenen Lehr- und Lerninhalte repräsentieren einen thematischen Querschnitt zum Thema Künstliche Intelligenz (s. Abb. 1). Das Spektrum der ausgewählten Themen orientiert sich an dem fachlichen und inhaltlichen Aufbau des ZAKKI Projektes in den Bereichen Technik, Analytik und Ethik. Eine klassische, harte Trennung der drei genannten Bereiche wurde für die Entwicklung des Kurses und die Wissensvermittlung als nicht zielführend erachtet. Vielmehr wurden in der Ausgestaltung eine inhaltliche Gewichtung vorgenommen.

Somit wurden inhaltliche und gedankliche Schranken aufgelöst und eine gesamtheitliche, reflektierte Sichtweise angestoßen. Es sei betont, dass es sich um eine prototypische Umsetzung eines Grundlagenkurses zur KI handelt. Teilnehmerinnen und Teilnehmer erhielten einen Überblick über die Themengebiete und wurden mit niederschweligen Übungen (u.a. Erstellung eines eigenen Chatbots, Entwicklung VCIO-Label, Fallstudie zur Merkmalsklassifikation innerhalb eines KI-Tools) erstmalig mit dem Thema KI zusammengebracht. Übungen wurden als Use Case formuliert, Leitfragen mit einer gleichbleibenden Struktur (Was wird thematisiert? Welche Lösungen werden aufgezeigt? Welche Kritik existiert?) unterstützten den eigenständigen Lösungsprozess via Lerntagebuch. Durch Gruppeneinteilung und Nutzung des Lerntagebuches konnten die adressierten KI-Themen und Übungen mit hohem Interesse, hoher Eigenmotivation und nachhaltigem Verständnis gelehrt werden. Der Kurs und das eingesetzte Lerntagebuch haben es geschafft, Teilnehmerinnen und Teilnehmer ohne Vorwissen zur KI, für das Thema und dessen Inhalte zu motivieren. Einzelne Schwerpunkte können danach gezielt in weiterführenden Kursen vertieft werden. So sind unter anderem Selbstlernkurse für die Python-Programmierung oder Innovationsentwicklung mit KI geplant. Die Nutzung von geeigneten Kollaborationswerkzeugen (u.a. moodle, Miro) und KI-Tools (u.a. KNIME, Orange, TensorFlow) ist ein weiterer wesentlicher Bestandteil zum Gelingen des Grundlagenkurses.

3 Lerntagebuch als Lernmethode

Neben der Aneignung von Fachwissen, gewinnen Handlungs- und Problemlösungskompetenzen immer mehr an Bedeutung [KS19]. Gerade MINT-Studierende werden durch das Einüben von Berechnungen und Verfahren sowie durch die Anwendungen starrer Lösungswege zwar darauf konditioniert, komplexe quantitative Probleme zu lösen; häufig fehlt es ihnen aber an einem qualitativen Verständnis der zumeist abstrakten Lerninhalte, was einen Wissenstransfer in die Praxis und somit nachhaltige Wissensvermittlung erschwert [Br21, Wa21a]. Um in der Lehrveranstaltung „Einführung in die KI – Grundlagen und Anwendungsfelder“ die KI-Lehrinhalte den Lernenden der ingenieurwissenschaftlichen Studiengänge zugänglicher zu gestalten und den Lernerfolg zu maximieren, wurde die Lernmethode „Lerntagebuch“ als Prüfungsform gewählt. Bei der Lernmethode „Lerntagebuch“ handelt es sich um eine aktivierende, didaktische Methode zur Unterstützung der Selbstlernprozesse des Lernenden [Sc22, Wa21b]. Mit der Methode des Lerntagebuchs sollten in der Lehrveranstaltung drei Ziele verfolgt werden: Erstens sollten die Studierenden durch eine eigene kurze Zusammenfassung der Sitzungsthemen einen Überblick über die Lerninhalte erhalten, zweitens sollten die Studierenden befähigt werden, wesentliche Thesen und Erkenntnisse der Lerninhalte benennen zu können und drittens sollten die Studierenden durch das Bearbeiten verschiedener Lernaufgaben (Übungen), so genannter Prompts, einen Anwendungsbezug zwischen dem erlangten Fachwissen und praxisnaher Anwendungsfälle (Theorie-Praxis-Transfer) herstellen, wodurch ihre Handlungs- und Lösungskompetenz geschult werden sollte.

Insgesamt soll die Methode des Lerntagebuchs in der Lehrveranstaltung nicht nur den Lernprozess des Lernenden nachvollziehbar machen, sondern den Lernenden eine intensivere Auseinandersetzung mit den Lerninhalten und somit ein nachhaltigeres Lernen ermöglichen.

3.1 Evaluationsergebnisse

Der Verlauf der Lehrveranstaltung und der Prozess der Wissensaneignung wurde evaluativ durch eine quantitative Fragebogenerhebung der teilnehmenden Studierenden (N=20) an vier Zeitpunkten (Anfangs- und Enderhebung sowie zwei Zwischenbefragungen) und einem abschließenden Workshop begleitet. Die Evaluation hat ergeben, dass die Teilnehmenden die Veranstaltung vorrangig aufgrund ihres Interesses an KI, aber auch wegen der Prüfungsform gewählt haben. So gaben in dem abschließenden Workshop 35% der Teilnehmenden an, dass sie das Wahlpflichtmodul aufgrund der Prüfungsform „Lerntagebuch“ gewählt zu haben. Als Vorteile am Lerntagebuch wurde die Möglichkeit zur umfangreichen Bearbeitung und intensiven Auseinandersetzen mit den Lerninhalten und deren Reflexion genannt. Auch wurde von den Studierenden der hohe Lerneffekt durch den inhaltlichen Praxis- und Realitätsbezug positiv hervorgehoben. Die durch das Lerntagebuch begünstigte Dokumentation der Lerninhalte habe zudem den positiven Nebeneffekt, dass ein späteres Wiederaufarbeiten der Inhalte einfach möglich sei. Gerade im Vergleich zur Prüfungsform „Klausur“ und dem oft damit zusammenhängenden Prüfungsstress am Ende des Semesters, betonten die Studierenden den angenehmen, über das Semester in kleine Arbeitspakete verteilten Arbeitsaufwand. Zusätzlich hätte das Lerntagebuch auch ein eigenständiges und strukturiertes Arbeiten in der Gruppe gefördert. Als nachteilig wurde die anfangs nicht eindeutig formulierten Prüfungsanforderungen des Lerntagebuchs, die zum Teil sehr umfangreichen Übungsaufgaben, das teilweise stark schwankende Engagement der Gruppenmitglieder sowie die technische begrenzte Umsetzung des Lerntagebuchs in moodle bewertet. Insgesamt bewerten 90% der Teilnehmenden die Prüfungsform „Lerntagebuch“ als „gut“ oder „sehr gut“ und 95% haben keinen Wunsch nach einer anderen Prüfungsform.



4 Zusammenfassung und Ausblick

Die Bedeutung von KI in der Ausbildung an Hochschulen wird voraussichtlich stetig zunehmen. Die rasante Entwicklung von KI-Technologien erfordert kontinuierliche Aktualisierung und Weiterentwicklung von Lehr-Lern-Konzepten. Aufbrechen und verbinden von Wissenssilos sowie die praxisbezogene Interaktion zum Thema KI ist ein wesentlicher Schritt zur Gestaltung aktivierender Lehr- und Lernformate. Das beschriebene Lerntagebuch kann eine Möglichkeit sein, Teilnehmerinnen und Teilnehmer für das Thema KI zu begeistern. Die im vorgestellten Modul erhobene Befragung motiviert dazu weitere innovierende Formate zu entwerfen oder zu kombinieren.

Literaturverzeichnis

- [Br21] Brunnhuber, M. et al.: Nicht ins Leere lehren. Feedback und Interaktion in MINT-Lehrveranstaltungen. In: (Waldherr, F., Walter, C, Hrsg.): Didaktisch und praktisch. Methoden und Medien für die Präsenz- und Onlinelehre, Schäffer-Poeschel, Stuttgart, 2021, S. 169–180
- [Ho22] Hochschule Magdeburg-Stendal: Ergebnisse der Online-Befragung zur Studien- & Lehrsituation sowie zu den Bedarfen an Lern- und Lehrmaterial, https://www.h2.de/fileadmin/user_upload/Innovative_Hochschullehre/ZAKKI/Umfrageergebnisse_Bedarfserhebung_StudLehr_FB_anonym_2022.pdf, Stand: 04.06.2023
- [KS19] Klopsch, B. ; Sliwka, A.: Service Learning als „deeper learning“: Durch soziales Engagement (über-)fachliche Kompetenzen fördern. In: (Jahn, D. et al., Hrsg.): Kritische Hochschullehre. Impulse für eine innovative Lehr- und Lernkultur, Wiesbaden : Springer, 2019, S. 163–182
- [De23] De La Higuera, C.: A report about Education, Training Teachers and Learning Artificial Intelligence: Overview of key issues, https://www.k4all.org/wp-content/uploads/2019/11/Teaching_AI-report_09072019.pdf, Stand: 25.06.2023
- [Sc22] Schellenbach-Zell, J.: How can teacher candidates be supported in the science-based reflection of self-experienced school situations? A quasi-experimental study on the learning effectiveness of prompts and feedback in internships, *Unterrichtswissenschaft* 50/4, Springer VS, S. 689–715, 2022.
- [Ve23] Verein Deutscher Ingenieure: Künstliche Intelligenz im Ingenieuralltag | VDI. URL <https://www.vdi.de/ueber-uns/presse/publikationen/details/kuenstliche-intelligenz-im-ingenieuralltag>, Stand: 04.07.2023
- [Wa21a] Walter, C.: Lernen ist nicht machbar. Eine Begründung für die Forderung nach aktivierenden Lehrmethoden. In: (Waldherr, F., Walter, C. Hrsg.): Didaktisch und praktisch. Methoden und Medien für die Präsenz- und Onlinelehre, Schäffer-Poeschel, Stuttgart, 2021, S. 105–114.
- [Wa21b] Waldherr, F. ; Wendorff, J. ; Kipp, M.: Methoden zur Begleitung der studentischen Selbstlernphase. In: (Waldherr, F., Walter, C. Hrsg.): Didaktisch und praktisch. Methoden und Medien für die Präsenz- und Onlinelehre, Schäffer-Poeschel, Stuttgart, 2021, S. 91–96.

Artificial Intelligence meets Serious Gaming

Maria Freese ¹, Birgit Zürn² and Julia Arlinghaus ³


1 Introduction & Cards for Artificial Intelligence

Hardly any other topic is currently being discussed as much as the use of artificial intelligence (AI). Considering studies [Li20], [Si22] showing that the use of AI still evokes skepticism and fear among the population, as well as the increasing tendency to cooperate with AI, tools are needed to inform humans about AI, make AI and associated risks and measures to deal with these risks transparent by providing a safe space for discussion. Serious games provide this framework. They are defined as games that, in addition to entertainment, pursue serious objectives [Sc23] and can be used in different application areas. Cards for Artificial Intelligence⁴ is an analogue round-based card game that can be experienced with up to eight players. The goal of this game is to raise awareness about topics around AI by reflecting on risks and subsequently measures to mitigate these risks. At the beginning of each round, one scenario is chosen by the game facilitator. The goal in the game is to score as many points as possible by choosing and arguing the most appropriate risks (first phase) for the introduced scenario and measures (second phase) to mitigate these risks. This game is seen as a discussion starter. Theoretical feedback loops are important to guarantee learning (e.g., ranking the selected risks in a risk matrix).

2 Method and Results

In two rapid prototyping workshops during this year's European Simulation and Gaming Forum ($N = 8$) as well as International Simulation and Gaming Association Conference ($N = 7$), the content of the Cards for Artificial Intelligence game has been developed by serious gaming and AI experts. Based on a previously introduced scenario (Use of ChatGPT), the rapid prototyping was mainly related to the generation of possible risks (e.g., Supports laziness) and per risk at least two measures to mitigate the risks (e.g., Penalty, Training). To guarantee comparability between the workshops, each followed the

¹ Otto-von-Guericke-Universität Magdeburg, Lehrstuhl für Produktionssysteme und -automatisierung,

Universitätsplatz 2, 39106 Magdeburg, maria.freese@ovgu.de, 
<https://orcid.org/0000-0001-8700-6250>

² Duale Hochschule Baden-Württemberg, Zentrum für Managementsimulation (ZMS), Paulinenstr. 50, 70178 Stuttgart, birgit.zuern@dhw-stuttgart.de

³ Fraunhofer-Institut für Fabrikbetrieb und -automatisierung, Sandtorstraße 22, 39106 Magdeburg,

julia.arlinghaus@iff.fraunhofer.de, 
<https://orcid.org/0000-0002-0287-2086>

⁴ The game is based on the serious game Cards for Biosafety [FTV22].

same structure: warm-up, explanation of game, distribution of consent forms, rapid prototyping, game play, debriefing, and post-game questionnaire. In total, 12 participants completed the questionnaire ($f = 4$, $m = 8$; $R = 28 - 58$ years, $M = 42.58$, $SD = 9.99$). The participants highlighted the combination of the rapid prototyping process and the actual game play in order to learn more about AI. Furthermore, the added value of the heterogeneous group composition was emphasized, which made it possible to perceive different perspectives. Eleven out of 12 participants (91.67%) state that serious games can help reducing fears in dealing with AI. They see the potential of AI in serious games primarily in the automation and individualization of processes as well as generation of adaptive scenarios. In contrast, the participants see risks in using AI in serious games especially with regard to the elimination of the human factor and associated challenges, such as loss of creativity and innovation. The participants found the combination of AI and an analogue card game “exciting”, “interesting”, and “the perfect solution”.

3 Discussion and Future Research

The goal of Cards for Artificial Intelligence is to define risks based on concrete AI-related scenarios and to derive measures to mitigate these risks. Even if in a first step the use of ChatGPT was chosen as a scenario, the addition of other AI-related scenarios is easily possible. In future, the game can also be used to empirically analyze the extent to which participants' attitudes change after experiencing the game and to what extent the game was able to broaden horizons in terms of using AI.

4 References

- [FTV22] Freese, M.; Tiemersma, S.; Verbraeck, A.: “Risk Management Can Actually Be Fun” - Using the Serious Cards for Biosafety Game to Stimulate Proper Discussions About Biosafety. In (Dhar, U.; Dubey, J.; Dumblekar, V.; Meijer, S.; Lukosch, H., eds.): *Gaming, Simulation and Innovations: Challenges and Opportunities*. ISAGA 2021. Lecture Notes in Computer Science, 13219. Springer, Cham, pp. 124-133, 2022.
- [Li20] Lichtenthaler, U.: Extremes of acceptance: employee attitudes toward artificial intelligence. *Journal of Business Strategy* 41/5, pp. 39-45, 2020.
- [Sc23] Schrader, C.: Serious Games and Game-Based Learning. In (Zawacki-Richter, O.; Jung, I., eds.): *Handbook of Open, Distance and Digital Education*. Springer, Singapore, pp. 1255-1268, 2023.
- [Si22] Sindermann, C. et.al.: Acceptance and Fear of Artificial Intelligence: associations with personality in a German and a Chinese sample. *Discover Psychology* 2/8, 2022.

Integration ethischer Kompetenzen in einen Kurs zu Grundlagen Künstlicher Intelligenz

Lisa Hopf¹, Maria Hinsche² und Christina B. Class³

Abstract: Die Bedeutung und das Potential von Künstlicher Intelligenz (KI) erfordert die Integration ethischer Kompetenzen in alle Lernangebote im Bereich KI, um ein ethisches Bewusstsein zu schärfen. Welche Kompetenzen hierzu zählen und wie diese zu vermitteln sind, ist noch weitgehend offen. Der vorliegende Beitrag gibt einen groben Überblick über notwendige Kompetenzen und stellt ein didaktisches Konzept vor, um diese zu vermitteln und eine Verzahnung von grundlegenden KI-Kompetenzen mit ethischen Kompetenzen zu erreichen.

Keywords: Ethik, Didaktik, Ethik und KI, Kompetenzen, Unterrichtskonzept

1 Einleitung

KI-Anwendungen beeinflussen viele Bereiche unserer Gesellschaft. Die damit verbundenen ethischen Herausforderungen sind vielfältig und betreffen nicht nur Fragen von personenbezogenen Daten und Fälle möglichen Missbrauchs oder direkter Diskriminierung. Technische Systeme haben Einfluss auf uns, unsere zwischenmenschlichen Beziehungen und daher unsere Gesellschaft als Ganzes. So thematisiert die Enquete-Kommission Künstliche Intelligenz innerhalb ihres Abschlussberichts zum Beispiel sowohl die Gefährdung von Autonomie und Selbstbestimmung, als auch die Veränderung des menschlichen Selbstverständnisses durch KI [De20]. Um die Auswirkungen von KI-Systemen einschätzen und diese verantwortungsvoll und zielgerichtet nutzen zu können, bedarf es daher neben eines grundlegenden Verständnisses von KI-Systemen auch ethische Kompetenzen im Blick auf KI.

Eine allgemein gültige Zusammenstellung von grundlegenden KI-Kompetenzen existiert zurzeit noch nicht. Definitionen im Blick auf AI Literacy sind durch unterschiedliche Schwerpunkte geprägt. Auch wenn Kompetenzen im Bereich Ethik erwähnt werden, sind diese nicht ausreichend detailliert beschrieben. Auch die didaktische Forschung bezüglich kompetenzorientierter KI-Lehre steht noch am Anfang [WRP20]. Es fehlt daher auch an

¹ Ernst-Abbe-Hochschule Jena, Fachbereich Grundlagenwissenschaften, Carl-Zeiss-Promenade 2, 07745 Jena, e-mail: lisa.hopf@eah-jena.de

² Ernst-Abbe-Hochschule Jena, Fachbereich Grundlagenwissenschaften, Carl-Zeiss-Promenade 2, 07745 Jena, e-mail: maria.hinsche@eah-jena.de

³ Ernst-Abbe-Hochschule Jena, Fachbereich Grundlagenwissenschaften, Carl-Zeiss-Promenade 2, 07745 Jena, e-mail: christina.class@eah-jena.de

Empfehlungen, wie die notwendigen ethischen Kompetenzen vermittelt werden können. Hier setzt der vorliegende Beitrag an.

2 Ethische Kompetenzen

Im Rahmen des Projekts "Modularer Lehrbaukasten zur Vermittlung von KI-Kompetenzen für angewandte Wissenschaften an der Ernst-Abbe-Hochschule Jena (MoVeKI2EAH)"⁴ werden grundlegende ethische Kompetenzen in Bezug auf KI mithilfe einer umfangreichen Literaturrecherche identifiziert und in einer Kompetenzmatrix zusammengestellt. In dieser Matrix werden die Kompetenzen sortiert und dabei verschiedenen Kompetenzniveaus zugeordnet. Hierbei hat sich gezeigt, dass die identifizierten Kompetenzen in die vier Themenbereiche "ethische Grundlagen", "spezifische Eigenschaften von KI-Systemen", die "Entwicklung und Analyse von KI-Anwendungen" sowie den "Einsatz und die Grenzen von KI" untergliedert werden können.

Um ethische Problemstellungen erkennen und über diese sprechen zu können, bedarf es *Ethischer Grundlagen*. Diese umfassen Kenntnisse relevanter Grundbegriffe wie Ethik, Moral und Werte sowie einiger grundlegender Ansätze der Moralphilosophie. Es ist notwendig, dass Studierende lernen, in einfachen Beispielen ethische Problemfelder zu identifizieren. Ein wichtiges Ziel ist, gemeinsam mit anderen Lösungsansätze diskutieren zu können, und sich dabei generell mit ethischen Fragen auseinanderzusetzen. KI-Systeme basieren auf Software und müssen Kriterien von guter und zuverlässiger Software erfüllen. Da sich die Entwicklung von KI-Anwendungen gleichzeitig grundlegend von der Entwicklung klassischer Systeme unterscheidet, müssen weitere Kriterien ins Auge gefasst werden. KI-Systeme weisen *spezifische Eigenschaften* auf, die es erschweren bestimmte Anforderungen sicherzustellen. Hiermit zusammen hängen Themen wie Transparenz und Erklärbarkeit, Verantwortlichkeit und Rechenschaftspflicht, Fairness und Gerechtigkeit, das Prinzip des Nicht-Schadens, Privatheit und Privatsphäre, Datenschutz sowie Datenqualität. Wie bei allen Softwaresystemen beinhaltet auch die Entwicklung von KI-Systemen verschiedene Schritte. Hierzu gehören Analyse, Entwicklung sowie der Einsatz. In all diesen Schritten werden Entscheidungen getroffen und Weichen gestellt, die ethische Relevanz haben können. Als Beispiele für ethisch relevante Kompetenzen für die *Entwicklung und Analyse* von KI-Systemen können die Identifikation von Stakeholdern und die Beachtung von ethischen Richtlinien während des gesamten Entwicklungsprozesses genannt werden. Neben grundlegenden Kenntnissen der Risikoanalyse müssen mit Blick auf KI-Systeme weitere Kompetenzen betrachtet werden, um den *Einsatz und die Grenzen zu bewerten*. Hier müssen insbesondere die Auswirkungen von KI-Systemen auf Menschen, die Interaktion von Menschen sowie die Gesellschaft als Ganze in das Bewusstsein rücken. Aber auch Auswirkungen auf die Umwelt und insbesondere Fragen der Nachhaltigkeit, der Ökologie und des Ressourcenverbrauchs bei der Entwicklung, dem Training großer Modelle sowie dem Einsatz müssen beachtet werden.

⁴ Website des Projekts: <https://www.eah-jena.de/moveki2eah>

3 Didaktische Grundlagen

Ziel der Vermittlung der genannten ethischen Kompetenzen ist die Stärkung eines ethischen Bewusstseins. Hierfür ist eine Verzahnung von ethischen und technischen Kompetenzen außerordentlich wichtig, da die Studierenden ansonsten Ethik von der technischen Entwicklung von KI-System abgetrennt erlernen und wahrnehmen. Studierende müssen daher lernen, zu erkennen, dass in jedem Schritt neben technischen auch ethische Entscheidungen zu treffen sind und sie sich daher auch immer in einem ethischen Kontext befinden [BLM22]. Die Thematisierung ethischer Fragen sollte daher in technischen Kursen und folglich auch in KI-Kursen zu einer alltäglichen Praxis werden, um so die Nachhaltigkeit des Ethikunterrichts zu gewährleisten [FGB20]. Eine inhaltliche Verzahnung legt auch wichtige Grundlagen für die Übertragung ethischer Kompetenzen aus dem universitären in den beruflichen Kontext.

In diesem Zusammenhang ist die gemeinsame Diskussion lebensweltrelevanter Beispiele hilfreich, anhand derer nicht nur die Fähigkeit zu reflektieren, sondern auch Selbstreflexion trainiert werden kann. Ein Lernen in der Gruppe trainiert darüber hinaus auch das Einüben konstruktiver sozialer Interaktion innerhalb von Problemlöseprozessen [AS19; Mi07]. Damit werden auch notwendige Kompetenzen eingeübt, um ethische Fragen in Bezug auf KI in einem Diskurs zu behandeln und zu gemeinsamen Entscheidungen zu gelangen.

4 Vorschlag eines Unterrichtskonzepts

Das folgende Unterrichtskonzept wurde für ein Modul entwickelt, in dem KI-Grundlagen für Bachelor-Studierende mit unterschiedlichen Vorkenntnissen vermittelt werden. Die Zeit, die für die Vermittlung von ethischen Kompetenzen zur Verfügung steht, ist, wie in vielen anderen ähnlichen Modulen, begrenzt. Ein wesentliches Ziel des Konzepts besteht in der Verzahnung der Behandlung von Grundlagen der KI mit ethischen Fragestellungen und Diskussionen, um Schubladendenken zu vermeiden.

Eine Verzahnung in jeder einzelnen Unterrichtseinheit ist in der Regel nicht möglich. Auch wenn wir alle im Alltag regelmäßig moralische Entscheidungen treffen, fühlen sich viele Lehrende bei dem Gedanken unwohl, in ihren Kursen die Diskussion ethischer Fragestellungen zu integrieren [FGB20; No21; SI20]. Aus diesem Grund werden für die Diskussion ethischer Themen wenige separate Präsenzveranstaltungen vorgesehen, die durch interessierte Lehrpersonen übernommen werden. Im aktuellen Modul planen wir hierfür zwei Termine à 90 Minuten ein. Im ersten Termin wird anhand von konkreten Beispielen in gemeinsamen Diskussionen das Verständnis geschult, dass Technik nicht neutral ist und immer Werte und Entscheidungen der Entwickler:innen enthält. Im zweiten Termin setzen sich die Studierenden mit Fragen der Diskriminierung durch KI-Systeme auseinander. Um fundierte Diskussionen führen zu können, erwerben die Studierenden technische Grundkenntnisse in verschiedenen Themengebieten der KI. Begleitend hierzu werden darauf bezugnehmende *Aufgaben für das Selbststudium* zur Verfügung gestellt, welche die ethische Reflexion der thematisierten Inhalte zum Ziel haben.

Einige Aufgaben nehmen direkten Bezug zu den behandelten theoretischen Grundlagen. Andere Aufgaben dagegen zielen auf das Nachdenken in größeren Zusammenhängen ab, das über das Reflektieren des Einsatzes einzelner Methoden hinausgeht. Konkrete Fragen öffnen den Blick für mögliche Probleme beim Einsatz bestimmter Maße oder Methoden. Rechercheaufträge ermuntern dazu, weitergehende Begriffe und Konzepte kennenzulernen, zusätzliche Aspekte zusammenzutragen oder reale Fälle zu sammeln. Kleinere Fallbeispiele oder Fallstudien werden mit Fragen angereichert und laden zur Diskussion ein. Diese unterschiedlichen Aufgabentypen ermöglichen die Auseinandersetzung mit ethischen Fragen aus unterschiedlichen Perspektiven und fördern unterschiedliche Kompetenzen.

Die Ergebnisse der Aufgaben werden von den Studierenden in die Lernplattform hochgeladen. Manche Aufgaben werden in Teams bearbeitet, einige Aufgaben werden von Peers kommentiert, um in den Dialog zu kommen. Andere Aufgaben werden kurz von der Lehrperson kommentiert.

So entsteht im Laufe des Semesters eine kleine Sammlung von individuellen oder gemeinsamen Dokumenten und schriftlich niedergehaltenen Reflexionen. Dieser Ansatz stellt eine prozessorientierte ethische Auseinandersetzung mit KI-bezogenen Themen in den Mittelpunkt. Daher werden die entstehenden Produkte auch nicht benotet. Die Bearbeitung sowie die aktive Teilnahme an den Ethik-Präsenzveranstaltungen wird jedoch für das Bestehen des Moduls vorausgesetzt.

5 Ein Beispiel zum Thema Clustering

Im Modul lernen die Studierenden die Funktionsweise zweier Clustering-Methoden, des hierarchischen und des k-Means Clustering, experimentell kennen und beobachten die Auswirkungen von Änderungen von Hyperparametern. Daneben setzen sie sich in Übungen mit der Frage nach der Bestimmung der optimalen Anzahl von Clustern auseinander und beobachten am Beispiel selbstgewählter zufälliger Daten, dass der k-Means Clustering Algorithmus auch dort Cluster ausmacht/bestimmt, wo keine vorhanden sind.

Hier knüpft nun das Selbststudium an. Für das Thema "Clustering" wurde ein Beispiel aus dem Bereich des Marketings gewählt. In dem Beispiel verwendet ein Sportbedarfshändler einen Cluster-Algorithmus, um anonymisierte Daten von Kund:innen in drei Marktsegmente einzuteilen und spezifische Angebote zu erstellen. Um diese Angebote zu definieren müssen nun im Nachhinein grundlegende Eigenschaften identifiziert werden, welche die einzelnen Gruppen ausmachen.

Die Studierenden sind nun gefordert, das Beispiel zu reflektieren. Dabei werden sie anhand kleinschrittiger Instruktionen angeleitet. Die Aufgaben werden gemeinsam von zwei bis drei Studierenden bearbeitet.

Zunächst wird auf die Interpretation der Cluster eingegangen. Clustering wird immer mit dem Ziel der Segmentierung eingesetzt. Diese muss interpretiert werden, um darauf

basierend zu handeln, in diesem Fall Angebote zu entwickeln. Die Studierenden tragen mögliche verwendete, nicht personenbezogene Attribute zusammen und beschreiben daraus ableitbare Interpretationen der drei Cluster. Hieran schließt sich eine Diskussion über die Möglichkeiten, realistische Interpretationen für die Cluster zu finden, an. Anschließend tragen die Studierenden aus Sicht der Kund:innen und des Händlers Vor- und Nachteile von Clustering für die Entwicklung von kundennahen Angeboten zusammen und geben eine "Empfehlung" für den vorliegenden Fall, wie die Daten interpretiert und genutzt werden könnten.

In diesem Beispiel kommt es nicht darauf an, dass die Studierenden spezifische Antworten finden und Themen identifizieren. Die Fragen dienen dazu, Denkprozesse sowie anschließende Diskussionen anzuregen. Die Studierenden transferieren in diesem Beispiel eingeführte KI-Methoden auf ein Anwendungsgebiet mit vielen Dimensionen. Sie erkennen, wie schwierig die Interpretation von Clustern im realen Fall ist und werden dafür sensibilisiert, Ergebnisse nicht unreflektiert zu übernehmen. Die Diskussion von Vor- und Nachteilen und die Entwicklung einer Einsatzempfehlung lädt dazu ein, das Anwendungsbeispiel aus unterschiedlichen Perspektiven umfassend zu betrachten und Ideen für einen verantwortungsbewussten Einsatz zu entwickeln.

Wir verzichten in diesem Beispiel bewusst auf die Nutzung personenbezogener Daten oder die Konstruktion eines Anwendungsfalls, z.B. im Bereich Medizin, bei dem Menschen offensichtlich zu Schaden kommen können. Stattdessen sollen die Studierenden darauf sensibilisiert werden, die ethische Perspektive stets zu berücksichtigen. Ein ethisches Bewußtsein und die Bereitschaft, sich mit diesen Fragen auseinanderzusetzen, wird insbesondere auch dann gefördert, wenn ethische Aspekte nicht als Ausnahme, sondern als allen Anwendungen inhärent wahrgenommen werden.

6 Zusammenfassung und Weiteres Vorgehen

KI-Kompetenzen sind für alle Studierenden von Relevanz. Hierzu gehören grundlegende Kompetenzen im Bereich KI und Ethik, damit die Studierenden KI-Systeme verantwortungsvoll entwickeln und nutzen sowie sich am gesellschaftlichen Diskurs über KI beteiligen können. Um das notwendige ethische Bewusstsein auch in Lehrangeboten mit begrenzter Zeit entwickeln zu können, bedarf es eines didaktischen Konzepts, das die Verzahnung der Vermittlung von technischen und ethischen Grundlagen ermöglicht. In diesem Beitrag werden die Grundlagen eines solchen didaktischen Konzeptes dargelegt und das Konzept beschrieben. Es wird durch ein Beispiel erläutert.

Das Konzept ist ein Resultat des Projekts MoVeKI2EAH, in dem u.a. ein Modul entwickelt wird, das grundlegende KI-Kompetenzen für Studierende unterschiedlicher Studiengänge vermittelt. Das beschriebene Beispiel wird aktuell gemeinsam mit den technischen Grundlagen des Clustering ausgearbeitet, um im kommenden Studienjahr eingesetzt zu werden.

Das Modul wird im Wintersemester 2023/2024 das erste Mal angeboten. Die dazugehörigen Lehrmaterialien werden in Zukunft als OER (Open Educational Resources) zur Verfügung gestellt. Diese beinhalten auch Selbststudiumsaufgaben zu ethischen Kompetenzen.

Danksagung

Das Projekt MoVeKI2EAH wird durch das BMBF (Bundesministerium für Bildung und Forschung) und den Freistaat Thüringen im Rahmen der Bund-Länder-Initiative zur Förderung von Künstlicher Intelligenz in der Hochschulbildung gefördert (12/2021 bis 11/2025, Förderkennzeichen 16DHBKI081). Wir danken allen Mitgliedern des Projektteams für das Feedback zu ethischen Kompetenzen und dem vorgestellten Konzept. Wir danken auch den anonymen Reviewern für die konstruktiven Kommentare.

Literatur

- [AS19] Arnold, R.; Schön, M.: *Ermöglichungsdidaktik: Ein Lernbuch*. hep, 2019, ISBN: 978-3-0355-1091-1.
- [BLM22] Bruneault, F.; Laflamme, A. S.; Mondoux, A.: *AI Ethics Training in Higher Education: Competency Framework, 2022*, URL: <https://eduq.info/xmlui/bitstream/handle/11515/38628/AI-Ethics-Competency-Framework-PIA-CLaurendeau-UQAM-EN-2022.pdf>.
- [De20] Deutscher Bundestag: *Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale*, Drucksache 19/23700, 2020.
- [FGB20] Fiesler, C.; Garrett, N.; Beard, N.: *What Do We Teach When We Teach Tech Ethics?* In (Zhang, J., Hrsg.): *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. ACM Digital Library, Association for Computing Machinery, New York, NY, United States, S. 289–295, 2020, ISBN: 9781450367936.
- [Mi07] Mikula, R.: *Lebenslanges Lernen – emanzipatorische Potenziale veränderter Lernkulturen*. MAGAZIN. Erwachsenenbildung.at/2, S. 5, 2007.
- [No21] Nourbakhsh, I. R.: *AI ethics: A call to faculty*. *Communications of the ACM* 64/9, S. 43–45, 2021.
- [SI20] Slavkovik, M.: *Teaching AI Ethics: Observations and Challenges*. *Norsk IKT-konferanse for forskning og utdanning/4*, 2020, ISSN: 1892-0721.
- [WRP20] de Witt, C.; Rampelt, F.; Pinkwart, N.: *Künstliche Intelligenz in der Hochschulbildung*. *Whitepaper, KI-Campus*, 2020, URL: <https://ki-campus.org/publications/whitepaper-ki-hochschulbildung>.


Erhebung einer Selbsteinschätzung zu KI-Kompetenzen sowie zu genutzten KI-Funktionen und Fortbildungsinteressen von Lehrkräften

Uwe Lorenz  ¹

Abstract: Es wird eine Fragebogenerhebung vorgestellt, mit der auf Grundlage des AI-PACK Rahmens Selbsteinschätzungen und Fortbildungsinteresse hinsichtlich der KI-bezogenen Anforderungsbereiche der Lehrerprofessionalität erfasst werden sollen. Außerdem soll untersucht werden, ob es fächerspezifische Unterschiede hinsichtlich der als relevant erachteten KI-Kompetenzen gibt.

Keywords: AI-PACK, KI-Kompetenzen, KI-Lehrkräftebildung

Digitalkompetenz und entsprechend „KI-Kompetenz“ von Lehrkräften muss im Zusammenspiel mit fachlichen und pädagogischen Kompetenzen diskutiert werden [Dö18]. Die KI-Lehrkräfteaus- und weiterbildung stellt sich als eine fächerübergreifende Querschnittsaufgabe, weil diese Technologie deutliche Auswirkungen auf zentrale Handlungsfelder aller Lehrkräfte hat. Wie werden die eigenen KI-Kompetenzen allgemein und in Bezug auf die Anforderungsbereiche der Lehrkräfteprofessionalität selbst eingeschätzt und wie groß ist das Interesse an entsprechenden Fortbildungen? Gibt es fachspezifische Unterschiede hinsichtlich der als relevant angesehenen KI-Funktionen? Zur Beantwortung dieser Fragen wurde eine Quantitative Fragebogenmethode mit Likert-Skalen gewählt, für die ein Pretest durchgeführt wurde, dessen Resultate wir hiermit vorstellen und zur Diskussion stellen möchten. Der Fragebogen, der sich an Lehrkräfte richtet, besteht aus 4 Teilen. Während im ersten Abschnitt allgemeine KI-Kompetenz aus technischer, anwendungsorientierter und gesellschaftlicher Perspektive [MRS22] behandelt wird, werden im zweiten Teil die lehrbezogenen Kompetenzbereiche adressiert. Es werden jeweils allgemein Selbsteinschätzungen und Weiterbildungsinteresse erfragt. Um die lehrbezogenen Felder zu spezifizieren, nutzen wir den Rahmen „AI-PACK“ [LR], in dem das „DPACK-Modell“ [Dö18] angewendet wird, welches auf dem weithin akzeptierten Rahmen „TPACK“ aufbaut. Im dritten Teil wird nach der fachspezifischen Relevanz bestimmter KI-Funktionen gefragt. Um zentrale KI-Funktionen zu ermitteln, wurde explorativ eine Liste mit fachspezifischen KI-Anwendungsfällen für die Schulfächer der allgemeinbildenden Schulen angefertigt, die ca. 100 Einträge umfasst. Aus diesen Anwendungsfällen wurden unter Einbeziehung von Ähnlichkeiten aus technischer Perspektive sechs zentrale KI-Funktionen extrahiert „klassifizieren/erkennen“, „korrigieren/aufbereiten“, „erzeugen“, „anpassen/variiieren“, „zusammenfassen/reduzieren“ und „übersetzen/interpretieren“ (vgl. Abb. 2). Die Funktionen

¹Freie Universität Berlin, Königin-Luise-Str. 24-26, 14195 Berlin, uwe.lorenz@fu-berlin.de,
 <https://orcid.org/0000-0002-4897-6411>

sollen unabhängig vom verarbeiteten Datentyp betrachtet werden, weshalb wir nach der Relevanz der Funktionen in Bezug auf Text, Bilder, Töne, sonstiges im Einzelnen fragen. Im vierten Teil wird offen nach Ideen für Anwendungen von aktiven KI-Systemen gefragt. Der Pretest wurde auf der Bildungsmesse „didacta“ durchgeführt, wobei die Teilnehmenden darauf angesprochen wurden, den Fragebogen auf Papier auszufüllen (n=33). Durch diese Art der Erhebung ergibt sich ein entsprechender Bias, der sich u.a. durch die Zusammensetzung der Besucher der Messe sowie auch des Teils, der bereit war den Fragebogen auszufüllen, ergibt. Bei der Darstellung der Ergebnisse stellen wir die Selbsteinschätzung von Lehrkräften des Primar- und Sekundarbereichs nebeneinander (vgl. Abb. 1). Bei den zentralen KI-Funktionen analysieren wir die Daten im Sekundarbereich auch im Hinblick auf Unterschiede zwischen Geisteswissenschaften und MINT-Fächer.

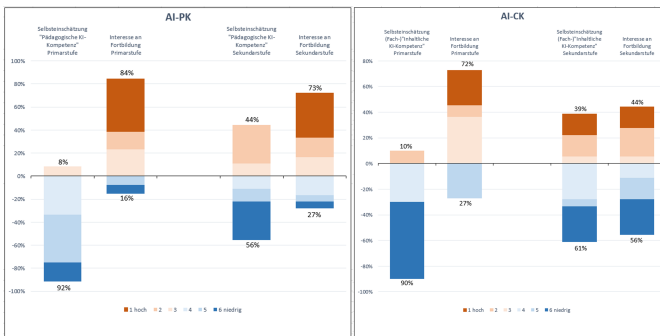


Abb. 1: Selbsteinschätzung und Fortbildungsinteresse in Bezug auf zwei AI-PACK Kompetenzfelder: KI-bezogene pädagogische Kompetenz (AI-PK) und KI-bezogene fachinhaltliche Kompetenz (AI-CK)

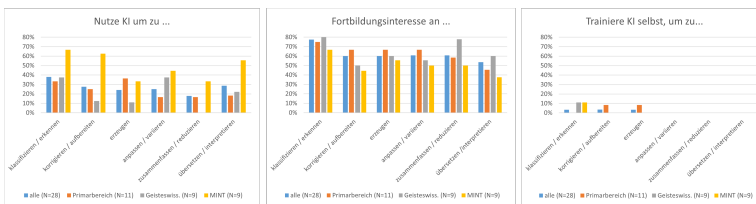


Abb. 2: Nutzung von zentralen KI-Funktionen und Fortbildungsinteresse

Literaturverzeichnis

- [Dö18] Döbeli Honegger, B.: Das DPACK-Modell. <https://mia.phsz.ch/DPACK/WebHome>, Stand: 13.01.2023.
- [LR] Lorenz, Uwe; Romeike, Ralf: AI-PACK – Eine Annäherung an den KI-bezogenen Bereich der Digitalkompetenz für Lehrkräfte mit dem DPACK-Modell. In: Tagungsband INFOS23.
- [MRS22] Michaeli, Tilman; Romeike, Ralf; Seegerer, Stefan: What students can learn about artificial intelligence - recommendations for K12 computing education. In: Proceedings of World Conference on Computers in Education, WCCE 2022. 2022.

KI-Technologieentwicklung und Bildung als miteinander in Beziehung stehende Aufgaben für die Lehrerbildung

Stefan Seegerer¹, Michaela Gläser-Zikuda², Florian Hofmann³ und Ralf Romeike⁴

Abstract: Der Beitrag stellt einen Forschungsansatz zur Adressierung von Chancen und Herausforderungen des Einsatzes von KI in der Lehrerbildung im Portfoliokontext vor. Im Fokus steht die Frage, wie die Professionalisierung von Lehramtsstudierenden sowie die Hochschullehre durch den Einsatz Künstlicher Intelligenz in Abstimmung mit der Kompetenzentwicklung der Lernenden und Lehrenden im Bereich KI unterstützt werden können. Neben der forschungsbasierten Entwicklung geeigneter KI-Methoden gilt hierbei ein besonderes Augenmerk der KI-Bildung der Studierenden, um auch die notwendigen motivationalen und volitionalen Voraussetzungen für den Einsatz eines KI-Systems zu berücksichtigen.

Keywords: Lehrerbildung, Portfolio, KI-gestützte Reflexion, Datenanalyse, KI-Bildung, Feedback

1 Einleitung

Trotz wachsender Potenziale von Anwendungen Künstlicher Intelligenz (KI) zeichnet sich für ihren Einsatz in der Hochschullehre ein Dilemma ab: Einerseits werden den Technologien umfangreiche Chancen, bspw. in Hinblick auf eine stärker individualisierte Gestaltung von Studienangeboten zugeschrieben, um u.a. Studienzufriedenheit zu fördern und Studienabbruchtendenzen entgegenzutreten; andererseits lassen sich bei Studierenden Abwehrhaltungen beobachten, die mit technologiebezogenen Unsicherheiten einhergehen bzw. auf diese zurückgehen. Eine besondere Herausforderung ergibt sich für die Lehrerbildung: Obwohl durch den Einsatz von KI-Technologien erprobte individuelle Konzepte zur Studienunterstützung einen signifikanten Beitrag zur Beseitigung des Lehrkräftemangels leisten könnten, lässt sich bei Lehramtsstudierenden eine größere Skepsis gegenüber neuen Technologien wahrnehmen – was sich durch ihre zukünftige Multiplikatorenrolle nochmal auf weitere Generationen manifestieren könnte. Im vorliegenden Beitrag wird ein Lösungsansatz beschrieben, in dem Technologieentwicklung und Bildung als miteinander in Beziehung stehende Aufgaben verstanden werden. Ziel ist die interdisziplinäre Erforschung von KI-Methoden für die Lehrerbildung und deren Realisierung innerhalb eines „KI-Portfolios“. Es wird untersucht, inwiefern Transparenz und Bildungsangebote hinsichtlich der eingesetzten KI-Verfahren zur Akzeptanz, KI-Bildung und -Mündigkeit

¹ FU Berlin, Didaktik der Informatik, Königin-Luise-Str. 24-26, 14195 Berlin, stefan.seegerer@fu-berlin.de

² FAU Erlangen-Nürnberg, Institut für Erziehungswissenschaft, Regensburger Straße 160, 90478 Nürnberg, michaela.glaeser-zikuda@fau.de

³ FAU Erlangen-Nürnberg, Institut für Erziehungswissenschaft, Regensburger Straße 160, 90478 Nürnberg, florian.hofmann@fau.de

⁴ FU Berlin, Didaktik der Informatik, Königin-Luise-Str. 24-26, 14195 Berlin, ralf.romeike@fu-berlin.de

von Lehramtsstudierenden und somit dem erfolgreichen Einsatz von KI-Technologien beitragen können.

2 Chancen und Herausforderungen von KI in der Lehrerbildung

Die Lehrerbildung steht bundesweit vor enormen Herausforderungen. Neben einem wachsenden Bedarf an qualifizierten Lehrkräften, die zunehmend überfachliche Aufgaben, wie die Umsetzung von Inklusion, Demokratiebildung oder digitale Bildung übernehmen sollen, nimmt die Heterogenität bei Studierenden zu. Hinzu kommt die steigende Dropout-Quote im Lehramtsstudium, nicht zuletzt aufgrund niedriger Studienzufriedenheit. Um die angeführten Herausforderungen zu fokussieren, ist der Portfolio-Ansatz [HWK16] vielversprechend. Insbesondere die mit alternativen, dialogbasierten Lern- und Leistungsinstrumenten verknüpfte Intention die Reflexionsfähigkeit der Studierenden zu fördern, kann einen wesentlichen Beitrag zur Professionalisierung von Lehramtsstudierenden liefern [PSM19]. Reflexion kann sowohl während (bspw. als reflection-in-action), als auch in der Vor- und Nachbereitung einer Situation stattfinden [Ko19]. Das Potenzial des Portfolios in der Lehrerbildung lässt sich mit den beiden Kategorien „Professionalisierungsinstrument“ (auf der Ebene der Lehramtsstudierenden) und „Strukturinstrument“ (auf der Ebene der inhaltlich-konzeptionellen Ausrichtung der Lehrerbildung) zusammenfassen [FC18]. Derartig operationalisierte Portfolioarbeit kann als Innovation verstanden werden [ZG16], die mit selbstreguliertem Lernen [GVR10], dem Erwerb professionsbezogener und digitaler Kompetenzen und damit der Förderung einer neuen Lernkultur sowie der Relationierung von Theorie und Praxis eng verknüpft ist.

Zusätzliche Möglichkeiten der Unterstützung reflexiven Lernens in der Lehrerbildung bieten mobile Technologien, bspw. mit Smartphones (vgl. [Pe19]). Mobile und multimediale Portfolios stellen eine erfolgsversprechende Umsetzungsalternative zu rein textbasierten Systemen mit hohem Cognitiven Load dar und werden von Studierenden meist als unaufdringlich und weniger ablenkend empfunden [MO14]. Das Metapholio-Konzept mit der gleichnamigen App [Pe19] bietet verschiedene Möglichkeiten Reflexion im Lehramtsstudium zu unterstützen (ebd.): Bspw. Wahrnehmen und Festhalten für „Reflection in Action“; kollaborativer Austausch und Diskussion festgehaltener „Momente“ bzw. Anregungen und Anlässe als Impuls für Reflexionsprozesse. Klassische und KI-basierte Verfahren der Learning Analytics ermöglichen es, Daten der Studierenden zu erfassen und hinsichtlich des Lernfortschritts und möglicher Unterstützungsangebote auszuwerten. Weitere Verbesserungsmöglichkeiten werden in der Unterstützung des selbstgesteuerten bzw. selbstregulierten Lernens gesehen [Ha18], da viele Studierende Schwierigkeiten haben, ohne äußeren Anlass ihren Lernprozess selbst zu steuern, was durch KI-Methoden, die auf die individuellen Lernerdaten zurückgreifen und so eine Personalisierung ermöglichen, unterstützt werden könnte. Während personalisierte Anreden, einfache inhaltliche Empfehlungen sowie auf aktuellen Leistungsdaten basierende adaptierte Lernangebote bereits als relativ problemlos realisierbar erachtet werden, ergibt sich ein individuell-adaptives Lernen anhand der vorliegenden Nutzerdaten, auch durch das Setzen von Lern- bzw.

Kompetenzziele, individualisierten Rückmeldungen und komplexeren Lerndialogen [Sa17]. Das Potenzial von individualisierten „Lernbegleitern“ wird auch für Präsenzphasen gesehen, welche die Dozierenden entlasten und den Lernprozess der Studierenden unterstützen können [SLF19]. Grundsätzlich ist festzustellen, dass je höher der Grad der beabsichtigten Lernunterstützung ist, umso elaboriertere Datenauswertungen und KI-Methoden notwendig werden. Modelle zur Analyse reflexiven Schreibens bauen auf zwei verschiedenen Merkmalsklassen auf, (1) die Tiefe (Ebenen) und (2) die Breite (Inhalt) der Reflexion, wobei i.d.R. eine Kombination dieser beiden Klassen genutzt wird [U115; CYZ16]. Unterschieden werden kann in diesem Zusammenhang zwischen wörterbuchbasierten, regelbasierten und auf maschinellem Lernen basierenden Ansätzen, welche regelmäßig kombiniert werden und nicht nur auf englischsprachige Portfolios, sondern auch auf reflexive Texte anderer Sprachen angewandt werden können (bspw. [Ch18]). Durch die Anwendung solcher Auswertungen für die KI-basierte Erzeugung von konstruktivem Feedback könnte die Qualität der Hochschullehre gefördert werden. Positive Effekte KI-generiertem Feedbacks auf das Schreiben reflexiver Texte sind bereits belegt [Gi17]. Neben der Förderung der Reflexionsfähigkeit und der Lernmotivation von Studierenden kann das Vertrauen in das gewählte Studium, speziell in naturwissenschaftlichen Fächern, gesteigert werden [K117].

Aus Sicht der informatischen Bildung und Medienbildung stellt sich zunehmend das Problem, dass KI-Technologien zwar in verschiedensten Alltagsgegenständen integriert verfügbar werden, hierzu allerdings nur wenige Bildungsmöglichkeiten existieren. Als Konsequenz werden gesellschaftliche Diskussionen über den Einsatz von KI häufig wenig fundiert und irrational geführt, und zwar sowohl in der gesellschaftlichen Meinungsbildung (wie bspw. hinsichtlich ethischer Aspekte des autonomen Fahrens) als auch bzgl. persönlicher Entscheidungen, wie bspw. zur Verwendung digitaler Assistenten (z.B. Siri und Co.). Resultierende Abwehrhaltungen stehen dem effektiven Einsatz KI-basierter Unterstützungssysteme in der Hochschullehre entgegen. So zeigt eine Studie zu Kompetenzen und Einstellungen von Informatiklehrkräften, dass ihr KI-Wissen maßgeblich durch oberflächliche „Hype-Begriffe“ geprägt ist und sie nur vage Vorstellungen von den tatsächlichen Grundprinzipien, Möglichkeiten und Grenzen der KI besitzen [LR19]. Aktuelle Ansätze der informatikdidaktischen Forschung zeigen, dass KI-Grundprinzipien sich so didaktisch aufbereiten lassen, dass deren grundlegende Ideen nachvollziehbar und die Prozesse transparent werden [LSR19; SMR20]. Es kann angenommen werden, dass ein besseres KI-Verständnis bewusstere Entscheidungen im Umgang mit KI-Technologien nach sich zieht. Da von angehenden Lehrkräften fächerübergreifend erwartet wird, sowohl einen Beitrag zur Entwicklung der Digitalisierung als auch zur Förderung der Medienkompetenz der Schülerinnen und Schüler zu leisten (bspw. gem. KMK-Strategie und der Gesellschaft für Fachdidaktik), sind Maßnahmen zu eruieren, wie diese zukünftigen Multiplikatoren zunächst ein eigenes Verständnis von KI-Technologien erwerben können. Derzeit in der Umsetzung befindliche Projekte zur Lehrerbildung und Digitalisierung, wie im Rahmen der Qualitätsoffensive Lehrerbildung, widmen solchen Potenzialen und Herausforderungen Künstlicher Intelligenz noch keine Aufmerksamkeit.

3 KI-Technologieentwicklung und Bildung gemeinsam denken

Im interdisziplinären Projekt PetraKIP werden die skizzierten Herausforderungen und Chancen zusammengeführt, um KI als unterstützende Technologie für die Lehrerbildung näher zu untersuchen. Ziel ist die interdisziplinäre Erforschung von KI-Methoden (insbesondere Textanalyse, Maschinelles Lernen, Benutzermodellierung, kontrollierter Dialog) mittels der anwendungsbezogenen Adaption, Implementation und Evaluation des bereits in der Lehrerbildung implementierten mobilen persönlichen Portfolio-Systems *Metapholio* [Pe19], erweitert um Künstliche Intelligenz. KI-Methoden sollen hierbei die individualisierte Lernunterstützung optimieren, aber auch transparent gemacht werden, um so als Reflexionsgegenstand mit dem Ziel der KI-Bildung und Mündigkeit in Bezug auf KI-Systeme zu fungieren. Dies soll erreicht werden durch die folgenden Entwicklungs- und Forschungsschwerpunkte:

- **Analyse von Reflexionstexten mit einem Mixed-Methods Ansatz:** Basierend auf einem Datensatz von digitalen Reflexionstexten von Informatik- und Lehramtsstudierenden werden inhaltliche, formale und linguistische Merkmale sowie die Qualität der schriftlichen Reflexionen nach einer Klassifikation von Hatton und Smith [HS95] mit qualitativen und quantitativen sozialwissenschaftlichen Methoden sowie mit verschiedenen maschinellen Analyseverfahren ausgewertet, um diese Ergebnisse für die Entwicklung der KI-basierten Portfolio-App sowie der adaptiven Feedback-Funktionen zu nutzen.
- **Studierendenbegleitung durch Personal Coaching (geleitete Reflexion):** Volltext-Reflexionen werden durch die Anwendung mehrerer NLP-Modelle (bspw. Klassifikationen, Word-Embeddings, Sentimentanalysen) in einen Vektorraum eingebettet bzw. auf eine Reihe von Zielvariablen abgebildet. Dies ermöglicht es zu „verstehen“, wie tief ein Studierender das Thema reflektiert hat, welche Stimmungslagen vorliegen und welche Analogien zu Reflexionen von anderen Studierenden oder zu eigenen früheren Erfahrungen vorliegen. Aus den gewonnenen Erkenntnissen sollen individualisierte Empfehlungen und Fragestellungen zur vertieften und vergleichenden Reflexion sowie motivationsunterstützende Maßnahmen abgeleitet bzw. initiiert werden.
- **Studierendenbegleitung durch Personal Mentoring (Orientierung und Entwicklungsanstöße):** Volltext-Reflexionen und zusätzliche inhaltsbezogene Dokumente werden durch die Anwendung mehrerer NLP-Modelle (bspw. Klassifikation, Named-Entity-Recognition, Topic Modelling) analysiert, um auf Grundlage von Studierendenideen, -reflexionen und kursspezifischer Literatur relevante und motivierende Inhalte (bspw. für Unterrichtseinstiege) vorzuschlagen, basierend auf positiv reflektierten Themen, Angebote für die persönliche Weiterentwicklung vorzuschlagen, professionsbezogene relevante Kompetenzdefizite zu ermitteln und fachübergreifende sowie -spezifische Lernprozessunterstützung und Entwicklungsanstöße zu geben.
- **Dozierendenunterstützung bei der Individualisierung der universitären Lehre:** Dozierende erhalten (durch Aggregation) Feedback über studentische

Lernprozesse, um bspw. auf geringe Lernmotivation und Studienzufriedenheit oder potentielle Studienabbruchneigungen rechtzeitig reagieren zu können. Hierzu werden Auswertungen der affektiven Auswirkungen der Lehrveranstaltungsinhalte und Leistungsrückmeldungen vorgenommen, individuelle Studierendenreflexionen indikatorbasiert ausgewertet, bei Bedarf adäquate Hilfsangebote (bspw. Studienberatung) vorgeschlagen und Lösungsmöglichkeiten zur Vereinbarkeit von Datenschutz und Profiling gefunden und transparent gemacht werden.

- **Strukturelle Integration von KI-Bildung bei Lernenden und Lehrenden:** Studierende und Dozierende wenden KI-Technologien nicht nur an, sondern erwerben am Beispiel des KI-Portfolios ein Verständnis der Grundprinzipien Künstlicher Intelligenz. Hierzu werden angewandte Verfahren transparent dargestellt, in kleinen überschaubaren Lerneinheiten erklärt, zugrundeliegende ethisch-rechtliche Fragestellungen der KI-basierten Analysen formuliert, diskutiert und auf Basis eines erworbenen Grundverständnisses konfident beantwortet (Privatheit von Daten, Datensicherheit, Voreingenommenheit, Transparenz, Erklärbarkeit, Manipulierbarkeit, usw.).

4 Diskussion

Die starke Verbreitung von Large Language Models wie ChatGPT hat gezeigt, dass KI-Innovationen Bildungssysteme rasant verändern und (zukünftige) Lehrkräfte unvorbereitet treffen können. Entsprechend ist es absolut notwendig im Studium nicht nur produktive, reflektierte Lernerfahrungen *mit* KI zu erwerben, sondern auch Grundkenntnisse *über* KI zu erlernen. Künstliche Intelligenz sollte also nicht nur als unterstützendes Instrument in der Hochschulbildung, sondern auch als Bildungsgegenstand für die involvierten Personen adressiert, und somit auch mit Blick auf deren motivationale und volitionale Lernvoraussetzungen und -prozesse für den Einsatz von KI-Systemen berücksichtigt werden.

Literaturverzeichnis

- [Ch18] Chiriatti, G. et al.: A NLP-based Analysis of Reflective Writings by Italian Teachers, Computational Linguistics CLiC-it 2018, S. 118-124, 2018.
- [CYZ16] Chen, Y.; Yu, B.; Zhang, X.; Yu, Y.: Topic modeling for evaluating students' reflective writing: A case study of pre-service teachers' journals. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. S. 1-5, 2016.
- [FC18] Feder, L.; Cramer, C.: Potenziale von Portfolioarbeit in der Lehrerbildung. Eine Analyse der Portfolioliteratur, DDS – Die Deutsche Schule, 110(4), S. 354–368, 2018.
- [Gi17] Gibson, A. et al.: Reflective writing analytics for actionable feedback. In: Proceedings of the seventh international learning analytics & knowledge conference. S. 153-162, 2017.

- [GVR10] Gläser-Zikuda, M.; Voigt, C.; Rohde, J.: Förderung selbstregulierten Lernens bei Lehramtsstudierenden durch portfoliobasierte Selbstreflexion. *Lerntagebuch und Portfolio aus empirischer Sicht*, S. 142-163, 2010.
- [Ha18] Handke, J.: Digitale Hochschullehre–Vom einfachen Integrationsmodell zur Künstlichen Intelligenz. In U. Dittler & C. Kreidl (Hrsg.), *Hochschule der Zukunft. Beiträge zur zukunftsorientierten Gestaltung von Hochschulen*, Springer VS, S. 249-263, 2018.
- [HS95] Hatton, N.; Smith, D.: Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11(1), S. 33–49, 1995.
- [HWK16] Hofmann, F.; Wolf, N.; Klauf, S.; Grassmé, I.; Gläser-Zikuda, M.: Portfolios in der LehrerInnenbildung. Ein aktueller Überblick zur empirischen Befundlage. In: M. Bos, A. Krämer & M. Kricke (Hrsg.), *Portfolioarbeit phasenübergreifend gestalten. Konzepte, Ideen und Anregungen aus der LehrerInnenbildung*. Waxmann, Münster & New York, S. 23-39, 2016.
- [K117] Klebanov, B. B.; Burstein, J.; Harackiewicz, J. M.; Priniski, S. J.; Mulholland, M.: Reflective Writing About the Utility Value of Science as a Tool for Increasing STEM Motivation and Retention–Can AI Help Scale Up?. *International Journal of Artificial Intelligence in Education*, 27(4), S. 791-818, 2017.
- [Ko19] Korthagen, F.: Linking reflection and technical competence: The logbook as an instrument in teacher education. *Europ. Jour. of Teacher Education* 22(2/3), S. 191-207, 1999.
- [LR19] Lindner, A.; Romeike, R.: Teachers' Perspectives on Artificial Intelligence. In: Pozdniakov, S., Dagiene, V. (Eds.), *Proceedings of ISSEP 2019*, S. 22 – 29, 2019.
- [LSR19] Lindner, A.; Seegerer, S.; Romeike, R.: Unplugged Activities in the Context of AI. In: *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives*, Springer, S. 123-135, 2019.
- [MO14] Mueller, P. A.; Oppenheimer D. M.: The pen is mightier than the keyboard: Advantages of longhand over laptop note-taking. *Psych. Science* 25(6), S. 1159-1168, 2014.
- [Pe19] Petko, D.; Schmid, R.; Müller, L.; Hielscher, M.: Metapholio: A mobile app for supporting collaborative note taking and reflection in teacher education. *Technology, Knowledge and Learning*, 24(4), S. 699-710, 2019.
- [Sa17] Satow, L.: Chatbots as Teaching Assistants: Introducing a Model for Learning Facilitation by AI Bots, 2017. Online: <https://blogs.sap.com/2017/07/12/chatbots-as-teaching-assistants-introducing-a-model-for-learning-facilitation-by-ai-bots/>, Stand: 05.05.2020.
- [SLF19] Schmohl, T.; Löffl, J.; Falkemeier, G.: Künstliche Intelligenz in der Hochschullehre. In: T. Schmohl; D. Schäffer (Hrsg.), *Lehrexperimente in der Hochschulbildung. Didaktische Innovationen aus den Fachdisziplinen*. wbv, Bielefeld, S. 117-122, 2019.
- [SMR20] Seegerer, S.; Michaeli, T.; Romeike, R.: So lernen Maschinen. *LOG IN - Informatische Bildung und Computer in der Schule*, 193-194, S. 25-29, 2020.
- [Ul15] Ullmann, T. D.: Keywords of written reflection-a comparison between reflective and descriptive datasets. In: *CEUR Workshop Proceedings 1465*, S. 83-96, 2015.
- [ZG16] Ziegelbauer, S.; Gläser-Zikuda, M. (Hrsg.): *Portfolio als Innovation in Schule, Hochschule und LehrerInnenbildung*. Waxmann, Münster, 2016.

Ein Projektkurs Data Science und Künstliche Intelligenz für die Sekundarstufe II

Untertitel des Beitrages

Yannik Fleischer¹, Sven Hüsing², Rolf Biehler³, Lukas Höper⁴, Susanne Podworny⁵, Carsten Schulte⁶

1 Einleitung

Wir berichten in diesem Beitrag vom Konzept und der Durchführung eines einjährigen Data-Science-Kurses für die Sekundarstufe II. Dieser Kurs wurde im Rahmen des von der Deutsche Telekom Stiftung initiierten und geförderten Projekts Data Science und Big Data in der Schule (ProDaBi, www.prodabi.de) entwickelt. Dabei handelt es sich um ein interdisziplinäres Projekt der Didaktik der Mathematik und der Didaktik der Informatik an der Universität Paderborn.

Im Schuljahr 2022/2023 wurde der sogenannte Projektkurs “Data Science und Künstliche Intelligenz” zum fünften Mal durchgeführt, dessen Konzeption in dieser Zeit ein Entwicklungsziel von ProDaBi darstellte. Gleichzeitig wird der Kurs weiterhin als eine Art „Entwicklungslabor“ für die Erprobung von kleineren Unterrichtsmodulen und -materialien zum Thema Data Science genutzt, die im Rahmen des Projektkurses eingesetzt und evaluiert werden. Der Kurs wird in Kooperation mit zwei Paderborner Schulen organisiert und durchgeführt, wobei er von allen interessierten Schüler:innen beider Schulen angewählt werden kann, die zuvor Informatik in der Jahrgangsstufe 11 gewählt haben.

2 Projektkurskonzept (Ziel und Vorgehen)

Im Projektkurs durchlaufen die Schüler:innen im Laufe eines Schuljahres ein authentisches Datenprojekt - von der ersten Arbeit mit Rohdaten bis zu einem Endprodukt. Das didaktische Konzept des Kurses sieht dabei sowohl eher offene Phasen als auch stärker angeleitete Phasen vor. In den offenen Phasen findet eine durch die Schüler:innen

¹ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, yanflei@math.upb.de,
<https://orcid.org/0000-0003-0318-0329>

² Universität Paderborn, Warburger Str. 100, 33098 Paderborn, sven.huesing@uni-paderborn.de

³ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, biehler@math.upb.de

⁴ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, lukas.hoepfer@upb.de

⁵ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, podworny@math.upb.de

⁶ Universität Paderborn, Warburger Str. 100, 33098 Paderborn, carsten.schulte@upb.de

selbst gesteuerte Arbeit am Datenprojekt statt, während sie in den angeleiteten Phasen Kenntnisse und Fertigkeiten in Bezug auf Datenexploration, maschinelles Lernen und datengetriebenes Programmieren mit Python erwerben.

Das Datenprojekt zieht sich dabei von Anfang an als “roter Faden” durch den Projektkurs und bietet verschiedene Lernanlässe. An geeigneten Stellen werden dann die thematisch passenden Unterrichtsmodule durchgeführt, um anschließend mit dem neu erworbenen Wissen das Datenprojekt fortführen zu können. Auf diese Weise wird einem “Vorratslernen” vorgebeugt. Zwei Unterrichtsmodule befassen sich mit dem Thema der Datenexploration, während zwei andere maschinelles Lernen (Entscheidungsbäume, Künstliche Neuronale Netze) als Thema haben. Ein weiteres Modul befasst sich einleitend mit dem Programmieren in Python.

Im Kurs werden zwei digitale Werkzeuge eingesetzt. Das erste ist die webbasierte interaktive Data-Science-Plattform CODAP (codap.concord.org), mit der die Schüler:innen Daten in einer menügesteuerten Umgebung intuitiv erkunden können. Das zweite Tool stellen Jupyter Notebooks (JNBs) [To17] dar, die als Programmierumgebung und Werkzeug mit der Programmiersprache Python für die Durchführung und Dokumentation von Datenprojekten genutzt werden. Neben der Erstellung eigener Dokumente durch die Schüler:innen werden zuvor erstellte CODAP- und JNB-Dokumente auch als Lernumgebungen für die Schüler:innen genutzt. Durch diese sogenannten Worked Examples [At00] sollen die Schüler:innen darin unterstützt werden, Datenexplorationen und maschinelles Lernen eigenständig umsetzen zu können.

Im Verlauf des Kurses erstellen die Schüler:innen zwei sogenannte Computational Essays [Di00] in Form von JNBs. Dabei dokumentieren sie ihr Vorgehen als interaktive Datenanalyseberichte, indem sie Code, Programmiererergebnisse und erläuternden Text in einem adaptierbaren Dokument festhalten. Im ersten Computational Essay wird die Datenexploration dokumentiert und das zweite bezieht sich auf die Dokumentation eines ersten Vorhersagemodells, das auf maschinellem Lernen basiert. Die beiden Computational Essays sind essentieller Teil des Endproduktes. Weiterhin besteht das Endprodukt in jedem Jahr aus einer Web-App (einschließlich eines auf maschinellem Lernen basierenden Entscheidungsmodells) und einer Präsentation, in der die genutzten Daten und Methoden sowie die Web-App erläutert und präsentiert werden.

3 Projektergebnisse

Die Schüler:innen können sich den Kontext ihres Datenprojektes selbst aussuchen. In den ersten vier Jahren arbeiteten die Kurse jeweils an einer KI, die die Parkplatzbelegung auf Paderborner Parkplätzen vorhersagen konnte. Dafür standen Parkbelegungsdaten über 12 Jahre zur Verfügung. Das Hauptziel des Projekts bestand hierbei darin, mit Hilfe von maschinellem Lernen datenbasierte Modelle zur Vorhersage der Belegung verschiedener Parkplätze in den kommenden Stunden zu entwickeln und zu dokumentieren. Der Zweck

war Menschen dabei zu helfen, schneller einen freien Parkplatz zu finden, um den Verkehrsfluss zu verbessern und Emissionen zu reduzieren.

Im jüngsten Projektkurs haben die Schüler:innen hingegen eine KI zur Erkennung verschiedener Begriffe der Gebärdensprache entwickelt. Dafür nahmen sie selbst mehrere tausend Bilder von Handgesten auf und trainierten damit eine künstliche Intelligenz mit Methoden des maschinellen Lernens, um ein Modell zu erstellen, das diese Gesten erkennt. Anhand der Projektergebnisse lässt sich erkennen, dass der Projektkurs zu verschiedenen datenbasierten Projekten anregt und Schüler:innen dazu befähigt, mithilfe von Daten gesellschaftlich relevante Projekte anzugehen.

Literatur

- [At00] Atkinson, R. K. et al.: Learning from Examples: Instructional Principles from the Worked Examples Research. *Review of Educational Research*, 70(2), 181–214, 2000.
- [Di00] DiSessa, A. A.: *Changing minds: Computers, learning, and literacy*. MIT Press, 2000.
- [To17] Toomey, D.: *Jupyter for Data Science – Exploratory analysis, statistical modeling, machine learning, and data visualization with Jupyter*. Packt Publishing, 2017.

KI lehren, KI lernen an der DSHH – Potentiale für die Bildungslandschaft

Malte Prieß ¹ und Kerstin Prechel ²

Abstract: Das Forschungsprojekt zielt darauf ab, den bildungspraktischen Bedarf im Bereich künstlicher Intelligenz (KI) zu erfüllen und die KI-Lehre in der Bildungsforschung zu stärken. Ziel ist der Ausbau von Lernressourcen und didaktischen Konzepten zum Themenkomplex KI an der Dualen Hochschule Schleswig-Holstein (DSHH) und damit ein Beitrag zur Digitalisierung an der Hochschule, indem Lehrende den Umgang mit digitalen Lehr-Formaten und Studierende (unterstützt durch KI-Campus Ressourcen) KI-Einsatzkontexte erlernen. Es wird ein Lehrkonzept für ein aktuell laufendes Grundlagenmodul zu Data Science und Maschinellem Lernen dargelegt, das Ansätze des Blended-Learning und öffentlich zugängliche Ressourcen nutzt. Um zu validieren, welche Elemente und Strukturen gut angenommen werden, ist im Nachgang eine differenzierte Evaluation seitens der Studierenden geplant. Aktuelle Beobachtungen aus dem laufenden Kurs wie die Veränderung der Rolle der Lehrenden sowie die Chancen und Herausforderungen für die Lernenden werden als Forschungshypothesen für eine anschließende Validierung und Konkretisierung dargelegt.

Keywords: Künstliche Intelligenz, Digitalisierung, Lernressourcen, Blended-Learning

1 Einleitung

Die Digitalisierung der Wissensvermittlung hat die Bildungslandschaft grundlegend verändert und stellt Hochschulen vor neue Herausforderungen [Me19]. Insbesondere die Lehre und das Lernen von Künstlicher Intelligenz spielen eine entscheidende Rolle in diesem Wandel. Die Einführung von Plattformen für Massive Open Online Courses (MOOCs) im Jahr 2012 war eng mit dem Ausbau der KI-Lehre verbunden [Ro12]. Seitdem wurden zahlreiche Online-Kurse zu KI entwickelt [Hi23; HR23; Ng23].

In diesem Kontext zielt unser Forschungsprojekt darauf ab, den bildungspraktischen Bedarf im Bereich KI systematisch zu erfüllen und die KI-Lehre an der Dualen Hochschule Schleswig-Holstein (DSHH) zu stärken. Dabei stellen sich uns folgende Forschungsfragen: Wie kann die Lehre und das Lernen von KI-Kompetenzen an der DSHH systematisiert und verbessert werden? Wie können digitale Lehrformate und KI-Einsatzkontexte effektiv in den Lehrplan integriert werden? Welche Rolle spielen Online-Kurse in diesem Prozess und wie werden sie von den Studierenden angenommen? Basierend auf diesen Forschungsfragen formulieren wir die folgenden Hypothesen:

¹ DSHH, Dekan Wirtschaftsinformatik, Hans-Detlev-Prien-Straße 10, 24106 Kiel malte.priess@dhsh.de

² DSHH, Dekanin Betriebswirtschaftslehre, Hans-Detlev-Prien-Straße 10, 24106 Kiel kerstin.prechel@dhsh.de

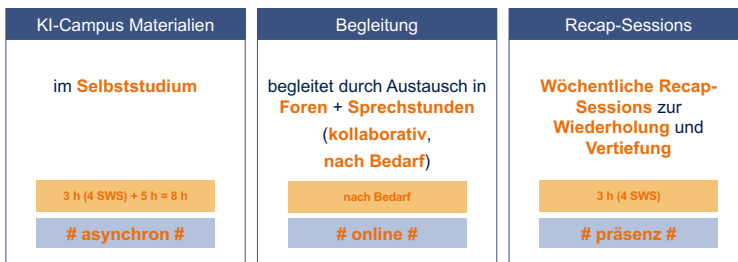


Abb. 1: Struktureller Aufbau und Workload des Moduls

H1: Die Integration von Online-Kursen in den Lehrplan verbessert das Lernen von KI-Kompetenzen bei den Studierenden.

H2: Die Nutzung digitaler Lehrformate und KI-Einsatzkontexte trägt zur Digitalisierung an der DSHS bei und wird positiv von den Lehrenden aufgenommen.

H3: Lehrende werden durch die Integration von Online-Kursen zu Lernbegleitern.

Um diese Hypothesen zu überprüfen, haben wir ein Lehrkonzept für ein Grundlagenmodul zu Data Science und Maschinellem Lernen entwickelt und implementieren es aktuell, das auf modernen Blended-Learning-Methoden und öffentlich zugänglichen Lernressourcen basiert. Darüber hinaus werden wir die Akzeptanz und Wirksamkeit dieses Ansatzes durch eine detaillierte Evaluation seitens der Studierenden überprüfen.

2 Auswahl der Bausteine

Um den geplanten Einstieg in das Thema KI, ML und Data Science zu unterstützen, bieten sich verschiedene Phasen sowie dazu passend ausgewählte Einheiten des KI-Campus³ an (s. Abb. 2). Die Integration dieser Ressourcen verringert die Arbeitslast in der Materialerstellung, wodurch es möglich wird, eine enge Abstimmung und inhaltliche Begleitung der KI-Campus-Angebote auf die eigene Lehrveranstaltung und die spezifischen Anforderungen des dualen Studiums, das fortlaufend den Rückbezug des vermittelten Wissens auf die praktischen Anforderungen in den Partnerunternehmen zum Ziel hat, sicherzustellen. Dies hat für die Studierenden den Nutzen, Informationen auf unterschiedlichen Kanälen zu erhalten, um Wissen besser zu verfestigen und zudem ein Hinterfragen des Erlernten zu ermöglichen.

3 Umsetzung

Im Sinne eines Blended-Learning-Ansatzes ist das Modul grundsätzlich in Phasen des Selbststudiums sowie in Präsenzeinheiten zur Wissenssicherung strukturiert (s. Abb. 1).

³ Der KI-Campus ist eine Lernplattform für künstliche Intelligenz (KI-Campus: Über den KI-Campus, <https://ki-campus.org/about>, Stand: 14.7.2023)

Woche	Selbststudium	Recap
Woche 1	Phase 1: Kick-Off + Überblick KI ▪ Kick-Off ▪ Künstliche Intelligenz in 2 Minuten erklärt: Was ist eigentlich KI? ▪ deeplearning.ai: AI for Everyone	
Woche 2	Phase 2a: Grundlagen des ML ▪ Data Literacy – Daten interpretieren durch Data Mining (Modul 1 + 2)	Recap Phase 1
Woche 3	Phase 2b: Grundlagen des ML ▪ Data Literacy – Daten interpretieren durch Data Mining (Modul 3 + 4)	Recap Phase 2a
Woche 4	Phase 2c: Grundlagen des ML ▪ Data Literacy – Daten interpretieren durch Data Mining (Modul 5 + 6)	Recap Phase 2b
Woche 5	Phase 2d: Grundlagen des ML ▪ Data Literacy – Daten interpretieren durch Data Mining (Modul 7 + 8)	Recap Phase 2c
Woche 6	<i>Puffer</i>	Recap Phase 2d
Woche 7	Phase 3: Ausblick auf KNNs, DL ▪ Videoserie Neural Networks ▪ Deep Learning mit TensorFlow, Keras und TensorFlow.js	<i>Puffer</i>
Woche 8	Phase 4: Gesellschaftliche Folgenabwägung, kritische Reflexion der Technologien ▪ Daten- und Algorithmenethik	Recap Phase 3
Woche 9	<i>Puffer</i>	Recap Phase 4
Woche 10	<i>Prüfungswoche</i>	

Abb. 2: Lehrphasen und inhaltliche Feingliederung des Moduls

Weiterhin dienen ein kollaborativer Austausch über Foren (der Lernplattform Moodle), Padlets⁴ sowie nach Bedarf zu vereinbarende Sprechstunden der Sicherung der erlernten Inhalte. Die einzelnen in Abbildung 2 angegebenen Phasen seien nun näher beschrieben:

- *Phase 1 – Kick-off und Überblick KI:* Im Rahmen einer Kick-Off-Veranstaltung in Präsenz bieten sich ausgewählte Einheiten zur Aktivierung und Motivation der Studierenden an. Im Anschluss werden die Studierenden mit den grundlegenden Bedeutungen der KI-Terminologie vertraut gemacht. Hierbei liegt der Fokus darauf, einen Eindruck zu vermitteln, was KI zu leisten vermag und wie sie zur Lösung von Problemen in der eigenen Organisation eingesetzt werden kann. Um eine sinnvolle Verknüpfung der digitalen Lerninhalte mit Präsenzeinheiten zu erreichen, werden – für diese und folgende Phasen – die asynchronen Lerneinheiten durch zeitlich kompakte Recap-Einheiten in Präsenz begleitet (s. Abb. 1 und 2). Diese tragen zur Verstetigung von angeeignetem Wissen sowie zur Fähigkeit der Studierenden bei, dieses Wissen selbstständig auf unbekannte Sachverhalte übertragen zu können.
- *Phase 2 – Vertiefung ML:* Ziel ist es, etablierte Techniken und Algorithmen sowie ausgewählte theoretischen Grundlagen im Bereich des ML zu vermitteln (etwa Datenvorbereitung, Modellbildung, Evaluation, Test- und Validierung, Merkmalsauswahl, Klassifikation, Regression, Clusterverfahren, Assoziationsanalyse, Gütemaße). Zudem lernen die Studierenden, diese Techniken praktisch zu implementieren und sie schnell und wirksam auf neue Probleme anzuwenden.
- *Phase 3 – Ausblick Künstliche Neuronale Netze (KNN), Deep Learning (DL):* Diese Phase zielt auf einen kurzen Ausblick auf den Bereich KNN. Zum Einsatz kommen

⁴ Padlets sind visuelle Boards für das Organisieren und Teilen von Inhalten (Padlet: So funktioniert's, <https://padlet.com/site/product/howitworks>, Stand: 14.07.2023)

ausgewählte Inhalte aus dem Kurs *Deep Learning mit TensorFlow, Keras und TensorFlow.js, KI-Campus-Original*, etwa die kompakte Einführung in DL.

- *Phase 4 – Gesellschaftliche Folgenabwägung, kritische Reflexion*: Um den Studierenden eine gesellschaftliche Folgenabwägung und kritische Reflexion der eingesetzten Technologien nahezubringen, sollen ausgewählte Teile aus der Einheit *Daten- und Algorithmenethik* eingesetzt werden.

4 Fazit und Ausblick

Unser Forschungsprojekt hat einen Rahmen vorgeschlagen, wie die Lehre und das Lernen von KI-Kompetenzen an der Dualen Hochschule Schleswig-Holstein systematisiert und verbessert werden können. Die Integration von Online-Kursen in den Lehrplan und die Nutzung digitaler Lehrformate und KI-Einsatzkontexte kann sich als effektive Strategien erweisen, um die Digitalisierung voranzutreiben und das Lernen von KI-Kompetenzen bei den Studierenden zu verbessern. Unsere Hypothesen werden durch die Ergebnisse unserer Studie, die nach Abschluss des ersten Durchlaufs stattfinden wird, geprüft. Insbesondere wird sich zeigen, wie sich die Rolle der Lehrenden durch die Integration von Online-Kursen hin zu Lernbegleitern verändern wird. Es gibt noch viele offene Fragen und Herausforderungen, die in zukünftigen Forschungsarbeiten angegangen werden müssen. Insbesondere ist es wichtig, die langfristigen Auswirkungen unserer Lehrstrategien auf die Lernergebnisse und Karriereaussichten der Studierenden zu untersuchen. Darüber hinaus sollten weitere Studien durchgeführt werden, um zu untersuchen, wie digitale Lehrformate und KI-Einsatzkontexte am besten in verschiedene Arten von Kursen und Studiengängen integriert werden können.




In Bezug auf die nächsten Schritte planen wir, das Lehrkonzept zu verfeinern und zu erweitern, basierend auf dem Feedback und den Erfahrungen der Studierenden und Lehrenden.

Literatur

- [Hi23] Hill, K.: AI for Executives, URL: <https://www.experfy.com/training/courses/ai-for-executives>, Stand: 14. 07. 2023.
- [HR23] University of Helsinki, Reaktor: Elements of AI - join the movement!, URL: <http://www.elementsofai.com/eu2019fi>, Stand: 14. 07. 2023.
- [Me19] Meyer-Guckel, V.; Klier, J.; Kirchherr, J. W.; Winde, M.: Future Skills: Strategische Potenziale für Hochschulen (Diskussionspapier Nr. 3), 2019.
- [Ng23] Ng, A.: AI For Everyone, Original work published, URL: <https://www.deeplearning.ai/ai-for-everyone/>, Stand: 14. 07. 2023.
- [Ro12] Rodriguez, C. O.: MOOCs and the AI-Stanford Like Courses: Two Successful and Distinct Course Formats for Massive Open Online Courses. European Journal of Open, Distance and E-Learning 15/02, 2012.

Entwicklung interdisziplinärer Module in der Hochschulbildung

Synthese von Künstlicher Intelligenz und Ingenieurwissenschaften

Stefanie Krause ¹, Simon Adler ², Johannes Bühl, René Schenkendorf, Kerstin Schneider, Frieder Stolzenburg ³ und Fabian Transchel

Abstract: Um der zunehmenden Bedeutung der künstlichen Intelligenz (KI) für Unternehmen Rechnung zu tragen, entwickelt die Hochschule Harz interdisziplinäre Module im Rahmen des kooperativen Studiengangs AI Engineering – Künstliche Intelligenz in den Ingenieurwissenschaften, unter anderem in den Bereichen Mobile Systeme und Telematik. Im vorliegenden Beitrag wird ein Einblick in die geplante Umsetzung für ausgewählte Module gegeben.

Keywords: Domänenspezifische KI-Bildung, Mobile Robotik, Telematik, Ingenieurwissenschaften, Industrieller Einsatz von KI

1 Einleitung

Spätestens seit den Entwicklungen im Kontext der Industrie 4.0 hat sich der Blick auf die Betriebs- und Prozessdaten geändert. Die Analyse von Betriebsdaten ermöglicht unter anderem die Optimierung der Anlagen und der Arbeitsprozesse. Die aktuelle Entwicklung von KI-Methoden sowie Hardware und Programmierparadigmen zu deren parallelisierter Berechnung, ermöglichen immer umfangreichere Analysen. Die Analyse technischer Systeme wird häufig dadurch erschwert, dass Expert*innen im Bereich der Datenwissenschaften und Künstlichen Intelligenz das Grundlagenwissen der Anwendungsdomäne fehlt und andererseits Expert*innen der Automatisierung zunehmend komplexere Anforderungen bei der Bereitstellung von Daten zu berücksichtigen haben. Die richtige Erstellung, Pflege, Dokumentation und Bereitstellung von Daten im Produktlebenszyklus technischer Anlagen hat signifikanten Einfluss auf den Aufwand bei der Datennutzung in der späteren Betriebsphase [Lö18]. Darüber hinaus stellen sich organisatorische Herausforderungen bei der Integration datengetriebener

1 Hochschule Harz, Fachbereich Automatisierung und Informatik, Friedrichstraße 57-59, 38855 Wernigerode, skrause@hs-harz.de, <https://orcid.org/0000-0002-1271-7514>

2 Hochschule Harz, Fachbereich Automatisierung und Informatik, Friedrichstraße 57-59, 38855 Wernigerode, simonadler@hs-harz.de, <https://orcid.org/0000-0002-9653-9005>

3 Hochschule Harz, Fachbereich Automatisierung und Informatik, Friedrichstraße 57-59, 38855 Wernigerode, fstolzenburg@hs-harz.de, <https://orcid.org/0000-0002-4037-2445>

Arbeitsweisen im Gesamtprozess [KT22]. Des Weiteren profitieren insbesondere Projekte erheblich von Domänenwissen [We22] [Da22].

Es sind daher neue Studienangebote zur Ausbildung von Fachkräften erforderlich, die Methoden der Anwendungsdomänen synergetisch mit denen der Datenanalyse adressieren. Während bisher KI-Themen vorwiegend als Vertiefung angeboten werden, ist hierfür eine modulübergreifende Verzahnung der Lerninhalte erforderlich, um die wechselseitigen Auswirkungen von KI, Fachwissen der Domänen und deren Daten zu vermitteln. Aus dieser Anforderung entstand die Initiative für den hochschulübergreifenden Studiengang *AI Engineering*: Ein interdisziplinärer, projektorientierter Bachelorstudiengang der synergetisch Künstliche Intelligenz und Ingenieurwissenschaften verbindet.

2 Interdisziplinäres Studium

Das Projekt *AI Engineering* [AI23] bündelt die Kompetenzen der Otto-von-Guericke Universität sowie der Hochschulen Anhalt, Harz, Magdeburg-Stendal und Merseburg in einem interdisziplinären Bachelorstudiengang. Studierende werden hier von Anfang an die Grundlagen der Ingenieurwissenschaften und der KI im Zusammenhang lernen. Jede Hochschule bringt die Expertise in je einem ingenieurtechnischen Schwerpunkt ein, aus dem die Studierenden anschließend wählen können. Die Hochschule Harz repräsentiert hierbei den Schwerpunkt *Mobile Systeme und Telematik*. Studierende erlernen u.a. den Einsatz von KI für die dynamische Prozesssteuerung, intelligente mobile Robotik, Fahrsimulation, Telematik, Recommender Systeme sowie Methoden des Data Engineering für mobile Assistenzsysteme und digitaler Zwillinge technischer System [AE20].

Die Rückmeldungen aus der Industrie wurden während der Planung des Curriculums (Abb. 1) berücksichtigt. Studierende lernen Methoden und deren praktische Anwendung unter Berücksichtigung gesellschaftlicher und sozialer Verantwortung. Sie werden die zielgerichtete Datenanalyse für den problemspezifischen Einsatz grundlegender statistischen Verfahren bis hin zu komplexen KI-Architekturen beherrschen.

Eine Übersicht des Curriculums der Vertiefung *Mobile Systeme und Telematik* ist in Abb. 1 dargestellt. Im Folgenden wird eine Auswahl der Module kurz vorgestellt.

Grundlagenstudium an der OVGU				Vertiefung Mobile Systeme und Telematik (HS Harz)			
	Semester 1	Semester 2	Semester 3	Semester 4	Semester 5	Semester 6	Semester 7
SCP	Technische Darstellungslehre	Technische Mechanik 1	Technische Mechanik 2 / 3	Synergetisches Grundlagenmodul 1	Faszination KI in mobilen Systemen	Data Engineering und anpassungsfähige intelligente Systeme	
SCP	Mathematik 1a für Ingenieure	Mathematik 1b für Ingenieure	Mathematik 2 für Ingenieure	Synergetisches Grundlagenmodul 2	Mobile Assistenzsysteme im Betrieb technischer Anlagen	Hybrides maschinelles Lernen - Wissens- und datenbasierte Modelle techn. Systeme	Fachpraktikum mit Reflexionsseminar
SCP	Elektrotechnische Grundlagen	Messtechnik	Industrielle KI-Systeme	Signalverarbeitung	Mobile Systeme und Telematik	Mobile Robotik und KI-Methoden	
SCP	Data Engineering	Grundlagen des maschinellen Lernens	Deep Learning für Ingenieure	KI-basierte Überwachung und Steuerung von technischen Systemen und Prozessen	Wahlpflichtmodul	Wahlpflichtmodul	
SCP	Einführung in die Informatik für Ingenieure	BWL für Ingenieure	Software Engineering + IT-PM (7 CP)	Erklärbare und sichere KI	Wahlpflichtmodul	Interdisziplinäres Projekt im ingenieurwissenschaftlichen Schwerpunkt	Bachelorarbeit
SCP	Einführung ins AI Engineering	Projekt Prototyping von KI-Systemen	Projekt ML Programmierung	KI-Reflexion und Ethik	Projekt Modellentwicklung für technische Systeme	Projekt Modellbetrieb / MLOps	

Abb. 1: Skizze des Curriculums mit Vertiefung an der Hochschule Harz

2.1 Modul Mobile Robotik mit KI-Methoden

In der Pflege, der industriellen Fertigung und vielen anderen Bereichen spielen autonome, intelligente Roboter eine immer größere Rolle. Um solche Systeme zu entwickeln, muss die klassische Robotik mit Methoden der KI sinnvoll verknüpft werden. Im Rahmen des Moduls erlernen Studierende die Programmierung, Lokalisation und Navigation mit aktueller Robotertechnik. Diese Grundlagen werden mit KI-Methoden beispielsweise zur Erkennung von Hindernissen ergänzt. Studierende entwickeln Verständnis für Sensordaten und deren Datenaufbereitung, wie beispielsweise das Entfernen von Rauschen. Die Genauigkeit, die Qualität und der Umfang der verfügbaren Sensordaten ist abhängig vom Kontext des mobilen Robotersystems und muss von Beginn an bedacht werden und schließt Methoden zur Objekterkennung mittels Bild-KI und Bewegungssteuerung intelligenter autonomer Agenten ein.

2.2 Modul Mobile Assistenzsysteme im Betrieb technischer Anlagen

Daten technischer Anlagen bestehen aus Signaldaten von Sensoren, Aktoren, internen Zuständen, Fehlermeldungen sowie aus Dokumentationen und Strukturdaten. Eine zielführende Datenanalyse erfordert zudem ein detailliertes Prozessverständnis des Anlagenbetriebs. Zudem muss berücksichtigt werden, dass Anlagen instandhaltungsbedingt einem kontinuierlichen Wandel unterliegen.

Studierende lernen, wie Informationen vor Ort ergänzt, rollen-, situations- und ortsgerecht bereitgestellt und in betriebliche Datenmodelle integriert werden können. Sie lernen, wie einerseits KI-Modelle Betriebsentscheidungen beeinflussen und wie KI-Modelle im Anlagenlebenszyklus erfasst und gepflegt werden müssen.

3 Fazit

In der Industrie werden dringend Fachkräfte benötigt, die neben einem Verständnis von KI ebenfalls Wissen über spezifische Anwendungsdomänen haben. In unserem hochschulübergreifenden Projekt erproben wir deshalb einen neuartigen Weg, KI in Synthese mit verschiedenen Fachdisziplinen zu betrachten. Als Hochschule für angewandte Wissenschaften verknüpfen wir unserer Erfahrungen im Bereich Data Science und KI mit domänenspezifische Anforderungen und Problemstellungen der Ingenieurwissenschaften. Aktuell werden dafür interdisziplinäre Vertiefungsmodule im Bereich Mobile Systeme und Telematik entwickelt. Wir versprechen uns damit eine höhere Fachkompetenz im Umgang mit digitalen Daten und eine Antwort auf die Anforderungen der Industrie.

Danksagung

„AI Engineering“ – Ein interdisziplinärer, projektorientierter Studiengang mit Ausbildungsschwerpunkt auf künstlicher Intelligenz und Ingenieurwissenschaften (AiEng) wird durch das Bundesministerium für Bildung und Forschung (BMBF, FKZ: 16DHBKI010) gefördert.

Literaturverzeichnis

- [AE20] Adler, S.; Eisenträger, M.; Bayrhammer, E.: Digitaler Zwilling: Effizient in Engineering und Betrieb. Zeitschrift für wirtschaftlichen Fabrikbetrieb, Walter de Gruyter GmbH, S. 425–428, 2020.
- [AI23] Webseite des Studiengangs AI-Engineering, <https://ai-engineer.de>, Abruf: 27.08.2023
- [Ad20] Adler, S. et al.: Digitale Assistenzsysteme für die Prozessindustrie - Ergebnisse und Erfahrungen aus dem Forschungsprojekt CPPSPProcessAssist. Herausgeber: Simon Adler, & Axel Eckstein. Fraunhofer Verlag, 2020.
- [Da22] Dash, T. et al.: A review of some techniques for inclusion of domain knowledge into deep neural networks. Scientific Reports 12.1, 2022.
- [Lö18] Löwen, U. et al.: Anwendungsszenario DDA: Durchgängiges und dynamisches Engineering von Anlagen. VDI Statusreport, 2018.
- [KT22] Kraut, N.; Transchel, F.: On the Application of SCRUM in Data Science Projects. 7th International Conference on Big Data Analytics (ICBDA), S. 1-9, 2022.
- [We22] Webster, Z.: Domain knowledge is necessary but not sufficient. Journal of AI, Robotics & Workplace Automation 1.3. S. 272-277, 2022.

Cybersecurity Privatsphäre - 4.
Interdisciplinary Privacy Security
at Large Workshop

Digitalisierung: Chancen nutzen, Risiken vermeiden

Annika Selzer¹, Ingo J. Timm²

Eine große Herausforderung der Digitalisierung ist es, einerseits die Chancen, die durch die Entwicklung und Nutzung neuer Technologien sowie durch die Verknüpfung und Auswertung von Daten entstehen, nicht gänzlich durch normative Vorhaben „auszubremsen“ sowie andererseits während des Einsatzes neuer Technologien und während Datenverknüpfungen und -auswertungen für eine angemessene Umsetzung sicherheitstechnischer und datenschutzrechtlicher Vorgaben zu sorgen. Mittels dieses Ansatzes des Austarierens zwischen Chancen und Risiken der Digitalisierung lassen sich einerseits große Mehrwerte für die Gesellschaft erreichen. U.a. können Datenverknüpfungen und neue Technologien dazu genutzt werden, eine drohende Überlastung der öffentlichen Daseinsvorsorge rechtzeitig zu erkennen und gegenzusteuern, die Bedarfe einer Gesellschaft an Kinderbetreuung, Schulungsangeboten oder Parkmöglichkeiten zu identifizieren und zu bedienen sowie den Ausbruch bzw. die Verbreitung ansteckender Krankheiten frühzeitig zu erkennen und mit angemessenen Maßnahmen entgegenzuwirken. Andererseits lassen sich mittels dieses Ansatzes durch das Mitdenken von Datenschutz- und IT-Sicherheitsmaßnahmen unangemessene Eingriffe in die Grundrechte und -freiheiten der Bürger sowie unberechtigte Zugriffe auf Betriebsgeheimnisse und ähnlich schützenswerte Informationen vermeiden bzw. auf ein angemessenes Maß reduzieren.

Der „4. Interdisciplinary Privacy & Security at Large“ Workshop widmet sich vor diesem Hintergrund dem interdisziplinären Austausch rund um die Chancen und Risiken der Digitalisierung. Neben technischen Aspekten sollen auch rechtliche, wirtschaftliche, wirtschaftsinformatische und ethische Aspekte des Themengebiets diskutiert werden.

Für die Vorschläge des geladenen Vortrags und die Begutachtung der eingereichten Beiträge danken wir den Mitgliedern des Programmkomitees Prof. Dr. Ricardo Büttner (Universität Bayreuth), Xengie Cheng DOAN (Universität Luxemburg), Erdem Durmus (NOTOS Xperts) Dr. Matthias Enzmann (Fraunhofer SIT, ATHENE), Prof. Dr. Dennis-Kenji Kipker (Universität Bremen, IGMR), Dr. Maximilian Schnebke (Universität Bremen, IGMR), Prof. Dr. Thomas Wilmer (Hochschule Darmstadt, ATHENE), Ruben Wolf (Fraunhofer SIT, ATHENE), und Dr. Daniel Woods (Universität Edinburgh).

Teile des Workshops wurden vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

¹ Fraunhofer SIT | ATHENE, annika.selzer@sit.fraunhofer.de.

² Universität Trier, Professur für Wirtschaftsinformatik, itimm@uni-trier.de.

Post Pandemic Follow-Up

How Data Sharing Can Be Reached in a Privacy-Friendly Way to Prepare Us for a Smart World

Annika Selzer¹ und Ingo J. Timm²

Abstract: Now that we (hopefully) enter the post-pandemic era, we urgently need to cope with the lessons learned during the COVID-19-pandemic, e.g., that we are not ready to ensure a privacy-friendly sharing of data which could help to reduce the spread of contagious diseases in the future. This paper reflects on this challenge and proposes a way for the privacy-friendly sharing of data.

Keywords: Anonymous Data, Data Sharing, GDPR, Personal Data, Smart City, Smart World.

1 Introduction

In December 2019 the world and life as we knew it changed dramatically: The SARS-CoV-2 coronavirus started spreading, quickly becoming a worldwide problem with numerous individual strokes of fate – may it be weeks of being isolated from loved ones, may it be the loss of jobs, or may it be the loss of health or even the loss of life, there is hardly anyone who has not been affected by the pandemic that has been caused by SARS-CoV-2.

Now, after more than three years into the pandemic, and hopefully shortly before entering the post-pandemic era, we urgently need to cope with the lessons learned during the pandemic. Regarding privacy, one major lesson learned from the pandemic is that we are not yet ready to share (personal) data³ within, e.g., authorities and health care facilities, even though the ability to do so could play a crucial role in limiting the spread of future diseases: The authors of this paper [SeT21c] conducted an empirical data collection with 30 Smart City participants from health, (primary) care, research and education sectors. The results of this data collection revealed that although Smart City participants see great

¹ Fraunhofer-Institute for Secure Information Technology and National Research Center for Applied Cybersecurity ATHENE, Rheinstraße 75, 64295 Darmstadt, Germany
Annika.selzer@sit.fraunhofer.de, <http://orcid.org/0000-0001-8859-4808>.

² University of Trier and German Research Center for Artificial Intelligence (DFKI), SDS Branch Trier, Behringstr. 21, 54296 Trier, Germany, ingo.timm@dfki.de, <https://orcid.org/0000-0002-3369-813X>.

³ Personal data means any information relating to an identified or identifiable natural person (Art. 4 (1) General Data Protection Regulation, or short GDPR). When processing personal data for more than pure private reasons (such as writing down the birthday of one's family members) one needs to comply to the relevant data protection statutes such as the GDPR.

potential to reduce the spread of contagious diseases and – to some extent – even to combat their causes in the future, the potential of data sharing and analyzing remains largely untapped up until today.

A *Smart City* uses information technology to analyze information of core systems, such as information and communication systems and devices that are connected through the internet, in running cities [SLF11]. By using, connecting, and analyzing such technology and its data, a Smart City can be optimized in its efficiency of services and operations. To achieve these advantages for the Smart City and its citizens, Smart City participants – such as civil service offices, hospitals, basic public utility providers, public transport providers, schools, and universities – usually process personal and anonymous data⁴ [GBP19; SLF11]. Therefore, they are – commonly speaking – highly experienced in collecting and analyzing data. However, through the 30 qualitative interviews we conducted with Smart City participants we learned that sharing and connecting relevant (personal) data is dramatically slowed down mainly by the following factors:

- 1) Smart City participants could reach the benefits of reducing the spreading of diseases and finding its causes mainly by sharing, connecting, and analyzing anonymous data. However, in some exceptional cases it could be beneficial to be able to re-identify individuals, e.g., when new treatments for diseases were found that could help patients with specific health profiles. However, Smart City participants usually do not know if and how such a re-identification is technically possible and legitimate under the relevant data protection statutes.
- 2) Smart City participants are unsure when a data set can be considered anonymous.
- 3) Smart City participants, especially public authorities such as public health agencies, are not technically ready to share (personal) data due to lack of interfaces and data exchange platforms [SeT21c].

To overcome this situation, we introduce a possible solution for a privacy-friendly way of data sharing within Smart Cities within this paper.

2 Preliminary consideration of (re)-identifiability

As already mentioned, Smart City participants could reach a lot of benefits, e.g., reducing the spreading of diseases and finding its causes, mainly by sharing, connecting, and analyzing anonymous data even though in some exceptional cases it could be beneficial to be able to re-identify individuals.

From a data protection perspective, the requirements of processing anonymous data in conjunction with the possibility of restoring the personal reference in individual cases

⁴ Anonymous data is information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the individual is not or no longer identifiable (Recital 26 GDPR).

seem to contradict each other, as *anonymous data* is information that does not or no longer identify an individual (recital 26 GDPR) and is thus the exact opposite of *personal data*. The distinction between these two terms is important, as only personal data is subject to the strict rules of the GDPR and other relevant data protection statutes. For purely anonymous data processing, however, the GDPR and other relevant data protection statutes do not apply. If the organization that processes personal data (the so called “data controller”) pseudonymizes personal data but retains the attribution rule between the pseudonyms and the relevant individual, the GDPR fully applies to this processing as well.

The wording of the recital of the GDPR regarding anonymous data – namely that anonymous data may not be identifiable or may no longer be identifiable – sounds absolute. However, according to the prevailing legal opinion, the identifiability of an individual depends exclusively on the knowledge of the data-processing entity regarding the identity of an individual. Therefore, information for which the data-processing entity X can establish a reference to an individual (because it has the necessary additional knowledge) is personal. In contrast, for the data-processing entity Y, which does not have the additional knowledge and has no possibility to obtain the additional knowledge, the information is anonymous [JS18, KB20, SHS19, PP18, Se22].

Following the prevailing legal opinion, the sharing, connecting, and analyzing of data within Smart Cities could be reached if Smart City participants could be enabled to only process (from their perspective) anonymous data in the context of reaching benefits regarding disease/pandemic management. Following the need to re-identify individuals in exceptional and justified cases, a data trustee could be designated to store the additional knowledge that is necessary to re-identify an individual [SeT21a].

3 Sharing, connecting, and analyzing anonymous data

Therefore, as an exemplary way to enable sharing, connecting and analyzing anonymous data within a Smart City context, Smart City participants could transfer personal data to a data trustee (step 1 in figure 1) that would need to provide relevant interfaces in between its own data processing systems and the data processing systems of Smart City providers and that would need to pseudonymize the data (step 2) without assigning any further function to the additional knowledge that could be used to (re-)identify individuals other than confidentially storing it. The pseudonymized data could then be made available for access by any other Smart City participant, while the data trustee excludes the possibility that anyone can access the additional knowledge that would be necessary to re-identify individuals (step 3).

As a result, only the data trustee itself⁵ could re-identify individuals, so that, according to the prevailing legal opinion about re-identification and anonymization, the Smart City

⁵ With the exception of the personal data transferring Smart City participant who – however – already needs to have a legal basis for personal data processing since he is the one who transferred the data in the first place.

participants ultimately access *anonymous data*, to which the provisions of the GDPR do not apply and therefore they can share, connect, and analyze these data without any further limitations other than attempting to fraudfully de-anonymize the data (e.g., through hacking the server of the data trustee) [SeT21a].

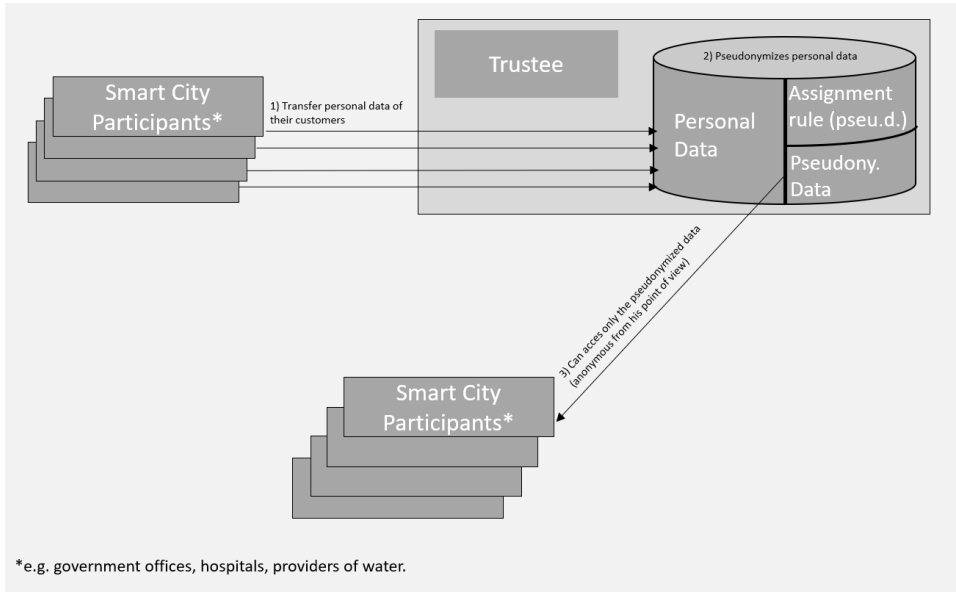


Figure 1: Sharing, connecting, and analyzing anonymous data

4 Re-Identifying individuals in exceptional cases

In order to meet the requirement of being able to re-identify individuals in exceptional cases, the data trustee could enable Smart City participants with the option to submit a request for re-identification for justified individual cases (step 1 in figure 2). The data trustee could then establish the personal reference for the individual data record (step 2) and contact the relevant individual to ask for consent to transfer his name and contact information to the enquiring Smart City participant (step 3). If the individual would consent to the disclosure of the data (step 4a), the Smart City participant would receive the personal data of the individual (step 5a). If, in contrast, the individual would refuse consent (step 4b), the personal data would not be disclosed to the Smart City participant (step 5b).

In case the Smart City participants request for re-identification is granted, he then needs to comply to the GDPR when processing the data set of the (one) individual that gave consent to his re-identification [SeT21a].

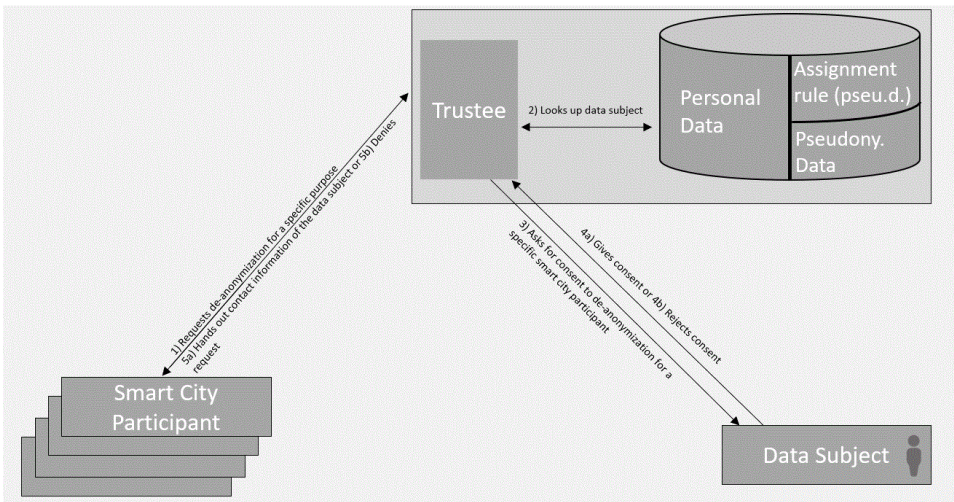


Figure 2: Re-Identifying individuals in exceptional cases

5 Requirements and Limitations

To benefit from the advantages of processing anonymous data with the opportunity to re-identify individuals in exceptional cases through the use of a data trustee, the following requirements and limitations need to be considered:

- 1) There needs to be a relevant legal basis for the initial transfer of personal data by the Smart City participant to the data trustee.
- 2) A Smart City participant who transfers personal data to the data trustee needs to have a verified e-mail address of all individuals whose personal data he transfers to the data trustee and needs to transfer this e-mail address to the data trustee to ensure that the data trustee can contact the correct individual for consent later.
- 3) The transfer of data to the data trustee is limited to data that can be considered anonymous based on the erasure of personally identifying characteristics of pseudonymized data. If the data trustee receives data that does not fulfill this requirement, he needs to erase the data set (or parts of it) to ensure anonymity towards data receiving Smart City participants.
- 4) The pseudonymized data sets cannot contain quasi-identifiers (e.g., the combination of an individual's date of birth, place of residence and gender – without any further information such as name or postal address – could be sufficient to establish a personal reference and therefore could be “quasi-identifiers”).

- 5) The data trustee needs to ensure that each data record sent to him is stored under a new pseudonym so that the Smart City participants cannot recognize whether a data set 1 and 2 is to be assigned to the same or two different individuals.
- 6) The Smart City participants do not have any additional knowledge about the de-anonymization of the data (neither from their own sources nor from public sources accessible to them) [SeT21b].

6 Outlook to a Smart World

“Smart City” has been used as a term for about 20 years now. While in some cities the efforts to link various information and communication technologies and to evaluate various data in a data protection-friendly manner for the improvement of the city and its citizens are being actively used and therefore true "Smart Cities" have emerged as a result, many other cities are still trying to follow up. The reason for this is, among other issues, that they have not yet created suitable interfaces for data exchange between the individual participants and are unsure about meeting the requirements for data anonymization.

The pandemic has shown us the importance of enabling the privacy-friendly exchange of relevant data in the event of an emergency. For the future, we therefore should take precautions to enable data exchange in a privacy-friendly manner, so that we can, among other things, detect and contain contagious diseases more quickly and, within its possibilities, even identify the causes of new diseases more quickly.

This paper showed one possible way to enable the exchange of relevant data in a privacy-friendly way. In the upcoming future, we need to prioritize discussing this and other possible solutions to eventually implement a privacy-friendly data exchange system that leads us into the future. In this regard, it is no longer enough to talk about "Smart Cities". We need efforts to create "Smart Countries," "Smart Continents," and – especially with regard to pandemics – a "Smart World" without restricting the personal rights of individuals, including their right to privacy, in an inappropriate manner.

The current breakthroughs and developments of artificial intelligence, especially in the field of large-scale language models, bring new complexity to this context: Not only data but also information and knowledge automatically derived from data, such as experiences and expectations, potentially based on linking originally personal data must be considered in future. Current top policy makers, such as the German Minister of Health, are discussing extensive use of large-scale language models for treatment of patients.⁶ Next to the mandatory societal discourse, analog solutions, as shown in this paper, should be developed for the extended use of innovative AI systems.

⁶ See <https://www.golem.de/news/digitalstrategie-lauterbach-erwartet-medizinische-revolution-durch-chatbots-2304-173690.html>.

Acknowledgements

This research work has been funded by the German Federal Ministry of Education and Research (BMBF) and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE as well as by BMBF within their support of SEMSAI (FKZ 031L0295A) as a consortium project within the MONID Modeling Network for Severe Infectious Diseases.

References

- [GBP19] Gassmann, O., and Böhm, J., and Palmie, M. (2019). *Smart Cities: Introducing Digital Innovation to Cities*. 1st Edition. Bingley: Emerald Publishing.
- [JS18] Jandt, D. and Steidle, R. (2018). *Datenschutz im Internet*. 1st Edition. Baden-Baden: Nomos.
- [KB20] Kuehling, J. and Buchner, B. (2020). *Datenschutz-Grundverordnung / Bundesdatenschutzgesetz, Kommentar*. 3rd Edition. München: C.H.Beck.
- [PP18] Paal, B. and Pauly, D. (2018). *Datenschutz-Grundverordnung und Bundesdatenschutzgesetz*. 2nd Edition. München: C.H.Beck.
- [Se22] Selzer, A. (2022). *Datenschutzrecht. Ein Kommentar für Studium und Praxis*. 1st Edition. Stuttgart: Fraunhofer Verlag.
- [SHS19] Simitis, S., Hornung, G. and Spiecker, I. (2019). *Datenschutzrecht*. 1st Edition. München: C. H. Beck.
- [SLF11] Su, K. and Li, J., and Fu, H. (2011). “Smart City and the Applications.” 2011 International Conference on Electronics, Communications and Control (ICECC), 1028-1031.
- [SeT21a] Selzer, A. and Timm, I. J. (2021). „Potenziale anonymer Datenverarbeitungen nutzen – Ein Vorschlag für Smart Cities“. *DuD 2021*, p. 816–820.
- [SeT21b] Selzer, A. and Timm, I. J. (2021). “Gestaltung eines Treuhand-Systems zum anonymen Datenaustausch in einer Smart-City-Umgebung -- Gewährleistung angemessener Schutzmaßnahmen“. *DuD 2021*, p. 826–830.
- [SeT21c] Selzer, A. and Timm, I. J. (2021). “Chances and Limitations of Personal and Anonymized Data Processing – Implementing Appropriate Technical and Organizational Measures and Creating Added Value in Smart Cities”. *GI Informatik 2021*, p. 773-788.

Towards a Taxonomy of Cyber Defence in International Law

Tanya Gärtner¹

Abstract: The increasing number of cyberattacks and their financial ramifications require legal certainty for cybersecurity experts and researchers that are faced with the task of defending against them. The variable use of language as well as the ongoing debate concerning the legality of certain measures of cyber defence further complicate this already critical problem. It is for these reasons that the following article attempts to establish clarity of terminology regarding cyber defence measures, specifically by disassociating hack backs from active cyber defence, before juxtaposing a proposal for a taxonomy of cyber defence measures with the international law on cyber activities and further unresolved questions in the legal debate.

Keywords: Active cyber defence, hack backs, classification, state sovereignty, international law

1 Introduction

Research has shown that cyberattacks occur at a near-constant rate [RBC07]. To be exact, a computer connected to the internet is attacked on average 2,805 times a day, or about every 30 seconds [RBC07]. Global annual cybercrime costs are expected to reach 8 trillion USD in 2023 and 10.5 trillion USD by 2025 [Mo22]. However, not only the financial ramifications of cyberattacks are increasing, but also the sophistication and scale with which attackers are employing malicious cyber activities [LXY13]. It is therefore particularly disadvantageous that the debate on cyber defence is seemingly characterized by an obscure and variable use of language.² For example, the UK National Cyber Security Strategy 2022 understands “active cyber defence” to be a tool that “helps organisations to find and fix vulnerabilities, manage incidents or automate disruption of cyberattacks (...)” [HM22]. Furthermore, in 2019 a bill was introduced to the United States Congress that aims to provide a defence to prosecution for fraud and related activity in connection with computers for persons defending against unauthorized intrusions into their computers, and for other purposes (so-called “Active Cyber Defence Certainty Act”) [Un19]. It defines the term “active cyber defence measure” as “any measure undertaken by, or at the direction of, a defender; and consisting of accessing without authorization the computer of the attacker to the defender’s own network to gather information in order to establish attribution of criminal activity (...); disrupt continued unauthorized activity against the

¹ Fraunhofer Institute for Secure Information Technology SIT | National Research Center for Applied Cybersecurity ATHENE, IT Law & Interdisciplinary Privacy Research, Rheinstr. 75, 64295 Darmstadt, tanya.gaertner@sit.fraunhofer.de.

² For an overview of the domestic understanding of technical terms in selected countries, see [Co19].

defender's own network; or monitor the behaviour of an attacker to assist in developing future intrusion prevention or cyber defence techniques (...)" [Un19]. Other countries have not put forward official definitions of these and various other relevant terms. This complicates not only the identification and comprehension of new technological developments, but also legislative and judicial processes that aim to determine the legality of specific cyber defence measures.

For this reason, the following article attempts to clearly differentiate and categorise measures of cyber defence based on characteristics that are generally relevant for the legal assessment of actions. As many cyberattacks are carried out internationally or have international effects, this attempts to aid the determination of whether specific cyber measures are permissible under international law. The *first characteristic* cyber defence measures will be sorted by, is whether they are taken preventatively or counteractively. This results in the differentiation between passive cyber defence and active cyber defence (chapter 2). The *second characteristic* cyber defence measures will be categorized by, is whether their effects are limited to the defended IT infrastructures and their associated services (hereinafter: "IT infrastructures") or if they extend beyond these limits. This creates the categories of internal cyber defence and external cyber defence (chapter 3). Within the category of external cyber defence, a further distinction is made according to a *third characteristic*, namely whether external cyber defence measures compromise the confidentiality, integrity or availability of the IT infrastructures of others, for instance by gaining access to these IT infrastructures, or not. This creates the categories of non-intrusive external cyber defence and intrusive external cyber defence (chapter 4). Consequently, these categories are then summarised in an overall view of cyber defence measures (chapter 5). Then follows an introduction to the international law of cyber activities (chapter 6), which allows for a concluding juxtaposition with the proposed taxonomy of cyber defence measures and an outlook on further unresolved questions (chapter 7).

The subsequent analysis will not take into account the effectiveness, practicability or cost of any cyber defence measures. Furthermore, no regard will be given to whether specific cyber defence measures are only available to certain parties, such as the military, government law enforcement or private individuals.

2 Passive vs. Active Cyber Defence

In the following, cyber defence measures will be determined to be either of a passive nature, if they are taken preventatively, or of an active nature, if they are taken counteractively. This is done for two reasons. The first reason is to seize the opportunity to clearly disassociate "hack backs" from the term "active cyber defence". Hack backs are understood to be retaliatory cyberattacks, conducted by victims of previous malicious cyber activities with the aim of punishing the original attacker [SW22b]. Hack backs in this sense are widely rejected by experts for an array of reasons, ranging from the

promotion of vigilante justice to the possibility of inadvertently affecting the resources of uninvolved third parties and the risk of conflict escalation [Wi20]. Consequently, hack backs will be disregarded in the following analysis. The second reason is that it is a general legal principle across various jurisdictions, that actions which are usually considered wrongful can be taken rightfully in certain situations, such as self-defence, see for example [Sc17], [PE19]. However, for the wrongfulness of such an action to be precluded, it cannot be taken preventatively but must rather be a reaction to a preceding attack. This means that emergency situations, such as an ongoing attack, permit further measures to be taken than would be the case under normal circumstances. Thusly, there is a possibility that cyber defence measures that are deemed to be generally incompatible with international law will still be practicable, but only if they are employed in reaction to a specific cyberattack.

Conventional cyber defence includes, for example, firewalls and anti-malware tools [LXY13], but also good workplace practices such as strong passwords and greater situational awareness [De17]. The common denominator between these measures is that they are passive in nature, meaning that they are implemented as a preventative safeguard against future cyberattacks that could possibly occur [DS17]. Their fundamental purpose is to increase the cybersecurity of a specific system overall [He21a]. Measures in this sense therefore generally have an introspective focus [De14]. Considering the varying aims of passive cyber defence, it becomes clear that some measures, such as firewalls, serve to establish defensive perimeters around IT infrastructures and prevent unauthorized access [De14]. These measures can be labelled as “fortified cyber defence” [De14]. In addition, other measures of cyber defence focus on guaranteeing the operability and recovery of IT infrastructures in case of a cyberattack [De14], such as regularly creating backups of data and software [De17]. These measures can be named “resilient cyber defence” [De14]. Although passive cyber defence is indispensable in securing IT infrastructures, modern threats in cyberspace are said to have made it an insufficient safeguard when used by itself [FZ14], [SW22a]. This has led to the call for additional, more pro-active measures, namely active cyber defence [FZ14]. Active cyber defence takes a different approach to cyber defence and offers additional advantages in combatting ongoing cyberattacks [LXY13]. Defence measures are taken in real-time and in reaction to an ongoing cyberattack [SW22b]. This allows measures to be tailored to the specific threat at hand and be adapted as needed. In contrast to passive cyber defence, active cyber defence is triggered by a specific threat and focusses merely on neutralizing or mitigating this threat. Active cyber defence measures can be automated in many cases, so that actual human interaction with the threat is not decisive.

A demonstrative example for this differentiation is the concept of user authentication by password [DS17]. If an incorrect password is entered and whoever is attempting to bypass the user authentication is simply denied access, this can be considered to be (fortified) passive cyber defence [DS17]. This measure will not hinder an attacker from gaining access by repeatedly entering possible passwords until the correct one is identified [DS17]. If, however, the user authentication mechanism additionally locks the respective account after a certain number of incorrect passwords have been entered, this can be considered as

a deliberate reaction to a specific threat and as an active element in user authentication [DS17]. This is even more true for user authentication mechanisms that continuously validate users' identity by using biometrics such as mouse and typing patterns and terminate access as soon as irregular behaviour is registered [DS17]. Another example for a cyber defence measure that seemingly sits on the border between passive and active cyber defence is the honeypot. Honeypots are decoys which are planted in the defended IT infrastructure and are intended to lure the attacker away from specific resources and, for example, stop or slow down automated attacks or give early warnings [Sp03]. Although honeypots are usually employed preventatively, their intended effect is triggered by and focused on a specific, ongoing cyberattack, which justifies their classification as active cyber defence.

3 Internal vs. External Cyber Defence

Cyber defence measures can be categorised according to their scope of effects. Measures that aid in the defence against cyberattacks can either be limited to the IT infrastructures being defended and therefore take place internally, or have effects that reach beyond those IT infrastructures and consequently be of an external nature, see similarly [DS17]. Other authors make a similar distinction by using the term “defensive” for measures that are limited to the defended IT infrastructures and the term “offensive” for measures that take place outside of the defended IT infrastructure or affect the IT infrastructures of others [Le15], [He21b]. This terminology however has certain disadvantages. Firstly, as the term “offensive” is generally defined as describing the action of starting an attack [Me23], its utilization to name a category of cyber measures that are used in a strategically defensive manner could lead to confusion. The logical conclusion would be to name a category of cyber defence measures “offensive cyber defence”. Moreover, the term “offensive”, in the logic of language, requires a counterparty – someone the offence is directed towards. Using this term sweepingly for cyber defence measures that do not take place inside of the defended IT infrastructures disregards the existence of such cyber defence measures which do not necessarily implicate the IT infrastructures of other parties. On the contrary, they can take place on a neutral level, for example by influencing the internet's control plane [SW22b]. Lastly, the term “offensive cybersecurity” refers to ethical hacking techniques that are used to discover and mitigate previously unknown cyber vulnerabilities and attack methods, for example in [ASL20]. Cybersecurity experts and researchers simulate cyberattacks by mimicking the cyber techniques used by hackers, but with the opposite aim of creating more dependable cyber defence tools [ASL20]. In this context, the term “offensive” is consequently used in a manner true to its definition. Thusly, it is advisable not to label external cyber defence methods as “offensive”, as to also avoid confusion with offensive cybersecurity.

When it comes to their employment, external cyber defence measures are commonly approached with caution due to the uncertainty surrounding their legality [Br21]. For example, the U. S. Computer Fraud and Abuse Act prohibits accessing computers without

authorization or exceeding authorized access and the German Criminal Code criminalizes accessing specifically protected data without authorization as well as preparatory acts thereto. Nonetheless, from a purely technological point of view, taking into account some external measures as valid options for cyber defence is expedient due to the fact that particular cyberattacks take place entirely beyond the limits of the defended IT infrastructures and therefore require defence measures that take place beyond these limits as well [SW22b]. This illustrates the need for clear determinations on which external defence measures should be deemed lawful and which external defence measures are not justifiable, rather than disregarding all measures of external cyber defence. A possible differentiation in this sense is proposed in the following chapter. As is illustrated in chapter 4, external cyber defence measures can either be non-intrusive or intrusive, which could possibly be useful in determining the legality of such measures.

A suitable example for internal active defence is the elimination of vulnerabilities and malware in a victim's IT infrastructure [SW22b]. Hackers oftentimes strive to corrupt as many devices as possible with malware to integrate them into a so-called botnet [Si13]. This is done without the knowledge of the device's owner [Ba09]. Botnets are thusly networks of hijacked devices and the controller of a botnet can direct the activities of the compromised devices [Si13]. These botnets are then used to automate large-scale attacks, such as spam email generation, data theft and distributed denial-of-service attacks (DDoS) [Ba09]. When attempting to take down a botnet, it is not efficient to shut down singular bots due to the sheer size a botnet usually has [Si13]. The preferred strategy of defence is therefore to delete the malware and remediate weak spots that were used for its installation from all bots at once [SW22b]. Depending on the cooperation of the manufacturers of the corrupted devices, the release of new patches and updates can be used to disinfect bots from the malware [SW22b]. Examples for external cyber defence, such as BGP hijack counteraction, are discussed in the following chapter.

4 Non-Intrusive vs. Intrusive External Cyber Defence

External cyber defence is at times erroneously reduced to measures that penetrate the IT infrastructures of attackers or uninvolved third parties. Within this category, it is however possible to further differentiate between non-intrusive external cyber defence measures and intrusive external cyber defence measures. An intrusion is the unauthorized access to IT infrastructures, or the compromise of their confidentiality, integrity or availability (so-called CIA-Triad, see [BA10]) or an attempt thereto [Na15]. Therefore, external cyber defence measures are only intrusive if they infiltrate the attacker's internal IT infrastructure or compromise the principles of the CIA-Triad in relation to the attacker's IT infrastructure. For example, without gaining access to the attacker's IT infrastructure, the destruction, deletion or encryption of data located therein will compromise the availability of resources and thusly constitute an intrusive external cyber defence measure. Non-intrusive external cyber defence measures do not require penetration of the attacker's IT infrastructures or compromise the principles of the CIA-Triad thereto, but for example

influence the internet's control plane by changing configuration data in routers, internet exchange points, internet service providers, internet registries or internet registrars [SW22b]. In doing so, the internet traffic going to or coming from IT infrastructure identified as belonging to an attacker can be manipulated, for example to conduct BGP hijack counteraction [SW22b]. Similar to identity theft, BGP hijacking is the illegitimate takeover of IP addresses [HM07]. Information sent via the internet finds its intended recipient based on the recipient's unique IP address [Ri02]. The route that the information takes is determined by the Border Gateway Protocol (BGP). The routing information announced in this way is not validated, meaning that attackers can falsely announce that they own the IP address of a chosen victim [Fe12]. The traffic addressed to the victim's IP address will now be either routed to the attacker instead of the victim or subsequently forwarded to the victim, allowing the attacker to eavesdrop and even modify the transported information [Ch15]. One of the known incidents of BGP hijacking occurred on the 24th of February 2008, when Pakistan Telecom used this method to censor YouTube on a national level [Fe12]. However, the false routing information was inadvertently communicated worldwide, resulting in YouTube being not reachable globally for around two hours [Fe12]. BGP hijacking can be thwarted by using its own method – a defender can announce that he owns the hijacked IP address and thereby direct the diverted traffic back to its intended recipient [SW22b]. BGP hijack counteraction is a prime example of active cyber defence that does not intrude into an attacker's IT infrastructures. Examples of intrusive external cyber defence measures include accessing attackers' IT infrastructures with passwords, for instance that were obtained on the darknet, or by taking advantage of misconfigured systems and protocols, outdated cryptography or vulnerabilities in hardware and software [SW22b]. Strictly speaking, utilizing so-called zero-day exploits in attackers' IT infrastructures in order to stop or mitigate on-going cyberattacks can also be considered a method of intrusive external cyber defence, irrespective of the on-going debate on the reasonableness of stockpiling zero-day exploits [SW22b]. Once access to attackers' IT infrastructures has been gained, further steps can be taken, for example data that was either stolen or is used for cyberattacks can be deleted or encrypted [St18]. Furthermore, attackers' IT infrastructures can be disrupted for the time of their attack, for example by overwriting the BIOS firmware or the defender initiating DDoS attacks as a defensive measure [St18].

5 Taxonomy of Cyber Defence

In summarizing the categorisations made so far, the table below gives an overview of how cyber defence could be understood with a view to assessing individual measures pursuant to international law. This does not constitute an exhaustive list of cyber defence measures.

		Passive Cyber Defence	Active Cyber Defence
Internal Cyber Defence		Mere user authentication, for example by password, without further consequences of multiple failed attempts	User authentication and in addition reactions to failed authentication, such as the locking of the respective account
		Firewalls, anti-malware tools, strong passwords, user training	Remediation of vulnerabilities in the defended IT infrastructures
		Port scans ³	Honeypots White worms ⁴
External Cyber Defence	Non-Intrusive External Cyber Defence	Sharing of threat information, for example between law enforcement and commercial service providers, see [DS17]	BGP hijack counteraction Address hopping ⁵
	Intrusive External Cyber Defence	Persistent engagement ⁶	Accessing of attackers' IT infrastructures (Temporary) disruption of attacker's IT infrastructure

Tab. 1: Taxonomy of Cyber Defence Measures

³ Port scans are used to automatically detect open ports and vulnerabilities in applications ports are used for, see [Gr21].

⁴ White worms are viruses which are deployed in defended IT infrastructures in order identify intrusions and can, for example, destroy or analyze the intruding malware, see [De17].

⁵ Address hopping means that a cyber defender changes the IP address of a victim of a cyberattack on a quasi-random basis and attackers and eavesdroppers therefore constantly have to search for and identify target data packets, see [De17].

⁶ Persistent engagement is a strategy in cyberspace that entails continuously intercepting cyber threats and degrading the cyber capabilities of adversaries in order to prevent them from launching cyberattacks in the first place, thereby strengthening domestic cybersecurity, see [U.18].

6 The International Law on Cyber Defence

International law, also known as public international law or law of nations, governs the relations between its subjects, which are predominantly states [Ip18]. Its guiding principle is the sovereign equality of all states, as held in Art. 2 (1) UN Charta [Vi19]. A much-cited definition of state sovereignty was put forward in the Island of Palmas arbitral award and denotes that “Sovereignty in the relation between States signifies independence. Independence in regard to a portion of the globe is the right to exercise therein, to the exclusion of any other State, the functions of a State” [Pe28]. This means that no state can impose binding rules on another state and no state is subject to the jurisdiction of an international court, without ratifying its statute willingly [Vi19]. Law making and law enforcement is dependent on consensus of the subjects of international law [Vi19], which in turn is a basis for its binding nature [Ip18]. Many institutions and principles of international law stem from state sovereignty, such as the principle of non-intervention and the prohibition of the use of force, Art. 2 (4) UN Charta [Ko11].

A state is responsible for all violations of international law that can be attributed to it [Sc19]. Although debates on the individual’s role in international law are ongoing, the responsibility of individuals for internationally wrongful acts is to date primarily held in the Rome Statue of the International Criminal Court, which only addresses genocide, crimes against humanity, war crimes and the crime of aggression [Ar16]. When it comes to states’ responsibility in the cyber context, the question arises which norms of international law could be breached by cyber activities [Sc19]. As of yet, there is no international treaty which exhaustively regulates this topic and the emergence of one seems far off [Za21]. Therefore, customary international law and the general principles of international law come to the fore. Although the applicability of international law in cyberspace has been agreed upon, many specifics are still unclear, especially the problem of whether state backed cyber operations, that encroach on the legally protected rights and interests of another state, can be justified [La21]. The prevailing opinion regarding the prohibition of the use of force is that cyber activities only then constitute a breach of this rule if their scale and effects are comparable to non-cyber operations that would be considered a use of force, which is the case if the cyber activities cause physical harm that extends beyond the targeted IT infrastructure [La21]. An infraction of the principle of non-intervention requires cyber activities to exert coercion on the internal affairs of a state [La21], [Sc19]. As these norms constitute high barriers for the responsibility of states and in turn leave room for a wide range of antagonistic cyber activities below the threshold of an internationally wrongful act, the violation of state sovereignty through cyber activities is increasingly discussed [La21]. The International Group of Experts that summarized international law on cyberspace in the legally non-binding but nevertheless very influential Tallinn Manual 2.0 assessed the lawfulness of remote cyber activities that manifest on a state’s territory according to two factors [Sc17]. The *first factor* is the degree of infringement upon the target state’s territorial integrity [Sc17]. This means, that damage or a loss of functionality of the cyber infrastructure within a state’s territory can in some

cases constitute a violation of that state's sovereignty [La21]. Whether cyber activities that result in effects below the threshold of loss of functionality constitute a violation of sovereignty remains disputed [Sc17]. The *second factor* is whether there has been an interference with or usurpation of inherently governmental functions [Sc17]. This means, that a violation of state sovereignty can occur if no physical damage or loss of functionality was caused, but, for example, the changing or deletion of data leads to the interference with the conduct of elections or the delivery of social services [Sc17]. An example for the usurpation of inherently governmental functions is a state taking over a botnet's command-and-control servers located in another state, without that state's approval. This operation curtails the inherently governmental function of law enforcement of the latter state [Sc17]. This violation of the territorial state's sovereignty could however be a justified countermeasure if the respective conditions are met.

7 Conclusion and Outlook

As there is no consensus as to whether cyber activities that manifest in a state's territory but do not result in either physical damage or the loss of functionality constitute a violation of sovereignty, the legal admissibility of most of the cyber defence measures discussed in this article is undetermined. States looking to defend against cyberattacks are therefore currently faced with a great degree of uncertainty. What would it mean for international relations if these measures were to be deemed legally permissible, especially considering that many defensive measures can be used offensively (for example port scans, BGP hijack counteraction and DDoS operations)? Would the alternative prohibition of certain defence measures leave states ill-equipped in the face of increasing cyber threats? Although the potential wrongfulness of cyber activities could be precluded in certain situations such as self-defence and countermeasures [Sc17], the respective conditions that need to be subsequently met are very high. As such, countermeasures are not permissible in response to cyberattacks conducted by a non-state actor, unless the cyberattacks are attributable to a state [Sc17]. Furthermore, international law traditionally requires an injured state to announce countermeasures to the responsible state before they are implemented [Sc17]. Both of these conditions are difficult to meet due to the general problem of attributing cyberattacks to their originators. Until more clarity is established regarding the legal handling of the different genera of cyber defence measures, their employment will have to be decided on a case-by-case basis. Linguistic clarity is especially important in this regard and could possibly be achieved on the basis of the taxonomy proposed in this article. This taxonomy therefore represents a useable approach towards further grasping cyber activities in international law and identifying the boundary between lawful and unlawful behaviour in cyberspace. Perhaps the degree of intrusiveness of cyber activities that do not amount to either damage or loss of functionality could be a useful distinguishing characteristic. Furthermore, cyber defence measures that can be taken reactively could possibly pose further courses of action in the emergency situation of an ongoing cyberattack.

Acknowledgements

This research work has been funded by the German Federal Ministry of Education and Research and the Hessen State Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE. This article reflects the personal opinion of the author.

References

- [Ar16] Arnauld, A. von: *Völkerrecht*, C.F. Müller, Heidelberg, 2016.
- [ASL20] Aiyanyo, I. D.; Samuel, H.; Lim, H.: A Systematic Review of Defensive and Offensive Cybersecurity with Machine Learning. *Applied Sciences* 17/10, p. 5811, 2020.
- [Ba09] Bailey, M. et al.: A Survey of Botnet Technology and Defenses. In (Institute of Electrical and Electronics Engineers (IEEE) Ed.): *2009 Cybersecurity Applications & Technology Conference for Homeland Security*. IEEE, pp. 299–304, 2009.
- [BA10] Bedner, M.; Ackermann, T.: *Schutzziele der IT-Sicherheit. Datenschutz und Datensicherheit (DuD)* 5/34, pp. 323–328, 2010.
- [Br21] Broeders, D.: Private Active Cyber Defense and (International) Cyber Security—Pushing the Line? *Journal of Cybersecurity* 1/7, pp. 1 - 14, 2021.
- [Ch15] Chiesa, M. et al.: Computational Complexity of Traffic Hijacking under BGP and S-BGP. *Theoretical Computer Science* 600, pp. 143–154, 2015.
- [Co19] Cordey, S. et al.: *National Cybersecurity and Cyberdefense Policy Snapshots: Updated Collection 2*. Center for Security Studies (CSS), ETH Zurich, Zürich, 2019.
- [De14] Dewar, R. S.: The “Triptych of Cyber Security”: A Classification of Active Cyber Defence. In (NATO CCD COE Ed.): *2014 6th International Conference on Cyber Conflict (CyCon)*. Tallinn, pp. 7–21, 2014.
- [De17] Dewar, R. S.: *Active Cyber Defense. CSS Cyber Defence Trend Analysis 1*. Center for Security Studies (CSS), ETH Zürich, Zurich, 2017.
- [DS17] Denning, D.; Strawser, B. J.: Active Cyber Defense: Applying Air Defense to the Cyber Domain. In (Perkovich, G.; Levite, A. E. Eds.): *Understanding Cyber Conflict. Fourteen Analogies*. Georgetown University Press, Washington, pp. 193–210, 2017.
- [Fe12] Fedler, R.: *Prefix Hijacking-Angriffe und Gegenmaßnahmen*. Chair for Network Architectures and Services, Department of Computer Science, Technische Universität München, 2012.
- [FZ14] Flowers, A.; Zeadally, S.: US Policy on Active Cyber Defense. *Journal of Homeland Security and Emergency Management* 2/11, pp. 289 - 308, 2014.

- [Gr21] Grieger, F.: Portscans im Lichte des Rechts: eine straf- und zivilrechtliche Analyse. *International Cybersecurity Law Review* 2/2, pp. 297–316, 2021.
- [He21a] Herpig, S.: *Active Cyber Defense Operations. Assessment and Safeguards*, Stiftung Neue Verantwortung (SVN), Berlin, 2021.
- [He21b] Herpig, S.: Ausschussdrucksache 19(12)985. Stellungnahme Thema "Verfassungs- und völkerrechtliche Fragen im militärischen Cyber- und Informationsraum unter besonderer Berücksichtigung des Parlamentsvorbehaltes, der Zurechenbarkeit von Cyberangriffen sowie einer möglichen Anpassung nationaler und internationaler Normen", Berlin, 2021.
- [HM07] Hu, X.; Mao, Z. M.: Accurate Real-time Identification of IP Prefix Hijacking. In (Institute of Electrical and Electronics Engineers (IEEE) Ed.): 2007 IEEE Symposium on Security and Privacy (SP '07). IEEE, pp. 3–17, 2007.
- [HM22] HM Government: *National Cyber Strategy 2022. Pioneering a cyber future with the whole of the UK*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1053023/national-cyber-strategy-amend.pdf, 2022, accessed 29/04/2023.
- [Ip18] Ipsen, K.: Regelungsbereich, Geschichte und Funktion des Völkerrechts. In (Epping, V.; Heintschel von Heinegg, W. Eds.): *Völkerrecht. Ein Studienbuch*, C.H. Beck, München, pp. 1–46, 2018.
- [Ko11] Kokott, J.: States, Sovereign Equality, *Max Planck Encyclopedia of International Law*, <https://opil.ouplaw.com/display/10.1093/law:epil/9780199231690/law-9780199231690-e1113>, 2011, accessed 29/04/2023.
- [La21] Lahmann, H.: Die völkerrechtliche Dimension der IT-Sicherheit. In (Hornung, G.; Schallbruch, M. Eds.): *IT-Sicherheitsrecht. Praxishandbuch*, Nomos, Baden-Baden, pp. 1–61, 2021.
- [Le15] Lee, R. M.: *The Sliding Scale of Cyber Security*, SANS Institute, <https://sansorg.egnyte.com/dl/GJEumszLQX>, 2015, accessed 29/04/2023.
- [LXY13] Lu, W.; Xu, S.; Yi, X.: *Optimizing Active Cyber Defense*. In (Das, S. K., Nita-Rotaru, C., Kantarcioglu, M. Eds.): *Decision and Game Theory for Security*. Springer International Publishing, pp. 206–225, 2013.
- [Me23] Merriam-Webster Dictionary: *Offensive*. <https://www.merriam-webster.com/dictionary/offensive>, accessed 02/04/2023.
- [Mo22] Morgan, S.: *Cybercrime To Cost The World 8 Trillion Annually In 2023*, *Cybercrime Magazine*, <https://cybersecurityventures.com/cybercrime-to-cost-the-world-8-trillion-annually-in-2023/>, accessed 29/04/2023.
- [Na15] National Institute of Standards and Technology: *Intrusion. Glossary*, [https://csrc.nist.gov/glossary/term/intrusion#:~:text=Definition\(s\)%3A,See%20intrusion.](https://csrc.nist.gov/glossary/term/intrusion#:~:text=Definition(s)%3A,See%20intrusion.), accessed 21/04/2023.
- [PE19] Perron, W.; Eisele, J.: § 32. In (Eser, A. Ed.): *Schönke / Schröder Strafgesetzbuch Kommentar*. C. H. Beck, München, 2019.

- [Pe28] Permanent Court of Arbitration: Island of Palmas arbitral award, 1928.
- [RBC07] Ramsbrock, D.; Berthier, R.; Cukier, M.: Profiling Attacker Behavior Following SSH Compromises. In (Institute of Electrical and Electronics Engineers (IEEE) Ed.): 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2007. Proceedings. IEEE, pp. 119 -124, 2007.
- [Ri02] Rinker, M.: Strafbarkeit und Strafverfolgung von „IP-Spoofing“ und „Portscanning“. Zeitschrift für IT-Recht und Recht der Digitalisierung (MMR) 10, pp. 663–666, 2002.
- [Sc17] Schmitt, M. N. Ed.: Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations. Second Edition, Cambridge University Press, New York, NY, 2017.
- [Sc19] Schröder, M.: Verantwortlichkeit, Völkerstrafrecht, Streitbeilegung und Sanktionen. In (Vitzthum, W.; Proelß, A. Eds.): Völkerrecht. De Gruyter, Berlin, Boston, pp. 698–761, 2019.
- [Si13] Silva, S. S. et al.: Botnets: A Survey. Computer Networks 2/57, pp. 378–403, 2013.
- [Sp03] Spitzner, L.: Honeypots: catching the insider threat. In (Institute of Electrical and Electronics Engineers (IEEE) Ed.): 19th Annual Computer Security Applications Conference, 2003. Proceedings. IEEE, pp. 170–179, 2003.
- [St18] Stiftung Neue Verantwortung (SVN): Aktive Cyber-Abwehr / Hackback. Impulse, <https://www.stiftung-nv.de/de/publikation/hackback-ist-nicht-gleich-hackback>, accessed 23/04/2023.
- [SW22a] Shulman, H.; Waidner, M.: Der Weg zur aktiven Cyberabwehr. Frankfurter Allgemeine Zeitung. <https://www.faz.net/aktuell/wirtschaft/digitec/cybersicherheit-der-weg-zur-aktiven-cyberabwehr-17980091.html>, accessed 06/04/2023.
- [SW22b] Shulman, H.; Waidner, M.: Aktive Cyberabwehr. ATHENE Whitepaper, Darmstadt, 2022.
- [U.18] U. S. Cyber Command: Achieve and Maintain Cyberspace Superiority. Command Vision for US Cyber Command, 2018, <https://www.cybercom.mil/Portals/56/Documents/USCYBERCOM%20Vision%20April%202018.pdf>, accessed 29/04/2023.
- [Un19] United States Congress: H. R. 3270. Active Cyber Defense Certainty Act, 2019.
- [Vi19] Vitzthum, W.: Begriff und Geltung des Völkerrechts. In (Vitzthum, W.; Proelß, A. Eds.): Völkerrecht. De Gruyter, Berlin, Boston, pp. 5–78, 2019.
- [Wi20] Winstead, N.: Hack-Back: Toward A Legal Framework For Cyber Self-Defense. <https://www.american.edu/sis/centers/security-technology/hack-back-toward-a-legal-framework-for-cyber-self-defense.cfm>, accessed 29/04/2023.
- [Za21] Zabierek, L. et al.: US-Russian Contention in Cyberspace. Are “Rules of the Road” Necessary or Possible?, <https://www.belfercenter.org/publication/us-russian-contention-cyberspace>, accessed 14/03/2023.

Der geplante US Active Cyber Defense Certainty Act

Ein Vorbild für Deutschland?

Jessica Kriegel, Alina Boll¹

Abstract: Durch die Gesetzesinitiative Cyber Defence Certainty Act versucht man in den USA, ein höheres Maß an strafrechtlicher Rechtssicherheit für den Bereich der aktiven Cyberabwehr zu schaffen. Die darin angestrebten strafrechtlichen Neureglungen und die Frage, ob diese uns in Deutschland ein Vorbild sein könnten, sind Gegenstand dieses Artikels.

1 Aktive Cyberabwehr in Deutschland

Wird ein deutsches Unternehmen Opfer eines Cyberangriffs, so ist für das angegriffene Unternehmen aufgrund der derzeitigen Rechtslage kaum erkennbar, wie es den Angriff durch Maßnahmen der aktiven Cyberabwehr technisch verhindern bzw. unterbrechen kann, ohne dabei selbst gegen geltendes Recht zu verstoßen, sofern sich die Maßnahmen der aktiven Cyberabwehr nicht ausschließlich auf Eingriffe in die eigene IT-Infrastruktur beziehen, sondern z.B. auf die IT-Infrastruktur des Angreifers oder gar unbeteiligter Dritter (z.B. den Internet Service Provider des Angreifers) zugegriffen werden muss, um den Cyberangriff zu stoppen [Br22; Se22]. Grund hierfür ist, dass bei der Durchführung von Maßnahmen aktiver Cyberabwehr oftmals erforderlich ist die gleichen technischen Vorgehensweisen zu nutzen, wie Cyberkriminelle dies bei dem Angriff selbst tun. Aus gerade diesem Vorgehen drohen den IT-Sicherheitsforschenden jedoch oftmals – trotz gänzlich anderer Zielsetzung – Strafbarkeitsrisiken, denn das Computerstrafrecht differenziert unzureichend zwischen den verfolgten Absichten [Br22; Se22]. Dies resultiert daraus, dass die Normen des deutschen Computerstrafrechts nicht vor dem Hintergrund der aktiven Cyberabwehr entstanden sind, sondern es primäres Ziel des Gesetzgebers war, die Grundformen des Hackens zu sanktionieren [Ko18].

Ebenso ist es für Cybersicherheitsforschende ein schwieriges Unterfangen, neue Maßnahmen der aktiven Cyberabwehr zu erforschen sowie ggf. selbst bei der Ausführung aktiver Cyberabwehr zu unterstützen, ohne dabei der Unsicherheit zu unterliegen, eventuell gegen geltendes Recht zu verstoßen. Insbesondere im Lichte des andauernden

¹ Diese Forschungsarbeit wurden vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt. Der vorliegende Beitrag gibt die persönliche Meinung der Autorinnen wieder.

Angriffskrieges auf die Ukraine bestehen darüber hinaus Unsicherheiten, ob Maßnahmen der aktiven Cyberabwehr ggf. in internationale Konflikte eingreifen könnten [Se22].

Diese unsichere Rechtslage in Bezug auf Maßnahmen der aktiven Cyberabwehr besteht ebenso in vielen anderen Ländern, so auch in den USA. Eine Gesetzesinitiative möchte in den USA nun ein höheres Maß an Rechtssicherheit für den Bereich der aktiven Cyberabwehr schaffen. Der vorliegende Beitrag stellt die wichtigsten Inhalte des Gesetzesentwurfs vor, benennt Kritikpunkte des Entwurfs und diskutiert, inwiefern die Regelungen des Gesetzesentwurfs für mehr Rechtssicherheit sorgen können.

2 Der US Active Cyber Defense Certainty Act im Überblick

Die Gesetzesinitiative „Active Cyber Defense Certainty Act“ (ACDC) hat in den letzten Jahren nicht nur für Aufsehen in den USA gesorgt, sondern auch die Diskussion über die rechtlichen und praktischen Auswirkungen der aktiven Cyberabwehr neu entfacht. Vorgelegt wurde die US-Gesetzesinitiative erstmals 2017, als H.R. 4036, um die proaktiven Möglichkeiten privater Unternehmen und Organisationen, die Opfer von Hacking-Angriffen geworden sind, zu erweitern und ihnen rechtliche Klarheit und Schutz bei der aktiven Verteidigung gegen Cyberangriffe zu bieten [Ch19]. Hintergrund des Gesetzentwurfs von 2017 war der Anstieg von Cyberkriminalität und die mangelnde Rechtssicherheit für Angegriffene in Bezug auf die Abwehr der Angriffe [Po20]. Gemäß 18 USC § 1030 CFAA, dem US-Gesetz gegen Computerbetrug und Missbrauch, wird in den USA derzeit nämlich strafrechtlich und zivilrechtlich geahndet, wenn sich eine Verteidigungsstrategie außerhalb des eigenen Netzwerks des Angegriffenen bewegen [Po20]. Genau hier wollte die Gesetzesinitiative aus dem Jahr 2017 ansetzen, um den CFAA in gewissem Umfang einzuschränken [Ac20]. Dadurch sollte es Opfern von Hacking-Angriffen erlaubt sein, Angreifer selbst zu identifizieren und Angriffe (auch außerhalb des eigenen Netzwerks) selbst zu beenden [Po20].

Obwohl die Gesetzesinitiative von 2017 im Ausschuss gescheitert war, wurde sie im Jahr 2019 als H.R. 2370 mit einigen wichtigen Änderungen gegenüber der ursprünglichen Version erneut eingereicht. Bis heute gibt es noch keine Entscheidung über eine Verabschiedung oder Ablehnung dieser neuen Gesetzesinitiative. Aktuell liegt der Gesetzentwurf wohl noch dem “Subcommittee on Crime, Terrorism, Homeland Security, and Investigations” vor und könnte von Cyberangriffen betroffene Unternehmen und Organisationen zumindest in strafrechtlicher Hinsicht Rechtssicherheit für die Umsetzung von Maßnahmen der aktiven Cyberabwehr geben [Co23].

2.1 Definition der aktiven Cyberabwehr

Die Definition der „aktiven Cyberabwehr“ ist von entscheidender Bedeutung, wenn es darum geht, festzulegen, welche Maßnahmen als Reaktion auf Hacking-Angriffe künftig in den USA erlaubt sein sollen und welche weiterhin unter Strafe gestellt werden.

Abschnitt 4 des ACDC-Entwurfs von 2019 definiert vor diesem Hintergrund zunächst die „aktive Cyberabwehr“ als den Einsatz von technischen Mitteln gegen Angreifer und dessen Computer, um einen unerlaubten Zugriff auf Informationen zu identifizieren, die Kontrolle über Systeme oder Daten wiederzuerlangen, die durch einen Cyberangriff kompromittiert wurden und angemessene Maßnahmen zu ergreifen, um Schäden oder weitere Angriffe zu verhindern. Durch Zugriff auf die Informationen des Angreifers soll eine Überwachung des weiteren Vorgehens des Angreifers ermöglicht werden und Informationen aus dem Netzwerk des Angreifers gesammelt werden, um die kriminellen Aktivitäten an die verantwortlichen Behörden für Cybersicherheit zu übertragen.

2.2 Erlaubte Maßnahmen der aktiven Cyberabwehr

Der Gesetzentwurf sieht eine Reihe von Maßnahmen vor, die von einer Haftung nach dem Computer Fraud and Abuse Act ausgenommen und somit künftig erlaubt sein sollen. Im Wesentlichen soll der ACDC-Act Unternehmen und Organisationen (im Folgenden auch zusammenfassend „Unternehmen“ genannt) autorisieren, außerhalb des eigenen Netzwerks Informationen über Hacker zu Attributionszwecken zu sammeln. Gemäß Abschnitt 3 des ACDC-Acts von 2019 soll künftig hierfür auch der bisher verbotene Einsatz von sog. Beacons zur Attribution straffrei sein. Darunter versteht man das Einfügen eines getarnten Codes in die eigene Software, der im Falle eines Cyberangriffs aktiviert werden und dann bspw. Informationen über den Standort des Angreifers liefern kann [Ac20]. Neben dem vorgenannten Sammeln von Daten über einen Angreifer soll u.a. auch die Überwachung des eigenen Netzwerks auf Angriffe sowie die Nutzung von Technologien zur Identifizierung von Bedrohungen und die Entfernung von Schadsoftware erlaubt sein.

Mit den eben aufgelisteten Maßnahmen will der ACDC-Act Unternehmen mehr Handlungsspielraum bei der aktiven Verteidigung gegen Cyberangriffe ermöglichen und einen Rechtsrahmen für die aktive Cyberabwehr schaffen. Berücksichtigt werden muss jedoch stets, dass die aktive Verteidigungsmaßnahme auch angemessen ist. Wann eine Maßnahme angemessen ist, lässt sich nicht abschließend aus dem Gesetzentwurf nehmen. Grundsätzlich lässt sich aber festhalten, dass die Maßnahmen insbesondere der Zurechnung strafbarer Handlungen dienen sollen, um diese an Strafverfolgungsbehörden weitergeben zu können. Außerdem sollen die Maßnahmen speziell auf die Unterbrechung der unerlaubten Aktivitäten gegen das Netzwerk des Verteidigers abzielen [Ch19].

2.3 Unerlaubte Maßnahmen der aktiven Cyberabwehr

In Abschnitt 4 des ACDC-Acts von 2019 werden einige unerlaubte Maßnahmen der aktiven Cyberabwehr aufgelistet, um den Maßnahmen aktiver Cyberabwehr und der Einschränkung des CFAA wiederum Schranken zu geben. Straffbar bleiben insbesondere solche Maßnahmen, die den Angreifer schädigen, wozu beispielsweise auch das vorsätzliche Zerstören oder Unbrauchbarmachen von Daten Dritter gehört. Hiervon

ausgenommen sind die versehentliche Zerstörung von Daten Dritter sowie die vorsätzliche Zerstörung der eigenen, durch den Angreifer gestohlenen Daten. Zudem soll es verboten sein, durch Maßnahmen der aktiven Cyberabwehr fahrlässig finanzielle Verluste oder körperliche Verletzungen zu verursachen. Die Maßnahme der aktiven Cyberabwehr darf außerdem keine Bedrohung für die öffentliche Gesundheit oder Sicherheit darstellen. Was genau darunter zu verstehen ist, wird im Gesetzesentwurf jedoch nicht weiter definiert. Zuletzt sieht der ACDC-Act vor, dass Maßnahmen der aktiven Cyberabwehr nicht über das erforderliche Maß hinausgehen dürfen. Über das erforderliche Maß hinaus gehen alle absichtlichen Maßnahmen, die ein „Mehr“ als das darstellen, was im Rahmen der Aufklärung eines Cyberangriffs erforderlich ist, um Zurechnungszwecke zu erfüllen. Die Maßnahme soll nicht weiterreichender als nötig sein und nur die begrenzten Zwecke verfolgen.

2.4 Privilegien von Forschungseinrichtungen zur Erforschung der aktiven Cyberabwehr

Im Gegensatz zur alten Fassung von 2017 enthält die Fassung des ACDC-Acts von 2019 keine Regelungen bzgl. der Privilegien von Forschungseinrichtungen zur Erforschung der aktiven Cyberabwehr.

2.5 Strafmaß

Die Fassung des ACDC-Acts von 2019 enthält keine spezifische Regelung hinsichtlich des Strafmaßes und der Faktoren zur Festlegung des Strafmaßes bei Überschreitungen der erlaubten Maßnahmen zur aktiven Cyberabwehr. Lediglich in Abschnitt 4 ist kurz festgehalten, dass aktive Cybermaßnahmen künftig keinen Verstoß darstellen und somit vom Straftatbestand des CFAA ausgenommen werden sollen. Eine zivilrechtliche Haftung des Verteidigers soll hingegen gemäß Abschnitt 4 des ACDC-Acts von 2019 weiterhin möglich bleiben.

Ungewollte Überschreitungen der Maßnahmen zur aktiven Cyberabwehr durch angegriffene Unternehmen sollen wiederum durch eine Anzeigepflicht und behördliche Aufsicht verhindert werden.

2.6 Anzeigepflicht und behördliche Aufsicht

Abschnitt 5 des ACDC-Acts von 2019 regelt die Anzeigepflicht und die behördliche Aufsicht. Er sieht vor, dass Unternehmen, die nach einem gegen sie gerichteten Angriff aktive Cyberabwehr betreiben wollen, ihre Aktivitäten der zuständigen Behörde melden müssen und anschließend von dieser in Bezug auf die aktive Cyberabwehr weiterbegleitet werden. Die (primär) zuständige Behörde soll die FBI National Cyber Investigative Joint Task Force sein. Bei der Meldung an die zuständige Behörde handelt es sich um eine

prozedurale Voraussetzung des Ergreifens von Maßnahmen der aktiven Cyberabwehr [Ch19]. Nur wenn der Angegriffene eine Antwort des FBI erhalten hat und der Empfang der Nachricht bestätigt wurde, dürfen die entsprechenden aktiven Maßnahmen verwendet bzw. angewendet werden. Durch dieses Vorgehen soll gewährleistet werden, dass die Abwehrmaßnahmen im Einklang mit dem Gesetz stehen und keine unerwünschten Nebenwirkungen (z.B. hinsichtlich einer Haftung des sich durch aktive Maßnahmen gegen einen Cyberangriff wehrenden Unternehmens) nach sich ziehen. Die behördliche Aufsicht soll die Einhaltung der Anzeigepflicht sicherstellen und dafür sorgen, dass Unternehmen und Regierungsbehörden die erforderlichen Maßnahmen ergreifen, um zukünftige Angriffe zu verhindern oder gegenwärtige Angriffe zu beenden. Die Aufsicht soll von der Regierung und unabhängigen Sicherheitsexperten unterstützt werden. Neben der Überwachung der Maßnahmen soll der behördlichen Aufsicht aber auch ein Eingreifen oder Stoppen von Maßnahmen erlaubt sein, sollten diese gegen das geltende Gesetz verstoßen oder unverhältnismäßig sein.

Die Anzeigepflicht und die behördliche Aufsicht sollen primär dem Schutz der handelnden Unternehmen sowie der Öffentlichkeit dienen. Anzeigepflicht und behördliche Aufsicht sollen also gerade nicht zu einer Einschränkung der aktiven Cyberabwehr für Unternehmen führen. Vielmehr soll der ACDC-Act von 2019 sicherstellen, dass diese sich aktiv gegen Cyberangriffe verteidigen können, solange sie sich an die Vorgaben des Gesetzes halten. Ziel ist es, dass Unternehmen ihre Aktivitäten transparent machen und die Weiterbegleitung (bzw. Überwachung) durch die behördliche Aufsicht akzeptieren, um die Rechtskonformität ihrer Abwehrmaßnahmen zu gewährleisten [Ch19].

3 Der US Active Cyber Defense Certainty Act in der Kritik

Der ACDC-Act von 2019 ist nicht unkommentiert geblieben und hat in den USA die Diskussion rund um das Thema aktive Cyberabwehr neu entfacht.

3.1 Positive Stimmen

Die Befürworter des Gesetzentwurfs sehen in der Verabschiedung des ACDC-Acts von 2019 die Möglichkeit, das wachsende Problem der Cyberkriminalität zu lösen, indem betroffenen Unternehmen Maßnahmen an die Hand gegeben werden, mit welchen sie sich selbst schnell und effektiv vor Cyberangriffen schützen und sich selbstständig gegen solche verteidigen können [Po20,483]. Indem durch den ACDC-Act genau festgelegt wird, welche Maßnahmen künftig erlaubt sind und folglich nicht mehr unter Strafe gestellt werden, könne Rechtssicherheit geschaffen werden. Zudem sind die Befürworter der Ansicht, dass der ACDC-Act für neue Gerechtigkeit sorgen könne, indem künftig vermehrt Hacker verfolgt und bestraft werden [Na20]. Die Einführung des ACDC-Acts eröffne weiterhin die Möglichkeit, Cyber-Bedrohungsinformationen zu sammeln und an

die zuständigen Behörden weiterzugeben. Gleichzeitig erziele man einen Abschreckungseffekt, welcher die Zahl der Cyberangriffen reduzieren könne [Na20].

3.2 Negative Stimmen

Der erste ACDC-Act scheiterte 2017 u.a. deswegen, weil er nicht genug Unterstützung im Kongress der Vereinigten Staaten hatte. Viele Experten kritisierten, dass es nicht ausreichend Anhörungen und Diskussionen gab, bevor der Gesetzesentwurf dem Kongress vorgelegt wurde. Auch gegenüber dem ACDC-Act von 2019 bestehen erhebliche Bedenken. Unter anderem befürchten Kritiker, dass die Verabschiedung des Gesetzes zu einem Wildwest-Szenario führen könne, in dem Unternehmen ohne angemessene Schulung und Regulierung voreilige und möglicherweise illegale Maßnahmen ergreifen könnten [Po20]. Eine solche Eskalation könne wiederum zu einer Bedrohung der nationalen Sicherheit führen, insbesondere dann, wenn Unternehmen mit Hackern aus anderen Ländern konfrontiert werden und durch Maßnahmen der aktiven Cyberabwehr außerhalb des eigenen Computer-Netzwerks gegen die Gesetze anderer Länder verstoßen [Ro20]. Letztendlich sind sich die Kritiker des ACDC-Acts sicher, dass der aktuelle Gesetzesentwurf aufgrund der sprachlichen Mehrdeutigkeit zu viele Lücken offenlassen und dadurch mehr Verwirrung stifte, als tatsächlich Antworten auf die aktuellen Probleme der Cyberkriminalität liefere. Vielmehr vereitle das Gesetz das eigentliche Ziel, die Rechtsunsicherheit zu beseitigen, da es keinen klaren Rahmen für die Abwehr von Cyberangriffen biete [Co18].

3.3 Eigene Stellungnahme

Eine Orientierung an den hier dargestellten Regelungsinhalten des ACDC-Acts kann die derzeitige Rechtsunsicherheit deutscher Unternehmen nur bedingt verbessern. Zunächst gilt es klarzustellen, dass der ACDC-Act unter dem Begriff der aktiven Cyberabwehr auch sog. Hackbacks subsumiert. Unter dem Begriff der Maßnahmen der aktiven Cyberabwehr sollte der deutsche Gesetzgeber jedoch keinesfalls Hackbacks (digitale Vergeltungsangriffe nach Beendigung eines Cyberangriffs) verstehen, sondern vielmehr technische Maßnahmen, die akute Cyberangriffe stoppen, proaktiv verhindern und deren Aufklärung dienen sollen [Sh22]. Offensive Hackbacks würden vielmehr dazu führen, dass sich Konflikte verstärken und weitere Angriffe zu erwarten wären, denn ein Vergeltungsangriff würde zumeist eine Gegenreaktion hervorrufen und in den seltensten Fällen einfach unbeantwortet bleiben.

Zum Inhalt der Regelung kann dennoch gesagt werden, dass durch eine grundsätzliche Erlaubnis von Maßnahmen der aktiven Cyberabwehr mit konkreten Verbotsvorbehalten in Verbindung mit vorgelagerten behördlichen Meldepflichten und der Begleitung (bzw. Überwachung) der durchgeführten, erlaubten Maßnahmen zur aktiven Cyberabwehr, durch eine hierfür zuständige Behörde, das strafrechtliche Risiko der Unternehmen und Cybersicherheitsforschenden deutlich reduziert werden könnte. Andererseits ist bei der

Durchführung einer Maßnahme der aktiven Cyberabwehr i.d.R. dringend erforderlich, dass diese schnellstmöglich umgesetzt wird, um schwerwiegende Auswirkungen von Cyberangriffen (wie z.B. großflächige Identitätsdiebstahle oder Angriffe gegen Kritische Infrastrukturen) erfolgreich verhindern und schnellstmöglich stoppen zu können. Das in dem ACDC-Act vorgeschlagene Vorgehen (Anzeigepflicht und behördliche Aufsicht) ist mit dieser Eilbedürftigkeit nur schwierig zu vereinbaren. Ebenso ist den negativen Stimmen zum vorgeschlagenen ACDC-Act dahingehend zuzustimmen, dass der Gesetzesentwurf aufgrund seines offenen Wortlauts und der nicht eindeutigen Abgrenzung zwischen erlaubten und nicht erlaubten Maßnahmen aktiver Cyberabwehr, die Gefahr mit sich bringt, dass Unternehmen ungeschult und voreilig möglicherweise illegale Maßnahmen ergreifen könnten. Dem lässt sich jedoch wiederum entgegenhalten, dass die angeordnete Meldepflicht und die behördliche Begleitung der Maßnahme zur aktiven Cyberabwehr, eine solche Missbrauchsgefahr weitestgehend ausschließen dürfte. Der Gesetzesentwurf schweigt zudem völlig zu der Frage der Erforschung der Maßnahmen aktiver Cyberabwehr und kann für den Bereich der Cybersicherheitsforschung daher kein Vorbild sein. Eine strafrechtliche Unsicherheit ließe sich im Bereich der aktiven Cybersicherheitsforschung bspw. durch punktuelle Ergänzungen der strafrechtlichen Normen des Computerstrafrechts erreichen. So könnten Tatbestandsausschlüsse, ähnlich wie bereits in § 86 Abs. 4, § 91 Abs. 2 Nr. 1 und § 201a Abs. 4 StGB, für Handlungen eingeführt werden, die zur Wahrnehmung überwiegender Forschungszwecke durchgeführt werden. Dies hätte den Vorteil, dass sich IT-Sicherheitsforschende bei der Durchführung von Handlungen, an denen ein konkretes Forschungsinteresse belegt ist, nicht zunächst in die Strafbarkeit begeben müssten, um dieser unter strengen Voraussetzungen (i.d.R. dem Eingreifen von Rechtfertigungsgründen) wieder entkommen zu müssen [Go21].

4 Fazit

Es bleibt festzustellen, dass das Ergreifen von Maßnahmen aktiver Cyberabwehr ein technisch, rechtlich und ethisch verantwortungsvolles Handeln der die Abwehrmaßnahmen durchführenden Stellen erfordert. Sowohl das Ergreifen als auch das Erforschen relevanter Maßnahmen darf nicht ohne klare Regeln erfolgen, da ein rechtsfreier Raum für Zwecke ausnutzbar ist, der nicht mit ethischen Grundsätzen (z.B. der wissenschaftlichen Forschung) vereinbar ist. Andererseits bedarf es der Sicherheit eines rechtskonformen Handelns, um derartige Maßnahmen umzusetzen und zu erforschen, um im Ergebnis u.a. die Grundrechte und -freiheiten der Bürger zu schützen [Go21,990; Se22].

Literaturverzeichnis

- [Ac20] Acalvio: Hacking Back with the Active Cyber Defense Certainty Act, <https://www.acalvio.com/hacking-back-with-the-active-cyber-defense-certainty-act/>, Stand: 15.03.2023.
- [Br22] Brodowski, D. / Golla, S.: Whitepaper zur Rechtslage der IT-Sicherheitsforschung, Version 1.0, S. 9, 2022.
- [Ch19] Chesney, R.: Hackback Is Back: Assessing the Active Cyber Defense Certainty Act, <https://www.lawfareblog.com/hackback-back-assessing-active-cyber-defense-certainty-act>, Stand: 15.03.2023.
- [Co23] CONGRESS.GOV, <https://www.congress.gov/bill/117th-congress/senate-bill/2292/all-info>, Stand: 15.03.2023.
- [Co] Cook, C.: Cross-Border Data Access and Active Cyber Defense: Assessing legislative options for a new international cybersecurity rulebook, STANFORD LAW & POLICY REVIEW.
- [SG21] Golla, S.: IT-Strafrecht und Strafrecht – Neukalibrierung eines belasteten Verhältnisses, JZ 02/2021, S. 985-990.
- [Ko18] Kochheim, D.: Cybercrime und Strafrecht in der Informations- und Kommunikationstechnik, 2. Auflage, 2018
- [Na20] Nachreiner, C.: The pros and cons of the proposed hack back bill, <https://www.scmagazine.com/perspective/policy/the-pros-and-cons-of-the-proposed-hack-back-bill>, Stand: 15.03.2023.
- [Po20] Porch, A. M.: Spoiling for a fight: Hacking Back with the Active Cyber Defense Certainty Act, South Dakota Law Review, 2020.
- [Ro20] Robert, A.: Hackback to the Drawing Board: Ambiguity and Risk in the Active Cyber Defense Certainty Act, B.C. INTELL. PROP. & TECH. F., 2020.
- [Se22] Selzer, A., Spiecker gen. Döhmman, I., Tagesspiegel Background, Rechtsrahmen der offensiven Cybersicherheitsforschung, veröffentlicht am 05.12.2022, <https://background.tagesspiegel.de/cybersecurity/warum-es-einen-rechtsrahmen-fuer-die-offensive-cybersicherheitsforschung-braucht>, Stand: 20.03.2023.
- [Sh22] Shulman, H. / Waidner, M.: Athene Whitepaper, Aktive Cyberabwehr, Stand: 10.10.2022.

Datenschutzrechtliche Rollen in Metaversen und im virtuellen Weiterleben

Von betroffenen Personen über Auftragsverarbeiter zu Verantwortlichen

Ines Geissler¹

Abstract: Das Metaverse und Umgebungen des virtuellen Weiterlebens bieten großes Potenzial der Datenökonomie für Unternehmen und auch Menschen nutzen die virtuellen Welten immer mehr, um mit anderen Personen zu interagieren oder um mit Verstorbenen „in Kontakt“ zu bleiben. Dadurch können eine Vielzahl an Akteuren an personenbezogenen Datenverarbeitungen beteiligt sein. Dieses Papier beschäftigt sich mit den datenschutzrechtlichen Verantwortlichkeiten und Rollen der Akteure im Metaverse und in Anwendungen des virtuellen Weiterlebens und diskutiert, ob KI-basierte Avatare mit einer eigenen Rechtspersönlichkeit ausgestattet werden sollten.

Keywords: Auftragsverarbeiter, Betroffene Personen, Datenschutz, Datenschutz-Grundverordnung, Virtuelles Weiterleben, Metaverse, Verantwortlichkeit.

1 Ein Einblick in zukünftige Formen des Lebens

1.1 Metaversen und Anwendungen des virtuellen Weiterlebens

Metaversen sind Räume des Internets, in der die virtuelle und analoge Realität miteinander verschmelzen. In diesen Räumen ist es Nutzern möglich, sich mit ihrer virtuellen Identität (in Form eines Avatars) zu bewegen und mit anderen Nutzern oder mit Organisationen zu interagieren, virtuelle Produkte zu erwerben oder virtuelle Dienstleistungen in Anspruch zu nehmen [Su22].

Anwendungen des virtuellen Weiterlebens ermöglichen dagegen das „Weiterleben“ einer Person nach deren Tod (nachfolgend „Avatar-Inspirator“), i. d. R. in Form eines Avatars. Der Avatar-Inspirator lernt hierfür einen KI-basierten Avatar zu Lebzeiten an, so dass er durch diesen nach seinem eigenen Tod weiterleben kann, seine Angehörigen (nachfolgend „Hinterbliebene“) weiter mit ihm interagieren können und so Unterstützung bei dem Trauerprozess oder bei alltäglichen Entscheidungen erhalten können [SM20].²

¹ Fraunhofer-Institut für Sicherer Informationstechnologie SIT | ATHENE – Nationales Forschungszentrum für angewandte Cybersicherheit, IT Law & Interdisciplinary Privacy Research, Rheinstr. 75, 64295 Darmstadt, ines.geissler@sit.fraunhofer.de, <https://orcid.org/0009-0001-3416-6768>.

² Das virtuelle Weiterleben findet nicht in Metaversen statt. Dies ist perspektivisch denkbar, im Folgenden werden Metaversen und Anwendungen des virtuellen Weiterlebens jedoch getrennt betrachtet.

In beiden Formen des virtuellen (Weiter-)Lebens werden eine Vielzahl personenbezogener Daten verarbeitet, in Metaversen insbesondere die personenbezogenen Daten der menschlichen Nutzer, in Umgebungen des virtuellen Weiterlebens u.a. die personenbezogenen Daten des Avatar-Inspirators und der Kommunikationspartner des weiterlebenden Avatars.

Insofern sind am virtuellen (Weiter-)Leben eine Vielzahl an Akteuren beteiligt, die es zu identifizieren und deren datenschutzrechtliche Rolle zu bestimmen gilt. Damit beschäftigt sich der nachfolgende Beitrag.

1.2 Datenschutzrechtliche Rollen aus der DSGVO

Das Datenschutzrecht kennt verschiedene datenschutzrechtliche Rollen der an einer Verarbeitung personenbezogener Daten Beteiligten. So unterscheidet es unter anderem zwischen dem Verantwortlichen, dem Auftragsverarbeiter und der betroffenen Person.

Verantwortliche sind gem. Art. 4 Nr. 7 DSGVO natürliche oder juristische Personen, Behörden, Einrichtungen oder andere Stellen, die allein oder gemeinsam über die Zwecke und Mittel der Verarbeitung entscheiden.

Auftragsverarbeiter sind gem. Art. 4 Nr. 8 DSGVO natürliche oder juristische Personen, Behörden, Einrichtungen oder andere Stellen, die personenbezogene Daten im Auftrag des Verantwortlichen verarbeiten.

Betroffene Personen sind gem. Art. 4 Nr. 1 DSGVO natürliche Personen, deren personenbezogene Daten verarbeitet werden. Der betroffenen Person kommt in der DSGVO die wohl zentralste Rolle zu. Diese werden durch die DSGVO im Zusammenhang mit den sie betreffenden Daten vor unrechtmäßigen Eingriffen in ihre Rechte und Freiheiten geschützt.

Je nachdem welche Rolle den einzelnen Akteuren zugeordnet wird, entstehen für sie verschiedene datenschutzrechtliche Rechte und Pflichten, wie bspw.

- das Recht auf Information, Widerruf einer Einwilligung und Recht auf Löschung (betroffene Personen),
- die Pflicht zur Identifizierung und Umsetzung einer Rechtsgrundlage vor Datenerhebung, die Pflicht zum Schließen eines Auftragsverarbeitungsvertrags und die Pflicht zur umfangreichen Datenschutzdokumentation (Verantwortlicher bzw. gemeinsam Verantwortliche) und
- die Pflicht zum Ergreifen technischer und organisatorischer Maßnahmen, das Schließen eines Auftragsverarbeitungsvertrags und die Pflicht zur Beachtung der Weisungsrechte des Verantwortlichen (Auftragsverarbeiter).

Vor diesem Hintergrund ist es die Grundvoraussetzung eines datenschutzkonformen Agierens in Metaversen und im virtuellen Weiterleben, die datenschutzrechtlichen Rollen der beteiligten Akteure zu definieren.

2 Datenschutzrechtliche Rollen des virtuellen Lebens im Metaverse

Zu den Akteuren in Metaversen gehören unter anderem der Metaverse-Anbieter („Plattformanbieter“), menschliche Metaverse-Nutzer (die im Metaverse durch Avatare abgebildet werden), Unternehmen als Metaverse-Nutzer, Anbieter von technischen Dienstleistungen im Metaverse sowie Stellen, die in einem Metaverse Daten erheben wollen (wie z. B. Strafverfolgungsbehörden). Diese können umfangreich personenbezogene Daten bereitstellen, verarbeiten und miteinander austauschen.

2.1 Plattformanbieter

Ein Plattformanbieter stellt die Infrastruktur und Technologie bereit, die es den Nutzern ermöglicht, in der virtuellen Welt zu agieren. Dazu gehören beispielsweise die Server, die die Verbindung zwischen den Nutzern aufrechterhalten, die Software, die für die Erstellung und Verwaltung von Avataren verwendet wird, sowie die Schnittstellen, über die die Nutzer mit dem Metaverse interagieren können [Bo22].

In einem zentral strukturiertem Metaverse kommt dem Plattformanbieter die Rolle des datenschutzrechtlich Verantwortlichen im Sinne des Art. 4 Nr. 7 DSGVO zu.

Der Plattformanbieter entscheidet sowohl über die Zwecke als auch über die Mittel der Datenverarbeitung [KSM22], z. B. darüber, welche Daten zur Registrierung erforderlich sind, welche Funktionalitäten die Plattform enthalten kann und mit welchen Schutzmaßnahmen den Risiken für die Rechte und Freiheiten der betroffenen Personen begegnet wird [Bi22].

Für Teile seines Metaverse-Angebots könnte der Plattformanbieter jedoch mit anderen Stellen gemeinsam verantwortlich sein (s. u.).

2.2 Organisationen als Nutzer

Auch Organisationen können Metaversen nutzen, indem sie sich als Unternehmen oder ihre Marke(n) und Produkte repräsentieren. So können etwa Events in Metaversen organisiert [Su22] sowie virtuelle Gegenstände zum Kauf angeboten werden [Bi22].

Diese Organisationen können dabei als datenschutzrechtlich Verantwortliche gem. Art. 4 Nr. 7 DSGVO handeln. Während der Nutzung von Metaversen verarbeiten diese Organisationen regelmäßig personenbezogene Daten der in den Metaversen agierenden Personen. Dazu gehören bspw. Anmelde Daten im Rahmen einer Anmeldung zu einem

Workshop. Für diese Datenverarbeitungen entscheiden sie über Zwecke und Mittel [BW23].

In Metaversen kann es allerdings auch vorkommen, dass mehrere Stellen an einer personenbezogenen Datenverarbeitung beteiligt sind. So ist es bspw. möglich, dass sich mehrere virtuelle Einzelhändler zu einem virtuellen Kaufhaus zusammenschließen [VAR12] und bspw. ein gemeinsames Treuepunktesystem anbieten. In diesem Fall entscheiden alle virtuellen Einzelhändler gemeinsam über die Zwecke und Mittel der Datenverarbeitung des Treuepunktesystems und sind somit hierfür gem. Art. 26 DSGVO gemeinsam verantwortlich, während sie für die sonstigen Datenverarbeitungen ihres Einzelhandels im Metaverse nicht mit den anderen Einzelhändlern gemeinsam verantwortlich sind.

Organisationen können als Nutzer jedoch auch eine gemeinsame datenschutzrechtliche Verantwortlichkeit mit dem Plattformanbieter eingehen.

Zwischen Organisationen als Nutzer sowie einem Plattformanbieter bestehen Parallelen zu dem Sozialen Netzwerk Facebook sowie den durch Organisationen betriebenen Facebook-Fanpages. Hierbei stellt Facebook die Plattform bereit und Organisationen können innerhalb der Plattform über ihre Fanpages eigene Dienste wie bspw. Umfragen, Marktanalysen oder Informationsaufbereitung anbieten. In diesem Zusammenhang werden Facebook weitere personenbezogene Daten der Nutzer bekannt, ohne dass die Betreiber der Fanpages dies beeinflussen können. Bei diesen Fanpages werden durch Facebook personenbezogene Daten der Fanpage-Besucher verarbeitet, während die Fanpage-Inhaber unter Umständen lediglich aggregierte Statistiken erhalten. Facebook-Fanpage-Inhaber sind laut einer EuGH-Entscheidung mit Facebook gemeinsam verantwortlich [Eu18].

In Metaversen ist dies ähnlich: Organisationen können als Nutzer ihre Dienste innerhalb der Metaverseplattform anbieten [Bi22]. So ist jeder virtuelle Einzelhändler wie eine Fanpage zu betrachten, durch den es dem Plattformanbieter ermöglicht wird, mehr Informationen über die betroffenen Personen zu erhalten. Dieser Parallele folgend wird hier die Meinung vertreten, dass auch der Plattformanbieter gem. Art. 26 DSGVO gemeinsam mit den Organisationen datenschutzrechtlich verantwortlich ist.³

³ Anders sehen dies: [BW23], die eine gemeinsame Verantwortlichkeit als unrealistisch ansehen, da eine gemeinsame Festlegung der Zwecke und Mittel bei der großen Vielzahl an nutzenden Organisationen praktisch schwer umsetzbar ist. Dies ist in der Sache zwar richtig, allerdings trifft dieser Umstand auch auf Facebook-Fanpage-Inhaber zu, so dass aufgrund der o. g. EuGH-Entscheidung Best-Practice-Lösungen gefunden werden mussten, um die rechtlichen Anforderungen an die gemeinsame Festlegung der Zwecke und Mittel umzusetzen. Facebook-Fanpage-Inhaber berufen sich regelmäßig darauf, dass die mit Facebook geschlossenen Nutzungsvereinbarungen einen Vertrag über die gemeinsame Verantwortlichkeit impliziert.

2.3 Organisationen als technische Dienstleister

In Metaversen gibt es eine Vielzahl technischer Dienstleister, die Unternehmen und Einzelpersonen unterstützen, um ihre virtuellen Präsenzen zu betreiben. Hierzu gehören unter anderem Hostingdienstleister, die beispielsweise dem Plattformanbieter Server bereitstellen, sowie Zahlungsanbieter, die Dienstleistungen anbieten, um Transaktionen innerhalb von Metaversen zu erleichtern, wie zum Beispiel den Kauf von virtuellen Gütern [Bi22].

Fraglich ist nun, ob es sich bei diesen Dienstleistern um Verantwortliche gem. Art. 4 Nr. 7 DSGVO oder Auftragsverarbeiter im Sinne des Art. 4 Nr. 8 DSGVO handelt. Dies hängt davon ab, wer die Entscheidung über die Mittel und Zwecke der Verarbeitung innehat. Dies hängt wiederum vom konkreten Einzelfall ab und ob der Dienstleister selbst über die Mittel und Zwecke der Verarbeitung entscheidet oder ob er weisungsgebunden entsprechend den Vorgaben des Verantwortlichen Daten verarbeitet [LWG21].

Hostingdienstleister verarbeiten Daten für die Zwecke des datenschutzrechtlich Verantwortlichen und sind lediglich der verlängerte Arm dessen. Dem steht auch nicht der Umstand entgegen, dass Hostingdienstleister i. d. R. in gewissem Umfang selbst über die zu treffenden technischen und organisatorischen Maßnahmen entscheiden. Denn sofern dem datenschutzrechtlich Verantwortlichen die grundlegenden Entscheidungen über die Mittel sowie alle Entscheidungen über die Zwecke der Verarbeitung vorbehalten sind, steht ein geringer Spielraum bei der Auswahl der Mittel der Datenverarbeitung durch den Hostingdienstleister einer Auftragsverarbeitung regelmäßig nicht entgegen [TG22]. Somit kommt Hostingdienstleistern im Metaverse die datenschutzrechtliche Rolle des Auftragsverarbeiters zu. Dagegen sind Zahlungsdienstleister regelmäßig keine Auftragsverarbeiter, sondern eigenständige Verantwortliche [Ba19, WB23], weil sie Daten zu eigenen Zwecken und unter Festlegung der Verarbeitungsmittel verarbeiten [WB23]. Zudem haben Plattformanbieter keine Weisungs- oder Einsichtsrechte in Bezug auf die vom Zahlungsdienstleister gespeicherten Daten [FHS19].

2.4 Organisationen als Datennutzer

Des Weiteren können Organisationen im Metaverse auftreten, um die dort zugänglichen Daten zu nutzen. Dies kann besonders interessant für Krankenkassen, Gesundheitsämter, Krankenhäuser und ähnliche Institutionen sein [At23], die Daten der Nutzer, die sie selbst z. B. über ihren aktuellen Gesundheitszustand und Symptome preisgeben, im Metaverse analysieren, um beispielsweise Pandemien vorherzusagen, Krankheitsverläufe zu untersuchen oder Versicherungsleistungen zu begünstigen oder zu verwehren.

Auch Datenzugriffe, -auswertungen und -verknüpfungen durch staatliche Stellen wie zum Beispiel Strafverfolgungsbehörden sind denkbar und werden zum Teil bereits vorbereitet [In22]. Strafverfolgungsbehörden könnten beispielsweise per Beobachtung von Avataren auffällige Verhaltensmuster erkennen und so Straftaten aufdecken oder gar verhindern.

Die vorgenannten Stellen erheben diese Daten i. d. R. zu eigenen Zwecken und mit von ihnen selbst gewählten Mitteln, so dass sie für derartige Datenverarbeitungen datenschutzrechtlich verantwortlich sind. Anders als bei Facebook-Fanpages entscheiden diese Organisationen selbst über die Zwecke und Mittel der Verarbeitung und nutzen das Metaverse lediglich als Datenerhebungsquelle. Die durch sie erfolgende Datenspeicherung und -auswertung erfolgt i. d. R. außerhalb des jeweiligen Metaverses und der Plattformanbieter erhält somit keinen zusätzlichen Zugang zu personenbezogenen Daten. Eine gemeinsame Verantwortlichkeit mit dem Plattformanbieter besteht daher regelmäßig nicht.

2.5 Menschliche, über Avatare abgebildete Nutzer

Durch ihre Nutzung der Metaversen werden ihre personenbezogenen Daten verarbeitet, so zum Beispiel die IP-Adresse, Bewegungsdaten ihres Avatars oder Daten über ihr Kaufverhalten [Bo22]. Der menschliche Nutzer ist eine betroffene Person im Sinne des Art. 4 Nr. 1 DSGVO. Menschen können Metaversen nutzen, indem sie als Avatare eine virtuelle Umgebung betreten und dort agieren und mit anderen Nutzern interagieren. Sie können beispielsweise Videospiele spielen, virtuelle Konferenzen abhalten und virtuelle Produkte und Dienstleistungen kaufen oder nutzen [Bi22].

Ein Avatar im Metaverse fungiert regelmäßig als digitale Identität eines Rechtssubjekts. Rechtssubjekte können natürliche und juristische Personen sein. Sofern sie durch natürliche Personen eingesetzt werden, werden sie also einer natürlichen Person zugerechnet. Diese Avatare können äußerliche Merkmale ihrer Nutzer haben und mit anderen Avataren und Akteuren im Metaverse in Kontakt treten [Bo22]. Die Avatare hinterlassen bei Nutzung der Metaversen Spuren in diesen. Ihr Verhalten und ihre Eigenschaften können personenbezogene Daten ihrer Nutzer aufweisen [KSM22]. Betreten Avatare bspw. virtuelle Kaufhäuser, können Informationen darüber erhoben werden, wie lange sich der Avatar dort aufhält und wofür er sich interessiert. Dabei handelt es sich um personenbezogene Daten ihrer Nutzer.

Da der Avatar von einem menschlichen Nutzer gestaltet und in dem betreffenden Metaverse bewegt wird, handelt der Avatar nicht autonom und ist somit weder eine juristische noch eine natürliche Person. Insofern kann er (selbst) weder Verantwortlicher, Auftragsverarbeiter noch betroffene Person sein.

3 Datenschutzrechtliche Rollen des virtuellen Weiterlebens

Auch im Kontext des virtuellen Weiterlebens bedarf es einer Identifizierung und Analyse der datenschutzrechtlichen Rollen.

3.1 Dienstanbieter

Dienstanbieter des virtuellen Weiterlebens sind Organisationen, die es technisch ermöglichen, mit einem Avatar zu kommunizieren, der zu Lebzeiten des Avatar-Inspirators von diesem KI-basiert trainiert wurde und sich sehr ähnlich verhält, wie der Avatar-Inspirator selbst [SM20]. Der Dienstanbieter entscheidet über die Zwecke und Mittel der Verarbeitung und stellt die technischen Möglichkeiten bereit, um die personenbezogenen Daten zu speichern und die Interaktion mit dem Avatar zu ermöglichen und ist insofern datenschutzrechtlich Verantwortlicher. Zu den personenbezogenen Daten, die von den Dienst Anbietern des virtuellen Weiterlebens verarbeitet werden, können Informationen wie Nutzernamen, E-Mail-Adressen und alle anderen Daten gehören, die der Avatar-Inspirator dem Dienstanbieter (gewöhnlich über eine von ihm betriebenen Plattform) zur Verfügung stellt, um seinen Avatar anzulernen, bzw. die im Rahmen der Kommunikation zwischen Avatar und den Hinterbliebenen entstehen.

3.2 Menschliche Nutzer

Die Dienste des virtuellen Weiterlebens werden von Menschen in Anspruch genommen, die entweder mit verstorbenen Angehörigen kommunizieren wollen oder in Vorbereitung auf das eigene Ableben den eigenen Avatar anlernen und trainieren wollen, damit dieser dann nach ihrem Tod mit den Hinterbliebenen kommunizieren kann [SM20].⁴ Dabei ist unter anderem zu klären, ob eine bereits verstorbene Person eine datenschutzrechtliche Rolle einnehmen kann.

3.2.1 Person, die über einen Avatar abgebildet wird

Sofern personenbezogene Daten des noch lebenden Avatar-Inspirators durch den Dienstanbieter verarbeitet werden, handelt es sich eben diesem gegenüber um eine betroffene Person im Sinne des Art. 4 Nr. 1 DSGVO. Er kann im Rahmen der Kommunikation beispielsweise Namen, Hobbies, Beruf und persönliche Erinnerungen preisgeben, die zum Anlernen der KI verarbeitet werden.

Im Zusammenhang damit, dass der Avatar-Inspirator persönlicher Erinnerungen zum Anlernen der KI benutzt, die sich auf seine Hinterbliebenen beziehen und die personenbezogenen Daten seiner Hinterbliebenen enthalten, stellt sich die Frage, ob der Avatar-Inspirator für diesen Teil der Datenverarbeitung selbst datenschutzrechtlich verantwortlich sein könnte. Jedoch ist dies – ausgehend von der Annahme, dass die Nutzung dieser Daten ausschließlich zu rein privaten bzw. familiären Zwecken erfolgt –

⁴ Der Prozess des Anlernens kann entweder zum Zeitpunkt des Todes des Avatar-Inspirators abgeschlossen sein oder noch nach dem Tod bspw. durch die Kommunikation mit Hinterbliebenen fortgeführt werden. Die vorliegende Veröffentlichung schließt letztgenannten Fall ein.

regelmäßig zu verneinen, weil der Avatar-Inspirator für diesen Datenverarbeitungsschritt unter die Haushaltsausnahme der DSGVO fällt, so dass für ihn die DSGVO nicht anwendbar ist und er somit auch nicht datenschutzrechtlich verantwortlich ist.

Die DSGVO regelt ausschließlich den Schutz personenbezogener Daten lebender Personen. Insofern ist der Avatar-Inspirator zwar vor seinem Tod, aber nicht mehr nach seinem Tod eine betroffene Person im Sinne der DSGVO.

3.2.2 Kommunikationspartner des Avatars

Kommunikationspartner des Avatars zu Lebzeiten können beispielsweise Angehörige und Freunde sein, die dem weiterlebenden Avatar Informationen, über die ihn inspirierende Personen weitergeben wollen. In diesem Zusammenhang werden von dem Dienstanbieter (durch den von ihm betriebenen, KI-basierten Avatar) personenbezogene Daten dieser Personen verarbeitet. Die Kommunikationspartner des Avatars zu Lebzeiten des Avatar-Inspirators sind gegenüber dem Dienstanbieter betroffene Person im Sinne des Art. 4 Nr. 1 DSGVO. Zu den verarbeiteten Daten können neben der IP-Adresse auch Informationen über die persönliche Beziehung zum Avatar-Inspirator gehören. Die Avatare können während ihrer Interaktionen personenbezogene Daten über die Kommunikationspartner sammeln und zu Trainingszwecken der KI verwenden [SM20]. Darüber hinaus können die Nutzer dem Dienstanbieter personenbezogene Daten wie ihren Namen, ihre E-Mail-Adresse und andere Informationen zur Verfügung stellen, damit sie mit den KI-generierten Avataren interagieren können.

Auch die Kommunikationspartner des Avatars eines verstorbenen Avatar-Inspirators sind gegenüber dem Dienstanbieter betroffene Personen gem. Art. 4 Nr. 1 DSGVO, da bei einer Kommunikation ebenso personenbezogene Daten wie IP-Adressen und z. B. die dem Avatar gestellten Fragen verarbeitet werden.

3.3 Avatar

Schließlich stellt sich die Frage, ob der KI-basierte Avatar, mittels dessen eine real lebende Person durch ein entsprechendes Dienstleistungsangebot des Dienstanbieters nach dem Tod virtuell weiterleben kann, selbst eine datenschutzrechtliche (ggf. mit dem Dienstanbieter gemeinsame) Verantwortung trägt, weil der Avatar selbst möglicherweise gegen geltendes Datenschutzrecht verstoßen könnte, z. B. in dem personenbezogene Daten von Hinterbliebenen anderen Kommunikationspartnern ungewollt offengelegt werden. So könnte der Avatar bspw. den Verlust des Arbeitsplatzes, eine Schwangerschaft oder eine Scheidung eines seiner Kommunikationspartner verkünden. Aus derartigen Datenschutzverstößen könnten sich gegebenenfalls Haftungsansprüche ergeben. Fraglich ist in dem Zusammenhang, ob diese dem Avatar (als datenschutzrechtlich Mitverantwortlichem) selbst zuzuordnen sind.

Die Auffassung, dass KI-basierten Avataren eine datenschutzrechtliche (Mit-)Verantwortlichkeit zukommen könnte, stützt sich auf folgende Überlegungen:

1. (K)eine Entscheidungsfähigkeit über die Zwecke der Verarbeitung

Ein datenschutzrechtlich Verantwortlicher muss in der Lage sein, den Verarbeitungsprozess zu steuern [PP21]. Dies könnte auf das KI-System zutreffen, wenn das KI-System selbst entscheiden würde, ob und zu welchen Zwecken Daten gespeichert oder gelöscht werden. KI-Systeme mit niedrigem Autonomiegrad beziehen die von ihm verarbeiteten Daten regelmäßig aus einer Datenbank [BI20], im Falle des virtuellen Weiterlebens i. d. R. aus einer Datenbank des Diensteanbieters, die er wiederum mit Daten des Avatar-Inspirators und seiner Kommunikationspartner befüllt bzw. von den Vorgenannten befüllen lässt, so dass der Diensteanbieter die Entscheidung trifft, auf welche Daten das KI-System zugreifen kann. Auch die Zwecke des Einsatzes des KI-Systems – und somit der Datenverarbeitung durch das KI-System – legt der Diensteanbieter fest. KI-Systeme mit hohem Autonomiegrad können dagegen selbst festlegen, welche personenbezogenen Daten sie zu welchem Zweck verarbeiten möchten [BI20]. Grundsätzlich wird also bei KI mit niedrigem Autonomiegrad davon auszugehen sein, dass der Diensteanbieter allein verantwortlich ist. Bei hohem Autonomiegrad könnte eine eigene Entscheidungsfindung hinsichtlich des Zwecks der Verarbeitung getroffen werden, so dass theoretisch eine gemeinsame Verantwortlichkeit mit dem Diensteanbieter in Betracht käme. Da die KI aber weder juristische noch natürliche Person ist, müsste sie zu einem Rechtssubjekt mit Rechtspersönlichkeit gemacht werden, die in der Lage wäre, Rechte und Pflichten zu erfüllen und datenschutzrechtliche Verantwortung zu übernehmen.

2. Einführung einer ePerson in der aktuellen Diskussion

Die Frage nach der Verantwortlichkeit und der Zuordnung zum menschlichen Handeln ist bei dem Einsatz von KI nicht immer einfach zu definieren. Um mehr Klarheit zu schaffen, wird insbesondere im Haftungsrecht über die Einführung einer elektronischen Person (ePerson) diskutiert. Bisher werden ausschließlich natürliche und juristische Personen als Rechtssubjekt mit Rechtspersönlichkeit angesehen. Diese sind rechtsfähig [Ri20]. Die ePerson wäre vergleichbar mit einer juristischen Person, so dass ihr eine eigene Rechtspersönlichkeit zugeschrieben und sie damit Rechtsfähigkeit besitzen würde. Die KI würde dadurch Inhaberin von Rechten und Pflichten und in Schadensfällen selbst zu Verantwortung gezogen werden und damit selbst für ihr Handeln haften [WZ03]. Die Einführung der ePerson würde dazu führen, dass Schäden haftungsrechtlich klar zugeordnet werden könnten und Beweisprobleme vermieden würden [LWG21].

Für eine Einführung einer ePerson spricht auch, dass KI-Systeme, genau wie Menschen, sich an die Umwelteinflüsse anpassen und ohne menschliche Mitwirkung Entscheidungen treffen können [KM15]. Der technische Fortschritt könnte sogar dazu führen, dass KI-Systeme die gleichen Denkprozesse wie Menschen aufweisen. Diese Vergleichbarkeit der Denkprozesse würde bedeuten, dass auch vergleichbare Rechte und Pflichten bestehen müssten [Be09].

Diesen Überlegungen kann nach der hier vertretenen Meinung aus den folgenden Gründen jedoch nicht gefolgt werden:

- a) Entscheidungen müssen auf natürliche Personen zurückgeführt werden können.

Grundsätzlich gilt, dass Entscheidungen über Zwecke und Mittel auf eine natürliche Person zurückgeführt werden müssen [HSH22]. Insbesondere bei der automatisierten Entscheidungsfindung inklusive Profiling muss sichergestellt werden, dass diese Entscheidung durch einen Menschen überprüft werden kann [WB23].

KI-Systeme können anfällig für Diskriminierungen sein, insbesondere dann, wenn sie auf nicht repräsentativen Datensätzen trainiert werden. Insofern sollten wichtige Entscheidungen nicht von KI-Systemen übernommen werden [Dj22].

- b) Eine eigene Verantwortlichkeit durch Gesetzgeber wird nicht vorgesehen.

Zwar hatte das EU-Parlament 2017 vorgeschlagen „langfristig einen speziellen rechtlichen Status für Roboter zu schaffen, damit zumindest für die ausgeklügelten autonomen Roboter ein Status als elektronische Person festgelegt werden könnte, die für den Ausgleich sämtlicher von ihr verursachten Schäden verantwortlich wäre [...]“ [Eu17], dies wurde allerdings durch die EU-Kommission nicht aufgegriffen [Eu19].

Die EU plant im Rahmen des durch die EU-Kommission vorgelegten Vorschlags einer KI-Verordnung den Rechtsrahmen für künstliche Intelligenz zu regeln. Der KI-Verordnung-Entwurf verfolgt einen risikobasierten Ansatz, der Regulierungsstufen aufweist. So sollen bestimmte, besonders risikobehaftete KI-Anwendungen verboten werden und andere KI-Anwendungen bestimmte technische und organisatorische Vorgaben erfüllen und einer Konformitätsbewertung unterliegen [BM21]. Eine eigene Rechtspersönlichkeit für KI ist auch hier nicht vorgesehen.

Art. 52 KI-Verordnung-Entwurf besagt, dass KI-Systeme, die für die Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden müssen, dass natürlichen Personen mitgeteilt wird, dass sie es mit einem KI-System zu tun haben, es sei denn, dies ist aufgrund der Umstände und des Kontexts der Nutzung offensichtlich. Weiterhin wird in Erwägungsgrund 53 der KI-Verordnung als angemessen empfunden, dass eine bestimmte als Anbieter definierte natürliche oder juristische Person die Verantwortung für das Inverkehrbringen oder die Inbetriebnahme von Hochrisiko-KI-System übernimmt.

- c) Die KI wird nicht verkörpert.

Insbesondere in Bezug auf KI-basierte Avatare im virtuellen Weiterleben bestehen Probleme in Bezug auf die Verkörperung. Während autonome Fahrzeuge und physische Roboter körperlich identifiziert werden können und somit ein Rechtssubjekt darstellen könnten, ist eine derartige Identifizierung bei virtuellen Avataren u. U. nicht möglich. Wenn im Kontext des virtuellen Weiterlebens beispielsweise mehrere Avatare auf dem

gleichen Algorithmus basieren, ist unklar, wie viele Rechtspersönlichkeiten vorliegen [Ri20]. Für jeden Avatar eine eigene oder eine zentrale Rechtspersönlichkeit für alle Avatare, die auf dem gleichen Algorithmus basieren?

d) KI hat nicht die (finanzielle) Motivation, sich an Gesetze zu halten.

Gegen die Einführung der ePerson spricht außerdem, dass KI-Systeme keine Bedrohung beziehungsweise keinen Anreiz in der Anpassung ihres Verhaltens hinsichtlich der Haftung sehen. Diese Verhaltensanpassung erfolgt bei natürlichen Personen, weil sie Sanktionen vermeiden möchten und auch das Verhalten juristischer Personen wird aufgrund der Gewinnerzielungsabsicht gesteuert. Ein KI-System verfolgt keines dieser Interessen und hat somit auch kein Interesse daran, sich Rechtsordnungen anzupassen. Mögliche Haftungsmassen müssten über Beiträge aus möglichen Haftpflichtversicherungen gezahlt werden, so dass KI-Systeme mit einem Mindestkapital ausgestattet werden müssen. Dieses müsste wiederum von Hersteller oder Betreiber oder anderen Akteuren gezahlt werden [Ri20].

Folglich kommt dem Avatar nach hier vertretener Ansicht keine eigene datenschutzrechtliche Rolle zu und die Verantwortung für die durch den Avatar erfolgende personenbezogene Datenverarbeitung ist grundsätzlich (allein) dem Dienstanbieter zuzuschreiben.

4 Zusammenfassung und Ausblick

Grundvoraussetzung für ein rechtskonformes Handeln im Zusammenhang mit dem virtuellen Weiterleben und dem Leben im Metaverse ist, dass allen Beteiligten klar sein muss, welche datenschutzrechtlichen Rechte sie innehaben, und welche Rechte und Pflichten sie entsprechend ihrer datenschutzrechtlichen Rollen erfüllen müssen. Der vorliegende Beitrag konnte folgende Akteure und datenschutzrechtlichen Rollen identifizieren und zuweisen:

Akteur	Datenschutzrechtliche Rolle
<i>Metaversen</i>	
Plattformanbieter	Kontextabhängig alleinige oder gemeinsame Verantwortliche
Organisationen als Nutzer	Kontextabhängig alleinige oder gemeinsame Verantwortliche
Organisationen als technische Dienstleister	Kontextabhängig gemeinsam Verantwortliche oder Auftragsverarbeiter
Organisationen als Datennutzer	Alleinige Verantwortliche

Menschliche, über Avatare abgebildete Nutzer	Menschen: Betroffene Personen Avatare: Keine datenschutzrechtliche Rolle
<i>Virtuelles Weiterleben</i>	
Dienstanbieter	Verantwortliche
Menschliche Nutzer	Lebende Personen: Betroffene Personen
Avatar	Keine datenschutzrechtliche Rolle

Tab. 1: Akteure und ihre datenschutzrechtlichen Rollen

Im Zusammenhang mit der Analyse der datenschutzrechtlichen Verantwortlichkeit des Avatars wurde darüber hinaus diskutiert, ob ein Avatar mit einer eigenen Rechtspersönlichkeit ausgestattet werden sollte. Dies ist insbesondere wegen der fehlenden finanziellen Motivation, die natürliche und juristische Personen haben, um geltendes (Datenschutz-)Recht einzuhalten, zu verneinen.

Perspektivisch scheint die Rolle des Avatar-Inspirators des virtuellen Weiterlebens aus datenschutzrechtlicher Sicht von besonderer Bedeutung zu sein, da dieser nach seinem Ableben keinen direkten Einfluss mehr auf den Avatar und die damit verbundenen Verarbeitungsprozesse nehmen kann. Aber auch in Bezug auf potenzielle Verletzungen seines postmortalen Persönlichkeitsschutzes scheint die Rolle des Avatar-Inspirators vor diesem Hintergrund besonders bedeutungsvoll zu sein, so dass mögliche Risiken für diese Rolle Gegenstand zukünftiger vertiefter Forschung sein sollte.

Danksagung

Diese Forschungsarbeiten wurden vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen des Projektes „Edilife“ (Förderkennzeichen: 6INS114B) sowie vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt. Der Beitrag gibt die persönliche Meinung der Autorin wieder.

Literaturverzeichnis

- [AAM23] Athar, A.; Ali, S. M.; Mozumder, M. A. I.; Ali, S.; Kim, H. - C.: "Applications and Possible Challenges of Healthcare Metaverse," 25th International Conference on Advanced Communication Technology (ICACT), S. 328-332, 2023.
- [Be09] Beck, S.: Grundlegende Fragen zum rechtlichen Umgang mit der Robotik, JR, S. 225-230, 2009.

- [Bi22] Bitkom e.V.: Wegweiser in das Metaverse – Technologische und rechtliche Grundlagen, geschäftliche Potenziale, gesellschaftliche Bedeutung, 2022.
- [Bl20] Bleckat, A.: Anwendbarkeit der Datenschutzgrundverordnung auf künstliche Intelligenz, DuD, S. 194-198, 2020.
- [BM21] Bomhard, D.; Merkle, M.: Europäische KI-Verordnung, Rdi, S. 276-283, 2021.
- [Bo22] Bossmann, O.: Das Metaverse – Schöne neue Zukunft oder Datenschutz-Albtraum, Neuss 2022.
- [BW23] Bender-Paukens, L.; Werry, S.: Datenschutz im Metaverse, ZD, S. 127-131, 2023.
- [Dj22] Djeflal, C.: „Soziale Medien und Kuratierung von Inhalten. Regulative Antworten auf eine demokratische Schlüsselfrage.“ In: Spiecker, I.: Demokratie und Öffentlichkeit im 21. Jahrhundert – zur Macht des Digitalen. Nomos Verlagsgesellschaft mbH & Co. KG, S. 177-189, 2022.
- [Eu17] EU-Parlament, Entschließung vom 16. Februar 2017 mit Empfehlungen an die Kommission zu zivilrechtlichen Regelungen im Bereich Robotik, 2017.
- [Eu18] EuGH, Urteil vom 05.06.2018 – C-210/16.
- [Eu19] EU-Kommission, Schaffung von Vertrauen in eine auf den Menschen ausgerichtete künstliche Intelligenz, COM, 2019.
- [HSH22] Hoeren, T.; Sieber, U.; Holznapel, B.: Handbuch Multimediarecht, 4. Auflage, C. H. Beck, 2022.
- [In22] Interpol Technology Assessment Report on the Metaverse, 2022.
- [KM15] Kirn, S.; Müller-Hengstenberg, C.: Technische und rechtliche Betrachtungen zur Autonomie kooperativ-intelligenter Softwareagenten, Künstliche Intelligenz 29, S. 59–74, 2015.
- [KSM22] Kaulartz, M.; Schmid, A.; Müller-Eising, F.: Das Metaverse – eine rechtliche Einführung, RD 2022, S. 521-532.
- [LWG21] Leupold, A.; Wiebe, A.; Glossner, S.: IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, 4. Auflage, C. H. Beck, 2021.
- [Ri20] Riehm, T.: Nein zur ePerson!, Rdi, S. 42-48, 2020.
- [SM20] Savin-Baden, M.; Mason-Robbie, V.: Digital Afterlife – Death Matters in a Digitale Age, CRC Press, 2020.
- [Su22] Sury, U.: Metaverse – parallele Welt(en), Informatik Spektrum, S. 407-409, 2022.
- [TG22] Taeger, J.; Gabel, D.: Kommentar DSGVO – BDSG – TTDSG, 4. Auflage, C. H. Beck, 2022.
- [VAR12] Vernaza, A; Armuelles, V. I.; Ruiz, I.: "Towards to an open and interoperable virtual learning enviroment using Metaverse at University of Panama," Technologies Applied to Electronics Teaching, S. 320-325, 2012.
- [WB23] Wolff, H. A.; Brink, S.: Beck'scher Online-Kommentar Datenschutzrecht, 43. Edition, C. H. Beck, 2023.

- [WZ03] Wettig, S.; Zehendner, E.: The Electronic Agent: A Legal Personality under German Law?, <https://beck-link.de/kc3tt>, Stand: 15.05.2023.

Leben in Metaversen und im Virtual Afterlife

Chancen und Risiken zukünftiger Formen des virtuellen (Weiter-)Lebens

Ines Geissler¹

Abstract: Metaversen – als Form des virtuellen Lebens – und Umgebungen des virtuellen Weiterlebens bieten großes Potenzial der Datenökonomie für Unternehmen. Auch Menschen nutzen die virtuellen Welten immer mehr, um mit anderen Personen zu interagieren oder um mit Verstorbenen „in Kontakt“ zu bleiben. Doch neben den vorgenannten Chancen bergen diese zukünftigen Formen des Lebens und Weiterlebens auch Risiken für die Rechte und Freiheiten natürlicher Personen. Diese Chancen und Risiken werden im vorliegenden Papier beleuchtet.

Keywords: Datenschutz, Datenschutz-Grundverordnung, DSGVO, Virtuelles Leben, Virtuelles Weiterleben, Metaversen.

1 Metaversen und virtuelles Weiterleben: Ein Einblick in zukünftige Formen des Lebens

Der Begriff „Metaverse“ hat seinen Ursprung in den 1990er Jahren und wird heute häufig als Überbegriff für eine virtuelle Welt verstanden, doch eine universelle Definition des Begriffes gibt es nicht [Bi22]. Allgemeine Einigkeit besteht darin, dass Metaversen virtuelle Umgebungen darstellen, die Menschen aus verschiedenen Lebensbereichen verbinden und ihnen die Möglichkeit geben, mittels sogenannter Avatare² miteinander zu interagieren und somit das soziale Miteinander der Nutzer zu bestärken [Ka22]. Das Ziel von Metaverse-Anbietern geht jedoch noch weiter: sie wollen i. d. R. eine virtuelle Welt erschaffen, die dem Nutzer neue Interaktionsmöglichkeiten mithilfe von Augmented³ und Virtual⁴ Reality Technologien ermöglichen sollen. Die Metaverse-Anbieter stellen dabei die Plattform bereit, auf die Nutzer⁵ zugreifen können und über die andere Dienstleister auf der virtuellen Metaverse-Umgebung weiterführende Dienste anbieten können. Diese

¹ Fraunhofer-Institut für Sicherer Informationstechnologie SIT | ATHENE – Nationales Forschungszentrum für angewandte Cybersicherheit, IT Law & Interdisciplinary Privacy Research, Rheinstr. 75, 64295 Darmstadt, ines.geissler@sit.fraunhofer.de, <https://orcid.org/0009-0001-3416-6768>.

² Avatare sind virtuelle Zwillinge der Nutzer, die es diesen ermöglichen, sich durch die virtuelle Welt zu bewegen und mit anderen Nutzern zu interagieren [Su22].

³ Mithilfe von Augmented Reality Technologien werden analoge Umgebungen abgebildet und durch virtuelle, computergenerierte Informationen ergänzt. Diese Umgebungen sind dreidimensional und kombinieren Objekte der analogen Welt mit virtuellen Objekten [Ca11].

⁴ Virtual Reality Technologien sollen Nutzer komplett in eine virtuelle Welt eintauchen lassen, in der sie sich präsent fühlen. Die Immersion in diese Welt erfolgt durch die Erzeugung von künstlichen Reizen für die visuelle, auditive und teilweise auch haptische Wahrnehmung [Dö16].

⁵ Alle Geschlechter sind hiermit gemeint.

virtuelle Welt kommt der analogen Welt in optischer und funktionaler Hinsicht sehr nahe. So gibt es bspw. virtuelle Wohnhäuser, Einkaufszentren und Cafés, in denen sich die Metaverse-Nutzer mithilfe der Augmented und Virtual Reality Technologien bewegen können. Nutzer können mittels Headsets oder anderer Ausstattung immersiv in virtuelle Welten eintauchen oder Ereignisse in der physischen Welt mit virtuellen Interaktionen verschmelzen lassen [KSM22]. Metaversen enthalten neben der sozialen Kommunikation somit i. d. R. auch ein funktionsfähiges Wirtschaftssystem, in dem sich Nutzer und an einem Metaverse teilnehmende Unternehmen bspw. Besitz erschaffen oder Investitionen tätigen können [Su22]. So wie das Internet selbst, werden auch die virtuellen Räume der Metaversen rund um die Uhr verfügbar sein [Bi22].

Mit der fortwährenden Digitalisierung und der Weiterentwicklung des Internets wird auch der Tod von Menschen zunehmend präsent in virtuellen Welten. Im Kontext des virtuellen Weiterlebens ermöglichen Technologieanbieter das virtuelle Weiterleben und die Interaktion von virtuellen Abbildnissen von Verstorbenen in Form von Avataren, Chatbots oder Ähnlichem [SB19]. Derartige Dienste setzen i. d. R. Technologien der Künstlichen Intelligenz (KI) ein, die mithilfe des Avatar-Inspirators⁶ einen weiterlebenden Avatar anlernen, der nach dem Tod des Avatar-Inspirators möglichst realgetreu mit den Hinterbliebenen⁷ interagiert [JB23]. Für das Anlernen werden zu Lebzeiten z. B. Sprachnachrichten, Textnachrichten oder Briefe von der KI ausgewertet. Dem Avatar können sodann bspw. Fragen gestellt werden, die dieser so beantwortet, wie es der Avatar-Inspirator vermutlich getan hätte [SM20].

Im Rahmen der Nutzung dieser neuen Formen des virtuellen (Weiter-)Lebens hinterlassen Menschen große Mengen an personenbezogenen Daten im Internet, die (ggf. auch nach ihrem Ableben) dort verbleiben und nicht nur die Daten der Profilinhaber in Metaversen bzw. der Avatar-Inspiratoren im virtuellen Weiterleben, sondern auch personenbezogene Daten der Kommunikationspartner dieser Personen umfassen. Dadurch entstehen umfangreiche Profile, die für Unternehmen besonders interessant sind, weil sie daraus das Verhalten, Interessen und Gewohnheiten lebender Personen analysieren können [He22]. Dies ermöglicht den Unternehmen, mit diesem Dienst personalisierte Werbung zu schalten und Angebote zu unterbreiten und das Dienstangebot fortlaufend zu verbessern und auszuweiten.

Metaversen und das virtuelle Weiterleben wurden bisher aus datenschutzrechtlicher Sicht nicht tiefgehend betrachtet, obwohl es notwendig wäre, die Rechtmäßigkeit der Datenverarbeitung in Metaversen und in Umgebungen des virtuellen Weiterlebens zu beleuchten. Es bedarf einer Analyse der Rechtmäßigkeit

- etwa in Bezug auf das Erstellen von Avataren zu einer lebenden oder einer bereits

⁶ Als Avatar-Inspirator bezeichnet dieser Beitrag alle Personen, die zu ihren Lebzeiten einen KI-gestützten Avatar anlernen, damit dieser nach ihrem Ableben mit den Hinterbliebenen des Avatar-Inspirators (weiter-)kommunizieren kann. Alle Geschlechter sind hiermit gemeint.

⁷ Als Hinterbliebene bezeichnet dieser Beitrag alle Personen, die mit einem weiterlebenden Avatar nach dem Tod des Avatar-Inspirators (weiter-)kommunizieren.

verstorbenen Person ohne Einverständnis des Avatar-Inspirators und ggf. seiner Hinterbliebenen,

- in Bezug auf die Erhebung und Auswertung von (Gesundheits-)Daten durch Dritte, wie bspw. Krankenhäuser, Gesundheitsämter und Krankenkassen,
- aber auch in Bezug auf Datenerhebungen und -auswertungen in Metaversen durch Strafverfolgungsbehörden,

um für die spezifischen Anwendungen mögliche Schutzlücken zu identifizieren, die in die Rechte und Freiheiten natürlicher Personen eingreifen können und denen insofern mit rechtlichen Schutzmechanismen und/oder Mechanismen des (technischen) Selbst Datenschutzes zu begegnen ist.

2 Stand der Wissenschaft

In Bezug auf Metaversen wurden bisher vor allem grundlegende Fragen zur Funktionsweise der Metaversen beantwortet und ein erster Überblick über die verschiedenen rechtlichen Probleme geschaffen [Bo22], [Ch22], [Su22]. Insbesondere erfolgten nur wenige, verhältnismäßig oberflächliche datenschutzrechtliche Betrachtungen der Rechtmäßigkeit der Metaversen. Ebenso wenig wurden rechtliche Schutzmaßnahmen und Mechanismen des Selbst Datenschutzes in Metaversen erörtert.

Aufarbeitungspotenzial besteht auch im Zusammenhang mit dem virtuellen Weiterleben. Bereits erfolgte wissenschaftliche Untersuchungen beziehen sich auf die Notwendigkeit bzw. auf den notwendigen Umfang eines postmortalen Datenschutzrechts im Allgemeinen [Ku20, Ra18, Ma15], [Ne16] auf die Umsetzung datenschutzrechtlicher Anforderungen im Kontext des Einsatzes von KI im Allgemeinen [Bo19], [Sc19] und vereinzelt im Speziellen auf die Betroffenenrechte bei der Verarbeitung von personenbezogenen Daten mittels KI [Wi18].

3 Chancen und (Datenschutz)risiken des virtuellen Lebens in Metaversen

3.1 Chancen des virtuellen Lebens in Metaversen

Metaversen bieten Menschen zunächst die Chance, sich in einem virtuellen gemeinsamen Raum mit anderen Menschen auf der ganzen Welt zu vernetzen und in computergenerierten, virtuellen Umgebungen wie Cafés, Restaurants und Bildungseinrichtungen in Echtzeit miteinander zu interagieren. So ergeben sich neue Möglichkeiten der Freizeitgestaltung und im Bereich Unterhaltung, insbesondere in Bezug auf Gaming, virtuelle Konzerte sowie virtuelle Kinobesuche [Bo22].

Doch nicht nur für Menschen ergeben sich in Metaversen Vorteile: Die COVID-19-Pandemie hat das Wachstum virtueller Erfahrungen beschleunigt. Viele Branchen erkunden das Potenzial der Metaversen und es entstehen neue Arten von elektronischem Handel und Produktion [Bo22]. Menschen sehen virtuelle Räume als gleichwertig zu realen Räumen an, so dass Erfahrungen in der virtuellen Welt (wie z. B. im Bereich des eSports) weiter an Bedeutung gewinnen werden [Bi22]. Auch können die virtuellen Repräsentanten der Metaverse-Nutzer in Form von Avataren virtuelle Produkte kaufen, was ebenfalls ein großes wirtschaftliches Potenzial für in Metaversen agierende Unternehmen darstellt. Darüber hinaus können in Metaversen u.a. umfangreiche Produktionsprozesse simuliert werden, um sie zunächst virtuell und sodann in der analogen Realität zu optimieren, was Unternehmen zu Kosteneinsparungen verhelfen kann [Ne22].

Des Weiteren können über Metaversen auch umfangreiche Weiterbildungsmöglichkeiten angeboten werden [Bi22], die sowohl den privaten Bereich (z. B. Teilnahme an einem virtuellen Workshop zur Gesundheitsvorsorge) als auch den beruflichen Bereich (z. B. Teilnahme an einem in Metaversen abgehaltenen Workshop über eine Programmiersprache) betreffen können.

Metaversen bieten somit neue Arbeitsbedingungen und verändern zugleich gesellschaftlich relevante Bereiche, wie etwa den Gesundheitssektor, Bildung, politische Teilhabe und öffentliche Dienste. Diese Möglichkeiten treiben die Tech-Industrie an, Produkte und Dienstleistungen für Metaversen zu entwickeln und es scheint, dass die Weiterentwicklung von Metaversen ein unvermeidlicher Schritt mit Hinblick auf die Konvergenz von Mensch und Technologie sein wird [Bo22].

3.2 (Datenschutz)risiken des virtuellen Lebens im Metaverse

Neben den Möglichkeiten, die Metaversen bieten, bestehen zahlreiche Risiken, insbesondere hinsichtlich des Datenschutzrechts und der Ethik.

Zunächst besteht die Gefahr der fehlenden Transparenz: in Metaversen können eine Vielzahl an Stellen personenbezogene Daten verarbeiten (neben dem Metaverse-Anbieter auch zahlreiche Unternehmen, die über die Metaverse-Plattform des Metaverse-Anbieters Dienste wie z. B. das Veranstalten von virtuellen Workshops, das Einrichten eines virtuellen Kaufhauses oder die Organisation eines virtuellen Konzerts anbieten) [Bi22]. Die jeweiligen Plattformen können Unmengen an personenbezogenen Daten ihrer Nutzer sammeln, die besonders sensibel sein können. Dazu gehören Daten über Standort, Verhalten, Interessen und Vorlieben, in Interaktionen mit anderen Metaverse-Nutzern preisgegebene Informationen zum Gesundheitszustand, zur sexuellen Orientierung oder zur politischen Einstellung, aber auch zum Teil Stimmtdaten, etwa wenn Avatare miteinander kommunizieren [Bo22]. Dabei handelt es sich besondere Kategorien personenbezogener Daten i. S. d. Art. 9 DSGVO, deren Verarbeitung aufgrund ihrer Sensibilität einem besonders hohen Schutz unterliegt. Dieser Umstand beruht auf der

Tatsache, dass bestimmte Arten personenbezogener Daten aufgrund ihrer höheren Schutzbedürftigkeit ein erhöhtes Risiko für die Verletzung der Rechte betroffener Personen bergen.

Durch die Vielzahl der datenverarbeitenden Stellen kann es zu fehlender Transparenz für betroffene Personen kommen, da diese nicht überblicken können, wer ihre personenbezogenen Daten verarbeitet. Da für die Datenverarbeitung Verantwortliche die Anforderungen aus der DSGVO auch in Metaversen erfüllen müssen, muss folglich auch im Zusammenhang mit der Datenverarbeitung in Metaversen Transparenz geschaffen werden [BW23]. Dies ist jedoch aufgrund der Masse an Verarbeitungskontexten, datenverarbeitenden Stellen und Produkten in Metaversen höchst problematisch, so dass für die in Metaversen agierenden natürlichen Personen das Risiko besteht, nicht einschätzen zu können, welche Stellen welche Daten über sie verfügen und verarbeiten. Dieser Umstand kann zu Unsicherheiten bei natürlichen Personen und zu Hemmungen in ihrem Verhalten führen, so dass sie sich in ihrer Persönlichkeit einschränken [Bu83]. Außerdem führt diese fehlende Transparenz dazu, dass betroffene Personen nicht die Möglichkeit haben, ihre Rechte geltend zu machen, weil ihnen mögliche Rechtsverstöße nicht bekannt sind [GH22].

Weiterhin besteht die Möglichkeit, dass diese Daten durch Dritte, wie Strafverfolgungsbehörden [In22] oder Versicherer [KS22], abrufbar sein können, um einerseits Straftaten zu verfolgen und andererseits bspw. Versicherungsleistungen zu gewähren oder zu verwehren. Unklar ist, ob und in welchem Umfang diese Stellen Daten aus Metaversen nutzen dürfen, um diese Zwecke zu erfüllen. Auch wenn die DSGVO grundsätzlich Anwendung auf Metaverseumgebungen findet, wurde sie dennoch nicht für diese konzipiert [Eu22]. Ein Risiko in einer umfangreichen Erhebung besteht darin, dass aus der großen Masse an Daten Persönlichkeitsprofile gebildet werden können. Diese Daten können neben den bereits genannten Informationen – im Kontext der Metaversen speziell – z. B. Reaktionen und Verhaltensweisen umfassen, die von (staatlichen) Institutionen verarbeitet werden können [Eu22]. Wenn eine umfangreiche Verarbeitung dieser Daten z. B. auch in privaten Räumen der Metaversen erfolgt, geht damit das Risiko einher, dass Menschen sich zurückhalten, auch im privaten Umfeld ihre Meinung zu äußern und ihr Verhalten einschränken aus Angst, sie könnten negativ bei Strafverfolgungsbehörden oder Versicherungsträgern auffallen.

Des Weiteren ist das Interagieren in Metaversen mit anderen Avataren oder Unternehmen grundsätzlich mit einem Diskriminierungspotenzial verbunden [Dj22]. So ist es z. B. vorstellbar, dass Plattformanbieter Chat-Nachrichten auswerten, um so Rückschlüsse auf mögliche psychische Erkrankungen zu ziehen. Diese Art der Datenanalyse könnte dazu führen, dass bestimmte Nutzer diskriminiert werden, bspw. weil ihnen bestimmte Dienstleistungen oder Angebote verwehrt werden. Andererseits könnten betroffene Personen gezielt aufgrund dieser persönlichen Informationen durch Werbetreibende angesprochen werden.

Außerdem ergeben sich Risiken in Bezug auf das Allgemeine Persönlichkeitsrecht aus

Art. 2 Abs. 1 in Verbindung mit Art. 1 Abs. 1 GG. Im Kontext der Metaversen sind Einschränkungen des Rechts denkbar, z. B., wenn die eigene Identität unerlaubt durch Dritte in Form eines Avatars nachgebildet wird. Die Risiken ergeben sich unter anderem daraus, dass unklar ist, welche Mittel und Maßnahmen den Nutzern zur Verfügung stehen, um derartige Nachbildungen zu verhindern oder für die Zukunft zu unterbinden.

Aus ethischer Sicht ergibt sich das Risiko, dass Menschen aus ihrem Alltag in die virtuelle Welt flüchten und sich dort ein ideales Selbst und ein ideales Leben aufbauen und kaum noch am Leben in der analogen Welt teilnehmen. Dies kann gesundheitliche Schäden verursachen und sich negativ auf das soziale Leben auswirken [SM20] und den Verlust des Freundeskreises oder sogar den Verlust des Arbeitsplatzes bedeuten.

4 Chancen und (Datenschutz)risiken des virtuellen Weiterlebens

4.1 Chancen des virtuellen Weiterlebens

Digitale Technologien bestimmen zunehmend das menschliche Leben und damit auch den Umgang mit dem Tod, dem Erinnern und dem Trauern. Einige Dienstleistungen des virtuellen Weiterlebens sollen Hinterbliebenen helfen, mit den Verstorbenen „in Kontakt“ zu treten bzw. das Gefühl entstehen lassen, mit der verstorbenen Person z. B. in Form eines KI-basierten Avatars oder eines Chatbots zu interagieren. Nach dem Ableben kann der Avatar den Hinterbliebenen Trost spenden und sie durch ihr Leben begleiten und sie bei wichtigen Entscheidungen unterstützen [SM20].

Die bei der Nutzung dieser Technologien hinterlassenen Daten bieten Technologieunternehmen die Chance, neue Märkte zu erschließen und Gewinne zu erzielen. Diese Unternehmen können die Daten nutzen, um bestehende Angebote zu erweitern und die Dienstleistungen so zu verbessern, dass sie den Wünschen der Hinterbliebenen entsprechen und diese weiter mit den Avatar-Inspiratoren in Kontakt bleiben können [VU19].

Dienste des virtuellen Weiterlebens sind bereits von Bedeutung in der Erinnerungskultur. So können bspw. Museumsbesucher mit Avataren von Holocaustüberlebenden sprechen und Fragen stellen, die diese beantworten [St20]. Die Fortschritte und Möglichkeiten, die durch KI bereits bestehen, beschleunigen die Entwicklung der Angebote im virtuellen Weiterleben weiterhin. Zukünftig wird daher damit zu rechnen sein, dass immer mehr Menschen diese Dienste auch im privaten Bereich nutzen.

4.2 (Datenschutz)risiken des virtuellen Weiterlebens

Die Nutzung von Angeboten des virtuellen Weiterlebens birgt verschiedene Risiken, darunter auch in Bezug auf den Datenschutz und die Ethik.

Ein datenschutzrechtliches Risiko besteht in der Tatsache, dass bei Diensten des virtuellen Weiterlebens personenbezogene Daten – und sogar eine hohe Anzahl besonderer Kategorien personenbezogener Daten im Sinne des Art. 9 DSGVO – verarbeitet werden können, ohne dass dies den betroffenen Personen bewusst ist. Grundsätzlich ist es möglich, einen KI-basierten Avatar neben Textdaten auch mit Video- und Audiodaten anzulernen, damit dieser möglichst wie die abzubildende Person aussieht und sich auch wie diese anhört [SM20]. Der Prozess des Anlernens wird i. d. R. von dem Avatar-Inspirator initiiert.⁸ Unklar ist in dem Zusammenhang, ob und in welcher Form der Avatar-Inspirator die Hinterbliebenen, die ebenfalls in den Text-, Video- und Audiodateien enthalten sein können, darüber informiert bzw. informieren muss. Diese Dateien können Informationen über die physischen Merkmale, die Gesundheit, die Emotionen oder die sozialen Interaktionen einer Person offenbaren. Diese Informationen sind sehr sensibel und können (seitens der Hinterbliebenen) nur für den familiären Kreis bestimmt gewesen sein, so dass eine unfreiwillige Offenlegung dieser Daten die Privatsphäre unrechtmäßig verletzen könnte. Dieser umfangreichen Verarbeitung personenbezogener Daten und besonderer Kategorien personenbezogener Daten muss auch beim Anbieter des Avatars durch angemessene Schutzmaßnahmen begegnet werden.

Weiterhin besteht das (rechtliche und ethische) Risiko, dass nach dem Tod des Avatar-Inspirators durch die KI unkontrollierbare Wesensveränderungen des Avatars entstehen und dieser sich negativ verändert und nicht mehr den Wünschen und Vorstellungen des Avatar-Inspirators entspricht [SM20]. Da der Avatar-Inspirator nach seinem Tod keinen unmittelbaren Einfluss auf den Avatar hat, könnte dies die Würde und die Persönlichkeit beeinträchtigen. Zwar besteht das postmortale Persönlichkeitsrecht gemäß Art. 2 Abs. 1 in Verbindung mit Art. 1 Abs. 1 GG, das die Fortwirkung des Persönlichkeitsschutzes über den Tod hinaus schützt. Vor diesem Hintergrund ist allerdings unklar, ob dieser Schutz für diese neuartigen Risiken ausreicht oder ob der Schutz ausgeweitet werden muss, um diesen entgegenzuwirken.

Ein weiteres (rechtliches und ethisches) Risiko besteht im Zusammenhang mit der Vorsorge beim virtuellen Weiterleben. Hier könnten ggf. die Interessen von Hinterbliebenen und Avatar-Inspiratoren an der Weiterführung der Datenverarbeitung nach dem Tod des Avatar-Inspirators entgegenstehen [KM15]. Z. B. könnte der Avatar-Inspirator im Rahmen einer Verfügung von Todes wegen nach § 2247 BGB vor seinem Tod festgelegt haben, ob und wann der Avatar und alle in dem Zusammenhang verarbeiteten Daten gelöscht werden sollen. Die Hinterbliebenen könnten jedoch – evtl. den Interessen des Avatar-Inspirators entgegenstehend – daran interessiert sein, dass der Avatar-Inspirator nicht durch den Avatar weiterlebt, weil sie nicht möchten, dass die Erinnerung an den Verstorbenen durch Erinnerungen mit der KI überschrieben werden. Denkbar ist andererseits, dass die Hinterbliebenen länger mit dem Avatar kommunizieren

⁸ Der Prozess des Anlernens kann entweder zum Zeitpunkt des Todes des Avatar-Inspirators abgeschlossen sein oder noch nach dem Tod bspw. durch die Kommunikation mit Hinterbliebenen fortgeführt werden. Im Folgenden werden jene Varianten betrachtet, bei denen der Trainingsprozess auch nach dem Tod weitergeführt wird.

möchten, als dies vom Avatar-Inspirator vorgesehen wurde. Uneinigkeit könnte auch in Bezug auf den Personenkreis, der dem Avatar zur Verfügung gestellt werden soll, bestehen. Der Avatar könnte bspw. nur ausgewählten Personen oder gar der breiten Öffentlichkeit verfügbar gemacht werden.

Darüber hinaus bestehen weitere rein ethische Risiken. Der Einsatz von Diensten des virtuellen Weiterlebens kann einerseits den Trauerprozess von Hinterbliebenen verlängern [ÖF17] und andererseits den Avatar-Inspirator so sehr einschränken, dass er nicht mehr am eigentlichen Leben teilnimmt, weil er sich nur noch auf den idealen, weiterlebenden Avatar, nicht mehr aber auf das analoge Leben seine realen Bezugspersonen fokussiert.

5 Zusammenfassung

Metaversen und das virtuelle Weiterleben bergen einerseits eine Vielzahl an Chancen für Menschen und Unternehmen, wie z. B.:

1. neue Gestaltungsmöglichkeiten des virtuellen Lebens und des virtuellen Weiterlebens.
2. zusätzliche Markterschließungsmöglichkeiten für Unternehmen sowie
3. das Vorantreiben des technologischen Fortschritts.

Andererseits sind Metaversen und das virtuelle Weiterleben mit Risiken für die Rechte und Freiheiten der von personenbezogenen Datenverarbeitungen betroffenen Personen verbunden, wie z. B.

1. durch die große Menge an Verarbeitungen von personenbezogenen Daten besonderer Kategorien,
2. bei Eingriffen in die Persönlichkeitsrechte natürlicher Personen durch Nachbildungen von lebenden oder verstorbenen Personen und
3. in Bezug auf die fehlende Transparenz bei komplexen Datenverarbeitungsprozessen.

Zukünftig wird es notwendig sein, Maßnahmen zu ergreifen, um den in diesem Beitrag identifizierten Risiken zu begegnen und so den betroffenen Personen die genannten Chancen zu ermöglichen, ohne dass dabei ihre Rechte und Freiheiten unangemessen eingeschränkt werden.

Danksagung

Diese Forschungsarbeiten wurden vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen des Projektes „Edilife“ (Förderkennzeichen: 6INS114B) sowie vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt. Der Beitrag gibt die persönliche Meinung der Autorin wieder.

Literaturverzeichnis

- [Bi22] Bitkom e.V.: Wegweiser in das Metaverse - Technologische und rechtliche Grundlagen, geschäftliche Potenziale, gesellschaftliche Bedeutung, 2022.
- [Bo19] Bohnstedt, J.: Vom Personenbezug zum Gerätebezug – KI und Datenschutz, DSRITB, S. 409-421, 2019.
- [Bo22] Bossmann, O.: Das Metaverse – Schöne neue Zukunft oder Datenschutz-Albtraum, Neuss 2022.
- [Bu83] BVerfG, Urteil des Ersten Senats vom 15. Dezember 1983 - 1 BvR 209/83, 1983.
- [BW23] Bender-Paukens, L.; Werry, S.: Datenschutz im Metaverse, ZD, S. 127-131, 2023.
- [Ca11] Carmigniani, J.; Furht, B.; Anisetti, M.; Ceravolo, P.; Damiani, E.; Ivkovi, M.: Augmented reality technologies, systems and applications, Multimedia Tools and Applications, S. 341–377, 2011.
- [Ch22] Cheong, B. C.: Avatars in the metaverse: potential legal issues and remedies, International Cybersecurity Law Review, 2022, S. 467-494, 2022.
- [Dö16] Dörner, R.; Broll, W.; Grimm, P.; Jung, B.: Virtual Reality und Augmented Reality (VR/AR), Informatik-Spektrum, S. 30-37, 2016.
- [Dj22] Djefal, C.: "Soziale Medien und Kuratierung von Inhalten. Regulative Antworten auf eine demokratische Schlüsselfrage." In: Spiecker, Indra: Demokratie und Öffentlichkeit im 21. Jahrhundert–zur Macht des Digitalen. Nomos Verlagsgesellschaft mbH & Co. KG, S. 177-189, 2022.
- [Eu22] European Parliamentary Research Service: Metaverse Opportunities, risks and policy implications, 2022
- [GH22] Gola, P.; Heckmann, D.: DSGVO, BDSG, C. H. Beck, 2022.
- [He22] Heesen, J.: Verstorbene als Medienprodukt, In: Wolfgang George & Karsten Weber (Hrsg.): Fehlendes Endlichkeitsbewusstsein und die Krisen im Anthropozän, 161–172, 2022.
- [In22] Interpol Technology Assessment Report on the Metaverse, 2022.
- [JB23] Jiménez-Alonso, B.; Brescó de Luna, I.: Griefbots. A New Way of Communicating With

- The Dead?, Integrative Psychological and Behavioral Science volume, S. 466–481, 2023.
- [Ka22] Kammler, P.; Gravemeier, L. S.; Göritz, L.; Kammler, F.; Gembariski, P. C.: Values of the Metaverse: Hybride Arbeit in virtuellen Begegnungsräumen, HMD, Praxis der Wirtschaftsinformatik, S. 1062–1074, 2022.
- [KM15] Klas, B.; Möhrke-Sobolewski, C.: Digitaler Nachlass – Erbenschutz trotz Datenschutz, NJW, 3473, 2015.
- [Ku20] Kubis, M.; Mayr, M.; Selzer, A.; Sperlich, T.; Steiner, S.; Waldmann, U.: Der digitale Nachlass - Eine Untersuchung aus rechtlicher und technischer Sicht, Stuttgart 2020.
- [Kr14] Kreße, B.: Entschädigungshöhe bei Persönlichkeitsrechtsverletzungen im Internet, NJ, S. 159-162, 2014.
- [KS22] Knoll, M.; Stieglitz, S.: Augmented Reality und Virtual Reality – Einsatz im Kontext von Arbeit, Forschung und Lehre, HMD, S. 6-22, 2022.
- [KSM22] Kaulartz, M.; Schmid, A.; Müller-Eising, F.: Das Metaverse – eine rechtliche Einführung, RDi, S. 521-532, 2022.
- [Ma15] Martini, M.: Trauer 2.0 – Rechtsfragen digitaler Formen der Erinnerungskultur, WiVerw, S. 35 ff., 2015.
- [Ne16] Netzwerk Datenschutzexpertise: Postmortaler Datenschutz - Auskunftsansprüche von Erben und Angehörigen zu personenbezogenen Internetdaten eines Verstorbenen, 2016.
- [Ne22] Neuburger, R.: Arbeiten in und mit dem Metaverse: Eine realistische Vision?. Wissensmanag. 4, S. 16–21, 2022.
- [ÖF17] Öhman, C.; Floridi, L.: The Political Economy of Death in the Age of Information: A Critical Approach to the Digital Afterlife Industry. Minds & Machines 27, S. 639–662, 2017.
- [Ra18] Rasta, G.: Personal Data and Digital Assets after Death: a Comparative Law Perspective on the BGH Facebook Ruling, EuCML S, 201 ff., 2018.
- [SB19] Savin-Baden, M.; Burden, D.: Digital Immortality and Virtual Humans, Postdigit Sci Educ 1/2019, S. 87-103, 2019.
- [Sc19] Schindler, S.: Künstliche Intelligenz und (Datenschutz-)Recht. ZD-Aktuell, 06647, 2019.
- [SM20] Savin-Baden, M.; Mason-Robbie, V.: Digital Afterlife – Death Matters in a Digital Age, CRC Press, 2020.
- [St20] Stahl, L.: Artificial intelligence is preserving our ability to converse with Holocaust survivors even after they die, <https://www.cbsnews.com/news/holocaust-stories-artificial-intelligence-60-minutes-2020-04-05/>, Stand: 20.04.2023.
- [Su22] Sury, U.: Metaverse – parallele Welt(en), Informatik Spektrum, S. 407-409, 2022.
- [VU19] Voinea, C.; Uszkai, R.: An ethical framework for digital afterlife industries. In Proceedings of the 13th International Management Conference „Management Strategies for High Performance” 31st October – 1st November 2019, Bucharest, 2019.

- [Wi18] Wieder, C.: Datenschutzrechtliche Betroffenenrechte bei der Verarbeitung von personenbezogenen Daten mittels künstlicher Intelligenz, Datenschutzrechtliche Betroffenenrechte bei der Verarbeitung von personenbezogenen Daten mittels künstlicher Intelligenz, in: Taeger, J. (Hrsg.). Rechtsfragen digitaler Transformationen. Edewecht, S. 505–518, 2018.

Data Privacy and Security in the Context of Corporate Digital Responsibility: A Scoping Review


K. Valerie Carl ¹

Abstract: The continuous digitalization affects private and professional lives alike, providing new chances but also threats for companies, consumers, and society. Particularly data privacy and security risks remain central for consumers. In this evolving setting, the concept of Corporate Digital Responsibility (CDR) gains traction. CDR provides a framework for the responsible application of digital technologies, thereby putting, inter alia, data privacy and security in a larger context. A remarkable amount of prior Information Systems (IS) research can be linked to facets of the concept of CDR. Hence, this study pursues the goal to evaluate the comprehensiveness of current research in IS, also guiding future research efforts. To address this goal, this study grounds on a scoping review to evaluate previous popularity in IS research. The results illustrate a need for a more comprehensive view on data privacy and security in the larger context of CDR in the IS discipline.

Keywords: Corporate Digital Responsibility, Information Systems, Scoping Review

1 Introduction

The ongoing digitalization allows for a plethora of digital products and services ranging from Big Data to more sophisticated Artificial Intelligence (AI) applications. These digital products share new chances for value creation but also condition new obstacles and threats [e.g., Sp22] in part due to the increasing amount of processed consumer data. In the course of the ongoing digitalization, central ethical and societal challenges concerning autonomy, balance of power, human dignity, justice, privacy, and security emerged [Ro18]. Consumers demand for more responsible and ethical corporate behavior, particularly concerning data privacy and security [e.g., Ma86, Mi21], thereby highly valuing, e.g., access and correction opportunities to personal data beyond legal requirements as well as an individualized information approach to data practices [Ca23]. Such activities bear the ability to influence consumers' perceptions of companies positively. Hence, research and practice account for that by developing rich insights into a comprehensive approach towards ethical corporate behavior. Within this evolving context, the concept of Corporate Digital Responsibility (CDR) gains traction. CDR provides guidance for companies on how to exploit the opportunities of digitalization while adequately addressing its risks by putting ethical implications of corporate behavior in the broader context of manifold corporate responsibilities. Above all, CDR provides a more comprehensive view on

¹ Goethe University Frankfurt/Main, Chair of Information Systems and Information Management, Theodor-W.-Adorno-Platz 4, D-60323 Frankfurt/Main, Germany, kcarl@wiwi.uni-frankfurt.de, 
<https://orcid.org/0000-0003-4655-1046>

corporate responsibilities in the digital context and allows a linked consideration and evaluation of the well-known, established individual sub-fields (e.g., data privacy and security). CDR applies to various kinds of digital technologies, products, and services despite their diverse characteristics and ethical requirements. The concept of CDR was highly practice-driven while now being an emerging topic in research [Lo21]. Yet, no research focused on providing a status-quo review of previous research on data privacy and security in the light of CDR, scoping the highly dispersed research field on corporate responsibilities in the digital context. Rather prior endeavors concentrated on summarizing research directly related to data privacy and security [e.g., SDX11]. To the best of knowledge, this study is the first to provide a scoping overview on existing research in IS that relates to the concept of CDR. This study evaluates the popularity and comprehensiveness of current research in the sense of the interlinked consideration of data privacy and security with various components of corporate responsibilities. Aiming to move data privacy and security forward in a broader context, the awareness where relatable research already resides is crucial. Hence, this study scopes previous, highly dispersed research efforts and links them to the new concept of CDR. Besides, this research provides guidance on future research paths in IS regarding data privacy and security in the light of CDR. Both contribute to the overarching aim of enabling a more comprehensive view on data privacy and security by establishing CDR in IS research. To achieve this, the next section introduces the concept of CDR and domain-specific sub-fields. Section three presents the methodology, while section four maps current research according to CDR sub-fields and provides a discussion on this study's results. Finally, this work concludes with the main insights and its limitations.

2 Corporate Digital Responsibility

CDR shares goals and closely relates to the concept of Corporate Social Responsibility (CSR). The concept of CDR especially accounts for the “exponential growth in technological development, malleability of technologies and data in use, and pervasiveness of technology and data” [Lo21, p.876], therefore requiring the consideration as a separate concept, since these particularities are not explicitly covered by CSR [Mih22]. CDR gains increasing traction in research and a consensus of the concept understanding emerges [Mih22]. Many different approaches to CDR developed with differing nomenclature and foci, sharing a common understanding [Mih22, Mu22]. This study employs an approach consisting of eight sub-fields [Th17] developed in the current CDR debate (see Table 1). Some of the sub-fields are subject to regulations (e.g., the GDPR). However, CDR describes the voluntary assumption of responsibilities beyond the legal minimum.

CDR sub-field	Sub-field description
Data privacy and security	The protection of consumers' data privacy and security should be ensured (e.g., restricted data usage).

Education and awareness	Consumers should be educated, including raised awareness of ecological, social, societal, and economic consequences of their consumption (e.g., resource consumption in usage).
Information and transparency	Consumers should have access to appropriate information adapted to their individual needs (e.g., product information).
Economic interests	The economic interests of consumers should be protected and promoted (e.g., through the deployed business model).
Product safety and liability	Consumers should be protected from risks to their health and safety (e.g., their mental safety especially in social media).
Access	Consumers should have access to basic digital products and services: mentally and physically (e.g., access to software).
Dispute resolution and awareness	Consumers should have access to effective dispute settlement and redress procedures (e.g., complaint handling).
Governance and participation mechanisms	Companies should provide appropriate participation mechanisms for interested parties (e.g., in the product development).

Tab. 1: Overview of CDR sub-fields [following Th17, adapted from Mi21]

The individual CDR sub-fields are not new to IS research. CDR puts the manifold corporate responsibilities in a broader, interconnected context to provide a more holistic approach to corporate responsibilities. Especially in practice, corporate responsibilities are interwoven and do not appear isolated but activities related to one sub-field can have a cross-effect. The concept of CDR ensures the integrated evaluation of corporate responsibilities (e.g., data privacy and security) across different subfields, thereby better reflecting reality. The concept applies to various stakeholder groups (e.g., employees, consumers) [e.g., Lo21]. To ensure a solid research approach, this study focuses on CDR commitment directed at consumers as one key stakeholder group.

3 Allgemeine Formatierung

Scoping reviews, also called mapping studies, approach current research on a distinct topic. They share the same basic methodology with systematic literature reviews [KBP11]. Albeit, they differ in their designated goal. Systematic literature reviews aim at aggregating research results in detail, whereas scoping reviews should provide an overview and classification of previous research activities to serve as an initial orientation for other researchers [e.g., KBP11], e.g., concerning popularity. Besides, scoping reviews aim at identifying areas with a need for additional research [e.g., KBP11], thus guiding towards future research paths. Scoping reviews are increasingly used in IS research to give an overview of scattered research and to motivate future research on emerging topics [e.g., Li18].

Conducting the scoping review, this study follows the aggregated guidelines by Petersen

and colleagues [PVK15] to provide a systematic and transparent process also ensuring rigor, reliability, and trustable results. During the whole conceptualization, actual search, and evaluation phase a detailed review protocol keeps record of the process and corresponding decisions. The keyword selection grounds on the aim of CDR: ethical and responsible behavior of companies in the digital context. To determine appropriate keywords, we drew on prior knowledge, yearlong experience with the topic, and an initial unstructured analysis. The concept-related keywords form one set of keywords and aim to cover ethical behavior (i.e., (("corporate" OR "business") AND "responsibility") OR "ethic*"). In this vein, we opted against searching for the respective eight sub-fields derived from the CDR debate not to limit the search to these eight fields, rather searching more openly. Besides, this study employs an object-related keyword set to narrow the results to the relevant digital context (i.e., "product" OR "good" OR "service" OR "tech*" OR "econ*" OR "platform*" OR "market" OR "commerce"). This study uses wildcards and Boolean operators. The search term consists of the domain keyword sets linked with "AND" searching title, classification codes, abstract, keyword, and subjects.

This scoping review focuses on prior publications in IS research to provide guidance on data privacy and security research in the light of CDR within this specific research domain. This study incorporates only the eight top-ranked IS journals, the Senior Scholars' Basket Journals, focusing on IS outlets and ensuring a high reputation for quality. The search process covers every journal independently using the same search criteria. The exclusion of publications in other disciplines like computer science is due to the goal of the publication to stimulate and guide research in IS. The search should provide a picture as up-to-date as possible, thus conducted in autumn 2022 incorporating publications of the first half year. In total, the search yields 224 publications. Two researchers independently performed each step of the scoping review process, discussed their results after each step, and formed consensus on the final assessment. The search process includes the initial screening of the derived publications according to their title, abstract, and keywords. One criteria of exclusion applies to publications that are not adaptable to or focus on corporate responsibilities in the digital context. Still, this study includes research not designated to CDR due to the concept's novelty. Additionally, the final set of publications only consists of research with a business-to-consumer focus compared to a business-to-business one. Summing up, the initial evaluation leads to the exclusion of 107 publications due to the misfit with the aforementioned exclusion criteria. Thus, the second step of the review process features 117 publications for a detailed analysis of the articles' full texts according to the same exclusion criteria, leading to the exclusion of another 71 publications. Hence, the third step of the scoping review includes 46 articles and aims at coding them according to acknowledged practices. Yet, there is no suiting coding scheme available for CDR sub-fields. For the development of a suiting coding scheme, this study employs the nomenclature of the eight sub-fields of CDR [Th17]. The scheme is a multiple classification scheme (i.e., coders can classify publications into more than one sub-field). This classification approach serves as an orientation framework and therefore should lead to a discursive and iterative process. Both coders performed coding independently first, discussed their results and deviations in coding, and solved them in mutual agreement

following established procedures in IS research [e.g., Li18], thereby developing a codebook and applying it to the sample in a second coding round based on the same nomenclature but with a broader understanding of each of the eight sub-fields.

4 Data Privacy and security in the context of CDR

The aim of this study is to scope prior research relating to the concept of CDR and therefore to enable a more comprehensive view on data privacy and security in IS research. The results of the conducted scoping review indicate a growing interest in CDR-relevant research (see Table 2). While in the early years only a few publications addressed responsibilities in the digital context, research interest since 2007 has been at a comparatively higher level (except for 2013 and 2014). Particularly in 2021, research interest in responsibilities in the digital context has been extremely high. The results highlight the timeliness of research in this area, thereby motivating continued research concerning a comprehensive approach to corporate responsibilities (e.g., data privacy and security) in future.

		Data privacy & security	Education & awareness	Information & transparency	Economic interests	Product safety & liability	Access	Dispute resolution & awareness	Governance & participation mechanisms
1986	[Ma86]	x					x		
1992	[Oz92]	x	x		x				
1993	[Cu93]	x							
1994	[BO94]		x						x
	[MK94]		x				x		
	[WT94]								x
1997	[CS97]			x					
2000	[AW00]								x
2001	[VB01]						x		
2002	[SS02]	x							
	[SSN02]			x				x	
2003	[HM03]						x		
2007	[CD07]							x	
	[CI07]			x					

	[OD07]		x				x			x
	[PW07]				x					
	[St07]	x					x			
2008	[SK08]	x		x			x	x		
	[THS08]	x		x				x		
	[TYC08]							x		
2009	[LSL09]	x		x		x	x			
2010	[MW10]	x		x			x			
	[WBC10]		x	x	x					
2011	[SDX11]	x								
	[XB11]	x	x	x	x					
2012	[SDS12]	x		x						
	[St12]	x								
	[Te12]	x		x	x					
2015	[GCC15]	x	x	x						
	[Hu15]									x
2016	[Ra16]	x			x	x	x			
	[XXL16]			x		x				
2017	[ABK17]	x		x	x					
2018	[SRK18]					x				
2019	[Ch19]	x		x						
2020	[WSM20]	x		x	x					
2021	[Be21]	x		x	x	x				
	[DJR21]					x		x		
	[DT21]	x					x			
	[MNH21]			x						
	[MVL21]	x			x		x			
	[NPN21]	x		x	x	x				x
	[SM21]	x			x	x				
	[Wo21]			x		x				
2022	[KG22]			x	x					

[Mik22]	x		x	x	x			
---------	---	--	---	---	---	--	--	--

Tab. 2: Classification of the scoped publications according to CDR sub-fields covered

When assessing research interest on sub-field-level, our findings indicate that (isolated) research on data privacy and security is well established. This sub-field is most dominantly researched in this study (see Table 2). Research interest was stable throughout the whole study period. In the recent years, however, research on the topic intensified.

Despite the high importance of data privacy and security for companies and consumers alike [e.g., Ca23], future research should go beyond an isolated approach to this sub-field. Table 2 indicates that research on data privacy and security in the light of the overarching concept CDR is not yet comprehensive. While comparably many publications address data privacy and security simultaneously with sub-fields like information and transparency or access [e.g., Ch19, DT21], only few publications address data privacy and security and dispute resolution and awareness or education and awareness [e.g., GCC15, THS08] in parallel. These results indicate a lacking interconnected understanding of data privacy and security with further sub-fields, particularly since dispute resolution and awareness and education and awareness have a strong applicability in the context of data privacy and security. Hence, future research should develop an enhanced understanding of these CDR sub-fields and their influences on one another.

Above all, the total coverage in Table 2 indicates that only relatively few publications adopt a broader view of corporate responsibilities across different issues [e.g., NPN21, Ra16] and none the whole concept spectrum, thereby not discussing data privacy and security in the broadest possible context. IS research on CDR-related topics is more isolated and reveals a more dispersed research field, implying the danger of disconnected research. Albeit recent IS publications follow a broader approach to digital responsibility with more sub-fields covered (on average) in parallel per publication. This emphasizes the broadening approach to digital responsibility, underpinning the usefulness of a concept such as CDR, which places individual digital responsibility areas such as data privacy and security in the broader context and thus enables a more comprehensive approach to them. Future research should adopt this broader perspective. In this way, a realistic understanding of the interwoven and influencing responsibilities can be developed. This helps researchers to understand these interrelationships better, but also supports practitioners in the implementation. In this way, the understanding of corporate responsibilities concerning data privacy and security can evolve, particularly becoming more comprehensive.

5 Conclusion

This research provides an initial systematic overview on previous highly dispersed research efforts in leading IS outlets. Aim of this study is to assess the popularity of data privacy and security in the light of the umbrella concept CDR in IS research. This study's

results corroborate that research on corporate responsibilities in the digital context, particularly data privacy and security, is well established in IS research. Rather, research remains incoherent and disconnected. Ultimately, this study sheds light on future research paths worth considering in IS research. None of the evaluated publications addresses the full range of CDR sub-fields simultaneously. However, corporate responsibilities never occur in isolation in practice. Hence, a holistic approach to data privacy and security, e.g., employing CDR, can help to understand the extent of corporate responsibilities and the influences of the sub-fields on another. Summing up, this study theoretically contributes to the existing knowledge on data privacy and security as well as CDR in IS research. The scoping of the examined publications contributes to a better understanding of the status-quo of research relating to CDR and its distinctive sub-fields focusing on data privacy and security research. Additionally, this study's results motivate and inform IS researchers concerning future research paths in the context of data privacy and security employing a comprehensive approach such as CDR.

Despite best efforts, some limitations apply. Firstly, this study is not and was not meant to be exhaustive. There is a deliberate focus on IS research. Secondly, this study focuses on CDR activities geared to consumers. Still, CDR provides a broad field of action applicable to, e.g., employees, business-to-business activities, or society itself. Future research should incorporate these differing stakeholder groups and their respective viewpoints.

To this end, goal of this publication is to motivate future work on data privacy and security in the context of CDR. This will shed light on yet underexplored research areas but also stimulate the discovery of novel fields worth considering.

Acknowledgement

The Hessian State Chancellery – Hessian Minister of Digital Strategy and Development supported this work under the promotional reference 6/493/71574093 (CDR-CAT).

Bibliography

- [ABK17] Anderson, C.; Baskerville, R. L.; Kaul, M.: Information Security Control Theory: Achieving a Sustainable Reconciliation Between Sharing and Protecting the Privacy of Information. *Journal of Management Information Systems* 34/4, pp. 1082-1112, 2017.
- [AW00] Adman, P.; Warren, L.: Participatory Sociotechnical Design of Organizations and Information Systems – An Adaptation of Ethics Methodology. *Journal of Information Technology* 15/1, pp. 39-51, 2000.
- [Be21] Berente, N. et al.: Managing Artificial Intelligence. *MIS Quarterly* 45/3, pp. 1433-1450, 2021.

- [BO94] Beath, C. M.; Orlikowski, W. J.: The Contradictory Structure of Systems Development Methodologies: Deconstructing the IS-User Relationship in Information Engineering. *Information Systems Research* 5/4, pp. 350-377, 1994.
- [Ca23] Carl, K. V. et al.: Consumer Perspective on Corporate Digital Responsibility—An Empirical Evaluation of Consumer Preferences. *Journal of Business Economics*, Forthcoming, 2023.
- [CD07] Costello, G. J.; Donnellan, B.: The Diffusion of WOZ: Expanding the Topology of IS Innovations. *Journal of Information Technology* 22/1, pp. 79-86, 2007.
- [Ch19] Chanson, M. et al.: Blockchain for the IoT: Privacy-Preserving Protection of Sensor Data. *Journal of the Association for Information Systems* 20/9, pp. 1274-1309, 2019.
- [CI07] Clemons, E. K.: An Empirical Investigation of Third-Party Seller Rating Systems in E-Commerce: The Case of buySAFE. *Journal of Management Information Systems* 24/2, pp. 43-71, 2007.
- [CS97] Choudhury, V.; Sampler, J. L.: Information Specificity and Environmental Scanning: An Economic Perspective. *MIS Quarterly* 21/1, pp. 25-53, 1997.
- [Cu93] Culnan, M. J.: “How Did They Get My Name?”: An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use. *MIS Quarterly* 17/3, pp. 341-363, 1993.
- [DJR21] Dunn, B.; Jensen, M. L.; Ralston, R.: Attribution of Responsibility After Failures Within Platform Ecosystems. *Journal of Management Information Systems* 38/2, pp. 546-570, 2021.
- [DT21] Díaz Andrade, A.; Techatassanasoontorn, A. A.: Digital Enforcement: Rethinking the Pursuit of a Digitally-Enabled Society. *Information Systems Journal* 31/1, pp. 184-197, 2021.
- [GCC15] Greenaway, K. E.; Chan, Y. E.; Crossler, R. E.: Company Information Privacy Orientation: A Conceptual Framework. *Information Systems Journal* 25/6, pp. 579-606, 2015.
- [HM03] Heng, M. S. H.; de Moor, A.: From Habermas’s Communicative Theory to Practice on the Internet. *Information Systems Journal* 13/4, pp. 331-352, 2003.
- [Hu15] Hutter, K. et al.: Machiavellianism or Morality: Which Behavior Pays Off in Online Innovation Contests? *Journal of Management Information Systems* 32/3, pp. 197-228, 2015.

- [KBP11] Kitchenham, B. A.; Budgen, D.; Pearl Brereton, O.: Using Mapping Studies as the Basis for Further Research – A Participant-Observer Case Study. *Information and Software Technology* 53/6, pp. 638-651, 2011.
- [KG22] Kordzadeh, N.; Ghasemaghahi, M.: Algorithmic Bias: Review, Synthesis, and Future Research Directions. *European Journal of Information Systems* 31/3, pp. 388-409, 2022.
- [Li18] Li, Y. et al.: Blockchain Technology in Business Organizations: A Scoping Review. In: *Proceedings of the 51st Hawaii International Conference on System Sciences (HICCS)*, pp. 4474-4483, 2018.
- [Lo21] Lobschat, L. et al.: Corporate Digital Responsibility. *Journal of Business Research* 122, pp. 875-888, 2021.
- [LSL09] Lee, S.; Shin, B.; Lee, H. G.: Understanding Post-Adoption Usage of Mobile Data Services: The Role of Supplier-Side Variables. *Journal of the Association for Information Systems* 10/12, pp. 860-888, 2009.
- [Ma86] Mason, R. O.: Four Ethical Issues of the Information Age. *MIS Quarterly* 10/1, pp. 5-12, 1986.
- [Mi21] Mihale-Wilson, C. et al.: Corporate Digital Responsibility – Extended Conceptualization and a Guide to Implementation. In: *Proceedings of the European Conference on Information Systems (ECIS)*, 2021.
- [Mih22] Mihale-Wilson, C. et al.: Corporate Digital Responsibility: Relevance and Opportunities for Business and Information Systems Engineering. *Business & Information Systems Engineering* 64/2, pp. 127-132, 2022.
- [Mik22] Mikalef, P. et al.: Thinking Responsibly About Responsible AI and “the Dark Side” of AI. *European Journal of Information Systems* 31/3, pp. 257-268, 2022.
- [MK94] Mirani, R.; King, W. R.: Impacts of End-User and Information Center Characteristics on End-User Computing Support. *Journal of Management Information Systems* 11/1, pp. 141-166, 1994.
- [MNH21] Marabelli, M.; Newell, S.; Handunge, V.: The Lifecycle of Algorithmic Decision-Making Systems: Organizational Choices and Ethical Challenges. *Journal of Strategic Information Systems* 30/3, pp. 1-15, 2021.
- [Mu22] Mueller, B.: Corporate Digital Responsibility. *Business & Information Systems Engineering* 64, pp. 689-700, 2022.
- [MVL21] Marabelli, M.; Vaast, E.; Li, J. L.: Preventing the Digital Scars of COVID-19. *European*

Journal of Information Systems 30/2, pp. 176-192, 2021.

- [MW10] Mingers, J.; Walsham, G.: Toward Ethical Information Systems: The Contribution of Discourse Ethics. *MIS Quarterly* 34/4, pp. 833-854, 2010.
- [NPN21] Nussbaumer, A.; Pope, A.; Neville, K.: A Framework for Applying Ethics-by-Design to Decision Support Systems for Emergency Management. *Information Systems Journal* 33/1, pp. 34-55, 2021.
- [OD07] Olphert, W.; Damodaran, L.: Citizen Participation and Engagement in the Design of E-Government Services: The Missing Link in Effective ICT Design and Delivery. *Journal of the Association for Information Systems* 8/9, pp. 491-507, 2007.
- [Oz92] Oz, E.: Ethical Standards for Information Systems Professionals: A Case for a Unified Code. *MIS Quarterly* 16/4, pp. 423-433, 1992.
- [PVK15] Petersen, K.; Vakkalanka, S.; Kuzniarz, L.: Guidelines for Conducting Systematic Mapping Studies in Software Engineering: An Update. *Information and Software Technology* 64, pp. 1-18, 2015.
- [PW07] Parameswaran, M.; Whinston, A.: Research Issues in Social Computing. *Journal of the Association for Information Systems* 8/6, pp. 336-350, 2007.
- [Ra16] Ransbotham, S. et al.: Special Section Introduction—Ubiquitous IT and Digital Vulnerabilities. *Information Systems Research* 27/4, pp. 834-847, 2016.
- [Ro18] Royakkers, L. et al.: Societal and Ethical Issues of Digitization. *Ethics and Information Technology* 20/2, pp. 127-142, 2018.
- [SDS12] Stahl, B. C.; Doherty, N. F.; Shaw, M.: Information Security Policies in the UK Healthcare Sector: A Critical Evaluation. *Information Systems Journal* 22/1, pp. 77-94, 2012.
- [SDX11] Smith, H. J.; Dinev, T.; Xu, H.: Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly* 35/4, pp. 989-1015, 2011.
- [SK08] Son, J.-Y.; Kim, S. S.: Internet Users' Information Privacy-Protective Responses: A Taxonomy and a Nomological Model. *MIS Quarterly* 32/3, pp. 503-529, 2008.
- [SM21] Stahl, B. C.; Markus, M. L.: Let's Claim the Authority to Speak Out on the Ethics of Smart Information Systems. *MIS Quarterly* 45/1, pp. 485-488, 2021.
- [Sp22] Spiekermann, S. et al.: Values and Ethics in Information Systems. *Business & Information Systems Engineering* 64, pp. 247-264, 2022.

- [SRK18] Seymour, M.; Riemer, K.; Kay, J.: Actors, Avatars and Agents: Potentials and Implications of Natural Face Technology for the Creation of Realistic Visual Presence. *Journal of the Association for Information Systems* 19/10, pp. 953-981, 2018.
- [SS02] Stewart, K. A.; Segars, A. H.: An Empirical Examination of the Concern for Information Privacy Instrument. *Information Systems Research* 13/1, pp. 36-49, 2002.
- [SSN02] Shim, J. P.; Shin, Y. B.; Nottingham, L.: Retailer Web Site Influence on Customer Shopping: Exploratory Study on Key Factors of Customer Satisfaction. *Journal of the Association for Information Systems* 3/1, pp. 53-76, 2002.
- [St07] Stahl, B. C.: ETHICS, Morality and Critique: An Essay on Enid Mumford's Socio-Technical Approach. *Journal of the Association for Information Systems* 8/9, pp. 479-490, 2007.
- [St12] Stahl, B. C.: Morality, Ethics, and Reflection: A Categorization of Normative IS Research. *Journal of the Association for Information Systems* 13/8, pp. 636-656, 2012.
- [Te12] Temizkan, O. et al.: Patch Release Behaviors of Software Vendors in Response to Vulnerabilities: An Empirical Analysis. *Journal of Management Information Systems* 28/4, pp. 305-338, 2012.
- [Th17] Thorun, C. et al.: Indicators of Consumer Protection and Empowerment in the Digital World. Results and Recommendations of a Feasibility Study. https://www.conpolicy.de/data/user_upload/Studien/ConPolicy_Indicator_study.pdf, 2017.
- [THS08] Tang, Z.; Hu, Y.; Smith, M. D.: Gaining Trust through Online Privacy Protection: Self-Regulation, Mandatory Standards, or Caveat Emptor. *Journal of Management Information Systems* 24/4, pp. 153-173, 2008.
- [TYC08] Turel, O.; Yuan, Y.; Connelly, C. E.: In Justice We Trust: Predicting User Acceptance of E-Customer Services. *Journal of Management Information Systems* 24/4, pp. 123-151, 2008.
- [VB01] Venkatesh, V.; Brown, S. A.: A Longitudinal Investigation of Personal Computers in Homes: Adoption Determinants and Emerging Challenges. *MIS Quarterly* 25/1, p. 71-102, 2001.
- [WBC10] Watson, R. T.; Boudreau, M.-C.; Chen, A. J.: Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. *MIS Quarterly* 34/1, pp. 23-38, 2010.
- [Wo21] Wong, R. Y. M. et al.: Standing Up or Standing By: Understanding Bystanders' Proactive Reporting Responses to Social Media Harassment. *Information Systems*

Research 32/2, pp. 561-581, 2021.

- [WSM20] Wiener, M.; Saunders, C.; Marabelli, M.: Big-Data Business Models: A Critical Literature Review and Multiperspective Research Framework. *Journal of Information Technology* 35/1, pp. 66-91, 2020.
- [WT94] Wong, E. Y. W.; Tate, G.: A Study of User Participation in Information Systems Development. *Journal of Information Technology* 9/1, pp. 51-60, 1994.
- [XB11] Xiao, B.; Benbasat, I.: Product-Related Deception in E-Commerce: A Theoretical Perspective. *MIS Quarterly* 35/1, pp. 169-196, 2011.
- [XXL16] Xu, B.; Xu, Z.; Li, D.: Internet Aggression in Online Communities: A Contemporary Deterrence Perspective. *Information Systems Journal* 26/6, pp. 641-667, 2016.

Deceptive patterns in consent dialogs on children’s websites

Suvi Lehtosalo¹ and Daniel W. Woods²

Abstract: Many privacy and data protection laws, such as Article 8 GDPR and the CCPA, establish different requirements when establishing a legal basis for collecting personal data about children. Our study asks whether and how children’s websites collect consent. We conduct an automated analysis of 2,066 educational and gaming websites, and a manual analysis of 13 large sites. We measure the prevalence of deceptive patterns identified in prior work, plus a new design consideration, whether the dialog is addressed to the child user’s parent or guardian. A small minority of websites address dialogs for children, which suggest the majority of children’s websites in our sample may not comply with Article 8 GDPR.

Keywords: Online consent, Data protection, Internet measurement, Child protection

1 Introduction

Under the paradigm of privacy self-management [So12], users are responsible for defining what websites can do with their personal data. Empirical studies demonstrate the challenges in doing so in a meaningful way. One barrier is the decision volume resulting from users browser multiple websites with diverse and complex privacy policies [MC08, BP10, So12]. Even when the user has made a decision, dialogs can be designed to make opting-out difficult [MBS20, No20]. Dialogs frequently contain *deceptive patterns*, which are design choices intended to influence users’ decisions; for example, the opt-in option may be highlighted, and the opt-out option hidden behind a ‘More Options’ button.

Discussions around online consent implicitly assume users are adults with functioning cognitive abilities. Yet, child users have a limited understanding of data collection and privacy risks [Su21]. This view is echoed by debates related to whether child patients can provide medical consent [Sh73].

Such considerations motivated policymakers to create different online consent requirements for children. The General Data Protection Regulation (GDPR) in the EU requires that consent is “given or authorised by the holder of parental responsibility over the child”, whereas consent is given by the data subject if they are an adult [MK17]. The difference is even starker under the California Consumer Privacy Act (CCPA), which creates an opt-out

¹ University of Edinburgh, School of Informatics, 10 Crichton St, Edinburgh, EH8 9AB, UK

² University of Edinburgh, School of Informatics, 10 Crichton St, Edinburgh, EH8 9AB, UK. Contact email: daniel.woods@ed.ac.uk

consent regime for adults but an opt-in regime for children (Under 16s), which must be obtained from a parent or guardian if the child is under 13 [AE22].

This raises the question of how children’s websites collect opt-in consent. Our paper answers:

RQ1 How does the prevalence of dark patterns in consent dialogs vary across adult and children’s websites?

RQ2 How does the prevalence of dark patterns in consent dialogs vary across the adult and children’s sections of a single website?

RQ3 Who are consent dialogs on children’s websites addressed to?

To answer *RQ1*, we sample adult and children’s websites from search engines. Dark patterns in this sample are detected by applying a system developed in prior work [KVV23]. To answer *RQ2* and *RQ3*, we craft a sample of popular websites that have a website section specifically for children. We then run both the automated and manual analysis on these websites, comparing the consent dialogs on the adult and children’s sections.

Our automated analysis shows that adult and children’s websites are similarly likely to display a consent dialog. Further, there is no major difference in the prevalence of dark patterns. Just 7% of dialogs on children’s games websites asked for consent from the user’s parent or guardian. Our manual analysis of 13 major websites revealed that 5 websites collect consent and personal data in a different way on the children’s section of the website, such as displaying a different consent dialog or not setting ID-like cookies on the children’s section.

We introduce the necessary background to this study in Section 2. This motivates and guides our methodology, which is described in Section 3. We present our results in Section 4. We reflect on the study in Section 5, and offer a conclusion in Section 6.

2 Background

Legal Online consent has been voluntarily collected since the 1990s when browsers requested user permission before installing cookies [MFF01]. This became a legal requirement in the EU in 2009, as the reformed ePrivacy Directive required websites to obtain consent before “storing or accessing cookies on a device” [Bo13]. The 2016 General Data Protection Regulation further expanded the role of online consent by allowing it to serve as a legal basis for data processing (beyond simply installing cookies). The role of online consent in US privacy law is harder to summarize because the law is fractured across states, agencies and standalone federal laws. To avoid these discrepancies, we primarily focus on studies in the GDPR context.

The GDPR placed particular restrictions on the processing of children’s data, recognising that children may be less aware of risks and as such require special protection. Most notably,

children below the age of 16 cannot consent to data processing, though this age threshold can be replaced by a national threshold for children between 13 and 16 [MK17] if the member state decides to do so.

Human Computer Interface Design choices, such as hiding the opt-out behind a ‘More Options’ button, have been repeatedly shown to influence the users’ consent decisions [KPW21]. These designs are known as *deceptive patterns* because they undermine the user’s autonomy, and may be non-compliant with European law [SBM20]; as such, it is relevant for us to look at dark patterns as a quantifiable measure of the level of compliance when comparing consent dialogs addressed to children and adults.

Scraping Studies Studies of consent interfaces found on real-world websites were surveyed by Kretschmer et al. [KPW21]. Most studies sample websites from a list of the most popular websites (e.g. Tranco [Po21]) [KPW21]. We are aware of three studies that extract arbitrary dialogs [Ei19, KS21, KVW23]. We use the *DarkDialogs* system, which correctly extracted 99% of dialogs in a hand-labelled sample [KVW23].

Children and Online Tracking Prior work has found that tracking is pervasive on children’s websites. A sample of 20 children’s websites all contained invisible tracking images [V118]. Similarly, children’s websites were observed to have the second-highest percentage of websites that perform at least some kind of tracking [Sa19]. Tracking is also prevalent in children’s apps [Zh20, Bi18]. Our contribution is the first study focused on consent dialogs on children’s websites.

3 Methodology

3.1 Sample of Websites

Automated We need a sample of websites that differ only in whether the users were children or adults. We cannot simply sample from children’s websites and compare that to a sample of adult websites. The adult sample would include websites with no children’s equivalent, such as financial services or tax advice. This would complicate statistical analysis because consent interfaces vary considerably across website category (see Fig. 4 [HWB21]).

We instead sample from categories for which both adult and children’s websites exist, namely games and education websites. We collected these using a search engine, following previous studies of children’s websites [Ca03, CZ13, NKA21]. We automatically extracted the results from the search terms: “online games” (GameGen), “online games for kids” (GameKid), “online games for adults” (GameAdult), and equivalently “educational website” (EdGen), “educational website for kids” (EdKid) and “educational website for adults” (EdAdult). This resulted in a sample of 2,066 websites after removing any Google Books results and sites that occurred multiple times under the same search term. Figure 1 shows the number of websites that occur under multiple search terms, which we chose not to remove from the analysis.

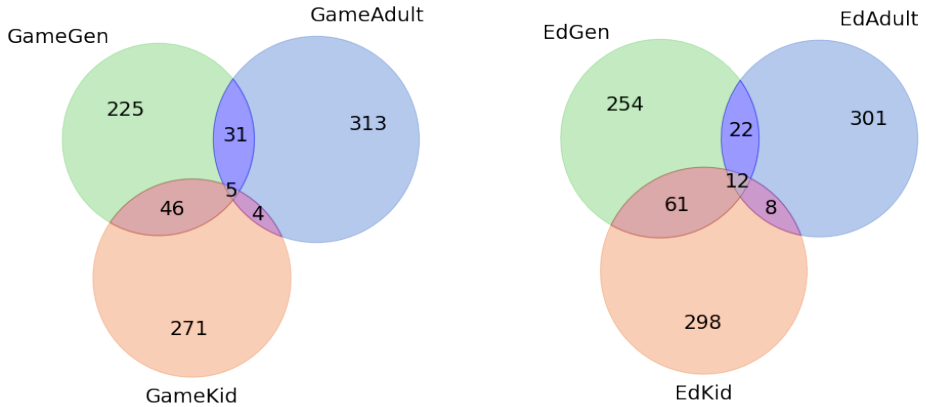


Fig. 1: A minority of websites were sampled under multiple categories.

Manual We collected a sample of 13 websites that had a specific children’s section, in addition to the website’s generic homepage (see Table 1 for the full sample). The pairs of websites were found based on the author’s knowledge, as well as searching “children’s websites” and “websites for kids”. The websites were accessed on (March 24th 2023) through Google Chrome with the most permissive privacy settings. All previous cookies were deleted before accessing each website.

3.2 Analysis

Automated We used an open-source tool, DarkDialogs, to automatically extract dialogs and detect dark patterns [KVV23]. DarkDialogs detects 10 different dark patterns (*DP1–10*) [KVV23]. We also introduce a new dark pattern that is specific to children’s websites:

DP11: ConsentFromKids A cookie dialog on a website aimed for children or for a general audience asks for consent from the user, without specifying that children should have their guardians read and respond to the dialog.

This was implemented by searching for and then manually checking extracted dialogs that contain the words parent/guardian/adult.

We conducted additional evaluation to ensure the system’s performance had not degraded over time on websites sampled from Tranco and Google. In terms of dialog extraction, the system correctly extracted a dialog in 93.5% of cases (compared to 98.7% in the original paper [KVV23]). In terms of dark pattern detection, we removed the *ObstructsWindow* dark pattern because of the high false positive rate (40%). After removing this, the system

	Websites		Dialogs		Children’s dialogs	
	Total	Loaded	n	% of sites	n	% of dialogs
EdAdult	344	343	215	62.7	1	0.5
EdGen	351	349	232	66.5	0	0
EdKid	381	379	200	52.8	1	0.5
GameAdult	354	353	174	49.3	0	0
GameGen	309	307	199	64.8	8	4.0
GameKid	327	326	180	55.2	12	6.7
Total	2,066	2,057	1,200	58.3	22	1.8

Tab. 1: The majority of education and gaming websites contain consent dialogs. Dialogs that reference the user’s parent or guardian are rare, especially among children’s educational websites.

had a False Positive rate of 2.45% and a False Negative rate of 7.35%. This highlights how quickly the performance of web measurement systems deteriorate over time.

Manual Each cookie dialog was manually checked for each of the 10 Dark Patterns. We checked whether dialogs on children’s websites appeared more child-friendly (using simplified language, a larger font, or bright colours), and whether they asked for consent from a parent. Any other significant difference was also noted. Cookies set after clicking each option were collected via Chrome’s Developer tools.

4 Results

4.1 Automated Analysis

We ran the DarkDialogs on 2,066 websites in March 2023. Table 1 shows that 2,057 of these websites successfully loaded. The majority (58%) of websites in our sample had a consent dialog, with the highest prevalence among general gaming and education websites. We use the term *children’s dialog*, which corresponds to *DP11*, to denote dialogs specifying that children should ask their guardians to respond to the dialog. Table 1 shows that less than 2% of dialogs are children’s dialogs. They are most prevalent on children’s gaming websites, although this is still just 7% of websites. Only two educational websites contained a children’s dialog.

The tool, DarkDialogs [KVW23], detects additional dark patterns, which are displayed in Figure 2b. We could not identify any generic statements like ‘dialogs on children’s websites contain fewer dark patterns’ because this varied by the type of website. For example, children’s gaming websites had a lower prevalence of *ComplexText* than general and adult gaming websites. However, this trend is reversed for education websites.

Dialogs on adult gaming websites display a higher prevalence of each dark pattern than dialogs on children’s gaming websites, with the exception of *MoreOptions*. The comparison

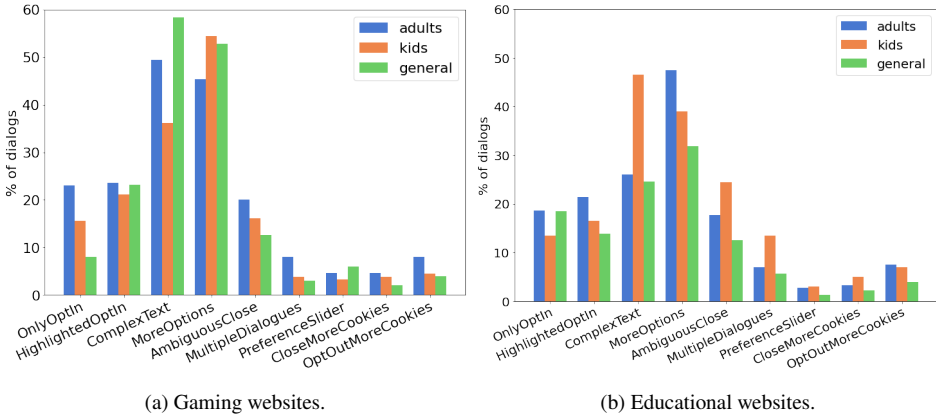


Fig. 2: Occurrence of different DPs in the found cookie dialogs.

	Mean	Median	Std deviation	Max	Min
GameGen	14.22	3.0	31.74	179	0
GameAdult	31.35	10.5	48.56	205	0
GameKid	12.26	4.0	23.12	138	0
EdGen	7.49	2.0	18.06	155	0
EdAdult	7.57	4.0	16.94	189	0
EdKid	10.07	5.0	17.42	110	0

Tab. 2: Third-party cookies set after opting in.

to general gaming websites is more variable. We could not identify any comparable trend among education websites.

Differences can be seen in the number of cookies set by websites in each sample. Table 2 shows the statistics of third-party cookies set after the opt-in option was selected. Children’s gaming sites set less than half as many third-party cookies as adult’s gaming websites. This is reversed for educational websites, in which children’s educational websites set more cookies than both general and adult educational websites.

4.2 Manual Analysis

We identified 13 websites with separate children’s subsites. Three had no cookie dialog on either website, and a further five had exactly the same cookie dialog on both sites. Notably, Table 3 shows the Ted and Time websites set fewer ID-like cookies (defined as cookies whose value is unique enough to be used as an identifier) on the children’s website, possibly because they do not track these users. We were left with five websites that display different cookie dialogs in each section:

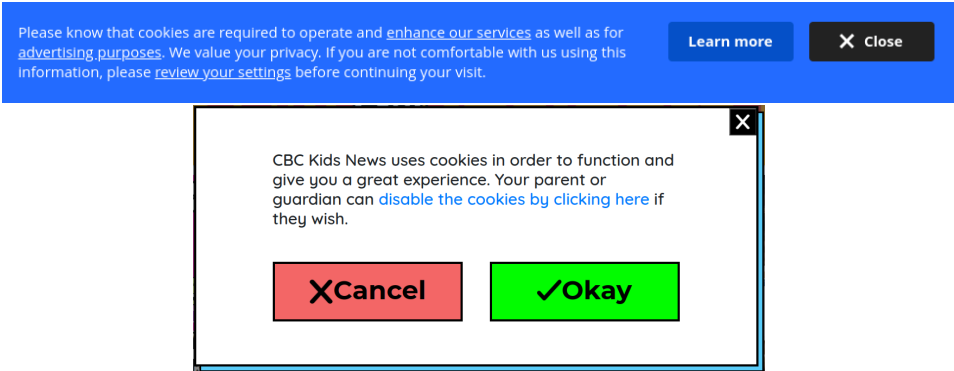


Fig. 3: Contrasting the cookie dialogs on *cbc.ca* and *cbc.ca/kidsnews*, we see that the former is specifically designed for children.

The website *Nationalgeographic.co.uk* has a one-click opt-out, whereas the dialog on the children's page *natgeokids.com/uk* only has a *MoreOptions* button with no 1-click opt-out. The children's page also asks whether the user is a teacher, kid or parent. It then displays the same dialog regardless of which option is chosen, whereas it should prompt the child to ask a parent or guardian to provide consent if 'Kid' is selected. However, Table 3 shows the children's website sets only one ID-like cookie, regardless of the consent decision. It could be that the website does not use this to track users browsing the children's site.

The website *Youtube.com* has a cookie dialog, whereas *Youtubekids.com* does not. Further, the children's site requires users to allow third-party cookies. When it is accessed with third-party cookies blocked, a warning is displayed and the user is not allowed to sign in.

Comparing the dialogs on a Canadian news website (*cbc.ca* to *cbc.ca/kidsnews*), the dialog on the children's website is simplified with a large font and bright colours. It also contains a reference to the child's guardians, but not in the expected way. Rather than request for the child's guardian to *consent* to cookies, they instead say the guardian can "disable the cookies ... if they wish" (shown in Figure 3). The other dialogs contains dark patterns like having no 1-click opt-out.

The children's website *dkfindout.com/uk* has two dialogs: the first one has opt-in and reject options, as well as an ambiguous close button and a "More options" button. After the first dialog has been closed, another one behind it becomes visible. This dialog states that by continuing to browse the site, the user is agreeing to its use of cookies; only a close option and a link to the cookie policy are provided. The recorded number of cookies set (see Table 3) further indicates that the first dialog actually has no impact on the number of cookies set. The general website *dk.com/uk* only has one cookie dialog, which is similar to the second dialog on the children's website, giving the user no option besides opting in.

Out of the two Guinness world records websites, only the children's section has a cookie

Website	Initial	Opt-in	Opt-out	Close
bbc.co.uk	2	2	2	
bbc.co.uk/cbeebies	3	3	2	
cbc.ca	8			11
cbc.ca/kidsnews	12	12	12	11
cia.gov and cia.gov/spykids	0			
nasa.gov and nasa.gov/kidsclub	0			
dk.com/uk	1	1		
dkfindout.com/uk	2	2	2	
guinnessworldrecords.com	7			
kids.guinnessworldrecords.com	0	0		
nationalgeographic.co.uk	0	10	1	
natgeokids.com/uk	1	1	1	
pbs.org and pbskids.org	2			
tate.org.uk and tate.org.uk/kids	0	11	0	
ted.com	9			9
ed.ted.com	4			5
time.com	7	35	18	
timeforkids.com	3	4	3	
vogue.com	6	37	7	
teenvogue.com	4	46	5	
youtube.com	0	2	1	
youtubekids.com	1			

Tab. 3: The children’s website set fewer ID-like cookies for Guinness World Records, National Geographic, Ted, and Time. There is no obvious difference for the other websites.

dialog, which only contains an opt-in button. This is likely because they do not track users browsing the children’s site. Table 3 shows that the children’s site sets no ID-like cookies.

5 Discussion

We discuss whether these websites are compliant, potential alternative solutions, and the limitations of our study.

Compliance Evaluating website compliance would be simpler if websites could be reliably divided into those websites used by children and those used by adults, and then further divided by jurisdiction. In such a world, children’s websites in the EU are non-compliant

with Article 8 of the GDPR if the dialog requests consent from a child user, and not the child's parent or guardian. In reality, most websites are used by both adult and children from a mixture of jurisdictions, which makes it unclear which standard to apply. This motivates two approaches to deciding whether to display an adult or children's dialog.

The most technically difficult approach is to infer the user's age and to present a children's dialog to children. This is analogous to how websites present cookie dialogs based on the user's jurisdiction, such as not showing GDPR-compliant dialogs to users from the US [Ei19, HWB20]. Another approach is to use the observed or intended demographics of users to choose a dialog that applies to all users. This leads to an arbitrary threshold for when a children's dialog should be offered—should a website for which 5% of the users are children? How about 50%?

These differences in potential approaches are why we lack confidence in declaring individual sites as non-compliant. Some websites targeted at adults or with majority adult user bases may have been erroneously included in our search term “online games for kids”. It is also possible (although unlikely given the technical challenge) that some websites accurately inferred that all researchers in this study were adults, and hence the website did not display a children's dialog. Nevertheless, it seems reasonable to conclude that the low prevalence of children's dialogs (0.5%/7% respectively of dialogs on children's educational/gaming websites) suggests multiple websites do not comply with Article 8 of the GDPR, given that we accessed these websites from a UK IP address,

Solutions We can either discuss solutions within the current legal paradigm or discuss solutions to the underlying problem. In terms of complying with the current legal regime, we recommend against trying to infer a user's age because doing so likely necessitates collecting additional data about users—this would erode the user's privacy before they are offered a chance to consent. Offering the same dialog to all users avoids this privacy erosion. Websites with a significant share of child users should adapt consent dialogs to address Article 8. Even in such cases, Article 8 also requires the controller to “make reasonable efforts to verify in such cases that consent is given or authorised by the holder of parental responsibility over the child”. It remains unclear what this involves. Given compliance with Article 8 is still ambiguous and that many US states are introducing privacy laws, it may be wiser to simply stop collecting data from users who are most likely children. This is presumably the choice Guinness World Records made given they set no ID like cookies on the children's section of their website.

In terms of the actual problem—that users lack of control over their privacy, and child users have even less awareness of privacy issues [Su21]—it is unclear that consent dialogs help users. The notice and consent approach to privacy law imposes a large decision burden on users [So12]. A parent or guardian interrupting household tasks to respond to a consent dialog on behalf of their child would waste even more time.

One alternative is for users to make a decision that applies to all websites, which would be

automatically communicated by a browser. The Global Privacy Control (GPC) is one such signal [ZA20, HWB21, Zi23], which represents an opt-out of data processing when turned on. Allowing parents and guardians to set the GPC on the device used by their children preserves both parental responsibility and time. However, websites have little incentive to adopt the GPC unless it is effectively enforced; few websites currently respect GPC signals [Zi23]. Further regulation and enforcement is therefore needed before the GPC can be relied on.

Limitations Measuring the Web is hard because the content served changes over time and across vantage points. For example, all of our measurements were made from the UK, which likely impacted which dialog was shown [Ei19]. This issue will become more important when measuring from the US given that privacy laws vary by state. We also discovered that the performance of the automated system that we used degraded since it was designed one year ago [KVV23]. Maintaining open-source web measurement systems remains a challenge for this reason.

It was also challenging to collect a sample of children's websites. Our manual analysis allowed us to find the children's section of large websites, which are undeniably targeted to children. This was more difficult when it came to finding children's websites, given website classifications often contain errors [Va20]. We used Google's search results as a pragmatic solution, which is vulnerable to biases resulting from how Google prioritises websites.

There is no existing research on how children perceive and are affected by dark patterns. This study did not include child participants, so the methods were validated solely from an adult perspective. Future work could include children in the process of assessing dialogs and designing more child-friendly alternatives.

6 Conclusion

Our measurement study asked whether and how children's websites collect consent for data processing from users. Our sample of educational and gaming websites revealed that children's websites were just as likely to collect consent from users. Around 7% of the dialogs on children's gaming websites were addressed to the user's parent or guardian. Children's websites displaying dialogs that are not designed for children may be violating Article 8 of the GDPR, although we are not aware of a legal judgement. Prior work showed non-compliance with other design patterns [No20, MBS20, KPW21].

Bibliography

- [AE22] Alomar, N.; Egelman, S.: Developers Say the Darnedest Things: Privacy Compliance Processes Followed by Developers of Child-Directed Apps. *Proceedings on Privacy Enhancing Technologies*, 4(2022):24, 2022.

- [Bi18] Binns, R.; Lyngs, U.; Van Kleek, M.; Zhao, J.; Libert, T.; Shadbolt, N.: Third Party Tracking in the Mobile Ecosystem. In: Proceedings of the 10th ACM Conference on Web Science. WebSci '18, Association for Computing Machinery, New York, NY, USA, p. 23–31, 2018.
- [Bo13] Borgesius, F. Z.: Behavioral targeting: A European legal perspective. *IEEE Security & Privacy*, 11(1):82–85, 2013.
- [BP10] Bonneau, J.; Preibusch, S.: The privacy jungle: On the market for data protection in social networks. In: *Economics of Info. Sec. and Privacy*, pp. 121–167. Springer, 2010.
- [Ca03] Cai, X.; Gantz, W.; Schwartz, N.; Wang, X.: Children's website adherence to the FTC's online privacy protection rule. *Journal of Applied Communication Research*, 31(4):346–362, 2003.
- [CZ13] Cai, X.; Zhao, X.: Online advertising on popular children's websites: Structural features and privacy issues. *Comp. in Human Behavior*, 29(4):1510–1518, 2013.
- [Ei19] van Eijk, R.; Asghari, H.; Winter, P.; Narayanan, A.: The impact of user location on cookie notices (inside and outside of the European union). In: *Workshop on Technology and Consumer Protection (ConPro'19)*. 2019.
- [HWB20] Hils, M.; Woods, D. W.; Böhme, R.: Measuring the emergence of consent management on the web. In: *Proc. of the ACM Internet Measur. Conf.* pp. 317–332, 2020.
- [HWB21] Hils, M.; Woods, D. W.; Böhme, R.: Privacy preference signals: Past, present and future. *Proc. on Privacy Enhancing Technologies*, 2021(4):249–269, 2021.
- [KPW21] Kretschmer, M.; Pennekamp, J.; Wehrle, K.: Cookie banners and privacy policies: Measuring the impact of the GDPR on the web. *ACM Trans. on the Web*, 15(4):1–42, 2021.
- [KS21] Kampanos, G.; Shahandashti, S. F.: Accept all: The landscape of cookie banners in Greece and the UK. In: *IFIP International Conference on ICT Systems Security and Privacy Protection*. Springer, pp. 213–227, 2021.
- [KVV23] Kirkman, D.; Vaniea, K.; Woods, D. W.: DarkDialogs: Automated detection of 10 dark patterns on cookie dialogs. In: *8th IEEE European Symposium on Security and Privacy*. IEEE, 2023.
- [MBS20] Matte, C.; Bielova, N.; Santos, C.: Do cookie banners respect my choice?: Measuring legal compliance of banners from iab europe's transparency and consent framework. In: *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, pp. 791–809, 2020.
- [MC08] McDonald, A.; Cranor, L. F.: The cost of reading privacy policies. *Journal of Law and Policy for the Information Society*, 4:543, 2008.
- [MFF01] Millett, L. I.; Friedman, B.; Felten, E.: Cookies and web browser design: Toward realizing informed consent online. In: *In Proceedings of the 2001 CHI Conference on Human Factors in Computing Systems*. pp. 46–52, 2001.
- [MK17] Macenaite, M.; Kosta, E.: Consent for processing children's personal data in the EU: following in US footsteps? *Information & Communications Technology Law*, 26(2):146–197, 2017.

- [NKA21] Norouzi, Y.; Keshavarz, H.; Athar, Z.: Evaluating children’s websites from an information visualization perspective: findings of a comparative mixed-methods study. *The Electronic Library*, ahead-of-print, 11 2021.
- [No20] Nouwens, M.; Liccardi, I.; Veale, M.; Karger, D.; Kagal, L.: Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In: *In Proc. of Human Factors in Computing Systems* (pp. 1-14). pp. 1–13, 2020.
- [Po21] Pochat, V. L.; Van Goethem, T.; Tajalizadehkhoob, S.; Korczyński, M.; Joosen, W.: , Tranco A Research-Oriented Top Sites Ranking Hardened Against Manipulation. <https://tranco-list.eu/>, 2021. Last accessed 19 October 2021.
- [Sa19] Sanchez-Rola, I.; Dell’Amico, M.; Kotzias, P.; Balzarotti, D.; Bilge, L.; Vervier, P.; Santos, I.: Can i opt out yet? GDPR and the global illusion of cookie control. In: *Proc. of the ACM Asia Conf. on Comp. and Comm. Sec.* pp. 340–351, 2019.
- [SBM20] Santos, C.; Bielova, N.; Matte, C.: Are cookie banners indeed compliant with the law? <https://arxiv.org/pdf/1912.07144.pdf>, 2020. Last accessed 27 August 2023.
- [Sh73] Shaw, A.: Dilemmas of “informed consent” in children. *The New England Journal of Medicine*, 1973.
- [So12] Solove, D.: Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.*, 126:1880, 2012.
- [Su21] Sun, K.; Sugatan, C.; Afnan, T.; Simon, H.; Gelman, S. A.; Radesky, J.; Schaub, F.: “They See You’re a Girl If You Pick a Pink Robot with a Skirt”: A Qualitative Study of How Children Conceptualize Data Processing and Digital Privacy Risks. In: *Proc. of the CHI Conference on Human Factors in Computing Systems*. 2021.
- [Va20] Vallina, P.; Le Pochat, V.; Feal, Á.; Paraschiv, M.; Gamba, J.; Burke, T.; Hohlfeld, O.; Tapiador, J.; Vallina-Rodriguez, N.: Mis-Shapes, Mistakes, Misfits: An Analysis of Domain Classification Services. In: *Proc. of the ACM Internet Measurement Conference. IMC ’20*, p. 598–618, 2020.
- [V118] Vlajic, N.; El Masri, M.; Riva, G. M.; Barry, M.; Doran, D.: Online Tracking of Kids and Teens by Means of Invisible Images: COPPA vs. GDPR. In: *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security. MPS ’18*, Association for Computing Machinery, New York, NY, USA, p. 96–103, 2018.
- [ZA20] Zimmeck, S.; Alicki, K.: Standardizing and implementing Do Not Sell. In: *Proceedings of the 19th Workshop on Privacy in the Electronic Society*. pp. 15–20, 2020.
- [Zh20] Zhao, F.; Egelman, S.; Weeks, H. M.; Kaciroti, N. A.; Miller, A. L.; Radesky, J. S.: Data Collection Practices of Mobile Applications Played by Preschool-Aged Children. *JAMA pediatrics*, p. e203345, 2020.
- [Zi23] Zimmeck, S.; Wang, O.; Alicki, K.; Wang, J.; Eng, S.: Usability and Enforceability of Global Privacy Control. *PETS*, 2:1–17, 2023.

Privacy-Preserving Stress Detection Using Smartwatch Health Data

Lucas Lange¹, Borislav Degenkolb² and Erhard Rahm³

Abstract: We present the first privacy-preserving approach for stress detection from wrist-worn wearables based on the Time-Series Classification Transformer (TSCT) architecture and incorporating Differential Privacy (DP) to ensure provable privacy guarantees. The non-private baseline results prove the TSCT to be an effective model for the given task. Our DP experiments then show that the private models suffer from reduced utility but can still be used for reliable stress detection depending on the application. Our proposed approach has potential applications in smart health, where it can be used to monitor smartwatch users' stress levels without compromising their privacy and provide timely interventions or suggestions to prevent adverse health outcomes. Another primary contribution is our evaluation, which studies and shows negative effects of DP regarding model training. The results of this work provide perspectives for future research and applications whenever the fields of stress detection and data privacy intervene.

Keywords: Stress Detection, Stress Recognition, Smartwatch, Time-Series Classification Transformer, Differential Privacy, Privacy-Preserving Machine Learning

1 Introduction

Stress is a prevalent issue that can have negative effects on both physical and mental health, leading to various health problems, such as cardiovascular disease, depression, and anxiety [Ya17]. Therefore, early detection is crucial for effective stress management and reduction, which can improve the quality of life for affected individuals.

In recent years, wearable devices such as smartwatches have emerged as a promising tool for monitoring stress levels, as they can collect various physiological signals such as heart rate, skin conductance, and accelerometer data, among others [Gi22a]. Machine learning techniques are widely used to analyze this data and detect patterns associated with stress [Sc18a]. However, the use of sensitive health data raises concerns about user privacy [Ja22; PZ18]. Through providing a provable privacy guarantee, users can benefit from health reports without stressing over privacy concerns.

In this work, we present the first approach to stress detection using smartwatch health data that preserves user privacy through Differential Privacy (DP) [Dw06]. We employ

¹ Leipzig University, Database Group, Augustusplatz 10, 04109 Leipzig, lange@informatik.uni-leipzig.de

² Leipzig University, Augustusplatz 10, 04109 Leipzig, bd79howo@studserv.uni-leipzig.de

³ Leipzig University, Database Group, Augustusplatz 10, 04109 Leipzig, rahm@informatik.uni-leipzig.de

a transformer-based architecture that achieves high accuracy in detecting stress levels from collected time-series data. The used WESAD dataset by Schmidt et al. [Sc18b] is a common standard dataset for this task (e.g. in related work presented in Sect. 3). Our privacy-preserving approach ensures that sensitive health information is protected while still allowing accurate predictions. A primary focus and contribution of our experimental evaluation are the provided insights into numerous drawbacks when training with DP. We thereby contribute to the growing body of research on privacy-preserving machine learning and its potential applications in healthcare.

Sect. 2 provides background information on relevant concepts. We then give an overview of related work in the field of stress detection in Sect. 3. Sect. 4 details our proposed approach and describes the conducted experiments. Their evaluation and results are then presented in Sect. 5. Finally, Sect. 6 summarizes the main contributions, discusses potentials and limitations, and suggests areas for future work.

2 Background

The following section goes over some of the fundamental concepts used in this work.

2.1 Stress detection

Stress detection, also known as stress recognition, is the process of detecting and identifying patterns in physiological signals that are associated with stress. Physiological signals such as heart rate variability, skin conductance, and respiration can provide valuable information about a person's stress level [Gi22a].

Machine learning techniques can be used to analyze these physiological signals and detect patterns associated with stress [Sc18a]. Wearable devices such as smartwatches have become increasingly popular for monitoring physiological signals related to stress and can collect data continuously throughout the day, providing a rich source of data for detection algorithms.

2.2 Time-Series Classification Transformers

Transformers as introduced by Vaswani et al. [Va17], are a type of neural network architecture that has shown excellent performance in natural language processing tasks. They use a self-attention mechanism to weigh the importance of different parts of the input sequence when making predictions.

Time series transformers extend the transformer architecture to handle time-series data [We23]. They incorporate temporal information into the self-attention mechanism by

operating on a sliding window of fixed length that moves along the time axis. This allows them to better handle long-term dependencies and more efficiently train on longer sequences [We23]. Time-Series Classification Transformers (TSCT) are a specific type of time series transformer architecture designed for classification problems. It has no decoder part and is better suited for our classification-based task of stress detection in physiological signals.

2.3 Differential Privacy

The concept of DP [Dw06] is a mathematical definition of privacy that ensures that the inclusion or exclusion of any individual’s data in a dataset does not significantly change the output of statistical queries on that dataset. The framework is based on the idea of adding noise to query answers (including sensitive person-related information) to protect individual privacy while preserving the utility of the dataset for analysis.

Formally, an algorithm A training on a set S is called (ϵ, δ) -differentially-private, if for all datasets D and D' that differ by exactly one record:

$$Pr[A(D) \in S] \leq e^\epsilon Pr[A(D') \in S] + \delta \quad (1)$$

The ϵ -parameter measures the level of privacy protection provided by the mechanism. It thus determines the amount of random noise that is added to the output of queries on the dataset to protect individual privacy. The smaller the value of ϵ , the stronger the privacy protection, but also the greater the amount of noise.

2.4 Differentially Private Stochastic Gradient Descent

Differentially Private Stochastic Gradient Descent (DP-SGD) [Ab16] is a variant of the stochastic gradient descent optimization algorithm that provides differential privacy guarantees. DP-SGD works by adding controlled noise to the gradients computed on each mini-batch of data during the training process. The amount of noise added is controlled by the DP privacy parameter ϵ , which determines the strength of privacy protection. Tuning the privacy parameter can be challenging, and careful selection and calibration of the noise level are required to achieve the desired privacy-utility trade-off.

3 Related Work

Out of the numerous publications on stress detection (e.g. as reviewed in [Sc18a]), we want to evaluate approaches from works also using the WESAD dataset by Schmidt et al. [Sc18b]. Further, we focus on solely using sensors available to wearable wrist devices like

Related work	Signals	Method	Accuracy	F1-score	ε
Schmidt et al. [Sc18b]	chest+wrist	LDA	92.28	90.74	∞
	wrist	Random Forest	87.12 ± 0.24	84.11 ± 0.31	∞
Siirtola [Si19]	wrist	LDA	87.40 ± 10.4	N/A	∞
Gil-Martin et al. [Gi22b]	chest+wrist	CNN	96.62 ± 0.11	96.63 ± 0.11	∞
	wrist	CNN	92.70 ± 0.16	92.55 ± 0.16	∞
Lange et al. (ours)	wrist	Transformer	91.89 ± 0.15	91.61 ± 0.34	∞
	wrist	DP-Transformer	78.88 ± 4.32	76.14 ± 4.27	10
	wrist	DP-Transformer	78.16 ± 8.40	71.26 ± 4.42	1
	wrist	DP-Transformer	71.15 ± 10.5	68.71 ± 5.70	0.1

Tab. 1: Classifier evaluation on the WESAD dataset for modalities collected from either chest or wrist devices regarding the binary (stress vs. non-stress) classification task. We compare accuracy (%) and F1-score (%) and include our achieved models on their ε -settings from DP.

smartwatches as per our goal, and compare performance for the binary task of classifying stress vs. non-stress.

In Tab. 1 we present the collected related works and also include our results in the overview. Approaches conforming to our criteria use the signals designated under the collective term *wrist*, while *chest+wrist* additionally uses the chest wearable data also available in the WESAD dataset. Currently, the best performing approach from related work for both signal collections is a CNN by Gil-Martin et al. [Gi22b]. It also reigns over the other models when omitting chest signals, where we see about 4% reduction in both its accuracy and its F1-score to 92.70% and 92.55%, respectively. Our results are included for comparison but evaluated later in Sect. 5.

We also find transformer models in related work, which however perform worse than the CNN approach and do not conform to our goals: transformer-based model on ECG [Be21] without wrist data, transformer-based model (Husformer) on wrist and chest signals [Wa23] without binary task. In summary, there are no prior transformer stress detection models solely based on wrist device data for our task, which is a gap we want to close.

We further find only one private approach for stress detection by Can; Ersoy [CE21]. However, they employ federated learning without additional privacy implementations, which provides only limited privacy while processing data [Bo23] and also leaves the resulting classification model open to attacks [Sh17]. With DP, we can guarantee provable privacy for our classification model, thus making this work a valuable contribution [DR+14].

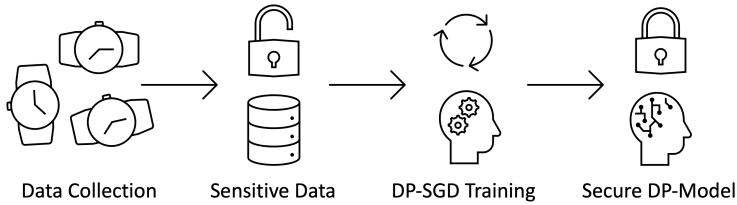


Fig. 1: Illustration of our DP-ensuring machine learning workflow. Instead of privatizing the data itself, we utilize DP-SGD training to directly create a private classification model ready for publishing.

4 Experimental Setup

In this section, we outline our study and give background on the experiments. Reference code is available from our GitHub repository: <https://github.com/BDegenkolb/Privacy-Preserving-Stress-Transformer>.

4.1 Methodology

The general scenario for this work is presented as a basic workflow illustration in Fig. 1. To provide smart health services, a trusted institution or data owner collects the sensitive health data from their users’ smartwatches to form a unified dataset for analysis. This data is then funneled into training machine learning models, like in our case for stress detection. To provide good results to all users, especially new device owners, it is important to achieve a generalizing model and this one-for-all approach is therefore favorable compared to creating only a personal model for each user. The final models are later delivered back to the devices for operation, leading to privacy concerns regarding the sensitive training data. A provable protection through DP is achieved by applying the DP-SGD training algorithm. Incorporating this workflow, the final private model is then safe for publishing and can provide health insights at a decreased privacy risk.

Our main objective is to investigate the feasibility of using transformer-based architectures for stress detection using smartwatch health data while preserving user privacy through DP. To evaluate the addition of DP, we first establish a baseline classification model for comparison and as a starting point for our private model. In our baseline performance we want to compete with the current best model by Gil-Martin et al. [Gi22b] (for more related work see Sect. 3). We choose a transformer model instead of CNN because of its capability of capturing long distance dependencies in time-series data [Li22]. On the basis of this baseline, we then evaluate the consequences of using DP regarding model performance.

Our private approach takes the existing baseline model and applies DP-SGD to train for DP. To better quantify the performance loss from DP, we train multiple models for different privacy levels of $\epsilon = \{10, 1, 0.1\}$. As a non-private model, the baseline is designated with

the privacy level $\epsilon = \infty$. Values of $\epsilon \leq 1$ are a common choice for good privacy [Ca19b; Na21] and through also including ϵ -values a magnitude larger and smaller than $\epsilon = 1$, we can better analyze the utility-privacy trade-off regarding our task [La23]. Our privacy accounting needs specific considerations due to the time series data and subject-oriented evaluation, which are given in Sect. 4.6.

4.2 Environment

This section details the software and hardware environment in which all experiments were implemented and run. Our description ensures reproducible results, which is also achieved by using a random seed set to a fixed number of 42 to get comparable training results.

Software-wise, we employ the Python programming language with machine learning workflows implemented using the Tensorflow⁴ and Keras⁵ libraries. For DP-SGD and other privacy-related implementations, we use Tensorflow Privacy⁶. For running the experiments we relied on Google Colab⁷, which enabled us to use a NVIDIA V100 Tensor GPU and 25 GB of system memory for our Jupyter notebooks.

4.3 Dataset

We study the performance of our proposed approach on the publicly available multimodal WESAD dataset [Sc18b], which is the standard dataset for our task. This dataset contains 15 healthy subjects (12 male, 3 female) with about 30 minutes of recorded health data each, which was acquired during a lab study. The physiological and motion data is sampled continuously and simultaneously from a wrist- and a chest-worn device, with the Empatica E4 wrist device offering the following modalities in different sampling frequencies: blood volume pulse (BVP), electrodermal activity (EDA), body temperature (TEMP), and three-axis acceleration (ACC). The three affective states recorded are stress, neutral, and amusement.

We focus on binary classification as stress vs. non-stress and only consider signals available from a wrist-worn wearable like smartwatches—i.e. the Empatica E4 device modalities from the WESAD dataset.

4.4 Preprocessing

An important prerequisite for working with continuous time-series data is signal preprocessing. Our methods are based on prior work by [Gi22b], who presented a throughout process.

⁴ Available at <https://github.com/tensorflow/tensorflow>

⁵ Available at <https://github.com/keras-team/keras>

⁶ Available at <https://github.com/tensorflow/privacy>

⁷ Available at <https://colab.research.google.com>

Signal	Sampling frequency	Frequency range	Subwindow length	Model inputs
ACC	64 Hz	0–30 Hz	7 seconds	210
BVP	64 Hz	0–7 Hz	30 seconds	210
EDA	64 Hz	0–7 Hz	30 seconds	210
TEMP	64 Hz	0–6 Hz	35 seconds	210

Tab. 2: Processing details for each signal sampled by the Empatica E4, based on [Gi22b].

Block	Layer	Settings	Data shape
Input	Input layer	N/A	(6, 210)
Encoder (x8)	Layer normalization	epsilon 1e-6	(6, 210)
	Multi head attention	256 head size, 4 heads	(6, 210)
	Dropout	0.25	(6, 210)
	Add	residual connection	(6, 210)
	Layer normalization	epsilon 1e-6	(6, 210)
	Convolution 1D	4 filters of 1x1, ReLU	(6, 4)
	Dropout	0.25	(6, 4)
	Convolution 1D	6 filters of 1x1, linear	(6, 210)
Output	Add	residual connection	(6, 210)
	Avg-Pooling 1D	reduce feature space	(6)
	Fully connected	128 units, ReLU	(128)
	Dropout	0.25	(128)
	Fully connected	2 units, sigmoid	(2)

Tab. 3: Detailed view of our TSCT architecture based on the basic model introduced by Keras [Nt]. Data shape is given as (features, sequence length), which is a ordering we found to be desirable.

In general, our preprocessing workflow is divided into three steps: 1. upsampling (Fourier method), 2. segmentation (windows), and 3. Fast Fourier Transformation (FFT).

The Empatica E4 wristband collects its modalities in differing heterogeneous sampling rates, which for consistency are upsampled to 64 Hz for all signal types using the Fourier method. This adjustment is important to have the signals available at equal time intervals when inputting in our transformer architecture. Afterwards, the data streams are segmented into 60-second windows and subdivided into different length P-second subwindows with a shift of 0.25 seconds. The subwindow length depends on the signal’s frequency range and is used to average the 60-second window’s spectrum (Fourier transform module) along the subwindows using FFT. This signal-specific spectrum transformation allows to gain a consistent 210 frequency points per 60-second window. These 210 frequency points represent the input for our machine learning model, which then classifies the data at the 60-second window level. We give an overview of preprocessing settings for each signal in Tab. 2. The overall preprocessing result for each subject in the dataset are time series inputs with 6 signals and a sequence length of 210 frequency points per signal.

4.5 None-Private Baseline

The basic structure of our TSCT architecture is shown in Tab. 3 and based on the Keras time series transformer [Nt], where the projection layers are implemented as 1D Convolutions. The model underwent a rigorous hyperparameter tuning process delivering the presented model architecture and settings. We found an optimum at 8 stacked encoder blocks inside the model. We train for 110 epochs with a batch size of 50 and a learning rate of $1e-4$. Our input shape takes the form of (features, sequence length), with features being the 6 available signals and 210 points of measurement each. While this is an unusual ordering compared to (sequence length, features), we found this setup to produce better classification results.

For effectively measuring performance in a small dataset of different patients, the evaluation is subject-based using the Leave One Subject Out (LOSO) method. We cycle through the 15 subjects and in each step take one subject as test and the other 14 as training sets for our machine learning model. The overall metrics are then averaged over all 15 single test results. The LOSO method is also used in related work which helps our comparability, especially to the CNN model by [Gi22b].

4.6 Privacy-Preserving Approach

We have to define multiple relevant variables when transferring a non-private baseline model to DP. Total epochs, training set size, batch size, and δ all influence the needed noise scale to achieve our desired ϵ -guarantees. While the batch size is fixed at 50, there are further calculations needed for the other values. Due to our evaluation strategy, we specifically consider the DP for our averaged LOSO model over all subjects, instead of a single subject.

For the relationship between epochs and DP, we take into account how often a single sample from the training set is seen during the whole training, since each appearance increases privacy risk [DR+14]. For us, this single sample is one data point in the time series. One epoch including the data point translates to one appearance and we therefore assume the maximum epochs of one subject. This is due to the implementation of our LOSO method, which means each of our 15 subject taking part in 14 training sets for the other subjects. Thus the subject's data points are involved in 14 training processes with 110 epochs each, leading to a total epoch count of $e_{total} = 14 * 110 = 1,540$.

Regarding our training set size, preprocessing results in multiple windows with 210 data points and 6 measurements each for our 15 subjects. We want privacy on the data point level of our time series, which is called event-level DP [Ca19a; Mi23]. In our setting however, in addition to hiding just single data points, we additionally want to hide the correlations between windows. We therefore only consider one window per subject to be unique, while the others are treated as copies, which is the worst-case regarding their sampling rate [DR+14]. Again, increasing the sampling rate of an item leads to higher noise needs, because the influence of the single sample on the model is rising proportionally.

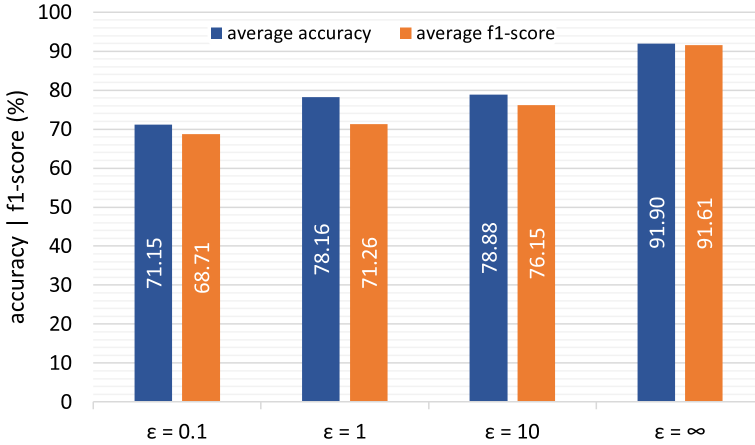


Fig. 2: Average performance over multiple runs each for our non-private and private models (DP).

ϵ	Accuracy	F1-score	Precision	Recall	Poor models
∞	91.89 ± 0.15	91.61 ± 0.34	91.90 ± 0.42	91.35 ± 0.49	None
10	78.88 ± 4.32	76.14 ± 4.27	77.58 ± 10.1	78.78 ± 16.0	$\approx 33\%$
1	78.16 ± 8.40	71.26 ± 4.42	68.05 ± 10.2	79.09 ± 14.6	$\approx 50\%$
0,1	71.15 ± 10.5	68.71 ± 5.70	60.77 ± 9.30	85.39 ± 18.2	$\approx 75\%$

Tab. 4: Comparison of average model results for non-private and private DP-models. Accuracy, F1-score, precision, and recall are given as %. We also consider their standard deviation. The ratio of poor models tells us how many of the trained model runs failed to reach more than 30% performance.

Meaning, a smaller dataset leads to higher DP noise. Our training set size for determining DP noise is thus given as $n = 210 * 6 * 15 = 18,900$. From that we can also determine $\delta = 1e - 5$ according to $\delta \ll \frac{1}{n}$ [Dw06].

5 Results

To stabilize the results and get a more representative view, we run reach LOSO experiment multiple times and retrieve the best and averaged performance over these runs. On our machine learning task some models either classify all given samples as only stress or only non-stress, thus failing to learn a useful representation of the underlying data. For ease of use we denote these non-representative models as *poor* models in the following. Poor models, in our case, hover around low performance levels of under 30%. When presenting averaged metrics in our evaluation we therefore only consider models achieving more than 30% to not disproportionately decrease the average.

Fig. 2 shows the averaged results for each evaluated privacy level over their tracked experiments. For our non-private performance, we did not manage to outperform the related work by Gil-Martin et al. [Gi22b] but came close with under 1% difference, see Tab. 1 for a full comparison to related work. Our averaged model results show a general trend of decreasing utility when training for better privacy guarantees. The actual reduction in performance from non-private to $\epsilon = 10$ (max. 15.45%), however, is more significant than inbetween the private models (max. 7.43%). In our case, F1-scores generally takes a bigger hit than accuracy when introducing DP.

In Tab. 4, we present more details in addition to the averages from Fig. 2. We include the recall and precision for each model, which shows that the lowering F1-scores at $\epsilon = \{1, 0.1\}$ are mostly due to a reduction in precision, while recall keeps higher values. Additionally, the now given standard deviation for our averages is significantly higher when training with DP and for the most part increases further when tightening the ϵ -setting. We also provide information on the ratio of poor models, i.e. models with less than 30% performance, and see a stern increase when lowering ϵ for better DP-guarantees. While in non-private training all models were able to achieve feasible results, we recorded between 33–75% failures on $\epsilon = \{10, 1, 0.1\}$. These results mark the importance of running several rounds of experiments when working with random noise from DP to get an idea of actual performance. We find the induced noise from DP significantly influences model convergence on all levels.

6 Conclusion

We presented the first differentially-private approach to smartwatch stress detection and considering a transformer architecture. While not being able to surpass them, we can however confirm to be close to the related best CNN results from Gil-Martin et al. [Gi22b] regarding our averaged non-private baseline results with under 1% less performance (91.89% accuracy and 91.61% F1-score). The utility-privacy loss recorded for our DP-enabled models lies at over 15%, which is a significant drop to consider. An accuracy of 78.16% and an F1-score of 71.26% at the strong privacy guarantee of $\epsilon = 1$ is a fair result but the model's feasibility depends on the actual usage scenario. In critical healthcare settings it could be intolerable to lose over 15% utility, while simple lifestyle recommendations from smart health applications on smartwatches might still be reasonable. Regarding the F1-score, we mainly saw a rapid lowering in precision when tightening the DP-guarantees. A low precision score can contribute to increased false positives, which could in the end could annoy users to a point of not using the stress detection feature. We can thus not recommend using very strong privacy guarantees in precision-dependent applications.

A central limitation is the unclear generalizability because of the WESAD dataset characteristics regarding its small size and distribution of gender, age, etc. [Sc18b]. Thus, a highly relevant future venture would be the transition from research to real world studies, where the larger amounts of user data available could also result in better performance for our transformer. In the same vein, generating more training data through synthetization methods

could have a positive impact. Another promising opportunity for future work would be the inclusion of other mobile data sources that could be relevant for stress detection, e.g. mobile phone usage [SP13].

Acknowledgments. We first thank the reviewers for their helpful feedback. We thank Nils Wenzlitschke for his contributions to our preprocessing implementation. The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany and by the Sächsische Staatsministerium für Wissenschaft Kultur und Tourismus in the program Center of Excellence for AI-research "Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig", project identification: ScaDS.AI.

Bibliography

- [Ab16] Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H. B.; Mironov, I.; Talwar, K.; Zhang, L.: Deep learning with differential privacy. In: ACM SIGSAC. 2016.
- [Be21] Behinaein, B.; Bhatti, A.; Rodenburg, D.; Hungler, P.; Etemad, A.: A Transformer Architecture for Stress Detection from ECG. In: ISWC '21. 2021.
- [Bo23] Boenisch, F.; Dziedzic, A.; Schuster, R.; Shamsabadi, A. S.; Shumailov, I.; Papernot, N.: When the Curious Abandon Honesty: Federated Learning Is Not Private, 2023, arXiv: 2112.02918 [cs.LG].
- [Ca19a] Cao, Y.; Yoshikawa, M.; Xiao, Y.; Xiong, L.: Quantifying Differential Privacy in Continuous Data Release Under Temporal Correlations. *IEEE Transactions on Knowledge and Data Engineering* 31/7, 2019.
- [Ca19b] Carlini, N.; Liu, C.; Erlingsson, Ú.; Kos, J.; Song, D.: The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks. In: *USENIX Security Symposium*. Vol. 267, 2019.
- [CE21] Can, Y. S.; Ersoy, C.: Privacy-preserving federated deep learning for wearable IoT-based biomedical monitoring. *ACM TOIT* 21/1, 2021.
- [DR+14] Dwork, C.; Roth, A., et al.: The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science* 9/3–4, 2014.
- [Dw06] Dwork, C.: *Differential Privacy*. In: *Automata, Languages and Programming*. Springer Berlin Heidelberg, 2006.
- [Gi22a] Giannakakis, G.; Grigoriadis, D.; Giannakaki, K.; Simantiraki, O.; Roniotis, A.; Tsiknakis, M.: Review on Psychological Stress Detection Using Biosignals. *IEEE Transactions on Affective Computing* 13/01, 2022.
- [Gi22b] Gil-Martin, M.; San-Segundo, R.; Mateos, A.; Ferreiros-Lopez, J.: Human Stress Detection With Wearable Sensors Using Convolutional Neural Networks. *IEEE Aerospace and Electronic Systems Magazine* 37/1, pp. 60–70, 2022.

- [Ja22] Jafarlou, S.; Rahmani, A. M.; Dutt, N.; Mousavi, S. R.: ECG Biosignal Deidentification Using Conditional Generative Adversarial Networks. In: Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2022.
- [La23] Lange, L.; Schneider, M.; Christen, P.; Rahm, E.: Privacy in Practice: Private COVID-19 Detection in X-Ray Images, 2023, arXiv: 2211.11434 [cs.LG].
- [Li22] Li, C.; Huang, X.; Song, R.; Qian, R.; Liu, X.; Chen, X.: EEG-based seizure prediction via Transformer guided CNN. *Measurement* 203/, 2022.
- [Mi23] Miranda-Pascual, À.; Guerra-Balboa, P.; Parra-Arnau, J.; Forné, J.; Strufe, T.: SoK: Differentially Private Publication of Trajectory Data. *Proceedings on Privacy Enhancing Technologies* 2/, pp. 496–516, 2023.
- [Na21] Nasr, M.; Songi, S.; Thakurta, A.; Papernot, N.; Carlin, N.: Adversary instantiation: Lower bounds for differentially private machine learning. In: 2021 IEEE Symposium on security and privacy (SP). IEEE, pp. 866–882, 2021.
- [Nt] Ntakouris, T.: Timeseries classification with a Transformer model, https://keras.io/examples/timeseries/timeseries_transformer_classification/, accessed: 05/03/2023.
- [PZ18] Perez, A. J.; Zeadally, S.: Privacy Issues and Solutions for Consumer Wearables. *IT Professional* 20/4, 2018.
- [Sc18a] Schmidt, P.; Reiss, A.; Duerichen, R.; Laerhoven, K. V.: Wearable affect and stress recognition: A review, 2018, arXiv: 1811.08854 [cs.HC].
- [Sc18b] Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K.: Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. In: ACM international conference on multimodal interaction. 2018.
- [Sh17] Shokri, R.; Stronati, M.; Song, C.; Shmatikov, V.: Membership Inference Attacks against Machine Learning Models, 2017, arXiv: 1610.05820 [cs.CR].
- [Si19] Siirtola, P.: Continuous Stress Detection Using the Sensors of Commercial Smartwatch. In: *UbiComp/ISWC '19 Adjunct*. 2019.
- [SP13] Sano, A.; Picard, R. W.: Stress Recognition Using Wearable Sensors and Mobile Phones. In: *Humaine Association Conference on ACII*. 2013.
- [Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I.: Attention is all you need. *NeurIPS* 30/, 2017.
- [Wa23] Wang, R.; Jo, W.; Zhao, D.; Wang, W.; Yang, B.; Chen, G.; Min, B.-C.: Husformer: A Multi-Modal Transformer for Multi-Modal Human State Recognition, 2023, arXiv: 2209.15182 [cs.HC].
- [We23] Wen, Q.; Zhou, T.; Zhang, C.; Chen, W.; Ma, Z.; Yan, J.; Sun, L.: Transformers in Time Series: A Survey, 2023, arXiv: 2202.07125 [cs.LG].
- [Ya17] Yaribeygi, H.; Panahi, Y.; Sahraei, H.; Johnston, T. P.; Sahebkar, A.: The impact of stress on body function: A review. *EXCLI journal* 16/, 2017.

Privacy Aware Processing

Sharing Uncritical Information to Enable Data Understanding and Preparation on Sensitive Data for Machine Learning


Marian Eleks¹ , Jonas Rebstadt² , Henrik Kortum³  and Oliver Thomas⁴

Abstract: In many machine learning (ML) applications, the provision of data and the training as well as the analysis of machine learning systems are performed by distinct actors, a data owner and a data consumer. To protect sensitive information in these ML-scenarios, privacy aware machine learning (PAML) methods are often applied to the data before sharing. Based on the type of PAML methods used, data understanding and preparation as defined in the CRISP-DM model become more difficult if not impossible. To enable these steps, we propose a method to share a variety of uncritical information with the data consumer who is then able to define the necessary processing steps on a meta-level. These are then applied to the data in the data owners local trusted environment before the PAML-methods whereupon the prepared and protected data is shared.

Keywords: Privacy Aware Machine Learning, Data Understanding, Data Preparation, CRISP-DM, Artificial Intelligence, Machine Learning

1 Introduction

Privacy aware machine learning (PAML) is becoming increasingly relevant due to the widespread adoption of machine learning (ML) in various application domains, as well as growing concerns about data privacy and security. PAML use cases with distinct data owners and data consumers find growing interest as ML permeates domains where ML knowledge is limited but data is highly sensitive. In a practical example. private households or the housing industry have an urge to reduce their usage of energy but don't want to or are not allowed to share the relevant information for promising ML-services [Ko23]. In these cases, the sensitive information in the patient data would be protected through various PAML-techniques before being sent to the researchers, fulfilling the

¹ Stratagon GmbH, Albert-Einstein-Straße 1, 49076 Osnabrück, marian.eleks@stratagon.de, 
<https://orcid.org/0000-0002-1516-5129>

² Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Smart Enterprise Engineering, Hamburger Straße 24, 49084 Osnabrück, jonas.rebstadt@dfki.de, 
<https://orcid.org/0000-0001-8531-3273>

³ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Smart Enterprise Engineering, Hamburger Straße 24, 49084 Osnabrück, henrik.kortum@dfki.de, 
<https://orcid.org/0000-0002-1089-711X>

⁴ Universität Osnabrück, Informationsmanagement und Wirtschaftsinformatik, Hamburger Straße 24, 49084 Osnabrück, oliver.thomas@uni-osnabrueck.de

privacy goals but preventing meaningful data understanding and preparation. However, a well-founded data understanding and preparation is essential for an optimal modelling outcome and are found in most AI-Lifecycle frameworks [Am19], [AS08], [Er17] including CRISP-DM [WH00]. To enable data understanding and preparation in PAML use cases with distinct actors, a technique of extracting and sharing uncritical information from the data called “privacy aware preview” is developed. Based on this, “privacy aware processing” is constructed with the core idea of the data consumer defining data preparation instructions based on the data understanding achieved in the privacy aware preview, which are then executed in the data owner’s local environment to achieve prepared and private data that is then shared with the data consumer. First, related work for privacy aware machine learning and its impact on CRISP-DM is introduced. Then, existing approaches for the previously outlined problem are collected in a literature review followed by the development of this paper’s solution, based on the findings in the literature. Lastly, the impact of the solution on CRISP-DM is addressed, followed by a discussion of the results and a conclusion.

2 Privacy Aware Machine Learning and its Consequences for CRISP-DM

The term Privacy Aware Machine Learning (PAML) describes “[...] the discipline of applying Machine Learning techniques in such a way as to protect and retain personal identities during the process.” [MKH17]. The goal of the discipline is to share and analyse data while protecting sensitive attributes [El22]. PAML might also be used to describe similar undertakings [AC19]. The literature mainly mentions six techniques: K-Anonymity generalizes and suppresses data, resulting in a dataset where any record has at least k records with identical quasi-identifiers [RMS18], [SD18], [Sw02a]. Differential Privacy’s basic idea is to add noise to the data before releasing it, in order to make it difficult to determine whether any particular individual’s data is included in the dataset [AC19], [Dw06], [LZ20], [RMS18], [SD18]. Synthetic Data, while mimicking the statistical properties of sensitive data, does not contain any real-world information, making it privacy preserving [Ab19], [HEM19], [Pa18]. Homomorphic Encryption enables computations on encrypted data so that model training may be accomplished on said data without decryption [AC19], [LZ20]. Similarity Preserving Hashing transforms the data to remove any interpretable information while preserving utility for ML [El22]. Lastly, Federated Learning keeps sensitive data on local devices, training individual models locally. A central model is achieved by aggregating the locally trained models without ever combining or seeing local sensitive data [Me21].

While PAML techniques make model training on sensitive data possible while preserving privacy, most ML procedure model feature a data understanding as well as a data preparation stage [Am19], [AS08], [Er17], [WH00]. By applying PAML methods to the data in the distinct actor use case, data understanding and preparation are complicated if not impossible since the data consumer receives incomplete (k -Anonymity), perturbed

(Differential Privacy) or incomprehensible (SPH, Homomorphic Encryption) information. The cycle of procedure models like CRISP-DM is broken in PAML use cases, raising a need for techniques that enable data understanding and preparation in a privacy aware context. This need is further justified in the literature [Br21], [Fa20], [JLE14], [LP21].

3 Privacy Aware Preview and Processing

3.1 Existing approaches and current limitations

To find existing solutions and ideas concerning data understanding, preprocessing and feature engineering in a privacy aware environment, a literature review using the search term ("*privacy*" OR "*security*" OR "*sensitive*") AND "*machine learning*" AND ("*preprocessing*" OR "*feature engineering*" OR "*data understanding*") is carried out. It yields approaches for privacy aware data understanding and privacy aware data preparation as well as current challenges and needs.

The simplest approach for privacy aware data understanding is to remove all sensitive data before sharing [SBG21], which is not practical since it entails large losses in data utility. Differential Privacy can also be used to share aggregate information about the data while preserving the individuals privacy [JLE14], [Sh23], which is beneficial to assess data distribution, value spaces, imbalances and missing values. Additionally, [JLE14] mention data generation as a possible solution for private data release, raising the possibility to share synthetic plain text data as a proxy for understanding the original data. [Jo19] invert the use case and use feature engineering to hide sensitive information, e.g., transforming GPS-coordinates into a place label like "Walgreens parking lot". This idea promises great results in use cases where a trusted entity is able to prescribe these kinds of transformations. This entity is not present in the use case of this paper, so it is not applicable. A final approach is identifying similar public datasets to the private data [JLE14], but since these are seldomly available, they will also not be used further.

All approaches for privacy aware data preparation found in the literature implicitly assume that some kind of data understanding has already taken place to be able to clearly define necessary preprocessing and transformation instructions. Thus, privacy aware data understanding is a prerequisite for privacy aware data preparation. The most promising proposition is to have the data consumer define a list of preprocessing and transformation instructions, which are then applied to the data in a trusted environment [HH22], [Jo19]. Another idea is to use homomorphic encryption in a federated learning environment to apply data preparation algorithms to encrypted data [Fa20], [HH22]. Going along with the first approach, these algorithms need to either be applicable to all kinds of raw data or be defined based on a previous data understanding. Finally, [JLE14] mention the use of principal component analysis or forms of feature selection and feature importance calculation to reduce the number of attributes and/or achieve dimensionality reduction in a privacy aware manner.

As mentioned in the data preparation, a need for privacy aware data understanding techniques is present in the literature, either explicitly [Br21], [JLE14] or implicitly [Fa20], [HH22], [Jo19]. Building from that, some papers mention a necessity for clean and complete data, implying a need for privacy aware preprocessing techniques. In some cases, experience in data science is assumed for the data owner [Fa20], [HH22], [Jo19], since they are expected to perform data manipulation. For this paper, the data does not need to have advanced knowledge of data manipulation for the solution to function. Conversely, [HH22], [JLE14] mention that the data needs to be in a relational format for some solutions to function, which is also true for this paper.

3.2 Sharing Uncritical Information for a well-founded Data Understanding

Being a prerequisite for all subsequent steps, a process for privacy aware data understanding based on sharing uncritical information is designed. A preview component generates as much uncritical information as possible from the raw data inside of the data owner's environment. Types of uncritical information from the literature review include synthetic data, which at this point does not need to come from an optimized model to provide an understanding of the data, saving resources and time. Moreover, aggregated statements fulfilling some degree of differential privacy based on requirements from the data owner can be shared to assess data distribution, value spaces, imbalances, and missing values. Using the ideas for dimensionality reduction, the data user can also be informed about the importance of attributes for certain use cases. In addition to the types found in the literature, uncritical meta information can be shared without breaching the privacy of any individual in the data. This includes attribute names, grammatical data types (int, float, string, ...) and statistical data types (categorical, ordinal, binary, ...). By sharing these types of information about the data with the data user, they gain an understanding of the data without acquiring sensitive information, as is shown in Figure 1.

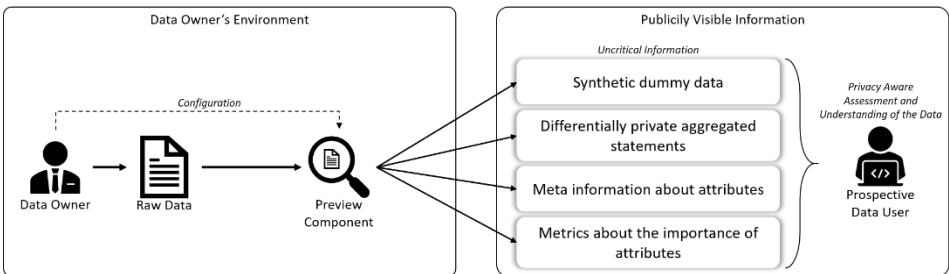


Figure 1: Privacy Aware Preview

Depending on the use case, the data understanding might be sufficient in and of itself, like in data matching or when the data user purchases usage rights to the private data after the preview. In other cases, it is used as a foundation for the following privacy aware processing.

3.3 Processing of Private Data for Data Preparation

Following [HH22], [Jo19], the privacy aware processing component allows the data user to define a list of processing steps that are then executed in the data owners trusted environment on the raw data before applying PAML-methods. This enables data users to perform data preparation on the raw data without having to breach privacy.

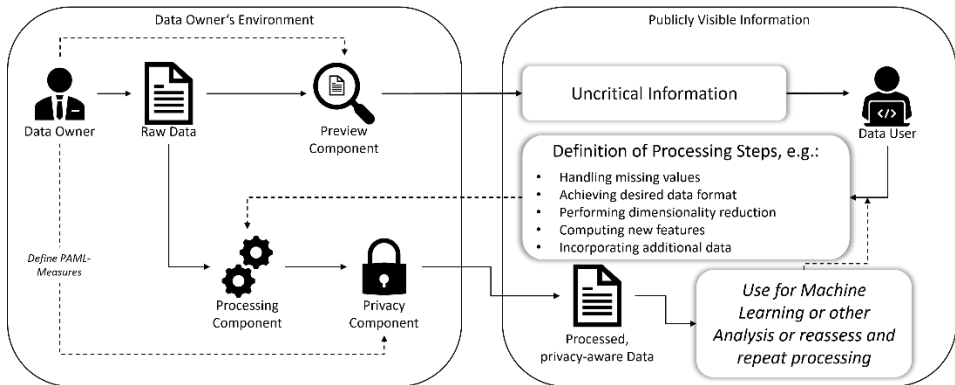


Figure 2: Privacy Aware Processing

In the full process as shown in Figure 2, data owners provide their raw data to the preview component, which they can adjust based on their notion of necessary privacy. The preview component then generates some or all of the possible types of uncritical information defined in 3.2 for the prospective data user to receive. The data user is then able to reach a decision on whether to use the data for their endeavours as well as gain an understanding of the inherent properties and necessary processing steps to undertake for data preparation. Following this, by executing a list of instructions defined by the data user in the environment of the data owner, data processing or data preparation in CRISP-DM terms can be carried out in a privacy preserving manner. Processing instructions can range from dealing with missing values or unclean data through reformatting data all the way to computing new features from the data or incorporating additional attributes from other datasets owned by the data user into the resulting dataset.

To illustrate the practical use of privacy aware processing, an example based on an insurance cost prediction dataset [Wa23] is established in Figure 3. Instead of receiving e.g. a homomorphically encrypted dataset that is uninterpretable for humans and not prepared to be used in ML, the data user receives the previously described preview. From this, they are able to deduce that with this dataset using the label “charges” and dropping any empty labels (1), a supervised training use case to predict insurance cost is possible. Additionally, they are able to deduce that the attribute region is unnecessary (2) and that the categorical attribute sex needs to be dummy encoded (3). Smoker is a string variable so it needs to be cast to binary/boolean (4) and age contains empty values that can be imputed (5). Listing further examples, the data user might choose to add normalization

instructions to age, children and BMI after seeing their value ranges, correct imbalances identified through the differentially aggregated percentiles or compute additional features (e.g. a special flag for increased COPD risk from the BMI and smoker attributes). All of these instructions are represented in source code, sent to the data owner and executed in their environment.

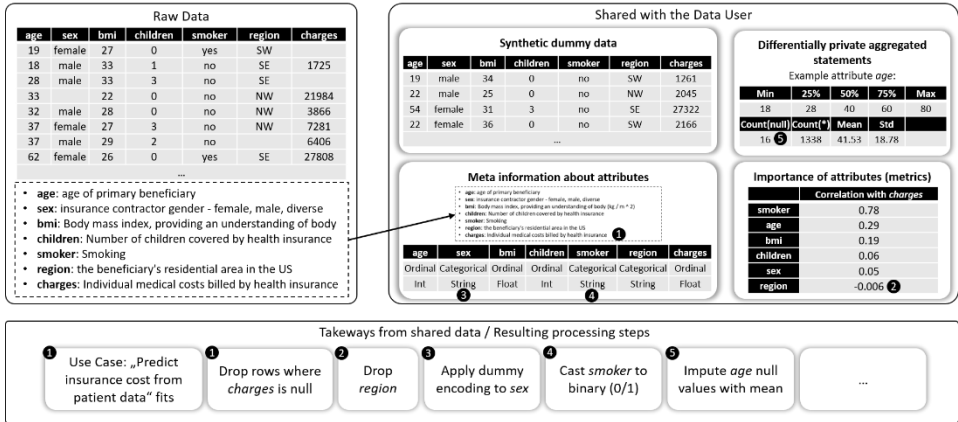


Figure 3: Practical example (insurance cost)

After the data preparation is finished, a privacy component applies PAML-measures to the data depending on the privacy requirements defined by the data owner. These can take the form of achieving k-Anonymity [RMS18], [SD18], [Sw02a] or Differential Privacy [AC19], [Dw06], [LZ20], [RMS18], [SD18] through generalization, suppression, or perturbation. Homomorphically encrypting the dataset is also possible [AC19], [LZ20] as well as training an optimized generative model to synthesize data [Ab19], [HEM19], [Pa18]. A most extreme measure would be the application of Similarity Preserving Hashing Anonymization algorithms to the data to remove any human interpretability [El22]. At this point, the resulting dataset is preprocessed, prepared and void of sensitive information and can be shared with the data user who is then able to use it as they please. They might also choose to revise their processing instructions based on newly gained insights achieved through modelling and evaluation using the data. This creates a cyclic process to achieve increasingly superior results through iterative revision, adhering to the CRISP-DM model.

4 Closing the Loop: Implications for Development Processes with CRISP-DM

In a PAML-context, the CRISP-DM cycle is broken through the harder and sometimes impossible understanding and preparation of privacy aware data. This paper mends the cycle through the ideas of privacy aware data preview and privacy aware data processing.

By providing uncritical information through various privacy aware data sharing techniques, all data understanding activities can be enabled in a privacy aware context. Synthetic initial data is provided and meta information about attributes as well as differentially private aggregated statements enable data description, exploration, and quality verification in a privacy preserving context. Data users are able to define a list of instructions for the processing component based on their data understanding and domain knowledge, enabling all data preparation activities. Additionally, privacy aware data understanding enables an informed modelling technique selection and test design, further facilitating CRISP-DM in a privacy aware environment.

5 Discussion and Conclusion

In this paper, the challenge of performing data understanding and preparation analogue to CRISP-DM in privacy aware use cases is uncovered and established through relevant literature. Based on current approaches, a solution consisting of a privacy aware data preview based on sharing uncritical information and a privacy aware data processing based on executing predefined preparation instructions in a trusted environment is designed. The solution promises to close the loop opened by applying CRISP-DM to privacy aware machine learning use cases. Its applicability is currently limited to CRISP-DM. While other AI process models seem promising, the solutions utility remains to be assessed. Additionally, it is limited to relational or tabular data and not suitable for e.g. images. As a main limitation, this paper stops at the design stage, leaving implementation and evaluation open for future research. Thus, utility assessments and acceptance by data consumers and owners respectively remain to be explored with special attention given to the most used and most useful processing steps. The designed solution does not reinvent data preprocessing or preparation but rather shifts the surrounding factors and underlying conditions to achieve better security for the data owner as well as the possibility for the data user to process raw data in their own image without having to breach privacy. As for the practical implications, this paper provides a groundwork on which domains with limited ML knowledge and highly sensitive data can build their training processes. It creates a starting point for new business models based on fully privacy aware machine learning as a service, providing solutions in domains like healthcare or smart living that fulfil the previously mentioned criteria. By extending an existing framework, service providers are able to build upon existing infrastructure and established pipelines, reducing redundant work and lowering the barrier to entry for these novel business cases. Overall, this paper solves a problem present in the literature and in established process models and provides a solution applicable to practical use cases in specific domains. Future research should focus on implementing and evaluating the privacy aware preview and processing designs in relevant applications domains. Based on the evaluation results and requirements for data consumers and owners, business models can be derived and realised in practice.

Bibliography

- [Ab19] Abay, N. C.; Zhou, Y.; Kantarcioglu, M.; Thuraisingham, B.; Sweeney, L.: Privacy preserving synthetic data release using deep learning. *Machine Learning and Knowledge Discovery in Databases*, pp. 510–526, 2019.
- [AC19] Al-Rubaie, M.; Chang, J. M.: Privacy-Preserving Machine Learning: Threats and Solutions. In: *IEEE Security and Privacy 2/19*, pp. 49–58, 2019.
- [Am19] Amershi, S.; Begel, A.; Bird, C.; DeLine, R.; Gall, H.; Kamar, E.; Nagappan, N.; Nushi, B.; Zimmermann, T.: Software Engineering for Machine Learning: A Case Study. *Proceedings - 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice, ICSE-SEIP 2019, Montréal*, pp. 291–300, 2019.
- [AS08] Azevedo, A.; Santos, M. F.: KDD, SEMMA and CRISP-DM: a parallel overview. In: *IADS-DM, 2008*.
- [Br21] Briguglio, W.; Moghaddam, P.; Yousef, W. A.; Traoré, I.; Mamun, M.: Machine learning in precision medicine to preserve privacy via encryption. In: *Pattern Recognition Letters 1/21*, pp. 148–154, 2021.
- [Dw06] Dwork, C.: Differential Privacy. *International Colloquium on Automata, Languages, and Programming*, Bd. 4052 LNCS, pp. 1–12, 2006.
- [El22] Eleks, M.; Rebstadt, J.; Fukas, P.; Thomas, O.: Learning without Looking: Similarity Preserving Hashing and Its Potential for Machine Learning in Privacy Critical Domains. *INFORMATIK 2022, 2022*.
- [Er17] Ericson, G.; Rohm, W.; Martens, J.; Harvey, B.; Schonning, N.: Team data science process documentation. Microsoft, 2017.
- [Fa20] Fang, P.; Cai, Z.; Chen, H.; Shi, Q.: FLFE: A Communication-Efficient and Privacy-Preserving Federated Feature Engineering Framework. *arXiv preprint arXiv:2009.02557 [cs.LG]*, 2020.
- [HEM19] Hittmeir, M.; Ekelhart, A.; Mayer, R.: Utility and Privacy Assessments of Synthetic Data for Regression Tasks. *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019, Los Angeles, CA*, pp. 5763–5772, 2019.
- [HH22] Hsu, R.-H.; Huang, T.-Y.: Private Data Preprocessing for Privacy-preserving Federated Learning. *2022 IEEE 5th International Conference on Knowledge Innovation and Invention (ICKII)*, pp. 173–178, 2022.
- [JLE14] Ji, Z.; Lipton, Z. C.; Elkan, C.: Differential Privacy and Machine Learning: a Survey and Review. *arXiv preprint arXiv:1412.7584 [cs.LG]*, 2014.
- [Jo19] Jones, M.; Johnson, M.; Shervey, M.; Dudley, J. T.; Zimmerman, N.: Privacy-Preserving Methods for Feature Engineering Using Blockchain: Review, Evaluation, and Proof of Concept. In: *Journal of Medical Internet Research 8/19*, pp. e13600, 2019.
- [Ko23] Kortum, H.; Hagen, S.; Eleks, M.; Rebstadt, J.; Remark, F.; Lowin, M.; Mihale Wilson, C.; Eberhardt, B.; Roß, A.; Maihöfner, D.; Hinz, O.; Thomas, O.: SECAI – Sustainable Heating through Edge-Cloud-based AI Systems. In: *HMD Praxis der Wirtschaftsinformatik 4/23*, pp. 1-22, 2023.

- [LP21] Lau, A.; Passerat-Palmbach, J.: Statistical Privacy Guarantees of Machine Learning Preprocessing Techniques. arXiv preprint arXiv:2109.02496 [cs.LG], 2021.
- [LZ20] Lisin, N.; Zapechnikov, S.: Methods and Approaches for Privacy-Preserving Machine Learning. *Advanced Technologies in Robotics and Intelligent Systems*. Bd. 80Springer, pp. 141–148, 2020.
- [Me21] Mercier, D.; Lucieri, A.; Munir, M.; Dengel, A.; Ahmed, S.: Evaluating Privacy-Preserving Machine Learning in Critical Infrastructures: A Case Study on Time-Series Classification. In: *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS X*, 2021.
- [MKH17] Malle, B.; Kieseberg, P.; Holzinger, A.: Interactive Anonymization for Privacy aware Machine Learning. Institute for Medical Informatics, Statistics & Documentation, Medical University Graz, Austria, Graz, 2017.
- [Pa18] Park, N.; Mohammadi, M.; Gorde, K.; Jajodia, S.; Park, H.; Kim, Y.: Data synthesis based on generative adversarial networks. In: *Proceedings of the VLDB Endowment* 10/18, pp. 1071–1083, 2018.
- [RMS18] Ram Mohan Rao, P.; Murali Krishna, S.; Siva Kumar, A. P.: Privacy preservation techniques in big data analytics: a survey. In: *Journal of Big Data* 1/18, pp. 1–12, 2018.
- [SBG21] Soni, M.; Barot, Y.; Gomathi, S.: A review on Privacy-Preserving Data Preprocessing. In: *Journal of Cybersecurity and Information Management* 2/21: Special Issue-RIDAPPH, pp. 16–30, 2021.
- [SD18] Salas, J.; Domingo-Ferrer, J.: Some Basics on Privacy Techniques, Anonymization and their Big Data Challenges. In: *Mathematics in Computer Science* 3/18, pp. 263–274.
- [Sh23] Shen, H.; Li, J.; Wu, G.; Zhang, M.: Data release for machine learning via correlated differential privacy. In: *Information Processing & Management* 3/23, pp. 103349, 2023.
- [Sw02a] Sweeney, L.: k-Anonymity: A Model for Protecting Privacy, *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems* 5/02, pp. 557–570, 2002.
- [Wa23] Wakefield, B.: Prediction of Insurance Charges, Kaggle, 2023. URL: <https://www.kaggle.com/datasets/thedevastator/prediction-of-insurance-charges-using-age-gender>, accessed 02/06/2023.
- [WH00] Wirth, R.; Hipp, J.: CRISP-DM: Towards a Standard Process Model for Data Mining. *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, London, pp. 29–39, 2000.

Cybersecurity Privatsphäre - 3.
International Workshop on Digital
Forensics / IWDF3

Analysing Distributions of Feature Similarities in the Context of Digital Anthropometric Pattern Matching Probability

Florian Heinke¹, Marie-Luise Heuschkel¹ and Dirk Labudde¹

Abstract: Digital anthropometric pattern matching encompasses biometric identification on the basis of a combination of anthropometric measurements depicting the proportions of the human body from image or video material. In a previous publication, maximum likelihood density estimation of distributions of anthropometric measurement distances allowed for estimation of the probability of a match to be in the order of 10^{-15} to 10^{-8} . However, the underlying nature and cause of these distributions remained unclear. This work represents an enhancement allowing for an analytical description of these distributions by assuming multivariate normals as distributions models, and by estimating distribution parameters that subsequently allow reasonable probability approximations. Thus the methodological groundwork presented here contributes to the evaluation of the probability for obtaining a match.

Keywords: Biometric features, anthropometric measurements, feature matching probability

1 Introduction and Motivation

The usage of surveillance cameras for the purpose of crime prevention or prosecution today rather is the rule than the exception. Conflicting to this is the fact that many a times the footage can hardly serve as criminal evidence due to perpetrators resorting to disguise their faces. For this reason, biometric research is expanding beyond traditional traits like fingerprints and DNA to include linear body measurements [AA14, ONS10, Da11] such as heights, lengths and widths, that together form anthropometric patterns, as possible passive biometric identifiers. While most of the evidence is anecdotal, a recent study [HHL22] comparing measurement sets of subsamples with individuals taken from a nationwide survey observed measurement dissimilarities to follow distinct distributions that were modelled to be gamma-distributed. Using maximum likelihood estimators and bootstrapping, these attained matching probabilities in the order of 10^{-15} to 10^{-8} . As with that study, research regarding the principles of human body measurements and human proportions is first and foremost empirical since the causes and effects of the various ecological influences cannot be unambiguously measured. Thus, it is not possible to develop first principle models that represent causal relations and closely adhere to reality. Capturing the statistical population, in this regard referring to the whole world population, is a physical impossibility and samples will always be a misrepresentation of some kind. In this context, the mathematical-analytical

¹ University of Applied Sciences Mittweida, Department of Natural and Computer Sciences, Technikumplatz 17, 09648 Mittweida, Germany heinke@hs-mittweida.de

description of distributions as the basis for probability calculations presents a potent approach to meaningfully model frequencies and correlations across samples. Such is the focus of this work that presents enhancements made to prior statistical models and revised probability estimates.

2 Materials & Methods

2.1 Data

The considered dataset consists of 340 measurement sets with 11 measurements per set from individuals in four subgroups taken from a nationwide survey of more than 13,000 subjects that was conducted and provided by Avalution GmbH. The subgroups are characterised by sex and approximate stature (body height): F1 corresponds to 85 females with a measured stature of 163 cm (sd = 0.13 cm, max. deviation: 0.2 cm) and F2 to 85 females at a stature of 173 cm (sd = 0.14 cm, max. dev.: 0.2 cm), while M1 comprises 85 men with a stature of 175 cm (sd = 0.17 cm, max. dev.: 0.3 cm) and M2 85 men at a stature of 185 cm (sd = 0.22 cm, max. dev.: 0.4 cm). Due to the data being identical with the one used in the previous study, the reader is referred to [HHL22] regarding further in-depth details such as measurement features, standards and resolutions.

2.2 Statistical Analyses

Elucidating the study's methodological groundwork requires a more formal treatment of the aforementioned probabilities to begin with: First, let the eleven measurements of some individual i be expressed as a vector $\mathbf{x}_i \in \mathbb{R}^{11}$, then dissimilarity to other measurements, \mathbf{x}_j , is defined as the euclidean distance between both.² Secondly, the vector $\mathbf{z}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ is referred to as the *difference vector* of these two. With $N = 85$ subjects in each data subset, one can obtain a set of 85^2 of such vectors by making an *inter-group comparison* – thus computing all pairwise difference vectors between subjects of two groups – or $85 \cdot (85 - 1)/2$ difference vectors in the case of an *intra-group comparison*. Next, with each measurement feature having a distinct resolution (see [HHL22] for concrete figures), the case of observing two matching measurement vectors \mathbf{x}_i and \mathbf{x}_j is equal to \mathbf{z}_{ij} holding absolute values which are all smaller than given measurement minima dictated by the resolutions. Thus, given one or two subject groups as well as a selection of features, techniques to estimate the probability for obtaining such vectors at random – to which we here refer to as *match probability*, denoted p_m – are introduced next.

² It is these sets of distances, we discussed to follow distinct distributions previously [HHL22] and which we subsequently modeled as gamma-distributed.

One fundamental aspect in statistically describing distance distributions is that squared euclidean distances of latent multivariate data are χ -squared distributed. Here we shall explore this notion. Consider a random variable Y distributed according to an n -dimensional multivariate Gaussian, $Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$, with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance matrix $\boldsymbol{\Lambda}$ with dimensions $n \times n$, then the weighted squared norms (the squared distance to the vector space origin) of N random realizations, which are themselves vectors in \mathbb{R}^n —formally denoted $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_N\}$ —are χ -squared distributed with n degrees of freedom, iff $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$. The diagonal entries are the variances of the n data features. Thus,

$$D_j = \|\mathbf{y}_j\|_{(\lambda)}^2 = \sum_{k=1}^n \frac{1}{\lambda_k} y_{j,k}^2 \sim \chi^2(n) \quad (1)$$

If $\boldsymbol{\mu} \neq \mathbf{0}$, then this relation can be achieved with $\|\boldsymbol{\mu}\|^2$ as the non-centrality parameter of the χ -squared distribution. Next, let $X \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is not necessarily a diagonal covariance matrix, then one can obtain latent linear mappings of realizations of X for which relation (1) holds. This can be achieved through diagonalizing $\boldsymbol{\Sigma}$ by factorization: $\boldsymbol{\Sigma} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^T$, whereas $\mathbf{Y} = \mathbf{E}^T \mathbf{X}$ are the linear mappings and \mathbf{E} is the matrix of n eigenvectors.³

Finally, we extend this relation to difference vectors of n -dimensional multivariate data. Let $Z = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N\}$ be a set of difference vectors as such that $\mathbf{z}_j = (\mathbf{a}_j - \mathbf{b}_j)^T$, $j = \{1, \dots, N\}$, with \mathbf{a}_j and \mathbf{b}_j being realizations of multivariate distributions $\mathcal{N}(\boldsymbol{\mu}_A, \boldsymbol{\Sigma}_A)$ respectively $\mathcal{N}(\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B)$. Then $Z \sim \mathcal{N}(\boldsymbol{\mu}_A - \boldsymbol{\mu}_B, \boldsymbol{\Sigma}_A + \boldsymbol{\Sigma}_B - 2 \cdot \mathbf{Cov}_{A,B})$. $\mathbf{Cov}_{A,B}$ denotes the cross-covariance matrix, which is theoretically expected to be a null matrix when A and B are considered independent. However, it is of practical concern when dealing with empirical data (whereas $\boldsymbol{\Sigma}_A$ and $\boldsymbol{\Sigma}_B$ are to be estimated from the data) and non-zero residual cross-covariance is to be expected and required to be estimated. Hence the here discussed relations between latent squared distances of multivariate data to the χ -squared distribution holds also for sets of difference vectors.

In our previous work we reported the euclidean distances of anthropometric measurement differences, here denoted as the norms of data vectors $\|\mathbf{x}_j\|$, to be well-distributed, upon which we modeled said distances as gamma-distributed as an *ad hoc* assumption [HHL22]. However, through the explanations given in this section, it becomes obvious that reported distances might indeed be 'skewed' χ -squared distributed, if the differences themselves are multivariate normally distributed. Re-scaling the squared distances, here denoted as $\|\mathbf{y}_j\|_{(\lambda)}^2 = \|\mathbf{x}_j\|_{(\lambda)}^2$, results in a natural description of the distance distribution. Furthermore, this allows hypothesis-driven statistical analysis substantiated by underlying theory, superseding our previous assumptions.

³ The process of linearly transforming the (multivariate) data into a latent space, as illustrated here, is strongly related to principal component analysis (PCA). Indeed, the coordinate system describing the latent space are the n principal components of the data \mathbf{X} . As such the latent vectors \mathbf{Y} can be interpreted as scores obtained by PCA, however only if the data X is centered.

Next, we elucidate that modeling the data at hand as multivariate normally distributed is indeed reasonable. Yet, there are various considerations concerning measurement dissimilarity warranting sound probability estimates, e.g. estimating the probability of randomly observing two sets of measurements being so similar that they can be regarded identical, or estimating the probability of observing dissimilarity in some value range of interest. The very basis for deriving such estimates is to formulate statistical models consistent with the data. In the context of this study, such a model is required to:

- be consistent with empirical marginal measurement distributions, notably for each feature,
- reflect correlations/covariances between features,
- be able to generate synthetic measurement data, whose distributions are statistically indistinguishable from the empirical data
- be able to generate data that, when counting matching features within such synthetic data of a subject group or data between groups, should provide match frequencies consistent with the empirical match frequencies,
- allow statistically tractable probability estimations through analytic techniques (f.e. integrating a probability density function) or, at least, by numerical approximations, such as Monte Carlo sampling.

Initial analyses of measurement distributions, made individually for each subject group as well as measurement features, indicated that it is reasonable to model said measurements as multivariate normally distributed. Based on that, for each subject group and intra- and inter-group comparison, feature mean (difference) vectors and covariance matrices were estimated. Therefore, Hamiltonian Monte Carlo sampling (HMC) was applied in order to obtain full posterior estimates [Ge13].⁴ Estimated distribution parameters were subsequently, as a test of consistency, used to generate synthetic data for each subject group. Hereby generated synthetic match frequencies showed to be well-consistent with data. It is thus that not only all of the above requirements are met by our model, it is also straightforward to model the distribution of measurement difference vectors, as discussed in paragraph 2.2, through the additive property of multivariate normals.

⁴ All statistical analyses as well as data processing steps were conducted in R [R 18] and Stan (version 2.31) [Ca17], using the *rstan* and *rethinking* [Mc20] packages as interfaces to Stan. For each (difference) vector set, 4000 distribution parameter samples were generated *a posteriori* by running eight parallel HMC chains utilizing No-U-Turn sampling, each conducting 500 warmup and 500 production iterations ($\delta_{\text{adapt}} = 0.95$, constrained sampling warnings were ignored, as these are to be expected when sampling such strongly constricted model setups). Proper convergence was verified visually through inspecting traceplots and checking chain diagnostics (all \hat{R} were within 0.1% relative error, n_{eff} values were well above 1,500 samples). Cholesky factorization was used in an effort to maintain numeric stability, whereas *a priori* Lewandowski-Kurowicka-Joe's η [LKJ09] was set to 2, aimed to provide a weakly informative, regularizing prior complementing observed mild correlations between measured features.

2.3 Estimating match probabilities

Although modeling the distributions of measurement difference vectors as multivariate models provide already powerful generative means to simulate synthetic data as well as match frequencies, from which meaningful match probability estimates can be inferred, if the sampled frequencies are sufficiently large, such Monte Carlo sampling techniques become numerically inefficient when match probabilities are vanishingly small. On the other hand, computing exact probabilities analytically involves solving the n -dimensional integral of the multivariate normal (where n is the number of considered features) which is known to be intractable [TG15]. To this point, two techniques were used to infer match probabilities in this study.

The first utilizes measurement distance weighting, as described in section 2.2. More precisely, the vector of n feature resolution minima \mathbf{r} gives the maximum weighted distance $D(\mathbf{r})$ at which identity is assumed for two sets of (highly similar) measurements and their respective difference vector. It is that the squared distance of n feature resolution minima—expressed through \mathbf{r} —, conditionally weighted by mean difference vector $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, yields the quantile of the cumulative χ -squared distribution function F_{χ^2} with n degrees of freedom that corresponds to the match probability; denoted $p_m = F_{\chi^2}(0 \leq D_j \leq D(\mathbf{r}) | n, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. Again, this estimate is directly conditional on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Propagating the approximated posterior distributions of both, results in posterior distributions of match probability estimates. This estimation technique was used for intra-group comparisons, where $\boldsymbol{\mu} = \mathbf{0}$ is expected. For inter-group comparisons, with $\boldsymbol{\mu} \neq \mathbf{0}$, the estimates are themselves biased by estimating non-centrality parameters, upon which we devised a second estimator, in order to support first made estimations.

This second technique, to which we refer to as *bounding box estimator*, utilizes a straightforward numeric geometric approximation: First, given the vector of n resolution minima, again denoted by vector \mathbf{r} , one computes the volume of the n -dimensional hyper-rectangle R spanned by \mathbf{r} . The probability of observing a below-resolution difference vector \mathbf{z} at random corresponds to the n -dimensional integral of the multivariate probability function F_N on R , which is equal to randomly obtaining a vector produced by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ located within R . More formally $p_m = p_R = \int_{-r_1}^{+r_1} \dots \int_{-r_n}^{+r_n} F_N(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) dx$. Next, R is mapped into the latent space of principal components, with scores $Y_R = \mathbf{E}^T R$. Following up, the hyper-rectangular bounding box B spanned by the n principal components, which encloses Y_R with minimal volume, is computed. Denoting the n lower and upper side bounds of B as $\{(b_{1,l}, b_{1,u}), \dots, (b_{n,l}, b_{n,u})\}$, one obtains the bounding box volume V_B simply by $V_B = \prod_{k=1}^n |b_{k,l} - b_{k,u}|$ and more importantly, with $\boldsymbol{\Sigma}$ being diagonalized to $\boldsymbol{\Lambda}$ such that $\Lambda_{i \neq j} = 0$, one can compute the probability $p_B = P(b_{k,l} \leq y_k \leq b_{k,u} | k = 1, \dots, n)$ from the product of marginal probabilities with the n bounds of B providing the marginal probability quantiles, thus $p_B = \prod_{k=1}^n P(b_{k,l} \leq y_k \leq b_{k,u} | y_{\mu,k}, \lambda_k)$. Assuming homogeneous probability density in B , which is especially reasonable when $(y_{\mu,k} - b_{k,l}) \approx (y_{\mu,k} - b_{k,u})$ and $|y_{\mu,k} - b_{k,l}| \gg 0$, the estimation of p_m involves scaling p_B through the ratio of volumina: $p_m = p_R \approx \frac{V_R}{V_B} \cdot p_B$. Further refinements of the approximation can be made by correcting for

density heterogeneity. Here we again utilized a Monte Carlo sampling strategy, by uniformly sampling a set of s difference vectors \mathbf{Z} and \mathbf{Y} located in R and B , respectively, followed by deriving a sampled set of s probability estimates

$$P^* = \left\{ p_{m,i} = \frac{F_{\mathcal{N}}(z_i | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \cdot V_R}{F_{\mathcal{N}}(y_i | y_{\boldsymbol{\mu}}, \boldsymbol{\Lambda}) \cdot V_B} \cdot p_B | i = 1, \dots, s \right\}. \quad (2)$$

In Figure 1 the estimated match probability posterior distributions are illustrated for all intra- and inter-group comparisons as 95 % highest posterior density intervals (HPDIs), juxtaposed to the intra-group match probability estimates made by previous *ad hoc* assumptions and sample bootstrapping [HHL22]. HPDIs were estimated here for all 11 features (left interval segments in subfigures a and b) as well as 10 features (right interval segments) by excluding shoulder heights, as the latter showed to be well correlated with stature heights and, thus, greatly lowering inter-group match probability estimates.

2.4 Plausibility Analyses

In order to verify that the modeling approach is consistent with the empirical data (see Section 2.2 for baseline requirements) and that valid match probability estimates can be achieved, inferred multivariate distribution parameters were utilized to generate synthetic measurement data of equal sample size with analogous per-feature resolutions, upon which pairwise synthetic difference vectors were computed. Match counts were subsequently derived. Full posterior inference of match count HPDIs was conducted from posterior distribution parameter samples.

Secondly, posterior match probabilities were used to directly simulate match count frequencies—thus synthetic data-free—utilizing a binomial model. In more detail, given an estimated match probability p_m for some intra- or inter-group comparison and feature combination, whereas the number of underlying corresponding pairwise comparisons shall be denoted as N , then we simulate match counts c from a binomial $c \sim \mathcal{B}(N, p_m)$. Subfigure 1c and Figure 2 illustrate empirical and simulated frequencies for exemplary comparisons with feature combinations of non-zero match count.

3 Results and Discussions

Estimating match probabilities required two strategies presented in section 2.3 in order to trace low-density regions of our distribution models in a numerically robust manner (as low-density estimations are vulnerable to numeric underflow), These estimators need to have underlying distribution models that properly describe the empirical data. In that respect, the four group datasets as well as difference vector sets were modelled as multivariate distributed, whereas corresponding distribution parameter estimations were obtained utilizing HMC.

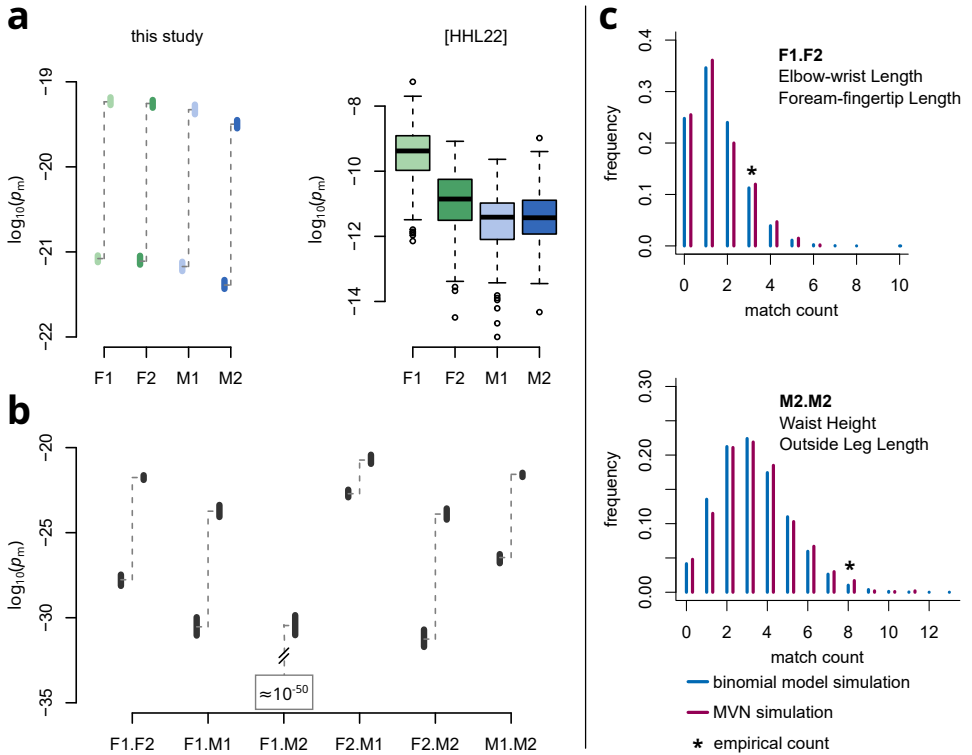


Fig. 1: Results of intra- and inter-group match probability estimations (a and b) and plausibility analyses (c). a: Segment pairs indicate 95% p_m HPDIs for all 11 features (left segment) as well as 10 features excluding shoulder height (right segment). The boxplots to the right show the estimates made by the *ad hoc* model reported in [HHL22], b: Here presented estimation techniques further allow approximating inter-class probabilities (left segments: 11 features, right segments: 10 features excluding shoulder height). c: simulated and observed match counts for two selected group comparisons (F1 vs F2 and M2 vs M2) and two exemplary feature subsets. Simulated match counts were obtained from binomial models using estimated match probabilities and by simulating measurement data from estimated multivariate normal distributions.

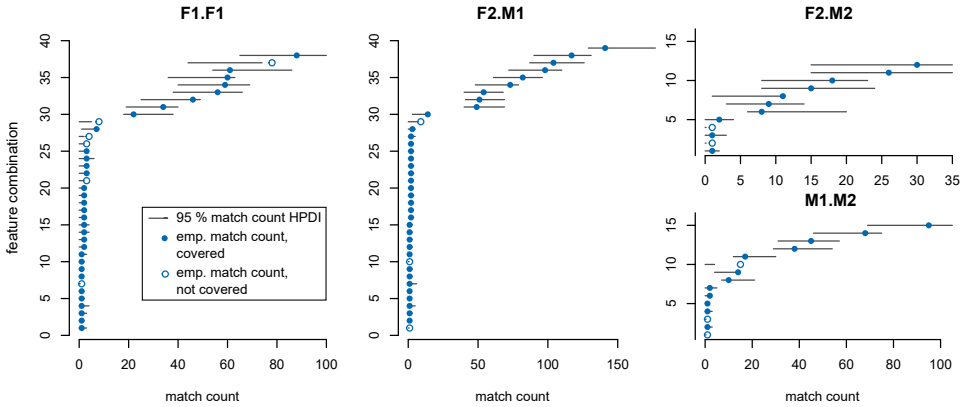


Fig. 2: Number of matches in four selected group comparisons with feature combinations for which non-zero match counts are observed. Blue dots and circles indicate these empirical counts. Grey segments report 95%-HPDIs computed from binomial model simulations. Circles are given for feature combinations where empirical counts lie outside respective HPDIs.

Synthesized measurement data, generated from retrieved parameter estimates, translated to theoretical match counts in agreement with empirical counts (see Figure 1c for exemplary depictions). Approximations of match probabilities were computed and utilized in a binomial model to simulate match counts directly. Similarly, these simulations produced match counts consistent with the data (Figure 2). Note that, by propagating *a posteriori* distribution parameter estimates into match probability estimators, the resulting simulated match counts are by themselves estimates of posterior count distributions. Overall, HPD intervals of match counts covered 84.72 % and 84.69 % of empirical non-zero match counts in intra-comparisons and inter-comparisons, respectively, on micro average.

The here presented distribution models provide a significant update to the assumptions made in [HHL22]. There, measurement dissimilarity densities were assumed to follow gamma-distributions. In contrast, the models presented here are established on knowledge-based assumptions which, as pointed out, are well-consistent with the empirical data and allowed to subsequently formulate match probability estimators beyond intra-group comparisons as discussed in [HHL22]. With respect to those, in Figure 1a the match probability estimates for intra-group comparisons are illustrated. These comparisons refer to the full datasets of all 11 measured features supplemented by estimates for datasets with only 10 features, where *Shoulder height* was excluded. The latter case is of special interest to depict, as shoulder heights showed to be correlated well to the statures (and thus to the subject groups), hence biasing the estimates as a confounder on a factor-like level. As shown, here estimated match probabilities are many orders smaller than reported in our previous study. Excluding shoulder height naturally increases the estimates by about 10^2 . The significant difference to our estimates reported previously is very much likely due to the fact that these were simply based on *ad hoc* dissimilarity distribution assumptions which were, in turn, used to extrapolate into the lower tail of the distributions, without considering the qualitative

connection between the model and the causal structures generating the data. Furthermore, the approach presented here allows to draw probability approximations for inter-group comparisons as well. Not surprisingly, obtained probability ranges are many orders smaller compared to intra-group comparisons, with the comparison of most dissimilar groups, namely F1 and M2, yielding estimates beyond meaningful interpretation. In this case, p_m is approximated to be in the range of $\approx 10^{-50}$. On average, inter-group match probabilities are estimated to fall in the range between 10^{-30} and 10^{-21} . It needs to be highlighted that there is no match between individuals in the dataset w.r.t. to combinations of four or more features. Thus there is no observational evidence that estimated figures are reasonable for cases with multiple features considered. Yet, in cases of feature combinations with observed intra- and inter-group matches, the empirical match counts are comparable to counts that were produced from simulations based on obtained matching probability estimates. This fact could be considered as indication that the novel estimations presented here might be successful extrapolations and, if so, utilized approaches might provide the means to draw generalized statistical conclusions about the applicability of novel biometric identifiers, which is currently work in progress.

4 Conclusion and Outlook

In a previous work we first started to provide more tangible evidence to support the distinctiveness of the anthropometric pattern, confirming what so far had only been indicated or exemplified in the literature [HHL22]. The feasibility of anthropometric pattern matching, as discussed in this work, rests on the premise of distinctiveness along with the other criteria of a passive biometric identifier (universality, permanence or persistence, and measurability). Only if this proof succeeds, can anthropometric pattern matching achieve sufficient conclusiveness. This in turn presents an opportunity to extend the scope of established passive biometric identifiers to include anthropometric rig matching. What is more, it may lead to increased utilisation of surveillance footage depicting the perpetrator at the crime scene as forensic evidence, which has not been fully leveraged thus far. While the study already allows to investigate relevant research questions such as the minimum number of features required to form an anthropometric pattern for matching or the definition of threshold values for determining the level of agreement that constitutes a match or a non-match in the context of samples, it marks a first step towards statistical generalisations to that affect that would facilitate assertions about the potential for misclassification [RSV19, JR08].

Bibliography

- [AA14] Andersson, V. O.; Araújo, R. M.: Full Body Person Identification Using the Kinect Sensor. In: 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, Lymassol, Cyprus. pp. 627–633, 2014.
- [Ca17] Carpenter, B.; Gelman, A.; Hoffman, M. D.; Lee, D.; Goodrich, B.; Betancourt, M.; Brubaker, M.; Guo, J.; Li, P.; Riddell, A.: Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 2017.

- [Da11] Dantcheva, A.; Velardo, C.; D'Angelo, A.; Dugelay, J.-L.: Bag of soft biometrics for person identification. *Multimedia Tools and Applications*, 51(2):739–777, 2011.
- [Ge13] Gelman, A.; Carlin, J. B.; Stern, H. S.; Dunson, D. B.; Vehtari, A.; Rubin, D. B.: *Bayesian data analysis*, third edition. 2013.
- [HHL22] Heinke, F.; Heuschkel, M.-L.; Labudde, D.: A frequentist estimation of duplicate probability as a baseline for person identification from image and video material using anthropometric measurements. *INFORMATIK 2022*, 2022.
- [JR08] Jain, A.K.; Ross, A.: *Introduction to Biometrics*. In (Jain, Anil K.; Flynn, Patrick; Ross, Arun A., eds): *Handbook of biometrics*. Springer International Publishing Cham, Switzerland, 2008.
- [LJK09] Lewandowski, D.; Kurowicka, D.; Joe, H.: Generating random correlation matrices based on vines and extended onion method. *J. Multivar. Anal.*, 100:1989–2001, 2009.
- [Mc20] McElreath, R.: *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press, 2020.
- [ONS10] Ober, D. B.; Neugebauer, S. P.; Sallee, P. A.: Training and feature-reduction techniques for human identification using anthropometry. In: *2010 4th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)* Washington, DC, USA. IEEE, pp. 1–8, 2010.
- [R 18] R Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [RSV19] Ramalingam, S.; Shenoy, A.; Viet, N. T.: Fundamentals and Advances in 3D Face Recognition. In (Obaidat, M. S.; Traore, I.; Woungang, I., eds): *Biometric-Based Physical and Cybersecurity Systems*, pp. 125–162. Springer International Publishing Cham, Switzerland, 2019.
- [TG15] Trinh, G.; Genz, A.: Bivariate conditioning approximations for multivariate normal probabilities. *Statistics and Computing*, 25:989–996, 2015.

Vorbetrachtung zur Entwicklung einer Analysesoftware zur Auswertung von videografierten Beschuldigtenvernehmungen im Strafermittlungsverfahren

Stefan Kellermann ¹ und Dirk Labudde ²

Abstract: Die videographische Aufzeichnung von (Beschuldigten-) Vernehmungen gewinnt in Deutschland zunehmend an Bedeutung für das Strafermittlungsverfahren. Ziel ist anhand der Aufzeichnungen das normgerechte Zustandekommen der Aussagen der Befragten für alle Verfahrensbeteiligten nachvollziehen zu können. In den letzten Jahren hat der Gesetzgeber die gesetzlichen Vorschriften zur verpflichtenden Aufzeichnung von Beschuldigtenvernehmungen in bestimmten Fallkonstellationen erweitert. Mithilfe der Bild-/Tonaufnahmen ergeben sich umfangreiche Möglichkeiten zur qualitativen Analyse. Eine technische Lösung zur Umsetzung der Analyse könnte, neben der Überprüfung der Einhaltung rechtlicher Vorgaben, als Evaluationsinstrument für das noch nicht abgeschlossene Ermittlungsverfahren sowie im Rahmen der Aus- und Fortbildung von Befragenden unterstützend Anwendung finden. Es wird die konzeptionelle Entwicklung einer Software vorgestellt, die anhand einer Kombination aus linguistischer und Bildanalyse Videoaufzeichnungen auswertet.

Keywords: Vernehmung, Videographie, Bilderkennung


1 Forschungsanliegen und Motivation

„Die Vernehmung ist eine der wichtigsten und am häufigsten gebrauchten kriminaltaktischen Methoden zur Untersuchung von Straftaten, kriminalistisch relevanten Sachverhalten sowie zu Ermittlungen bei Ordnungswidrigkeiten.“³ Die Vernehmung des Beschuldigten nimmt im polizeilichen Ermittlungsverfahren eine bedeutende Rolle ein. Der/die Beschuldigte steht in der Regel im Mittelpunkt des zu klärenden strafrechtlich relevanten Ereignisses, so dass zur Aufklärung oftmals auf ihre/sein Ereigniswissen kaum verzichtet werden kann.⁴ Die Vernehmung des/der Beschuldigten dient jedoch nicht nur der Aufklärung des kriminalistisch relevanten Ereignisses, sondern gibt dem/der Beschuldigten auch die Möglichkeit, sich zu den vorgebrachten Anschuldigungen

¹ Universität Greifswald, Ernst-Lohmeyer-Platz 1, Greifswald, 17489, ste-

fan.kellermann@stud.uni-greifswald.de,  <https://orcid.org/0009-0002-0536-0673>

² Hochschule Mittweida, Technikumplatz 17, Mittweida, 09648, labudde@hs-mittweida.de,

 <https://orcid.org/0000-0003-0466-0017>

³ Ackermann, R.; Clages, H.; Roll, H.: Handbuch der Kriminalistik. Kriminaltaktik für Praxis und Ausbildung, 6. Aufl, Boorberg, Stuttgart, S. 614, 2022.

⁴ Mohr, M.; Schimpel F.; Schröer, N.: Die Beschuldigtenvernehmung. Lehr- und Studienbriefe Kriminalistik / Kriminologie 5, 1. Aufl, Verl. Deutsche Polizeiliteratur, Hilden/Rhld, S. 5, 2006.

gen zu verteidigen. Die Bemühungen der Beamt*innen um die Ermittlung des tatsächlichen Sachverhaltes und um die gerichtsverwertbare Aufzeichnung der Beschuldigtenvernehmung in einem Protokoll, stehen dem Interesse des/der Beschuldigten auf eine wirksame Verteidigung – tendenziell – entgegen.⁵ Die Umstände der Beschuldigtenvernehmung und somit das rechtmäßige Zustandekommen von Aussagen des/der Beschuldigten werden in Gerichtsprozessen regelmäßig in Frage gestellt.

Der Gesetzgeber hat das Strafprozessrecht in Bezug auf die Vernehmung von Beschuldigten in den letzten Jahren beständig fortentwickelt.⁶ Die in bestimmten Fallkonstellationen verpflichtende technische Aufzeichnung von Vernehmungen hielt verstärkt Einzug in die StPO. Das aktuell bestehende Gesetzgebungsverfahren, Gerichtsverhandlungen technisch aufzuzeichnen zeigt, dass der Anspruch an Transparenz zur Nachvollziehbarkeit von Aussagen über die innerpolizeiliche Anwendung hinaus an Bedeutung gewonnen hat.⁷ Der aktuelle Koalitionsvertrag der Regierungsparteien sieht zudem eine grundsätzlich verpflichtende Aufnahme von Vernehmungen in Bild- und Ton vor.⁸

Das Zustandekommen der Aussage des/der Beschuldigten ist von großem Interesse, da bei der Anwendung von verbotenen Vernehmungsmethoden gem. § 136a Abs. 1, 2 StPO zwingend gem. § 136a Abs. 3 StPO Beweisverwertungsverbote drohen. Zudem können infolge ausgeübten psychologischen Druckes der Vernehmungspersonen im Rahmen der Beschuldigtenvernehmung erlangte falsche Geständnisse generiert werden, welche darüber hinaus ihrerseits zu Fehlurteilen führen können.⁹

2 Aufbau und Ablauf einer Vernehmung

Weltweit bestehen unterschiedliche Vernehmungsmodelle und Herangehensweisen zur Durchführung einer Vernehmung.

⁵ Ebd., S. 15.

⁶ Gesetz zur effektiveren und praxistauglicheren Ausgestaltung des Strafverfahrens – Bundesgesetzblatt Jahrgang 2017 Teil I Nr. 58, Bonn, Stand: 23.08.2017;

Gesetz zur Stärkung der Verfahrensrechte von Beschuldigten im Jugendstrafverfahren - Bundesgesetzblatt Jahrgang 2019 Teil I Nr. 47, Bonn, Stand: 16.12.2019;

Gesetz zur Modernisierung des Strafverfahrens – Bundesgesetzblatt Jahrgang 2019 Teil I Nr. 46, Bonn, Stand: 12.12.2019;

Gesetz zur Neuregelung des Rechts der notwendigen Verteidigung - Bundesgesetzblatt Jahrgang 2019 Teil I Nr. 46, Bonn, Stand: 12.12.2019;

Gesetz zur Fortentwicklung der Strafprozessordnung und zur Änderung weiterer Vorschriften - Bundesgesetzblatt Jahrgang 2021 Teil I Nr. 37, Bonn, 30.06.2021.

⁷ Gesetzentwurf der Bundesregierung: Entwurf des Gesetzes zur digitalen Dokumentation der strafgerichtlichen Hauptverhandlung, Stand: 10.05.2023.

⁸ SPD, Bündnis 90/Die Grünen u. FDP: Mehr Fortschritt wagen. Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit. Koalitionsvertrag 2021, S. 106, 2021.

⁹ Siehe hierzu ergänzend: Deutscher Bundestag: Entwurf eines Gesetzes zur effektiveren und praxistauglicheren Ausgestaltung des Strafverfahrens, BT-Drs. 18/11277, S. 24ff, 2017.

Die Méndez-Prinzipien¹⁰ wurden auf UN-Ebene international von verschiedenen Staaten mit Anerkennung zur Kenntnis genommen.¹¹ Das PEACE-Vernehmungsmodell erfüllt die Anforderungen der Méndez-Prinzipien.¹² Das norwegische CREATIVE-Modell¹³ kann als Leitfaden zur Umsetzung des PEACE-Modells angesehen werden. Im Kern beschreibt es eine dem Menschen zugewandte, nicht vorverurteilende, empathische, auf Vertrauen basierende Kommunikation, unter Einhaltung rechtlicher Vorgaben.¹⁴ Der phasenhafte Verlauf des PEACE-Modells wird kurz dargestellt.

1. Planning and preparation (fallbezogen, mentale, physische Vorbereitung)
2. Engage and explain (Aufzeichnung, Erklären, Belehrung, Vernehmungsgrund- und Ablauf)
3. Account (Freie Schilderung, Aktives Zuhören durch Vernehmenden, anschließend Klärung offener Fragen, strategische Beweispräsentation)
4. Closure (Abschluss der Vernehmung)
5. Evaluation (Auswertung)

Das Modell zielt darauf ab, durch einen professionellen Beziehungsaufbau, transparente Kommunikation und Empathie zuverlässige Aussagen über relevante Inhalte zu generieren. Gleichzeitig wird das Risiko durch Falschaussagen aufgrund eines nicht-geständnisorientierten Vernehmungsstils minimiert. Das PEACE-Modell wurde bereits wissenschaftlich evaluiert, weiterentwickelt und wird in seiner Umsetzung in Deutschland auch als “Untersuchende Vernehmungstechnik” bezeichnet.¹⁵

In Deutschland wird derzeit in der Polizei kein einheitliches Vernehmungsmodell angewandt. Die im Abschnitt 5.1 ebenfalls zum Bezug genommene Studie “BEST- Befragungsstandards für Deutschland” hat im Rahmen einer Vollerhebung von Curricula der Polizeien und ergänzenden Befragungen von Lehrenden festgestellt, dass mehrheitlich das PEACE-Modell, das Kognitive Interview oder die Strukturierte Vernehmung, neben weiteren Methoden, wie 3-K-Modell oder die Sondierungsmethode, unterrichtet werden. In über 80 % der untersuchten polizeilichen Fortbildungsangebote werden, neben der theoretischen Wissensvermittlung, auch praktische Lehrformen, wie Übungen oder Rollenspiele als Lehrform angewandt.¹⁶

¹⁰ Association for the prevention of torture u.a.: Principles on Effective Interviewing for Investigations and Information Gathering. 2021.

¹¹ United Nations: Torture and other cruel, inhuman or degrading treatment or punishment. S. 5, 2022, https://digitallibrary.un.org/record/3993571/files/A_C.3_77_L.45-EN.pdf?ln=en, Stand: 30.05.2023.

¹² Bull, R.; Milne, R.: Psychologie der Vernehmung. Huber, Bern, 2003.

¹³ CREATIVE ist ein Akronym. Aus Gründen des Umfangs wird weiterführend auf: Rachlew, A.; Loken, G.; Bergestuen, S.: A Guide to the Professional Interview. Anthem Press, S.5f, 2022 verwiesen.

¹⁴ Rachlew, A; Loken, G.; Bergestuen, S.: A Guide to the Professional Interview. Anthem Press, 2022.

¹⁵ Convention against Torture Initiative (CTI): Die untersuchende Vernehmungstechnik. S. 3f, 2017.

¹⁶ Stegemann, L; Kluth, L.: Vernehmungsförderung der Polizei im Vergleich mit außerbehördlichen Angeboten. Kriminalistik 12/2021, S. 681, 2021.

3 Informationsquellen für die Analyse

Die vorliegende Arbeit beschäftigt sich mit der Konzeptionierung einer Software zur inhaltlichen Analyse von Videoaufzeichnungen von Beschuldigtenvernehmungen unter qualitativen und quantitativen Aspekten befassen.

Hierbei sollen Aspekte berücksichtigt werden, die die Qualität der durchgeführten Vernehmung, sowohl aus der Perspektive der Vernehmenden als auch aus der Perspektive der Vernommenen, unter Berücksichtigung rechtlicher als auch fachlicher Gesichtspunkte, automatisiert einschätzt und somit unterstützt. Ausdrücklich soll die Inaugenscheinnahme durch die Prozessbeteiligten nicht ersetzt werden. Das Ziel der Lügendetektion wird ebenfalls nicht verfolgt, da dies unter bisher bekannten wissenschaftlichen Aspekten nicht sicher möglich ist.¹⁷ Denkbar wäre, dass die Ergebnisse der qualitativen Auswertung bei der Evaluierung von Befragungsmethoden durch Strafverfolgungsbehörden im Ermittlungsverfahren, zur Aus- und Fortbildung oder aber auch bei der rechtlichen Prüfung durch Verfahrensbeteiligte herangezogen werden könnte.

Zunächst müssen jedoch Merkmale herausgearbeitet werden, die mithilfe einer Software erkennbar gemacht werden können und zielführende Informationen liefern.

Als Quellen können einerseits die Tonspuren und hiermit verknüpfte linguistische Analysen und andererseits die Bilddaten, wiederum verknüpft mit Bilderkennungsmustern, dienen. Die wörtliche Transkription von Bild-/Tonaufnahmen wäre für eine einfache erste Statistik und weitere Analysen von Vorteil. Veränderungen der Tonlage der beteiligten Sprecher*innen sichtbar zu machen, lässt sich dagegen nur über Methoden der Stimmprofileanalyse nachweisen.

Die technische Feststellung der Verteilung des Sprechanteils der Beteiligten, die Auswertung von Sprechpausen sowie eine linguistische Analyse, die eine Zuordnung zur Frageart, z.B. offen/geschlossen/suggestiv ermöglicht, könnten Anhaltspunkte sein, um die Qualität der Vernehmung zu beurteilen.

4 Die Vernehmung als Kommunikationsmodell

Kommunikation ist weit mehr als nur verbal, linguistisch kanalisiert. Die Abbildung 1 zeigt eine mögliche Einteilung von Anteilen während einer Kommunikation. Im Zusammenhang mit Vernehmungen sind der verbale und Teile der non-verbalen Kommunikation von entscheidender Bedeutung. Die visuelle Kommunikation (als Teil der non-verbalen Kommunikation), mit Körperbewegung und Körperhaltung, kann wichtige

¹⁷ May, L.; Schneider, T.; Okulicz-Kozaryn, M.: Notwendigkeit einer kritischen Prüfung von Methoden zur Lügenerkennung und Vernehmungstaktiken zur Informationsgewinnung von Beschuldigten: ein Positionspapier aus Psychologie und Polizei. *Forensische Psychiatrie, Psychologie, Kriminologie*, 2022, S. 138-143, Springer, 2022.

Informationen liefern, um zum Beispiel ein Unbehagen oder Drucksituationen beim Vernommenen festzustellen. Hinter jeder Kommunikation steht immer auch ein Kommunikationsmodell. In der Literatur gibt es eine Reihe von Kommunikationsmodellen, beispielhaft seien hier das Vier-Seiten-Modell nach Schulz von Thun¹⁸, Kommunikationsmodell nach Watzlawick¹⁹ bzw. Berlo genannt.²⁰

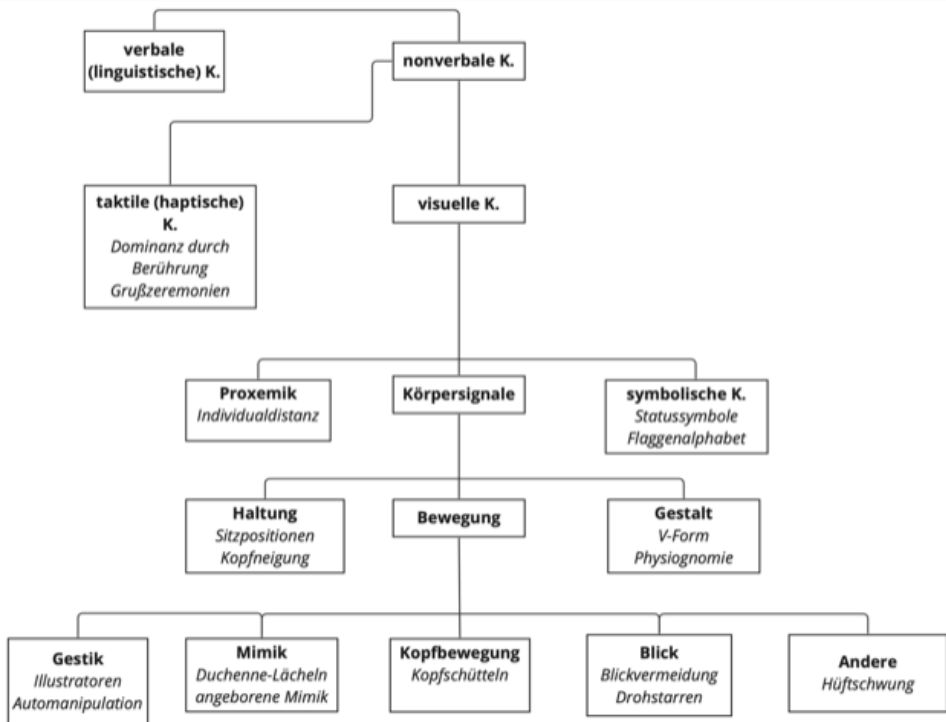


Abb. 1: Kommunikation - Kommunikation setzt sich aus verschiedenen Anteilen zusammen. Verbale, Non-verbale Kommunikation sowie paraverbal). Die non-verbale Kommunikation beinhaltet die visuelle, taktile und chemische Kommunikation.

Eine Vernehmung lässt sich durch das Kommunikationsmodell von Berlo recht gut beschreiben (Abb. 2). David Berlo stellte 1960 seine Kommunikationstheorie auf. Diese ist auch als SMCR-Modell bekannt, weil sie aus vier Komponenten besteht: Quelle

¹⁸ Schulz von Thun, F.: Miteinander Reden. Band 1: Störungen und Klärungen, 48. Auflage, Rowohlt Taschenbuch Verlag, Reinbek bei Hamburg, 2010.

¹⁹ Watzlawick, P.; Beavin J.; Jackson, J: Menschliche Kommunikation. Formen, Störungen, Paradoxien, 13. Auflage, Hogrefe AG, Bern u.a., 2016.

(source), Nachricht (message), Kanal (channel) und Empfänger (receiver).²¹

Vernehmungen sind sowohl aus der Sicht des Befragten, aber auch zum Teil aus der Sicht des Fragenden eine Stresssituation. Durch Gefühle zeigen beiden Kommunikationspartner ihre Sichtweise auf den Inhalt und signalisieren dem Anderen, wie wichtig das angesprochene Thema ist. Aus der Literatur ist bekannt, dass Emotionen in Videos detektierbar und klassifizierbar sind.²²

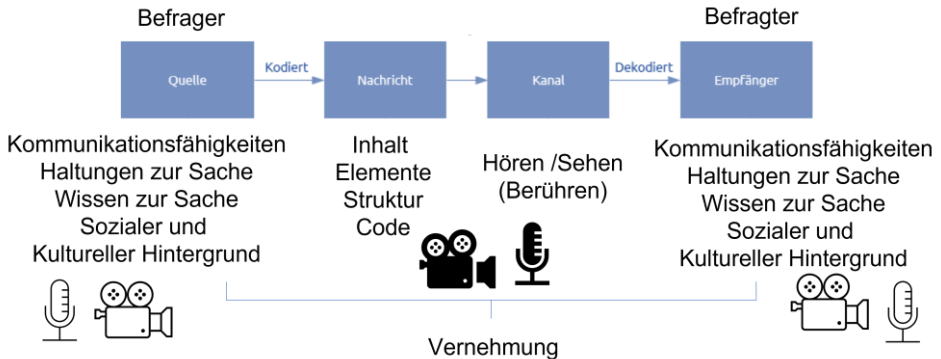


Abb. 2: Kommunikationsmodell für eine Vernehmung – Basierend auf dem Kommunikationsmodell von Berlo können die 4 Komponenten benannt werden. Die Nachricht setzt sich aus den Bausteinen Inhalt, Struktur, Elementen und Code zusammen. Die Kanäle lassen sich als Hören und Sehen sowie berühren (in besonderen Fällen) definieren. Der Befragter und der Befragte verfügen über individuelle Kommunikationsfähigkeiten, wie Haltung und Wissen zur Sache, sowie soziale und kulturelle Hintergründe.

Einige Informationen sind direkt aus den Aufzeichnungen ableitbar (Abb. 2 schwarze Symbole), z.B. gab es Berührungen oder die verbale Kommunikation. Jedoch müssen andere Informationen, wie der non-verbale und para-verbale Inhalt abgeleitet und interpretiert werden (weiße Symbole). Würde man die Kommunikation als ein Muster verstehen, besteht die Möglichkeit nach Anomalien zu suchen. Orientieren kann man sich an den Arbeiten von Paul Ekman zum FACS (Kodierungssystem des Gesichtes).²³ Als

²¹ Berlo, D.: The Process of Communication. An Introduction to Theory and Practice, Holt, Rinhart and Winston, New York, 1960.

²² Fiehler, R.: Kommunikation und Emotion. Theoretische und empirische Untersuchungen zur Rolle von Emotionen in der verbalen Interaktion, Walter de Gruyter, Berlin u.a., 1990;

Bühler, K.: Sprachtheorie. Die Darstellungsfunktion der Sprache. 3.Auflage, UTB, Stuttgart, 1999.

²³ Tian Y.; Kanade, T; Cohn, J.: Recognizing Action Units for Facial Expression Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.23, no.2, 2001,

<http://www.cs.cmu.edu/~face/Papers/112006-1.pdf>, Stand: 01.06.2023;

Bartlett, M.; Hager J; Ekman, P. ;Sejnowski, T.: "Measuring Facial Expressions by Computer Image Analysis, Psychophysiology, vol. 36, 1999, S. 253-264, 1999;

Donato, G; Bartlett, M.; Hager J.; Ekman P.; Sejnowski T.:Classifying Facial Actions. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 10, 1999, S. 974-989, 1999;

Ekman, P.: Facial Expression and Emotion. American Psychologist, vol. 48, 1993, S. 384-392, 1993;

Ergebnis erhält man ein analog zum Gesicht (FACS) ein De-Kodierungssystem für eine Vernehmung. Dieses System muss sich an unterschiedlichen Komponenten orientieren, welche sich aus Kategorien der Sprache und Verhalten zusammensetzen.

5 Stand in Deutschland und International

5.1 Deutschland

Neben dem Vernehmungsmodell könnte ein Anknüpfungspunkt zur Erstellung eines Kodierungssystems für eine entsprechende Analysesoftware das von 2019 bis 2021 unter anderem an der Fachhochschule für öffentliche Verwaltung, Polizei und Rechtspflege Mecklenburg-Vorpommern (FHöVPR MV) durchgeführte Forschungsprojekt „BEST: Befragungsstandards für Deutschland“ bilden. BEST sollte Hinweise für Standards in Vernehmungen im behördlichen sowie unternehmerischen Kontext erarbeiten. Es wurden verschiedene Befragungsmethoden untersucht und häufig auftretende Fehler analysiert. Die Datenerhebung erfolgte anhand einer bundesweiten Auswertung von Curricula zur Vernehmungs-/Befragungsausbildung, Befragungen von Vertretern von Polizeibehörden und Unternehmen sowie Strafgefangenen, teilnehmender Beobachtung an Beschuldigtenvernehmungen sowie der Auswertung von audiovisuellen Aufzeichnungen von simulierten Vernehmungen. An dem Forschungsprojekt waren verschiedene Akteure und assoziierte Partner beteiligt. Die FHöVPR MV hatte dabei einen eigenen Forschungsauftrag innerhalb des Projektes.²⁴

Kernergebnisse des Forschungsprojektes sind u.a. (siehe Abschnitt 2):²⁵

- Vernehmungslehre ist in allen polizeilichen Erstausbildungen enthalten
- Umfang, Art und Inhalt der Lehre sehr heterogen
- der zeitliche Umfang wird durch die befragten Lehrenden insgesamt als zu gering eingeschätzt
- Fortbildungsangebote sind nicht in allen Bundesländern vorhanden

Im Rahmen der Analyse der simulierten Befragungen wurde eine umfangreiche Kodierung erstellt. Die technische Operationalisierung der verwendeten Kodierung könnte

Ekman, P.; Friesen, W.: Pictures of Facial Affect, Consulting Psychologists Press, Palo Alto (CA.), 1976;
Ekman, P.; Friesen, W.: The Facial Action Coding System: A Technique for The Measurement of Facial Movement, Consulting Psychologists Press, San Francisco, 1978.

²⁴ FHöVPR MV: BEST - Befragungsstandards für Deutschland – Teilvorhaben „Analyse der polizeilichen Aus- und Fortbildungssituation und Auswertung polizeilicher Vernehmungen“. <http://www.fh-guestrow.de/forschung/abgeschlossene-forschungsprojekte/best-befragungsstandards-fuer-deutschland>, Stand: 30.05.2023.

²⁵ Präsentation der Abschlusstagung (online) des Forschungsprojektes BEST vom 16.11.2021, Folie 35, 2021.

Ausgangspunkt bei der Entwicklung der Software sein. Zudem könnten die durch May, Fahsing und Milne gebildeten Kategorien zur hierarchischen Einordnung von informativ-erhellenden Äußerungen bei der untersuchenden Vernehmung genutzt werden.²⁶

5.2 Wissenschaftliche Anknüpfungspunkte international

Im kanadischen Forschungsprojekt PRESEL „Projet de recherche sur l’exploitation sexuelle des enfants en ligne“²⁷, wurden im Zeitraum von 2018 bis 2020 Beschuldigtenvernehmungen von Sexualstraftätern untersucht. Im Rahmen der Untersuchung wurde ebenfalls eine Kodierung entwickelt, die mit dem hiesigen Kenntnisstand abgeglichen und für die Erstellung einer Software beigezogen werden könnte.

Das Training Tool der Convention against Torture Initiative (CTI) aus dem Jahre 2017 wird bereits im Regierungsprogramm Österreichs 2020²⁸ aufgegriffen und soll dort als Basis zur Prüfung auf Bestehen von vertrauensbasierten Vernehmungsmethoden herangezogen werden. Dies könnte hier ebenso erfolgen.

Ferner beschreiben die Méndez-Prinzipien zu „Effektiven Vernehmungen in Ermittlungen und Informationssammlungen“ grundsätzliche Standards für Vernehmungen aus Sicht der association for prevention of torture.²⁹ Im Rahmen der 77. Sitzung der Generalversammlung der UN, hat auch die Bundesrepublik Deutschland die Méndez-Prinzipien mit Anerkennung zur Kenntnis genommen.³⁰ Eine zu entwickelnde Software könnte die Ziele der Méndez-Prinzipien forcieren, indem die technisch automatisierte Erkennung zur Einhaltung wissenschaftlicher Standards, die Aufdeckung von Verstößen gegen verbotene Vernehmungsmethoden verbessert sowie die Auswertung zu Ermittlungs- und Trainingszwecken unterstützt werden.

²⁶ May, L.; Fahsing, I.; Milne, B.: Die untersuchende Vernehmung von Zeug*innen und Beschuldigten – ein internationaler forschungsbasierter und praxiserprobter Vernehmungsansatz. In: Handbuch Polizeipsychologie, Springer Gabler, Wiesbaden, S. 717ff., 2023.

²⁷ Paquette, S., Fortin, F.; Bergeron, A.; Deslauriers-Varin, N.: Online Child Sexual Exploitation & Solicitation: The PRESEL Projekt.
https://www.researchgate.net/profile/Sarah-Paquette/publication/338434963_Online_Child_Sexual_Exploitation_Solicitation_The_PRESEL_Project/links/5e14d4c44585159aa4bcd916/Online-Child-Sexual-Exploitation-Solicitation-The-PRESEL-Project.pdf, Stand: 30.05.2023.

²⁸ Republik Österreich: Aus Verantwortung für Österreich. Regierungsprogramm 2020–2024, Wien, S. 213, 2020.

²⁹ association for the prevention of torture: Principles on Effective Interviewing for Investigations and Information Gathering, Genf, 2021.

³⁰ United Nations: Torture and other cruel, inhuman or degrading treatment or punishment. S. 5, 2022, https://digitallibrary.un.org/record/3993571/files/A_C.3_77_L.45-EN.pdf?ln=en, Stand: 30.05.2023.

6 Technische Rahmenanforderungen einer Analysesoftware für Vernehmungen

Welche technischen Anforderungen stellt ein De-Kodierungssystem einer Vernehmung an eine Software? Sicher ist es möglich ein intelligentes System zur Analyse der Emotionen zu verwenden. Erste Ansätze sind exemplarisch in der Arbeit von Michael et.al.³¹ dargestellt. Jedoch bleibt die Suche nach der „baseline“ eines Menschen die größte Herausforderung. Als „baseline“ wird das normale Erscheinungsbild eines Menschen bezeichnet, welches einen neutralen Emotionsausdruck besitzt. Diese „baseline“ muss auf Sprache und Verhalten erweitert werden.

Das Konzept sollte aus den folgenden Modulen (Abb. 3) aufgebaut sein:

- Data Akquise Modul: Erhebung der Daten (Protokolle, Videoaufzeichnungen) während einer Vernehmung
- Prozessierung der Daten Modul: Strukturierung der Daten und Überführung in ein verarbeitbares zeitlich homogenes Format
- Data Analyse Modul: Kodierung der Daten – Abgleich Verhalten, gesprochene Wort, Tonfall, Gesichtsausdruck – Suche nach Widersprüchen (Anomalieerkennung)
- Bewertung der Informationen: Bewertung durch eine Person

Eine derartige Analysesoftware erfordert ein gewisses Maß an Standardisierung und einheitliche Anforderungen an die technische Umsetzung. So kann gewährleistet werden, dass die Aufnahmen Mindestanforderungen bezogen auf die Qualität der Aufnahmen erfüllen. Die technische Umsetzung muss Faktoren, wie Kameraperspektive, Ausleuchtung, Sitzanordnungen und die Gestaltung des Raumes beinhalten und optimieren. In die Betrachtungen müssen auch akustische Fragen einbezogen werden, um verwertbares Aufzeichnungsmaterial zu erhalten. An dieser Stelle sei darauf hingewiesen, dass für die Standards technische Versuche erforderlich sind und in Empfehlungen ausgesprochen werden müssen.

³¹ Michael, V.; Klöden, M; Labudde, D.: KI-basierte Emotionserkennung in Videos. Polizei-Informatik 2020, RediromaVerlag, S 56- 74, 2020.

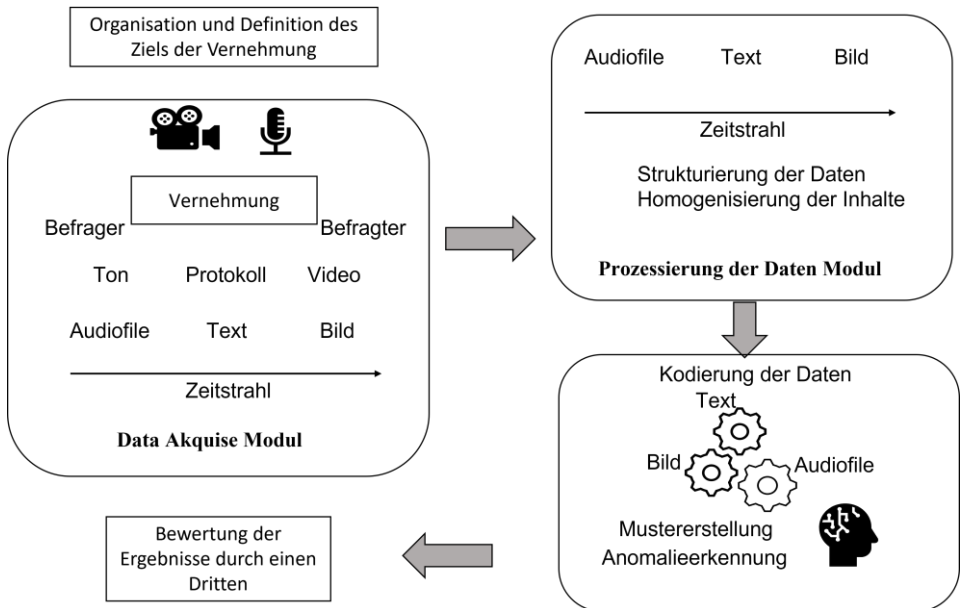


Abb. 3: Module der Analysesoftware – Der Gesamtprozess startet mit der Organisation und Zielsetzung der Vernehmung. Die Datenakquise stellt die direkte Aufnahme (audio und video) und das Protokoll als Quellen dar. Durch die Prozessierung werden die Daten strukturiert und in chronologisch verarbeitbare Daten überführt. Die Suche nach Muster auf der Grundlage der definierten Kodierung analog einer Anomalie Suche liefert die Informationen für die Bewertung.

7 Aussichten der Umsetzung

Dass die Masse an auswertbaren Vernehmungsvideos in Deutschland zunehmen wird, lässt sich an den aufgezeigten Gesetzesänderungen und -initiativen sowie anhand des bestehenden Koalitionsvertrages der Bundesregierung prognostizieren. Die gesetzlichen Verpflichtungen zur Aufnahme werden auch die Frage nach der Auswertung der Videos in der Anzahl steigen lassen. Eine hierbei unterstützende technische Lösung könnte die zur Auswertung notwendigen personellen Ressourcen schonen. Weiterhin kann die Auswertung zu Ermittlungszwecken erleichtert werden. Gleichzeitig wirkt eine technische Auswertung grundsätzlich unterstützend bei der Evaluation der Vernehmungsqualität. Darüber hinaus könnte eine derart ausgestaltete Software zu Aus- und Fortbildungszwecken sowie für wissenschaftliche Analysen eingesetzt werden. Das Ziel, die Auswertung und fortlaufende Evaluation von Vernehmungsprozessen anhand wissenschaftlicher Erkenntnisse zu forcieren, steht zudem im Einklang mit den Méndez-Prinzipien.

Bibliography

- [ACR22] Ackermann, R.; Clages, H.; Roll, H: Handbuch der Kriminalistik. Kriminaltaktik für Praxis und Ausbildung, 6. Aufl., Boorberg, Stuttgart, 2022.
- [AP21] AP, association for the prevention of torture: Principles on Effective Interviewing for Investigations and Information Gathering. Genf, 2021.
- [BHE99] Bartlett, M.; Hager J; Ekman, P. ;Sejnowski, T.: Measuring Facial Expressions by Computer Image Analysis. *Psychophysiology*, vol. 36, S. 253-264, 1999.
- [BM03] Bull, R.; Milne, R.: *Psychologie der Vernehmung*. Huber, Bern, 2003.
- [Bü99] Bühler, K.: *Sprachtheorie. Die Darstellungsfunktion der Sprache*, 3.Auflage, UTB, Stuttgart, 1999.
- [CT17] CT, Convention against Torture Initiative (CTI): *Die untersuchende Vernehmungstechnik*, 2017.
- [DBH99] Donato, G; Bartlett, M.; Hager J.; Ekman P.; Sejnowski T.:Classifying Facial Actions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, S. 974-989, 1999.
- [EF76] Ekman, P.; Friesen, W.: *Pictures of Facial Affect*. Consulting Psychologists Press, Palo Alto (CA.), 1976.
- [EF78] Ekman, P.; Friesen, W.: *The Facial Action Coding System: A Technique for The Measurement of Facial Movement*. Consulting Psychologists Press, San Francisco, 1978.
- [Ek93] Ekman, P.: Facial Expression and Emotion. *American Psychologist*, vol. 48, S. 384-392, 1993.
- [En17] Entwurf eines Gesetzes zur effektiveren und praxistauglicheren Ausgestaltung des Strafverfahrens, BT-Drs.18/11277, Bundestag, 2017.
- [En23] Entwurf des Gesetzes zur digitalen Dokumentation der strafgerichtlichen Hauptverhandlung, Bundesregierung, 2023.
- [FH23] FH, FHöVPR MV: BEST - Befragungsstandards für Deutschland – Teilvorhaben „Analyse der polizeilichen Aus- und Fortbildungssituation und Auswertung polizeilicher Vernehmungen“, <http://www.fh-guestrow.de/forschung/abgeschlossene-forschungsprojekte/best-befragungsstandards-fuer-deutschland>, Stand: 30.05.2023.
- [Fi90] Fiehler, R.: *Kommunikation und Emotion Theoretische und empirische Untersuchungen zur Rolle von Emotionen in der verbalen Interaktion*, Walter de Gruyter, Berlin u.a., 1990.
- [Ge17] Gesetz zur effektiveren und praxistauglicheren Ausgestaltung des Strafverfahrens, Bundesgesetzblatt, Bonn, 2017.
- [Ge19a] Gesetz zur Stärkung der Verfahrensrechte von Beschuldigten im Jugendstrafverfahren Bundesgesetzblatt, Bonn, 2019.
- [Ge19b] Gesetz zur Modernisierung des Strafverfahrens, Bundesgesetzblatt, Bonn, 2019.

- [Ge19c] Gesetz zur Neuregelung des Rechts der notwendigen Verteidigung, Bundesgesetzblatt, Bonn, 2019.
- [Ge21] Gesetz zur Fortentwicklung der Strafprozessordnung und zur Änderung weiterer Vorschriften, Bundesgesetzblatt, Bonn, 2021.
- [MFM23] May, L.; Fahsing, I.; Milne, B.: Die untersuchende Vernehmung von Zeug*innen und Beschuldigten – ein internationaler forschungsbasierter und praxiserprobter Vernehmungsansatz, In: Handbuch Polizeipsychologie, Springer Gabler, Wiesbaden, S. 717ff, 2023.
- [MKL20] Michael, V.; Klöden, M; Labudde, D.: KI-basierte Emotionserkennung in Videos, In: Polizei-Informatik 2020, RediromaVerlag. S 56- 74, 2020.
- [MSO22] May, L.; Schneider, T.; Okulicz-Kozaryn, M.: Notwendigkeit einer kritischen Prüfung von Methoden zur Lügernerkenntnis und Vernehmungstaktiken zur Informationsgewinnung von Beschuldigten: ein Positionspapier aus Psychologie und Polizei. Forensische Psychiatrie, Psychologie, Kriminologie 16, S. 138-143, 2022.
- [MSS06] Mohr, M., Schimpel, F., Schröer N.: Die Beschuldigtenvernehmung. 1. Aufl., Deutsche Polizeiliteratur, Hilden/Rhld, 2006.
- [PA21] PA, Präsentation der Abschlusstagung (online) des Forschungsprojektes BEST vom 16.11.2021, Folie 35, 2021.
- [PFB19] Paquette, S., Fortin, F.; Bergeron, A.; Deslauriers-Varin, N.: Online Child Sexual Exploitation & Solicitation: The PRESEL Projekt.
https://www.researchgate.net/profile/Sarah-Paquette/publication/338434963_Online_Child_Sexual_Exploitation_Solicitation_The_PRESEL_Project/links/5e14d4c44585159aa4bcd916/Online-Child-Sexual-Exploitation-Solicitation-The-PRESEL-Project.pdf, Stand: 30.05.2023.
- [RLB22] Rachlew, A.; Löken, G.; Bergestuen, S.: A Guide to the Professional Interview. Anthem Press, London u.a., 2022.
- [RÖ20] RÖ, Republik Österreich: Aus Verantwortung für Österreich. Regierungsprogramm 2020–2024, Wien, S. 213, 2020.
- [Sc10] Schulz von Thun, F.: Miteinander Reden. Band 1: Störungen und Klärungen, 48. Auflage, Rowohlt Taschenbuch Verlag, Reinbek bei Hamburg, 2010.
- [SK21] Stegemann, L.; Kluth, L.: Vernehmungsfortbildung der Polizei im Vergleich mit außerbehördlichen Angeboten. Kriminalistik 12/2021; S. 679-684, 2021.
- [SP21] SP, SPD, Bündnis 90/Die Grünen u. FDP: Mehr Fortschritt wagen. Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit. Koalitionsvertrag, Berlin, 2021.
- [TKC01] Tian Y.; Kanade, T; Cohn, J.: Recognizing Action Units for Facial Expression Analysis. IEEE Transactions on pattern analysis and machine intelligence, vol.23, no.2, 2001, <http://www.cs.cmu.edu/~face/Papers/112006-1.pdf>, Stand: 01.06.2023.
- [UN22] UN, United Nations: Torture and other cruel, inhuman or degrading treatment or punishment, S. 5, 2022,
https://digitallibrary.un.org/record/3993571/files/A_C.3_77_L.45-EN.pdf?ln=en, Stand: 30.05.2023.

- [WBJ16] Watzlawick, P.; Beavin J.; Jackson, J.: Menschliche Kommunikation. Formen, Störungen, Paradoxien, 13. Auflage, Hogrefe AG, Bern u.a., 2016.

Post-mortem path correlation based on the NT Object Manager in Windows 1x systems

Dominic Helfer¹, Felix Rothe² and Ronny Bodach³

Abstract: The specifications of file and directory paths in forensic artifacts of Windows 1x systems are not uniform. A correlation of paths is needed to prove the hypothesis that two paths in different artifacts describe the same file. During runtime of Windows, this correlation is managed inside the NT Object Manager [A122]. The available information of the NT Object Manager is lost when Windows is shut down, so an analyst with the appropriate knowledge and experience must perform the correlation of paths manually. A mapping of the NT Object Manager is required to develop forensic tools that allow an automated correlation of paths. The mapping was used to develop a reconstruction approach based on an empirical study of differently configured Windows 1x systems. This allows for post-mortem path correlation using non-volatile data.

Keywords: Digital Forensics, Path Correlation, Windows Artifacts, NT-Object Manager

1 Introduction

File and directory paths are crucial in every digital forensic investigation. They can be found in almost all forensic artifacts. The Windows operating system supports a variety of different path formats, namely DOS-path, DOS-Device-path, and UNC-path [Mi22]. Different artifacts also use different formats, in addition to those already mentioned. These formats must be correlated manually by a skilled and well trained digital forensic expert. This impedes the automatic creation of timelines in a consistent format and complicates searching through the artifacts.

A traditional DOS path is the format that users encounter most regularly in Windows Explorer when accessing files on the local computer. A UNC path is used for network shares in the form `\\server\directory\file`. The DOS-Device path consists of the device path specifier `\\?\` or `\\.\\`, followed by the volume or the drive. I.e., it can be in the form:

- `\\?\c:\directory\file.extension`
- `\\?\HarddiskVolume1\directory\file.extension`
- `\\?\Volume{{GUID}}\directory\file.extension`

¹ ConSecur GmbH, Cyber Defense Center, Nödiker Straße 118, 49716 Meppen, helfer@consecur.de

² ConSecur GmbH, Cyber Defense Center, Nödiker Straße 118, 49716 Meppen, rothe@consecur.de

³ University of Applied Sciences Mittweida, Faculty of Computer and Life Sciences, Technikumplatz 17, 09648 Mittweida, bodach@hs-mittweida.de

It is evident that there are numerous ways to point to the same volume using various formats. During the runtime of Windows, the correlation between these different formats is performed using the NT Object Manager [Al22]. It is part of the NT kernel and is handling Windows resources. For example, it performs operations like creating new object types or creating and deleting objects. Objects can be named objects or symbolic links which point to other objects. Other parts of the NT kernel rely on this NT Object Manager, so they don't have to do the resource handling themselves. If the objects are named, they exist in the so-called NT Object Manager namespace which organizes them in a hierarchical directory system. The different directories serve different purposes.

In their work, Helfer et. Al. [HRB22] created a mapping of the NT Objects and symbolic links in the NT Object Manager based on two Windows 10 and one Windows 11 machines. Each machine got a single SATA hard drive assigned. With the help of this mapping, they proposed a reconstruction approach which utilizes information located in the SYSTEM registry hive.

This paper aims to verify and extend the work previously done by creating the NT Object Manager maps for more test scenarios. Especially the test for two different hard drive types, SATA and NVMe, is introduced.

2 Method

To map the NT Objects and symbolic links of hard drives and partitions, we virtualized 16 different Windows 10 (19045) and 11 (22621) systems each using KVM (7.2) and Proxmox VE (7.4-3). Following configuration scenarios were created for each Windows version, with Windows 10 using MBR and Windows 11 using GPT as partition scheme:

- Scenario 1: Windows is installed on an emulated SATA hard drive.⁴
- Scenario 2: Windows is installed on an emulated NVMe hard drive.⁴
- Scenario 3: Two emulated SATA hard drives, Windows installed on first one.
- Scenario 4: Two emulated SATA hard drives, Windows installed on second one.
- Scenario 5: Two emulated NVMe hard drives, Windows installed on first one.
- Scenario 6: Two emulated NVMe hard drives, Windows installed on second one.
- Scenario 7: One emulated SATA, one NVMe drive. Windows on SATA drive.
- Scenario 8: One emulated SATA, one NVMe drive. Windows installed on NVMe.

A PowerShell script utilizing the NtObjectManager module (1.1.20) [Go22] is used to acquire the NT Object Manager namespace. To retrieve the registry hives a volume

⁴ This scenario was already tested in [HRB22]

for Registry hives the RegistryExplorer (2.0.0.0) where used [Ez23].

3 Results

The NT Objects and the symbolic links found in the NT Object Manager namespace depend on the device to which they belong and its capabilities. The NT Objects and symbolic links of CD-ROM, partition and hard drive devices are described in the following sections. The following convention is used to describe variables in the names of objects and links:

- [x] is the number of a hard drive or CD-ROM in the system
- [y] is the number of a partition on a hard drive
- [z] is the number of a volume in the system
- [GUID] is an alpha-numerical Global Unique Identifier [Mi23a].

3.1 CD-ROM devices

In the experiments, three symbolics links pointing to a CD-ROM device were found (see Fig. 1). All symbolic links are in the `\GLOBAL??` directory. If the CD-ROM device is mounted to a drive letter, a link can be found containing the mount point.

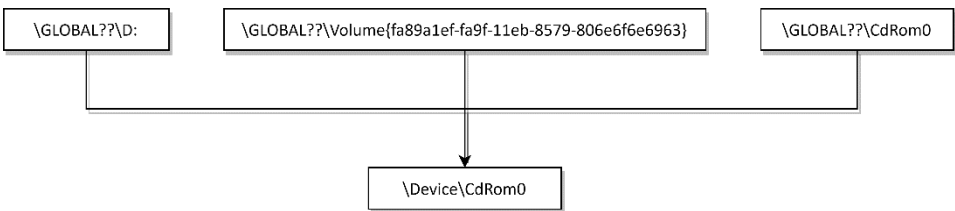


Fig. 1: Example of the symbolic links pointing to the NT Object of a CD-ROM device (Reprint from [HRB22] with permission).

3.2 Hard Drives

In all experiments, four symbolic links point towards the object `DR[x]` in `\Device\Harddisk[x]` which in turn refers to a hard drive (see Fig. 2). Two of them, `PhysicalDrive[x]` and `Disk{[GUID]}`, are in the `\GLOBAL??` directory. The link `multi(0)disk(0)rdisk([x])` in `\ArcName` doesn't point directly to `DR[x]`. Instead, it points to `\Device\Harddisk[x]\Partition0`, which in turn links to the device object.

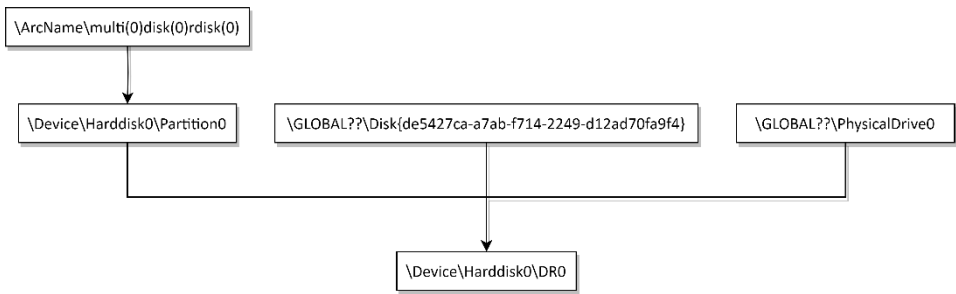


Fig. 2: Example of the symbolic links pointing to the NT Object of a hard drive (Reprint from [HRB22] with permission).

3.3 Partitions

The symbolic links pointing to the NT Object of a partition depend on the nature of the partition. The device driver creates `\Device\HarddiskVolume[z]`. In the experiment we determined that the following properties of a partition influence the symbolic links:

- Does it have a drive letter?
- Is it a boot partition?
- Is it the Windows system partition?
- Is it the Windows Recovery partition? (only on GPT formatted drives)
- Is it the Windows Reserved partition? (only on GPT formatted drives)

All partition Objects have the same structure. They are named `HarddiskVolume[z]` and are in the `\Device` directory. The symbolic link `Partition[y]` is in the subdirectory `Harddisk[x]`. It points directly to the partition object. Like for hard drives, the link `\ArcName\multi(0)disk(0)rdisk(x)partition(y)` is not pointing directly to the NT Object, instead it points to the link `Partition[y]`. The `\GLOBAL??` directory contains multiple links pointing directly to the object (see Fig. 3). For unmounted partitions these are:

- `HarddiskVolume[z]`
- `Harddisk[x]Partition[y]`
- `Volume{[GUID]}`
- `STORAGE#Volume#[[GUID 1]]#[16 digits]#[[GUID 2]]`

If the partition is mounted to a drive letter, an additional symbolic link `\GLOBAL??\drive letter]`: can be found.

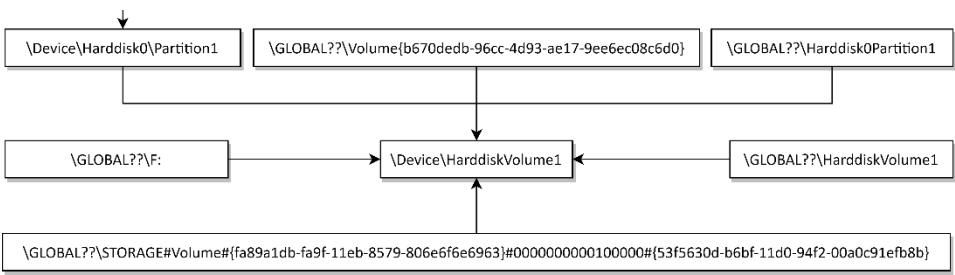


Fig. 3: Example of symbolic links pointing to the NT Object of a mounted partition (Reprint from [HRB22] with permission).

The partition containing the Windows operating system has additional symbolic links (Fig. 4). The `\Device` directory contains the additional links `BootDevice`, `BootPartition` and `OSDataDevice`. The last one doesn't point directly to the partition's NT Object. Instead, it points to the partition's symbolic link in `\ArcName`. The root of the NT Object Manager contains the `OSDataRoot` link which points to the previously mentioned `OSDataDevice` link. For access via the Win32 API, `\GLOBAL??` contains the symbolic link `BootPartition`. It points directly to the partition's object in `\Device`.

The bootloader partition has symbolic links like an unmounted partition. Also, there are the symbolic links `\GLOBAL??\SystemPartition` and `\Device\SystemPartition`. The same symbolic links as for an unmounted partition point to the NT Object of a Windows recovery partition. The Microsoft Reserved Partition, which can only be found on GPT formatted hard drives is missing the `Volume{GUID}` link in the `\GLOBAL??` directory.

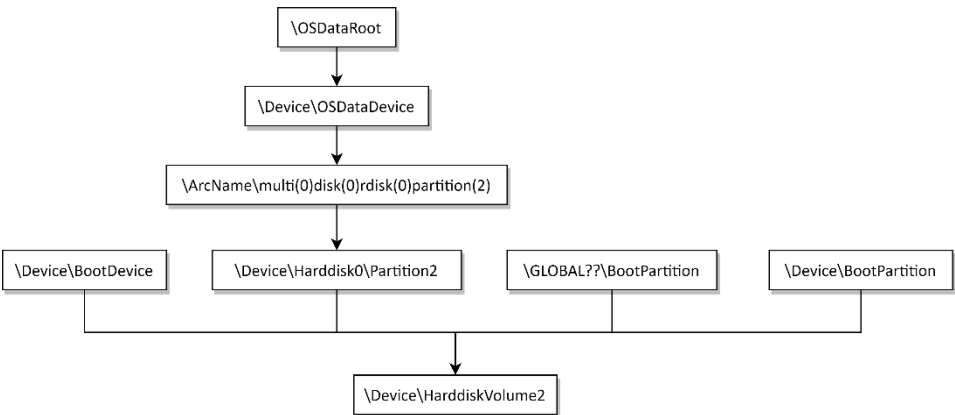


Fig. 4: Additional symbolic links of a partition containing the operation system Windows 10 or 11. The figure only contains symbolic which are not already shown in Fig. 3 (Reprint from [HRB22] with permission).

4 Evaluation

The experiments conducted in this paper showed similar results as the three conducted by Helfer et. Al. [HRB22]. The reconstruction approach described by them must be adopted for NVMe drives because the SCSI Port number is for those due to technical reasons not available.

4.1 CdRom[x]

The variable [x] in the NT Object `\Device\CdRom[x]` and in the symbolic link in `\GLOBAL??` depends on the number of detected CD-ROM devices in the system and their order of detection. The first CD-ROM device is assigned the value 0. This value is incremented for any further devices. The experiments show that the CD-ROM device attached to the lowest SCSI Port is detected first.

The SCSI port can be determined from the value *LocationInformation* in the registry key `SYSTEM:\ControlSet[c]\Enum\SCSI\CdRom&Ven_[vendor]&Prod_[product]\[instance id]` as already described in [HRB22].

4.2 \GLOBAL??\[drive letter] and \GLOBAL??\Volume{[GUID]} (CD-ROM)

Each mounted CD-ROM device is assigned a drive letter which can be used by the user to access the data on this device. The drive letter is assigned by the Mount Manager who creates the symbolic link `\GLOBAL??\Volume{[GUID]}`. For persistence reasons it creates a key-value pair in the registry key `SYSTEM:MountedDevices` which maps the drive letter to a volume GUID [Mi21]. This key-value pairs in the registry can be used to determine the drive letter and the volume GUID for CD-ROM devices as it is detailed described in [HRB22].

4.3 Harddisk[x], PhysicalDrive[x] and multi(0)disk(0)rdisk([x])

The variable [x] in the NT Object `\Device\Harddisk[x]\DR[x]` and in the symbolic links `\Device\Harddisk[x]\Partition0`, `\GLOBAL??\PhysicalDrive[x]` and `\ArcName\multi(0)disk(0)rdisk([x])` depends on the number and order of detected hard drives (like CD-ROM devices in 0). The value [x] starts at 0 and is incremented for each hard drive which was recognized by Windows.

The detection order for the hard drives depends on their type. The experiments showed that NVMe drives are detected earlier than SATA drives. If a system contains only SATA drives, the SCSI port can be used to determine the detection order like it is described in detail in [HRB22]. This information can be found in the *LocationInformation* value in the registry key `SYSTEM:\ControlSet[c]\Enum\SCSI\Disk&Ven_[vendor]&Prod_[product]\[instance id]`.

address rather than the SCSI port. To determine the PCI address of a NVMe drive, the registry key `SYSTEM:\ControlSet[c]\Enum\SCSI\Disk&Ven_[vendor]&Prod_[product]\[instance id]\Properties\{83da6326-97a6-4088-9453-a1923f573b29}\000A` is first used to determine the corresponding PCI device. This can then be looked up under `SYSTEM:\ControlSet[c]\Enum` where the `LocationInformation` contains the PCI address. The experiments showed that the NVMe drive with the lower PCI address is detected first.

If systems have both NVMe and SATA drives, it seems that the NVMe drive is always detected first.

4.4 \GLOBAL??\Disk{[GUID]}

The origin of the disk GUID seen in the symbolic link `\GLOBAL??\Disk{[GUID]}` is still unknown.

4.5 Harddisk[x]Partition[y] and multi(0)disk(0)rdisk([x])partition([y])

The determination of the value `[x]` for the symbolic links `\Device\Harddisk[x]Partition[y]`, `\GLOBAL??\Harddisk[x]Partition[y]` and `\ArcName\multi(0)disk(0)rdisk([x])partition([y])` is the same as described in 4.3. The value `[y]` is incremented for each partition of hard drive starting at 1, since `y = 0` is already used in the symbolic link `\Device\Harddisk[x]\Partition0` to reference the hard drive device. The order the partitions are counted in is the same in which they are listed in the partition table.

4.6 HarddiskVolume[z]

The device description `\Device\HarddiskVolume[z]` is created by the driver for volumes in the NT Object Manager. The value `[z]` is a numerical value starting at 1 which is incremented for every volume device `[On99]`. The value `[z]` in the symbolic link `\GLOBAL??\HarddiskVolume[z]` has the same meaning.

4.7 \GLOBAL??\[drive letter]:

Each mounted partition gets a drive letter assigned by the Mount Manger. It creates a symbolic link of the form `\GLOBAL??\Volume{[GUID]}` and assigns a drive letter to this link. A database (`SYSTEM:\MountedDevices`) in the registry stores the assignment to keep it persistent throughout a system reboot. `[Mi21]` The drive letter is stored in the registry as a value name in the form `\DosDevices\[drive letter]:`. The value itself contains the volume GUID in one of two possible binary formats:

- **DMIO:ID format:** The binary value is 24 bytes long where the first 8 bytes contain the signature DMIO:ID. The remaining 16 bytes represent the volume GUID. `[Op22]`

- **Disk Signature – Offset format:** The second format is 12 bytes long. The first 4 bytes contain the MBR disk signature of the hard drive, and the remaining 8 bytes show the offset of the partition in bytes. Both information’s can be used to reconstruct the volume GUID, as described in 4.8.

4.8 \GLOBAL??\Volume{[GUID]}

The volume GUID referenced in the symbolic link `\GLOBAL??\Volume{[GUID]}` is a 16-byte identifier, following the structure described in [Mi23a]. The experiments verified the findings of [HRB22] that the structure of the volume GUID varies between MBR and GPT formatted drives.

For MBR formatted drives, the first block of the GUID represents the MBR disk signature. This is a four-byte-large value at offset 0x1B8 of the MBR. The following three four-byte-large blocks of the GUID have the value “0000”. The last GUID block encodes the partition start offset in bytes. To determine the offset, the twelve bytes are interpreted as little endian and left shifted by 0x10. In Fig. 5 the corresponding MBR record of a partition with the volume GUID `{6eed82d3-0000-0000-0000-100000000000}` is shown.

```

000001b0: 65 6d 00 00 00 63 7b 9a d3 82 ed 6e 00 00 80 20 em...c{....n...
000001c0: 21 00 07 7f 39 06 00 08 00 00 00 90 01 00 00 7f !...9.....
000001d0: 3a 06 07 fe ff ff 00 98 01 00 7c f3 2d 06 00 fe :.....|.-...
000001e0: ff ff 27 fe ff ff 00 90 2f 06 00 60 10 00 00 00 ..'...../..`....
000001f0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 55 aa .....U.
    
```

Fig. 5: MBR partition table of a hard drive. The MBR drive signature is highlighted in yellow and the start position (LBA) of the partition with the volume GUID `{6eed82d3-0000-0000-0000-100000000000}` is highlighted in black (Reprint from [HRB22] with permission).

Partitions located on GPT formatted drives get the unique partition GUID assigned as volume GUID. To transform a binary GUID into a string the first three GUID blocks (8 + 4 + 4 Bytes) are interpreted as little endian. The last two blocks (4 + 12 bytes) are interpreted as big endian. The Windows Mount Manager assigns every partition a volume GUID and therefor a symbolic link in `\GLOBAL??`. The Microsoft Reserved Partition (partition type: `{e3c9e316-0b5c-4db8-817d-f92df00215ae}`) is an exception [Mi23b].

```

00000420: 00 08 00 00 00 00 00 00 ff 27 03 00 00 00 00 .....'......
00000430: 00 00 00 00 00 00 00 80 45 00 46 00 49 00 20 00 .....E.F.I. .
00000440: 73 00 79 00 73 00 74 00 65 00 6d 00 20 00 70 00 s.y.s.t.e.m. .p.
00000450: 61 00 72 00 74 00 69 00 74 00 69 00 6f 00 6e 00 a.r.t.i.t.i.o.n.

```

Fig. 6: GUID partition table record of a partition with the volume GUID {b670dedb-96cc-4d93-ae17-9ec6ec08c6d0}. The unique partition GUID was highlighted in yellow (Reprint from [HRB22] with permission).

4.9 Windows Partition and Boot partition

The partition containing the Windows operating system has five and the one containing the boot loader has two additional symbolic links in the NT Object Manager. As described in detail in [HRB22], the registry keys *SystemBootDevice* and *FirmwareBootDevice* in the registry key *SYSTEM:\ControlSet[x]\Control* can be used to assign them a hard drive and a partition.

5 Conclusion

This paper aims to verify the reconstruction approach of the NT Objects and symbolic links present in [HRB22] by analysing the NT Object Manager namespace of eight different machine configurations each for Windows 10 and 11. We have shown that the previous reconstruction approach was not suitable for systems with NVMe hard drives. This resulted in incorrectly reconstructed NT objects and symbolic links. The approach presented in this paper allows a reconstruction of these objects also for pure NVMe or mixed SATA and NVMe configurations.

The structure build by NT Objects and symbolic links of CD-ROM, hard drives, and their partitions is essential for normalizing file paths. This normalization is in turn needed for the automated correlation of forensic artifacts from varying sources. It would be desirable to be able to reliably perform this reconstruction on dead systems. Currently, the structure of the NT Object Manager can only be acquired from a live system using for example the PowerShell module *NtObjectManager* [Go22] or a memory dump which is for example analysed using the *symlinkscan* plugin of *Volatility Framework* [Tv20]. The current reconstruction approach doesn't consider PnP storage devices or network shares. This should be addressed in future works to complete the reconstruction approach.

6 Bibliography

- [Al22] A. Allievi, A. Ionescu, M. E. Russinovich and D. A. Solomon: Windows Internals - Seventh Edition - Part 2, Pearson Education Inc., 2022.
- [Ez23] E. Zimmerman: Eric Zimmerman Tools - Registry Explorer, 28/02/2023, Available: <https://ericzimmerman.github.io/#!index.md>, Accessed: 30/06/2023
- [Go22] Google LCC: sandbox-attacksurface-analysis-tools, 23/01/2022, Available: <https://github.com/googleprojectzero/sandbox-attacksurface-analysis-tools>, Accessed: 20/05/2023
- [Hö21] M. Hörz: HxD - Freeware Hex-Editor und Disk-Editor, 11/02/2021, Available: <https://mh-nexus.de/de/hxd/>, Accessed: 28/06/2023
- [HRB22] D. Helfer, F. Rothe and R. Bodach: Korrelation von Zeitstempeln und Pfadangaben von Ausführungsartefakten eines Windows 1x Systems, Mittweida, Hochschule Mittweida, 2022.
- [Mi21] Microsoft Corporation: Supporting Mount Manager Requests in a Storage Class Driver, 15/12/2021, Available: <https://learn.microsoft.com/en-us/windows-hardware/drivers/storage/supporting-mount-manager-requests-in-a-storage-class-driver>, Accessed: 20/05/2023
- [Mi22] Microsoft Corporation: Microsoft Docs - Formate von Dateipfaden, 15/12/2022, Available: <https://learn.microsoft.com/de-de/dotnet/standard/io/file-path-formats>, Accessed: 20/05/2023
- [Mi23a] Microsoft Corporation: Guid Struct, 15/05/2023. Available: <https://learn.microsoft.com/de-de/dotnet/api/system.guid?view=net-7.0>, Accessed: 20/05/2023
- [Mi23b] Microsoft Corporation: UEFI/GPT - based hard drive partitions, 10/02/2023, Available: <https://learn.microsoft.com/en-us/windows-hardware/manufacture/desktop/configure-uefigpt-based-hard-drive-partitions?view=windows-11>, Accessed: 20/05/2023
- [On99] W. Oney: Basic Structure of a WDM Driver - Programming the Microsoft Windows Driver Model, Redmond, Washington, Microsoft Press, 1999.
- [Op22] Open Text Corporation: Windows Drive Letter Assignments, 2022, Available: <https://security.opentext.com/appDetails/Windows-Drive-Letter-Assignments>, Accessed: 20/05/2023
- [Tv20] T. V. Foundation: Volatility Framework - Volatile memory extraction utility framework, 11/12/2020. Available: <https://github.com/volatilityfoundation/volatility>, Accessed: 30/06/2023

Geschwindigkeitsanalyse bei der Sicherung von verschiedenen Speichermedien

Eileen Köhler¹ und Daniel Spiekermann²

Abstract: In der heutigen Zeit ist Speicher keine Kostenfrage mehr. Es ist daher nicht verwunderlich, dass immer mehr Speichermedien mit größerer Speicherkapazität verbaut werden. Für den Endbenutzer ist es eine Erleichterung, da er sich nicht mehr überlegen muss, welche Daten er wirklich braucht und welche nicht. Ein Aufräumen auf dem Handy beispielsweise, um Speicherplatz für neue Bilder oder Nachrichten zu schaffen, entfällt fast vollständig. Unter anderem für Ermittlungsbehörden kann diese Datenflut ein großes Problem darstellen und im Falle der Beweissicherung wie die Suche nach der Nadel im Heuhaufen enden. Bei kritischen oder priorisierten Fällen ist dies ein nicht hinnehmbarer Zustand. Diese Ausarbeitung beschäftigt sich daher mit der Dauer einer Sicherung von Speichermedien mit unterschiedlichen Größen. Daraus soll im weiteren Verlauf erarbeitet werden, ob es Einsparpotenzial bei der Dauer einer Sicherung gibt, damit zum Beispiel Kriminalfälle schneller aufgeklärt werden können. Hierfür werden freie und kommerziell genutzte Programme untersucht, verglichen und hinsichtlich ihrer zeitlichen Sicherungsmöglichkeiten forensisch bewertet.

Keywords: IT-Forensik; Speichermedien; Massendatenanalyse

1 Einleitung

Betrachtet man die Kapazität von Speicherträgern wird ersichtlich, dass eine technische Weiterentwicklung im Gange ist. Die Kapazität eines modernen Smartphones zum Beispiel reicht von 16 Gigabyte bis hin zu einem Terabyte [Ma21]. Nicht nur bei Smartphones ist dieses Phänomen zu beobachten, auch im Bereich der Festplatten werden immer größere Kapazitäten erreicht, nämlich teilweise bis zu einer Kapazität von 18 Terabyte [Ha21].

Was für den Endnutzer praktisch ist, kann für Strafverfolgungsbehörden eine Herausforderung werden. Alle digitalen Beweismittel können am Ende dazu beitragen, eine Straftat vollumfänglich aufzuklären [Ma19]. Betrachtet man neben Strafverfolgungsbehörden auch Unternehmen, dann haben diese ebenfalls Interesse, an einer schnellen Analyse ihrer Daten. Beispielsweise in Situationen, wo ein IT-Zwischenfall zu einem Systemausfall geführt hat. In beiden Fällen ist der Faktor Zeit eine der wichtigsten Komponenten, wenn es um eine schnellere Klärung des Sachverhalts geht. Nimmt die Sicherung eines Datenträgers eine längere Zeit in Anspruch, können die Ergebnisse und gefunden Spuren auch erst später ausgewertet werden.

¹ FernUniversität Hagen, Fakultät für Mathematik und Informatik, Universitätsstraße 11, 58097 Hagen, Deutschland
eileen.koehler@studium.fernuni-hagen.de

² Fachhochschule Dortmund, FB Informatik, Emil-Figge-Straße 42 44227 Dortmund, Deutschland daniel.spiekermann@fh-dortmund.de

Diese Ausarbeitung soll einen Überblick geben, wie lange eine Sicherung mit zwei klassischen forensischen Tools dauert und wie sich die Kapazität des Speichermediums auf die Geschwindigkeit auswirkt. Nach der Einleitung in Kapitel 1 folgen in Kapitel 2 diverse Ausarbeitungen, die sich bereits mit einer ähnlichen Fragestellung beschäftigt haben. In Kapitel 3 werden unsere Ergebnisse einer Versuchsreihe vorgestellt und das letzte Kapitel 4 beinhaltet eine Zusammenfassung der Arbeit und gibt einen Ausblick auf weitere Ansätze.

2 Stand der Forschung

Auf der Webseite von X-Ways wird mit einer effizienten Software geworben, die einzigartig auf dem Markt scheint – vor allem was die zeitliche Performance bei einer Sicherung angeht [AG23]. In [F 23] aus 2011 wird die Aussage getroffen, dass X-Ways mehr als doppelt so schnell sei, wie ein Konkurrenzprodukt. Betrachtet man [Ri13] aus dem Jahr 2013, dann wird X-Ways im Gegensatz zu FTK als deutlich schneller eingestuft bei der Erstellung eines e01 Images. Auch in diesem Report [Bo23], wird X-Ways als nahezu schnellste Software eingestuft. Ebenfalls wird in diesem Bericht aus 2020 [Er20] demonstriert, dass X-Ways fast doppelt so schnell ist, was die Zeit einer Sicherung angeht, als der FTK Imager. Weitere Aspekte für eine schnellere Analyse von digitalen Beweismitteln lassen sich auch in diesen Ausarbeitungen aus dem Jahr 2013 finden [RQM13], [EC13]. Genaue Einstellungen der Programme oder der genaue Versuchsaufbau gehen aus diesen Studien nicht immer hervor.

3 Ergebnisse

Um einen aktuellen Vergleich für 2023 zu bekommen, wurden erneute Tests mit dem FTK Imager und X-Ways Forensics durchgeführt. Dabei wurden USB-Sticks in verschiedenen Größen verwendet. Die USB-Sticks wurden zu Beginn in exFAT formatiert, enthielten keinerlei Daten und wurden auf einem Windows Rechner jeweils mit FTK Imager in der Version 4.3 und X-Ways Forensics in der Version 20.7 als e01 Image gesichert. Die Applikationen zum Sichern wurden jeweils mit Administratorrechten gestartet und es wurden die Standardeinstellungen beider Programme gewählt. Es gab nur einen Sicherungsdurchlauf.

Tab. 1: Ergebnisse der Versuchsreihe für die Erstellung eines e01 Images

	16 GB	32 GB	64 GB	128 GB	256 GB	512 GB
FTK Imager	00:05:07	00:07:52	00:10:21	00:26:10	01:26:02	02:29:34
X-Ways	00:05:36	00:09:36	00:12:10	00:29:48	01:31:48	02:43:14
Zeitliche Differenz	00:00:29	00:01:44	00:01:49	00:03:38	00:05:46	00:13:40

In Abbildung 1 sind die Ergebnisse der Sicherung mit dem FTK und X-Ways dargestellt. Wie zu erwarten, dauert die Sicherung des größten Datenträgers am längsten und des kleinsten Datenträgers am kürzesten. In Tabelle 1 sind die Sicherungszeiten in Stunden

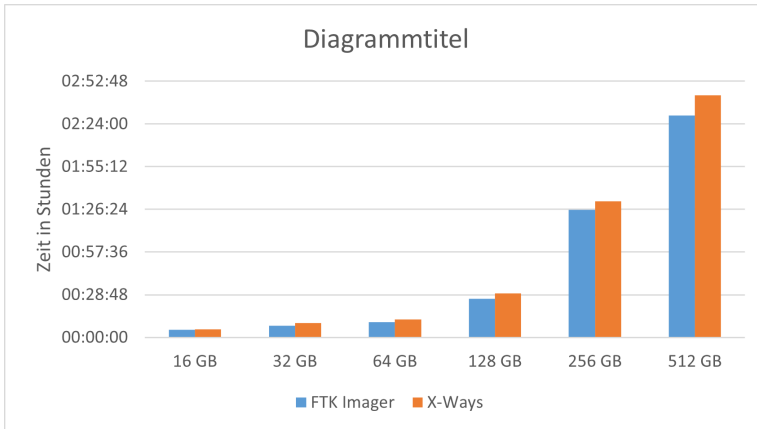


Abb. 1: Ergebnisse der Versuchsreihe im Vergleich

abzulesen. Bei den Zeiten ist die reine Erstellzeit des e01 Images berücksichtigt worden. Andere Image-Formate wurden nicht erstellt.

Die Sicherung des 16 GB USB-Sticks dauerte mit FTK 05:07 Minuten wohingegen der 512 GB USB-Stick eine Sicherungszeit von 02:29:34 Stunden benötigte. Ähnliche Ergebnisse erhält man für X-Ways. Hier dauerte der 16 GB USB-Stick 05:36 Minuten und der 512 GB große Stick 02:43:14 Stunden. Ein stichprobenartiger Vergleich konnte keinen Unterschied zwischen den Dateisystemen feststellen. Es wurde kein externer Writeblocker verwendet.

4 Zusammenfassung

In dieser Ausarbeitung wurden insgesamt sechs verschiedene USB-Sticks forensisch physikalisch gesichert. Dabei wurde von jedem Stick mit ein e01 Image erstellt. Herausgearbeitet werden sollte dabei, welches Programm die effizientere Sicherung im Bezug auf die Zeit liefert und wo die Unterschiede sind.

Es wird deutlich, dass sich der FTK Imager und X-Ways Forensics sehr ähnlich sind, was die Zeiten der Sicherung angeht. Hierbei ergeben sich nur minimale Unterschiede. Die maximale Differenz in der Sicherungszeit beträgt knapp 14 Minuten.

Zusammenfassend kann daher gesagt werden, dass der FTK Imager insgesamt immer schneller war in der Sicherungszeit eines e01 Images als X-Ways Forensics. Die in Kapitel 2 aufgeführten Ergebnisse aus diversen Ausarbeitungen und Berichten, dass X-Ways Forensics zumindest was die zeitliche Komponente angeht performanter ist, konnten damit abschließend nicht bestätigt werden. Es sollte jedoch berücksichtigt werden, dass weitere umfassende Tests durchgeführt werden müssen, um genauere Aussagen treffen zu können.

Gerade Performance Unterschiede des Computersystems müssen hierfür berücksichtigt werden und weitere Sicherungsdurchläufe sind zwingend nötig. Auch sollten die Einstellungen innerhalb der Programme besser aufeinander abgestimmt werden.

Durch diese Ausarbeitung ergeben sich weitere Arbeitsschritte und Untersuchungsoptionen, die in der Zukunft abgearbeitet werden sollten. Spricht man insbesondere die IT-Forensik und die damit verbundene noch durchzuführende Aufbereitung und Auswertung an, dann spielt die Zeit der Sicherung dennoch eine wichtige Rolle bei der Strafverfolgung. Umso schneller eine Sicherung durchgeführt werden kann, umso schneller kann auch eine Ermittlungsperson die Daten weiterverarbeiten.

Literaturverzeichnis

- [AG23] X-Ways Imager: Höchste Geschwindigkeit, intelligenteste Kompression.
- [Bo23] Imaging Speed.
- [EC13] Honing digital forensic processes.
- [Er20] AccessDataGroup FTK Imager 4.3 vs XWaysSoftware X-Ways 19.9 SR-3 imaging a 1 TB NVME.
- [F 23] F-Response Speed Test Results.
- [Ha21] Externe Festplatten - Empfehlungen der Redaktion im Juli 2021.
- [Ma19] Moderne POLIZEI: Polizist der Zukunft: kompetent – digital – vernetzt.
- [Ma21] Wie viel Speicher braucht ein Smartphone wirklich? - 1&1 Magazin.
- [Ri13] Forensic imaging tools.
- [RQM13] Real-time digital forensics and triage.

The digital skeleton in modern video analysis - inter- and intraspecific comparison of individual rigs

Elena Pistorius,¹ Sabine Richter,² and Dirk Labudde³

Abstract: New methods for identifying individuals based on their digital skeleton (rig) gain attention in digital forensics. To ensure a solid comparison, understanding the development of this digital skeleton is crucial. One approach is manually creating a rig using a parametrized 3D room and cameras with embedded pictures, capturing multiple angles of the same person with markers on defined spots. Using these pictures and the room, a movable marker-based rig can be built in software like Blender (<https://www.blender.org>). It can be used to fit the rigged doll into another scene or compare proportions. For a similar process, algorithm-based software like OpenPose (<https://github.com/CMU-Perceptual-Computing-Lab/openpose>) estimates a digital skeleton from pictures or videos. The output includes JSON files with joint coordinates and pictures with embedded skeletons, which can manually be processed into a rig. While these procedure can save time and expand applications, precision and supervision differ from marker-based rigs. A comparative study assessing similarity is discussed in the following paper.

Keywords: video analysis, comparison, individual rigs

1 Introduction

“[On November 25, 2019, a break-in occurred at the Green Vault in Dresden]“ [Ti21]. The perpetrators forcibly gained access to the exhibition rooms, where jewels from the time of Augustus the Strong were on display, setting off a security alarm. However, they managed to flee with their loot before the arrival of the police. In classical forensics, there are various methods available to identify the perpetrators. If fingerprints are found, they can be matched to a previously recorded individual based on the characteristics of the ridge patterns (minutiae). The number of required minutiae depends on the respective country. However, in the aforementioned case, dactyloscopy could not be applied because the perpetrators wore gloves. Moreover, since the crime scene is a public museum, it was expected that there would be a multitude of unknown fingerprints. Another method of personal identification in classical forensics is the analysis of found DNA traces. However, due to the public space and the tactical approach and disguises used by the perpetrators, no reliable leads could be found

¹ University of Applied Sciences Mittweida, Department of Natural and Computer Sciences, Technikumplatz 17, 09648 Mittweida, Germany pistoriu@hs-mittweida.de

² University of Applied Sciences Mittweida, Department of Natural and Computer Sciences, Technikumplatz 17, 09648 Mittweida, Germany richter8@hs-mittweida.de

³ University of Applied Sciences Mittweida, Department of Natural and Computer Sciences, Technikumplatz 17, 09648 Mittweida, Germany labudde@hs-mittweida.de

in this case either. In the field of digital forensics, which does not rely on the comparison of biological and material evidence, facial analysis could have been possible. However, since the perpetrators wore masks, this method could not be applied either. In summary, neither biological nor digital traces were found that could have led to the identification or even the investigation of the perpetrators using commonly employed methods of information retrieval in Germany. The face and hands were covered, and other biological traces were not relevant. However, one individual characteristic that was not concealed is the body proportions and movements. The following method describes how it is possible to identify, assign, include, or exclude a person based on the digital skeleton (digital rig). The approach to the creation and comparison of multiple variations is pursued in this paper.

2 The digital rig and the pose of a person

A digital skeleton, also known as a rig, refers to the abstraction of the human motion or skeletal system and its associated kinematics. “Rigging is the process of setting up a group of controls to operate a 3D model, analogous to the strings of a puppet.” [Or12] Therefore, selected major bones are created along a defined template, such as a video recording or a photograph, following the depiction of the bones shown there. These bones are then connected to each other: “For human animation, landmarks connect weights to positions (joints and sites) on the virtual skeleton (rig) to drive humanoid movement.” [Sc23] In order to simulate human motion, specific properties and constraints are assigned to the defined bones. For example, the knee is treated as a hinge joint, which limits the range of motion of a digital rig between the tibia and femur using inverse kinematics, allowing movement in only one direction. Despite this limitation, it is still possible to rotate the entire leg by rotating the femur in relation to the pelvis. Due to these limitations of the bone movements and relations to each other, it becomes possible to manipulate the digital rig in a way that can assume natural postures. These rigs can then be fitted into postures or movements within a video frame or photo. Such a posture of the digital skeleton is referred to as a pose. Fitting the rig into a video frame or adjusting a pose, not only involves changing the position of the bones relative to each other but also affects the objectively measured size of a person within the frame. If a compact, bent, or squatting posture is evident, the fitting process will result in a smaller measured size compared to an upright pose. That’s why a distinction is made between “pose height“ (body height) and “body size“ (maximum height). Body size refers to the maximum attainable height of a person with legs nearly hip-width apart (forming an imaginary orthogonal line from the hip to the ankle), a straight head and back, and an upright stance. On the other hand, body height refers to the measurable distance from the ground to a well-defined endpoint of the head, based on the currently assumed pose. After clarifying the basic terms used, the following will delve into the methodology employed for rig creation. It should be noted that the created rigs are representations of real individuals. The creation process is based on two different approaches:

- The creation based on well-defined markings placed at selected points of the joint system during a special identification procedure. (marker-based rig)
- Derivation of a digital skeleton based on artificial intelligence. (OpenPose-based rig / OP-based rig)

For this purpose, a mobile treatment room is set up following standardized guidelines (cf. [Be22]). It includes a rotating platform (<https://www.stageonair.com/>) with markings indicating the defined foot position that the person being measured will assume later on. At a well-defined distance, at least one camera is positioned frontally or two cameras are slightly offset from each other and aligned with the rotating platform. The height of the camera systems is determined by the previously measured body size of the person under examination. The center of the camera lens is directed towards the person's body center, which corresponds to half of their body height (see Fig. 1). The camera parameters need to be recorded for subsequent use.

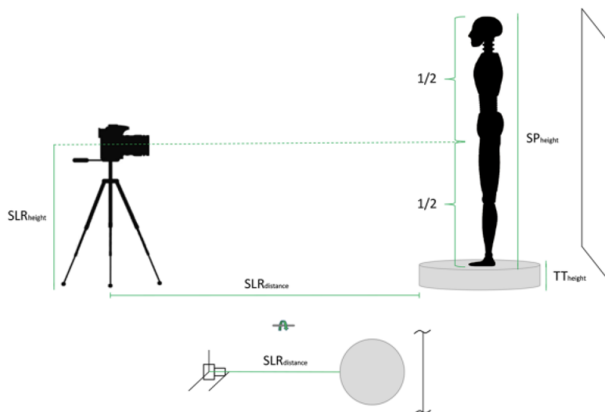


Abb. 1: Exemplary structure of the workstation for photogrammetric data collection. (from [Be22]).

Before processing the person under examination, this setup is captured by a terrestrial laser scanner (FARO FOCUS 3D X130), which creates a 3D point cloud. In order to create a marker-based rig during photogrammetric treatment, it is essential to set defined points of the joint system as anatomical markers on the skin of the person being treated. Through palpation, bony joint points of the subjects are identified and marked on their skin surface to make stable points of the passive musculoskeletal system visible. The marked areas include:

- Jugular notch (Jugulum)
- The shoulder with the acromion of the scapula as a reference
- The elbow joint, with the olecranon of the ulna as the main reference

- The wrist, with the joint space between the forearm bones and the proximal row of carpal bones as a reference
- Bilateral pelvic crest (highest point of the iliac crest)
- The knee, with the patella as the main reference
- The ankle joint, specifically the outermost points on each side of the fibula (lateral) and tibia (medial) as described above
- Tip of the longest toe (specific to each individual, *Digitus pedis I-V*)

With these markers in place, the person under examination is positioned on the rotating platform in such a way that they stand with their feet following the foot markings, hip-width apart and upright. The platform is then rotated 360 degrees while the camera system captures a video or triggers at a defined interval. This process results in a video or a series of individual frames that provide a 360-degree view of the person being measured. Following the data capture, the data processing takes place in several steps. Using the “FARO Scene Software” (<https://www.faro.com/de-DE/Products/Software/SCENE-Software>), “Cloud Compare” (<https://www.cloudcompare.org/main.html>) and “MeshLab” (<https://www.meshlab.net>), the point cloud is processed into a model that can be imported into Blender. During the process, the scale of the model is preserved, ensuring that measurements in the model correspond to real-world measurements. Virtual cameras are placed in the model to match the imaging characteristics of the original cameras. By aligning reference points and markers, the captured image is integrated into the 3D scene. To create a rig, multiple viewpoints of the person are needed, including frontal and lateral views. Examining the person from different perspectives enables the determination of the individual marked joint points in 3D space. These person-specific defined joint points are connected by a mesh. A rig is then applied to this mesh, enabling movement of the bones and joints. This results in a simplified representation of the person’s skeletal system. For the OpenPose-based rig, the person is also treated with photogrammetry, but anatomical markers are not required. The Keypoint Detection Algorithm OpenPose [Ca19] [Si17] is used to predict and visualize body, foot, and facial keypoints for each image. These images, along with the OpenPose predictions, are aligned with the 3D reference model. Using multiple images with different rotations, a mesh and rig can be generated in Blender based on the predicted keypoints from OpenPose.

3 Comparison and evaluation of the creation methods and clothing conditions

3.1 Possible dimensions of the comparison

Possible measures for evaluating the similarity or dissimilarity of two rigs are a length comparison and the Root Mean Square Deviation (RMSD) in normalized and unnormalized

forms. In the length comparison, bone lengths of the two rigs are determined by calculating the Euclidean distances between the starting and ending points of each bone. This measure allows for a pose-independent comparison of the rigs since the bone lengths remain unchanged regardless of the rig's pose. The calculation involves subtracting each bone length of the second rig (Fig. 2, rig B), from the corresponding bone length of the first rig (Fig. 2, rig A). The resulting differences are squared, summed, and divided by the number of bone lengths per rig, n ($n=13$). Taking the square root of the resulting value provides the length comparison result.

$$E^{AB} = \sqrt{\sum_{i=1}^n \frac{(d_{Ai} - d_{Bi})^2}{n}} \tag{1}$$

This measure represents a dissimilarity metric. A value of zero indicates two identical rigs.

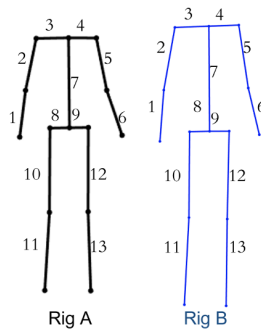


Abb. 2: Exemplary representation of two rigs and the number of determinable bone lengths (modified excerpt from [Be22]).

As the value increases, the two compared rigs become more dissimilar. This measure has been used in previous work [Be22]. The other two measures (RMSD, normalized RMSD) are pose-dependent metrics. They are used when a 2D image is available, and a rig posed in the 3D reference space needs to be compared with this 2D image. For each joint point, the coordinates are first determined, and then the Euclidean distances between corresponding joint points, $d(a, b)$, are calculated.

$$d(a, b) = \sqrt{(X_a - X_b)^2 + (y_a - y_b)^2} \tag{2}$$

The Euclidean distances are then squared, summed, and divided by the number of joint points per rig. Taking the square root of the resulting value provides the RMSD.

$$RMSD = \sqrt{\frac{1}{n} \sum_i d_i^2} \tag{3}$$

The classical RMSD may not provide meaningful comparisons when analyzing results from different captures due to its dependence on image resolution and the number of pixels

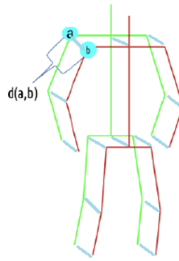


Abb. 3: Exemplary Figure for determining the Euclidean distances between analogous joint points of two rigs for subsequent calculation of the RMSD.

occupied by the person. To address this issue, the normalized RMSD is used. A measuring stick (ms) with a known height is created in the 3D reference space and positioned at the location of the person in the photo or video material. By capturing an image of the measuring stick using the parametrized camera of the 3D reference model, the number of pixels occupied by the measuring stick can be determined. This information, along with the known size of the measuring stick in 3D space, allows for the calculation of a conversion factor in cm/px. Based on the example shown in Figure 4, the following conversion factor can be calculated:

$$ConversionFactor = \frac{h_{ms} \text{ in cm}}{h_{ms} \text{ in px}} = \frac{200cm}{3092px} = 0,065 \frac{cm}{px} \quad (4)$$

By using the conversion factor [cm/px], the normalized RMSD can be calculated. This is done by multiplying the classic RMSD [px] with the factor [cm/px]:

$$RMSD[px] \cdot ConversionFactor[cm/px] = NormalizedRMSD[cm] \quad (5)$$

The normalized RMSD is now independent of factors such as the resolution of the underlying images and the area occupied by the person in the recordings. Thus, the normalized RMSD in cm becomes comparable and interpretable.



Abb. 4: Exemplary visulisation of a measuring stick within the 3D scene (200 cm).

3.2 Inter- and intra-individual comparison of marker-based and OP-based rigs

Although the individuality of digital rigs has been established [HHL22], there is a scientific question regarding the derivability of digital rigs under different conditions, such as clothing and the influence of different creation methods. Based on a small sample of 30 rigs, including 5 subjects with 6 rigs each, the influence of 3 clothing conditions and 2 creation methods will be tested. The 5 subjects are part of the COMBI study, which includes a total of 65 analyzed individuals (A01-A65). The clothing conditions are defined as follows: clothed (regular clothing, no markers placed), unclothed (sportswear, no markers placed), and taped unclothed (sportswear, markers placed). The comparison of the different rigs was performed using the similarity measure of Length RMSD introduced (section 3.1)). However for the calculation, 18 measurement distances in centimeters from the created rigs were considered. In figure 5, the measurement distances are named and depicted on a rig.

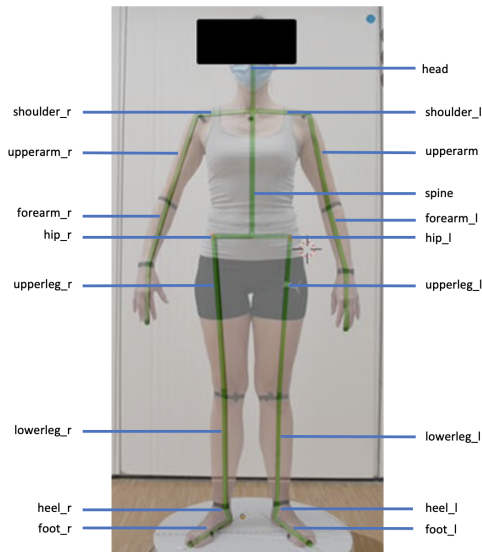


Abb. 5: Exemplary marker-based rig with 18 measurement distances.

Let's first consider the comparison of different clothing conditions within the marker-based rigs. It is evident that the assignment of rigs in the same clothing condition works accurately (green diagonal). This means a one-to-one assignment functions without errors. Furthermore, it can be observed that the assignment of rigs within the clothing conditions of a subject produces smaller RMSD values compared to the comparison between different subjects. This indicates that the influence of the worn clothing within the comparison of a subject is noticeable but always smaller than the comparison to a different subject.

Within the comparisons of OpenPose-based rigs, the one-to-one assignment is also accurately visible. The assignment within a subject under different clothing conditions is mostly present

		Hand Rig															
		A03			A26			A30			A32			A35			
		abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	
H a n d R i g	A03	abgelenkt	0,000	0,008	0,010	0,044	0,048	0,039	0,028	0,020	0,025	0,029	0,030	0,021	0,026	0,027	
		unabgelenkt		0,000	0,011	0,044	0,048	0,039	0,027	0,019	0,023	0,025	0,029	0,031	0,021	0,028	0,028
		bekleidet			0,000	0,045	0,049	0,038	0,029	0,021	0,024	0,025	0,029	0,029	0,023	0,028	0,026
	A26	abgelenkt				0,000	0,009	0,015	0,038	0,043	0,040	0,033	0,023	0,024	0,031	0,026	0,024
		unabgelenkt					0,000	0,019	0,039	0,045	0,040	0,037	0,028	0,029	0,035	0,031	0,028
		bekleidet						0,000	0,032	0,036	0,038	0,026	0,016	0,018	0,026	0,022	0,017
	A30	abgelenkt							0,000	0,014	0,014	0,036	0,032	0,033	0,026	0,029	0,024
		unabgelenkt								0,000	0,012	0,032	0,032	0,032	0,022	0,026	0,025
		bekleidet									0,000	0,035	0,034	0,035	0,026	0,033	0,028
	A32	abgelenkt										0,000	0,013	0,014	0,017	0,022	0,023
		unabgelenkt											0,000	0,010	0,017	0,018	0,018
		bekleidet												0,000	0,018	0,019	0,016
	A35	abgelenkt													0,000	0,012	0,016
		unabgelenkt														0,000	0,014
		bekleidet															0,000

Abb. 6: Representation of the results of the RMSD calculations between 5 marker-based rigs in each of the 3 clothing conditions. Green cells represent small values, while red cells represent high values.

		OP Rig															
		A03			A26			A30			A32			A35			
		abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	abgelenkt	unabgelenkt	bekleidet	
O p e n P o s e R i g	A03	abgelenkt	0,000	0,011	0,016	0,030	0,029	0,025	0,018	0,017	0,022	0,028	0,025	0,032	0,027	0,033	0,026
		unabgelenkt		0,000	0,014	0,025	0,025	0,022	0,019	0,018	0,023	0,024	0,023	0,027	0,020	0,026	0,018
		bekleidet			0,000	0,031	0,032	0,026	0,024	0,024	0,026	0,029	0,030	0,029	0,030	0,029	0,033
	A26	abgelenkt				0,000	0,011	0,011	0,040	0,039	0,044	0,010	0,015	0,014	0,022	0,028	0,023
		unabgelenkt					0,000	0,014	0,038	0,037	0,041	0,015	0,019	0,018	0,020	0,025	0,020
		bekleidet						0,000	0,036	0,035	0,038	0,011	0,015	0,012	0,022	0,027	0,022
	A30	abgelenkt							0,000	0,012	0,020	0,039	0,035	0,042	0,031	0,038	0,030
		unabgelenkt								0,000	0,015	0,036	0,034	0,040	0,027	0,031	0,027
		bekleidet									0,000	0,042	0,040	0,045	0,034	0,035	0,033
	A32	abgelenkt										0,000	0,012	0,011	0,020	0,025	0,021
		unabgelenkt											0,000	0,016	0,025	0,031	0,026
		bekleidet												0,000	0,025	0,029	0,025
	A35	abgelenkt													0,000	0,013	0,008
		unabgelenkt														0,000	0,015
		bekleidet															0,000

Abb. 7: Representation of the results of the RMSD calculations between 5 OpenPose-based rigs in each of the 3 clothing conditions. Green cells represent small values, while red cells represent high values.

as well. However, some discrepancies can be observed in the table. For example, the calculated RMSD between positions A32 clothed and A32 unclothed is 0.016. In the same column, a smaller RMSD is observed between A32 clothed and A26 clothed (0.012) or A26 unclothed (0.014). This means that in this case, the rig of subject A32 clothed could be assigned to the rig of A32 unclothed, but in subsequent positions of similarity, A26 clothed and A26 unclothed would follow. This phenomenon occurs at multiple points in the comparison between the OpenPose rigs. From this, it can be inferred that the different clothing conditions have a greater influence on the defined bone lengths of the rigs compared to the marker-based rigs. Additionally, the presence of markers leads to differing predictions by OpenPose. Finally, let's aim to compare the OpenPose-based rigs with the marker-based rigs in this context.

It is noticeable that there is no discernible pattern of similarity based on the obtained RMSD values. This means that the OpenPose-based rigs and the marker-based rigs are not comparable in all clothing conditions. The RMSD values for the same clothing conditions of the subjects are not significant, nor is the comparison between the subjects. The reason lies within the different methods of creation. The joint points always deviate from the position of the markers.

		OP Rig																
		A03			A26			A30			A32			A35				
		abgeklebt	unabgeklebt	bekleidet	abgeklebt	unabgeklebt	bekleidet	abgeklebt	unabgeklebt	bekleidet	abgeklebt	unabgeklebt	bekleidet	abgeklebt	unabgeklebt	bekleidet		
H # n d R i g	A03	abgeklebt																
		unabgeklebt	0,032															
		bekleidet		0,030	0,034	0,044	0,040	0,042	0,038	0,033	0,034	0,042	0,043	0,047	0,036	0,033	0,033	
	A26	abgeklebt			0,044	0,040	0,042	0,043	0,037	0,037	0,044	0,046	0,050	0,039	0,036	0,036		
		unabgeklebt		0,032	0,036	0,046	0,043	0,045	0,041	0,037	0,037	0,044	0,046	0,050	0,039	0,036	0,036	
		bekleidet			0,044	0,040	0,042	0,039	0,035	0,032	0,043	0,044	0,048	0,038	0,035	0,034		
	A30	abgeklebt			0,059	0,055	0,051	0,074	0,068	0,067	0,058	0,065	0,063	0,053	0,044	0,053		
		unabgeklebt			0,059	0,055	0,051	0,074	0,068	0,067	0,058	0,065	0,063	0,053	0,044	0,053		
		bekleidet			0,061	0,067	0,078	0,072	0,072	0,064	0,070	0,068	0,058	0,049	0,058			
	A32	abgeklebt																
		unabgeklebt																
		bekleidet						0,055	0,067	0,062	0,060	0,054	0,060	0,058	0,049	0,040	0,048	
	A35	abgeklebt								0,058	0,052	0,049	0,062	0,065	0,068	0,053	0,046	0,052
		unabgeklebt								0,041	0,038	0,054	0,056	0,060	0,045	0,039	0,044	
		bekleidet								0,041	0,042	0,059	0,061	0,064	0,051	0,046	0,050	
											0,035	0,039	0,040	0,036	0,030	0,035		
												0,049	0,048	0,042	0,034	0,040		
													0,047	0,041	0,033	0,040		
														0,032	0,025	0,032		
															0,023	0,032		
																0,039		

Abb. 8: Representation of the results of the RMSD calculations between 5 marker-based rigs and 5 OpenPose-based rigs in each of the 3 clothing conditions. Green cells represent small values, while red cells represent high values.

3.3 Joint points and their detection of OpenPose

The use of OpenPose allows for the prediction of joint points and enables (partial) automation. OpenPose also provides the opportunity to utilize additional measures for evaluating the results. The normalized RMSD can now be used as a similarity measure (chapter 3.1). This measure is independent of the display size of the observed person in the video, allowing for the comparison of OP-based rigs of subjects represented with different pixel counts in the video. However, this raises the question of the quality of joint point detection at different distances. In order to determine the required baseline, the subjects were marked with markers at the analyzed joint points. Figure 9 exemplifies the procedure for further testing. In this process, 5 measurement markers are placed on a pre-attached (normalized) grid, corresponding to well-defined distances, which are the same for all 5 subjects. Corresponding frames (individual images) are extracted from the recorded video material and undergo a standardized analysis process. Each selected frame serves as input for the AI - OpenPose, where the joint points (keypoints) described in [Be22] are detected and output both visually and as coordinates. Subsequently, the differences between the anatomical landmarks and the joint points detected by OpenPose are calculated. In comparing the manually placed joint points with the predicted joint points by OpenPose, differences can be observed for all subjects and at all distances (Fig. 10). Due to the various poses and distances to the camera, an average deviation between OpenPose joint points and manually placed joint points can be determined for each joint point. Particularly high deviations between manually placed joint points and OpenPose predictions are observed for the hip points. Deviations are also observed for the shoulder points. The goal should be to define a (partial) rig pattern that uses the predicted joint points with the highest stability and minimal dependence on distance and posture.

4 Discussion and conclusion

The study focused on the role of a digital skeleton (rig) and its application in the forensic context. Forensics can be seen as a comparative science, where traditional forensic methods

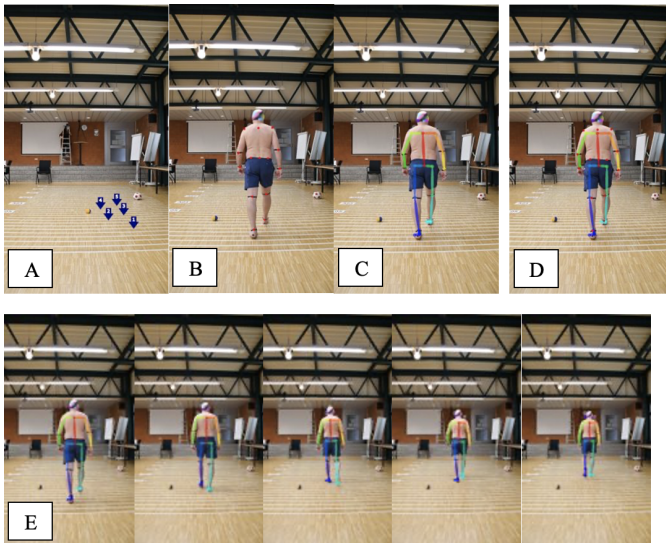


Abb. 9: Procedure - On a pre-attached (normalized) grid, 5 measurement markers are placed (A), which correspond to well-defined distances and are the same for all 5 participants. From the recorded video material, the corresponding frames (individual images) are extracted and subjected to a standardized analysis process. Each selected frame serves as input for the AI - OpenPose, where the keypoints described in Becker et al. are detected and outputted both visually and as coordinates (C, E). Subsequently, the differences between the anatomical landmarks (red markings in B and D) and the detected joint points by OpenPose are calculated (D).

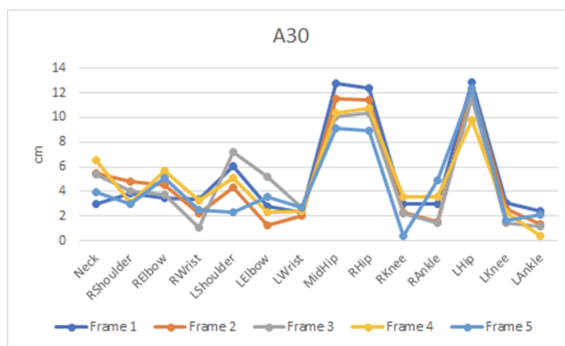


Abb. 10: The figure shows the difference in distances between the anatomical markers and the detected joint points, with varying distance of the person from the camera.

involve a crime scene sample and a reference sample. Similar procedures are followed for digital rigs, using marker-based and AI methods (OpenPose) to create rigs of subjects (suspects) and compare them with rigs from crime-relevant subjects in a well-defined population (perpetrators). Various similarity measures were introduced and discussed in this study. Comparing traces, in this case rigs, from different sources often involves variations in quality. This raises the question of comparability. While the individuality of rigs has been proven under ideal conditions [HHL22], it is necessary to establish limits related to the quality of the underlying data, such as resolution, which represents the depiction of a criminal subject in pixels. Position-based adjustments of rigs were made to demonstrate that associations can still be made even with lower resolution. The influence of different clothing conditions was also tested and yielded similar results in the associations. The possibility of automating rig creation through artificial intelligence methods (OpenPose) simplifies the process, but errors must be carefully determined. Manual markers were placed on anatomically correct joint points of 5 subjects to compare them with the results obtained from AI-based joint point detection (OpenPose). The individual deviations per joint point varied significantly, but they were comparable and fell within the same intervals when comparing rigs from different subjects. This allows for the exclusion of critical joint point predictions through the comparison of perpetrator rigs and suspect rigs, enabling a statistically reliable comparison. This study represents another step towards establishing the digital rig as a biometric characteristic in forensic science.

Literaturverzeichnis

- [Ti21] Timtschenko, F.: Professionelles Sicherheitsmanagement für Unternehmen - Leitfaden für erfolgreiche Corporate Security. p. 96. Springer Gabler, Wiesbaden, 2021.
- [Or12] Orvalho, V. et.al.: A facial rigging survey. EUROGRAPHICS, 2012.
- [Sc23] Scott, E. et.al.: Proposed landmarking for improved digital product creation. CDAPT2023 Vol.4, No. 1, pp.70-87. 2023
- [Be22] Becker, S.; Heuschkel, M.; Labudde, D.: COMBI: Artificial Intelligence for Computer-Based Forensic Analysis of Persons. KI - Künstliche Intelligenz, pp. 1-10, 2022.
- [Ca19] Cao, Z.; Hidalgo Martinez, G.; Simon, T.; Wie, S.; Sheikh, Y. A.: OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.
- [Si17] Simon, T.; Joo H.; Matthews I.; Sheikh Y.: Hand Keypoint Detection in Single Images using Multiview Bootstrapping. CVPR, 2017.
- [HHL22] Heinke, F.; Heuschkel, M. L.; Labudde, D.: A frequentist estimation of duplicate probability as a baseline for person identification from image and video material using anthropometric measurements. INFORMATIK 2022, pp 91-98, 2022.

Digital Twin – Potenziale für die Polizei

Dirk Volkmann¹, Sabine Schildein² und Roman Povalej³

Abstract: Für die moderne Polizeiarbeit haben Themen wie GIS, BIM, CAD, Digital Twin u. a. eine immer größer werdende Bedeutung. Neue Technologien helfen, polizeiliche Aufgaben effektiver und effizienter zu bewältigen. So kann ein Digital Twin z. B. bei der Bewältigung von Einsatzlagen, im ermittelnden Bereich bei On the Case oder Cold Cases die Polizeiarbeit nachdrücklich unterstützen. Allgemein ist das Ziel solch einer kriminalistischen Arbeit die Rekonstruktion und wahrheitsgetreue Repräsentation einer Tathandlung bzw. eines Tatgeschehens. Im Paper werden relevante Begrifflichkeiten betrachtet sowie Einsatzmöglichkeiten geografischer Informationssysteme unter Implementierung von Digital Twins als Referenzmodelle am Beispiel der Vernehmung aufgezeigt. Dabei wird auf die GIS- und BIM-Methoden reflektiert, die sich als vernehmungsunterstützende Elemente anbieten und einen enormen Mehrwert erzeugen können. Abgeschlossen wird das Paper mit einer Reflektion der gewonnenen Erkenntnisse hinsichtlich Bedeutung für die moderne Polizeiarbeit.

Keywords: Geografisches Informationssystem (GIS), Building Information Modeling (BIM), Computer Aided Design (CAD), Digital Model, Digital Shadow, Digital Twin, Veneman, Wahrnehmungs-/Beobachtungs-/Urteilsfehler

1 Einleitung

Die stetig zunehmende Datenflut stellt die Polizei weiterhin vor große Herausforderungen. Neben der zu bewältigenden Quantität muss auch eine ausreichend hohe Qualität der Daten realisiert werden. Vor allem der Umgang mit räumlichen Daten und Informationen bedingt, sich verstärkt mit Begriffen wie GIS (Geografisches Informationssystem), BIM (Building Information Modeling), CAD (Computer Aided Design) oder bzw. und auch Digital Twin auseinander zu setzen. Die Visualisierung raumbezogener Zusammenhänge, die Durchführung von Analysen, die Optimierung von Einsatzplanungen sowie Präventionsmaßnahmen und nicht zuletzt das statistische Berichtswesen haben entsprechend bei den Polizeien Einzug gehalten. Die Assoziation der digitalisierten mit georeferenzierten Daten und Informationen formiert somit ein zentrales Bindeglied zwischen den derzeitigen und den künftigen Aufgabenwahrnehmungen insbesondere im ermittelnden Bereich bei den Polizeien dar.

¹ Polizeiakademie Niedersachsen, Bürgermeister-Stahn-Wall 9, 31582 Nienburg
dirk.volkmann@polizei.niedersachsen.de

² Polizeiakademie Niedersachsen, Gimter Straße 10, 34346 Hann. Münden
sabine.schildein@polizei.niedersachsen.de

³ Polizeiakademie Niedersachsen, Gimter Straße 10, 34346 Hann. Münden
roman.povalej@polizei.niedersachsen.de

Neue Technologien und Systeme, die durch eine digitale Transformation provoziert werden, können dabei assistieren, polizeiliche Anforderungen effektiver und effizienter zu kanalisieren, zu regulieren, zu explizieren bzw. zu partizipieren, z. B. bei der Bewältigung von Einsatzlagen, im ermittelnden Bereich bei On the Case oder Cold Cases durch die Modellierung, Visualisierung und Simulation mittels BIM-CAD-Modells in Form eines Digital Twin.

Nachfolgend wird auf essenzielle Grundlagen und Einsatzmöglichkeiten geografischer Informationssysteme eingegangen, vor allem hinsichtlich der Implementierung von Digital Twins als Referenzmodelle. Die Potenziale für polizeiliche Ermittlungsarbeit sollen am Beispiel der Vernehmung mit den damit ggf. initiierten Wahrnehmungs- bzw. Beobachtungsfehlern demonstriert werden. Abschließend wird auf die mittlerweile unverzichtbaren Instrumente derartiger Systeme reflektiert.

2 Begriffserörterung

2.1 Geografisches Informationssystem (GIS)

Mit einem *GIS*, werden die räumlichen und geografischen Daten extrahiert, integriert, kombiniert, analysiert und präsentiert, d. h. Systematisierung, Registrierung, Organisation, Analyse, Visualisierung bzw. Präsentation räumlicher Informationen i. S. v. Geoinformationen [Bi16, S. 8] – s. Abb. 1 [VSP23].



Abb. 1: Beispiel für eine georeferenzierte Visualisierung eines Objektes als Digital Twin aus drei verschiedenen Perspektiven und mit drei verschiedenen Programmen: Vectorworks, EliteCAD, Lumion

Demzufolge profitiert in einem GIS neben dem Raum- auch der Zeitbezug [St20, S. 5; BRJ06, S. 25 f.]. Besonders bei der Generierung von räumlichen geometrischen, topologischen oder geostatistischen Analysen (z. B. Mobilitäts-, Sensor- oder Netzwerkanalysen, oder deren explorative Visualisierung) wird das ersichtlich [Bi16, S. 193].

2.2 Building Information Modeling (BIM)

Der Begriff *BIM* repräsentiert in seinem Ursprung eine Planungsmethode, die eine Informationsdatenbank rund um ein Bauwerk modelliert und implementiert [NI15]. Heute charakterisiert BIM eine Arbeitsmethode, die für kontinuierliche, zentrale und objektbasierte Administration und Koordination projektrelevanter Informationen steht [Eg13, S. 18]. Das Ziel von BIM formiert sich in der konsequenten digitalen Modellierung aller Eigenschaften und Informationen eines Objektes zur Unterstützung sämtlicher damit verbundenen Prozesse in einem gemeinsamen, ganzheitlichen Modell sowie dem darauf basierenden transparenten Datenaustausch zwischen allen beteiligten Akteuren und Fachdisziplinen [BC20, S. 21]. Dabei kann nachfolgende Differenzierung intendiert werden:

- *BIM-Level 1*: das Digital Model⁴ als zwei- und dreidimensionales CAD⁵-Modell,
- *BIM-Level 2*: den sogenannten Digital Shadow⁶ als dreidimensionales CAD-Modell mit den Daten aus der tatsächlichen physischen Konstruktion (s. Abb. 2 [VSP23]) und

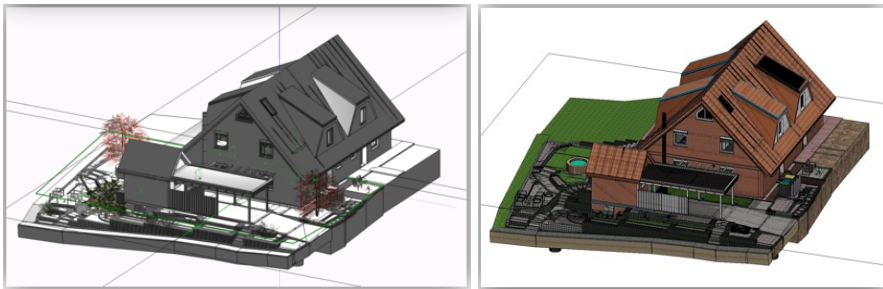


Abb. 2: Modellierung und Visualisierung eines interoperablen CAD-Modells (BIM-Level 2) mit zwei verschiedenen CAD-Programmen: links mit Vectorworks, rechts mit EliteCAD

- *BIM-Level 3*: den Digital Twin⁷ als dreidimensionales CAD-Modell mit bidirektionalen Datenverbindungen (s. Abb. 3 [VSP23]), d. h. BIM-Level 2 unter

⁴ Digital Model – beschreibt eine digitale Darstellung eines bestehenden oder geplanten physischen Objektes, ohne einen automatisierten Datenaustausch zwischen dem physischen und dem digitalen Objekt zu gewährleisten [Kr18, S. 1017]

⁵ CAD illustriert das computergestützte Konstruieren z. T. noch in einer zweidimensionalen Illustration in Ebenen und mit zunehmender Digitalisierung in einer dreidimensionalen Illustration unter Implementierung von Kanten-, Flächen- und Volumenmodellen, u. a. den Entwurf von Produkten, Objekten oder Gegenständen durch computergestützte Grafikerstellung [La18], vornehmlich in BIM-Level 2 und 3

⁶ Digital Shadow – basierend auf Definition eines Digitalen Modells besteht ein automatisierter Datenfluss zwischen dem Zustand eines bestehenden physischen und digitalen Objektes [Kr18, S. 1017]

⁷ Digital Twin – d. h. Digital Shadow unter globaler Integration der Datenflüsse zwischen einem bestehenden physischen und einem digitalen Objekt in beide Richtungen [Kr18, S. 1017]

Implementierung interoperabler Daten, illustrieren bzw. repräsentieren [Kr18, S. 1016 f.].

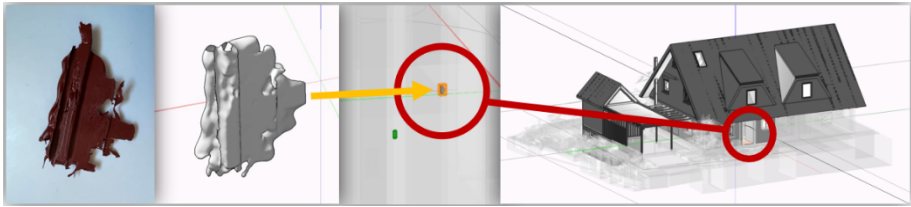


Abb. 3: Silikonabdruck als Negativbild (AccuTrans AB) einer Hebelmarke, sein 3D-Modell als Digital Twin und die georeferenzierte Positionierung seines Auffindeorts an der Eingangstür der Doppelhaushälfte [Umsetzung: Vectorworks]

2.3 Digital Twin

Ein *Digital Twin* wird u. a. definiert als eine globale virtuelle Beschreibung eines physischen Produkts, sowohl auf der Mikro- als auch auf der Makroebene [BAM18]. Es ist somit ein Twinning zwischen dem physischen und dem virtuellen Raum und kombiniert ein physisches Einheit i. S. v. Objekt, ein virtuelles Abbild und die damit implementierten Datenverbindungen [Jo20]. Es handelt sich hierbei um eine komplexe physische und funktionale Charakteristik eines Systems oder einer Komponente, in der die kompletten Informationen impliziert sind [Sc17; Kr18, S. 1017]. Die Technologie des Digital Twins ist Teil der neuen digitalen Technologien und assistiert dabei der digitalen Transformation [VM21]. Häufig handelt es sich um sogenannte Kerntechnologien der Industrialisierung 4.0 (z. B. Internet of Thing, Big Data – siehe auch Abb. 3), insbesondere Simulationsmethoden (z. B. kontinuierliche Simulation) oder Kommunikationsprotokolle (z. B. MQTT, OPC-UA) [Kr18, S. 1018].

2.4 Vernehmung

Eine *Vernehmung* repräsentiert einen rechtlich geregelten Kommunikationsprozess unter Implementierung spezieller Techniken und Methoden und unter Beachtung kriminalistischer Gesichtspunkte, psychologischer Anforderungen und kriminaltaktischer Methoden [ACR19, S. 600]. Sie reflektiert v. a. auf die Untersuchung von Straftaten, kriminalistisch relevanter Sachverhalte sowie Ordnungswidrigkeiten. Dabei fokussiert sich die Vernehmung auf die Generierung des subjektiven Befundes als Personalbeweis, d. h. möglichst objektive i. S. v. unverfälschte Informationen zu handelnden bzw. über andere Personen, zur Sache selbst, zu Handlungsabläufen i. S. v. Tatgeschehnissen und/oder zu Handlungsorten zu projizieren [ASF21, S. 99]. Die Vernehmung steht somit im Kontext v. a. zu Wahrnehmung und Erinnerung [Zi95, S. 207 bzw. S. 345 ff.], sodass entsprechende Verzerrungen durch personen- oder sachbezogene Faktoren den

Informationsgehalt entscheidend hinsichtlich Wahrheit, Irrtum oder gar Lüge dokumentieren [ASF21, S. 27f f.]. Letztlich skizziert der Begriff Vernehmung i. S. d. StPO, dass Polizei als vernehmende Institution der Auskunftsperson in amtlicher Funktion gegenübertritt und in Ausübung dieser Funktion von dieser Person eine Auskunft verlangen kann.

2.5 Wahrnehmung / Wahrnehmungsfehler

Die Vernehmung reflektiert somit auf diverse Transformationsebenen, insbesondere bei der Fähigkeit der sozialen Wahrnehmung, wonach Individuen in der Lage sind, persönliche Eigenheiten bzw. Eigenschaften anderer Individuen zu apperzipieren, zu lokalisieren und zu identifizieren [Zi95, S. 699]. Gerade hier können sich Fehler oder Verzerrungen der sozialen Wahrnehmung (*Wahrnehmungsfehler*), z. B. durch fundamentale Attributionsfehler, motivationsbedingte oder Akteur-Beobachter-Verzerrungen, [Zi95, S. 704 ff.] bzw. als *Beobachtungs-* [Wi19] oder *Urteilsfehler* [UP15, S. 111] manifestieren. Diese lassen sich als Fehler zu Lasten des Beobachters (im polizeilichen Kontext die vernehmende Institution, d. h. wahrnehmungsbedingt, interpretationsbedingt, erinnerungsbedingt und/oder wiedergabebedingt) und der Beobachtung (d. h. Realitäts- und Erwartungseffekte, Beobachtungs- und Untersuchungsbedingungen oder Mängel des Beobachtungssystems [GW97, S. 56 f.; Wi19] bzw. des Akteurs, im polizeilichen Kontext die zu vernehmende Person) kategorisieren.

Wahrnehmung und die damit verbundenen Wahrnehmungssysteme sind somit ein konstruktiver Prozess, der sowohl einen Vorgang als auch ein Ergebnis einer Reizverarbeitung charakterisiert [Wi21]. Dieser Prozess ermöglicht somit neben der einfachen Informationsaufzeichnung als Reaktion auf Umweltreize als Standard- oder Vergleichsreiz auch deren aktive Organisation und Interpretation. Der Prozess kann mittels Stufensystem skizziert bzw. modelliert werden, sodass auf der sensorischen Stufe die Sinnesempfindungen i. S. v. Sinnesdaten, z. B. Reize bzgl. Licht, Helligkeit oder Farbe sowie Reize durch Hören hinsichtlich Frequenz, Amplitude oder Komplexität, kodiert und analysiert, auf der Stufe der Wahrnehmung i. e. S. die gewonnenen Informationen integriert und organisiert und auf der Stufe der Klassifikation dann entsprechend Objekte bzw. Gegenstände identifiziert und kategorisiert werden [Kö12, S. 68 ff.; SS17, S. 18 f.]. Die Wahrnehmung ist dabei von einem Bezugsrahmen abhängig, der sowohl einen realen räumlichen und zeitlichen als auch einen komplementären i. S. v. unabhängig von den Realitäten implementieren kann [Kö12, S. 68 ff.; SS17, S. 102ff.; Zi95, S. 208].

3 Potenzial für die Vernehmung

Das Ziel kriminalistischer Arbeit formiert sich in der Rekonstruktion und wahrheitsgetreuen Repräsentation einer hypothesengeleiteten bzw. -orientierten

Tathandlung bzw. eines Tatgeschehens. Gerade bei Tatörtlichkeiten mit komplexen Raumstrukturen und daraus resultierenden Handlungsabläufen hängt die visuelle Wahrnehmung i. S. v. Verarbeitung eines Individuums u. a. von der Sensation, Kognition bzw. Perzeption ab. Dies bedeutet, dass die Reproduktion der gespeicherten Informationen bzw. Daten mit einem Raum-Zeit-Bezug, die zumeist nur in einer zweidimensionalen Ebene (z. B. Foto vom Tatort) vorliegen, sehr stark von der individuellen kognitiven Leistungsfähigkeit des Beobachters (vernehmende Person) und des Akteurs (zu vernehmende Person) bestimmt wird. Dieser Wahrnehmungs- bzw. Verarbeitungsprozess ist sehr fehleranfällig.

Gerade hier liegen die Potenziale der georeferenzierten Informationsaufbereitung mit der BIM-Methode in Form eines Digital Twins. So wird ein Tatort rekonstruiert und in einem realitätsgetreuen BIM-Modell (idealerweise BIM-Level 3) implementiert und für die Vorbereitung und Durchführung einer Vernehmung bereitgestellt. Durch eine entsprechende Visualisierung des Tatorts im Vernehmungsraum können v. a. die Fehleranfälligkeit in der visuellen Verarbeitung aufgrund eines suboptimal ausgeprägten visuellen Gedächtnisses oder Vorstellungsvermögens beim Akteur und Beobachter, weitestgehend kompensiert werden kann. Die Visualisierung des BIM Modells kann dreidimensional mittels Stereo-, Virtual-(VR-) oder Augmented-(AR)-, auch Mixed Reality-Projektion [DBGJ19, S. 1ff. u. 12 ff.] realisiert werden.

In der Reflexion der GIS- und BIM-Methoden als vernehmungunterstützende Elemente bieten sich vor allem die untersuchenden Interviewtechniken in Kombination mit der Rapport-Methode an, in der Literatur auch PEACE-Methode getitelt. Somit können durch die Anwendung der möglichen Frage- bzw. Vernehmungstechniken und -methoden auf die Reproduktion der ggf. im Gedächtnis gespeicherten Informationen i. S. d. Wahrnehmungs- und Erinnerungskontext abgerufen werden.

Hierbei lässt sich mit Hilfe eines 3D-Modells als Digital Twin der Tatort in Kombination mit den durch die digitale Fallanalyse implementierten ermittlungsrelevanten Daten (Fotos, Videos, Bilder aus diversen Scanverfahren, Spuren des objektiven Befundes usw.) identifizieren, präparieren, organisieren, analysieren und visualisieren.

Auf Basis vorliegender digitaler und analoger Spuren wie zum Beispiel Entfernungs- und Positionsangaben, Mobilitäts- und/oder Livingverhalten bzw. zu Inventarisierung i. F. v. Objekten oder Gegenständen (s. Abb. 4 und Abb. 5) handelnder Akteure oder zu den zum Tatzeitraum vor Ort vorherrschenden Sichtverhältnissen inklusive Helligkeitsanalyse (siehe auch [VSP23]), kann im Rahmen der Erhebung des subjektiven Befundes eines Akteurs (z. B. Zeugin vs. Beschuldigter) der Wahrheitsgehalt der Angaben in der Sache nahezu in Echtzeit v. a. in der Reflektion auf einen Raum-Zeit-Bezug, mit Hilfe hypothesengeleiteter bzw. -orientierter Rekonstruktion des Tatgeschehens bewertet werden.

Die Abb. 4 zeigt das Ergebnis einer hypothesengeleiteten/-orientierten Rekonstruktion eines Tatgeschehens. Der Tatort wurde als Digital Twin erstellt. Videoaufzeichnungen, Daten der vorliegenden SmartHome-(Person-)IoT-Devices und Kamerapositionen wurden

im Digital Twin implementiert und visualisiert. Auf dieser Basis können Plausibilitätsprüfungen ggf. während einer Vernehmung in Real Time vorgenommen werden. Zudem kann der Digital Twin ohne Inventarisierung visualisiert werden (s. Abb. 5). So können retrospektive Schilderungen der wahrgenommenen Objekte oder Gegenstände aus der Erinnerung heraus durch einen Akteur in den Digital Twin integriert werden. Dies ist eine Erweiterung der möglichen Kausilitäts- und Plausibilitätsprüfungen. Mittels einer HoloLens lassen sich visualisierte Soll-Ist-Abgleiche vor Ort vornehmen, um eine sogenannte Nullhypothesenprüfung vorzunehmen. Ein HoloLens-Bild in in der Abb. 5 enthalten.



Abb. 4: Ergebnis-Beispiel für eine hypothesengeleitete/-orientierte Rekonstruktion eines Tatgeschehens unter Zuhilfenahme von Videos und digitaler Spuren zur Implementierung in einem Digital Twin [Umsetzung: Lumion]

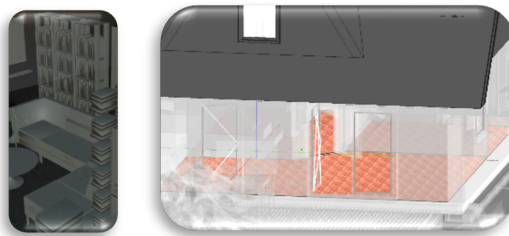


Abb. 5: Links: HoloLens-Bild von einem Raum zwecks Soll-Ist-Abgleich bei erneuter Tatortaufsuche; Rechts: Digital Twin ohne Inventar [Umsetzung: Lumion/Vectorworks]

In einer Vernehmungssituation sollte ein Kind (Angehörige einer vulnerablen Gruppe) als Zeugin im Rahmen einer Vernehmung ihr Wahrgenommenes aus dem Gedächtnis heraus als Erinnerungsleistung wiedergeben. Dabei sollte sie u. a. Angaben zu einer Person (Täter) und der entsprechenden Tageszeit der Kontaktaufnahme machen. Dabei stilisierten sich Wahrnehmungslücken, vor allem bei der Beschreibung der Person und der Tageszeit. Allerdings konnte das Kind ein Detail zu einem Gegenstand sehr klar wiedergeben: „*etwas Blinkendes im Ohr!*“. Im Rahmen der Vernehmung konnte einzig auf einen glänzenden Gegenstand im Ohr, z. B. ein Ohrstecker bzw. -ring, prospierrt werden. Diese relevanten

Informationen implementiert in einen Digital Twin hätten zur Optimierung des Vernehmungszieles beitragen können, sodass durch eine 24/7 Sichtfeld- bzw. Sonnenstandanalyse aus dem Blickwinkel der Zeugin die offenen Fragen hinsichtlich der von der Zeugin beschriebenen Reflektion durch einen natürlichen Lichteinfall (Sonneneinstrahlung) effektiv und effizient geklärt werden könnten – theoretisch in nahezu Real Time.



Abb. 6: Ergebnis einer Sichtfeld- bzw. Sonnenstandanalyse, fokussiert auf den durch das Kind (Zeugin) wahrgenommenen Gegenstand (Ohrstecker) [Umsetzung: Lumion – fotobasierte 3D-Visualisierung; Basis: BIM-Modell – EliteCAD (BIM-Level 3)]

Die Abb. 6 zeigt das Ergebnis der Sichtfeld- bzw. Sonnenstandanalyse aus der Position des Kindes als Zeugin (Kind) im Türbereich. Die Position des Blickwinkels wurde um 90° mit dem Uhrzeigersinn gedreht, um insbesondere den durch das Kind wahrgenommenen Gegenstand (Ohrstecker) am Ohr der betreffenden Person (rot markiert) durch die Reflektion von natürlichem Licht (Sonneneinstrahlung) zu visualisieren und mit dem zeitlichen Ablauf des Sonnenstandes zu reflektieren. In diesem geschilderten Sachverhalt konnte die Frage nach dem Zeitpunkt der Begegnung des Kindes mit der betroffenen Person nicht weiter eingegrenzt werden als auf die Umschreibung „nachmittags“.

Fehler und Verzerrungen können insgesamt leichter identifiziert, revidiert und unter Beachtung der gewählten Vernehmungsstrategie und -technik (bspw. dem *Kognitiven Interview*) vermieden werden. Hinsichtlich der Reflexion auf Wahrnehmungs- und Erinnerungskontexte, inklusive Perspektivwechseln, können Widersprüche zielorientiert aufgedeckt werden. Dadurch wird es möglich, bei zusätzlicher Einbindung spezifischer Vernehmungs- bzw. Fragemethoden (exemplarisch sind hier antiquierte jedoch bekannte Methoden aufgeführt wie Zick-Zack-Methode, Überrumpelung- oder Überraschungsmethode) aufkommende Widersprüche verzerrungsfrei schnell reflektieren zu können.

Die in der Vernehmung visualisierten bzw. modellierten Raum-Zeit-Daten lassen sich in einem Digital Twin ggf. in Real Time problemlos integrieren (s. Abb. 6 [VSP23]), sodass das Ziel der Vernehmung durch den Einsatz der 3D-Technik nicht nur von der kognitiven Leistung der handelnden Personen i. S. v. Akteur und Beobachter fixiert wird, sondern objektiv von der Minimierung der Fehler- und Verzerrungsfaktoren partizipiert.

4 Ausblick

Resümierend bleibt festzustellen, dass parallel zum Thema Digitalisierung 4.0 auch die Methoden der 3D-Modellierung und Visualisierung insbesondere von Digital Twins immens prosperieren und sich somit auch in den polizeilichen Arbeitsabläufen bzw. -prozessen stärker manifestieren sollte.

Die Implementierung dieser Methoden postuliert einen Paradigmenwechsel, der in der Praxis nicht abrupt funktionieren wird. Er sollte in aufeinander abgestimmten Stufen realisiert werden. Aus diesem Grund sind beim Wissensmanagement i. S. v. Transferleistung über GIS und BIM im Besonderen zu Digital Shadow und Digital Twin prägnante und plausible Wirkmechanismen erforderlich. So könnte über den angeregten Kommunikationsprozess ein progressiver Wissenstransfer initiiert und somit eine hohe bzw. offene Transparenz dokumentiert werden und würde folglich auch im progressiven Datenqualitätsmanagement reflektieren.

Gleichbedeutend wird mit der GIS- und BIM-Methode i. F. e. Digital Twin und dem damit einhergehenden Datenqualitätsmanagement die Realisierung von Transparenz, Konsistenz und Integration aller Daten postuliert, die erst einen automatisierten Datenfluss bzw. Workflow garantiert - gleichbedeutend mit der Distanzierung von sogenannten Individuallösungen.

Die Lebenswirklichkeit ist dreidimensional, die Realität der Verfahrensakte, auch als elektronische Akte, nur zweidimensional [La13, S. 321 f.]. Mit Orientierung auf die Zukunft werden sich sowohl die GIS- und BIM-Methoden und insbesondere Digital Twins als auch das progressive Datenqualitäts-, Kommunikations- und Wissensmanagement dauerhaft in der Polizeiarbeit zur Gewährleistung eines automatisierten Workflows etablieren und somit direkt auch auf die Vernehmungsstrategie, z. B. durch Echtzeitreflektion des subjektiven mit dem objektiven Befund, Einfluss haben und somit immens zur Qualitätssteigerung beitragen.

Literaturverzeichnis

- [ACR19] Ackermann, R.; Clages, H.; Roll, H.: Handbuch der Kriminalistik, 5. Auflage, Boorberg Verlag, Stuttgart, 2019.
- [ASF21] Artkämper, H.; Schilling, K.; Floren, Th.: Vernehmungen. Taktik - Psychologie - Recht, 3. Auflage, Verlag Deutsche Polizeiliteratur GmbH, Hilden, 2021.
- [BAM18] Baruffaldi, G.; Accorsi, R.; Manzini, R.: Warehouse management system customization and information availability in 3pl companies: A decision-support tool. *Industrial Management & Data Systems*, Emerald Publishing Ltd, Bingley, Band 119, Nr. 2, S. 251-273, 2018, <https://doi.org/10.1108/IMDS-01-2018-0033>, Stand: 12.06.2023.
- [Bi16] Bill, R.: Grundlagen der Geoinformationssysteme, 6. Auflage, Wichmann-Verlag, Berlin, 2016.

- [BC20] Blankenbach, J.; Clemen, C.: Grundlagen: Die Methode BIM. In (DVW e. V., Runder Tisch GIS e. V., Hrsg.): Leitfaden Geodäsie und BIM, Version 2.1 (2020), Wichmann Verlag, Bühl/München, 2020, <https://dvw.de/images/anhang/2757/leitfaden-geodaesie-und-bim2020onlineversion.pdf>, Stand: 22.02.2021.
- [BRJ06] Booch, G.; Rumbaugh, J.; Jacobsen, I.: Das UML Benutzerhandbuch - Aktuell zur Version 2.0. Addison-Wesley Verlag, München, 2006.
- [DBGJ19] Dörner, R.; Broll, W.; Grimm, P.; Jung, B.: Virtual und Augmented Reality (VR/AR). Grundlagen und Methoden der Virtuellen und Augmentierten Realität, 2. Auflage, Springer Vieweg Verlag, Berlin, 2019.
- [Eg13] Egger, M.; Hausknecht, K.; Liebich, T.; Przybylo, J.: BIM-Leitfaden für Deutschland. Information und Ratgeber. Endbericht. Bundesinstitut für Bau-, Stadt- und Raumforschung (Hrsg.), Bonn, 30.11.2013, <https://bmdv.bund.de/SharedDocs/DE/Anlage/DG/Digitales/bim-leitfaden-deu.html>, Stand: 23.12.2020.
- [GW97] Greve, W.; Wentura, D.: Wissenschaftliche Beobachtung. Eine Einführung, 1. Auflage, Beltz Psychologie VerlagsUnion, Weinheim, 1997, <https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/wentura/GreveWentura1997.pdf>, Stand: 06.01.2022.
- [Jo20] Jones, D. S: Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology, Volume 29, Part A, Pages 36-52, ISSN 1755-5817, Mai 2020, <https://doi.org/10.1016/j.cirpj.2020.02.002>, Stand: 12.06.2023.
- [Kö12] Körber, B.: Visuelle Wahrnehmung. In (Lorei, C., Sohnemann, J., Hrsg.): Grundwissen Eigensicherung, 2. Auflage, Verlag für Polizeiwissenschaft, Frankfurt, S. 67-82, 2012.
- [Kr18] Kritzing, W. K.; Karner, M.; Traar, G.; Henjes, J.; Sihm, W.: Digital Twin in manufacturing: A categorical literature review and classification. Conference Paper Archive. (I. PapersOnline, Hrsg.) Amsterdam, Elsevier BV, 51-11, S. 1016-1022, 2018, <https://doi.org/10.1016/j.ifacol.2018.08.474>, Stand: 12.03.2023.
- [La18] Lackes, R.: CAD. Gabler Wirtschaftslexikon. Springer Nature Verlag, Wiesbaden, 19.02.2018, <https://wirtschaftslexikon.gabler.de/definition/cad-29408/version-253016>, Stand: 12.06.2023.
- [La13] Larl, W.: Virtuelle Realität. In (Artkämper, H., Clages, H., Hrsg.): Kriminalistik gestern - heute - morgen. Festschrift zum 10-jährigen Bestehen der Deutschen Gesellschaft für Kriminalistik, 1. Auflage, Boorberg Verlag, Stuttgart, Bd. 4, S. 317-328, 2013.
- [NI15] NIMBS, N. B.: National BIM Standard-United States. (N. I. Staaten, Hrsg.) Washington, 05.2015, https://www.nationalbimstandard.org/files/NBIMS-US_FactSheet_2015.pdf, Stand: 23.12.2020.
- [Sc17] Schleich, B.; Anwer, N.; Mathieu, L.; Wartzack, S.: Shaping the digital twin for design and production engineering. CIRP Annals, Volume 66, Issue 1, Pages 141-144, 2017, <https://doi.org/10.1016/j.cirp.2017.04.040>, Stand: 12.07.2023.
- [SS17] Spering, M.; Schmidt, T.: Allgemeine Psychologie 1. kompakt: Wahrnehmung, Aufmerksamkeit, Denken, Sprache, 3. Auflage, Beltz Verlag, Weinheim/Basel, 2017.

- [St20] Strobl, J.: Dynamische Geoinformatik - Geleitwort. In (Zagle, B., Loidl, M., Hrsg.): Geo-IT in Mobilität und Verkehr, 1. Auflage, Wichmann Verlag, Berlin/Offenbach, S. 5-7, 2020.
- [UP15] Unkelbach, C.; Plessner, H.: "Category-Split" Effekte bei Urteilen über Sportlerinnen, Sportler und Sportarten. Zeitschrift für Sozialpsychologie, Hogrefe Verlag, Göttingen, Vol. 38, No. 2, 05.03.2015, <https://doi.org/10.1024/0044-3514.38.2.111>, Stand: 07.01.2022.
- [VM21] VanDerHorn, E.; Mahadevan, S.: Digital Twin: Generalization, characterization and implementation. Decision Support Systems, Volume 145, 113524, June 2021, <https://doi.org/10.1016/j.dss.2021.113524>, Stand: 12.06.2023.
- [VSP23] Volkman, D.; Schildein, S.; Povalej, R.: Smart City / Smart People – mit Digital Twins dem Täter auf der Spur. In (Honekamp, W., Fährdrich, J., Hrsg.): Polizei-Informatik 2023, Rediroma Verlag, Remscheid, to be published 2023.
- [Wi19] Wirtz, M. A.: Beobachtungsfehler. Dorsch Lexikon der Psychologie. (M. A. Wirtz, Hrsg.) Bern, Hogrefe Verlag, 02.07.2019, <https://dorsch.hogrefe.com/stichwort/beobachtungsfehler#search=f5899511924b3ba0ae1bb53edb18af4b&offset=0>, Stand: 06.01.2022.
- [Wi21] Wirtz, M. A.: Beobachtungsfehler. Dorsch Lexikon der Psychologie. (M. A. Wirtz, Hrsg.) Bern, Hogrefe Verlag, 08.03.2021, <https://dorsch.hogrefe.com/stichwort/wahrnehmung#search=6ef2c31a519d43d355bf466931345363&offset=0>, Stand: 06.01.2022.
- [Zi95] Zimbardo, P. G.: Psychologie. (bearbeitet und herausgegeben von Hoppe-Graff, S., Keller, B.), 6. Auflage, Springer Verlag, Berlin/Heidelberg/New York, 1995.

Möglichkeiten, Grenzen und Konzepte für eine digitalbasierte Täterbewertung im Bereich der §§ 174 – 184

Mirijam Steinert¹, Sabrina Herchel² und Dirk Labudde³

Abstract: Im Zusammenhang mit Straftaten gegen die sexuelle Selbstbestimmung (§§ 174 - 184) kommt es üblicherweise zu einer Auswertung von digitalen Beweismitteln. Diese Datenmengen wachsen stetig. Durch die Zusammenführung aller verfügbaren Daten innerhalb des Verfahrens ergeben sich neue Chancen, um weitere Erkenntnisse zu gewinnen. Bei der Analyse von Datenträgern werden vorwiegend multimediale Inhalte wie Bilder, Videos und Chatnachrichten ausgewertet und in einem Gutachten verschriftlicht. Oftmals stellt sich die Frage, wie hoch das Gefährdungspotenzial durch den Beschuldigten ist. Bei der Begutachtung von Sexualstraftätern werden Risikoanalysen durchgeführt. Aus welchen Elementen der verfügbaren digitalen Daten können bei der Risikobewertung von Sexualstraftätern Schlüsse gezogen werden und ergeben sich dabei weitere Möglichkeiten der Risikobewertung? Kann ein Täterprofil bei hinreichender Nutzeridentifizierung (Täter = Nutzer) auf Basis der verfügbaren digitalen Daten automatisiert erstellt werden, und welche spezifischen Informationen sind dafür relevant? Welche gesetzlichen und technischen Einschränkungen müssen bei der Beantwortung dieser Fragen beachtet werden, und welche Risiken sind damit verbunden?

Keywords: Risikoanalyse, Automatisierung, Sexualstraftäter, Beweismittelanalyse, Datenselektion, Digitale Forensik

¹ FZ forensic.zone GmbH, Freiburger Straße 11, 09648 Mittweida, info@forensiczone.de

² sabrina.herchel@outlook.com

³ Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida, labudde@hs-mittweida.de

1 Digitale Forensik und die Risikoanalyse von Sexualstraftätern

Sexualstraftaten gehören zu den schwersten Straftaten, die gegen die körperliche Integrität von Personen gerichtet sind und gegen die Selbstbestimmung anderer Personen verstoßen [18]. Eine wichtige Maßnahme, um potenzielle Opfer zu schützen und die öffentliche Sicherheit zu gewährleisten, ist die Beurteilung des Gefährdungspotenzials von Sexualstraftätern. Dabei wird untersucht, inwiefern der Täter weiterhin eine Bedrohung darstellt und welche Maßnahmen zur Risikoreduzierung ergriffen werden können. In diesem Zusammenhang gewinnt die Auswertung digitaler Daten von Beschuldigten immer mehr an Bedeutung, da sie eine wichtige Informationsquelle darstellen können.

Durch einfache Zugriffsmöglichkeiten im Internet werden Sexualdelikte gegen Kinder durch verschiedene Arten und Begehensweisen von Cyberkriminalität ermöglicht, wie zum Beispiel Kinderpornografie, sexuelle Belästigung und Cyber-Grooming [RT07].

Zum gegenwärtigen Zeitpunkt geht man von verschiedenen Arten von Cyber-Sexualstraftätern aus, die sich nach ihren Online-Aktivitäten (z.B. Anschauen oder Anwerben von Kindern), nach der Häufigkeit und dem Alter ihrer Opfer unterscheiden lassen. Zum Beispiel gibt es „Gelegenheitsnutzer“, die aus Neugier Kinderpornografie konsumieren, und „präferentielle“ Täter, die sie gezielt suchen. McLaughlin stellte 1989 unter anderem fest, dass viele der Cyber-Sexualstraftäter in Berufen arbeiten, die entweder direkt oder indirekt mit Kindern zu tun haben (Lehrer, Polizisten, Anwälte, medizinisches Personal, Fotografen, religiöse Führer und viele mehr) [RT07].

1.1 Gefährdungspotenzial von Sexualstraftätern

Das Risikoprofil eines Sexualstraftäters im Netz beschreibt die individuellen Merkmale, die das Risiko einer erneuten Sexualstraftat beeinflussen [RE07]. Dazu gehören zum Beispiel die Art und Schwere der begangenen oder geplanten Straftat, die Motivation und Einstellung des Täters, die Verfügbarkeit von Opfern oder Tatgelegenheiten, die Anwesenheit oder Abwesenheit von Schutzfaktoren wie soziale Unterstützung oder Therapiebereitschaft, sowie die Reaktion auf Interventionen oder Sanktionen [Kr09].

Die Risikoanalyse von Sexualstraftätern ist ein kritischer Bereich der Kriminalpsychologie und forensischen der Beurteilung. Ziel dieser Analyse ist es, die Wahrscheinlichkeit zu bestimmen, dass ein Individuum in der Zukunft weitere sexuelle Straftaten begeht. Sie ist sowohl ein diagnostisches als auch ein präventives Werkzeug und kann dazu beitragen, den Kreislauf der Wiederholungstäter zu durchbrechen [Kr09].

Die Risikoanalyse erfolgt in der Regel in mehreren Schritten und beinhaltet sowohl statische als auch dynamische Risikofaktoren. Statische Risikofaktoren sind solche, die nicht verändert werden können, wie die Vergangenheit eines Individuums, einschließlich vorheriger Verurteilungen oder früher sexueller Straftaten. Sie sind wichtige Indikatoren

für das Risiko, da Studien gezeigt haben, dass Personen, die in der Vergangenheit sexuelle Straftaten begangen haben, ein höheres Risiko haben, in der Zukunft weitere Straftaten zu begehen [GZ13].

Dynamische Risikofaktoren hingegen sind Merkmale oder Umstände, die sich im Laufe der Zeit ändern können. Dazu gehören Aspekte wie der Umgang mit Stress, die Fähigkeit zur Empathie, die Verwendung von Drogen oder Alkohol und die Einstellung zur Sexualität. Diese Faktoren können mit therapeutischer Intervention verändert werden und bieten daher Möglichkeiten zur Risikominimierung [GZ13].

Zusätzlich zu diesen Faktoren werden in der Risikoanalyse auch Schutzfaktoren berücksichtigt. Dies sind Aspekte, die das Risiko für weitere Straftaten reduzieren können, wie etwa stabile familiäre Beziehungen, eine positive Arbeitsgeschichte oder Teilnahme an Behandlungsprogrammen.

Die Methoden zur Durchführung von Risikoanalysen variieren, können psychologische Bewertungen, Interviews, Überprüfung von Strafregistern und andere Dokumente, Beurteilungen durch Fachleute und standardisierte Bewertungsinstrumente umfassen. Einige der bekanntesten Instrumente zur Risikoeinschätzung bei Sexualstraftätern sind der Static-99 [KB20], der Sex Offender Risk Appraisal Guide (SORAG) [RE07] und der Risk for Sexual Violence Protocol (RSVP) [HB20].

Es ist wichtig zu betonen, dass Risikoanalysen, obwohl sie auf fundierten Methoden und Forschungen basieren, Vorhersagen sind und daher eine Unsicherheit aufweisen. Sie sollten immer als Teil eines umfassenden Ansatzes zur Risikomanagement und -minimierung gesehen werden, der auch Interventionen zur Verhaltensänderung, Überwachung und Unterstützung beinhaltet [St06].

Insgesamt spielt die Risikoanalyse von Sexualstraftätern eine wesentliche Rolle bei der Vorbeugung von Sexualstraftaten und dem Schutz von Gemeinschaften. Sie ist ein kritischer Faktor in der Entscheidungsfindung bezüglich Strafverfolgung, Gerichtsverfahren, Strafzumessung, Bewährung und der Bereitstellung von Behandlungsprogrammen.

Seto et al. untersuchten in ihrem Artikel die Frage, wie wahrscheinlich es ist, dass Online-Sexualstraftäter (insbesondere Online-Kinderpornografie-Täter) entweder schon einmal oder in Zukunft Offline-Sexualstraftaten mit Kontakt zu einem Opfer begangen haben oder begehen werden. Das Paper führt dazu zwei Meta-Analysen durch: Die erste analysiert die Kontakt-Sexualstraftatengeschichte von Online-Tätern, während die zweite die Rückfallraten von Online-Tätern in Folgestudien untersuchte. Die erste Meta-Analyse ergab, dass etwa 1 von 8 Online-Sexualstraftätern (12%) zum Zeitpunkt ihrer Index-Straftat eine offiziell bekannte Kontakt-Sexualstraftatengeschichte haben ($k = 21$, $N = 4.464$). Etwa einer von zwei (55%) Online-Sexualstraftätern gab in den sechs Studien, die Selbstberichtsdaten hatten ($N = 523$), eine Kontakt-Sexualstraftat zu. Die zweite Meta-Analyse zeigte, dass 4,6% der Online-Sexualstraftäter während einer 1,5- bis 6-jährigen Nachbeobachtung eine neue Sexualstraftat irgendeiner Art begingen ($k = 9$, $N = 2.630$); 2,0% begingen

eine Kontakt-Sexualstraftat und 3,4% begingen eine neue Kinderpornografie-Straftat. Die Ergebnisse dieser beiden quantitativen Überprüfungen legen nahe, dass es möglicherweise eine unterschiedliche Untergruppe von Online-only-Tätern gibt, die ein relativ geringes Risiko haben, in Zukunft Kontakt-Sexualstraftaten zu begehen [SKB11].

Zu den in der Literatur am häufigsten gezählten Risikoindikatoren für eine Wiederholungstat bei Sexualstraftätern gehören:

Vorherige Straftaten: Das Vorhandensein von früheren Sexualstraftaten ist ein starker Prädiktor für zukünftige Sexualstraftaten. Je mehr und je schwerer die vorangegangenen Delikte waren, desto höher ist das Rückfallrisiko [KB20].

Antisoziale Persönlichkeitsmerkmale: Merkmale wie Impulsivität, Aggressivität, mangelnde Empathie und Verantwortungslosigkeit können das Risiko erhöhen. Diese Merkmale können auch mit anderen Problemen wie Drogen- oder Alkoholmissbrauch, psychischen Störungen oder sozialer Isolation einhergehen [Ho20].

Sexuelle Devianz: Interesse an oder Beteiligung an sexuellen Aktivitäten, die als gesellschaftlich inakzeptabel oder illegal angesehen werden, kann ebenfalls ein Risikofaktor sein. Dazu gehören zum Beispiel Pädophilie, Paraphilien oder sexuelle Gewaltfantasien [HH20].

Einstellungen und Überzeugungen: Überzeugungen, die sexuelle Gewalt legitimieren oder rechtfertigen, können ebenfalls das Risiko erhöhen. Dazu gehören zum Beispiel die Verharmlosung oder Leugnung der eigenen Taten, die Abwertung oder Schuldzuweisung der Opfer oder die Rationalisierung der eigenen Bedürfnisse [AV03].

Drogen- oder Alkoholmissbrauch: Drogen- und Alkoholkonsum kann das Risiko erhöhen, da sie die Hemmungen senken und die Impulskontrolle beeinträchtigen können. Außerdem können sie zu anderen Problemen wie Arbeitslosigkeit, Schulden oder Konflikten mit dem Gesetz führen [HMH06].

Soziale Isolation oder mangelnde soziale Unterstützung: Dies kann das Risiko erhöhen, da es die Wahrscheinlichkeit verringert, dass die Person positive soziale Einflüsse oder Unterstützung hat. Eine stabile Partnerschaft, eine gute Beziehung zu Familie oder Freunden oder eine sinnvolle Beschäftigung können hingegen das Risiko senken [GL16].

Nicht-Einhaltung von Behandlungsplänen oder Bewährungsaufgaben: Dies kann ein Indikator dafür sein, dass die Person nicht bereit oder in der Lage ist, Verhaltensänderungen vorzunehmen. Eine erfolgreiche Therapie kann das Risiko senken, indem sie die Person dazu befähigt, ihre Probleme zu erkennen und zu bewältigen, alternative Strategien zu entwickeln und ihre Einstellungen und Überzeugungen zu verändern [BBH21].

2 Digitale Beweismittelanalyse und Automatisierung

Im realen Raum können physische Beweise, wie handgeschriebene Notizen, Drucke von Bildern oder Videos und andere physische Gegenstände, gesammelt und analysiert werden. Techniken wie die Handschriftenanalyse können zur Identifizierung von Tätern eingesetzt werden.

Digitale Forensik ist ein wachsender und zunehmend wichtiger Bereich der forensischen Wissenschaft. Im Gegensatz zur analogen Forensik, die sich auf physische Beweise konzentriert, befasst sich die digitale Forensik mit Informationen, die in digitaler Form gespeichert sind.

Analysten nutzen eine Vielzahl von Werkzeugen und Techniken, um Daten von digitalen Geräten zu extrahieren, zu analysieren und zu interpretieren. Dies kann sowohl existente als auch gelöschte Daten umfassen. Die gesammelten Daten können eine breite Palette von Informationen umfassen, einschließlich Chat-Nachrichten, E-Mails, Dokumente, Bilder, Videos, Suchverläufe und vieles mehr.

Einer der Schlüsselbereiche der digitalen Forensik ist die Untersuchung von gespeicherten digitalen Daten. Hierbei können Experten tiefer in das Betriebssystem des Geräts eindringen, um versteckte oder gelöschte Daten zu finden. In vielen Fällen kann diese Art von Untersuchung dazu führen, dass inkriminierte Daten gefunden werden, die sonst übersehen worden wären.

Auch die sozialen Medien spielen in der digitalen Forensik eine immer größere Rolle. Viele Menschen hinterlassen umfangreiche digitale Fußabdrücke in sozialen Netzwerken, und diese Daten können am Ende oft aufschlussreiche Beweise liefern. Dazu können zum Beispiel Posts, Fotos, Standortdaten und Freundeslisten gehören.

2.1 Erkenntnisse aus Datenselektion und Methoden der Risikoanalyse

Die Auswertung digitaler Beweise kann helfen, das Risikoprofil eines Sexualstraftäters im Netz zu bestimmen, indem sie Informationen über das Verhalten, die Interessen, die Kontakte und die Absichten des Täters oder der Täterin liefert:

Suchverhalten: Wenn beispielsweise ein Individuum nach spezifischen, potenziell problematischen Themen oder Inhalten sucht, könnte dies auf ein erhöhtes Risiko hindeuten. Dies könnte mit einer klinischen Beurteilung korrelieren, die ähnliche Interessen oder Neigungen identifiziert hat. Bestimmte Suchbegriffe und besuchte Internetseiten könnten auf ein Interesse an sexuell devianten Inhalten hinweisen und Hinweise auf seine Absichten oder Präferenzen liefern. Zum Beispiel könnte ein Sexualstraftäter nach Informationen suchen, die mit seinen Straftaten in Zusammenhang stehen oder diese erleichtern könnten.

Dateien und Multimedia: Die Anwesenheit bestimmter Arten von Dateien, wie z.B. pornografisches Material, könnte als Risikofaktor angesehen werden. Darüber hinaus könnte die forensische Analyse solcher Dateien zusätzliche Informationen liefern, wie z.B. den Zeitpunkt des Herunterladens oder die Quelle der Dateien.

Missbrauchsabbildungen sind digitale Darstellungen von sexuellen Handlungen mit oder an Kindern oder Jugendlichen. Sie sind nach § 184b StGB strafbar und gelten als schwere Form der sexualisierten Gewalt gegen Kinder und Jugendliche. Missbrauchsabbildungen können sowohl als Beweismittel für eine begangene oder geplante Sexualstraftat dienen, als auch als Auslöser oder Verstärker für eine solche Straftat wirken.

Soziale Medien und Kommunikation: Verhaltensmuster in sozialen Medien und digitaler Kommunikation könnten Aufschluss über die Einstellungen und Überzeugungen eines Individuums geben. Beispielsweise könnten aggressive oder manipulative Verhaltensweisen in diesen Kontexten auf ein erhöhtes Risiko hindeuten. Auf der anderen Seite kann eine geringe Aktivität in sozialen Medien oder Mangel an Kommunikation mit anderen auf soziale Isolation hinweisen.

Standortdaten: Standortdaten könnten Aufschluss darüber geben, wo sich ein Individuum aufhält und welche Orte es regelmäßig besucht. Wenn beispielsweise ein Individuum häufig Orte besucht, die mit früheren Straftaten in Zusammenhang stehen, könnte dies auf ein erhöhtes Risiko hindeuten.

App-Nutzung: Die Nutzung bestimmter Arten von Apps könnte ebenfalls auf ein erhöhtes Risiko hindeuten. Beispielsweise könnte die häufige Nutzung von Apps, die den Zugang zu potenziell problematischen Inhalten ermöglichen, ein Risikoindikator sein.

Eine manuelle Analyse ist dabei sehr zeitaufwendig, mühsam und fehleranfällig. Deshalb ist es sinnvoll, Automatisierungsmethoden zu verwenden, die diese Aufgabe übernehmen oder in der Vorselektierung unterstützen können. Informationen können dabei schneller und effizienter extrahiert, gefiltert und interpretiert werden.

2.2 Automatisierung der Datenselektion

Vor allem die Analyse von Mediendateien und Kommunikationsspuren sowie deren Interpretation haben bereits etablierte Möglichkeiten der Automatisierung, die eine Selektion deutlich vereinfachen. Timeline-Analysen helfen bei einer zeitlichen Einschätzung von Taten und Standortdaten bringen die Ergebnisse und Ereignisse in einen räumlichen Kontext.

Klassifikation von Bildern

Die automatisierte Auswertung von Bildern erfordert spezielle Algorithmen und Techniken, die auf maschinellem Lernen und Deep learning basieren, um die Bilder zu verarbeiten

und zu interpretieren. Einige Beispiele für die automatisierte Auswertung von Bildern mit verschiedenen Algorithmen sind:

Das *PhotoDNA-Verfahren*, das von Microsoft entwickelt wurde und einen Hash-Wert für jedes Bild erzeugt, der als digitaler Fingerabdruck dient. Dieser Hash-Wert kann mit einer Datenbank von bekannten illegalen Bildern verglichen werden, um diese schnell zu identifizieren [Po21].

Das *C4All-Verfahren*, das vom Bundeskriminalamt (BKA) entwickelt wurde und eine Kombination aus verschiedenen Algorithmen nutzt, um Bilder nach verschiedenen Kriterien zu filtern und zu kategorisieren. Diese Kriterien sind zum Beispiel die Anzahl der Personen auf dem Bild, das Alter der Personen, die Art der sexuellen Handlung oder die Qualität des Bildes [FG16].

Das *FaceMatch-Verfahren*, das vom National Center for Missing & Exploited Children (NCMEC) entwickelt wurde und eine Gesichtserkennungssoftware nutzt, um Opfer auf verschiedenen Bildern wiederzuerkennen und zu identifizieren. Dieses Verfahren kann auch helfen, neue Opfer zu entdecken oder bekannte Opfer ausfindig zu machen [Ba21].

2.3 Analyse von Texten

Die Textanalyse, auch als Text Mining bekannt, ist eine Methode, die bei der Auswertung von Beweismitteln in strafrechtlichen Untersuchungen eingesetzt wird, um Muster und Trends in Textdaten zu erkennen. In Bezug auf Online-Sexualstraftäter können Methoden des Text Mining verwendet werden, um Kommunikationsmuster in Chat-Nachrichten, E-Mails und anderen Textnachrichten zu analysieren.

Es gibt verschiedene Techniken, die im Text Mining verwendet werden, einschließlich, aber nicht beschränkt auf:

Sentiment-Analyse: Diese Methode wird verwendet, um die Gefühlslage oder den Ton eines Textes zu bestimmen. Sie ist ein Teilgebiet der Textanalyse, das Methoden des Natural Language Processing und des maschinellen Lernens verwendet. Bei der Untersuchung von Internet-Sexualstraftätern könnte dies verwendet werden, um die Absichten oder Motivationen des Täters besser zu verstehen [ZM16].

Themenmodellierung: Hierbei handelt es sich um eine statistische Methode zur Entdeckung von abstrakten "Themen", die in einer Sammlung von Dokumenten oder Textnachrichten auftreten. Dies kann verwendet werden, um häufige Themen oder Muster in der Kommunikation eines Täters zu erkennen [ZM16].

Entity Recognition: Diese Methode identifiziert und klassifiziert Namen von Personen, Organisationen, Orten, Ausdrücken von Zeiten, Mengen, prozentualen Angaben, usw. in einem

Text. Diese Technik kann dazu beitragen, Schlüsselinformationen in den Kommunikationen eines Täters zu erkennen [Su18].

2.4 Digitale Timeline-Analyse

Die digitale Timeline-Analyse ist ein Verfahren, das die zeitliche Abfolge von Ereignissen auf einem Datenträger oder einem Computersystem rekonstruiert und visualisiert. Die digitale Timeline-Analyse kann zum Beispiel verwendet werden, um die angegebenen oder vermuteten Aktivitäten eines Beschuldigten oder eines Opfers vor, während oder nach einer Straftat nachzuweisen oder zu widerlegen. Die digitale Timeline-Analyse basiert auf verschiedenen Arten von Zeitstempeln, die von Betriebssystemen, Anwendungen oder Dateisystemen erzeugt werden. Diese Zeitstempel können Informationen über die Erstellung, Änderung, Zugriff oder Löschung von Dateien oder Ordnern enthalten. Die digitale Timeline-Analyse kann mit speziellen Software-Tools durchgeführt werden, die die Zeitstempel auslesen, filtern, sortieren und darstellen können [LD23].

2.5 Geografische Analyse von Tatorten und Täterverhalten

Geografisches Profiling ist die praktische Anwendung von verschiedenen geographischen, kriminologischen und psychologischen Prinzipien, um anhand von extrahierten Geokoordinaten Aussagen über das Verhalten einer Person treffen zu können. Obwohl Geoprofiling nicht fest in empirisch getesteten Theorien verankert ist, nutzt es mehrere theoretische Ansätze: die Crime Pattern Theory [SRS21], den Routine Activity Approach [Ti13] und die Rational Choice Perspective [BPR05; GVW14; SK11].

Geo-Profiling wird als Analysemethode vor allem bei Sexualstraftätern angewandt, die mehrere Opfer ohne Vorbeziehung überfallen. Die Analyse der Tatorte kann dabei Hinweise auf die Persönlichkeit, die Motivation und das Verhalten des Täters liefern [Ca13; GVW14].

3 Vorteile und Herausforderungen bei der automatisierten Risikoanalyse

Automatisierung in der Datenanalyse bietet zahlreiche Vorteile, darunter die Steigerung der Effizienz und Geschwindigkeit der Datenanalyse, die Erschließung neuer Erkenntnisse aus bisher ungenutzten oder verborgenen Daten, die Verbesserung der Qualität und Genauigkeit der Analyse sowie die Unterstützung von Entscheidungsprozessen und Handlungsoptionen.

Allerdings gibt es auch Herausforderungen zu bewältigen, wie den Schutz der Privatsphäre und der Persönlichkeitsrechte der Betroffenen, die Sicherstellung der Rechtmäßigkeit und Transparenz der Datenverarbeitung, die Gewährleistung der Zuverlässigkeit und

Nachvollziehbarkeit der Datenanalyse sowie die Vermeidung von Diskriminierung oder Verzerrung durch die Datenanalyse.

Es ist wichtig zu beachten, dass die digitale Forensik wohldefinierten rechtlichen und ethischen Anforderungen unterliegt. Dazu gehören neben den Fragen des Datenschutzes und der Privatsphäre, auch die Notwendigkeit, Beweise auf eine Art und Weise zu analysieren, die ihre Integrität für die Verwendung in Gerichtsverfahren gewährleistet.

Die Analyse von Datenträgern, die als Beweismittel in amtliche Verwahrung genommen werden, unterliegt dem Grundsatz der Verhältnismäßigkeit und dem Schutz des informationellen Selbstbestimmungsrechts der Betroffenen. Das bedeutet, dass die Analyse nur dann zulässig ist, wenn sie für die Erfüllung einer gesetzlichen Aufgabe erforderlich ist und die schutzwürdigen Interessen der Betroffenen nicht überwiegen. Außerdem muss die Analyse auf das erforderliche Maß beschränkt werden und darf nicht zu einer unzulässigen Verknüpfung oder Auswertung von personenbezogenen Daten führen [Cz17].

Wenn einer der Kommunikationspartner des Beschuldigten eine zeugnisverweigerungs-berechtigte Person ist, wie zum Beispiel ein Ehegatte oder ein naher Verwandter, kann sich daraus ein Verwertungsverbot ergeben. Das heißt, dass die aus den Chatverläufen gewonnenen Erkenntnisse nicht als Beweismittel gegen den Beschuldigten oder den Kommunikationspartner verwendet werden dürfen. Der Schutz der Beschuldigten bezüglich der Auswertung von digitalen Asservaten ist also nicht absolut, sondern abhängig von den Umständen des Einzelfalls und den gesetzlichen Vorgaben [Bu20].

Die DSGVO regelt den Umgang mit personenbezogenen Daten, also Daten, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen. Dazu können auch digitale Asservate gehören, wie zum Beispiel Chatverläufe auf Mobiltelefonen, die im Rahmen einer strafrechtlichen Ermittlung beschlagnahmt werden. Die DSGVO gilt grundsätzlich auch für die Verarbeitung von personenbezogenen Daten durch Behörden wie die Polizei oder die Staatsanwaltschaft, es sei denn, es gibt nationale Gesetze, die spezielle Regelungen für diesen Bereich vorsehen. In Deutschland ist das zum Beispiel das Bundesdatenschutzgesetz (BDSG), das in den §§ 45 bis 84 BDSG besondere Vorschriften für die Verarbeitung von personenbezogenen Daten für Zwecke der Strafverfolgung enthält [VV18].

Der Umgang mit digitalen Asservaten muss also sowohl den Anforderungen der DSGVO als auch den Anforderungen des BDSG entsprechen. Das bedeutet unter anderem, dass die Verarbeitung von personenbezogenen Daten nur dann zulässig ist, wenn sie für die Erfüllung einer Aufgabe erforderlich ist, die im öffentlichen Interesse liegt oder in Ausübung öffentlicher Gewalt erfolgt (Art. 6 Abs. 1 lit. e DSGVO). Außerdem müssen die Grundsätze der Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz, Zweckbindung, Datenminimierung, Richtigkeit, Speicherbegrenzung, Integrität und Vertraulichkeit beachtet werden (Art. 5 DSGVO). Die betroffenen Personen haben grundsätzlich das Recht auf Auskunft, Berichtigung, Löschung, Einschränkung der Verarbeitung und Widerspruch gegen die Verarbeitung ihrer personenbezogenen Daten (Art. 15 bis 21 DSGVO). Allerdings können

diese Rechte eingeschränkt werden, wenn dies zur Wahrung von wichtigen öffentlichen Interessen wie der Strafverfolgung erforderlich ist (Art. 23 DSGVO). Die Verantwortlichen für die Verarbeitung von personenbezogenen Daten müssen zudem geeignete technische und organisatorische Maßnahmen treffen, um ein dem Risiko angemessenes Schutzniveau zu gewährleisten (Art. 24 und 32 DSGVO). Sie müssen auch eine Datenschutz-Folgenabschätzung durchführen, wenn die Verarbeitung ein hohes Risiko für die Rechte und Freiheiten der betroffenen Personen birgt (Art. 35 DSGVO) [VV18].

4 Schlussfolgerungen

Zusammenfassend lässt sich sagen, dass die digitale Forensik und die Risikoanalyse bei Sexualstraftaten im Internet wichtige Werkzeuge zur Aufklärung von Straftaten und zum Schutz von Opfern sind. Die automatisierten Kategorisierungsmechanismen können die Artefakte nach verschiedenen Kriterien wie Sprache, Inhalt, Genre, Thema oder Stimmung sortieren und so einen Teil der Auswertung übernehmen. Die Strafverfolgung kann von einer automatisierten Risikoanalyse von Sexualstraftätern profitieren, aber sie kann diese nicht vollständig ersetzen.

Auf der einen Seite kann die automatisierte Risikoanalyse dazu beitragen, schneller und effizienter Straftäter mit einem hohen Risiko zu identifizieren und ein schnelleres Eingreifen zu ermöglichen. Klinische Risikoanalysetests könnten mit zusätzlichen Informationen unterlegt werden um die Informationsdichte zu erhöhen und so eine bessere Beurteilung erst möglich zu machen. Dazu kann künstliche Intelligenz einen wesentlichen Beitrag leisten in dem Semantik aus Bildern und Texten zusammengeführt werden und andere, sonst unerkannte, Zusammenhänge aufgedeckt werden.

Weiterhin kann eine Automatisierung die Arbeitsbelastung und den Zeitaufwand von Experten verringern, die große Mengen digitaler Beweise analysieren müssen. Dabei kann sie den Schutz von Opfern und potenziellen Opfern verbessern, indem sie weiteren Schaden verhindert oder verringert.

Auf der anderen Seite kann es zu falsch-positiven oder falsch-negativen Ergebnissen kommen, was zu falschen Anschuldigungen oder verpassten Gelegenheiten zum Eingreifen führt. Grundsätzlich sollte ein Experte die Ergebnisse überprüfen, da Sexualstraftaten ein wichtiges und sensibles Thema sind. Die Automatisierung einzelner Teilgebiete ist daher eine nützliche Hilfe, um die Artefakte vorzuselektieren und die Auswertungszeit zu verkürzen.

Literatur

- [18] Sexualstraftat – Schreibung, Definition, Bedeutung, Beispiele, de, Apr. 2018, URL: <https://www.dwds.de/wb/Sexualstraftat>, Stand: 20. 05. 2023.
- [AV03] Arkowitz, S.; Vess, J.: An Evaluation of the Bumby RAPE and MOLEST Scales as Measures of Cognitive Distortions With Civilly Committed Sexual Offenders. en, *Sexual Abuse* 15/4, S. 237–249, Okt. 2003, ISSN: 1079-0632, 1573-286X, URL: <http://journals.sagepub.com/doi/10.1177/107906320301500402>, Stand: 31. 05. 2023.
- [Ba21] Banaschak, S.; Ohlogge, S.; Froch-Cortis, J.; Mayer, F.: Grundlagen der Begutachtung von kinderpornografischen Abbildungen in der Rechtsmedizin. de, *Rechtsmedizin* 31/5, S. 463–473, Okt. 2021, ISSN: 0937-9819, 1434-5196, URL: <https://link.springer.com/10.1007/s00194-021-00503-7>, Stand: 20. 05. 2023.
- [BBH21] Brankley, A. E.; Babchishin, K. M.; Hanson, R. K.: STABLE-2007 Demonstrates Predictive and Incremental Validity in Assessing Risk-Relevant Propensities for Sexual Offending: A Meta-Analysis. en, *Sexual Abuse* 33/1, S. 34–62, Feb. 2021, ISSN: 1079-0632, 1573-286X, URL: <http://journals.sagepub.com/doi/10.1177/1079063219871572>, Stand: 31. 05. 2023.
- [BPR05] Beauregard, E.; Proulx, J.; Rossmo, D. K.: Spatial patterns of sex offenders: Theoretical, empirical, and practical issues. en, *Aggression and Violent Behavior* 10/5, S. 579–603, Juli 2005, ISSN: 13591789, URL: <https://linkinghub.elsevier.com/retrieve/pii/S1359178905000030>, Stand: 24. 05. 2023.
- [Bu20] Bundestag: Die Auswertung von auf digitalen Endgeräten gespeicherten Chatverläufen in strafrechtlichen Ermittlungsverfahren Strafverfahrensrechtliche Rahmenbedingungen unter besonderer Berücksichtigung von Zeugnisverweigerungsrechten, *Techn. Ber.*, Jan. 2020, URL: <https://www.bundestag.de/resource/blob/796960/502f104acf21e11c88bf054cda3e0b32/WD-7-096-20-pdf-data.pdf>, Stand: 20. 05. 2023.
- [Ca13] Canter, D.; Hammond, L.; Youngs, D.; Juszczak, P.: The Efficacy of Ideographic Models for Geographical Offender Profiling. en, *Journal of Quantitative Criminology* 29/3, S. 423–446, Sep. 2013, ISSN: 0748-4518, 1573-7799, URL: <http://link.springer.com/10.1007/s10940-012-9186-6>, Stand: 20. 05. 2023.
- [Cz17] Czerner, F.: Digitale Forensik zwischen (Online-)Durchsuchung, Beschlagnahme und Datenschutz. In (Labudde, D.; Spranger, M., Hrsg.): *Forensik in der digitalen Welt*. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 265–300, 2017, ISBN: 9783662538005 9783662538012, URL: http://link.springer.com/10.1007/978-3-662-53801-2_10, Stand: 20. 05. 2023.

- [FG16] Franke, I.; Graf, M.: Kinderpornografie: Übersicht und aktuelle Entwicklungen. de, Forensische Psychiatrie, Psychologie, Kriminologie 10/2, S. 87–97, Mai 2016, ISSN: 1862-7072, 1862-7080, URL: <http://link.springer.com/10.1007/s11757-016-0361-8>, Stand: 20. 05. 2023.
- [GL16] Giguère, G.; Lussier, P.: Debunking the psychometric properties of the LS\CMI: An application of item response theory with a risk assessment instrument. en, Journal of Criminal Justice 46/, S. 207–218, Sep. 2016, ISSN: 00472352, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0047235216300393>, Stand: 31. 05. 2023.
- [GVW14] Goodwill, A. M.; Van Der Kemp, J. J.; Winter, J. M.: Applied Geographical Profiling. In (Bruinsma, G.; Weisburd, D., Hrsg.): Encyclopedia of Criminology and Criminal Justice. Springer New York, New York, NY, S. 86–99, 2014, ISBN: 9781461456896 9781461456902, URL: https://link.springer.com/10.1007/978-1-4614-5690-2_207, Stand: 20. 05. 2023.
- [GZ13] Göbbels, S.; Zimmermann, L.: Rehabilitation von Straftätern: Das „Risk-need-responsivity“-Modell. de, Forensische Psychiatrie, Psychologie, Kriminologie 7/1, S. 12–21, Feb. 2013, ISSN: 1862-7072, 1862-7080, URL: <http://link.springer.com/10.1007/s11757-012-0199-7>, Stand: 01. 06. 2023.
- [HB20] Hart, S. D.; Boer, D. P.: Structured Professional Judgment Guidelines for Sexual Violence Risk Assessment. In (Douglas, K. S.; Otto, R. K., Hrsg.): Handbook of Violence Risk Assessment. 2. Aufl., Routledge, Second edition. | New York, NY : Routledge, 2021. | Series: International perspectives on forensic mental health, S. 322–358, Dez. 2020, ISBN: 978-1-315-51837-4, URL: <https://www.taylorfrancis.com/books/9781315518367/chapters/10.4324/9781315518374-18>, Stand: 20. 05. 2023.
- [HH20] Hörburger, T. A.; Habermeyer, E.: Zu den Zusammenhängen zwischen paraphilen Störungen, Persönlichkeitsstörungen und Sexualdelinquenz. de, Forensische Psychiatrie, Psychologie, Kriminologie 14/2, S. 149–157, Mai 2020, ISSN: 1862-7072, 1862-7080, URL: <http://link.springer.com/10.1007/s11757-020-00597-x>, Stand: 31. 05. 2023.
- [HMH06] Hanson, R. K.; Morton, K. E.; Harris, A. J. R.: Sexual Offender Recidivism Risk: What We Know and What We Need to Know. en, Annals of the New York Academy of Sciences 989/1, S. 154–166, Jan. 2006, ISSN: 00778923, 17496632, URL: <https://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2003.tb07303.x>, Stand: 31. 05. 2023.
- [Ho20] Hollerbach, P.; Habermeyer, E.; Nitschke, J.; Sünkel, Z.; Mokros, A.: Construct Validity of the German Version of the Hare Psychopathy Checklist – Revised. en, European Journal of Psychological Assessment 36/5, S. 805–816, Sep. 2020, ISSN: 1015-5759, 2151-2426, URL: <https://econtent.hogrefe.com/doi/10.1027/1015-5759/a000566>, Stand: 31. 05. 2023.

- [KB20] Kube, S.; Banse, R.: Literaturübersicht zur prädiktiven Validität des Static-99 im deutschsprachigen Raum. de, *Forensische Psychiatrie, Psychologie, Kriminologie* 14/3, S. 300–314, Aug. 2020, ISSN: 1862-7072, 1862-7080, URL: <https://link.springer.com/10.1007/s11757-020-00616-x>, Stand: 20. 05. 2023.
- [Kr09] Kröber, H.-L.: Kriminalprognose bei Sexualstraftätern. de, *Psychotherapeut* 54/4, S. 237–244, Juli 2009, ISSN: 0935-6185, 1432-2080, URL: <http://link.springer.com/10.1007/s00278-009-0676-4>, Stand: 20. 05. 2023.
- [LD23] Lyublinskaya, I.; Du, X.: Annotated digital timelining: Interactive visual display for data analysis in mixed methods research. en, *Methods in Psychology* 8/, S. 100108, Nov. 2023, ISSN: 25902601, URL: <https://linkinghub.elsevier.com/retrieve/pii/S2590260122000194>, Stand: 20. 05. 2023.
- [Po21] Povalej, R.; Rittelmeier, H.; Fähndrich, J.; Berner, S.; Honekamp, W.; Labudde, D.: Die Enkel von Locard: Analyse digitaler Spuren in der forensischen Informatik. de, *Informatik Spektrum* 44/5, S. 355–363, Okt. 2021, ISSN: 0170-6012, 1432-122X, URL: <https://link.springer.com/10.1007/s00287-021-01393-5>, Stand: 20. 05. 2023.
- [RE07] Rettenberger, M.; Eher, R.: Aktuarische Kriminalprognosemethoden und Sexualdelinquenz: Die deutsche Version des SORAG. en, *Monatsschrift für Kriminologie und Strafrechtsreform* 90/6, S. 484–497, Dez. 2007, ISSN: 2366-1968, 0026-9301, URL: <https://www.degruyter.com/document/doi/10.1515/mks-2007-900603/html>, Stand: 20. 05. 2023.
- [RT07] Robertiello, G.; Terry, K. J.: Can we profile sex offenders? A review of sex offender typologies. en, *Aggression and Violent Behavior* 12/5, S. 508–518, Sep. 2007, ISSN: 13591789, URL: <https://linkinghub.elsevier.com/retrieve/pii/S1359178907000262>, Stand: 20. 05. 2023.
- [SK11] Schubert, K.; Klein, M.: *Das Politiklexikon: Begriffe, Fakten, Zusammenhänge*. Bundeszentrale für Politische Bildung, Bonn, 2011, ISBN: 9783838901749.
- [SKB11] Seto, M. C.; Karl Hanson, R.; Babchishin, K. M.: Contact Sexual Offending by Men With Online Sexual Offenses. en, *Sexual Abuse* 23/1, S. 124–145, März 2011, ISSN: 1079-0632, 1573-286X, URL: <http://journals.sagepub.com/doi/10.1177/1079063210369013>, Stand: 20. 05. 2023.
- [SRS21] van Sleeuwen, S. E. M.; Ruiter, S.; Steenbeek, W.: Right place, right time? Making crime pattern theory time-specific. *Crime Science* 10/1, S. 2, Jan. 2021, ISSN: 2193-7680, URL: <https://doi.org/10.1186/s40163-021-00139-8>, Stand: 24. 05. 2023.
- [St06] Stadtland, C.; Hollweg, M.; Kleindienst, N.; Dietl, J.; Reich, U.; Nedopil, N.: Rückfallprognosen bei Sexualstraftätern — Vergleich der prädiktiven Validität von Prognoseinstrumenten. de, *Der Nervenarzt* 77/5, S. 587–595, Mai 2006, ISSN: 0028-2804, 1433-0407, URL: <http://link.springer.com/10.1007/s00115-005-1945-2>, Stand: 20. 05. 2023.

- [Su18] Sun, P.; Yang, X.; Zhao, X.; Wang, Z.: An Overview of Named Entity Recognition. In: 2018 International Conference on Asian Language Processing (IALP). IEEE, Bandung, Indonesia, S. 273–278, Nov. 2018, ISBN: 978-1-72811-175-9, URL: <https://ieeexplore.ieee.org/document/8629225/>, Stand: 31.05.2023.
- [Ti13] Tilley, N., Hrsg.: Handbook of Crime Prevention and Community Safety. Willan, 2013, ISBN: 9781134014637.
- [VV18] Voigt, P.; Von Dem Bussche, A.: EU-Datenschutz-Grundverordnung (DSGVO). Springer Berlin Heidelberg, Berlin, Heidelberg, 2018, ISBN: 9783662561867 9783662561874.
- [ZM16] Zhai, C.; Massung, S.: Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining. Association for Computing Machinery und Morgan & Claypool, New York, NY, USA, 2016, ISBN: 9781970001174.

Requirements for a public digital forensics cloud

Martin Morgenstern¹ and Wilfried Honekamp ²

Abstract: The acquisition of digital evidence in criminal proceedings has become considerably more important in recent years. At the same time, the amount of data has also increased. The need to use cloud or big data solutions for digital forensics to be able to efficiently process the permanently increasing amount of data and number of cases has been recognised for years. By using public cloud providers, such as Amazon AWS and Microsoft Azure, to secure and analyse digital evidence, resources could be used in a scalable and flexible way. Forensic service providers have had to keep a large number of data carriers for forensic backups because they have to be available immediately in case of an emergency and cannot be procured only when needed.

Keywords: Digital Forensics, cloud, big data, digital evidence

1 Introduction

Due to the rapid increase in data quantity, source and diversity, an efficient evaluation of digital evidence with traditional technologies and methods is often no longer possible [TB16]. Leimbach and Bachlechner state that there is no exact limit to when a quantity of data becomes big data. They say that big data means a quantity of data that is so large that traditional technologies and methods for data processing and analysis can no longer be used efficiently. In addition to the amount of data, processing speed and heterogeneity are also decisive for the classification of data as big data [TL14, OD23]. Although the use of big data can make the analysis of unstructured mass data much more efficient, large amounts of data must first be backed up and then transferred to the big data environment. From the authors' point of view, current methods, such as the transport of data carriers, are not a practicable solution. An alternative to transporting data carriers can be the use of cloud services. Cloud computing is not to be confused with a simple classic client-server application. The National Institute of Standards and Technology has defined five characteristics of cloud services. These are self-service, access via a network, sharing of resources, rapid scalability and measurability of IT resources [TL14]. In practice, big data and cloud technologies can be combined. Long-term planning, storage and provisioning of data media would not be necessary when using public cloud providers, as the necessary resources can be scaled up in a few seconds. However, the use of public cloud providers

¹ Hochschule Stralsund, IACS, Zur Schwedenschanze 15, Stralsund, 18435, martin.morgenstern@hochschule-stralsund.de

² German Police University (DHPol), Institute for Police Technology (PTI), Zum Roten Berge 18-24 Münster, 48165, wilfried.honekamp@dhpol.de,  <https://orcid.org/0000-0003-2931-7047>

for securing and analysing digital evidence seems to have been virtually ruled out in professional circles thus far due to legally unresolved issues or other reservations [TB16] [Ri22]. One conceivable solution could be the development of cloud-in-cloud software, including corresponding upload clients. The solution to be developed would have to implement all the requirements that still have to be collected to preserve the forensic usability of digital traces. For an efficient analysis of digital evidence, it is necessary to be able to react to changed requirements at short notice. This is currently a great challenge for the police for reasons of public procurement law [Ho23]. By using a public cloud infrastructure, the current need for short-term adaptability of analysis environments can be realised in a short time.

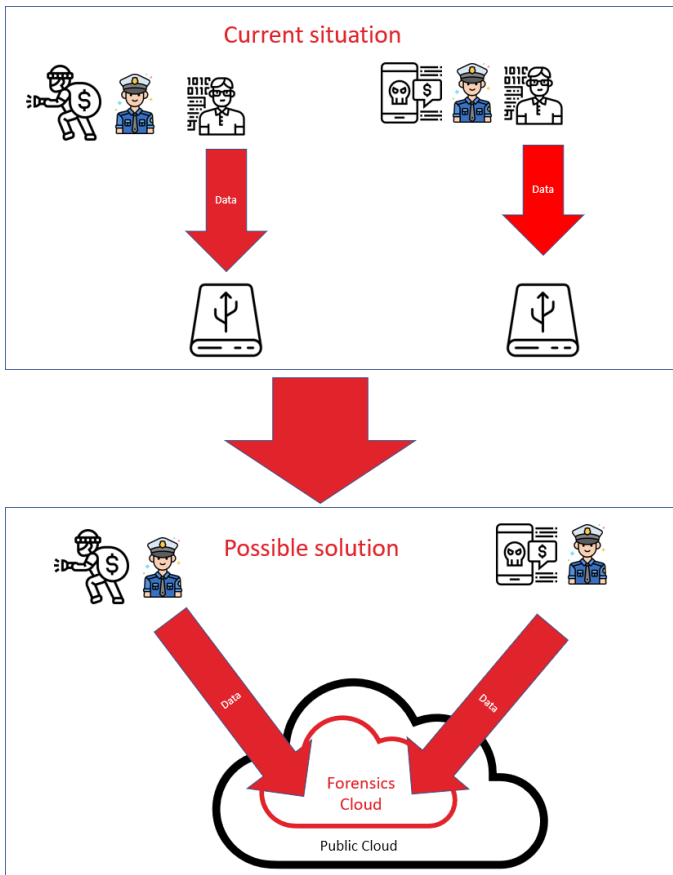


Fig. 1: Idea of a forensics cloud
 (Icons made by RaftelDesign, Flat Icons, and Eucalypt from flaticons.com)

The need for digital forensics is also constantly increasing outside of law enforcement agencies. For example, in January 2023, the National Association of Statutory Health

Insurance Physicians invited tenders for incident response and digital forensics services with an estimated value of 500,000 euros for a period of 48 months. Digital forensics has developed into an important business area for private companies in recent years. Leading auditing firms as well as various specialised service providers in a wide range of company sizes offer computer forensic services [KB22].

Due to the increasing number of digital forensics experts, it is foreseeable that the number of private forensics service providers will increase in the coming years. However, especially for small forensics companies, the procurement of powerful hardware and software is a major hurdle. By providing cloud solutions for digital forensics that include not only analysis tools but almost the entire infrastructure needed, starting with the forensic backup of data, the entry hurdle into self-employment for digital forensic specialists can be significantly lowered [TB16].

There are many special fields in digital forensics, the number of which is constantly increasing. In addition to the quantity of data, the number of possible data sources is also increasing annually. This has led to smaller forensic service providers, in particular, often specialising in one area of digital forensics. These can be specialists for the analysis of certain types of data, e.g., multimedia forensics but also specialists for specific data sources such as motor vehicle forensics [MFH22] [Ha21] [Bö09].

To date, the securing of digital evidence has largely been done by forensic experts. However, this means that these experts must be available at all times. Spontaneous digital preservation of digital traces without digital forensic experts is therefore not possible at the present time. In this contribution, we analyse the requirements of a centrally available forensics cloud as the primary storage medium for digital evidence, to which all police officers entrusted with the case have access, and whether such a forensic cloud could lead to an increase in the efficiency of police work. The public digital forensics cloud to be established and introduced in Figure 1 offers a provider-independent possibility for collaboration between a large number of forensic experts. Since no requirements for such a cloud have been published yet, these are to be determined by expert knowledge acquisition.

2 Method

To capture the requirements for a forensics cloud, an expert survey was conducted using a structured questionnaire [Mo23]. The survey contained both open and closed questions. Closed questions were evaluated quantitatively, and open questions were evaluated with a qualitative content analysis according to Mayring [Ma22]. In some cases, additional information was provided for the closed questions, which was then also evaluated qualitatively.

In May 2023, four practitioners of digital forensics with several years of professional experience were interviewed as experts. Expert 1 is head of a digital forensics unit in a

public security authority, has many years of experience in the field of IT forensics and is a publicly appointed expert in this field. Expert 2 has a professorship in IT forensics and works as a publicly appointed expert for digital forensics. Expert 3 works as an analyst and scientific coordinator for digital forensics in a state criminal investigation office. Expert 4 works as a course instructor for computers and cybercrime at a police training institution and has practical experience in securing digital traces.

A category system was developed analogous to [He14] for the analysis of the answers. It is essentially analogous to the questions asked.

Category	Name	Description
C 1	Qualification	This category records the relevant professional qualification of the participant.
C 1.1	Education	Contains the participant's vocational training.
C 1.2	Work experience	Contains the relevant professional experience of the participant.
C 2	Necessity as primary storage medium	The category analyses the need for a forensics cloud.
C 2.1	Lack of resources as a reason for not preserving evidence	It was asked whether there were any known cases in which digital evidence was not secured due to a lack of resources or skilled personnel.
C 2.2	Crime clearance rate	The experts were asked to assess whether more crimes could be solved if nontechnical personnel were enabled to secure digital traces in a spontaneous spontaneously, practicably and in a way that can be used in court.
C 3	Cooperation	
C 3.1	Development to date Specialisation	The experts were asked to indicate how, in their view, the role of specialisation has changed in recent years.
C 3.2	Importance of cooperation	The experts were asked about their experience of the role that cooperation between various governmental and nongovernmental experts already plays today.
C 3.3	Expectation of an increase in specialisation	It was asked whether an increase in specialisation is expected in the next few years.

C 4	Other	
C 4.1	Expected added value	The experts were asked to state whether, from their point of view, further added value that was not addressed in the previous questions can be expected from the planned forensics cloud.
C 4.2	Acceptance requirements	The experts were asked, what the important factors for the acceptance of a forensics cloud as the primary storage medium for criminally relevant data are.
C 4.3	Other comments	Finally, the participants were able to provide further comments on the topic.

Tab. 1: Categories for analysing

3 Results

This section presents the results of the evaluation of the interviews. The question of whether the participants were aware of cases in which digital traces were not secured due to a lack of resources was answered differently by all participants. One participant each answered this question with a yes or a no without further explanation. The other answers indicated that there were cases where traces were not saved. However, both experts wrote that the reasons for this were not a lack of technical equipment but mainly a lack of expertise on site. Expert 3 also stated that expertise available in other authorities may not always be shared due to official regulations. Another problem with new technologies is that the possibilities for data extraction have not yet been developed.

Another reason given for not securing evidence was poor preclearance, as well as simply not requesting IT forensics experts. Only one respondent believes that more crimes could be solved if nontechnical personnel were given the opportunity to secure data spontaneously, practically and in a court of law. All other respondents stated that at least some technical understanding is necessary to identify digital traces as such. Furthermore, one answer referred to the already existing nationwide information portal, where data can also be uploaded by citizens in case of special damage situations.

All participants confirmed that specialisation in IT forensics has increased in importance in recent years. In this context, two answers pointed out that despite specialisation, solid basic knowledge is necessary. In regard to specialisations in IT forensics, a distinction must be made between specialisations in data origin and data evaluation. The evaluation of data should be carried out independently of the data origin. Nevertheless, the type of possible data sources will continue to increase, which is why there will be an increase in specialisations here as well.

Due to specialisations but also for other reasons, cooperative work is already taking place between different IT forensics experts. In several federal states, state and nonstate investigators work together. This practice is not without controversy, as law enforcement is a sovereign task. An increase in specialisations is expected by all participants. Expert 1 added, "There will always be more specialisations, but they will require more and more cooperative collaboration for successful investigations. As things stand now, definitely also between governmental and nongovernmental forensic experts". The experts were asked what added value - not named in the previous questions - they expect from a vendor-independent forensics cloud as the primary storage medium for digital evidence. In one interview, this question was not answered. One expert stated that a corresponding solution is already being set up for law enforcement in Saxony.

Expert 1 sees a possible added value in only using the technology of large cloud providers to operate a private cloud for police investigative work with its own infrastructure. At present, there are still justified political reservations about the direct use of large cloud providers. These could also not be solved by using a cloud-in-variant solution.

Expert 3 sees cost savings as a possible advantage since any necessary licences could be used across authorities. Furthermore, the processing of the data can be more efficient due to the available computing power. The lack of transport of data carriers to analysts and investigators was also mentioned as an added value.

With regard to the implementation of necessary encryption measures, all experts stated that current and the most secure procedures should be used. One participant pointed out that even with current encryption procedures in a cloud-in-cloud solution, it cannot be ruled out that the respective cloud operator gains access to sensitive data. At the latest, when the data are processed, they must be decrypted in the working memory. Research is currently being conducted into encryption methods that should enable processing without decryption, and the first test variants already exist. However, it will be several years before these could be used productively, which is still uncertain at present.

Almost all participants mentioned high security levels as an acceptance requirement for a forensics cloud. These were described in more detail in individual responses. Security must not be based on the self-commitment of individual cloud providers, which is currently the case with many public cloud providers. Security measures must not be outsourced but must be reflected in the technical solution used. Further acceptance requirements are simple usability and a fast internet connection of the cloud.

Two of the experts used the opportunity to make further comments on the topic. Expert 1 thinks that it would make sense to think about whether it might not also be more economical to build up one's own infrastructure. From a certain size upwards, this could be irrelevant in terms of costs, but this would have to be calculated. Furthermore, the normal police staff would have to be put in a position to recognise when experts might be necessary to secure data. A solution that is to be developed must also be politically feasible, particularly when data are secured by laypersons, as it can then be assumed that not all available data are secured.

Expert 3 states that direct data extraction to a cloud environment seems difficult because most tools need a direct connection to the device. Furthermore, some tools would require special hardware that could not be virtualised. The possibility that all police officers would be able to back up digital data themselves could only possibly lead to a relief of the forensic specialists. In return, however, the amount of secured data would increase, which would lead to a higher time requirement for its evaluation. Securing digital traces is not the everyday task of police officers. Therefore, a significantly higher error rate and increased probability of loss of evidence must be expected if all police members secure data themselves.

4 Discussion and Conclusions

A forensics cloud that can be used by all police staff has the potential to significantly improve the efficiency of securing and evaluating digital evidence. An increase in efficiency could be realised in particular through the elimination of transport routes, as well as a possible immediate backup of data. However, the results of the survey showed that when digital evidence is secured by nonforensic experts, it can be assumed that backups are not always carried out completely and professionally. In the worst case, there is even a fear of a break in the chain of custody. One of the main arguments against securing digital evidence by nonforensic experts is that they will overlook more digital traces than IT forensic experts do.

However, the problem listed is not entirely new. One result of the survey was that IT forensic experts were often not on site because the need for these experts was not recognised. The overlooking of digital traces in these cases is independent of whether the criminally relevant data are backed up in a cloud or on a local data carrier. A logical consequence of this experience should be to qualify all police officers to recognise digital traces as such. However, this qualification would require a basic understanding of digital traces [Ho23].

Another important aspect in connection with clouds is always the issue of security. Especially when processing data relevant to criminal law, a very high security standard must be guaranteed. Since direct storage and processing of data with a public cloud provider is considered too insecure by the authors, the basic idea to solve the problem is to develop a cloud-in-cloud solution. The approach should be that security measures can only be controlled by the owner of the forensics cloud. The security of the forensics cloud would need to be audited and certified by independent and trusted third parties. A residual risk of loss of confidentiality will never be completely excluded. At this point, a social discourse on what risks should be accepted would be appropriate.

The police infrastructure for the prosecution of criminal offences belongs to the critical infrastructure area. The status of Critical Infrastructure does not exclude that it is outsourced to external cloud providers. Following the general trend towards cloud computing, in recent years, operators of critical infrastructures have also begun to partially

outsource their computer resources and use them as a service. These are, for example, banks or companies from the energy supply sector [We20a] [We20b] [Kü23]. The question therefore arises whether higher requirements apply to the storage of digital evidence than to other critical infrastructure. In contrast to banks and energy service providers, the police belong to the authorities and organisations with security tasks (BOS). [BB23] Particularly in the area of the police, a leak of confidential information can pose a direct threat to people's lives. Regardless of the expected risks, however, it has been shown that provider-independent forensic clouds can represent added value. It therefore seems sensible to explore the topic further. In particular, research should be conducted into how the identified risks could be minimised. Since even internal police systems can never guarantee 100% security, a follow-up study could compare the security of police systems with that of a possible forensics cloud.

Bibliography

- [BB23] Bundesamt für Bevölkerungsschutz und Katastrophenhilfe (BBK): BOS, Behörden und Organisationen mit Sicherheitsaufgaben, <https://www.bbk.bund.de/SharedDocs/Glossareintraege/DE/B/BOS.html>, accessed: 30/06/2023.
- [Bö09] Böhme, R.; Freiling, F.; Gloe, T.; Kirchner, M: Multimedia-Forensik als Teildisziplin der digitalen Forensik. Presentation at the GI Workshop „Multimedia-Forensik“, Lübeck, 2009.
- [Kü23] Kümmerlein, K.: Ist die Cloud Teil der Energiewende? <https://www.cloudcomputing-insider.de/ist-die-cloud-teil-der-energiewende-a-7462505e5f33964c04d7133f41100c12/>, accessed: 30/06/2023.
- [He14] Helfferich, C.: Leitfaden- und Experteninterviews. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*. Springer Fachmedien, Wiesbaden, pp. 559–574, 2014.
- [Ho23] Honekamp, W.; Povalej, R., Rittelmeier, H., Fähndrich F.; Berner S.; Labudde, D.: Technologiegetriebene Polizeiausbildung im Umgang mit Digitalen Spuren. In: *Handbuch Cyberkriminalologie*. Springer Fachmedien, Wiesbaden. pp. 1-30, 2022.
- [KB22] Kassenärztliche Bundesvereinigung. Öffentliche Ausschreibung Berlin 2022 Rahmenvertrag digitale Forensik und Incident-Response. https://ausschreibungen-deutschland.de/1002019_Rahmenvertrag_digitale_Forensik_und_Incident-Response_2022_Berlin, accessed: 30/06/2023.
- [LB14] Leimbach, T.; Bachlechner, D.: Big Data in der Cloud. Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB), Berlin, 2014.
- [Ma22] Mayring, P.: *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Beltz, Weinheim, 2022.
- [MFH22] Morgenstern, M.; Fähndrich, F.; Honekamp, W.: Ontology in the Digital Forensics Domain: A Scoping Review. In: Demmler, D., Krupka, D. & Federrath, H. (Hrsg.),

INFORMATIK 2022. Gesellschaft für Informatik, Bonn. pp. 71-80, 2022.

- [Mo23] Morgenstern, M. Fragebogen zur Erfassung der Anforderungen an eine Forensik-Cloud. <https://paperswithcode.com/paper/fragebogen-zur-erfassung-der-anforderungen-an>, accessed: 11/07/2023.
- [OD23] Oracle Deutschland. Was versteht man unter Big Data? <https://www.oracle.com/de/big-data/what-is-big-data/> accessed: 11/07/2023.
- [Ri22] Rittelmeier, H.: Cloud-Konzepte für die digitale Forensik, Presentation, Interesting Times Forensics, Nüdlingen, 2022.
- [TB16] Tabona, O., Blyth, A.; A forensic cloud environment to address the big data challenge in digital forensics. In 2016 SAI Computing Conference (SAI), London, pp. 579-584, 2016.
- [We20a] Weidmann, T.: Solarisbank migriert als erste deutsche Bank vollständig in die Cloud, <https://www.it-finanzmagazin.de/solarisbank-arbeitet-als-erste-deutsche-bank-ausschliesslich-in-der-cloud-115060/>, accessed: 30/06/2023.
- [We20b] Weidmann, T.: Deutsche Bank geht in die Google Cloud: Warum der Schritt überfällig ist, <https://www.it-finanzmagazin.de/deutsche-bank-geht-in-die-google-cloud-warum-der-schritt-ueberfaellig-ist-115585/>, accessed: 30/06/2023.

Discussion of Basic Concepts of digital Forensics on the example of Volatility and Manipulability.

Johannes Fährndrich,¹ Lars Mechler²

Abstract: Digital forensics has many properties of evidence, which need to be defined formally. The goal of this paper is to formalize the properties of evidence with the application of digital forensics. The contribution is a discussion of the current definitions of volatility and manipulability. The analysis starts the discussion on the application of those concepts to the application of ontological definitions to digital forensics. With that, we want to discuss the trust and the need to understand complex AI models.

Keywords: cyber forensics, computer forensics

1 Introduction

The basic terms used in a scientific discussion need to be well-defined. Well-defined technical terms help to reduce misapprehension, which leads to different interpretation. Different interpretations lead to misunderstanding, which leads to different reasoning. With this argument, the basis of our scientific discussion is a common ground on the technical terms used. The technical terms are given by the definitions of the field of study. This idea of defining a category has been subject to research [Ka55]. The idea here is the discussion of objects themselves, their quality, quantity, modality, and relations.

Most definition of a chain of custody for digital forensics have similar steps [PS15]. For such a chain of custody, we analyze definitions of ontology and with that the definition of basics terms used in digital forensics.

The contribution of this paper is to collect the state-of-the-art of basic terms used in digital forensics and discuss their value. In cases where a new definition could be helpful, we suggest an updated formulation.

2 State of the Art

First, we look at the requirements of how a description of a definition is done in the academic discussion. Scholars have a discourse about a term of interest, and a prevailing opinion is

¹ Hochschule für Polizei Baden-Württemberg, Fachgruppe für Digitales Spuren und Künstliche Intelligenz, Sturmbülstr 250, 78054, Germany johannesfaehndrich@hfpol-bw.de

² Hochschule für Polizei Baden-Württemberg, Fachgruppe für Digitales Spuren und Künstliche Intelligenz, Sturmbülstr 250, 78054, Germany larsmechler@hfpol-bw.de

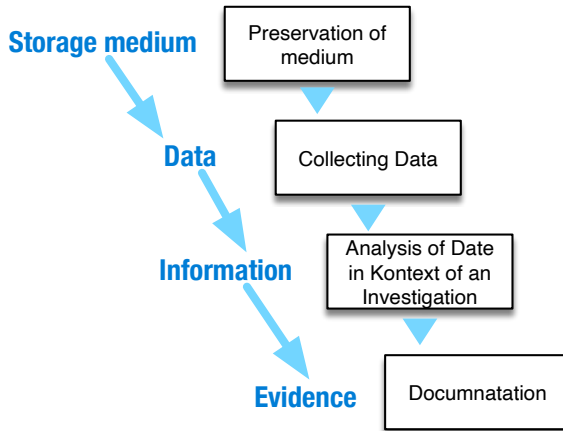


Abb. 1: Abstract description of the chain of custody for digital forensics.

reached. This process is repeated for each term used. This includes, e.g., the discussion of relationships. Making the result of such a discussion, an ontology, a model. Some structuring of the domain has been done [AI21]. Therefore, we look at the description of an ontology. Some discussion done here is taken from [Fä18]³ with the adaption to the discussion context (Digital Forensics).

A quite general description of ontologies has been given by Gruber, with “*an explicit specification of a conceptualization*” of entities of concern and their relationships⁴. This definition is quite abstract and does not formalize the notion of an ontology but rather explains what is done to create an ontology, which can be done by a multitude of modeling techniques.

To grasp the definition of ontology ready to hold up in court, we want a formal description of ontologies. At least more formal than the definition of Gruber. To find a fitting formalization, we need to look at similar terms first, to distinct them from ontologies:

Model: A model can be seen as an abstraction of an entity. Containing the relevant information regarding a given context. Favre [Fa04a; Fa04b] would describe it as “a system that enables to give answers about a system under study without the need to consider directly this system under study.” [Fa04b, p. 3] and further Miller [MM01, p. 3] describes a model as “... a model is a formal specification of the function, structure and/or behavior of a system”. This should be distinguished from a machine learning model [MRT18].

³ for more details see <https://shorturl.at/iCVZ6>

⁴ <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html> last visited 2017.09.01

Epistemology: Epistemology can be seen as “*the field of philosophy which deals with nature and source of knowledge*” [Nu87] (cited in [SKR07, p. 6]). This sounds like a good starting point on describing knowledge for the use in criminal court. But the problematic areas like the use of AI in investigations is part of this discussion. Regarding AI research, the knowledge is interpreted as consisting of propositions and logical reasoning upon those propositions to create new knowledge in the form of formal structures [Gu95]. With that, Epistemology can not be used to describe digital evidence, since the evidence can be created by software (including use of Machine learning and AI).

Taxonomy: A Taxonomy is the science of classification of entities into classes. Putting them into a category. Thus identifying common properties and properties which differential the entities. Taxonomies structure entities into an order which can hierarchical [SL97]. Further, Euzenat and Shavaiko [ES07, p. 27] define a taxonomy as “a partially ordered set of taxons (classes) in which one taxon is greater than another one only if what it denotes includes what is denoted by the other.” For digital investigations, a taxonomy is mostly not the discussion point. As part of structural science, computer science as well-defined terms which are mostly part of a taxonomy. It is the conceptual designing of how an ontology is created, which aspects are abstracted, and which are designed following a certain view point, which pose a problem.

The definition of a model is too general to be used in forensics, since the abstract way entities and relationships are discussed, seems unnecessary for forensics. Euzenat and Shavaiko [ES07] define an ontology as a conceptual model with features of an entity-relation model. The ontology is thereby seen as a logical theory with model theoretic semantics.

A Taxonomy seems too restricted for entities of digital forensics, since the classification of entities is not enough (“is-a” relationship). This is needed if we want to formalize relationships like “part-of” e.g., a message is part of a conversation under investigation.

A scoping review [MFH22] has extracted the projects of the state-of-the-art of ontologies in digital forensics. The projects UCO [Sy16] and CASE [CNH19] are two projects to design a united ontology out of the discussion of experts in digital forensics.

Figure 2 illustrates a simple taxonomy using the “has” relation, which describes the storage of data in a file system. This gives us an example of how an ontology can capture part of our beliefs and knowledge, depending on the viewpoint of the ontology creator.

This example of Fig. 2 shows how an ontology represents the view of the author. Many of the described entities and their relations can be seen differently. E.g., emphases can be given to the abstraction: Metadata can be further separated into internal and external meta-data.

However, first, we have to look at a different kind of formalization. There are two well-researched theories to formalize ontologies: Set theory and Mereology [Sm98]. We will

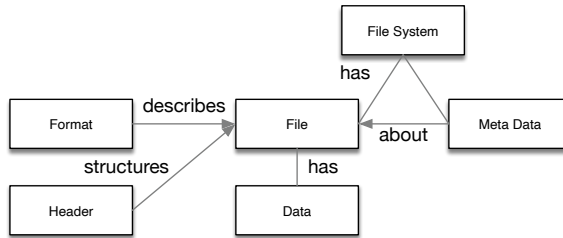


Abb. 2: An example ontology of data in a files

have a short look at both. Applications of such formalization are UCO [Sy16] another one is CASE [CNH19].

Mereotopology The theory of Mereology which became popular in the early 20th century. Mereology is based on the principle of parthood [Sc03] which is similar to the idea of a subset relation in set-theory. The fundamental element in Mereology is called "Object" and is separated into *thin* and *thick* objects. Where thick objects undergo change and, in consequence, are volatile, and the thin objects are invariant. For that reason, a thick object can be identified over time by its set of thin objects which are its parts. This leads to the question of identity: How much change can an Object undergo before becoming another object?. An example in a digital forensic investigation is the use of a refurbished Smartphone. The MAC-Address (thin object) might be the same, but the "thick" object might have changed.

Set-theoretic Ontologies A second and more common formalization of an ontology is a set-theoretic view [Sm98]. Set-theory as one of the basic concepts in mathematics is made up of entity in called *elements* which are joint to a group of entities called *sets* by using the *element relation* (\in). With the *subset* (\subset) relation, one can declare a hierarchy of sets. More formal approaches model set-theoretic ontologies as category⁵ like in the work of Patterson [Pa17], e.g., allows only relations between two sets of concepts.

The argument against Set-theoretic Ontologies models for formal ontologies is that they are inadequate since, e.g., we are not modeling any continuous universe [Sm98]. We argue that in computer science, every data is discretized. Thus, continuous models are approximated.

There are many drawbacks to set theory, which can be resolved with additional abstraction. One of them is that relations are basic entities. Building a taxonomy over each relation allows us to define, e.g., a hierarchy in relations like the "is-a" or "is-part" relations. The definition of an ontology by Maedche und Staab [Ma01] or Euzenat and Shavaiko [ES07]

⁵ Category theory formalized structures in directed graphs with labels on edges and nodes.

or of Pickert⁶ are described in [Fä18]. As well as the here used **Model of Conception** by Mahr [Ma97].

The definition of a model of conception of Mahr [Ma97] describes an abstract view on a conceptualization. This maps to natural language, where each concept can be described by a set of concepts [MP87]. With that, we can describe, e.g., the relation “is created by X” in “File created by user X” which in itself is a set of concepts which might be called “author” or might be a relationship between a file and a person ”authored-by”.

The idea here is, that in an investigation, even if the investigative team tries to be objective, there is always a conceptualization including a point of view ⁷.

Formally we defined an ontology similar to Euzenat and Shavaiko [ES07] but now we should even consider the mind set of the Model of conception from Mahr [Ma97] as shown in Figure 3.

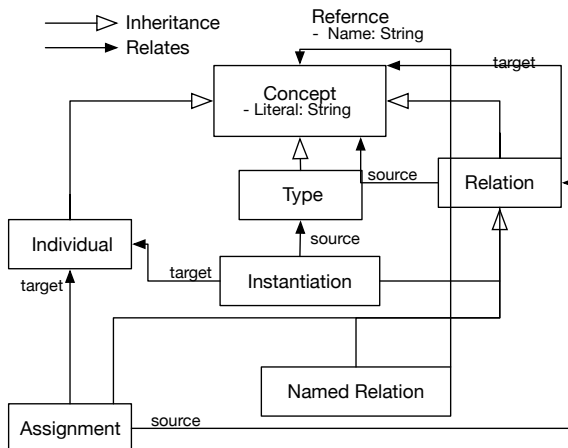


Abb. 3: Type graph of an ontology definition.

Figure 3 describes that everything is a concept. Especially, relations are concepts. We have specified three example relationships:

Named Relation: are relationships with a name like a ”send from” or “IP-Address”.

Instantiation: a relationship which relates concepts to individuals like ”192.168.1.1” or ”127.0.0.1”.

Assignment: a relationship which assigns concrete individuals to the sources or targets of a relationship. E.g. ”send form” and ”192.168.1.1”.

⁶ Taken from a technical report from http://www.dbis.informatik.hu-berlin.de/dbisold/lehre/WS0203/SemWeb/artikel/2/Pickert_Ontologien_final.pdf

⁷ e.g., the history of each person in the team.

We can see that this definition is more practical and is meant for the comparison of ontologies, since there is an explicit difference between classes and instances as well as between an instantiation and an assignment. Furthermore, Euzenat and Shavaiko declare data types and data values as part of the ontology, which is a kind of relations and entities with the purpose of comparing the individuals of an ontology.

The idea is similar to OWL [Ma04] with some extensions from Mahr: Everything is a concept. This means every relation is a concept. Meaning a relation can be seen as a concept, thus can have relationships with other concepts or relationships. This gives us the ability to further describe every aspect of the ontology without redefining its overall structure.

2.1 Conclusion

Like UCO [Sy16] another one is CASE [CNH19] the selection of OWL as ontology description language can be supported. The ontology with a model-theoretic semantics after Tarski [Ta44] which need a formal representation and a history of existing tools. Tools like in UCO and CASE, the existing of tools like reasoners allows a faster application into investigative processes. An interpretation allows us to define the meaning of an expression formulated in the language defined by the ontology. Tarski, e.g., does this for atomic well-formed formulas and defines that if A is a well-formed formula, so is $\text{not}A$.

All these formalizations have entities which might have some kind of relation with other entities or themselves. The relations are quite general and might differ depending on the language used to describe the ontology. The basic concepts are still subject to discussion and the ontological semantics and pragmatics need to be described. Discussion of such properties, their classification and how they can be related and described in a formal way, is part of the here presented contributions.

In general, we want to describe evidence as something which is not true or false, since both are impossible to prove. We rather formulate probabilities on a spectrum.

The idea of Fig. 4 is to describe evidence (or a hint) in Locard's Exchange principle [Sa12]. The idea is to bring the theory from the physical evidence described by Locard to the digital evidence [Po21]. The evidence is taxonomies with regard to properties in digital investigations.

3 Analysis of Definitions

With the definition of digital evidence, we analyze the two properties of evidence, volatility and Manipulability. Ruffly speaking, the meaning of the two terms are: **Volatility** is how long evidence preserver in a given environment. And **Manipulability** is the resources needed to manipulate evidence.

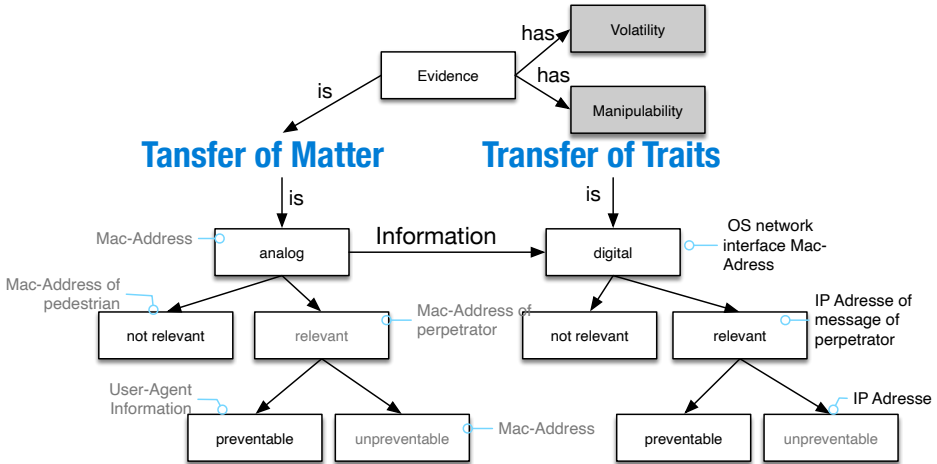


Abb. 4: Ontology representation of evidence of the perspective of digital forensics.

We discuss these two since other properties like preventability or relevance need additional discussion. Preventable e.g., is difficult since digital evidence can be not preventable like an IP-Address to send a request. But with enough manipulation afterward, the evidence can be tampered with. E.g., by deleting log files. Thus, even though we categorize a digital evidence a non-preventable, it might still unavailable for the investigation⁸. Additionally, we do not discuss “red herring” or feint evidence, with a similar argument: The creation of dummy evidence can be seen as manipulation of the evidence, e.g., by rewriting log files in a system to point to other IP-Addresses.

3.1 Volatility

Volatility is defined by the Cambridge Dictionary⁹ with: “the quality or state of being likely to change suddenly”. For digital forensics, this is transferred to memory. Memory can change its state from one bit to another. Corrupting the information stored in the memory. Corruption of information can lead to the loss of data. The loss of data can lead to the loss of evidence.

The definitions we want to discuss are the definitions of properties of digital forensics from [DF15]. They differentiate between persistent evidence (evidence which is conserved over a “relatively big time span” without power.) and volatile evidence, which needs power to be preserved. They further categorize volatile “in a broader sense” evidence if the evidence

⁸ One question here is: Does the preventability then influence the volatility? Since the information needed to be there for the system to work, e.g., the TCP/IP communication. But the manipulation of evidence, e.g., deadlines for deleting of data might destroy them.

⁹ <https://dictionary.cambridge.org/dictionary/english/volatility> visited: 30.6.2023

preserver with powered devices. Without power, they are lost. Evidence with volatility "in a strict sense" is evidence which is not persistent even though the power to the system is provided. Here we get into a gray area of definitions, using the "power on" as part of the definition. The problem here is, that even evidence volatile in a strict sense can be differently volatile. Fig. 5 shows examples of different evidence and its classification of a volatility as a spectrum.

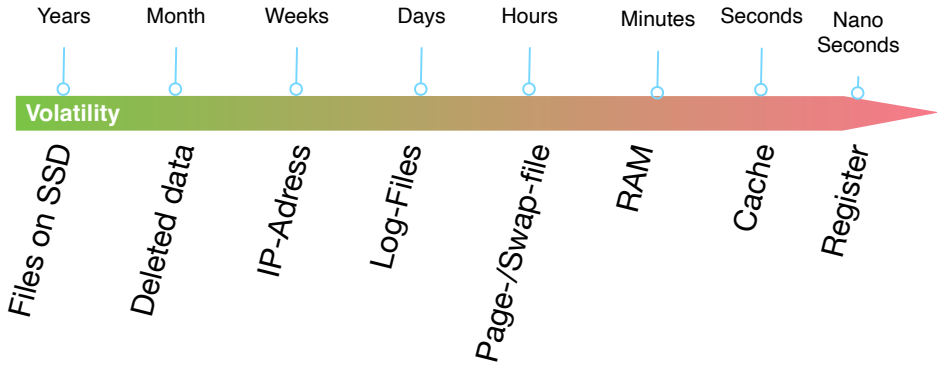


Abb. 5: Volatility as a spectrum.

Our model of volatility does take into account, that there are different reasons why evidence is at a certain level of volatility. The idea is that the evaluation of a piece of evidence can be volatile for many reasons. E.g., circular log files or ring buffer are overwritten after a reach a maximum of time or capacity. This classification of a volatility of log files therefore (because they could be written to a persistent memory) is still volatile, depending on how much is written in them, or how much time has elapsed.

Additionally, the volatility is dependent on the evidence collection process shown in Fig. 1. Depending on how the Preservation and analysis of data is done, the volatility of the evidence changes. One example here could be the analysis of Page-/Swap files. If a running system is analyzed, the page file might contain persistent versions of the data held in the RAM. If during the collection of efficient a laptop lit is closed, or the memory management of the system evacuates additional data from the ram, the page file is altered.

This example shows, that the definition of "something is stored as long as there is power to the system" does not apply here. The definition by [DF15] includes this with the notion of "strict sense" but does not quantify this idea.

Volatility: The time information is available for reading depending on the context.

$$volatility(D)_\theta \propto P(chage(D))_\theta, \tag{1}$$

with D being the data looked at and C the context. The function $chage(D)_\theta$ Likelihood of change

$$P(chage(D)_\theta) = \mathcal{L}(\theta|D) \quad (2)$$

With a fixed θ , this is a likelihood given the observed data D . For digital investigation, the question is how to act on this information. The basic concept of likelihoodist statics is the ratio of likelihoods. Thus, depending on the context, the ratio of different contest can be calculated:

$$A(\theta_1 : \theta_2|D) = \frac{\mathcal{L}(\theta_1|D)}{\mathcal{L}(\theta_2|D)} \quad (3)$$

This means for two contexts $\theta_{1,2}$ the possible context can be asses and decisions can be based on the ratio of likelihoods. E.g. if the operating system is known, and the amount of RAM memory in proportion to the available System memory. The likelihood of a write in a page-/swap file can be asses and help with further investigative decisions.

This kind of model for volatility opens the theory up to the use of further statistical analysis, e.g., the use of a survival model [Ow01] like the Cox proportional hazard model [Co72] With survival models we can estimate the expected duration of time until an event occurs. The event being, e.g., the writing of some data from RAM to the page-/swap file. An alternative here would be to use Bayesian statistics.

In our application of digital forensics, with such an analysis, the chain of custody could be time boxed, to maximize the evidence collected as well as critical decision points could be learned with every collection. Making the system learn for every one. E.g., there could be the system features stored in a database including successfully used modules. Thus showing if a Windows Build version changed something in the implementation of memory management for RAMA dress space layout randomization so that memory acquisition does no longer work and tools need adaption.

3.2 Manipulability

Manipulation in digital forensics has been discussed by, e.g., [DF15]. They state that the manipulation of digital evidence might not leaf evidence at the level of digital evidence itself. The idea here is, that some evidence can be manipulated with ease and some is more difficult to manipulate. But first, we define manipulation in the context of digital forensics.

Manipulation: The change of information with a purpose of creating misinterpretation.

The definition of manipulation as a type of change seems logical, since a manipulation without change would result in no change in the evidence. Thus, for manipulation of digital evidence, changes in the information or in its interpretation needs to be done.

The second part, changing the interpretation, is more complex. An interpretation of evidence includes its presentation before court. With that, standards like the change of custody for information processing need to be fulfilled. The chain of custody is needed, so that manipulation of the digital evidence can be minimized. The interpretation of the evidence can be changed, if the chain of custody is corrupted. Which might conclude in an exclusion of the evidence at trial. To formalize this kind of manipulation, some additional publications are needed.

Therefore, we concentrate on the first part of manipulation of the information which makes up the evidence. Fig.6 shows different qualities of manipulability.

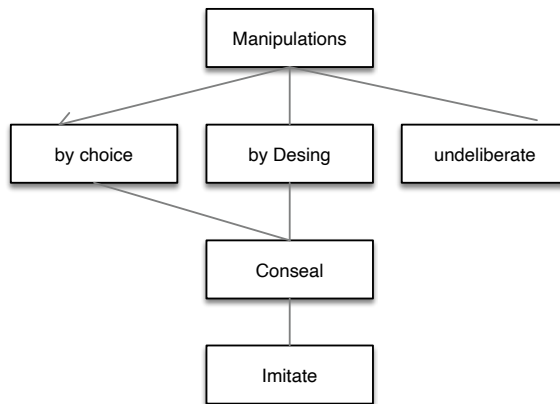


Abb. 6: Qualities of Manipulability.

Manipulation can have many reasons, which we classified in three categories:

by choice: the changes making up the manipulation have been done with intend. E.g., deleting Event-logs.

by design: the changes are done by the system itself, e.g., for privacy reasons in autonomous cars using cameras with ring buffers, which are overwritten after a defined time.

undeliberate: without intend, the information used as evidence can change. E.g., with an error in the transmission protocols.

For the manipulation, there can be different purposes. We classified them in two classes: To conceal information. This could be a simple deletion of information. The lack of information created by the deletion is evidence again. Thus, the second class inheriting concealing

information is imitated. By imitate, we mean imitating evidence. Imitating of evidence is harder to identify manipulation and might lead to false evidence.

The problem with a formal definition of manipulability is, that the time it takes to manipulate something, is dependent on the skills and tools of the manipulator. Furthermore, the quality of the evidence can be correlated to the manipulability. For Example, DNS Caches in a browser is specific to the system user (Easy to manipulate but specific). DNS-Caches of wireless LAN Router are specific to all users in Network. Meaning that the evidence is less useful, and it is harder to manipulate the DNS Cache of a Router. Even less specific is the DNS Cache of the ISP. It can be interpreted as: one of the customers of the ISP used the DNS Entry. This is even less specific to the use, and in general even harder to manipulate.

Without knowledge of the skills and tools or access of the manipulator, we define manipulability as a spectrum as well. Even though the in Fig.7 presented examples might be placed differently by others, we want to show, that each part of the spectrum can be used.

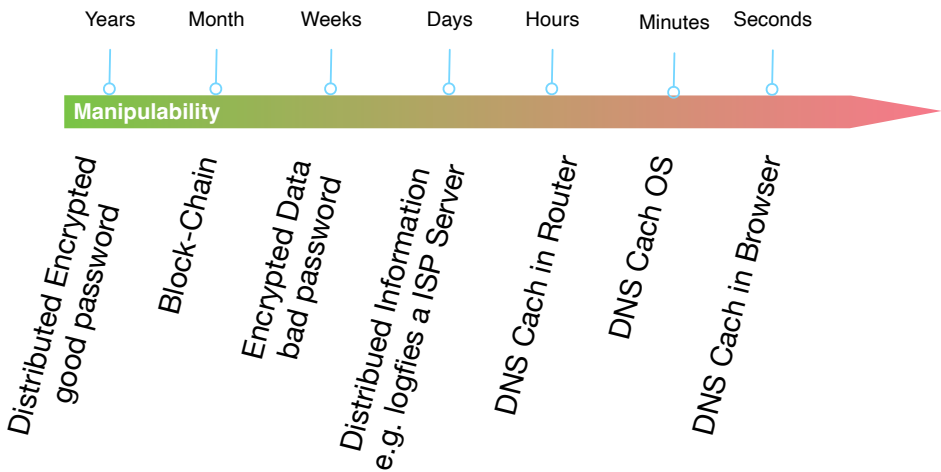


Abb. 7: Manipulability as a spectrum.

Fig. 7 depicts green as hard to manipulate and read as easy. Ranging from a 51% attack on a blockchain to the deletion of a browser history in the browser. In our future work, we want to discuss the other basic terms of digital forensics and correlate them. Correlation of properties of digital forensics¹⁰ needs to be shown by evidence. We want to collect evidence with experiments on manipulation of information, the possibility to identify and revoke a manipulation.

¹⁰ like done in the correlation of expressiveness and manipulability.

3.3 Future Work

With the beginning of seeing volatility and manipulability as a spectrum, we are able to classify digital evidence to their degree of volatility and manipulability. Furthermore, we are looking forward to the scientific discourse on the definition of basic forensics terms in the application of digital forensics. Next on our research agenda is to integrate the idea in first responder software, so that information about the analyzed systems might be collected and used to improve the collection of digital evidence.

Literatur

- [Al21] Al-Dhaqm, A.; Ikuesan, R. A.; Kebande, V. R.; Abd Razak, S.; Grispos, G.; Choo, K.-K. R.; Al-Rimy, B. A. S.; Alsewari, A. A.: Digital forensics subdomains: the state of the art and future directions. *IEEE Access* 9/, S. 152476–152502, 2021.
- [CNH19] Casey, E.; Nelson, A.; Hyde, J.: Standardization of file recovery classification and authentication. *Digital Investigation* 31/, S. 100873, 2019.
- [Co72] Cox, D. R.: Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* 34/2, S. 187–202, 1972.
- [DF15] Dewald, A.; Freiling, F. C.: *Forensische informatik. BoD–Books on Demand*, 2015.
- [ES07] Euzenat, J.; Shvaiko, P.: *Ontology Matching*. Springer-Verlag Berlin Heidelberg, 2007, ISBN: 3-540-49611-4.
- [Fa04a] Favre, J.-M.: Towards a basic theory to model model driven engineering. In: *Workshop in Software Model Engineering, WiSME*. S. 262–271, 2004, URL: http://scholar.google.com/scholar?q=related:pZ-67HCYDo4J:scholar.google.com/&hl=en&num=20&as_sdt=0,5.
- [Fa04b] Favre, J.-M.: Foundations of meta-pyramids: languages vs. metamodels. In: *Episode II. Story of Thotus the Baboon, Procs. Dagstuhl Seminar*. 2004.
- [Fä18] Fährdrich, J.: *Semantic Decomposition and Marker Passing in an Artificial Representation of Meaning*. Technische Universitaet Berlin (Germany), 2018.
- [Gu95] Guarino, N.: Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human-Computer Studies* 43/5-6, S. 625–640, 1995, URL: <http://www.sciencedirect.com/science/article/pii/S107158198571066X>.
- [Ka55] Kant, I.: *Critique of pure reason*, tr. by JMD Meiklejohn. 1855.
- [Ma01] Maedche, A.: Ontology Learning for the Semantic Web. *English, IEEE Intelligent systems* 16/2, S. 72–79, 2001, URL: http://books.google.de/books?id=Hm4jFCxk5VYC&printsec=frontcover&dq=intitle:Ontology+Learning+for+the+Semantic+Web&hl=&cd=1&source=gbs_api.

- [Ma04] Martin, D.; Paolucci, M.; McIlraith, S.; Burstein, M.; McDermott, D.; McGuinness, D.; Parsia, B.; Payne, T.; Sabou, M.; Solanki, M.; Srinivasan, N.; Sycara, K.: Bringing Semantics to Web Services: The OWL-S Approach. In: Semantic Web Services and Web Process Composition. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 26–42, 2004, ISBN: 978-3-540-24328-1, URL: http://link.springer.com/10.1007/978-3-540-30581-1_4.
- [Ma97] Mahr, B.: Gegenstand und Kontext - Eine Theorie der Auffassung. In (Eyferth, K.; Mahr, B.; Posner, R.; Wysotzki, F., Hrsg.): Prinzipien der Kontextualisierung. Technische Universität Berlin, 1997.
- [MFH22] Morgenstern, M.; Fähndrich, J.; Honekamp, W.: Ontology in the Digital Forensics Domain: A Scoping Review. INFORMATIK 2022/, 2022.
- [MM01] Miller, J.; Mukerji, J.: Model Driven Architecture (MDA), Techn. Ber., 2001.
- [MP87] Mel'čuk, I.; Polguère, A.: A Formal Lexicon in the Meaning-Text Theory (or How to Do Lexica with Words). Computational Linguistics 13/3-4, S. 261–275, 1987, URL: <http://dl.acm.org/citation.cfm?id=48166>.
- [MRT18] Mohri, M.; Rostamizadeh, A.; Talwalkar, A.: Foundations of machine learning. MIT press, 2018.
- [Nu87] Nutter, E.: Epistemology. Encyclopedia of Artificial Intelligence 1/, S. 460–468, 1987, URL: http://scholar.google.com/scholar?q=related:-HhWksfaAhoJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5&as_ylo=1987&as_yhi=1987.
- [Ow01] Owen, A. B.: Empirical likelihood. CRC press, 2001.
- [Pa17] Patterson, E.: Knowledge Representation in Bicategories of Relations, Techn. Ber., 2017, URL: <https://arxiv.org/abs/1706.00526>.
- [Po21] Povalej, R.; Rittelmeier, H.; Fähndrich, J.; Berner, S.; Honekamp, W.; Labudde, D.: Die Enkel von Locard: Analyse digitaler Spuren in der forensischen Informatik. Informatik Spektrum 44/, S. 355–363, 2021.
- [PS15] Prayudi, Y.; Sn, A.: Digital chain of custody: State of the art. International Journal of Computer Applications 114/5, 2015.
- [Sa12] Sammons, J.: The basics of digital forensics: the primer for getting started in digital forensics. Elsevier, 2012.
- [Sc03] Schneider, L.: How to Build a Foundational Ontology. In: Agents and Computational Autonomy. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 120–134, 2003, ISBN: 978-3-540-20059-8, URL: http://link.springer.com/10.1007/978-3-540-39451-8_10.
- [SKR07] Sharman, R.; Kishore, R.; Ramesh, R.: Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems. Springer Science + Business Media, LLC, New York, 2007, ISBN: 978-0387-37019-4.

- [SL97] Shanks, D.; Lamberts, K.: Knowledge, Concepts, and Categories. Psychology Press, 1997.
- [Sm98] Smith, B.: Basic Concepts of Formal Ontology. Formal Ontology in Information Systems/, S. 19–28, 1998, URL: <http://philpapers.org/rec/SMITBT>.
- [Sy16] Syed, Z.; Padia, A.; Finin, T.; Mathews, L.; Joshi, A.: UCO: A unified cybersecurity ontology. UMBC Student Collection/, 2016.
- [Ta44] Tarski, A.: The Semantic Conception of Truth: and the Foundations of Semantics. Philosophy and phenomenological research 4/3, S. 341, 1944, URL: <http://www.jstor.org/stable/2102968?origin=crossref>.

Cybersecurity Privatsphäre - Recht
und Technik. Datenschutz im
Diskurs (RuT2023)

Recht und Technik – Datenschutz im Diskurs

Rüdiger Grimm ¹, Gerrit Hornung², Christoph Sorge³ und Indra Spiecker genannt Döhmann⁴

Vorwort zu den Workshopbeiträgen

Der Workshop „Recht und Technik – Datenschutz im Diskurs“ ist nach zehn Jahren ein kontinuierlicher und fester Bestandteil der Jahrestagung der Gesellschaft für Informatik. Nach dem pandemiebedingten zweimaligen Ausweichen auf das Online-Format kann er nun zum zweiten Mal wieder in Präsenz stattfinden. Bei aller erfreulichen Lebendigkeit der Online-Treffen war es uns im letzten Jahre eine große Freude, Gastgeber eines sehr gut besuchten und intensiv diskutierenden Workshops der Jahrestagung sein zu dürfen.

Inhaltlich bleibt die Ausrichtung wie gewohnt: Wir bieten ein Forum für Beiträge von Informatiker/innen, Juristen/innen und Vertreter/innen benachbarter Fächer, die an Fragestellungen des technikbasierten Datenschutzes arbeiten. Wie schon in den Vorjahren werden Themen adressiert, die anwendungsorientiertes Potential für interdisziplinären Diskurs und Zusammenarbeit bieten und die Möglichkeiten aufzeigen, wie Datenschutz durch Technik präzisiert und umgesetzt werden kann.

Die Beiträge, die wir aus der Vielzahl an qualitativollen Einreichungen in einem *peer-reviewed*-Verfahren mit Unterstützung unserer Gutachter/innen auswählen konnten, zeigen die thematische und disziplinären Bandbreite der derzeit im Spannungsfeld von Recht und Technik diskutierten Themen. Die Einreichungen ordnen sich zu folgenden drei Themenfeldern:

- Governance und Enforcement des Datenschutzrechts am Beispiel der aktuellen Bußgeldpraxis der Aufsichtsbehörden
- Transparenzfragen an der Schnittstelle zwischen rechtlichen Informationspflichten und konkreter technischer Umsetzung
- Datenschutz und Künstliche Intelligenz am Beispiel der geplanten „KI-Reallabore“

¹ Fraunhofer SIT, Media Security and Forensics, Rheinstraße 75, 64295 Darmstadt, ruediger.grimm@sit.fraunhofer.de,  <https://orcid.org/0009-0005-5481-8419>

² Universität Kassel, Öffentliches Recht, IT-Recht und Umweltrecht, Henschelstraße 4, 34127 Kassel, gerrit.hornung@uni-kassel.de

³ Universität des Saarlandes, Rechtsinformatik, Gebäude C3 1, 66123 Saarbrücken, christoph.sorge@uni-saarland.de

⁴ Goethe-Universität Frankfurt am Main, Öffentliches Recht, Informationsrecht, Umweltrecht, Verwaltungswissenschaft, Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, spiecker@jur.unifrankfurt.de

Wir freuen uns außerdem sehr, dass wir Annika Selzer (Fraunhofer SIT Darmstadt) für einen eingeladenen Vortrag zu einigen grundlegenden Überlegungen zu Entwicklungslinien und Umbrüchen des Datenschutzes gewinnen konnten.

Wir wünschen den Beiträgen eine rege Diskussion auf dem Workshop und eine interessierte Leserschaft. Wir freuen uns darauf, auch im nächsten Jahr wiederum „Recht und Technik – Datenschutz im Diskurs“ anzubieten.

Unser besonderer Dank gilt den Mitgliedern des Programmkomitees, unter denen in diesem Jahr folgende Personen in die Begutachtung eingebunden waren:

- Matthias Bäcker, Universität Mainz
- Franziska Boehm, KIT
- Jens-Matthias Bohli, Hochschule Mannheim
- Matthias Enzmann, Fraunhofer SIT, Darmstadt
- Christian Geminn, Universität Kassel
- Nils Gruschka, Universität Oslo
- Christoph Gusy, Universität Bielefeld
- Marit Hansen, ULD Schleswig-Holstein
- Niko Härting, Rechtsanwalt, Berlin
- Walter Hötendorfer, Research Institute Digital Human Rights Center, Wien
- Thomas Kahler, DPOblog.eu
- Thomas Kehr, Kanzlei MMV - Martini, Mogg, Vogt PartGmbH
- Thomas Lapp, Rechtsanw, Mediator, Fachanwalt IT-Recht, IT-Kanzlei dr-lapp.de GbR
- Kai von Lewinski, Uni-Passau
- Ronald Petrlic, TH Nürnberg
- Hannah Ruschemeier, CAIS - Center for Advanced Internet Studies, Bochum
- Burkhard Schafer, Universität Edinburgh
- Tobias Singelstein, Ruhr-Universität Bochum
- Jürgen Taeger, Universität Oldenburg

Gesundheits-Apps auf dem Prüfstand – Überprüfung der Angaben in Datenschutzerklärungen zur Datenweitergabe

Matthias Kohn¹, Merle Freye², Mehrdad Bahrini³ und Alexander Herbst⁴

Abstract: Apps sind von einer großen Informationsasymmetrie zwischen den Anbieter:innen und den Nutzer:innen im Hinblick auf die Verarbeitung personenbezogener Daten durch die jeweiligen Apps gekennzeichnet. Datenschutzerklärungen sollen diese Informationsasymmetrie aufbrechen und Transparenz herstellen. Ob die Angaben in den Datenschutzerklärungen zutreffend sind und die Apps die Daten tatsächlich in der angegebenen Art und Weise verarbeiten, lässt sich für die Nutzer:innen jedoch kaum nachvollziehen. Bei Gesundheits-Apps werden naturgemäß besonders sensible Daten verarbeitet, weshalb die datenschutzrechtlichen Anforderungen hier besonders hoch anzusetzen sind. Der Beitrag stellt einen Ansatz zur technischen Überprüfung der Datenschutzkonformität von Smartphone-Apps vor und untersucht am Beispiel von Gesundheits-Apps, ob die in den Datenschutzerklärungen genannten Drittländer und Empfänger:innen mit der tatsächlichen Datenverarbeitung der Gesundheits-Apps übereinstimmen. Zusätzlich wird der Einfluss von Kategorienangaben auf die Transparenz in Datenschutzerklärungen kritisch betrachtet.

Keywords: App-Analyse, Transparenz, Datenschutzerklärungen, Informationsasymmetrie, Datenweitergabe

1 Einleitung

Die DSGVO enthält mit Art. 13 DSGVO einen umfangreichen Katalog an Informationen, die die Anbieter:innen den Nutzer:innen zum Zeitpunkt der Datenerhebung zur Verfügung stellen müssen. In der Praxis erfüllen die Anbieter:innen diese Verpflichtung durch die Verlinkung einer umfangreichen Datenschutzerklärung. Die meisten Nutzer:innen lesen diese Datenschutzerklärung jedoch nicht, da diese für juristische Laien zu lang und zu unverständlich ist.

Diejenigen Nutzer:innen, die die Datenschutzerklärung lesen, finden häufig nicht nur vage Informationen vor, sondern besitzen auch keine Möglichkeit, die Informationen auf ihren Wahrheitsgehalt zu prüfen. Wenn die Datenschutzerklärung beispielsweise besagt, dass keine Daten in den USA verarbeitet werden und die Daten nicht an Dritte weitergegeben werden, besteht für die Nutzer:innen letztendlich keine Möglichkeit, dies zu prüfen. Insbesondere im Hinblick auf Gesundheits-Apps, die in den App-Stores frei erhältlich

¹ Universität Bremen, Institut für Informations-, Gesundheits- und Medizinrecht (IGMR), Universitätsallee, 28359 Bremen, kohn@uni-bremen.de

² Universität Bremen, Institut für Informations-, Gesundheits- und Medizinrecht (IGMR), Universitätsallee, 28359 Bremen, mfreye@uni-bremen.de

³ Universität Bremen, Digital Media Lab, Bibliothekstr. 5, 28359 Bremen, mbahrini@uni-bremen.de

⁴ Universität Bremen, Digital Media Lab, Bibliothekstr. 5, 28359 Bremen, herbst1@uni-bremen.de

sind, findet keine Prüfung der Angaben der Anbieter:innen statt (im Gegensatz zu den „Apps auf Rezept“ bzw. digitalen Gesundheitsanwendungen des § 33a SGB V – DiGA, die ab 2024 zertifiziert werden müssen, § 139e XI SGB V). Dabei erfordert gerade die Verarbeitung von sensiblen, besonders schützenswerten Gesundheitsdaten ein hohes Schutzniveau. Nutzer:innen müssen damit in hohem Maße auf die Korrektheit der Angaben in der Datenschutzerklärung vertrauen, ohne diese in vollem Umfang überprüfen zu können.

Im Folgenden wird daher ein Ansatz zur technischen und juristischen Überprüfung von Angaben in Datenschutzerklärungen von Apps entwickelt, mit dessen Hilfe 20 Gesundheits-Apps beispielhaft untersucht werden. Dabei liegt der Schwerpunkt auf der Frage, ob die Angaben zu Drittländern und Empfänger:innen mit der tatsächlichen Datenverarbeitung der Apps übereinstimmen und ob diese Angabe für die Nutzer:innen transparent sind.

2 Methodik

Die vorliegende Untersuchung vereint interdisziplinär die technische und juristische Analyse von Apps und Datenschutzerklärungen. Konkret bedeutet dies, dass mithilfe eines Top-down und Bottom-up Ansatzes Untersuchungsziele erarbeitet werden, auf deren Grundlage 20 Gesundheits-Apps untersucht werden.

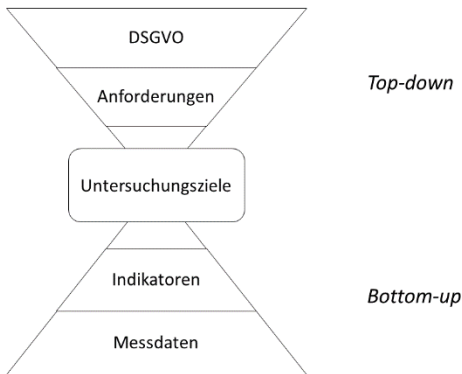


Abb. 1: Untersuchungsziele an der Schnittstelle zwischen Recht und Technik

Im *Top-down* Prozess wird zunächst herausgearbeitet, welche Informationen die Anbieter:innen basierend auf der DSGVO zur Verfügung stellen müssen und, welche dieser Informationen überhaupt technisch überprüft werden können. *Bottom-up* werden auf Basis der im Rahmen der technischen Analyse messbaren Daten (z.B. ein- und ausgehender Netzwerkverkehr, verwendete Bibliotheken und angeforderte

Zugriffsrechte), Indikatoren für verschiedene Datenschutz- und Datensicherheitsfaktoren entwickelt. Durch die Kombination von Top-down und Bottom-up-Ansatz ergeben sich die Untersuchungsziele für die Analyse von Datenschutzerklärungen und Apps (Abb. 1).

Gegenstand der sich anschließenden Analyse sind 20 Gesundheits-Apps der Kategorien „Gesundheit & Fitness“ und „Medizin“ des Google Play Stores, die von den Nutzer:innen häufig heruntergeladen wurden [Ar22a, Ar22b]. Für die Untersuchung werden zunächst die Datenschutzerklärungen der 20 Gesundheits-Apps im Hinblick auf die vorher bestimmten Untersuchungsziele analysiert. Anschließend werden die technischen Analysedaten erhoben und schließlich die Konformität von Datenschutzerklärung und tatsächlicher Datenverarbeitung der Gesundheits-Apps evaluiert.

3 Juristische Analyse

Die DSGVO stellt umfangreiche Informationspflichten für die Verarbeitung von personenbezogenen Daten auf. Zur Identifikation der juristischen Anforderungen müssen zum einen die relevanten Datengruppen identifiziert werden und zum anderen die Informationspflichten identifiziert werden, die anschließend technisch überprüft werden können.

3.1 Personenbezogene Daten

Die Informationspflichten nach Art. 13 DSGVO betreffen ausschließlich die Verarbeitung von personenbezogenen Daten. Personenbezogene Daten sind gemäß Art. 4 Nr. 1 DSGVO solche Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen. Entscheidend für die App-Analyse ist somit, ob die Nutzer:innen mit den erhobenen Daten identifizierbar sind (Schild in [Wo23] Art. 4 DSGVO Rn. 14 ff.; Arning/Rothkegel in [Ta22] Art. 4 DSGVO Rn. 22 ff.). Apps können neben einer direkten Identifizierung über Accounts auch weitere Daten erheben, die eine Identifizierung der Nutzer:innen ermöglichen. Das können etwa die öffentliche IP-Adresse ([Bg17] S. 115), Mobilfunknummer, Adressbuchdaten, Geräteidentifikationsnummern oder die Werbe-ID sein. Außerdem lassen sich aus Bewegungsmustern und aus Audioaufnahmen der individuellen Stimme (über einen Abgleich von Profilen) Personen identifizieren (Profiling) ([Sa13] S. 304). Durch Gesundheits-Apps werden naturgemäß zusätzlich Gesundheitsdaten und damit besonders sensible personenbezogene Daten im Sinne des Art. 9 DSGVO verarbeitet. Darunter fallen insbesondere Daten, die Aussagen über den derzeitigen oder künftigen Gesundheitszustand treffen, aber auch Angaben, die unmittelbar physische und psychische Zustände und Charakterisierungen einer Person betreffen (Weichert in [Kü20] Art. 9 DSGVO Rn. 36 f.; Frenzel in [Pa21] Art. 9 DSGVO Rn. 15). Da die Informationspflichten sich nur auf personenbezogene Daten beziehen, muss in der sich anschließenden technischen Analyse somit sichergestellt werden, dass die Apps personenbezogene Daten verarbeiten.

3.2 Informationspflichten zum Datenverhaltensverhalten

Art. 13 DSGVO enthält 12 Informationspflichten, welche die Anbieter:innen von Apps erfüllen müssen, wenn sie personenbezogene Daten bei den Nutzer:innen erheben. Diese Informationen werden in den Datenschutzerklärungen der jeweiligen App zur Verfügung gestellt.

Allerdings werfen insbesondere die Informationen über Drittländer (Art. 13 Abs. 1 lit. f DSGVO) und über die Empfänger:innen (Art. 13 Abs. 1 lit. e DSGVO) eine zentrale Frage auf. So ist im Rahmen der Information über die Empfänger:innen umstritten, ob die Angabe von Kategorien zulässig ist und, ob eine namentliche Nennung der Empfänger:innen erfolgen muss. Im Hinblick auf Art. 13 DSGVO entzündet sich der Meinungsstreit an dem Wortlaut des Art. 13 Abs. 1 lit. e DSGVO, da die Empfänger:innen oder Kategorien von Empfänger:innen „gegebenenfalls“ mitgeteilt werden sollen. Vertreten wird in diesem Zusammenhang ein eingeschränktes Wahlrecht: Nur für den Fall, dass eine konkrete Empfänger:in bei der Datenerhebung nicht absehbar ist, soll eine Umschreibung nach der Kategorie zulässig sein (Bäcker in [Kü20] Art. 13 DSGVO Rn. 30). Sobald die Empfänger:in also feststeht, würde das Wahlrecht erlöschen (Dix in [Si19] Art. 13 DSGVO Rn. 11; Franck in [Go22] Art. 13 DSGVO Rn. 20; [Lo19] S. 216). Andererseits ist der Wortlaut auch offen für ein uneingeschränktes Wahlrecht, sodass grundsätzlich zwischen einer Kategorie oder der konkreten Benennung gewählt werden könnte (Kamlah in [Pl18] Art. 13 Rn. 13).

Eine ähnliche Diskussion besteht für die Information in den Datenschutzerklärungen über Drittländer. Auch in diesem Zusammenhang ist nicht geklärt, ob die Drittländer nach Art. 13 Abs. 1 lit. f DSGVO namentlich genannt werden müssen, oder nicht. Der Wortlaut und die Systematik der DSGVO sprechen dagegen, da lediglich Art. 30 Abs. 1 S. 2 lit. e DSGVO die „Angabe des Drittlandes“ erfordert, nicht aber Art. 13 Abs. 1 lit. f DSGVO (Schmidt-Wudy in [Wo23] Art. 14 DSGVO Rn. 57). Allerdings vertreten die Art.-29-Datenschutzgruppe und auch der diese ablösende Europäische Datenschutzausschuss [Eu18] die Ansicht, dass das Drittland „normalerweise“ namentlich benannt werden soll ([Ar18b] S. 48 f.). Die namentliche Nennung entspricht auch dem Sinn und Zweck der Informationspflichten nach Art. 13 DSGVO, nachdem die Nutzer:innen in die Lage versetzt werden sollen, das Risiko der Datenverarbeitung abschätzen zu können. Wenn eine sehr weite Umschreibung verwendet wird – wie beispielsweise eine Datenverarbeitung „in anderen Ländern“ – ist eine Risikoeinschätzung kaum möglich.

Für die Information über Drittländer als auch für die Information über Empfänger:innen ist somit nicht klar, wann eine kategoriale Umschreibung zulässig ist. Die technische Untersuchung verspricht eine Argumentationsbasis für die Streitstände, da technisch erfasst wird, wie viele Empfänger:innen bzw. Drittländer in der Praxis unter eine Kategorie zusammengefasst werden.

3.3 Datenschutzrechtliche Anforderungen

Somit können die in den Datenschutzerklärungen angegebenen Drittländer und Empfänger:innen im Wege einer technischen Analyse überprüft werden, sofern sie personenbezogene Daten verarbeiten.

4 Technische Analyse

Für die technische Analyse der tatsächlichen Datenverarbeitung durch die Apps kann auf die bestehenden Ansätze der IT-Sicherheitsforschung zurückgegriffen werden, die auch für die Analyse von Malware verwendet werden ([Sa13] S. 307).

4.1 Bestehende Ansätze zur App-Analyse

Grundsätzlich liefern *Metadaten* Auskunft über das Verhalten der App. Beispielsweise geben die in der `AndroidManifest.xml` Datei enthaltenen Daten unter anderem Auskunft über die angeforderten Berechtigungen auf dem Gerät. Allerdings gilt hier zu beachten, dass Metadaten auch manipuliert werden können und daher lediglich Indizien für ein bestimmtes Verhalten der App liefern, nicht jedoch belastbare Belege. Eine Studie der Mozilla Foundation kam etwa zum Schluss, dass ca. 80% der hinterlegten Metadaten in Googles Data Safety Formular im Play Store nicht mit den Angaben in den Datenschutzerklärungen der Apps übereinstimmten [St23]. Um belastbare Ergebnisse zu erhalten, wird auf die statische Code Analyse sowie die dynamische Analyse inklusive der anfallenden Netzwerkdaten zurückgegriffen.

Für die *statische Code Analyse* muss die App zunächst dekompiert werden. Anschließend kann der Quellcode analysiert und ausgewertet werden. Dafür kann beispielsweise auf `jadx` zurückgegriffen werden; dabei handelt es sich um einen Dex zu Java Decompiler, der Android Dex und Apk Dateien in Java Quelltext rückübersetzt [Ka23]. Im nächsten Schritt wird der so gewonnene Quellcode teilautomatisiert analysiert, wobei dann die eingebundenen Softwarebibliotheken sowie die im Quellcode hinterlegten Domains und IP-Adressen extrahiert werden.

Bei der *dynamischen Analyse* wird das Verhalten des Programms während der Laufzeit analysiert. Hier können detaillierte Berichte zum Verhalten des Programms erstellt werden. So können die Datenflüsse innerhalb des Programms in zeitlicher Relation zu Ereignissen – wie dem Einholen einer Einwilligung oder der Eingabe bestimmter personenbezogener Daten – analysiert werden. Bei der *Netzwerkanalyse* werden der ein- und ausgehende Datenverkehr überwacht. Dafür wird ein Intercept-Proxy benötigt, der zwischen die Kommunikation der App mit den anderen Diensten geschaltet wird und so die Analyse des Netzwerkverkehrs ermöglicht. Damit eine Analyse des Netzwerkverkehrs auch bei verschlüsselten Verbindungen erfolgreich ist, muss in der Testumgebung ein

eigenes Root-Zertifikat installiert sein, sodass der Netzwerkverkehr beim Intercept-Proxy entschlüsselt werden kann.

4.2 Gewählter Ansatz

Die Analyse erfolgt in drei Schritten. Zunächst wird die Installationsdatei der App statisch analysiert, ohne dass die App ausgeführt wird. Neben dem Quelltext können dabei auch die Ressourcen und verwendeten Zertifikate betrachtet werden. Im zweiten Schritt wird eine dynamische Analyse durchgeführt und die App mit Testdaten betrieben. Dabei werden die Datenverarbeitung sowie die ein- und ausgehenden Netzwerkverbindungen überwacht. Im letzten Schritt werden die gemessenen Daten mit den Angaben in der Datenschutzerklärung abgeglichen.

Die Gesundheits-Apps werden mithilfe des Mobile Security Framework (MobSF) untersucht [Gi23]. MobSF wurde zur Aufspürung von Schadcode und Sicherheitsmängeln in Android, iOS und Windows-Apps entwickelt und bietet unter anderem eine ausführliche automatisierte statische sowie eine dynamische Analyseoption. Soweit die Gesundheits-Apps sich nicht oder nicht zuverlässig in einer virtualisierten Umgebung starten lassen, werden diese auf einem gerooteten Smartphone installiert. Auf dem Gerät ist ein Root Zertifikat hinterlegt, womit auch der verschlüsselte Netzwerkverkehr mithilfe eines Intercept-Proxies nachvollzogen werden kann.

Allerdings ist zu beachten, dass der gewählte Ansatz auch Limitationen unterliegt. Soweit eine Datenverarbeitung oder Weitergabe erst nachgelagert auf dritten Servern erfolgt, können diese nicht ohne Zugriff auf diese Server überprüft werden. Das bedeutet, dass beispielsweise Datenübertragungen, die nicht direkt von der App an einen Drittserver erfolgen, sondern erst nach einer Übertragung an die Server der Anbieter:innen stattfinden, technisch nicht nachvollzogen werden können. Darüber hinaus verwenden die meisten Anwendungen verschiedene Kodierungssysteme, was das Auffinden ursprünglicher und verständlicher Daten erschwert. Unter Kodierung versteht man den Prozess der Umwandlung von Daten oder einer bestimmten Folge von Zeichen, Symbolen, Alphabeten usw. in ein bestimmtes Format für die sichere Übertragung von Daten.

4.3 Messbare Daten

Mit der Kombination von statischen und dynamischen Analyseverfahren lassen sich verschiedene Daten zur Datenverarbeitung durch die App messen und analysieren. Dabei können die folgenden Einzeldaten erhoben werden: IP-Adressen und Domains der kontaktierten Server, verwendete Softwarebibliotheken, Inhalt der Datenübermittlungen, angefragte Zugriffsberechtigungen auf dem Endgerät sowie Zugriffe auf Zubehör, wie z.B. eine gekoppelte Smartwatch.

4.4 Indikatoren

Über die IP-Adressen der kontaktierten Server lassen sich die Serverstandorte und damit die Länder bestimmen, in die eine Datenübertragung erfolgt. Über die IP-Adressen und die Domains der kontaktierten Server können darüber hinaus mithilfe von Whois-Lookups und Untersuchungen des aus- und eingehenden Netzwerkverkehrs die dahinterstehenden Dienste und Unternehmen identifiziert werden. Die verwendeten Softwarebibliotheken liefern zusätzliche Hinweise zur Identifikation der hinter den kontaktierten Servern stehenden Datenverarbeiter:innen. Anhand des Netzwerkverkehrs lässt sich im Übrigen auch bestimmen, ob und welche personenbezogenen Daten übermittelt werden und an wen diese weitergegeben werden. Zusätzlich lässt sich auch bestimmen, ob diese Übertragung verschlüsselt oder unverschlüsselt erfolgt.

4.5 Mapping von Indikatoren und datenschutzrechtlichen Anforderungen

Diese Indikatoren ermöglichen konkrete Aussagen über die identifizierten juristischen Anforderungen. Ob und in welchen Ländern eine Datenverarbeitung vorgenommen wird, lässt sich überprüfen, indem die Standorte der kontaktierten Server und die dazugehörigen Länder mithilfe der IP-Adressen bestimmt werden. Die Empfänger:innen von Daten werden – zumindest, wenn die Daten direkt von der App an diese übermittelt werden – ebenfalls über die kontaktierten Server ermittelt. Allerdings ist die Dauer der Datenverarbeitung bzw. der Datenspeicherung auf den Servern der Anbieter:innen nur sehr eingeschränkt technisch überprüfbar. Eine zuverlässige Aussage über die Dauer der Datenverarbeitung kann damit nicht getroffen werden.

Somit sind die Untersuchungsziele zum einen die Überprüfung der angegebenen Datenverarbeitungen in Drittländern und zum anderen die Überprüfung der angegebenen Empfänger:innen der personenbezogenen Daten.

5 Untersuchungsergebnisse

Auf Grundlage der Untersuchungsziele wurden 20 Gesundheits-Apps untersucht und die Ergebnisse der technischen Analyse mit den Angaben in der jeweiligen Datenschutzerklärung abgeglichen.

5.1 Auswahl der Apps

Untersuchungsgegenstand sind 20 Apps, die am häufigsten aus den Kategorien „Gesundheit & Fitness“ und „Medizin“ des Google Play Stores heruntergeladen wurden. Als Grundlage für die Häufigkeit des Downloads dienen die von Androidrank erhobenen Downloadzahlen am 14. August 2022 [Ar22a, Ar22b]. Nicht berücksichtigt wurden dabei zwei Apps, die nicht auf dem europäischen Markt verfügbar waren und somit nicht den

Anforderungen der DSGVO unterfielen. Außerdem konnte eine App nicht überprüft werden, da sie nach dem USA-Bann von Huawei aus dem Google Play Store entfernt worden war. Die drei herausgefallenen Apps wurden mit den drei nachfolgenden Apps der Darstellung auf Androidrank ersetzt. Somit ergaben sich 16 Tracking-Apps (darunter sieben mit dem Fokus Schwangerschaft und Periode), zwei Training-Apps, ein Symptom Checker und eine App, die medizinische Informationen vermittelt.

5.2 Übereinstimmung angegebener Drittländer und Empfänger:innen mit dem realen Datenfluss

Bei der *Drittlandsübermittlung* von personenbezogenen Daten zeigte die Analyse ein gemischtes Ergebnis. Sofern Angaben zur Drittlandsübermittlung von Daten gemacht wurden, waren diese mit einer Ausnahme zutreffend. Bei sechs Apps wurden keine Angaben zu Drittlandsübermittlungen gemacht, obwohl Datenströme in Drittländer gemessen werden konnten. Damit waren die Angaben von 65% der Apps zutreffend. Relativiert werden muss dieses Ergebnis jedoch aus Transparenzgesichtspunkten, denn diese Übereinstimmungen ergaben sich vor allem daraus, dass die Drittländer in diesen Datenschutzerklärungen auf sehr generelle Art und Weise beschrieben wurden: So verwiesen einige Datenschutzerklärungen pauschal darauf, dass die Daten in anderen Ländern verarbeitet werden oder gaben an, dass Daten in den Ländern verarbeitet werden, in denen das Unternehmen Büros besitzt.

Bei der Analyse fiel darüber hinaus auf, dass häufig eine Kombination aus Kategorie von Drittländern und einem oder mehreren Beispielen verwendet wurde, z.B.: „USA und andere Drittländer“. Berücksichtigt man nur die nicht bereits beispielhaft genannten Länder, so kamen in neun Fällen auf eine in der Datenschutzerklärung genannte Kategorie an Drittländern auch nur ein nicht ausdrücklich genanntes Drittland. Ob die Angabe einer Kategorie an Stelle eines einzelnen Drittlands dem Transparenzerfordernis genügen kann, erscheint fraglich. Wenn sich lediglich ein Land hinter einer Kategorie verbirgt, sollte dieses aus Transparenzgründen auch konkret benannt werden, insbesondere wenn es sich um eine allgemeine Kategorie wie „andere Drittländer“ handelt.

Zu ähnlichen Ergebnissen gelangte die Analyse auch bei den *Empfänger:innen* von Daten. Die Angaben in der Datenschutzerklärung stimmen hier sogar bei 85% mit den ermittelten Empfänger:innen überein. Im Gegensatz zu den Drittlandsübermittlungen wurde in allen Datenschutzerklärungen auf die Empfänger:innen eingegangen. Aber auch hier fiel auf, dass die Apps sehr weite kategoriale Beschreibungen verwenden und konkrete Empfänger:innen – wenn überhaupt – nur beispielhaft benannt wurden. 60% der untersuchten Apps verweisen unter anderem auf abstrakte „Geschäftspartner“ oder „Diensteanbieter“ als mögliche Empfänger:innen, ohne diese genauer zu spezifizieren. Dabei wurden mit einer Kategorie durchschnittlich 10,24 unterschiedliche Empfänger:innen zusammengefasst. Allerdings variierte die Anzahl der Empfänger:innen pro referenzierter Kategorie stark, da bei 20% der untersuchten Datenschutzerklärungen weniger als zwei Empfänger:innen in einer Kategorie zusammengefasst werden. Auch hier

stellt sich die Frage, ob diese Datenschutzerklärungen dem Transparenzerfordernis gerecht werden können, wenn eine abstrakte Kategorie wie „Geschäftspartner“ anstelle eines oder zweier Dienste verwendet wird. Auf der anderen Seite fassten 40% der Apps durchschnittlich mehr als 10 Empfänger:innen in einer Kategorie zusammen. In einige Kategorien fallen sogar noch mehr Empfänger:innen, da für die Kategorie der Datenweitergaben an Regierungsorganisationen oder Strafverfolgungsbehörden bei keiner der untersuchten Apps zuordbare Datenströme gemessen werden konnten. Insbesondere bei Datenweitergaben an interessenbasierte Werbenetzwerke oder Trackingdienste wie „Google Analytics“ wurden diese nur bei 60% der Tracker-verwendenden Apps konkret benannt oder als eigene Kategorie aufgeführt. In allen anderen Fällen sind die Trackingdienste und Werbenetzwerke gemeinsam mit Hosting-Provider:innen und Provider:innen von In-App Funktionalitäten (z.B. Karten oder Wetterdiensten) in allgemeinen Kategorien wie „Geschäftspartner“, „Dritte“ oder „Diensteanbieter“ zusammengefasst.

Die überwiegende Übereinstimmung von Datenschutzerklärung und tatsächlicher Datenverarbeitung der Gesundheits-Apps basierte somit maßgeblich auf sehr weit und allgemein gefassten Informationen in den Datenschutzerklärungen. Allerdings muss in diesen Fällen hinterfragt werden, ob weite und kategoriale Umschreibungen den Nutzer:innen tatsächlich eine ausreichende Informationsgrundlage bieten können.

5.3 Rückschlüsse für die Erfüllung von Informationspflichten

Innerhalb der Informationspflichten des Art. 13 DSGVO ist strittig, wann die Empfänger:innen von Daten in Kategorien zusammengefasst werden dürfen und, ob Drittländer namentlich benannt werden müssen. Die Untersuchung hat ergeben, dass die meisten Apps weite Kategorien verwenden, die so allgemein gehalten sind, dass eine Vielzahl von Empfänger:innen und Drittländern darunter fällt und fallen könnte. Da hierdurch kein echter Informationsmehrwert entsteht, muss die Informationsvermittlung anders umgesetzt werden. Im Folgenden soll eine funktionale Trennung des Informationsmediums vorgeschlagen werden, die die informationelle Selbstbestimmung der Nutzer:innen situationsgerecht gewährleistet.

Dementsprechend soll die Datenschutzerklärung von der Information hinsichtlich einer informierten Einwilligung unterschieden werden – die Datenschutzerklärung ist also nicht das Informationsmedium für eine Informierte Einwilligung. Weiterhin sollen die Informationen in der Datenschutzerklärung andere sein als die Informationen, die den Nutzer:innen vor der Abgabe der Einwilligung zur Verfügung gestellt werden.

Wenn Nutzer:innen die Datenschutzerklärung gezielt aufsuchen und lesen, müssen die Angaben transparent genug sein, um Art und Ausmaß der Datenweitergabe einschätzen zu können. In der Datenschutzerklärung sollten die Empfänger:innen und Drittländer somit konkret benannt werden, soweit den Anbieter:innen diese Informationen vorliegen.

Im Hinblick auf die informierte Einwilligung sind die Informationen hingegen so aufzubereiten, dass die Nutzer:innen eine unmittelbar darauf aufbauende Entscheidung über die Einwilligung treffen können. Die Informationen für die Abgabe einer informierten Einwilligung sollten daher kurz sein und nicht alle Informationspflichten der Art. 13, 14 DSGVO abbilden, sodass die Datenschutzerklärung nach der hier vertretenen Ansicht nicht das entsprechende Informationsmedium der informierten Einwilligung sein kann. Diese funktionelle Teilung von Datenschutzerklärung und informierter Einwilligung bedeutet somit, dass allgemeine Kategorien und Umschreibungen im Rahmen einer Datenschutzerklärung zur Wahrung der Transparenz nicht genügen und hier eine konkrete Auflistung soweit möglich erfolgen sollte. Für die Information im Hinblick auf die Einwilligung können hingegen Kategorien und Umschreibungen gegebenenfalls zulässig sein – sofern sie konkreter gefasst sind als die in der Untersuchung gefundenen Beispiele; hier besteht ein weiterer Forschungsbedarf zur Konzeption solcher Kategorien.

Eine funktionelle Unterscheidung deutet auch das BfArM an, das für Gesundheits-Apps, deren Kosten als DiGA von den gesetzlichen Krankenkassen erstattet werden, ausdrücklich festgelegt hat, dass die Information hinsichtlich der Einwilligung nicht allein über einen Verweis auf die Datenschutzerklärung erfolgen darf ([Bf22] S. 19). Die Informationen in den Datenschutzerklärungen setzen zwar die Pflichten des Art. 13 DSGVO um, sind aber nicht gleichzusetzen mit den Anforderungen an eine informierte Einwilligung ([Co20] S. 28; Kamlah in [P118] Art. 13 DSGVO Rn. 17; [Fra16] S. 116; [Ka17] S. 27). Darüber hinaus scheint auch der Europäische Datenschutzausschuss von einer Unterscheidung zwischen Art. 13 DSGVO und der informierten Einwilligung auszugehen und verweist darauf, dass eine gültige Einwilligung „in informierter Weise“ vorliegen kann, selbst wenn nicht alle Elemente des Art. 13 DSGVO beim Einholen der Einwilligung genannt werden ([Eu18]; [Ar18a] S. 17). Somit kann ein Verstoß gegen die Informationspflichten des Art. 13 DSGVO lediglich isoliert sanktioniert werden; die Einwilligung hingegen kann trotzdem rechtmäßig sein, wenn eine entsprechende Information außerhalb der Datenschutzerklärung vorliegt.

Auch der EuGH entschied im Januar 2023, dass die bloße Angabe von Kategorien von Empfänger:innen der personenbezogenen Daten intransparent sei [Eu23]. Zwar betrifft die Entscheidung lediglich die Auskunftspflicht nach Art. 15 DSGVO und die Angabe von (Kategorien von) Empfänger:innen. Allerdings stützt das Urteil die funktionelle Trennung von Informationspflichten, da es für den Fall eines ausdrücklich kommunizierten Interesses, das hinter dem Auskunftsrecht steht, eine umfassende Aufzählung der Empfänger:innen verlangt. Das Auskunftsverlangen ist vergleichbar mit dem gezielten Aufsuchen der Datenschutzerklärung, sodass bei einem solchen gezielten Ansteuern der Datenschutzerklärung ebenfalls detaillierte Angaben überzeugen.

Es muss somit zwischen Informationen im Hinblick auf die informierte Einwilligung und weiteren Informationen im Hinblick auf die Informationspflichten des Art. 13 DSGVO unterschieden werden, um auf der einen Seite einer Informationsüberflutung in der Einwilligungssituation vorzubeugen, und auf der anderen Seite, den Anforderungen an die Genauigkeit und Transparenz gerecht zu werden.

6 Fazit

Mithilfe des interdisziplinären Ansatzes wurde erfolgreich ein Verfahren zur Überprüfung von Apps entwickelt. Durch die Analyse der Datenschutzerklärungen konnten empirische Beweise für praktische Probleme in der Umsetzung der Informationspflichten nach Art. 13 DSGVO gesammelt werden.

Die beispielhafte Anwendung auf 20 Gesundheits-Apps hat gezeigt, dass die tatsächliche Datenverarbeitung der Gesundheits-Apps im Hinblick auf die Empfänger:innen und die Drittländer überwiegend mit den Angaben übereinstimmen. Dies ist jedoch nur der Fall, da die Anbieter:innen weite Kategorien und ungenaue Umschreibungen nutzen. Die beiden Probleme die sich dabei am stärksten herauskristallisiert haben, waren, dass entweder eine abstrakte weite Kategorie genannt wurde anstatt der Nennung einer konkreten Stelle oder, dass unter einer weiten Kategorie sehr viele unterschiedliche Dienste zusammengefasst wurden.

Die Ergebnisse stellen in Frage, inwiefern die Verwendung von Kategorien in Datenschutzerklärungen überhaupt dem Transparenzerfordernis genügen kann. Angesichts dieser Bedenken sollte eine funktionelle Trennung von Datenschutzerklärung und Angaben für die informierte Einwilligung erfolgen: Die Nennung von Kategorien und Umschreibungen im Rahmen der Datenschutzerklärung sollte grundsätzlich unzulässig sein, während sie in Informationen zur informierten Einwilligung zulässig sein kann, sofern sie konkret genug ist.

7 Danksagung

Die Autor:innen bedanken sich für die Unterstützung des Leibniz-WissenschaftsCampus Bremen Digital Public Health, der gemeinsam von der Leibniz-Gemeinschaft (W4/2018), dem Land Bremen und dem Leibniz-Institut für Präventionsforschung und Epidemiologie - BIPS gefördert wird.

Dieser Beitrag wurde vom Bundesministerium für Bildung und Forschung (BMBF) unter dem Förderkennzeichen 16SV8503 (UsableSec@Home-Projekt) und von der Klaus Tschira Stiftung (www.klaus-tschira-stiftung.de) gefördert.

Literaturverzeichnis

- [Ar22a] Androidrank, https://www.androidrank.org/android-most-popular-google-play-apps?category=HEALTH_AND_FITNESS, Stand: 14.08.2022.
- [Ar22b] Androidrank, <https://www.androidrank.org/android-most-popular-google-play-apps?category=MEDICAL>, Stand: 14.08.2022.

- [Ar18a] Art.-29-Datenschutzgruppe, Leitlinien für Transparenz gemäß der Verordnung 2016/679, WP 259 rev. 01, 2018.
- [Ar18b] Art.-29-Datenschutzgruppe, Leitlinien für Transparenz gemäß der Verordnung 2016/679, WP 260 rev. 01, 2018.
- [Bf22] BfArM: Prüfkriterien für die von digitalen Gesundheitsanwendungen (DiGA) und digitalen Pflegeanwendungen (DiPA) nachzuweisenden Anforderungen an den Datenschutz, 2022.
- [Bg17] BGH: Urt. v. 16.05.2017 – VI ZR 135/13, GSZ, S. 113-120, 2018.
- [Co20] ConPolicy, Innovatives Datenschutz-Einwilligungsmanagement, 2020.
- [Eu18] Europäischer Datenschutzausschuss, Endorsement 1/2018, 2018.
- [Eu23] EuGH: Urt. v. 12.01.2023 – C-154/21, NJW, S. 973-976, 2023.
- [Fra16] Franck, Lorenz: Das System der Betroffenenrechte nach der DSGVO, RDV, S. 11-119, 2016.
- [Go22] Gola, Peter; Heckmann, Dirk: DSGVO, BDSG, 3. Aufl., C. H. Beck, 2022.
- [Gi23] Github, <https://github.com/MobSF/Mobile-Security-Framework-MobSF>, Stand: 28.04.2023.
- [Ka23] Kali, <https://www.kali.org/tools/jadx/>, Stand: 28.04.2023.
- [Ka17] Kamps, Michael; Schneider, Florian: Transparenz als Herausforderung: Die Informations- und Meldepflichten der DSGVO aus Unternehmenssicht, K&R, S. 24-29, 2017.
- [Kü20] Kühling, Jürgen; Buchner, Benedikt: DSGVO, BDSG, 3. Aufl., C. H. Beck, München, 2020.
- [Lo19] Lorenz, Bernd: Datenschutzrechtliche Informationspflichten, VuR, S. 213-221, 2019.
- [Pa21] Paal, Boris; Pauly, Daniel: DSGVO, BDSG, 3. Aufl., C. H. Beck, München, 2021.
- [Pl18] Plath, Kai-Uwe: DSGVO/BDSG, 3. Aufl., Otto Schmidt, Köln, 2018.
- [Sa13] Sachs, Andreas; Meder, Miriam: Datenschutzrechtliche Anforderungen an App-Anbieter, ZD, S. 303-308, 2013.
- [Si19] Simitis, Spiros; Hornung, Gerrit; Spiecker gen. Döhmman, Indra: Datenschutzrecht, Nomos, Baden-Baden, 2019.
- [St23] Stopper, Anne; Caltrider, Jen, <https://foundation.mozilla.org/de/campaigns/googles-data-safety-labels/>, Stand 28.04.2023.
- [Ta22] Taeger, Jürgen; Gabel, Detlev: DSGVO, BDSG, TTDSG, 4. Aufl., Deutscher Fachverlag GmbH, Frankfurt am Main, 2022.
- [Wo23] Wolff, Heinrich; Brink, Stefan: BeckOK Datenschutzrecht, 43. Ed., C.H. Beck, München, 2023.

Verantwortlichkeit im Unternehmen und Verschuldensmaßstab im Rahmen des Art. 83 DSGVO

Bußgeldpraxis im Fall Deutsche Wohnen


Kai Julian Kemmler ¹

Abstract: Bereits 2019 hat die Berliner Beauftragte für Datenschutz und Informationsfreiheit gegen die Deutsche Wohnen SE ein Bußgeld in Höhe von etwa 14 Mio. Euro erlassen. Grund dafür war die Speicherung von Daten für deren Speicherung kein Bedarf (mehr) bestand. Dabei wurde durch die Behörde kein konkreter Verantwortlicher im Unternehmen im Bescheid benannt. Es steht die Frage im Raum, ob es zur Zurechnung dieses bußgeldbewehrten Verhaltens einer konkret verantwortlichen natürlichen Person bedarf. Auch ist fraglich, ob in dieser Person ein vorsätzliches oder fahrlässiges Verhalten für ein Bußgeld im Rahmen von Art. 83 DSGVO festgestellt werden müssen. Der Beitrag setzt sich mit den Argumenten, die unter anderem vor den Instanzen vorgetragen wurden, auseinander und kommt letztlich zu dem Ergebnis, dass sowohl eine konkrete zurechenbare Verantwortlichkeit im Unternehmen als auch dabei Vorsatz oder Fahrlässigkeit festgestellt werden muss. Zum Ende werden die Auswirkungen der einander bedingenden Rechtsfragen auf die Höhe eines Bußgeldes abstrakt erläutert.

Keywords: Art. 83 DSGVO, Verschulden, Rechtsträgerprinzip, Funktionsträgerprinzip, Zurechnungsnorm, Schuldgrundsatz, verfassungsrechtlich verankerter Schuldgrundsatz, Ordnungswidrigkeitenrecht, echte Strafe, Verfassungsidentität, § 41 BSDG, § 30 OWiG, Kartellrecht, objektiver Pflichtverstoß, Rechtsträger, Sanktion sui generis, Strafcharakter, unmittelbare Verbandshaftung, effet utile, Datenschutzrecht, Verbandshaftung, funktionaler Unternehmensbegriff, Effizienzgebot, Verantwortliche und Auftragsverarbeiter, Analogieverbot, Art. 103 II GG, Vorsatz, Fahrlässigkeit, § 10 OWiG, Strict-liability, Gefährdungshaftung, Gesamtumsatz, konzernweiter Umsatz

1 Einleitung

Mit Art. 83 IV bis VI DSGVO hat der Gesetzgeber ein neuartiges Bußgeldregime geschaffen, um dem Datenschutzrecht zu seiner Durchsetzung zu verhelfen. Neben der für Aufsehen sorgenden Höhe der verhängten Bußgelder ergeben sich auch Einzelprobleme, die mit der Umsetzung in deutsches Recht, der Verknüpfung zu anderen europarechtlichen Regelungen und Grundprinzipien aus dem deutschen Recht zusammenhängen, wobei alle bisherigen gerichtlichen Entscheidungen eigene Problemfelder aufwerfen.

¹ Goethe-Universität, Fachbereich Rechtswissenschaft, Theodor-W.-Adorno-Platz 1, 60323 Frankfurt a.M., kemmler@jur.uni-frankfurt.de,  <https://orcid.org/0009-0007-5241-7085>

Vorliegend sollen die sich ergebenden Probleme im Kontext der Bußgeldentscheidung gegen die Deutsche Wohnen SE genauer unter die Lupe genommen werden.

2 Chronologie – der Fall „Deutsche Wohnen“

Zum zugrundeliegenden Sachverhalt: Im Juni 2017 und März 2019 fanden bei der Deutsche Wohnen SE (im Folgenden „Deutsche Wohnen“) Prüftermine im Rahmen einer Vor-Ort-Kontrolle, § 38 IV BDSG a.F., durch die Berliner Beauftragte für Datenschutz und Informationsfreiheit (im Folgenden „Behörde“) statt [Kg22, Rn. 3, Be19]. Es wurde festgestellt, dass das Archivsystem Mieterdaten (unter anderem Identitätsnachweise, Informationen über Vormietverhältnisse, Steuer-, Krankenversicherungs- und Sozialdaten) so speicherte, dass nicht sichergestellt war, dass diese bei Entfall der Erforderlichkeit einer Speicherung gelöscht werden können. Darüber hinaus wurden auch Daten gespeichert, bei denen keine Erforderlichkeit und damit keine Zulässigkeit der Speicherung bestand, so zumindest die Behörde [Be19, S. 1].

Aufgrund dessen wurde die Deutsche Wohnen aufgefordert, bis Ende 2017 die nicht mehr erforderlichen Daten aus dem Archiv zu entfernen. Bereits nach dem ersten Termin hat die Deutsche Wohnen versucht, die Mängel in der Datenspeicherung teilweise, im Rahmen des nach ihrer Meinung Möglichen, abzustellen. Dies hatte bis zum März 2019 das Problem jedoch nicht vollständig beheben können, sodass laut Behörde noch kein rechtmäßiger Zustand erreicht war [Be19].

Auf den gewonnenen Erkenntnissen basierend erließ die Behörde einen Bußgeldbescheid gegen die Deutsche Wohnen mit einem Bußgeld in Höhe von 14.385.000 Euro.

Gegen diesen Bescheid legte die Deutsche Wohnen Einspruch beim Landgericht (LG) Berlin ein und erreichte zunächst die Verfahrenseinstellung [Lg21, Nr. 1 im Tenor]. Als Begründung wurde angeführt, dass der Bescheid an so gravierenden Fehlern gelitten habe, dass er nicht Grundlage eines Verfahrens sein könne. Es könne, so das Gericht, vorliegend nicht das Unternehmen, sondern nur eine natürliche Person Betroffene eines Bußgeldverfahrens sein. Der Bescheid ist jedoch gegen das Unternehmen selbst gerichtet.

Gegen den Beschluss des LG Berlin legte die Staatsanwaltschaft Berlin Beschwerde ein [Be21]. Das nun befassende Kammergericht (KG) Berlin hat grundsätzliche Fragen über die Ausgestaltung der Zurechnung einer Ordnungswidrigkeit im Rahmen von Verstößen gegen die DSGVO im Raum stehend gesehen und die Rechtsfrage dem EuGH vorgelegt. Dessen Entscheidung steht nach der Anfang 2023 durchgeführten mündlichen Verhandlung und den im April 2023 veröffentlichten Schlussanträgen des Generalanwalts des EuGH (GA) [Ga23] noch aus.

3 Problemfelder

Aus dem geschilderten Sachverhalt ergeben sich verschiedene, im Folgenden beleuchtete, Problemfelder. Neben der Frage, wessen Verhalten der Anknüpfungspunkt eines Bußgeldbescheids nach Art. 83 IV bis VI DSGVO sein kann, ist auch fraglich, welcher Maßstab an ein Verschulden zu stellen ist. Auch soll auf Auswirkungen dieser Probleme für die Höhe des verhängten Bußgeldes in aller Kürze eingegangen werden.

4 Grundsätzlicher Anknüpfungspunkt im OWiG

In einem Bußgeldverfahren kommt nach den Regeln des Ordnungswidrigkeitenrechts die Bebußung einer juristischen Person nicht direkt in Betracht. Vielmehr handelt es sich um die Zurechnung eines bußgeldbewährten tatbestandlichen Verhaltens einer natürlichen Person zu einer juristischen Person nach dem so genannten Rechtsträgerprinzip [No12, S. 330, WV21b, S. 343]. Die entsprechende Zurechnungsnorm ist § 30 OWiG.

In § 30 I OWiG sind leitende Stellungen natürlicher Personen enumerativ aufgelistet, bei denen eine Haftung für die juristische Person oder Personenvereinigung, für die sie verantwortlich handeln, entstehen kann. Daneben sieht § 30 IV OWiG eine Bebußung in den Fällen vor, in denen trotz Vorliegens tatbestandlicher Voraussetzungen kein Bußgeldverfahren gegen eine verantwortliche natürliche Person geführt wird, sondern direkt gegen die juristische Person, für die sie verantwortlich ist [Lg21, Rn. 11].

Für Personen, die keine Leitungsfunktion innehaben, ist die Zurechnung einer Aufsichtspflichtverletzung über § 130 I OWiG möglich. So sorgen die §§ 30 I, IV, 130 I OWiG dafür, dass auf der Rechtsfolgenseite für die Verletzung einer Pflicht das Bußgeld gegen die juristische Person als Nebenbeteiligte festgesetzt werden kann. Grundsätzlich braucht es danach aber immer eine natürliche Person, an deren Handlung angeknüpft wird [No12, S. 330, WV21b, S. 343].

Das hinter dieser Ausgestaltung stehende Prinzip ist der verfassungsrechtlich verankerte Schuldgrundsatz, Artt. 20 III, 103 II, 2 I, 1 I GG [Me22, Rn. 48, De21, Rn. 57 ff.]. Dieser besagt, dass eigenverantwortliches menschliches Verhalten die Voraussetzung jeglicher Schuld ist, um deren Bebußung es gerade geht. Der Schuldgrundsatz ist im Strafrecht dabei zwingend [Me22, Rn. 1, Kg22, Rn. 5, Lg21, Rn. 11f.].

Fraglich ist jedoch, ob der Schuldgrundsatz vorliegend auch zwingend zu beachten ist.

Einer Ansicht zufolge handelt es sich bei dem Ordnungswidrigkeitenrecht nicht um Strafrecht, sodass das Schuldprinzip im Ordnungswidrigkeitenrecht gerade nicht zwingend sei und damit auch nicht im Bußgeldrecht. Grund dafür soll sein, dass Grundlage der Bebußung nicht oder zumindest nicht immer ein strafrechtlicher Schuldvorwurf ist [Bo19a, S. 537, Be17, S. 558, De21, Rn. 57, Be20b, Rn. 36].

Jedoch ist laut BVerfG das Schuldprinzip immer dann zu beachten, wenn die Strafe eine repressive Ahndung vorsieht, aufgrund derer ein angedrohtes Übel wegen einer rechtswidrigen Zuwiderhandlung verhängt wird und damit eine echte Strafe darstellt, was nicht nur Kriminalstrafen sein müssen [Bv66, Rn. 20f., Bv76, Rn. 7, Lg21, Rn. 21, De21, Rn. 59f., Be20b, Rn. 36]. Danach handelt es sich bei einem Bußgeld immer dann um eine echte Strafe, auf die der Schuldgrundsatz Anwendung findet, wenn eine Geldbuße verhängt werden kann, die dann ein Übel darstellt.

Dem BVerfG nach gehört der Schuldgrundsatz zudem zur Verfassungsidentität und führt dazu, dass jede Strafe Schuld voraussetzt, wie jüngst in dem Beschluss zum europäischen Haftbefehl erneut ausgeführt wurde [Bv15, Rn. 48f.].

Entsprechend wird überzeugend vertreten, dass es sich bei Bußgeldrecht um Strafrecht im weiteren Sinne handelt, weshalb für die Erfüllung eines Tatbestands ebenfalls nach dem Schuldgrundsatz eine vorsätzliche oder fahrlässige Begehung eines Tatbestands notwendig ist [Ho22, Rn. 18, Fr21b, Rn. 14, Me22, Rn. 1, Ro18, Rn. 5-6].

Es gilt daher grundsätzlich, dass zur Bebußung einer juristischen Person eine natürliche Person den Verstoß zunächst entweder vorsätzlich oder fahrlässig begangen haben muss.

Hinzu kommt das Rechtsträgerprinzip, §§ 9, 30 OWiG. Dieses besagt, dass im Unternehmen der Rechtsträger für das Handeln der einzelnen Unternehmensteile verantwortlich sein muss, an das im Bußgeldverfahren angeknüpft wird. In der Folge muss eine Geldbuße auch nach dem Rechtsträgerprinzip an das Handeln eines Repräsentanten anknüpfen [Bg11, Rn. 15, Me22, Rn. 17.2, Ac12, S. 1f., Lg21, Rn. 16, 23].

5 Anknüpfungspunkt im Bußgeldverfahren nach der DSGVO

Fraglich ist, ob dies auch für ein Verfahren im Rahmen von Art. 83 IV bis VI DSGVO gilt. Dies wird teilweise in Abrede gestellt [Kg22, Rn. 5, Rn. 26f., Bo19a, S. 537, Fr21b, Rn. 14, Be20b, Rn. 35f., Be17, S. 558, Ga23, Rn. 43ff.]. Grundsätzlich handelt es sich auch hier um ein Bußgeldverfahren.

5.1 Schuldgrundsatz

Für Bußgeldverfahren nach Art. 83 IV bis VI DSGVO enthält das Bundesdatenschutzgesetz eine wirksamkeitsunterstützende analogieverweisende Regelung, nämlich § 41 I BDSG [Fr21a, Rn. 1, Lg21, Rn. 18]. Nach deren Wortlaut findet das OWiG entsprechend Anwendung. Ausgenommen davon sind die §§ 17, 35, 36 OWiG, bei denen die DSGVO abschließende Regelungen enthält [Lg21, Rn. 16]. Im Umkehrschluss bedeutet das, dass also gerade § 30 OWiG nicht ausgenommen ist. Eben das spricht dafür, dass der Schuldgrundsatz auch in Art. 83 IV bis VI DSGVO nach dem Willen des Gesetzgebers in der in § 30 OWiG getroffenen Ausgestaltung gelten soll.

Der Schuldgrundsatz könnte nach den Ausführungen zur Aussage des BVerfG darüber hinaus sogar hier zwingend sein. Dafür muss es sich um eine echte Strafe handeln, was der Fall ist, wenn die Strafe eine repressive Ahndung vorsieht, aufgrund derer ein angedrohtes Übel wegen einer rechtswidrigen Zuwiderhandlung verhängt wird (siehe 4).

Vertreten wird, dass es sich bei den Sanktionen nach Art. 83 IV bis VI DSGVO nicht um eine echte Strafe handele, sondern um Sanktionen sui generis [Ne18, Rn. 17, Be20b, Rn. 20, Be17, S. 556]. Begründet wird dies damit, dass es im Unionsrecht um die Verhinderung eines verzerrten Wettbewerbs gehe. Das Unionsrecht habe zudem eine zum deutschen Recht völlig andere Struktur bezüglich der Sanktionen, schon weil ein Vorgehen gegen Unternehmen und nicht gegen einzelne Rechtsträger vorgesehen sei, wie beispielsweise im Kartellrecht [Ho22, Rn. 9]. Diesem systematischen Argument kann auch hier wieder die Systematik des deutschen Bußgeldrechts entgegengehalten werden. Die Autoren, die eine Sanktion sui generis sehen, verweisen zwar auf Art. 83 II S. 2 lit. b) DSGVO und ziehen Parallelen zum Kartellrecht, eine überzeugende Argumentation gegen die Überlegungen zum Schuldprinzip und der ausdrücklichen Regelung, dass § 30 OWiG auch für Bußgelder nach Art. 83 IV bis VI DSGVO gelten soll, gelingt ihnen hingegen nicht.

Das wird besonders dadurch deutlich, dass diejenigen, die sich gegen den Schuldgrundsatz stellen, nicht oder nur als unbeachtliche „Problematik“ erwähnen, dass der Schuldgrundsatz integrationsfest zur Verfassungsidentität des Grundgesetzes gehört [Be20b, Rn. 36].

Zudem wird man bei der Heranziehung der oben dargestellten Kriterien des BVerfG zum Strafcharakter sagen müssen, dass eine Geldbuße von hier ca. 14 Mio. Euro und generell von maximal 2 % bzw. 4 % des Umsatzes ein deutliches Übel darstellt, das repressiv durch den Staat verhängt wird und daher einer Strafe für ein rechtswidriges Verhalten entspricht. Das gilt umso mehr in Anbetracht dessen, dass es sich um „bloßes Ordnungsrecht“ handelt [TRE19, S. 787]. Das Bußgeld soll nach dem Normzweck und sogar nach dem Wortlaut, Art. 83 I DSGVO, zudem gerade eine abschreckende Wirkung haben, es muss also als Übel empfunden werden [Ne18, Rn. 1, Be20b, Rn. 31, Fr21b, Rn. 3, Lg21, Rn. 27, 29]. Was daran ein Übel anderer Art (sui generis) sein soll, findet keine Beantwortung. Folglich gilt das Schuldprinzip auch hier.

Daher müsste es auch hier einen Repräsentanten geben, der für den rechtswidrigen Zustand verantwortlich ist und aufgrund dessen Handeln dann ein Bußgeld gegen die juristische Person ergeht. Auf diese Argumentation stützte sich auch das LG Berlin, weshalb es zu dem Schluss kam, dass das Bußgeldverfahren nicht bloß gegen die Deutsche Wohnen gerichtet werden könne, sondern die Feststellung eines Repräsentanten notwendig gewesen wäre [Lg21, Rn. 12].

5.2 Verbandshaftung

Gegen die Auffassung des LG wird, abseits der Frage des Schuldprinzips, angeführt, mit der Einführung der DSGVO und des Art. 83 DSGVO speziell sei eine unmittelbare Verbandshaftung begründet worden. Dem steht zwar § 41 I BDSG entgegen, jedoch greife hier der Anwendungsvorrang des Europarechts. § 41 I BDSG und die §§ 30, 130 OWiG seien so auszulegen, dass es zu keiner Wettbewerbsverzerrung komme und damit die effektive Durchsetzung des europäischen Rechts nicht gefährdet ist, also der effect utile Grundsatz, Art. 197 AEUV, gewahrt bleibt [Lg21, Rn. 13]. Und tatsächlich besteht im Rahmen des Haftungsregimes des § 30 OWiG ein Problem mit großen Verbänden. Je größer der Verband, desto schwerer ist es, einen verantwortlichen Repräsentanten zu finden, an dessen Verhalten dann die Bebußung geknüpft werden kann.² Das hat zur Folge, dass kleinere juristische Personen oder Personengesellschaften oder aber natürliche Personen deutlich eher belangt werden können als große Unternehmen mit erheblich größeren Datenmengen in der Verarbeitung. Das mindert die Effektivität der Durchsetzung des Datenschutzrechts und damit des Europarechts [Me22, Rn. 3a.2, Kg22, Rn. 18, Lg21, Rn. 13, Lg20, Rn. 53].

Dafür, dass auf diese Zurechnungsproblematik mit der Schaffung einer Verbandshaftung in der DSGVO reagiert worden sei, sollen auch die Erwägungsgründe (EG) der DSGVO sprechen.

In 9, 10 und 11 der EG ist von einem einheitlichen Schutzniveau die Rede, welches durch diese Verordnung erreicht werden soll. In 13 wird über gleichwertige Sanktionen in allen Mitgliedsstaaten gesprochen. In 129 kommt zum Ausdruck, dass es dafür einheitliche Werkzeuge braucht und in 148, dass diese so ausgestaltet sein sollen, dass auch tatsächlich der gewährte Schutz effektiv durchgesetzt werden kann [Kg22, Rn. 16, Lg20, Rn. 55 bis 60, Lg21, Rn. 13 aE.]. Daraus wird geschlussfolgert, auch wenn die EG mangels Rechtssatzcharakters nur zur Auslegung herangezogen werden können, dass eine an das europäische Kartellrecht angelehnte Verbandshaftung geschaffen werden sollte [Lg21, Rn. 13, Lg20, Rn. 61]. Zusammen mit dem unionsrechtlichen Effizienzgebot wird so für eine Überformung des nationalen Rechts und damit auch des Schuldprinzips argumentiert [Kg22, Rn. 9].

5.3 Unternehmensbegriff

Ebenso wird gegenüber dem Rechtsträgerprinzip argumentiert, dass dieses durch das Funktionsträgerprinzip unionsrechtlich ersetzt worden sei. Das Funktionsträgerprinzip kommt aus dem europäischen Kartellrecht. Diesem liegt in Art. 101 AEUV und

² Bereits in jüngster Vergangenheit wurde unter anderem deshalb an einem Verbandssanktionengesetz gearbeitet, welches letztlich jedoch gescheitert ist, [BH22, S. 18f., Ro18, Rn. 129 mwN]. Das dem zugrundeliegenden Problem widersprechende Argument der deutschen Regierung, es wäre trotzdem ohne Probleme eine Zurechnung über das System der §§ 9, 30 und 130 OWiG möglich, ohne einen Verantwortlichen zu identifizieren [Ga23, Rn. 60], erscheint in diesem Kontext verfehlt.

Art. 102 AEUV ein Unternehmensbegriff zugrunde, der nicht an die einzelnen Organe oder Repräsentanten anknüpft, sondern an die wirtschaftliche Unternehmenseinheit und das unabhängig davon, ob und wie der Verband in der nationalen Rechtsordnung ausgestaltet ist [We22, Rn. 25, Ho22, Rn. 9, Kg22, Rn. 15]. Nach ständiger Rechtsprechung des EuGH ist ein Unternehmen danach „jede eine wirtschaftliche Tätigkeit ausübende Einheit, unabhängig von ihrer Rechtsform und der Art ihrer Finanzierung“ [Exemplarisch: Eu91, Rn. 21, Ho22, Rn. 9, FSW16, S. 121]. Sinn und Zweck dieses funktionalen Verständnisses, also das Umfassen von jeder wirtschaftlich tätigen Einheit, ist die Sicherstellung des Wettbewerbs [FSW16, S. 121, Ho22, Rn. 9]. Dass im europäischen Kartellrecht dieser funktionale Unternehmensbegriff Anwendung findet, ist unstreitig [We22, Rn. 25, KF18, Rn. 74, Eu91, Rn. 21, Ga23, Rn. 45]. Die DSGVO übernimmt diesen Unternehmensbegriff zudem in Art. 4 Nr. 18 DSGVO. Auch hat der Verordnungsgeber in S. 3 des EG 150 einen ausdrücklichen Gleichlauf mit dem Unternehmensbegriff aus den Artt. 101, 102 AEUV angedacht.

Fraglich ist infolge der Übernahme dieser Definition aus dem Kartellrecht in die DSGVO, ob der Begriff „Unternehmen“ in Art. 83 IV bis VI DSGVO so auszulegen ist, dass auf den funktionalen Unternehmensbegriff und das Funktionsträgerprinzip anstelle des Rechtsträgerprinzips und der nationalen Zurechnungsregeln abzustellen ist, dieses Prinzip also überlagert wird.

Diese Frage hat das KG dem EuGH zur Vorabentscheidung vorgelegt [Kg22, Tenor I. 1.].

Als Argument für eine Überlagerung wird, neben den bereits oben genannten Argumenten, noch der Vorrang des Europarechts angebracht. Da es sich bei der DSGVO um Sekundärrecht in Form einer Verordnung der EU handle, sei sie insoweit verbindlich und unmittelbar anwendbar, Art. 288 AEUV [Sc22, S. 208, Fr21b, Rn. 24, 27, Kg22, Rn. 9, WV21a, Lg20, Rn. 47f., Lg21, Rn. 13]. Auch dieses Argument wird erneut durch das Effizienzgebot flankiert [Kg22, R. 10].

Dagegen spricht, dass sich dem Wortlaut des Art. 83 DSGVO keine Übernahme des funktionalen Unternehmensbegriffs entnehmen lässt, also auch keine unmittelbare Haftung, sondern nur die Haftungssumme, sofern es um ein Unternehmen geht [WV21a, S. 206]. Zudem sind Adressaten der Norm, Art. 83 IV bis VI DSGVO, Verantwortliche und Auftragsverarbeiter, ergo nicht direkt die Unternehmen. Es geht also um ein Zuwiderhandeln dieser Stellen gegen die Verhaltensanweisungen [WV21a, S. 206, Be20b, Rn. 21]. Diese Argumentation lässt sich zudem auf eine ausdrückliche Definition dieser Verantwortlichen stützen, die anderslautend zu dem Unternehmensbegriff, Art. 4 Nr. 18 DSGVO, für Verantwortliche in Art. 4 Nr. 7 DSGVO und für Auftragsverarbeiter in Art. 4 Nr. 8 DSGVO definiert sind. Der sich auftuende Widerspruch im Wortlaut vermag aufgrund der Widersprüchlichkeit und damit verbunden Unklarheit das Außerachtlassen des Schuldprinzips (siehe 5.1), auf dem sich das Rechtsträgerprinzip begründet, nicht zu rechtfertigen.

Gegen die Übernahme des funktionalen Unternehmensbegriffs und des Funktionsträgerprinzips spricht zudem, dass sich der nationale Gesetzgeber in der

Umsetzung ausdrücklich gegen eine Verbandsverantwortlichkeit entschieden hat (siehe 5.2).

Dagegen wird für das Funktionsträgerprinzip erwidert, dass es sich bei der Ausgestaltung dann eben um eine unionsrechtswidrige Umsetzung handle und § 30 OWiG im Lichte des Unionsrechts so auszulegen sei, dass nicht ein Repräsentant, sondern der Verband ein vorwerfbares Verhalten begangen haben muss. Alternativ sei Art. 83 IV bis VI DSGVO so zu verstehen, dass in ihm eine selbstständige Adressierung von „Unternehmen“ im funktionalen Sinne zu sehen sei [WV21a, S. 208, Lg21, Rn. 27, Kg22, Rn. 13]. Dieser Auffassung hält das LG bezüglich der unionsrechtskonformen Auslegung entgegen, dass gerade im Bereich des Strafrechts, bei dem aus Art. 103 II GG ein absolutes Analogieverbot abgeleitet wird, jede Auslegung, die die Wortlautgrenze überschreitet, verfassungswidrig sei [Lg21, Rn. 28]. Die zweite Überlegung stützt sich in erster Linie auf die Erwägungsgründe. Diese haben aber im Gegensatz zu § 41 I BDSG gerade keinen Rechtssatzcharakter, können den Wortlaut also nicht verdrängen [Lg21, Rn. 13, Lg20, Rn. 61].

Aufgrund dieser Argumentation verbleibt allein die Möglichkeit, dass § 41 I BDSG durch den EuGH für europarechtswidrig erklärt wird, da dieser die effektive Durchsetzung des Europarechts verhindere. Wie die Entscheidung ausfallen wird, bleibt abzuwarten. Gezeigt hat sich in der Vergangenheit jedoch, dass der EuGH den einheitlichen Unternehmensbegriff für Bußgelder im Rahmen des europäischen Kartellrechts möglichst weit gefasst sehen will und er deshalb eine Überlagerung des § 30 OWiG durch Art. 83 DSGVO bevorzugen könnte [Ho22, Rn. 9f., Kg22, Rn. 15]. Einer bloßen Nichtanwendung im Zuge dieser Überlagerung der Vorschriften aus dem nationalen Recht, wie oben erläutert und wie es zudem der GA in seinen Schlussanträgen als Konsequenz vorschlägt [Ga23, Rn. 62], stehen jedenfalls die obigen verfassungsrechtlichen Erwägungen fundamental entgegen. Dies und die konstatierte Präferenz des EuGH könnte im Hinblick auf das Schuldprinzip, welches das BVerfG für eine strafrechtliche Ahndung voraussetzt und als integrationsfesten Verfassungsgrundsatz ansieht, neue Probleme der Kollision des EuGH mit dem BVerfG bedeuten [Ac12, S. 322, Bv66, S. 336].

6 Maßstab der Unternehmensschuld

Sollte der EuGH Art. 83 IV bis VI DSGVO jedoch so auslegen, dass ein Unternehmen unmittelbar in einem Verfahren betroffen sein kann, ist die Folgefrage bezüglich des Maßstabs des vermittelten Verschuldens, ob ein bloßer objektiver Pflichtverstoß für ein Bußgeld nach Art. 83 IV bis VI DSGVO ausreicht oder ob es Vorsatz beziehungsweise Fahrlässigkeit bedarf.

Einer Ansicht zufolge wird vertreten, dass aufgrund der Aufzählung in Art. 83 II DSGVO, in dessen lit. b) ausdrücklich von einem vorsätzlichen oder fahrlässigen Verstoß die Rede ist, weshalb für die Erfüllung eines bußgeldbewährten Tatbestands im Sinne von Art. 83 IV bis VI DSGVO gerade kein Vorsatzerfordernis bestehe. Art. 83 II DSGVO normiere

Zumessungskriterien und nur hier sei Vorsatz oder Fahrlässigkeit für die Höhe eines Bußgelds relevant, im Umkehrschluss also nicht als Voraussetzung eines Bußgelds im Tatbestand [Kg22, Rn. 28, Be20b, Rn. 35].

Die DSGVO enthält darüber hinaus keine allgemeine Festlegung, dass es neben den geregelten Tatbestandsmerkmalen auch noch Vorsatz oder Fahrlässigkeit bedarf wie etwa § 15 StGB oder § 10 OWiG. Eine solche Festlegung könne sich folglich nur aus dem Gesetz über Ordnungswidrigkeiten, welches nach § 41 I BDSG partiell entsprechend anzuwenden ist, also aus dem Verweis des § 41 I BDSG auf § 10 OWiG, oder aus allgemeinen Grundsätzen ergeben [Bo19a, S. 537, Kg22, Rn. 28]. Der Verweisung von § 41 I BDSG auf § 10 OWiG folgend, kann nur vorsätzliches Handeln geahndet werden, es sei denn, Fahrlässigkeit ist ausdrücklich durch eine Geldbuße bedroht, was wiederum der Fall sein muss, da ansonsten Fahrlässigkeit als Zumessungskriterium leerlaufen würde. Damit müsste auch für ein Bußgeld nach Art. 83 IV bis VI DSGVO Vorsatz oder Fahrlässigkeit festgestellt werden.

Nach der erstgenannten Ansicht wird entgegnet, dass es sich bei § 41 I BDSG auch hiernach um eine unionsrechtswidrige Regelung handle, da eine nationale Vorschrift nicht die effektive Durchsetzung des Unionsrechts hindern könne [Be17, S. 558f., Ga23, Rn. 53f.]. Schon aus diesem Grund wird durch den GA argumentiert [Ga23, Rn. 82], dass ein Erfordernis von Vorsatz und Fahrlässigkeit aus der Systematik des Art. 83 I bis III DSGVO herzuleiten sei [Ga23, Rn. 81], wodurch im Ergebnis auch dieser zu einem Erfordernis von Vorsatz und Fahrlässigkeit kommt. Er führt in seiner Argumentation aus, dass dies aus der Anforderung der Verhältnismäßigkeit in Abs. 1 und aus Vorsatz und Fahrlässigkeit als Zumessungskriterium, Abs. 2, abzuleiten sei. Ebenso aus Abs. 3, da dieser vorsieht, dass selbst wenn ein Verantwortlicher oder Auftragsverarbeiter vorsätzlich oder fahrlässig gegen mehrere Bestimmungen der Verordnung verstoßen habe, die Höhe der Geldbuße nicht den Betrag für den schwerwiegendsten Verstoß übersteigen kann, wodurch sich im Ergebnis also ergibt, „dass rein objektive Verstöße für die Sanktion insoweit ohne Bedeutung sind, als sie nicht kumulativ zu vorsätzlichen oder fahrlässigen Verstößen berücksichtigt werden“ [Ga23, Rn. 81]. Dadurch wird aber auch deutlich, dass der Verweis an sich auf das unionsrechtliche Effizienzgebot jedoch aufgrund des oben erläuterten Schuldgrundsatzes problematisch erscheint. Beim Ordnungswidrigkeitenrecht handelt es sich um Strafrecht im weiteren Sinne, bei dem schon nach dem Schuldgrundsatz Vorsatz oder Fahrlässigkeit nachgewiesen werden muss.

Hinzu kommt, dass der Schuldgrundsatz auch eine Ausprägung des Verhältnismäßigkeitsgrundsatzes darstellt und schon deshalb die Notwendigkeit vorsätzlichen oder fahrlässigen Handelns in Art. 83 IV bis VI DSGVO hineinzulesen ist [Ho22, Rn. 18, Be17, S. 558]. Der Grundsatz der Verhältnismäßigkeit findet ausdrücklich auch im Unionsrecht Anwendung, Art. 5 IV EUV.

Zu beachten ist auch, dass gerade die erstgenannte Ansicht, die nur ein Zumessungskriterium erblickt, übersieht, dass unionsrechtlich unter Vorsatz und Fahrlässigkeit nicht dasselbe verstanden wird wie im nationalen Recht. Unter

vorsätzlichem Verhalten versteht die Kommission als europäische Kartellbehörde einen weitreichenderen Maßstab als nach nationalem Recht. So lässt sie für die Frage, ob vorsätzlich gehandelt wurde, schon bloße Kenntnis von tatbestandserfüllenden Gegebenheiten ausreichen. Eines Bewusstseins bezüglich der Rechtsverletzung bedarf es danach nicht [Eu05, Rn. 73-75, EL21, 131 Rn. 18]. Das schließt im Umkehrschluss aber gerade die Lesart aus, die vertritt, dass es wegen des Erwähnens von Vorsatz und Fahrlässigkeit in Art. 83 II S. 2 lit. b) DSGVO eben dieser tatbestandlich nicht bedarf [Zur Begriffsbestimmung des Vorsatzes im Unionsrecht: Be20b, Rn. 37]. Dem Schuldgrundsatz ist also nicht schon dadurch genüge getan, weil Vorsatz und Fahrlässigkeit in der Zumessung Beachtung finden. Das schließt aus diesem Grund aber auch die Lesart aus, die aus der Systematik des Art. 83 I bis III DSGVO ein Erfordernis von Vorsatz und Fahrlässigkeit herleiten.

Dies spricht ebenfalls für ein aus dem nationalen Recht kommendes Vorsatzerfordernis, wobei teilweise von Verfechtern der erstgenannten Ansicht, die reine Zumessungskriterien erblickt, zumindest zugestanden wird, dass es tatbestandlich einer Fahrlässigkeit als notwendiges Merkmal zur Sanktionierung bedürfe, sonst könne schon keine Zumessung eröffnet sein [Be17, 558, Abs. 4.2].

Überaus problematisch erscheinen im Rahmen des Funktionsträgerprinzips zudem Verstöße in Kapitalgesellschaften, bei denen zwar auch irgendjemand für das Unternehmen formal verantwortlich ist, beispielsweise der Vorstand, dieser jedoch in der Regel schon aufgrund der hierarchischen Strukturen und einer dezentralisierten Beweisführung nicht die verantwortliche Person sein dürfte, aber Vorsatz oder Fahrlässigkeit doch eine Anknüpfungsperson bräuchte. Sogar die Kommission knüpft im Kartellrecht an ein vorhandenes Bewusstsein an [EL21, Rn.17]. In einer maximal weiten Lesart müssten Handlungen aller Mitarbeiter eines Teils des Unternehmens dem ganzen Unternehmen, ohne einen konkreten Mitarbeiter zu ermitteln, zugerechnet werden, wollte man keine konkrete Person herausstellen. Letztlich würde die Bebußung faktisch direkt an der juristischen Person anknüpfen müssen. Diese kann jedoch gerade nicht selbst, auch nicht im Kartellrecht, vorsätzlich oder fahrlässig handeln [EL21, Rn.17]. Nur natürliche Personen können Vorsatz entwickeln oder fahrlässig handeln. Eine Anknüpfung direkt an sämtliche Mitarbeiter, wie im europäischen Kartellrecht teilweise vertreten [Be17, 556 Abs. 3.1] oder nur die vertretungsberechtigten Mitarbeiter [EL21, Rn. 17, Eu83, Rn. 97] entfernt sich auch von dem Schuldgrundsatz.

Zusammenfassend bedeutet das, dass dem Unternehmen nach europäischem Kartellrecht alle begangenen Handlungen der nach außen (zumindest) handlungsberechtigten Personen pauschal zugerechnet werden und ein maximal weites Fahrlässigkeitsverständnis angelegt wird.

Sogar eine Zurechnung der Handlungen externer Dienstleister, die gar nicht Teil des Unternehmens sind, erscheint möglich [Eu16, S. 428, Ga23, Rn. 57 spricht von „Mitarbeitern im weiteren Sinne“]. Übertrüge man dieses Haftungsregime, dann käme es faktisch zu einer bloßen Gefährdungshaftung durch das Ausreichen einer auf einem

objektiven Pflichtverstoß begründeten Verantwortlichkeit. Dies hat sowohl das KG als auch der GA, als Strict-liability-Grundsatz bezeichnet, festgestellt [Kg22, Rn. 28, Ga23, Rn. 63].

Gerade jedoch die Bedenken, die ein Aussetzen des Schuldgrundsatzes, der Teil der Verfassungsidentität ist, hervorrufen und die ausdrückliche Anordnung im nationalen Recht, dass die OWiG anzuwenden sei sowie die bisherige, oben angesprochene Rechtsprechung des BVerfG eben zum Schuldgrundsatz, sind überzeugende Argumente, die gegen eine Verantwortlichkeit aufgrund eines bloß objektiven Pflichtverstoßes sprechen. In der praktischen Umsetzung wird es zudem wenig problematisch sein, mindestens Fahrlässigkeit festzustellen [Be20b, Rn. 37, Bo19a, S. 537]. Daher ist das Argument, durch das Erfordernis von Vorsatz oder Fahrlässigkeit sei dem *effet utile* Grundsatz nicht ausreichend Rechnung getragen, wenig überzeugend. Folgt man dieser Argumentation, so wird an das Verhalten eines Verantwortlichen die Frage zu stellen sein, ob er vorsätzlich oder fahrlässig gehandelt hat. Das bloße objektive Erfüllen tatbestandlicher Voraussetzungen reicht demnach nicht aus.

7 Höhe des Bußgelds

Wie erwähnt ist ein vielfach problematisierter Punkt in der Literatur und bereits ergangenen Entscheidungen zudem die Höhe der verhängten Bußgelder, auf die in diesem begrenzten Rahmen nur allgemein eingegangen und aufgezeigt werden soll, inwiefern sich die vorhergehenden Erläuterungen auch bei der Bußgeldbemessung auswirken.

Die Geldbuße ist zunächst nur eine Möglichkeit, wie die jeweilige Behörde auf einen rechtswidrigen Zustand reagieren kann. In Art. 83 II S. 1 DSGVO heißt es, dass die Geldbuße zusätzlich oder anstelle einer Maßnahme nach Art. 58 II lit. a bis h und j DSGVO verhängt werden kann [Ne18, Rn. 14]. Während Art. 83 IV bis VI DSGVO die Frage, ob überhaupt ein Bußgeld verhängt werden kann thematisiert und einen Rahmen für die maximale Höhe definiert, kann die jeweilige Behörde die schlussendliche Höhe des Bußgelds festsetzen, wobei die in Art. 83 II DSGVO definierten Zumessungskriterien zur Bestimmung des verwirklichten Unrechts zu beachten sind [BT20, S. 122, Bo19b, Rn. 37]. Die mögliche Höhe des Bußgeldes ist in der DSGVO dabei an das Kartellrecht angelehnt [FSW16, S. 120, RLE18, § 24, Rn. 52, Fr21b, Rn. 20]. Damit erlangt erneut die oben erläuterte Fragestellung zum funktionalen Unternehmensbegriff Relevanz, ob nämlich der konzernweite Umsatz Grundlage für die Bebußung sein soll, auch wenn der Wortlaut der Bußgeldvorschriften nur von einzelnen Unternehmen spricht [TRE19, S. 783, Ds19, S. 3, Lg20, Rn. 85]. Verkannt wäre jedoch, wenn dem funktionalen Verständnis gefolgt wird, dass im europäischen wie nationalen kartellrechtlichen Kontext nicht der Gesamtumsatz maßgeblich ist, sondern nur der konzernweite Umsatz, der mit dem Verstoß gegen das Kartellrecht erwirtschaftet wurde [Bk21, Rn. 9f., Eu06, Rn. 5]. Wird zur Bußgeldbemessung doch der Gesamtumsatz herangezogen, dann sei zudem in den meisten Fällen schon nicht möglich, zu ermitteln, wie hoch der Umsatz ist, der durch

einen konkreten Verstoß gegen die DSGVO entstanden ist. Es fehlt somit an einer Übertragbarkeit [TRE19, Rn. 19, Me21, S. 3]. Gelten müsste in diesem Verständnis daher, dass „Unternehmen“ wörtlich zu verstehen und nur der Umsatz der einzelnen handelnden Einheit umfasst ist, was zu deutlich niedrigeren Bußgeldern, auch im konkreten Fall, führen könnte.

Unabhängig davon wären auf der Rechtsfolgenseite Bußgelder in Höhe von maximal 2 % des Vorjahresumsatzes oder 10 Mio. Euro, je nachdem welcher der Beträge höher ist, Art. 83 IV, bzw. 4 % des Vorjahresumsatzes oder 20 Mio. Euro, ebenfalls der höhere Betrag, Art. 83 V, VI, möglich.

Im Rahmen der Zumessung ist dann bei der Ausübung des Ermessens nach Art. 83 I DSGVO zu beachten, dass eine Bebußung wirksam, verhältnismäßig und abschreckend sein muss. Die maximale Höhe ergibt sich hierbei nicht kumuliert aus der Summe der maximal möglichen Beträge, sondern Art. 83 III DSGVO sieht vor, dass nur ein Bußgeld, abhängig vom Maximalbetrag des schwersten Verstoßes, für jeden einzelnen Verstoß verhängt werden kann, auch wenn mehrere Bußgeldtatbestände des Art. 83 DSGVO erfüllt sind. Insofern werden mehrere Vorgänge in Fällen derselben verbundenen Verarbeitung in Tateinheit verklammert [Ed22, Rn. 16, TRE19, Rn. 31f., Fr21b, Rn. 16].

Im Ergebnis zeigt sich, dass die offene, durch den EuGH zu entscheidende Rechtsfrage nach dem anzuwendenden Unternehmensbegriff auch Einfluss auf die Höhe der Bußgeldbemessung haben wird.

8 Schlusswort

Wie der ergangene Bescheid schlussendlich aussehen wird, nachdem der EuGH geurteilt hat, ist damit offen. Jede Richtungsentscheidung dürfte das endgültige Ergebnis beeinflussen. Einen Ausblick könnten die Schlussanträge des GA jedoch bereits geben, wonach zwar die Möglichkeit einer Unternehmensverantwortlichkeit ohne an einer konkreten Person anzuknüpfen positiv gesehen [Ga23, Rn. 86] aber letztlich im Sinne der Deutschen Wohnen einer verschuldensunabhängigen Haftung eine Absage erteilt wird [Ga23, Rn. 81]. Es bleibt damit nicht auszuschließen, dass es letzten Endes bei der Aufhebung und Verfahrenseinstellung, die das LG Berlin durchgeführt hatte, bleibt.

Literaturverzeichnis

- [Ac12] Achenbach, H.: Verbandsgeldbuße und Aufsichtspflichtverletzung (§§ 30 und 130 OWiG) – Grundlagen und aktuelle Probleme. NZWiSt • Neue Zeitschrift für Wirtschafts-, Steuer- und Unternehmensstrafrecht 9/1, S. 321–328, 2012.
- [Be17] Bergt, M.: Sanktionierung von Verstößen gegen die Datenschutz-Grundverordnung. Datenschutz und Datensicherheit - DuD 9/41, S. 555–561, 2017.

- [Be19] Berliner Beauftragte für Datenschutz und Informationsfreiheit: Berliner Datenschutzbeauftragte verhängt Bußgeld gegen Immobiliengesellschaft. Pressemitteilung Nr. 711.412.1, Berlin, 2019.
- [Be20b] Bergt, M.: DSGVO Art. 83 Allgemeine Bedingungen für die Verhängung von Geldbußen. In (Kühling, J.; Buchner, B. Hrsg.): Datenschutz-Grundverordnung BDSG. Kommentar. C.H. Beck, München, 2020.
- [Be21] Berliner Beauftragte für Datenschutz und Informationsfreiheit: Bußgeldbescheid gegen Deutsche Wohnen SE: Beschwerde gegen Einstellung des Verfahrens eingelegt. Pressemitteilung Nr. 711.439.1, Berlin, 2021.
- [Bg11] BGHSt 57, 193.
- [BH22] Baur, A.; Holle, M.: Verbandssanktionierung – Zurück auf Los? Eine Zwischenbilanz zur gescheiterten Neuaufstellung. ZRP - Zeitschrift für Rechtspolitik 1/55, S. 18, 2022.
- [Bk21] Bundeskartellamt: Leitlinien für die Bußgeldzumessung in Kartellordnungswidrigkeitenverfahren. https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Leitlinien/Bußgeldleitlinien_Oktober2021.pdf;jsessionid=95ED16BE169A72C3DDD4DA7D243F0382.1_cid371?__blob=publicationFile&v=8, Stand: 02.06.2022.
- [Bo19a] Boms, W.: Ahndung von Ordnungswidrigkeiten nach der DS-GVO in Deutschland. ZD - Zeitschrift für Datenschutz 12/9, S. 536–540, 2019.
- [Bo19b] Boehm, F.: DSGVO Art. 83 Allgemeine Bedingungen für die Verhängung von Geldbußen. In (Simitis, S.; Hornung, G.; Spiecker Döhmann, I. Hrsg.): Datenschutzrecht. DSGVO mit BDSG. Nomos, Baden-Baden, 2019.
- [BT20] Behr, N.; Tannen, F.: Droht das Zeitalter der Datenschutzgeldbußen? – Haftungsrisiken wegen Datenschutzverstößen und was Unternehmen dagegen tun können. CCZ - Corporate Compliance Zeitschrift 3/13, S. 120–126, 2020.
- [Bv15] BVerfGE 140, 317.
- [Bv66] BVerfGE 20, 323.
- [Bv76] BVerfGE 42, 261.
- [De21] Degenhart, C.: GG Art. 103 Rechtliches Gehör, Grundrechte des Angeklagten. In (Sachs, M. Hrsg.): Grundgesetz. Kommentar. C.H. Beck, München, 2021.
- [Ds19] Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder: Konzept der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder zur Bußgeldzumessung in Verfahren gegen Unternehmen. [https://www.datenschutzkonferenz-online.de/media/ah/20191016_bußgeldkonzept.pdf](https://www.datenschutzkonferenz-online.de/media/ah/20191016_bu%C3%9Fgeldkonzept.pdf).
- [Ed22] EDPB - European Data Protection Board: Guidelines 04/2022 on the calculation of administrative fines under the GDPR. https://edpb.europa.eu/system/files/2022-05/edpb_guidelines_042022_calculationofadministrativefines_en.pdf, Stand: 31.05.2022.

- [EL21] Emmerich, V.; Lange, K. W.: Kartellrecht. Ein Studienbuch. C.H. Beck, München, 2021.
- [Eu05] EuG, Urteil vom 29.11.2005, T-52/02, ECLI:EU:T:2005:429.
- [Eu06] Europäische Kommission: Leitlinien für das Verfahren zur Festsetzung von Geldbußen gemäß Artikel 23 Absatz 2 Buchstabe a) der Verordnung (EG) Nr. 1/2003. [https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52006XC0901\(01\)&from=DE](https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52006XC0901(01)&from=DE), Stand: 02.06.2022.
- [Eu16] EuGH, Urteil vom 21.06.2016, C-542/14, ECLI:EU:C:2016:578.
- [Eu83] EuGH, Urteil vom 07.06.1983, 100 bis 103/80, ECLI:EU:C:1983:158.
- [Eu91] EuGH, Urteil vom 23.04.1991, C-41/90, ECLI:EU:C:1991:161.
- [Fr21a] Frenzel, M.: BDSG § 41 Anwendung der Vorschriften über das Bußgeld- und Strafverfahren. In (Paal, B. P.; Pauly, D. A.; Ernst, S. Hrsg.): Datenschutz-Grundverordnung. Bundesdatenschutzgesetz. C.H. Beck, München, 2021.
- [Fr21b] Frenzel, M.: DSGVO Art. 83 Allgemeine Bedingungen für die Verhängung von Geldbußen. In (Paal, B. P.; Pauly, D. A.; Ernst, S. Hrsg.): Datenschutz-Grundverordnung. Bundesdatenschutzgesetz. C.H. Beck, München, 2021.
- [FSW16] Faust, S.; Spittka, J.; Wybitul Tim: Milliardenbußgelder nach der DS-GVO? Ein Überblick über die neuen Sanktionen bei Verstößen gegen den Datenschutz. ZD - Zeitschrift für Datenschutz 3/6, S. 120–125, 2016.
- [Ga23] GA Manuel Campos Sánchez-Bordona, C-807/21, ECLI:EU:C:2023:360.
- [Ho22] Holländer, C.: DSGVO Art. 83 Allgemeine Bedingungen für die Verhängung von Geldbußen. In (Brink, S.; Wolff, H. A. Hrsg.): BeckOK Datenschutzrecht. C.H. Beck, München, 2022.
- [KF18] Koenig, C.; Förtsch, B.: AEUV Art. 107 [Zulässigkeit von staatlichen Beihilfen]. In (Streinz, R. et al. Hrsg.): EUV/AEUV. Vertrag über die Europäische Union - Vertrag über die Arbeitsweise der Europäischen Union - Charta der Grundrechte der Europäischen Union. C.H. Beck, München, 2018.
- [Kg22] KG Berlin, EuGH-Vorlage vom 06.12.2022 – 3 Ws 250/21 - Juris.
- [Lg20] LG Bonn, Urteil vom 11.11.2020 – 29 OWi 1/20 - Juris.
- [Lg21] LG Berlin, Beschluss vom 18.02.2021 – 526 OWi LG 1/20 - Juris.
- [Me21] Meyer, S.: Bußgelder bei Datenschutzverstößen: Alles nur halb so schlimm? ZVertriebsR - Zeitschrift für Vertriebsrecht 1/10, S. 1–3, 2021.
- [Me22] Meyberg, A.: OWiG § 30 Geldbuße gegen juristische Personen und Personenvereinigungen. In (Graf, J. Hrsg.): BeckOK OWiG. C.H. Beck, München, 2022.
- [Ne18] Nemitz, P.: DSGVO Art. 83 Allgemeine Bedingungen für die Verhängung von Geldbußen. In (Ehmann, E. et al. Hrsg.): DS-GVO. Datenschutz-Grundverordnung Kommentar. C.H. Beck; LexisNexis, München, Wien, 2018.

- [No12] Noak, T.: Einführung ins Ordnungswidrigkeitenrecht – Teil 2. Rechtsfolgen. Zeitschrift für das Juristische Studium (ZJS) 3, S. 329–334, 2012.
- [RLE18] Rüpke, G.; Lewinski, K. von; Eckhardt, J.: Datenschutzrecht. Grundlagen und europarechtliche Neugestaltung. C.H. Beck, München, 2018.
- [Ro18] Rogall, K.: § 30 Geldbuße gegen juristische Personen und Personenvereinigungen. In (Mitsch, W. Hrsg.): Karlsruher Kommentar zum Gesetz über Ordnungswidrigkeiten. C.H. Beck, München, 2018.
- [Sc22] Schnabel, C.: Kommentar zu KG Berlin, Beschluss vom 06.12.2021, 3 Ws 250/21. Kommunikation & Recht – K&R 2/25, S. 139–140, 2022.
- [TRE19] Timmer, H.; Radlanski, P.; Eisenfeld, A.: Die Bußgeldbemessung bei DSGVO-Verstößen. Computer und Recht 12/35, S. 782–788, 2019.
- [We22] Weiß, W.: AEUV Art. 101 (ex-Art. 81 EGV) [Verbot wettbewerbsbeschränkender Maßnahmen]. In (Calliess, C.; Ruffert, M. Hrsg.): EUV/AEUV. Das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta Kommentar. C.H. Beck, München, 2022.
- [WV21a] Wybitul, T.; Venn, N.: Die bußgeldrechtliche Haftung von Unternehmen nach Art. 83 DSGVO. NSTz - Neue Zeitschrift für Strafrecht 4, S. 204–209, 2021.
- [WV21b] Wybitul, T.; Venn, N.: Verteidigung von Unternehmen gegen Geldbußen nach Art. 83 DSGVO. ZD - Zeitschrift für Datenschutz 7/11, S. 343–348, 2021.

Umbruch im Datenschutz

Datenschutzvorsorge in der Cybersicherheitsforschung, Verifizierung von Anonymität und Berücksichtigung der Nutzerbedürfnisse

Annika Selzer¹ 

Abstract: Das im Zentrum des Datenschutzes stehende Ziel ist es, im Rahmen personenbezogener Datenverarbeitungen die Grundrechte und -freiheiten betroffener Personen zu schützen. Doch der Ruf des Datenschutzes scheint sich in den letzten Jahren stetig verschlechtert zu haben und selbst bei betroffenen Personen, deren Rechte und Freiheiten geschützt werden sollen, scheint der Datenschutz keinen hohen Stellenwert zu genießen, sondern fast schon als lästig empfunden zu werden. Ist es vor diesem Hintergrund Zeit für ein Umdenken im Datenschutz?

Keywords: Comic, Datenschutzinformation, Datenschutzmetriken, Datenschutzvorsorge, Einwilligung, Informationsgrafik, offensive Cybersicherheitsforschung

1 Herausforderung: Datenschutz zukunftsfähig machen

Das Datenschutzrecht hat das übergeordnete Ziel, Menschen im Rahmen der sie betreffenden Verarbeitung personenbezogener Daten vor Eingriffen in ihre Grundrechte und -freiheiten zu schützen. Hierbei wird durch die in Europa anwendbaren Rechtsakte zum Datenschutz unter Achtung dieser Schutzpflicht ein Ausgleich zwischen den Interessen der von einer personenbezogenen Datenverarbeitung betroffenen Personen, datenverarbeitenden Organisationen und der Gesellschaft als Ganzes geschaffen [SHS19, KB20]. Durch diese Rechtsakte – z.B. die Datenschutz-Grundverordnung (DSGVO) – wird die Verarbeitung personenbezogener Daten strengen Regeln unterstellt, so dass unter anderem die Datenverarbeitung nur mittels einer vorhandenen, einschlägigen Rechtsgrundlage rechtmäßig möglich ist, die betroffenen Personen umfangreich über die sie betreffenden Datenverarbeitungen informiert werden müssen und die datenverarbeitende Organisation umfangreiche technische und organisatorische Schutzmaßnahmen ergreifen muss, um die mit einer personenbezogenen Datenverarbeitung einhergehenden Risiken für die Rechte und Freiheiten betroffener Personen zu reduzieren. Um wiederum die Interessen der datenverarbeitenden Organisationen und der Öffentlichkeit zu adressieren, sehen die anwendbaren Rechtsakte u.a. unter bestimmten Voraussetzungen das (i.d.R. seitens der datenverarbeitenden Stelle bestehende) berechtigte Interesse sowie

¹ Fraunhofer-Institut für Sichere Informationstechnologie SIT und Nationales Forschungszentrum für angewandte Cybersicherheit ATHENE, Rheinstraße 75, 64295 Darmstadt, Deutschland,

Annika.selzer@sit.fraunhofer.de,  <http://orcid.org/0000-0001-8859-4808>.

öffentliche Interessen als rechtmäßige Rechtsgrundlage der Datenverarbeitung vor, enthalten Ausnahmen für die wissenschaftliche Forschung und erlauben bei der Implementierung technischer und organisatorischer Maßnahmen die Berücksichtigung der wirtschaftlichen Interessen der datenverarbeitenden Organisation.

In den letzten Jahren scheint sich der Ruf des Datenschutzrechts in Organisationen stetig verschlechtert zu haben: Von den ab dem Jahr 2018 in den datenverarbeitenden Organisationen bestehenden Unsicherheiten rund um die rechtskonforme Umsetzung der DSGVO², über Unmut und Unsicherheiten, die in den letzten fünf Jahren durch einige sehr hohe Bußgelder gegen DSGVO-Verstöße entstanden,³ bis hin zu der Wahrnehmung des Datenschutzes als Verhinderer des technologischen Fortschrittes – u.a. wegen Unsicherheiten in Bezug auf die Verarbeitung anonymisierter (ehemals) personenbezogener Daten [ST21a]. Aber auch bei den betroffenen Personen selbst scheint das Datenschutzrecht mittlerweile, insbesondere aufgrund der Vielzahl der von ihnen abverlangten Einwilligungen, z.B. im Zusammenhang mit Webtracking-Cookies [DHS21], einen zunehmend schlechteren Ruf zu haben. Schlussendlich monieren Forschungseinrichtungen, dass sich moderne Methoden der Cybersicherheitsforschung nicht rechtssicher mit geltendem Datenschutzrecht zusammenbringen lassen, da das Datenschutzrecht von überholten Gegebenheiten in der personenbezogenen (Cybersicherheits-)Forschung ausgeht [BSS23].

Vor diesem Hintergrund will der vorliegende Beitrag anhand von drei Beispielen diskutieren, inwiefern es Zeit für ein Umdenken im Datenschutzrecht ist. Diese drei Beispiele sind Datenschutzrecht in der offensiven Cybersicherheitsforschung (Kapitel 2), Datenschutz und Anonymität (Kapitel 3) und Nutzerbedürfnisse im Datenschutzrecht (Kapitel 4). Der Beitrag endet mit einem Fazit (Kapitel 5).

2 Datenschutz in der offensiven Cybersicherheitsforschung

Da neuartige Angriffe und Schwachstellen auf IT-Anwendungen, -Systeme und -Infrastrukturen immer komplexer werden, simulieren Cybersicherheitsforschende im Rahmen ihrer Forschung die Rolle von böswilligen Angreifenden, um akute Bedrohungen für IT zu verstehen. Aus den hierdurch gewonnenen Erkenntnissen leiten die Forschenden Abhilfemaßnahmen gegen neuartige Angriffe und Schwachstellen ab und unterstützen angegriffene Organisationen in der aktiven Angriffsabwehr. Hierbei sind Werkzeuge und Methoden der aktiven Angriffsabwehr häufig sehr ähnlich oder sogar gleich zu den für Angriffe gegen IT verwendeten Werkzeugen und Methoden, werden von den Forschenden

² Die DSGVO wurde am 25. Mai 2018 anwendbar. Zur Unsicherheit: <https://www.bitkom.org/Presse/Presseinformation/Datenschutz-deutsche-Wirtschaft-2022-DS-GVO-wenig-Wettbewerbsvorteile>.

³ Einen Überblick über Bußgelder, die aufgrund von DSGVO-Verstößen verhängt wurden, findet sich auf www.enforcementtracker.com.

jedoch eingesetzt, um Angriffe abzuwehren und die IT somit zu stärken, anstatt sie – wie ein böswilliger Angreifer – zu schwächen [BSS23].

Den hier skizzierten Bereich der Cybersicherheitsforschung, bei dem Forschende mit dem Ziel der Entwicklung von Abhilfemaßnahmen die Rolle böswilliger Angreifer simulieren und ggf. ähnliche oder gleiche Werkzeuge und Methoden wie Angreifer einsetzen, um jedoch im Gegensatz zu Angreifern die IT sicherer zu gestalten, nennt man *offensive Cybersicherheitsforschung*. Die offensive Cybersicherheitsforschung leistet einen wertvollen und notwendigen Beitrag dazu, unsere Gesellschaft vor personenbezogenen Datendiebstählen und Abgreifen von Geschäftsgeheimnissen, aber auch vor Angriffen auf unsere Kritischen Infrastrukturen zu schützen [SW22, SS22].

Die offensive Cybersicherheitsforschung ist in Bezug auf den Datenschutz mit großen Rechtsunsicherheiten verbunden. Dies liegt vor allem darin begründet, dass die DSGVO – und andere das Datenschutzrecht regelnde Rechtsakte – nicht vor dem Hintergrund entstanden sind, speziell die offensive Cybersicherheitsforschung zu regeln, sondern allgemeingültige Regeln für die Verarbeitung personenbezogener Daten zu schaffen. Die DSGVO gibt insofern einen Regelungsrahmen für (fast) jede denkbare Form der Verarbeitung personenbezogener Daten vor, enthält aber keine Regelungen, die sich speziell auf die offensive Cybersicherheitsforschung beziehen. Dieser Umstand führt dazu, dass die technischen Gegebenheiten der offensiven Cybersicherheitsforschung häufig nicht zu den in den für sie geltenden Rechtsakten normierten Pflichten passen [GS23].

Beispiel: Die DSGVO geht davon aus, dass jede personenbezogene Datenverarbeitung vor der Erhebung von Daten detailliert geplant und rechtlich vorbereitet wird. Offensive Cybersicherheitsforschung zielt hingegen in der Regel nicht darauf ab, personenbezogene Daten zu verarbeiten und kann derartige Datenverarbeitungen in der Regel auch nicht antizipieren. Gleichwohl kann es passieren, dass – insbesondere im Rahmen der Abwehr eines Cyberangriffs – personenbezogene Daten von Forschenden erhoben werden, ohne dass dies vorher absehbar oder gar datenschutzrechtlich vorbereitet war [GS23].

Im eben beschriebenen Beispiel sind die Anforderungen der DSGVO aufgrund der Unvorhersehbarkeit personenbezogener Datenverarbeitungen nicht oder nur verspätet umsetzbar, was wiederum die Rechtsunsicherheit des Bereichs der offensiven Cybersicherheitsforschung bedingt. Um diesen wichtigen Bereich der Cybersicherheitsforschung jedoch nicht – bedingt durch die Rechtsunsicherheiten, mit denen dieser Bereich der Forschung derzeit belastet wird – auszubremsen, bedarf es eines Umdenkens im Datenschutzrecht dahingehend, dass sich dieses auf unplanbare Datenverarbeitungen im Rahmen der offensiven Cybersicherheitsforschung einstellt und hierfür adäquate Regeln vorsieht, die eine rechtskonforme Datenverarbeitung in derartigen Fällen ermöglicht. Eine Möglichkeit hierfür könnte in einer sorgfältigen Datenschutzvorsorge liegen, im Rahmen derer eine ggf. *bevorstehende* personenbezogene Datenverarbeitung vorbereitet wird, indem *Annahmen* zu der Datenverarbeitung getroffen werden, die – insbesondere basierend auf Erfahrungen vorangegangener, ähnlicher offensiver

Forschungsaktivitäten – als *wahrscheinlich* anzunehmen sind. Insofern wäre für eine bevorstehende, offensive Forschungsaktivität beispielsweise eine einschlägige Rechtsgrundlage für die Verarbeitung personenbezogener Daten nach Art. 6 DSGVO zu identifizieren, nicht aber eine Rechtsgrundlage nach Art. 9 DSGVO, wenn die Erfahrungen aus der Vergangenheit gezeigt haben, dass im Rahmen ähnlicher Forschungsaktivitäten bisher nie besondere Kategorien personenbezogener Daten verarbeitet wurden. Unter den gleichen Annahmen sollten u.a. ein Mindestmaß an technischen und organisatorischen Maßnahmen umgesetzt und die Umsetzung der Datenschutzvorsorge – u.a. in Form eines Eintrags der offensiven Forschungsaktivität und der damit ggf. erfolgenden personenbezogenen Datenverarbeitung in das Verzeichnis der Verarbeitungstätigkeiten – dokumentiert werden [SS22, BSS23]. Sofern die Datenschutzvorsorge unter der Berücksichtigung der Gegebenheiten der offensiven Cybersicherheitsforschung und der Risiken für die Rechte und Freiheiten der betroffenen Personen näher ausgestaltet würde, könnte in dieser somit ein neues Instrument des Datenschutzrechts liegen, das es ermöglicht, unplanbare Datenverarbeitungen rechtskonform und -sicher umzusetzen.

Nur durch ein Umdenken im Datenschutzrecht, das die offensive Cybersicherheitsforschung in Zukunft rechtssicher ausführbar macht, kann letztlich sichergestellt werden, dass unsere Gesellschaft weiter von den Vorteilen dieser Forschung in Form des hohen Schutzes unserer IT profitieren kann. Hierbei ist zu beachten, dass Wissenschaft – sozusagen als Gegenstück zur Rechtsunsicherheit – nicht im rechtsfreien Raum durchgeführt werden darf und unbedingt adäquater Regeln bedarf, die das Wohl der Gesellschaft in den Vordergrund der Forschung stellen [GS23].

3 Datenschutzrecht und Anonymität

Das vorliegende Kapitel skizziert die Frage des Bedarfs eines Umbruchs im Datenschutzrecht am Beispiel der Anonymität und hiermit verbundener, datenschutzrechtlich motivierter Herausforderungen. Anonymisierte Daten sind Daten, die sich nicht mehr auf eine identifizierte oder identifizierbare natürliche Person beziehen (Erwgr. 26 DSGVO) und sind insofern das Gegenteil der personenbezogenen Daten. Sofern eine Datenverarbeitung mit ausschließlich anonymisierten Daten möglich ist, finden die Regelungen der DSGVO – und anderer für den Datenschutz einschlägige Rechtsakte – auf diese Datenverarbeitungen keine Anwendung.

Eine im Jahr 2021 durchgeführte Interviewreihe mit 30 Smart-City-Akteuren in Deutschland (u.a. Grundversorger, Bildungseinrichtungen, Behörden und Gesundheitsdienstleister) ergab, dass durch den Austausch anonymisierter Daten zwischen einzelnen Smart-City-Akteuren ein erheblicher Mehrwert für die Bürger einer Smart City erreicht werden könnte: 76,67% der Interviewten gaben an, durch den Austausch anonymisierter Daten mit anderen Smart-City-Akteuren *theoretisch* Mehrwerte wie u.a. eine Verbesserung der Krankheitsbekämpfung/-eindämmung, bessere Reaktion auf wachsende Bedarfe an Betreuungs- und Bildungseinrichtungen sowie verbesserte Warmmöglichkeiten im

Katastrophenfall erreichen zu können. Jedoch würde das Potenzial des Austauschs anonymisierter Daten laut der Interviewten derzeit nicht ausgeschöpft werden, weil die Smart-City-Akteure oft nicht beurteilen könnten,

- wann Daten aus rechtlicher Sicht rechtssicher als anonymisiert zu betrachten seien und
- welche Daten eines Datensatzes z.B. gelöscht, verallgemeinert werden oder erhalten bleiben dürften, um im Ergebnis einen Datensatz mit ausschließlich (DSGVO-konform) anonymisierten Daten zu erhalten,

so dass die Regeln der DSGVO auf die weitere Datenverarbeitung nicht mehr anwendbar wären. In der Konsequenz verarbeiten 43,33% der Interviewten ausschließlich personenbezogene Daten und machen somit von den Vorteilen des Datenaustauschs in anonymisierter Form keinen Gebrauch, obwohl dieser einerseits große Mehrwerte für die Bürger mit sich bringen könnte, andererseits den Smart-City-Akteuren selbst aber auch Möglichkeiten des Angebots neuer Dienstleistungen und Produkte im Zusammenhang mit diesen Mehrwerten bieten würde [ST21a].

Die DSGVO hält die Anforderungen an anonymisierte Daten sehr vage: Lediglich in Erwgr. 26 findet sich die Information, dass es sich bei anonymisierten Daten um personenbezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann, handelt. Doch die beschriebenen Unsicherheiten zur rechtssicheren Feststellung der Anonymität, die sich nicht auf die 30 befragten Smart-City-Akteure beschränken, sondern mindestens auch seitens weiterer Smart-City-Akteure [ST21a] – vermutlich aber auch seitens einer Vielzahl weiterer datenverarbeitender Organisationen außerhalb des Smart-City-Kontextes – besteht, zeigt den dringenden Bedarf an Konkretisierung der Anonymität aus datenschutzrechtlicher Sicht auf. Um die Generierung von Mehrwerten für die Gesellschaft sowie Möglichkeiten, basierend auf diesen Mehrwerten neue Dienstleistungen und Produkte entwickeln zu können, nicht durch Unsicherheiten in Bezug auf die datenschutzkonform durchgeführte Anonymisierung zu verhindern, bedarf es somit einer Möglichkeit, den Anonymitätsgrad eines Datensatzes rechtssicher zu bestimmen.

Ein möglicher Weg, die Anonymität eines Datensatzes zu überprüfen, ist die Entwicklung von Metriken. Metriken sind quantitative Messwerte, mittels derer Organisationen systematisch die Einhaltung externer Anforderungen messen können. Mithilfe derartiger Metriken ließe sich grundsätzlich der Anonymitätsgrad eines konkreten, auf Basis der zuvor definierten Metriken analysierten Datensatzes bestimmen. Gleichwohl lässt sich dies nur umsetzen, wenn die aus rechtlicher Sicht an anonymisierte Daten bestehenden Anforderungen zuvor konkretisiert würden – denn nur was zuvor konkret als Anforderung bestimmt wurde, lässt sich mittels Metriken überprüfen [St22].

Insofern bedarf es auch in Bezug auf Anonymität und die datenschutzrechtlichen Anforderungen an diese eines Umdenkens im Datenschutzrecht. Denn nur wenn die bisher sehr vage formulierten Anforderungen an Anonymität konkretisiert werden, wäre auch eine Überprüfung des Anonymitätsgrades eines Datensatzes – und somit die Feststellung der (Un-)Anwendbarkeit der DSGVO und anderer datenschutzrechtlicher Rechtsakte – möglich und könnten wir als Gesellschaft von Vorteilen der Verarbeitung rechtssicher anonymisierter Daten profitieren.

4 Nutzerbedürfnisse im Datenschutzrecht

Die DSGVO – und andere datenschutzrechtliche Rechtsakte – sehen als eine mögliche Rechtsgrundlage personenbezogener Datenverarbeitungen die Einwilligung vor und konkretisieren die datenschutzkonforme Umsetzung des Einholens und Widerrufens dieser. Bei einer Einwilligung handelt es sich um jede durch die betroffene Person freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung, wobei die Einwilligung in Form einer Erklärung oder in Form einer sonstigen eindeutigen bestätigenden Handlung abgegeben werden (Art. 4 Nr. 11 DSGVO). Darüber hinaus umfasst das Europäische Datenschutzrecht die Pflicht des Verantwortlichen, betroffene Personen umfangreich über die sie betreffende personenbezogene Datenverarbeitung zu informieren. Gegenstand dieser Informationspflicht sind u.a. der Name und die Kontaktdaten des Verantwortlichen, die Datenverarbeitungszwecke und die einschlägige(n) Rechtsgrundlage(n) (Art. 13, 14 DSGVO).

Die Texte, die zum Einholen von Einwilligungen oder zur Umsetzung der Informationspflicht genutzt werden, sind i.d.R. lang, kompliziert geschrieben und langweilig gestaltet und werden aus diesen Gründen nur in den seltensten Fällen gelesen und Einwilligungserklärungen sogar häufig ungelesen unterschrieben [Do22]. Jedoch ist es angesichts des oben aufgezeigten Ziels des Datenschutzrechts sehr sinnvoll, die betroffenen Personen zum Lesen der Texte zu „animieren“, so dass diese z. B. im Rahmen des Lesens einer Einwilligungserklärung verstehen, dass die Erteilung der Einwilligung freiwillig und widerrufbar ist.

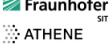
Darüber hinaus nimmt unter den betroffenen Personen der Unmut über die Vielzahl der von ihnen abverlangten Einwilligungen – z.B. im Zusammenhang mit Webtracking-Cookies und den sogenannten „Cookie-Bannern“, mit deren Hilfe die Einwilligung in das Webtracking umgesetzt werden soll – zu. Neben der Vielzahl der hierdurch täglich abverlangten Einwilligungen ist es u.a. auch problematisch, dass diese häufig – zu Lasten der betroffenen Personen – nicht datenschutz- bzw. nutzerfreundlich ausgestaltet sind [DHS21].

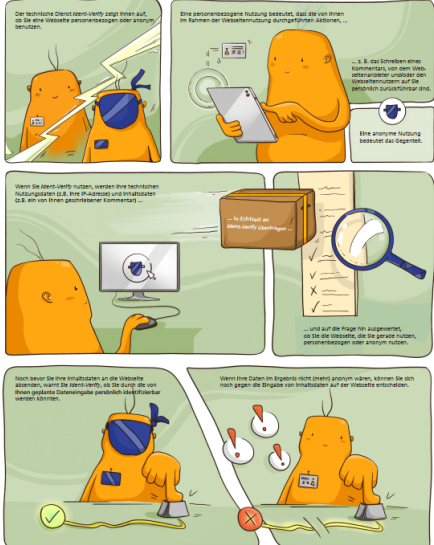
Vor diesem Hintergrund ist es dringend nötig, die (tatsächlichen) Bedürfnisse der betroffenen Personen in den Mittelpunkt der Ausgestaltung datenschutzrechtlicher Texte und Einwilligungsformen zu stellen und auch insofern ein Umdenken im Datenschutzrecht zu bewirken. Denn nur wenn die betroffenen Personen ihre datenschutzrechtlich

relevanten Entscheidungen informiert und ohne Verdruss treffen, lassen sich die Ziele des Datenschutzrechts, die betroffenen Personen vor Eingriffen in ihre Rechte und Freiheiten im Zusammenhang mit personenbezogener Datenverarbeitung zu schützen, langfristig erreichen. U.a. gilt es daher, die in Bezug auf Einwilligungserklärungen, Datenschutzinformationen und Einwilligungserteilung bestehenden Bedürfnisse betroffener Personen kennenzulernen und mit Hilfe empirischer Methoden zu erforschen, inwiefern bei betroffenen Personen Interesse und Neugier geweckt werden kann.

Wie nutze ich den technischen Dienst Ident-Verify? 

- 1. Ident-Verify im Überblick**
Der technische Dienst Ident-Verify zeigt Ihnen auf, ob Sie eine Webseite personenbezogen oder anonym benutzen.
- 2. Begriffsbestimmung**
Eine personenbezogene Nutzung bedeutet, dass die von Ihnen im Rahmen der Websitenutzung durchgeführten Aktionen, z.B. das Schreiben eines Kommentars, von dem Webseitenanbieter und/oder den Websitenutzern auf Sie persönlich zurückführbar sind. Eine anonyme Nutzung bedeutet das Gegenteil.
- 3. Personenbezogene oder anonyme Nutzung**
Wenn Sie Ident-Verify nutzen, werden Ihre technischen Nutzungsdaten (z.B. Ihre IP-Adresse) und Inhaltsdaten (z.B. ein von Ihnen geschriebener Kommentar) in Echtzeit an Ident-Verify übertragen und auf die Frage hin ausgewertet, ob Sie die Webseite, die Sie gerade nutzen, personenbezogen oder anonym nutzen.
- 4. Warnung**
Noch bevor Sie Ihre Inhaltsdaten an die Webseite absenden, warnt Sie Ident-Verify, ob Sie durch die von Ihnen geplante Dateneingabe persönlich identifizierbar werden könnten.
- 5. Entscheidung**
Wenn Ihre Daten im Ergebnis nicht (mehr) anonym wären, können Sie sich noch gegen die Eingabe von Inhaltsdaten auf der Webseite entscheiden.

Wie nutze ich den technischen Dienst Ident-Verify? 



Die technische Dienst Ident-Verify zeigt Ihnen auf, ob Sie eine Webseite personenbezogen oder anonym benutzen.

Eine personenbezogene Nutzung bedeutet, sagt die von Ihnen im Rahmen der Websitenutzung durchgeführten Aktionen, ... z.B. das Schreiben eines Kommentars, von dem Web-Anbieter und/oder den Websitenutzern auf Sie persönlich zurückführbar sind. Eine anonyme Nutzung bedeutet das Gegenteil.

Wenn Sie Ident-Verify nutzen, werden Ihre technischen Nutzungsdaten (z.B. Ihre IP-Adresse) und Inhaltsdaten (z.B. ein von Ihnen geschriebener Kommentar) ... in Echtzeit an Ident-Verify übertragen ...

... und auf die Frage hin ausgewertet, ob Sie die Webseite, die Sie gerade nutzen, personenbezogen oder anonym nutzen.

Noch bevor Sie Ihre Inhaltsdaten an die Webseite absenden, warnt Sie Ident-Verify, ob Sie durch die von Ihnen geplante Dateneingabe persönlich identifizierbar werden könnten.

Wenn Ihre Daten im Ergebnis nicht (mehr) anonym wären, können Sie sich noch gegen die Eingabe von Inhaltsdaten auf der Webseite entscheiden.

Abb. 1: Alternative Darstellungen datenschutzrechtlicher Texte in Form einer Informationsgrafik (links) und eines Comics (rechts)⁴

Abb. 1 zeigt mit einer Informationsgrafik und einem Comic zwei Alternativen zu einer rein textbasierten Datenschutzzinformation, die – abhängig vom Verarbeitungskontext und der Zielgruppe einer Datenschutzzinformation [DHS21] – gute Möglichkeiten darstellen können, um betroffenen Personen genau dies zu vermitteln. Derartige Ansätze gilt es in Zukunft weiterzuerfolgen und die mittels der Austestung der Ansätze gewonnenen Erkenntnisse in die Datenschutzpraxis zu überführen. Hierzu gehört auch die Berücksichtigung von Bedürfnissen an die Einwilligungshäufigkeit sowie an ein

⁴ Die auf den Abbildungen enthaltenen Informationen sind als Teilauszug einer Datenschutzzinformation zu verstehen und enthalten insofern nicht alle in Art. 13 DSGVO genannten Pflichtangaben.

(verbessertes) übergeordnetes Einwilligungsmanagement, z.B. in Bezug auf Cookie-banner.

5 Fazit

Das Datenschutzrecht und der damit verbundene Schutz betroffener Personen vor Eingriffen in ihre Rechte und Freiheiten im Rahmen personenbezogener Datenverarbeitungen ist wichtig und – nicht zuletzt in unserer von Smarten Geräten geprägten und immer vernetzter werdenden Welt – aktueller denn je. Doch der vorliegende Beitrag zeigt auf, dass das Datenschutzrecht vor einem Umbruch steht, in dem es gilt, Wege zu finden, technischen Fortschritt und die damit verbundenen gesellschaftlichen Vorteile zu ermöglichen, rechtssicher agieren zu können und die Verarbeitung personenbezogener Daten entlang der Bedürfnisse der betroffenen Personen zu gestalten. Diesen Herausforderungen gilt es in Zukunft zu begegnen, um das Datenschutzrecht in gewisser Weise in eine Ära zu überführen, in der Datenschutzrecht nicht mehr als lästiges Übel und Verhinderer wahrgenommen wird, sondern seinem Ziel dienlich und in dieser Zielerreichung sowohl von den betroffenen Personen als auch von den datenverarbeitenden Organisationen akzeptiert und unterstützt zu werden.

6 Danksagung

Der Beitrag gibt die persönliche Meinung der Autorin wieder. Dieser Beitrag wurden vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

7 Literatur

- [BSS23] Boll, A., Selzer, A., und Spiecker gen. Döhmman, I.: Datenschutz in der offensiven Cybersicherheitsforschung. Tagesspiegel-Online: <https://background.Tagesspiegel.de/cybersecurity/datenschutz-in-der-offensiven-cybersicherheitsforschung>, Stand: 25.8.2023.
- [DHS21] Diel, S., Heereman, W., und Spychalski, D.: The use of web tracking technologies and its compliance with the consent-requirement. GI Informatik 2021, S. 789-794, 2021.
- [Do22] Doan, X., Selzer, A., Rossi, A., Botes, M., and Lenzini, G: Context, Prioritization, and Unexpectedness: Factors Influencing – User Attitudes About Infographic and Comic Consent. Companion Proceedings of the Web Conference, 2022.

-
- [GS23] Gärtner, T., und Selzer, A.: Begriffsverwirrung verhindern. Tagesspiegel-Online: <https://background.tagesspiegel.de/cybersecurity/begriffsverwirrung-verhindern-was-massnahmen-aktiver-cyberabwehr-sind-und-was-nicht>, Stand: 25.8.2023.
- [KB20] Kühling, J. und Buchner, B.: Datenschutz-Grundverordnung / Bundesdatenschutzgesetz, Kommentar. 3. Auflage. München: C.H.Beck, 2020.
- [SHS19] Simitis, S., Hornung, G. und Spiecker gen. Döhmman, I.: Datenschutzrecht – DSGVO mit BDSG. 1. Auflage. München: C. H. Beck, 2019.
- [SS21] Selzer, A. und Spiecker gen. Döhmman, I.: Warum es einen Rechtsrahmen für die offensive Cybersicherheitsforschung braucht. Tagesspiegel-Online: <https://background.tagesspiegel.de/cybersecurity/warum-es-einen-rechtsrahmen-fuer-die-offensive-cybersicherheitsforschung-braucht>, Stand: 25.8.2023.
- [ST21a] Selzer, A. und Timm, I. J.: Chances and Limitations of Personal and Anonymized Data Processing – Implementing Appropriate Technical and Organizational Measures and Creating Added Value in Smart Cities. GI Informatik 2021, S. 773-788, 2021.
- [SW22] Shulman, H. und Waidner, M.: Aktive Cyberabwehr. ATHENE-Whitepaper: <https://www.athene-center.de/fileadmin/Downloads/aktive-cyberabwehr.pdf>, Stand: 25.8.2023.
- [St22] Stummer, S.: Issues of Verifying Anonymity: An Overview. GI Informatik 2021, S. 179-194, 2022.

Verarbeitung personenbezogener Daten in KI-Reallaboren nach dem KI-VO-E als Herausforderung für den datenschutzrechtlichen Zweckbindungsgrundsatz

Jan-Philipp Stroscher ¹ und Hendrik Link ²


Abstract: Reallabore wurden bisher insbesondere in der Finanzbranche eingesetzt, um unter Befreiung von Regulatorik neue Technologien auszuprobieren. Mit der KI-VO will die EU zur Innovationsförderung KI-Reallabore einführen. Art. 54 KI-VO-E enthält eine Ausnahmeregelung zum elementaren Zweckbindungsgrundsatz der DSGVO, um das Trainieren von KI mit personenbezogenen Daten zu erleichtern. Gleichzeitig wurde die Vorschrift mit der Intention geschaffen, Art. 6 Abs. 4 DSGVO gerecht zu werden, der Vorgaben für Gesetze, die eine Zweckänderung zulassen wollen, enthält. Aufgrund seines privilegierten Anwendungsbereichs und den umfangreichen Schutzanforderungen entspricht Art. 54 KI-VO-E zwar den Anforderungen aus Art. 6 Abs. 4 i. V. m. Art. 23 DSGVO trägt aber nur sehr eingeschränkt zur Innovation in diesem Bereich bei.

Keywords: KI-VO, KI-Reallabor, Regulatory Sandboxes, Zweckbindungsgrundsatz, DSGVO

1 Einleitung

Künstliche Intelligenz (KI) ist eines der kontrovers diskutierten Themen der Gegenwart und jüngeren Vergangenheit. Auf der einen Seite wird KI enormes Potenzial in vielen Bereichen nachgesagt (u. a. autonomes Fahren, Nutzung zur medizinischen Diagnostik und Therapie) und auf der anderen Seite wird vor den Gefahren dieser neuen Technologie gewarnt (u. a. Diskriminierung, Bedrohung der Meinungsvielfalt). In diversen Rechtsgebieten wird intensiv diskutiert, inwiefern KI von den bestehenden Gesetzen bereits reguliert wird und ob und welche Gefahren von KI ausgehen, die von neuen Gesetzen abgedeckt werden müssen. Diese Frage wird durch die an Tempo gewinnende Entwicklung von KI weiter verstärkt. Gesetzgeber stehen vor der Herausforderung, mit dem rapiden Fortschritt der technischen Entwicklung schrittzuhalten. Auf Unionsebene wird bei der Regulierung technischer Innovationen zunehmend ein agiler Ansatz favorisiert. Diese Agilität sollen Regulatory Sandboxes (deutsch: Reallabore) als neues Regulierungselement bringen, indem Potenziale nutzbar gemacht werden, ohne die Gefahren unreguliert zu belassen. So hat die Kommission am 21.04.2021 einen ersten Entwurf zur KI-Verordnung (KI-VO-KomE) [[Eu21]] vorgelegt, indem KI-Reallabore in

¹ Universität Kassel, Fachgebiet Öffentliches Recht, IT-Recht und Umweltrecht, Henschelstraße 4, 34127 Kassel, jan-philipp.stroscher@uni-kassel.de  <https://orcid.org/0000-0002-8208-4409> .

² Universität Kassel, Fachgebiet Öffentliches Recht, IT-Recht und Umweltrecht, Henschelstraße 4, 34127 Kassel, hendrik.link@uni-kassel.de  <https://orcid.org/0009-0006-0596-2120> .

Titel V als zentrale Maßnahme zur Innovationsförderung eingeführt werden. Ergänzend hierzu wird die Möglichkeit der Weiterverarbeitung von personenbezogenen Daten zu bestimmten Zwecken in KI-Reallaboren (Art. 54 KI-VO-KomE) vorgesehen. Ziel von Reallaboren im Allgemeinen ist es, Technologien in einer geschützten Anwendung zu entwickeln oder zu testen, wofür häufig Ausnahmen von bestehenden gesetzlichen Vorgaben gemacht werden. Mit Art. 54 KI-VO-KomE hat die Kommission eine gewichtige Problematik erkannt und eine solche Ausnahme statuiert: Wollen Entwickler ihre KI-Anwendungen mit personenbezogenen Daten trainieren, steht dies im Konflikt mit dem Zweckbindungsgrundsatz des Datenschutzrechts, da die gesammelten Daten im Regelfall zu anderen Zwecken erhoben wurden. Dies wirft einige Fragen auf, denen sich dieser Beitrag widmen möchte. Zunächst ist allgemein zu klären, wie der KI-VO-KomE zur DSGVO steht. Sodann ist konkret zu untersuchen, inwiefern sich der Art. 54 DSGVO in die Systematik des Art. 6 Abs. 4, Art. 23 DSGVO einfügt und ob Reallabore so einen Mehrwert für die Entwickler von KI bieten können. Neben dem Kommissionsentwurf werden ebenso die Fassungen des Rats (KI-VO-RatE) [[Ra22]] und die des Europäischen Parlaments (KI-VO-ParlE) [[Eu23b]] veröffentlichten Fassung berücksichtigt.

2 Das Regulierungsinstrument: Reallabore

Reallabore beschreiben zeitlich und räumlich begrenzte Testräume, in denen innovative Technologien und Geschäftsmodelle unter möglichst realen Bedingungen getestet werden können [[Le22], S. 433], wobei es keine allgemein anerkannte Definition gibt [[Bu19], S. 9]. Der Rat der Europäischen Union versteht unter den Reallaboren „konkrete Rahmen, die, indem sie einen strukturierten Kontext für Experimente vorgeben, es ermöglichen, innovative Technologien, Produkte, Dienstleistungen oder Ansätze – aktuell insbesondere im Zusammenhang mit der Digitalisierung – wo geeignet in einer realen Umgebung für einen begrenzten Zeitraum oder in einem begrenzten Teil einer Branche oder eines Gebiets unter regulatorischer Aufsicht und Gewährleistung angemessener Schutzmaßnahmen zu erproben“ [[Ra20], S. 4]. Ein Reallabor kann dabei physischer sowie digitaler Natur sein [[No22], S. 20].

Reallabore bieten einen pragmatischen Ansatz, um Innovationen rechtlich zu begleiten. Der Mehrwert für die Entwickler liegt darin, dass Ausnahmen von gesetzlichen Vorgaben eingeräumt oder eine individuelle Anleitung zur Erfüllung solcher gesetzlichen Vorgaben – vorwiegend durch Aufsichtsbehörden – gegeben wird [[Ra21a], S. 1]. Sie haben eine Doppelrolle, indem einerseits Entwickler Innovationen in einem möglichst realistischen Umfeld testen können („business learning“), andererseits Aufsichtsbehörden und Gesetzgeber Erfahrungswerte zur Regulierung technischer Neuerungen gewinnen können („regulatory learning“) [[Mv22], S. 2]. Reallabore stellen dabei kein Novum dar, da dieser Mechanismus bereits in diversen Sektoren eingesetzt wird, etwa in den Sektoren Gesundheit, Rechtsdienstleistungen, Luftfahrt, Verkehr und Logistik, Energie und vor allem im Finanzsektor [[Ra20], S. 3]. Der stetige technische Fortschritt sorgt in vielen Bereichen dafür, dass das technische Können das regulatorische Dürfen übersteigt

[[Le22], 433]. Sowohl der europäische Gesetzgeber als auch das BMWK auf nationaler Ebene setzen sich daher für eine verstärkte Schaffung dieser sog. Experimentierklauseln, z.B. in Form von Reallaboren, ein [[Ra20], S. 5, [Bu19], S. 15].

2.1 Chancen

Viele Vorteile der Reallabore resultieren aus dem unmittelbaren Zusammenwirken verschiedener interdisziplinärer Akteure und der damit einhergehenden vereinfachten Kommunikation zwischen Entwicklern, Experten, Konsumenten und Behörden [[Mv22], S. 2] sowie den dadurch entstehenden Synergieeffekten, insbesondere für Startups und Kleinanbieter, wenn diese ohne entsprechende Abteilungen eine Entwicklung von KI anstreben [[No22], S. 17]. Häufig stehen gesetzliche Vorgaben der Entwicklung neuer Technologie im Wege, was durch die Schaffung von Reallaboren überwunden werden könnte [[Bu19], S. 15]. Neben Rechts- und Planungssicherheit erhalten Entwickler auch durch den kontinuierlichen Austausch mit Aufsichtsbehörden Zugang zu regulatorischer Expertise [[Kr21], S. 438, [Ra21b], S. 96, [No22], S. 22]. Fragen zur Gesetzeskonformität von Innovationen können daher frühzeitig erkannt und bereits in der Entwicklungsphase der Technologien adressiert werden, wodurch ein schnellerer Markteintritt ermöglicht wird [[Mv22], S. 2]. Da KI-Anwendungen oft mit personenbezogenen Trainingsdaten arbeiten müssen, können die sich daraus ergebenden datenschutzrechtlichen Fragestellungen bereits im Entwicklungsstadium in KI-Reallaboren berücksichtigt werden, wodurch gleichermaßen die Anforderungen aus Art. 25 Abs. 1 DSGVO, Datenschutz durch Technikgestaltung, deutlich effizienter umgesetzt werden können [[PA21], S. 3]. Bezüglich KI betrifft dies die neuen Anforderungen aus der KI-VO-KomE selbst [[Ge21], S. 354], sowie den verbundenen Fragen nach der Haftung, die sich aus der in Überarbeitung befindenden Produkthaftungsrichtlinie und der geplanten Richtlinie zur KI-Haftung ergeben werden [[RL23], [Li22]].

Technische Neuerungen bringen im Regelfall auch Schwierigkeiten für die Regulatorik durch den Gesetzgeber sowie die Aufsichtsbehörden mit sich. Es ist häufig unklar, ob alte Regelungsmuster auf neue Technologien übertragen werden können bzw. sollten. Aufsichtsbehörden können in mehrfacher Hinsicht aus Reallaboren lernen. Einerseits können sie ein technisches Verständnis der jeweiligen Produkte entwickeln. So werden Probleme, Möglichkeiten und Grenzen von Innovationen nachvollzogen und in die Regulierungspraxis einbezogen, wodurch eine evidenzbasierte Regulatorik ermöglicht wird [[PA21], S. 2]. Ferner kann die Sinnhaftigkeit der Anwendung bestehender Gesetze auf neue Technologie kritisch hinterfragt werden, wenn regulatorische Vorgaben im geschützten Umfeld von KI-Reallaboren temporär außer Kraft gesetzt werden (Experimenting by derogation) [[PA21], S. 1]. Andererseits besteht Raum für die öffentlich-rechtliche Seite, neue Formen der Regulierung auszuprobieren (Experimenting by devolution) [[Bu19], S. 17, [PA21], S. 1], die durch eine starke Umwälzung von Wirtschaft und Gesellschaft durch technische Innovationen notwendig werden kann.

Durch das Zusammenspiel der verschiedenen Akteure und den damit einhergehenden Lerneffekten aufseiten der Entwickler, der Aufsichtsbehörden sowie des Gesetzgebers können Konsumenten von der Einführung neuer, potenziell sicherer und ausgewogenerer sowie evidenzbasiert regulierter Technologien profitieren [[PA21], S. 2].

2.2 Risiken

Neben den vielen Vorteilen von Reallaboren bestehen jedoch auch verschiedene Risiken, die bei der Regulierung zu berücksichtigen sind. Generell wird Reallaboren vorgeworfen, dass Innovationen über die Schaffung angemessener Schutzmaßnahmen gestellt werden [[Mv22], S. 3]. Im Einzelnen wird moniert, dass Reallabore die Gefahr einer willkürlichen Ausübung der Aufsicht durch Behörden („Arbitrage“) sowie eines gegenseitigen Unterbietens rechtlicher Anforderungen („race to bottom“ Wettbewerb) mit sich bringen würden [[Ro22], S. 373]. Auch können Reallabore den Wettbewerb verzerren, indem Verbraucher Unternehmen bevorzugen, die zum Teil in einem Reallabor tätig waren [[Gr18]]. In die entgegengesetzte Richtung denkend, könnten Reallabore durch strikte Vorgaben der Aufsichtsbehörden die Innovationskraft des Privatsektors vermindern [[Mv22], S. 3].

Neben den generellen Kritikpunkten zu KI-Reallaboren wurden spezielle Risiken in den Entwürfen der KI-VO identifiziert. Im Fokus stand dabei der weite Gestaltungsspielraum der Mitgliedstaaten bzw. deren Aufsichtsbehörden bei der Ausgestaltung der KI-Reallabore. Divergierende Vorgaben können zu einem „forum shopping“ zwischen den KI-Reallaboren der Mitgliedstaaten führen, wobei die geringster Regulatorik zur höchsten Popularität führen dürfte [[Mv22], S. 5]. Die stetige Begleitung von Unternehmen in KI-Reallaboren könnte zu einer Überlastung der Aufsichtsbehörden führen, wenn solche mit dünner Personaldecke zusätzlich noch KI-Reallabore begleiten müssten. Kleinere Mitgliedstaaten haben zudem die Befürchtung geäußert, nationale Aufsichtsbehörden nicht mit entsprechendem Know-How besetzen zu können. Weiter könnte durch den KI-VO-KomE Rechtsunsicherheit ausgelöst werden, wenn das Verhältnis zu weiteren Rechtsakten nicht vom Unionsgesetzgeber geregelt wird [[Ra21a], S. 7]. Hinsichtlich des Datenschutzrechts fordert das European Data Protection Board (EDPB) mögliche Widersprüche zwischen den Regeln für KI-Reallabore nach dem KI-VO-KomE und der DSGVO zu verhindern [[EE21], S. 17].

3 KI-Reallabore im System des KI-VO-KomE

Mit dem KI-VO-KomE hat die Kommission den ersten umfassenden Rechtsrahmen im Bereich der Künstlichen Intelligenz konzipiert [[RW21], S. 330]. Dem Kommissionsentwurf liegt ein dualistischer Ansatz zugrunde [[RW21], S. 336]. Einerseits sollen Verbraucher und deren Grundrechte geschützt werden [[Ge21], S. 356], andererseits sollen Investitionen in KI sowie innovative KI gefördert werden [[RW21],

S. 336]. Als Grundlage des Verbraucherschutzes differenziert der KI-VO-KomE zwischen verschiedenen Risikoklassen von KI [[Mv22], S. 5]. In Art. 5 KI-VO-KomE werden KI-Anwendungen genannt, die aufgrund eines inakzeptablen Risikos für die Allgemeinheit verboten sind. In Art. 6-29 KI-VO-KomE werden Hochrisikosysteme stark reguliert, während Anwendungen mit begrenztem Risiko nur die Transparenzanforderung aus Art. 52 KI-VO-KomE erfüllen müssen. KI-Anwendungen mit niedrigem oder minimalem Risiko treffen keine Vorgaben aus der KI-VO-KomE. Obgleich die Entwurfsfassungen des Rates und des Parlaments gewichtige Abweichungen enthalten, etwa bei der zentralen Definition von KI, wird von dem Grundkonzept nicht abgewichen. Auf Titel I - IV, die Vorgaben hinsichtlich der verschiedenen Risikostufen von KI-Anwendungen geben, folgt in Titel V sodann die Innovationsförderung. Mit Regelungen in Art. 53, 54 KI-VO-KomE bilden die KI-Reallabore das Hauptwerkzeug der Innovationsförderung.

3.1 Regelungsbestand der Art. 53, 54 KI-VO

Art. 53 KI-VO-KomE regelt die Errichtung von KI-Reallaboren. Der Parlamentsentwurf ist dabei deutlich umfangreicher als der Kommissionsentwurf. Der Ratsentwurf hat, neben kleineren Anpassungen in Art. 53 KI-VO-RatE, mit Art. 54a, 54b KI-VO-RatE Regelungen zur Ermöglichung unbeaufsichtigter Tests in der realen Welt eingefügt, um die Reallabore innovationsfreundlicher zu gestalten [[Ra22], S. 7]. Eine andere relevante Frage, bei der Uneinigkeit herrscht, ist, wer KI-Reallabore einrichten soll. Sowohl nach dem Rats- als auch nach dem Kommissionsentwurf soll diese Aufgabe den Aufsichtsbehörden i. S. d. Art. 59 KI-VO-KomE zustehen, während das Parlament diese Aufgabe den Mitgliedstaaten überträgt. Darüber hinaus muss jeder Mitgliedstaat zumindest ein KI-Reallabor bereitstellen. Einigkeit besteht wiederum bei dem zentralen Aspekt der Ausnahmen von Rechtsvorschriften. Hier bleiben in allen drei Fassungen die Entwickler nach den Vorschriften des Europarechts sowie dem Recht der Mitgliedstaaten haftbar. Parlament und Rat wollen allerdings, dass keine Bußgelder aufgrund von Verstößen gegen die KI-VO verhängt werden, wenn nach Treu und Glauben den Anweisungen der Aufsichtsbehörden gefolgt wird. Auch hier werden jedoch keine Ausnahmeregelungen und Experimentierklauseln aufgenommen. Einzig Art. 54 KI-VO-KomE kann als Ausnahmeregelung verstanden werden. Die Fassungen von Rat und Parlament gleichen dem Kommissionsentwurf in diesen Punkten stark. In Art. 54 Abs. 1 KI-VO-KomE wird der Zweckbindungsgrundsatz bezüglich der Entwicklung und Erprobung innovativer Systeme mit personenbezogenen Daten unter den Bedingungen des Abs. 1 lit. a - j KI-VO-KomE aufgehoben. In Art. 54 Abs. 1 lit. a KI-VO-KomE wird geregelt, dass nur solche KI-Anwendungen von der Ausnahme der Zweckbindung profitieren, die der öffentlichen Sicherheit, öffentlicher Gesundheit oder dem Umweltschutz dienen sollen. Die Kommission möchte diese Ausnahme auch auf die Gefahrenabwehr und Strafverfolgung erweitern, jedoch haben sowohl der Rat als auch das Parlament diesen Aspekt gestrichen. Somit bleibt festzuhalten, dass der Anwendungsbereich der einzigen Ausnahmeregelung, der Befreiung vom Zweckbindungsgrundsatz aus Art. 54 KI-VO-KomE, bereits stark eingeschränkt ist.

3.2 Das Verhältnis der Art. 53, 54 KI-VO zur DSGVO

Der Wortlaut der KI-VO-KomE sowie die Fassungen des Rats und Parlaments regeln das Verhältnis zur DSGVO nicht abschließend. Die Kritik des EDPB an dem unklaren Verhältnis sowie der einhergehenden Rechtsunsicherheit wurde hier – unverständlicherweise – nicht berücksichtigt [[EE21], S. 17]. Genereller Konsens der an der Gesetzgebung beteiligten Institutionen, die Vorschriften der DSGVO nicht beeinträchtigen zu wollen, kommt jedoch an verschiedenen Stellen zum Ausdruck. So spricht die Kommission in der Gesetzesbegründung explizit davon, die Vorschriften der DSGVO sollen „unberührt“ bleiben [[Eu21], S. 4]. EG 72 KI-VO-KomE besagt, dass Art. 54 KI-VO-KomE im Einklang mit Art. 6 Abs. 4 DSGVO stehen sollte, was gleichermaßen in EG 72a KI-VO-ParlE sowie EG -72a KI-VO-RatE zum Ausdruck kommt.

4 Art. 54 KI-VO-KomE und der datenschutzrechtliche Zweckbindungsgrundsatz

Art. 54 KI-VO-KomE erlaubt den Teilnehmern an den KI-Reallaboren eine Weiterverarbeitung von personenbezogenen Daten zu einem bestimmten geänderten Zweck und stellt demnach eine Zweckänderungsregelung i. S. v. Art. 6 Abs. 4 DSGVO dar [[ES21], S. 1192], was wiederum eine Ausnahme vom Zweckbindungsgrundsatz i. S. v. Art. 5 Abs. 1 lit. b Hs. 1 DSGVO bedeutet.

4.1 Der Zweckbindungsgrundsatz gem. Art. 5 Abs. 1 lit. b Hs. 1 DSGVO sowie die Ausnahmen nach Art. 6 IV DSGVO

Art. 54 KI-VO-KomE ist erforderlich, da der Zweckbindungsgrundsatz gem. Art. 5 Abs. 1 lit. b Hs. 1 DSGVO, bei dem es sich um ein das Datenschutzrecht prägendes und zugleich zentrales Grundprinzip handelt, festlegt, dass personenbezogene Daten nur für festgelegte, eindeutige und legitime Zwecke erhoben werden dürfen und nicht in einer mit diesen Zwecken nicht zu vereinbarenden Weise weiterverarbeitet werden dürfen. Das Spannungsverhältnis zwischen Zweckbindung und Weiterverarbeitung zu anderen Zwecken wird bereits seit vielen Jahren im Kontext von Big Data Anwendungen diskutiert (Datennutzung vs. Datenschutz) [[KB14], [RN15], [Ri16], [CD17]]. In diesen Fällen argumentieren Digitalkonzerne, dass Big Data Anwendungen zu Innovationen führen würden, die durch eine strikte Auslegung des Zweckbindungsgrundsatzes ausgebremst werden. Eine ähnliche Diskussion wird im Rahmen der Regulierung von KI geführt [[Br23], [AB22]].

Eine strenge Auslegung des bzw. Bindung an den Zweckbindungsgrundsatz stellt hingegen keinen Selbstzweck dar, sondern liegt in der verfassungsrechtlichen Grundlage dieses Prinzips begründet. Der Grundsatz der Zweckbindung ist eine Ausprägung des Rechts auf informationelle Selbstbestimmung gemäß Art. 2 Abs. 1, 1 Abs. 1 GG, welches

vor fast 40 Jahren vom BVerfG im wegweisenden Volkszählungsurteil entwickelt wurde und vor allem auch in Art. 8 Abs. 2 S. 1 GRCh explizit festgeschrieben ist. Jedoch stellte bereits das BVerfG ebenfalls fest, dass dieses Grundrecht und damit auch der Grundsatz der Zweckbindung nicht absolut gilt und die betroffene Person keinen Anspruch auf die Verarbeitungsfreiheit hat. Der Verarbeitung sind jedoch Grenzen zu setzen, die dem Einzelnen die Verteidigung seiner Subjektivität erlaubt [[Si84], S. 399]. Die Verarbeitungsprozesse müssen stets für die betroffene Person „durchschaubar und nachvollziehbar sein“ [[Si84], S. 400].

Auf Grund des Umstandes, dass die Verarbeitungsfreiheit nicht absolut gilt, sieht die DSGVO in Art. 6 Abs. 4 DSGVO die Möglichkeit vor, dass Verarbeitungen rechtmäßig sein können, auch wenn diese nicht dem ursprünglichen Erhebungszweck entsprechen.³ Art. 6 Abs. 4 DSGVO sieht grundsätzlich zwei Konstellationen vor: Die weitere Verarbeitung kann gem. Art. 6 Abs. 4, 1. Hs. DSGVO auf die Einwilligung des Betroffenen oder auf einer anderen (europäischen oder mitgliedstaatlichen [[KM16], S. 451]) Rechtsvorschrift gestützt werden. Wenn beides nicht einschlägig ist, kann gem. Art. 6 Abs. 4, 2. Hs. DSGVO noch die Vereinbarkeit des ursprünglichen mit dem weiteren Verarbeitungszweck geprüft werden (Kompatibilitätsprüfung). Bei dieser Prüfung sind vom Verantwortlichen die in dem Katalog des Abs. 4 lit. a bis e genannten Kriterien zu „berücksichtigen“ (vgl. EG 50).

Eine Einwilligung wird bei der Entwicklung von KI, innerhalb wie außerhalb von Reallaboren, kaum relevant sein, da die Einholung der Einwilligung zeit- und kostenintensiv ist. Darüber hinaus gefährdet der mögliche Widerruf inklusive der Löschung die Kontinuität der Datengrundlage; dies ist aber gerade erforderlich, um KI-Anwendungen zu trainieren, da es – soweit ersichtlich – nicht möglich ist, dass KI-Systeme die Erkenntnisse der Auswertung einzelner personenbezogener Daten vergessen. In der Folge könnten KI-Systeme mangels Einhaltung der DSGVO in Gänze nicht einsetzbar sein.

Die von der DSGVO selbst in Art. 5 Abs. 1 lit. b, 2. Hs., Art. 89 DSGVO vorgesehene Möglichkeit, zu im öffentlichen Interesse liegenden Archivzwecken, zu wissenschaftlichen oder historischen Forschungszwecken und zu statistischen Zwecken personenbezogene Daten rechtmäßig weiterverarbeiten zu dürfen, hat insbesondere bezogen auf die wirtschaftlichen Akteure in diesem Kontext keine Relevanz, da diese Ausnahme dann nicht greift, wenn rein kommerzielle Interessen verfolgt werden, wovon bei Entwicklungen in diesem Kontext auszugehen ist. Mit §§ 23, 24, 27 BDSG bestehen mitgliedstaatliche Rechtsgrundlagen, die für bestimmte Zwecke eine Weiterverarbeitung ermöglichen. Für § 27 BDSG, der auf Art. 9 Abs. 2 lit. j DSGVO beruht, gilt das zu Art. 89 DSGVO gesagt gleichermaßen. Daneben sehen §§ 23, 24 BDSG weitere mitgliedstaatliche Rechtsgrundlage i. S. v. Art. 6 Abs. 4 DSGVO für eine zweckfremde

³ In Art. 5 Abs. 1 lit. b, 2. Hs. DSGVO wird zudem eine Weiterverarbeitung für im öffentlichen Interesse liegende Archivzwecke, für wissenschaftliche oder historische Forschungszwecke oder für statistische Zwecke (gemäß Art. 89 Abs. 1 DSGVO) als mit dem ursprünglichen Zweck vereinbar angesehen und damit zulässig angesehen.

Weiterverarbeitung vor, wobei lediglich § 24 BDSG grundsätzlich Frage kommt, da § 23 BDSG die Weiterverarbeitung durch öffentlichen Stellen betrifft. Aber auch der sehr begrenzte § 23 BDSG ist vorliegend nicht einschlägig, da dieser in § 23 Abs. 1 BDSG Gefahrenabwehr und Strafverfolgung adressiert und in § 23 Abs. 2 BDSG die Weiterverarbeitung zur Geltendmachung zivilrechtlicher Ansprüche regelt.

Ein Rückgriff auf die Prüfung der Zweckvereinbarkeit wäre nach gegenwärtigen Rechtsstand grundsätzlich möglich, wobei die Anwendung dieser Kriterien auf Grund der eingeschränkten Bestimmtheit eine nicht unerhebliche Rechtsunsicherheit mit sich bringt. Das führt neben dem zeitlichen, personellen und damit auch finanziellen Aufwand für eine solche Kompatibilitätsprüfung vor allem dazu, dass die Entwickler die weitere Verarbeitung unter dem Vorbehalt der potenziellen Rechtswidrigkeit fortführen würden. In Anbetracht der drohenden Konsequenzen – vor allem eines potenziellen Bußgeldes – werden Entwickler von der Nutzung abhalten, was das Ziel der Innovationsförderung konterkarieren würde. Die Entbindung von der Kompatibilitätsprüfung dürfte daher einen bedeutenden Mehrwert für Entwickler darstellen.

Die vorangegangenen Ausführungen zeigen auf, dass eine gesonderte Regelung zum Training von KI-Systemen geschaffen werden muss und die Auslegung von Art. 6 Abs. 4 DSGVO demnach maßgeblich darüber entscheidet, wie weit „die Schleuse“ zur Weiterverarbeitung personenbezogener Daten geöffnet wird. Davon hängt ferner ab, welche Geschäftsmodelle und potenziellen technischen Entwicklungen zur Umsetzung gelangen können [[Ro19], Rn. 1], und konkret, ob KI-Reallabore in der Ausprägung des Art. 54 KI-VO-KomE ein wirksames Werkzeug darstellen kann.

4.2 Vereinbarkeit von Art. 54 KI-VO-KomE mit Art. 6 Abs. 4, 1. Hs. DSGVO

Art. 54 KI-VO-KomE müsste als unionsrechtliche Rechtsgrundlage i. S. v. Art. 6 Abs. 4, 1. Hs. DSGVO zudem den Wesensgehalt – schon aufgrund des Art. 52 Abs. 1 S. 2 GRCh – der Grundrechte und Grundfreiheiten achten [[Ro19], S. 26] und die in einer demokratischen Gesellschaft notwendigen und verhältnismäßigen Maßnahme zum Schutz der in Art. 23 Abs. 1 DSGVO genannten Ziele erfüllen.

Maßnahme zum Schutz der in Artikel 23 Absatz 1 genannten Ziele

Die Ziele des Art. 54 KI-VO-KomE müssen zunächst mit denen des Art. 23 DSGVO übereinstimmen. Art. 54 Abs. 1 KI-VO-KomE verfolgt das Ziel der Innovationsförderung (siehe 3.) für verschiedene Bereiche, die vom Unionsgesetzgeber als gesellschaftlich besonders relevant angesehen wurde. Die in Art. 54 Abs. 1 lit. a, i) KI-VO-KomE enthaltenen Ziele der „Verhütung, Ermittlung, Aufdeckung oder Verfolgung von Straftaten oder Strafvollstreckung, einschließlich des Schutzes vor und der Abwehr von Gefahren für die öffentliche Sicherheit unter der Kontrolle und Verantwortung der zuständigen Behörden, wobei die Verarbeitung auf der Grundlage des Rechts der

Mitgliedstaaten oder des Unionsrechts erfolgt“, findet sich im sowohl im Rats- als auch im Parlamentsentwurf nicht wieder, wenngleich Art. 23 Abs. 1 lit. d DSGVO dies als relevantes öffentliches Ziel darstellen würde. Nach Art. 54 Abs. 1 lit. a, ii) KI-VO-KomE dürfen ebenso personenbezogene Daten zum Zweck der „öffentliche Sicherheit und öffentliche Gesundheit, einschließlich Verhütung, Bekämpfung und Behandlung von Krankheiten“ weiterverarbeitet werden, was ebenfalls in Art. 23 Abs. 1 lit. c bzw. lit. e DSGVO erfasst ist. Dementsprechend ist die Erweiterung der Parlamentsfassung, die explizit auch die Erkennung, Diagnose und Kontrolle von Krankheiten benennt, nicht zu beanstanden. Im Parlamentsentwurf wurde neben einem hohen Umweltschutzniveau und der Verbesserung der Umweltqualität die „protection of biodiversity, pollution as well as climate change mitigation and adaptation“⁴ ergänzt. Diese Ziele lassen sich ebenfalls unter den Begriff des Umweltschutzes subsumieren und sind daher als sonstiges wichtiges Ziel von Art. 23 Abs. 1 lit. e DSGVO erfasst. Ferner schlägt der Parlamentsentwurf vor, als Ziele die „safety and resilience of transport systems, critical infrastructure and networks“⁵ zu ergänzen. Die Einfügung dieser Schutzziele dürfte auf das Erkennen der besonderen Vulnerabilität dieser Bereiche zurückzuführen sein, was insbesondere durch verschiedene, zuletzt vielfach auftretende Cyberangriffe offengelegt wurde. Art. 23 Abs. 1 lit. e DSGVO stellt einen sehr breiten Tatbestand dar, der jedes wichtige Gemeinwohlziel erfasst, das vom Gewicht her denen der übrigen Ziele gleichkommt. Die verschiedenen Cyberangriffe [[GMZ23]], wie auch physische Angriffe (z.B. das Durchtrennen der Leitungskabel der Deutschen Bahn [[Ta23]]) haben offengelegt, wie anfällig die entsprechenden Systeme für Angriffe sind, sodass dessen Schutzbedürftigkeit und in Folge die Einordnung als wichtiges Gemeinwohlziel zu bejahen ist. Gleiches dürfte für die vom Rat vorgeschlagene Ergänzung der Bereiche „energy sustainability, efficiency and quality of public administration and public services und „cybersecurity“ gelten. Insbesondere Covid- 19 hat gezeigt, welche Konsequenzen sich aus einer eingeschränkten Leistungsfähigkeit der öffentlichen Verwaltung und öffentlicher Dienstleistungen ergeben können.

Beachtung des Wesensgehalts des Art. 8 GRCh

Art. 54 KI-VO-KomE muss zudem den Wesensgehalt der Grundrechte und Grundfreiheiten achten, was Art. 23 Abs. 1 DSGVO als zusätzliche Voraussetzung festlegt, aber ohnehin von Art. 52 Abs. 1 S. 2 GRCh vorausgesetzt wird [[Ro19], Rn. 26, [AV23], Rn. 110]. Die Beachtung der Wesensgehaltsgarantie erstreckt sich ferner nicht alleine auf den Schutz personenbezogener Daten aus Art. 8 GRCh, sondern erfasst auch alle anderen grundrechtlich geschützten Positionen, für die sich mittelbar Einschränkungen ergeben können [[Di19], Rn. 13], wobei die Bedeutung des Wesensgehalts abstrakt schwierig zu bestimmen ist. Auch wenn die Wesensgehaltsgarantie nicht auf Art. 8 GRCh beschränkt ist, so ergibt sich aus Art. 54

⁴ Im Ratsentwurf wird die ergänzend in vergleichbarer Weise wie folgt ergänzt: „protection and improvement of the quality of the environment, including green transition, climate change mitigation and adaptation.“

⁵ Im Ratsentwurf findet sich eine vergleichbare Ergänzung mit folgenden Worten: „energy sustainability, transport and mobility“.

Abs. 2 KI-VO-KomE doch eine deutliche Fokussierung auf den Zweckbindungsgrundsatz aus Art. 8 Abs. 1 S. 2 GRCh. Indem Art. 54 KI-VO-KomE jedoch auf in rechtmäßiger Weise erhobene Daten zugreift und für einen begrenzten Bereich zu eine Weiterverarbeitung im erheblichen öffentlichen Interesse ermächtigt, wird den Entwicklern insbesondere kein generelles Zugriffsrecht auf elektronische Kommunikation gegeben und die betroffenen Personen werden auch nicht der Rechtsschutzmöglichkeiten beraubt [[Di19], Rn. 14]. Zudem werden in Art. 54 Abs. 1 lit. b-j KI-VO-KomE Anforderungen vorgesehen, die dem Schutz der betroffenen Personen dienen. Dadurch entsteht zumindest keine Zweckerweiterungsmöglichkeit, die den Wesensgehalt des Art. 8 Abs. 1 GRCh verletzt. Damit wahrt Art. 54 KI-VO-KomE den Wesensgehalt der Grundrechte und Grundfreiheiten.

Prüfung der Verhältnismäßigkeit

Darüber hinaus müsste Art. 54 KI-VO-KomE eine i. S. v. Art. 6 Abs. 4, 1. Hs. DSGVO notwendige und verhältnismäßige Maßnahme darstellen. Es muss demnach eine Verhältnismäßigkeitsprüfung durchgeführt werden, die nach der Rechtsprechung des EuGH voraussetzt, dass die Einschränkung im Hinblick auf das verfolgte Ziel geeignet, erforderlich und nicht außer Verhältnis zum verfolgte Ziel steht und auf ein zur Zielerreichung absolut notwendiges Maß beschränkt wird [[Ja21], Rn. 17 ff., [Be17], Rn. 4]. Es ist hierbei eine situationsorientierte Prüfung vorzunehmen, weswegen pauschale Anknüpfungen an bestimmte Behörden oder Aufgaben unzulässig sind [[Eu15], [Pa21], Rn. 9, [Be17], Rn. 4]. Indem personenbezogene Daten, abweichend vom ursprünglichen Erhebungszweck, zur KI-Entwicklung weitergenutzt werden können, ist die Aufhebung der Zweckbindung geeignet, um das Training von KI zu verbessern. Ein Training der KI mit personenbezogenen Daten wird für einige Entwickler allerdings nicht erforderlich sein, da anonymisierte oder synthetische Daten verwendet werden können. Deswegen ordnet Art. 54 Abs. 1 lit. b KI-VO-KomE an, dass nur solche Entwickler von dem Zweckbindungsgrundsatz befreit werden können, welche Daten zur Erfüllung der Anforderungen an Hochrisiko-KI benötigen.

Daher kommt es darauf an, ob Art. 54 Abs. 1 KI-VO-KomE außer Verhältnis zum verfolgten Ziel steht. Aufseiten der betroffenen Person steht das gewichtige Recht auf Schutz personenbezogener Daten aus Art. 8 GRCh, vorliegend insbesondere in der Ausprägung des Zweckbindungsgrundsatzes in Art. 8 Abs. 2 S. 1 GRCh. Dies stellt eines der fundamentalen Prinzipien des Datenschutzrechts dar, das der betroffenen Person ermöglicht, die vorgenommenen Verarbeitungen überblicken zu können und bei potenziellen Überschreitungen davon ausgehend Betroffenenrechte geltend zu machen. Dem stehen die nicht minder gewichtigen Gründe des Allgemeinwohls gegenüber, die in Art. 54 Abs. 1 KI-VO-KomE genannt werden. Bei denen handelt es sich – wie bereits dargestellt – um im besonderen öffentlichen Interesse liegende Ziele, die zudem Grundrechte Dritte berühren. Beispielsweise werden auf Grund der Auswirkungen des Klimawandels Starkwetterereignisse vermehrt auftreten, wodurch neben Eigentum vor allem Leib und Leben Dritter betroffen sein können. Zudem können sich aus nicht

ausreichender Sicherung der Kritischen Infrastrukturen im Falle eines Ausfalls Versorgungsengpässen der Bevölkerung ergeben, die, wenn sie beispielsweise Krankhäuser betreffen, erhebliche Konsequenzen haben können.

Zu berücksichtigen ist, dass die verfolgten Ziele allgemein gefasst sind und durch den Ratsentwurf und den Parlamentsvorschlag in Teilen erweitert, in Teilen aber auch begrenzt wurden, insbesondere was den sensiblen Bereich der Strafverfolgung und Strafvollstreckung anbelangt. Jedoch sind mit Gesundheitsschutz weiterhin äußerst grundrechtssensible Bereiche betroffen.

Es muss berücksichtigt werden, dass eine zu enge Begrenzung auf einzelne gegenwärtig bekannte Aspekte innovationshemmend wirken kann, wodurch das anvisierte Ziel des Titels V des KI-VO-KomE wiederum gefährdet wäre. Für eine verhältnismäßige Einschränkung des Zweckbindungsgrundsatzes streiten die Sicherungsmaßnahmen, die von Art. 54 Abs. 1 lit. b-j KI-VO-KomE vorgegeben werden. Diese Sicherungsmaßnahmen erinnern an technische und organisatorische Maßnahmen i. S. d. DSGVO und sind ähnlich abstrakt gefasst und können so auch auf diverse und sich besonders dynamisch entwickelnde KI-Anwendungen angewendet werden. Durch Art. 54 Abs. 1 lit. b KI-VO-KomE wird klargestellt, dass personenbezogene Daten nur in den zur Erfüllung der in Titel III Kapitel 2 genannten Anforderungen erforderlich ist. Andernfalls soll auf anonyme, synthetische oder nicht personenbezogene Daten zurückgegriffen werden, wodurch dem Grundsatz der Datenminimierung Rechnung getragen wird. In eine ähnliche Richtung geht die sich aus Art. 54 Abs. 1 lit. g KI-VO-KomE ergebende Pflicht zur Löschung personenbezogener Daten, soweit die Beteiligung am Reallabor beendet oder die Speicherfrist abgelaufen ist. Hierdurch wird der Grundsatz der Speicherbegrenzung berücksichtigt. In diesem Kontext ist problematisch, dass auch nach der erforderlichen Löschung der Daten eine Rekonstruktion der personenbezogenen Daten aus trainierten KI-Systemen möglich sein kann [[Bo22], S. 411]. Daher könnte man erwägen, den Einsatz von z. B. Differential Privacy oder Kryptografie vorzuschreiben, die sich jeweils am Stand der Technik orientieren [[WBH]]. Gleichzeitig werden die Erkenntnisse, die eine KI-Anwendung gewonnen hat, verstetigt (siehe 4.1). Hier besteht ein gewichtiger Unterschied zu Reallaboren in anderen Bereichen. Mit Beendigung des Testbetriebs in Reallaboren endet die rechtliche Privilegierung. Bei KI-Reallaboren wirkt die temporäre Privilegierung hingegen durch das „Behalten des Gelernten“ fort.

Im Ratsentwurf sowie im Parlamentsvorschlag wird zudem in Art. 54 KI-VO-KomE die Ergänzung von angemessenen technischen und organisatorischen Schutzmaßnahmen vorgeschlagen, wodurch auf Art. 25, 32 DSGVO angespielt wird. Dass die Entwickler Schutzmaßnahmen ergreifen sollen, kann indes schon aus Art. 54 Abs. 1 lit. d KI-VO-KomE herausgelesen werden, wenn es dort heißt, dass die Datenverarbeitungsumgebung geschützt sein muss. Durch die Erweiterung wird hingegen die besondere Bedeutung dieser Schutzmaßnahmen nochmal verdeutlicht.

Art. 54 Abs. 1 lit. f KI-VO-KomE betont zudem die Funktionsweise der KI-Reallabore, indem vorgeschrieben wird, dass die in diesem Umfeld getätigten Verarbeitungen keine Auswirkungen auf die betroffenen Personen haben dürfen.

Daneben werden Kontroll- und Eingreifmechanismen sowie Transparenzaspekte adressiert. Die Art. 54 Abs. 1 lit. c-e KI-VO-KomE sehen Pflichten zur Etablierung wirksamen Überwachungsmechanismen vor, um bei potenziellen hohen Risiken für betroffene Personen die ebenfalls vorzuhaltenden Reaktionsmaßnahmen umsetzen zu können. Der Ratsentwurf und der Parlamentsvorschlag ergänzen den Normtext durch einen expliziten Verweis auf Art. 35 DSGVO sowie Art. 39 Verordnung (EU) 2018/1725⁶, wodurch der Umfang der zu berücksichtigen potenziellen Risiken nochmals konkretisiert wird. Zudem muss die Nutzung der personenbezogenen Daten in einer funktional getrennten und besonders geschützten Umgebung vorgenommen werden, auf die Dritte keinen Zugriff haben und eine entsprechende Übermittlung ebenfalls nicht besteht. Ergänzt werden die Kontroll- und Reaktionsmechanismen durch umfangreiche Transparenzpflichten in den Art. 54 Abs. 1 lit. h) - j) KI-VO-KomE. Hierin werden Protokollpflichten sowie technische und Prozessdokumentation vorgeschrieben. Zudem muss eine Zusammenfassung erstellt werden, die von der zuständigen Behörde veröffentlicht werden kann, um die Öffentlichkeit mit einzubeziehen. Ergänzend hierzu hätte die Einbindung von Mechanismen zur Nachvollziehbarkeit von KI explizit mit vorgeschrieben werden können, um die Transparenz tatsächlich umsetzen zu können [[Ce]]. Zudem ergibt sich aus dem Umstand, dass Art. 54 KI-VO-KomE keinen Ausschluss der Betroffenenrechte vorsieht, dass diese weiterhin gelten [[Ho22], S. 563]. Hierdurch kann auf die weite Formulierung der Ziele mit angemessenen Maßnahmen reagiert werden. Zudem kann bei der Abwägung auch nicht vollkommen außer Acht gelassen werden, dass der freie Datenverkehr im Binnenmarkt auch als Schutzgut in Art. 1 Abs. 1, 3 DSGVO [[En22]] als gleichwertig neben dem Schutz der personenbezogenen Daten geregelt ist.

Dies führt zu dem Ergebnis, dass grundsätzlich von einer Einhaltung der Anforderungen des Art. 6 Abs. 4, 1. Hs. DSGVO ausgegangen werden kann, sodass es auf die Prüfung der Zweckvereinbarkeit nach Art. 6 Abs. 4, 2. Hs. DSGVO nicht mehr ankommt.

5 Zusammenfassung und Fazit

Der Unionsgesetzgeber hat richtigerweise erkannt, dass die Potenziale von KI nutzbar gemacht werden sollten und dass das bestehende Recht hierfür keinen angemessenen Regulierungsrahmen bereithält. Der Kommissionentwurf wurde unter vielen verschiedenen Gesichtspunkten juristisch diskutiert; verhältnismäßig zurückhaltend sind

⁶ Der Verweis auf Art. 35 Verordnung (EU) 2018/1725 anstatt Art. 39 Verordnung (EU) 2018/1725 dürfte ein redaktionelles versehen sein.

die juristischen Diskussionen zur Möglichkeit der Einrichtung von KI-Reallaboren ausgefallen.

Dass dieses aus anderen Bereichen bereits bekannte Instrument des Reallabors auch für KI vorgesehen wurde, ist zu begrüßen. Dem Datenschutz ist es förderlich, dass auch innerhalb der Reallabore die Vorgaben der DSGVO zu beachten sind und sogar die Ausnahme von der Zweckbindung nach Art. 54 KI-VO-E sich innerhalb der Vorgaben der DSGVO bewegt. Denn die besondere Bedeutung erlangt der Zweckbindungsgrundsatz aus der verfassungsrechtlichen Verankerung im Recht auf informationeller Selbstbestimmung aus Art. 2 Abs. 1 i. V. m. Art. 1 Abs. 1 GG und der unmittelbaren Formulierung in Art. 8 Abs. 2 S. 1 GRCh. Art. 54 Abs. 1 KI-VO-KomE sieht eine Ausnahme von diesem Zweckbindungsgrundsatz vor, dessen Zulässigkeit auf Grund der abschließenden Regelung in Art. 6 Abs. 4 i. V. m. Art. 23 DSGVO an dem dort aufgestellten Maßstab zu messen ist. Wenngleich die grundsätzlich als vereinbar mit Art. 23 Abs. 1 DSGVO anzusehenden Ziele weit formuliert sind, kann dieser Umstand durch die umfangreichen Pflichten in Art. 54 Abs. 1 KI-VO-KomE aufgefangen werden. Jedoch könnten weitere Maßnahmen wie die Berücksichtigung von Maßnahmen des Differential Privacy oder der Kryptografie oder solche zur Erhöhung der Transparenz von KI, in die Regulierung aufgenommen werden.

Da nunmehr das Europäische Parlament seine finale Fassung beschlossen hat [[Eu23a]], können die Trilogverhandlungen beginnen. Es wird spannend zu beobachten sein, welche Positionen der drei beteiligten Akteure sich durchsetzen werden und bzw. ob gänzlich neue Inhalte Eingang in die Gesetzesfassung finden werden. Hier sollte sich die Union ernsthaft mit der Innovationsförderung der KI-VO auseinandersetzen. Denn was den Datenschutz stärkt, wird die Attraktivität der Reallabore mindern. Art. 54 KI-VO-KomE war bereits die einzige Bereichsausnahme, die im Titel V zur Innovationsförderung vorgesehen war. Darüber hinaus ist der Anwendungsbereich des Art. 54 KI-VO-KomE denkbar klein. Wenn richtigerweise der Datenschutz auch in Reallaboren hochgehalten wird, so ist zu überlegen, inwiefern Entwickler anderweitig bei der Datengewinnung für das KI-Training unterstützt werden können. Mit der unechten Ausnahmeregelung des Art. 54 KI-VO-KomE wird keinem Entwickler geholfen sein. Anderenfalls lässt der Unionsgesetzgeber erhebliches Potenzial der Reallabore als neues Regelungselement ungenutzt.

Literaturverzeichnis

- [AB22] Accenture; Bitkom: Datenschutz als Herausforderung für die Digitalisierung. Studie zur Evaluierung von Anwendungsfällen von Datenschutzhürden in Europa mit Schwerpunkt Deutschland. <https://www.bitkom.org/sites/main/files/2023-01/Studie-Datenschutz-als-Herausforderung-fur-die-Digitalisierung-final.pdf>, Zugriff 25.08.2023.
- [AV23] Albers, M.; Veit, R.-D.: Art. 6. In (Wolff, H. A.; Brink, S.; von Ungern-Sternberg, A. Hrsg.): BeckOK Datenschutzrecht, 2023.
- [Be17] Bertermann, N.: Art. 23. In (Ehmann, E.; Selmayr, M. Hrsg.): Datenschutz-Grundverordnung, 2017.
- [Bo22] Botta, J.: Die Förderung innovativer KI-Systeme in der EU - Zum Kommissionsvorschlag der KI-Reallabore. ZfDR (Zeitschrift für Digitalisierung und Recht) 2022, S. 391–412, 2022.
- [Br23] Braun, T.: Warum das KI-Gesetz in der EU notwendig ist. https://www.sr.de/sr/home/nachrichten/politik_wirtschaft/eu_ki_gesetz_ei_nordnung_saarbruecken_dfki_ceo_krueger_100.html, Zugriff 25.08.2023.
- [Bu19] Bundesministerium für Wirtschaft und Klimaschutz: Freiräume für Innovationen. Das Handbuch für Reallabore. https://www.bmwk.de/Redaktion/DE/Publikationen/Digitale-Welt/handbuch-fuer-reallabore.pdf?__blob=publicationFile, Zugriff 25.08.2023.
- [CD17] Culik, N.; Döpke, C.: Zweckbindungsgrundsatz gegen unkontrollierten Einsatz von Big Data-Anwendungen. Zeitschrift für Datenschutz, S. 226–230, 2017.
- [Ce] Cerrato, Coronel, Cerrato et al.: Fair Interpretable Representation Learning with Correction Vectors. arXiv:2202.03078, 2022.
- [Di19] Dix, A.: Art. 23. In (Simitis, S. et al. Hrsg.): Datenschutzrecht. DSGVO mit BDSG. Nomos, Baden-Baden, 2019.
- [EE21] European Data Protection Board; European Data Protection Supervisor: Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). https://edps.europa.eu/system/files/2021-06/2021-06-18-edpb-edps_joint_opinion_ai_regulation_en.pdf, Stand 25.08.2023.
- [En22] Engeler, M.: Der Konflikt zwischen Datenmarkt und Datenschutz. Eine ökonomische Kritik der Einwilligung. Neue Juristische Wochenschrift, S. 3398–3405, 2022.

- [ES21] Ebert, A.; Spiecker gen. Döhmman, I.: Der Kommissionsentwurf für eine KI-Verordnung der EU. *Neue Zeitschrift für Verwaltungsrecht*, S. 1188–1193, 2021.
- [Eu15] EuGH, 2015.
- [Eu21] Europäische Kommission: Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>, Stand 25.08.2023.
- [Eu23a] Europäisches Parlament: Abänderungen des Europäischen Parlaments vom 14. Juni 2023 zu dem Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_DE.html, Stand 25.08.2023.
- [Eu23b] Europäisches Parlament: Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. <https://www.europarl.europa.eu/resources/library/media/20230516RES90302/20230516RES90302.pdf>, Stand 25.08.2023.
- [Ge21] Geminn, C.: Die Regulierung Künstlicher Intelligenz. *Zeitschrift für Datenschutz*, S. 354–359, 2021.
- [GMZ23] Gürkov, C.; Meyer-Fünffinger, A.; Zierer, M.: IT-Dienstleister des Bundes im Visier. <https://www.tagesschau.de/investigativ/br-recherche/it-dienstleister-bundesbehoerden-hacker-100.html>, Stand 25.08.2023.
- [Gr18] Groneman, T.: Die „Regulatory Sandbox“ – Ein Modell auch für das deutsche Bankaufsichtsrecht? <https://www.fch-gruppe.de/Beitrag/733/die-regulatory-sandbox-ein-modell-auch-fuer-das-deutsche-bankaufsichtsrecht>, Stand 25.08.2023.
- [Ho22] Hornung, G.: KI-Regulierung im Mehrebenensystem. Der KI-Verordnungsentwurf und regulatorische Perspektiven der Mitgliedstaaten. *DuD*, S. 561–566, 2022.
- [Ja21] Jarass, H. D.: Art. 8. In (Jarass, H. D. Hrsg.): *EU-Grundrechte-Charta*, 2021.

- [KB14] Katko, P.; Babaei-Beigi, A.: Accountability statt Einwilligung? Führt Big Data zum Paradigmenwechsel im Datenschutz? *Multimedia und Recht*, S. 360–364, 2014.
- [KM16] Kühling, J.; Martini, M.: Die Datenschutz-Grundverordnung: Revolution oder Evolution im europäischen und deutschen Detenschutzrecht? *EuZW*, S. 448–453, 2016.
- [Kr21] Krönke, C.: Sandkastenspiele - „Regulatory Sandboxes“ aus der Perspektive der Allgemeinen Verwaltungsrechts. *JuristenZeitung*, S. 434–443, 2021.
- [Le22] Leonetti, E.: Besonderheiten in Bezug auf fahrzeugführerlos-konzipierte Shuttlebussystem (ADS-DV). In (Chibanguza, K.; Kuß, C.; Steege, H. Hrsg.): *Künstliche Intelligenz. Recht und Praxis automatisierter und autonomer Systeme. Nomos, Baden-Baden, 2. Teil § 3 Q*, 2022.
- [Li22] Link, Hendrik: EU-Kommission: Entwurf einer Richtlinie zur KI-Haftung. *ZD-Aktuell 19/2022*, S. 1372, 2022.
- [Mv22] Madiega, T.; van de Pol, A. L.: Artificial intelligence act and regulatory sandboxes. *Members Research Service PE 733.544*. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPR_S_BRI\(2022\)733544_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPR_S_BRI(2022)733544_EN.pdf), Stand 25.08.2023.
- [No22] Noerr: Analyse der Potentiale und rechtlichen Umsetzungsmöglichkeiten von KI-Reallaboren auf europäischer und nationaler Ebene unter besonderer Berücksichtigung des Entwurfs der Europäischen Kommission für einen KI-Rechtsrahmen. https://www.bmwk.de/Redaktion/DE/Publikationen/Studien/gutachten-noerr-reallabore.pdf?__blob=publicationFile&v=6, Stand 25.08.2023.
- [Pa21] Paal, B. P.: Art. 23. In (Paal, B. P.; Pauly, D. A. Hrsg.): *Datenschutz-Grundverordnung*, 2021.
- [PA21] Pop, F.; Adomavicius, L.: *Sandboxes for Responsible Artificial Intelligence*. EIPA Briefing September 2021, 2021.
- [Ra20] Rat der Europäischen Union: *Schlussfolgerungen des Rates zu Reallaboren und Experimentierklauseln als Instrumente für einen innovationsfreundlichen, zukunftsicheren und resilienten Rechtsrahmen zur Bewältigung disruptiver Herausforderungen im digitalen Zeitalter*. Beratungsergebnisse 13026/20. <https://data.consilium.europa.eu/doc/document/ST-13026-2020-INIT/de/pdf>, Stand 25.08.2023.
- [Ra21a] Ranchordas, S.: *Experimental lawmaking in the EU: Regulatory Sandboxes*, 2021.

- [Ra21b] Ranchordas, S.: Experimental Regulations for AI: Sandboxes for Morals and Mores. *Morals & Machines* 1/1, S. 86–100, 2021.
- [Ra22] Rat der Europäischen Union: Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. <https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/de/pdf>, Stand 25.08.2023.
- [Ri16] Richter, P.: Big Data, Statistik und die Datenschutz-Grundverordnung. *Datenschutz und Datensicherheit* 9/40, S. 581–586, 2016.
- [RL23] Rossek, C.; Link, Hendrik: EU-Kommission: Entwurf einer überarbeiteten Produkthaftungs-RL – Relevanz hinsichtlich der Haftung für KI. *ZD-Aktuell*, S. 1105, 2023.
- [RN15] Roßnagel, A.; Nebel, M.: (Verlorene) Selbstbestimmung im Datenmeer. *Privatheit im Zeitalter von Big Data*. *DuD*, S. 455–459, 2015.
- [Ro19] Roßnagel, A.: Art. 6 Abs. 4. In (Simitis, S. et al. Hrsg.): *Datenschutzrecht. DSGVO mit BDSG*. Nomos, Baden-Baden, 2019.
- [Ro22] Roth, O. M.: Die "regulatory sandbox" als Regulierungsansatz für FinTechs - eine rechtsvergleichende Analyse. *Zeitschrift für Bankrecht und Bankwirtschaft*, S. 364–377, 2022.
- [RW21] Rostalski, F.; Weiss, E.: Der KI-Verordnungsentwurf der Europäischen Kommission. *ZfDR (Zeitschrift für Digitalisierung und Recht)*, S. 329–357, 2021.
- [Si84] Simitis, S.: Die informationelle Selbstbestimmung. Grundbedingung einer verfassungskonformen Informationsordnung. *Neue Juristische Wochenschrift*, S. 398–405, 1984.
- [Ta23] Tagesschau: Wegen Kabeldiebstahl Bahnstrecke bei Bochum erneut gesperrt. <https://www.tagesschau.de/inland/regional/nordrheinwestfalen/bochum-kabeldiebstahl-bahnstrecke-100.html>, Stand 25.08.2023.
- [WBH] Winter, C.; Battis, V.; Halvani, O.: Herausforderungen für die Anonymisierung von Daten. Technische Defizite, konzeptuelle Lücken und rechtliche Fragen bei der Anonymisierung von Daten. *ZD* 2019, S. 489–493.

Kritisches Denken - Konzeptionelle
Herausforderungen für die KI -
Zukünfte

Zukunft durch hybride IT- und KI-Innovationen

Klaus Mainzer^{1,2,3}

Abstract: Um Zukünfte zu gestalten, fordert dieser Artikel ein Innovationsportfolio von IT- und KI-Technologien, die als Basis-, Brücken- und Zukunftstechnologien aufeinander abgestimmt sind [M23]. Wie ein Portfolio von Aktien ist auch ein Innovationsportfolio dynamisch und muss ständig gestaltet werden. Methodisch wird dazu auf die mathematische Theorie komplexer Systeme und nichtlinearer Dynamik zurückgegriffen, mit der sich komplexe Systeme und Netzwerke in Natur, Wirtschaft und Gesellschaft modellieren lassen. Chaos und Risiken werden auf dieser Grundlage in Frühwarnsystemen abschätzbar und in strategisches Handeln umsetzbar.

Ziel ist ein nachhaltiges Innovationsportfolio (3.), in dem hybride KI mit symbolischer KI und subsymbolische KI des Machine Learning (3.1), hybrides Computing mit klassischem Computing und Quantencomputing (3.2), Hybridrechner mit Analogrechner und Digitalcomputer (3.3), hybride Robotik bzw. Embodied Robotik („Embodied Mind“) mit analoger Robotik und digitaler Robotik (3.4) und neurobiologisch orientiertem Computing (3.5) verbunden werden. Eine große Herausforderung für die Zukunft ist die Sicherung dieser Technologien durch Verifikation und Erklärbarkeit [MK22].

Keywords: Zukunftstechnologien, Innovationsportfolio, hybride KI, Machine learning, Quantencomputing, neuromorphes Rechnen, analoges Rechnen, symbolische KI, subsymbolische KI, statistisches Lernen, kausales Lernen, Verifikation.

1 Einführung

Zukunftstechnologien sind wie Aktien Wetten auf die Zukunft. Im Fall von Energietechnologien müssen für Europa Vor- und Nachteile von z.B. Solar- und Windenergie, Wasserkraft, Wasserstoff, Kern- und Fusionsenergie gegeneinander abgewogen und in einem „hybriden“ Energiesystem miteinander verbunden werden, um das europäische Innovationsportfolio auf eine nachhaltige Zukunft auszurichten. Ein Innovationsportfolio sorgt für Diversität und resilientes Verhalten, um auf die Risiken der Zukunft flexibel reagieren und sich bei punktuellen Rückschlägen insgesamt erholen zu können [Ma23].

Versorgungs- und Lieferketten der Energie sind nur ein Beispiel für komplexe Netzwerke moderner Zivilisation, deren Daten- und Informationsmengen ohne Digitalisierung und Künstliche Intelligenz (KI) nicht zu bewältigen sind. Allerdings benötigt die Digitalisierung ein gewaltiges Energieaufkommen, das bei verschiedenen Rechnertechnologien unterschiedlich ausfällt. Wie in der Energiefrage darf bei der

¹ TUM Senior Excellence Faculty, Technische Universität München

² Carl Friedrich von Weizsäcker Center, Universität Tübingen

³ Presidential Office, Europäische Akademie der Wissenschaften und Künste

Digitalisierung nicht auf eine einzige Lösung gesetzt werden, sondern es muss das gesamte technologische Potential in einem Innovationsportfolio gebündelt werden. In diesem Artikel werden klassische Digitalisierung und Künstliche Intelligenz zusammen mit Zukunftstechnologien wie neuromorphes Rechnen (nach dem energiesparenden Modus natürlicher Gehirne) ebenso berücksichtigt wie Quantenrechner, Quantenkommunikation und Quantentechnologie. Dazu müssen für Europa Vor- und Nachteile der digitalen und analogen Technologien gegeneinander abgewogen und in einer „hybriden“ IT und KI verbunden werden, damit auch dieses europäische Innovationsportfolio zusammen mit der Energiefrage auf eine nachhaltige Zukunft ausgerichtet wird.

Ein Innovationsportfolio setzt sich aus Basis-, Brücken- und Zukunftstechnologien zusammen, die sich kurz- und mittelfristig ändern, aufgegeben und durch neue ersetzt werden müssen. Wie ein Portfolio von Aktien ist auch ein Innovationsportfolio dynamisch und muss ständig gestaltet werden. Methodisch wird dazu auf die mathematische Theorie komplexer Systeme und nichtlinearer Dynamik zurückgegriffen, mit der sich komplexe Systeme und Netzwerke in Natur, Wirtschaft und Gesellschaft modellieren lassen. Chaos und Risiken werden auf dieser Grundlage in Frühwarnsystemen abschätzbar und in strategisches Handeln umsetzbar.

2 Von der klassischen KI zum Machine Learning

Informations- und Kommunikationsnetzwerke sind der Hintergrund einer weltweiten Automatisierung durch Künstliche Intelligenz. In Zukunft werden enorme Rechnerkapazitäten notwendig sein, um die gewaltigen Datenmengen dieser Zivilisation zu bewältigen. Die Komplexität des Lebens, seine unverstandenen Zusammenhänge, seine Empfindlichkeit und Gefährdung, die sich in Krankheiten wie Krebs ebenso zeigt wie in viralen Pandemien, erfordert neue Tools in Life Sciences und Medizin. Hier wird Bioinformatik zunehmend auf Machine Learning und geeignete Rechner- und Speicherkapazitäten zurückgreifen müssen. Das gleiche gilt zur Bewältigung weltweiter Finanz- und Wirtschaftskrisen, die Frühwarnsysteme erfordern. So wachsen Informations- und Kommunikationsnetze mit dem großen anderen HighTech Hype, der Künstlichen Intelligenz, zusammen.

Die Digitalisierung der menschlichen Zivilisation erfordert einen ungeheuren Bedarf an Energie. Rechnertechnologie und Informations- und Kommunikationsnetze sind daher nicht getrennt vom Energieverbrauch dieser Zivilisation zu sehen. Damit wird zugleich die ökologische Dimension der Digitalisierung deutlich, da fossile Energien die Umwelt belasten. Was können wir von der Natur lernen, die äußerst effektive und energiesparende Gehirne und Nervensysteme in der Evolution hervorgebracht hat? Hier kommt das neuromorphe Rechnen auf memristiven Architekturen und photonischen Rechnern ins Spiel, dessen Rechnerarchitekturen den Gehirnen der Evolution nachempfunden sind [Ne21]. Statt energiefressender von-Neumann Architekturen der klassischen Rechner vom Smartphone und PC bis zum Supercomputer kommen nun Rechereinheiten zum Einsatz,

die zwar noch auf Hardware, aber nach den effizienten Methoden von Neuronen und Synapsen in Gehirnen arbeiten. Statt einer totalen Digitalisierung nutzen sie zudem die Vorteile analoger Verfahren, wie man sie von lebenden Organismen kennt.

Nach Elektrifizierung und Digitalisierung im 20. Jahrhundert steht nun die Quantisierung der Kommunikations- und Versorgungsnetzwerke an. Das geschieht schrittweise und nicht „disruptiv“. Auch der universelle Quantencomputer wird nicht „disruptiv“ die klassische Rechnertechnologie ablösen, sondern zunehmend in klassische und neuromorphe Rechnerstrukturen eingebettet und neue Aufgaben lösen, die mit diesen Verfahren ausgeschlossen waren. Man spricht bereits von „ökologischen“ und hybriden Rechnernetzwerken, die sich schrittweise über die Welt ausbreiten.

2.1 Symbolische KI: Logik und Deduktion

In einer ersten Phase orientierte sich KI an formalen (symbolischen) Kalkülen der Logik, mit denen Problemlösungen regelbasiert abgeleitet werden können. Man spricht deshalb auch von symbolischer KI [Ma19]. Ein typisches Beispiel ist das automatische Beweisen mit logischen Deduktionen, die sich mit Computerprogrammen realisieren lassen. Automatisierung bedeutet bis zu einem bestimmten Grad auch Autonomie, da Computerprogramme die Beweistätigkeit eines Mathematikers übernehmen. Wissensbasierte Expertensysteme sind Computerprogramme, die Wissen über ein spezielles Gebiet speichern und ansammeln, aus dem Wissen automatisch Schlussfolgerungen ziehen, um zu konkreten Problemen des Gebietes Lösungen anzubieten. Im Unterschied zum menschlichen Experten ist das Wissen eines Expertensystems aber auf eine spezialisierte Informationsbasis beschränkt ohne allgemeines und strukturelles Wissen über die Welt [Pu88].

Um ein Expertensystem zu bauen, muss das Wissen des Experten in Regeln gefasst werden, in eine Programmiersprache übersetzt und mit einer Problemlösungsstrategie bearbeitet werden. Die Architektur eines Expertensystems besteht daher aus den folgenden Komponenten: Wissensbasis, Problemlösungskomponente (Ableitungssystem), Erklärungskomponente, Wissenserwerb, Dialogkomponente. In dieser Architektur werden zugleich die Grenzen symbolischer KI deutlich: Fähigkeiten, die nicht oder nur schwer symbolisch erfasst und regelbasiert simuliert werden können, bleiben der symbolischen KI verschlossen.

2.2 Subsymbolische KI: Statistik und Induktion

Sensorische und motorische Fähigkeiten werden nicht aus Lehrbuchwissen logisch abgeleitet, sondern aus Beispielen erlernt, trainiert und eingeübt. So lernen wir, uns motorisch zu bewegen und in einer Vielzahl sensorischer Daten Muster und Zusammenhänge zu erkennen, an denen wir unser Handeln und Entscheiden orientieren können. Da diese Fähigkeiten nicht von ihrer symbolischen Repräsentation abhängen, spricht man auch von subsymbolischer KI. An die Stelle der formalen Schlüsse der Logik

tritt nun die Statistik der Daten. Beim statistischen Lernen sollen allgemeine Abhängigkeiten und Zusammenhänge aus endlich vielen Beobachtungsdaten durch Algorithmen abgeleitet werden [Va98]. An die Stelle der Deduktion in der symbolischen KI tritt also in der subsymbolischen KI die Induktion. Dazu können wir uns ein naturwissenschaftliches Experiment vorstellen, bei dem in einer Serie von veränderten Bedingungen (Inputs) entsprechende Ergebnisse (Outputs) folgen. In der Medizin könnte es sich um einen Patienten handeln, der auf Medikamente in bestimmter Weise reagiert.

Die derzeitigen Erfolge des Machine Learning scheinen die These zu bestätigen, dass es auf möglichst große Datenmengen ankommt, die mit immer stärkerer Computerpower bearbeitet werden. Die erkannten Regularitäten hängen dann aber nur von der Wahrscheinlichkeitsverteilung der statistischen Daten ab.

Statistisches Lernen versucht, ein probabilistisches Modell aus endlich vielen Daten von Ergebnissen (z.B. Zufallsexperimente) und Beobachtungen abzuleiten.

Statistisches Schließen versucht umgekehrt, Eigenschaften von beobachteten Daten aus einem angenommenen statistischen Modell abzuleiten.

2.3 Lernen mit neuronalen Netzen

In der Automatisierung statistischen Lernens nehmen neuronale Netze mit Lernalgorithmen eine Schlüsselrolle ein. Neuronale Netze sind vereinfachte Rechenmodelle nach dem Vorbild des menschlichen Gehirns, in denen Neuronen mit Synapsen verbunden sind. Die Intensität der neurochemischen Signale, die zwischen den Neuronen ausgesendet werden, sind im Modell durch Zahlengewichte repräsentiert. Probabilistische Netzwerke haben experimentell eine große Ähnlichkeit mit biologischen neuronalen Netzen. Werden Zellen entfernt oder einzelne Synapsengewichte um kleine Beträge verändert, erweisen sie sich als fehlertolerant gegenüber kleineren Störungen wie das menschliche Gehirn z.B. bei kleineren Unfallschäden. Das menschliche Gehirn arbeitet mit Schichten paralleler Signalverarbeitung. So sind z.B. zwischen einer sensorischen Inputschicht und einer motorischen Outputschicht interne Zwischenschritte neuronaler Signalverarbeitung geschaltet, die nicht mit der Außenwelt in Verbindung stehen.

Tatsächlich lässt sich auch in technischen neuronalen Netzen die Repräsentations- und Problemlösungskapazität steigern, indem verschiedene lernfähige Schichten mit möglichst vielen Neuronen zwischengeschaltet werden. Die erste Schicht erhält das Eingabemuster. Jedes Neuron dieser Schicht hat Verbindungen zu jedem Neuron der nächsten Schicht. Die Hintereinanderschaltung setzt sich fort, bis die letzte Schicht erreicht ist und ein Aktivitätsmuster abgibt [HSW89].

Künstliche neuronale Netze sind zwar äußerst effektiv, um komplexe Probleme (real world problems) zu bearbeiten. Was aber fehlt, sind Spezifikationen und Standards für die Sicherheit ihrer Outputs. Dazu muss die Black Box neuronaler Netze besser verstanden, kontrolliert und verifiziert werden. Die Verifikation neuronaler Netze ist allerdings ein

hartes Erkenntnisproblem: Selbst der Nachweis einfacher Eigenschaften erweist sich im Rahmen der Komplexitätstheorie als NP-vollständig. Gründe dafür sind die Größe der praktisch angewendeten Netze (Skalierung) und die nichtlinearen Aktivierungsfunktionen ihrer Neuronen, die von Menschen in diesem Umfang und mit dieser Geschwindigkeit nicht nachvollzogen werden können. Da neuronale Netze zudem der Dynamik komplexer Systeme unterliegen, sind sie häufig empfindlich gegen kleine Störungen und Veränderungen ihrer Inputs, die sich zu unkontrollierbaren Effekten aufschaukeln können. Robustheit und Stabilität der Netze hängt also mit ihrer Sicherheit eng zusammen.

Neuronale Netze sind der Schlüssel zum Erfolg des Machine Learning in technischer Anwendung und kommerzieller Praxis. Allerdings kommen dort mittlerweile Netze mit Hunderten und Tausenden von Schichten und Millionen von Neuronen zum Einsatz. Mit Tiefe (depth) wird die Anzahl der neuronalen Schichten bezeichnet. Man spricht mittlerweile nicht mehr nur von Deep Learning, sondern bei dieser enormen Anzahl von Schichten auch von „Deepest Learning“. Training und Abläufe dieser Netze sind nur noch statistisch beschreibbar und erscheinen zufällig. Verifikationen ihres Verhaltens stützen sich in der Regel nur noch auf punktuelle Tests mit Trial-and-Error Verfahren. Damit verschärft sich das Black Box Problem dieser Netze mit ihrer wachsenden Tiefe und Breite in der Praxis.

2.4 Vom statistischen zum kausalen Lernen

Statistisches Lernen und Schließen aus Daten reichen aber nicht aus. Wir müssen vielmehr die kausalen Zusammenhänge von Ursachen und Wirkungen hinter den Messdaten erkennen [Pe09]. Diese kausalen Zusammenhänge hängen von den Gesetzen der jeweiligen Anwendungsdomäne unserer Forschungsmethoden ab, also den Gesetzen der Physik, den Gesetzen der Biochemie und des Zellwachstums im Beispiel der Krebsforschung, etc. Wäre es anders, könnten wir mit den Methoden des statistischen Lernens und Schließens bereits die Probleme dieser Welt lösen.

Statistisches Lernen und Schließen ohne kausales Domänenwissen ist blind – bei noch so großer Datenmenge (Big Data) und Rechenpower!

Neben der Statistik der Daten bedarf es zusätzlicher Gesetzes- und Strukturannahmen der Anwendungsdomänen, die durch Experimente und Interventionen überprüft werden. Kausale Erklärungsmodelle (z.B. das Planetenmodell oder ein Tumormodell) erfüllen die Gesetzes- und Strukturannahmen einer Theorie (z.B. Newtons Gravitationstheorie oder die Gesetze der Zellbiologie):

Beim kausalen Schließen werden Eigenschaften von Daten und Beobachtungen aus Kausalmodellen, d.h. Gesetzesannahmen von Ursachen und Wirkungen, abgeleitet. Kausales Schließen ermöglicht damit, die Wirkungen von Interventionen oder Datenveränderungen (z.B. durch Experimente) zu bestimmen.

Kausales Lernen versucht umgekehrt, ein Kausalmodell aus Beobachtungen, Messdaten und Interventionen (z.B. Experimente) abzuleiten, die zusätzliche Gesetzes- und Strukturannahmen voraussetzen.

2.5 Verifikation von Machine Learning

Eine große Herausforderung des modernen Machine learning ist Prüfung seiner Korrektheit. In der mathematischen Logik heißt ein Formalismus korrekt mit Bezug auf ein Anwendungsmodell, wenn die formalen Ableitungen des Formalismus auch nur wahre Aussagen des Anwendungsmodells erfassen. Ein Computerprogramm ist korrekt, wenn bewiesen werden kann, dass das Programm auch tatsächlich das tut, was es tun soll. Ein solcher Korrektheitsbeweis heißt Verifikation des Computerprogramms [Ma22].

Ein klassisches Beispiel für formale Verifikation von Computerprogrammen ist die SAT-Methode. Sie geht auf die Anfänge der symbolischen KI zurück und basiert vollständig auf der formalen (symbolischen) Logik. In der Logik geht es beim Erfüllbarkeitsproblem (SAT) um die Frage, ob es eine Interpretation gibt, die eine gegebene boolesche Formel erfüllt. In diesem Fall wird die Formel als erfüllbar bezeichnet. Gibt es eine solche Zuordnung nicht, ist die durch die Formel ausgedrückte Funktion für alle möglichen Zuordnungen von Variablen falsch und die Formel ist nicht erfüllbar. Nach Cooks Theorem ist SAT NP-komplett. Es gibt keinen bekannten Algorithmus, der jedes SAT-Problem effizient löst. Aber heuristische SAT-Algorithmen können zumindest eingesetzt werden, um begrenzte Problemklassen mit Tausenden von Variablen und Formeln zu lösen [Bi09].

Es stellt sich die Frage, ob SAT-Methoden auch auf das maschinelle Lernen mit neuronalen Netzen angewendet werden können [PT12]. In diesem Fall müsste man neuronale Netzmodelle finden, die durch Rand- und Seitenbeschränkungen eingeschränkt sind. Neuronale Netzknoten von neuronalen Netzen haben aber oft kein lineares Input-Output-Verhalten. Sie können jedoch linear approximiert werden. Zu diesem Zweck wird die gewichtete Summe der Eingangssignale der Knoten als Variable c bezeichnet. Die Variable d sei der Ausgang eines Knotens. Wenn es obere und untere Grenzen $[l, u]$ für c gibt, kann die Beziehung zwischen c und d durch Randbedingungen wie $d \geq 0$, $d \geq c$ und $d \geq u(c - 1)/(u - 1)$ angenähert werden. Offensichtlich werden diese Randbedingungen durch lineare Gleichungen für die Konstanten l und u dargestellt [Eh17].

Im nächsten Schritt werden die Randbedingungen durch boolesche Formeln dargestellt. Die Idee ist, einen linearen Programmierer und einen SAT-Löser zu kombinieren. Der SAT-Solver soll prüfen, ob diese booleschen Formeln eine Erfüllbarkeitszuweisung haben. Zu diesem Zweck werden sie in eine konjunktive Normalform (CNF) umgewandelt, die aus durch Konjunktionen (Disjunktionen von Literalen) verbundenen Klauseln besteht.

Ein SAT-Solver arbeitet so, dass er die booleschen Variablen nacheinander auswertet. Backtracking findet immer dann statt, wenn ein Konflikt zwischen der aktuellen Auswertung und einer Klausel festgestellt wird. SAT-Solver können durch verschiedene Lernheuristiken erweitert werden. Ein Beispiel sind SMT (Satisfiability Modulo Theory) Solver, die SAT Solver mit spezialisierten Entscheidungsverfahren für andere Theorien kombinieren.

Die Verifikation von Feed-Forward Neuronalen Netzen kann formal in folgenden Schritten realisiert werden: Gegeben sei ein Feed-forward neuronales Netz G mit einer Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, die durch einen Satz linearer Constraints φ über den reellen Variablen $V = \{x_1, \dots, x_n, y_1, \dots, y_m\}$ definiert ist. Das Verifikationsproblem von G besteht darin, entweder eine Bewertungsfunktion α für die Variablen aus V zu finden, die φ über die Eingangs- und Ausgangsknoten von G mit $f(x_1, \dots, x_n) = (y_1, \dots, y_m)$ erfüllt, oder zu zeigen, dass keine solche Bewertungsfunktion für die Knoten existiert.

Bei einem linearen Programmlöser (LP) geht man von einer gegebenen Menge linearer Ungleichungen über reellen Variablen und einer linearen Optimierungsfunktion (lineares Programm) aus. Das Verifikationsproblem der linearen Programmierung besteht darin, eine Zuordnung zu den Variablen zu finden, die die Zielfunktion minimiert und alle Nebenbedingungen erfüllt. Ein fortschrittlicheres Instrument ist die Kombination aus einem Solver für lineare Programmierung (LP) und einem Solver für Erfüllbarkeit (SAT, SMT).

Ein Beweisassistent oder interaktiver Theorembeweiser (Interactive Theorem Prover) ist ein Softwarewerkzeug, das den Prozess des formalen Beweisens durch Mensch-Maschine-Interaktion unterstützt [MSS18]. In Computerprogrammen werden Datentypen verwendet, um Softwarefehler (Bugs) zu reduzieren. Die Typentheorie CoC (Calculus of Construction) ist die Grundlage für den Beweisassistenten Coq [CH88]. Coq implementiert eine Programmspezifikation, die auf einer Erweiterung von CoC, dem Calculus of Inductive Constructions (CiC), basiert und eine Logik höherer Ordnung mit einer reich typisierten funktionalen Sprache kombiniert [CJ96].

In Coq wird die Verifikation von Beweisen auf die Verifikation von Typen in Typentheorien wie CiC reduziert. Der Kern von Coq ist daher ein Beweisalgorithmus für Typen in der Sprache von CiC. Weitere Details zu Coq sind in verschiedenen Tutorials enthalten. Schließlich wurden Coq und CiC in fortgeschrittenen, schwierigen Beweisen verwendet (z.B. dem Vier-Farben-Satz).

Die Extraktion von zertifizierten Programmen arbeitet mit rekursiven Schemata von terminierenden Algorithmen. Die Extraktion von Programmen setzt voraus, dass eine originalgetreue Coq-Version des Zielprogramms in funktionaler Sprache in Coq eingebettet ist. Korrektheitseigenschaften der Coq-Version des Zielprogramms können dann bewiesen werden und eine funktionale Programmversion automatisch extrahieren. Wenn die automatische Extraktion gesichert ist, erfüllt das resultierende funktionale Programm die erwarteten Korrektheitseigenschaften.

In der Praxis sind formale Korrektheitsbeweise im Machine Learning bisher nur begrenzt anwendbar [TB03]. Eine enorme Herausforderung ist die Parameterexplosion der praktisch angewendeten neuronalen Netze wie beim ChatGPT. Häufig sind bisher nur statistische Stichproben möglich. Das ist eine Herausforderung, aber zugleich auch Chance für die logisch-mathematische Grundlagenforschung, praxisrelevante Ergebnisse für das Machine Learning zu liefern.

3 Hybride IT- und KI-Systeme als nachhaltiges Innovationsportfolio

Jede Technologie weist Vor- und Nachteile auf, die gegeneinander abzuwägen sind. Zudem erweisen sich Technologien als Brücken zu neuen und effizienteren Lösungen. Sie sind also immer auch „Brückentechnologien“. In einer solchen Situation wäre es ausgesprochen unklug, nur auf eine existierende Technologie zu setzen oder in Zukunft zu hoffen. Wie bei Aktien kommt es darauf an, ein Portfolio verschiedener Technologien zu vernetzen, um bei Ausfällen und Fehleinschätzungen einzelner Teile einen guten Schnitt zu haben. Ein solches Innovationsportfolio ist zudem in dynamischer Veränderung und muss gewartet und gepflegt werden [Ma23].

3.1 Hybride KI = symbolische KI + subsymbolische KI

Ein dynamisches Innovationsportfolio ist für Energietechnologien erforderlich, aber auch für Computertechnologien und die darauf aufbauenden Informations- und Kommunikationssysteme mit Künstlicher Intelligenz. Das moderne Machine und Deep Learning löst keineswegs die klassische KI „disruptiv“ ab. Vielmehr kommt es darauf an, symbolische KI als Kontroll- und Verifikationsverfahren mit den statistischen Methoden der subsymbolischen KI zu verbinden. Subsymbologische KI zielt vor allem auf Mustererkennung, wie es in Organismen in Wahrnehmungsvorgängen auftritt, während symbolische KI das logische Denken und Schließen des menschlichen Verstands (z.B. in Expertensystemen) abbildet. Die Verbindung von beiden Ansätzen wird als hybride KI bezeichnet und kommt damit der menschlichen Intelligenz näher als die Reduktion auf eine der beiden Ansätze [65].

3.2 Hybrides Computing = Klassisches Computing + Quanten Computing

Eine hybride Verbindung nimmt auch das klassische Computing vom Smartphone und PC bis zum Supercomputer mit dem Quanten Computing ein [Ma20a]. Quanten Computing wird, wenn es ausgereift ist, nicht „disruptiv“ das klassische Computing ablösen. Vielmehr werden die Funktionen des Quanten Computing in klassische Großrechneranlagen eingebettet sein, um für sie typische Aufgaben im Verbund mit klassischen Prozeduren zu lösen. Das gilt auch für die bereits existierenden D-Wave Rechner, die nicht wie

Quantencomputer auf Schaltkreise mit quantenlogischen Gattern aufbauen, sondern durch adiabatisches Rechnen wie in der Thermodynamik Optimierungsaufgaben lösen. Der Unterschied zum klassischen adiabatischen Rechnen besteht darin, dass bereits Tausende von Quantenbits verwendet werden, allerdings simuliert auf herkömmlichen Großrechnern. Es geht also wieder um das Lösen spezieller Aufgaben („Optimierungsaufgaben“) mit speziellen Methoden („adiabatisches Rechnen mit Quantenbits“), die in klassischen Rechneranlagen eingebettet sind.

3.3 Hybridrechner = Analogrechner + Digitalcomputer

Bevor universelle programmierbare Digitalcomputer in den 1940er Jahren die Grundlage für die moderne Digitalisierung legten, kamen Analogrechner zum Einsatz, die ebenfalls für das Lösen spezieller Aufgaben konstruiert wurden. Die Idee der Analogrechner ist uralte [U110a]: Man baut z.B. ein technisch-mechanisches Modell („Analogon“) des physikalischen Planetensystems, um damit Konstellationen der Planeten zu bestimmen. Bereits im ersten vorchristlichen Jahrhundert hatten griechische Mathematiker mit dem Mechanismus von Antikythera einen solchen Analogrechner vorgestellt, der durch Einstellungen eines Zahnradmechanismus für angenommene Himmelsphären solche Berechnungen durchführen konnte. Mit Aufkommen der Elektrotechnik konnten Röhren Transistoren etc. zur Realisation von Rechenaufgaben eingesetzt werden. Bei einer Differential- oder Integralgleichung wurden die Rechenoperatoren z.B. des Multiplizierens, Addierens, Integrierens, Differenzierens durch entsprechende („analoge“) elektrotechnische Einheiten dargestellt und zum Lösen der Aufgabe miteinander technisch verbunden. Die mathematische Gleichung lässt sich als Schaltplan für das analoge technisch-physikalische Modell verstehen, mit dem die Lösungen berechnet werden.

Ein Analogrechner benötigt daher eine Vielzahl von Rechenelementen, um alle die verschiedenen Operationen, Größen und Verknüpfungen in Aufgaben „analog“ zu repräsentieren. Demgegenüber benötigt ein Digitalcomputer nur wenige Recheneinheiten, die extrem schnell mit sehr einfachen Befehlen nacheinander abgearbeitet werden müssen. Es sind die Bits 0 und 1 mit wenigen Operationen des Addierens, Subtrahierens und Multiplizierens von Bits, die in wenigen logischen Gattern zu hoch komplexen Schaltkreisen verknüpft werden können. Für eine Additionsaufgabe muss ein Digitalcomputer zunächst die Bitfolge für den Befehl „Addiere“ aus einem Speicher lesen und ausführen. Die Summanden müssen ebenfalls als Bitfolgen aus einem Speicher ausgelesen werden, bevor die Operation an den Summanden ausgeführt wird. All das kostet Zeit und Energie, die sich bei Abermillionen von Rechenschritten für einzelne Bits aufpotenziert. Ein Analogrechner benötigt keine von-Neumann Architektur mit getrennten Speicher-, Kontroll- und Recheneinheiten und sequentiellen Ausführungsbestimmungen der Befehle in Algorithmen und Programmen, sondern löst die Aufgaben „direkt“ in einem elektronischen Modell.

Aus heutiger Sicht sind Analogrechner daher energieeffizient und zeitsparend, aber beschränkt auf spezielle Anwendungen, die sich z.B. in Differential- und Integralgleichungen darstellen lassen. Dabei handelt es sich vor allem um spezielle Aufgaben aus Technik-, Natur- und Naturwissenschaften, für die jeweils ein analoges Rechnermodell hergestellt werden muss. Mathematisch benutzen solche Gleichungen reelle Zahlen wie die Dezimalbrüche (z.B. $\pi = 3,1415 \dots$) mit beliebig („unendlich“) kleinen Größen und kontinuierlichen Prozessen im Unterschied zu digitalen Zahlen wie den beiden Bits. Analoges Computing bezieht sich daher auch auf das Rechnen mit reellen Zahlen [Ma18].

In den 1970er Jahren schien es zunächst so, dass Digitalcomputer wegen ihrer universellen Anwendbarkeit auf allen möglichen Gebieten Analogrechner ablösen würden. Wegen der ungeheuren Geschwindigkeiten der Digitalcomputer vom Smartphone bis zum Supercomputer schienen auch die typischen Aufgaben von Analogrechnern simulierbar. Unter dem Gesichtspunkt der Energieeffizienz und der Umweltschonung rücken nun Analogrechner wieder in das Zentrum des Interesses.

Unter Hybridrechnern versteht man mittlerweile auch die Kopplung von Digital- und Analogrechnern [U110b]. Dabei wird der Analogrechner als leistungsfähiger Ko-Prozessor für den Digitalrechner benutzt. Dazu wird allerdings auch Software benötigt, um den Analogrechner in der Rechnerumgebung des Digitalrechners programmieren zu können. Ähnlich wie die Funktionen eines Quantencomputers sind also Analogrechner für das Lösen spezieller Aufgaben in digitale Umgebung eingebettet. Man spricht auch von „ökologischen“ Rechnersystemen und meint damit die vielfältige Vernetzung von Rechnertypen.

3.4 Hybride Robotik = Embodied Robotik („Embodied Mind“) = Analoge Robotik + Digitale Robotik

Auch in der Robotik werden digitale und analoge Funktionen wie in einem Organismus verbunden. Kognitive und intellektuelle Fähigkeiten lassen sich nicht in einer vom Körper getrennten Software abbilden. Um Zusammenhänge und Strukturen zu erkennen, bedarf es Erfahrung durch Wahrnehmungsprozesse mit Sinnesorganen und Erleben von Bewegungsabläufen mit organischer Motorik wie Hände und Finger. In der Kognitionspsychologie und Kognitionsphilosophie spricht man von „embodied mind“: Der menschliche „Geist“ (mind) ist nicht isoliert, sondern „verkörperlicht“ (embodied) im Organismus. Viele körperliche Abläufe finden analog über Sensoren statt, Kontroll- und Steuerungsfunktion eher digital. Humanoide Roboter werden zunehmend nach dem Vorbild von „embodied mind“ in hybrider Kopplung von analogen und digitalen Funktionen entwickelt [Ma20].

Analoge Funktionen nach dem Vorbild des Gehirns kommen auch beim neuromorphen Rechnen ins Spiel. Biologische Gehirne sind keineswegs digitale Schalterkästen, in denen Neuronen „feuern“ und „nicht-feuern“, also digital zwischen zwei Zuständen als Bits 0

und 1 hin- und her schalten. Vielmehr haben synaptische Verbindungen analoge Komponenten, da sie kontinuierlich und graduell Gewichtungen berücksichtigen. Analoge Komponenten werden mit digitalen Komponenten in memristiven und photonischen neuronalen Systemen verbunden. Die Hardware neuromorpher Systeme ist allerdings mit Silizium und nanotechnischen Materialien realisiert und nicht mit lebendem Gewebe („Wetware“) wie in biologischen Gehirnen [Ne21]. Daher ist diese Hardware sicher robuster und kann nicht von Krankheiten wie biologisches Gewebe befallen werden. Auch sind Leistungssteigerungen und Dauerbelastungen möglich, die lebendes Gewebe nicht aushält bis hin zur Ermüdung von natürlichen Gehirnen.

3.5 Neurobiologisch hybride Gehirne

Ein entscheidender Nachteil der bisherigen Hardware neuromorpher Systeme ist aber die Tatsache, dass sie sich von sich aus nicht verändern und wie biologische Gehirne entwickeln können. Das biologische Gehirn eines Kindes wächst, Hardware eines neuromorphen Systems nicht. Dazu kodieren Gene den Wachstumsprozess, der zur Vernetzung der lebenden neuronalen Zellen führt. Dieser Wachstumsprozess hängt von zahlreichen Wechselwirkungen ab und unterliegt der nichtlinearen Dynamik eines komplexen Systems. Damit entstehen und entwickeln sich auch kognitive Funktionen und Intelligenz, Emotionen und Bewusstseinszustände und sind nicht fertig vorgegeben. Die Entwicklungspsychologie untersucht die gengesteuerten Phasenübergänge dieser Wachstumsprozesse, von denen die Entstehung kognitiver Fähigkeiten wie das Lernen abhängen.

Demgegenüber sind künstliche neuronale Netze technisch fertig konstruiert. Sie simulieren kognitive Fähigkeiten entweder wie im Machine Learning auf digitalen Computern oder realisieren sie mit der Hardware neuromorpher Systeme. Die „Verdrahtung“ der künstlichen Gehirne ist also bisher in Software oder Hardware vorgeben. Da diese künstlichen Netze hochkomplexe Systeme sind, bedarf es aber enormer Information, um sie mit allen Details zu beschreiben. Man spricht auch von einer Parameterexplosion im Deep Learning künstlicher neuronaler Netze.

Um das Wachstum biologische Gehirne zu erfassen, genügt die weitaus geringere Informationsmenge, um die steuernden Genome und Gene des Wachstumsprozesses zu erfassen [Hi21]. Wachstum ist zeit- und energieabhängig. Je mehr Zeit und Energie im Wachstum aufgewendet werden, je mehr Information kann bereits in ein Netz untergebracht werden, bevor es anfängt zu lernen. So kommt es, dass in den Gehirnen von Tierbabies Informationen untergebracht sind, die nicht erst aufwendig gelernt werden müssen. Sie können wie z.B. junge Fohlen gleich nach der Geburt stehen und die Milchzitzen ihrer Mütter finden. Für aufwendiges Trainieren der neuronalen Netze durch viele Beispiele mit großen Datenmengen ist im Überlebenskampf der Evolution keine Zeit.

Beim Entwurf neuronaler Netze hatte man sich bisher auf vereinfachende Abstraktionen beschränkt, die von biologischen Details absehen. Bereits neuromorphe Systeme berücksichtigen zunehmend neurobiologische Details von Synapsen und Neuronen. Um die Energieeffizienz und Leistungsfähigkeit lebender Gehirne zu erreichen, wird man zunehmend die Ergebnisse der Entwicklungsbiologie und Entwicklungspsychologie berücksichtigen. Dann wird man wohl auch der erstaunlichen Fähigkeit lebender Gehirne auf die Spur kommen, die keineswegs „Big Data“ zur Entwicklung kreativer Innovationen benötigt. Diese Gehirne sind häufig in der Innovationsgeschichte mit wenig Informationen ausgekommen, um mit bisher unverstandenen kreativen Fähigkeiten Innovationen zu entwickeln.

3.6 Hybride IT-Systeme mit Mensch und Computer

Im dynamischen Innovationsportfolio darf also der Mensch mit seinen kreativen Fähigkeiten nicht vergessen werden. Hybride IT-Systeme bezieht sich daher auch auf die Kopplung von Menschen mit Algorithmen und Menschen. Während Algorithmen mit äußerster Präzision gigantische Informationsmengen blitzartig verarbeiten können, sind Menschen in der Regel kreativer und zeichnen sich sowohl durch Intuition als auch Empathie im Umgang mit ihrer Umwelt aus. Entscheidungen sollten zwar letztlich noch beim Menschen liegen. Wenn aber ein erfahrener Pathologe bei der Beurteilung von Tumoren in Gewebeschnitten auf Tausende von Beispielen zurückgreifen kann, Lernalgorithmen im Deep Learning aber blitzschnell auf Millionen von Beispielen, scheint es für den Menschen eng zu werden.

Ein menschlicher Experte sollte in der Lage sein, nicht nur seine Diagnose anzugeben, sondern auch Erklärungen und Abwägungen zu begründen. Diese Erklärungskomponente muss auch von Algorithmen eingefordert werden (explainable AI). Im Beispiel der Medizin müssen solche Erklärungen am Ende gerichtsfest sein. Hier werden Grenzen für einen Algorithmus deutlich, wenn nicht juristisches Vor- und Hintergrundwissen gegeben ist. Daher sind Letztentscheidung und die damit verbundene Verantwortung für den Menschen in risikoreichen und brisanten Situationen wie in der Medizin vorgeschrieben.

Informationsaustausch zwischen Mensch und Maschine (Algorithmus) werden in hybriden Systemen gefordert. Menschen lernen dadurch von der Maschine und umgekehrt. So entstehen Rückkopplungen des Lernens zwischen zwei Seiten, die sich komplementär ergänzen. Am Beispiel der Medizin wird deutlich: In die Letztentscheidung z.B. einer Therapieanwendung bei Krebs sind nicht nur Fragen der genauen Tumorbestimmung und eventuelle Rückkopplungen mit anderen Krankheiten des Patienten zu berücksichtigen, sondern psychische und soziale, rechtliche und ökonomische Faktoren der Lebensqualität, die für einen Algorithmus (noch?) nicht realisierbar sind. Selbst wenn es möglich wäre, alle diese Faktoren zu erfassen und in einer Gesamtbeurteilung nach einem Algorithmus zu erfassen, müsste dieser Algorithmus selbst hinterfragt werden. Deshalb läge die Letztverantwortlichkeit in diesem Fall beim Arzt.

4 Ziel: Hybride IT-Systeme als dynamisches Forschungsportfolio

Ziel ist also ein dynamisches Innovationsportfolio der derzeitigen Analog, IT- und KI-Systeme, die hybrid miteinander verbunden sind. Die verschiedenen Brückentechnologien können sich ergänzen, gegenseitig verstärken oder ablösen. Am Ende sollten sie aber für uns Menschen und diesen Planeten eine nachhaltige Dienstleistung auch mit Blick auf Energieverbrauch und Umweltbelastung sein.

Empfehlenswert wäre, wenn es ein Nachhaltigkeitsiegel für IT-Systeme geben könnte, das z.B. in Form von DIN-Normen, wie sie aus der Technik bekannt sind, vergeben wird. Eine KI-Normungroadmap wurde für Deutschland bereits vorgeschlagen, in der alle technischen, ökonomischen, ökologischen, sozialen, rechtlichen und ethischen Nachhaltigkeitsfaktoren miteinfließen [WW20]. Dabei muss klar zum Ausdruck kommen, dass diese Art von Nachhaltigkeitsforderung keine Innovationsbremse darstellt, sondern innovationsfördernd wirkt, da damit Rechtssicherheit und Orientierungshilfen gegeben sind.

Literaturverzeichnis

- [Bi09] Biere, A.; Heule, M.; van Maaren, H.; Walsh, T. (Hrsg.): Handbook of Satisfiability, Amsterdam, IOS Press, 2009.
- [CH88] Coquand, T.; Huet, G.: The calculus of constructions. In: Information and Computation 76/2–3, 95–120, 1988.
- [CJ96] Coupet-Grimal, S.; Jakubiec, L.: Coq and Hardware Verification: A Case Study. TPHOLs, LCNS 1125, 125-139, 1996.
- [Eh17] Ehlers, R.: Formal verification of piece-wise linear feed-forward neural networks. In: arXiv:1705.01320v3 [cs.LO], 2017.
- [Hi21] Hiesinger, P.R.: The Self-Assembling Brain. How Neural Networks grow smarter. Princeton University Press, 2021.
- [HSW89] Hornik, K.; Stinchcombe, M.; White, H.: Multilayer feedforward networks are universal approximators neural networks. In: Neural Networks 2, 359-366.
- [Ma18] Mainzer, K.: The Digital and Real World. Computational Foundations of Mathematics, Science, Technology, and Philosophy. World Scientific Singapore, Kap. 10-12, 2018.
- [Ma19] Mainzer, K.: Artificial Intelligence – When do machines take over? Springer 2nd edition, 2019.
- [Ma20a] Mainzer, K.: Quantencomputer. Von der Quantenwelt zur Künstlichen Intelligenz. Springer, 2020.
- [Ma20b] Mainzer, K.: Leben als Maschinen. Brill Mentis: Paderborn. 2. Auflage, 2020.

- [Ma22] Mainzer, K.: Verifikation und Sicherheit für Neuronale Netze und Machine Learning. In: Mainzer (Hrsg.) Philosophisches Handbuch der Künstlichen Intelligenz. Springer, 2022.
- [Ma23] Mainzer, K.: Zukunft durch nachhaltige Innovation im Wettkampf der Systeme, Springer-Gabler, S. 359-366, 2023.
- [MK22] Mainzer, K.; Kahle, R.: Grenzen der KI – theoretisch, praktisch, ethisch, Springer, 2022.
- [MSS18] Mainzer, K.; Schuster, P.; Schwichtenberg, H.: Proof and Computation. From Proof Theory and Univalent Mathematics to Program Extraction and Verification. World Scientific Singapore, 2018.
- [Ne21] NeuroSys: BMBF-Zukunftscluster neuromorphe Hardware für autonome Systeme der Künstlichen Intelligenz. Projektpapier, S. 14-15, 2021.
- [Pe09] Pearl, J.: Causality - Models, Reasoning, and Inference. Cambridge (Mass.), 2009.
- [PT12] Pulina, L.; Taccella, A.: Challenging SMT solvers to verify neural networks. In: AI Communications 25 2, 117-135, 2012.
- [Pu88] Puppe, L.F.: Einführung in Expertensysteme, Springer, 1988.
- [TB03] Tretmans, J.; Brinksma, E.: TorX: Automated model-based testing. In: Hartman, A.; Dussa- Zieger, K. (Hrsg.): Proceedings of the First European Conference on Model-Driven Software Engineering, 2003.
- [UI10a] Ulmann, B.: Analogrechner: Wunderwerke der Technik - Grundlagen, Geschichte und Anwendung. Oldenbourg, 2010
- [UI10b] Ulmann B.: Analog and Hybrid Computer Programming. De Gruyter: Berlin, 2010.
- [Va98] Vapnik, V.N.: Statistical Learning Theory. Wiley, 1998.
- [WW20] Wahlster, W.; Winterhalter, C. (Hrsg.), Deutsche Normungsroadmap Künstliche Intelligenz, DIN + DKE, 2020.

Zukunft der KI: Verantwortung und Vertrauen

Reinhard Kahle¹

Abstract: Bei der Diskussion um die Zukunft der Künstlichen Intelligenz finden sich auch dystopische Szenarien, die soweit reichen, daß sich die Menschheit von einer künstlichen „Superintelligenz“ in ihrer Existenz bedroht fühlen müßte. In diesem Beitrag argumentieren wir dafür, daß der Mensch die Zukunft der KI in erster Linie im Hinblick auf Verantwortung und Vertrauen hin kontrollieren muß.

Keywords: Künstliche Intelligenz, Zukunft, Verantwortung, Dystopie

1 Einleitung

Die diesjährige Jahrestagung der Gesellschaft für Informatik steht unter dem Thema *Zukünfte*. Das ist natürlich eine sprachliche Fehlkonstruktion, da es letztlich nur eine Zukunft gibt, die vom Menschen erfahren wird. Doch es ist berechtigt zu fragen, in welcher Form wir diese eine Zukunft heute schon vorhersehen mögen, und in welcher Form wir Weichen stellen sollten, um diese Zukunft positiv zu beeinflussen. Es ist eine vorzügliche Eigenschaft des Menschen, hierzu alternative Entwicklungen gedanklich „durchrechnen“ zu können und so die Folgen des menschlichen Handelns zumindest in einem gewissen Rahmen abzuschätzen. Wir malen uns sozusagen verschiedene „Zukünfte“ aus.

In diesem Artikel geht es konkret darum, die Rolle der Künstlichen Intelligenz in diesen möglichen Zukünften zu antizipieren und aus den vorstellbaren Szenarien Schlußfolgerungen für unseren Umgang mit der neuen Technologie zu ziehen.

2 Dystopische Szenarien

Seit es die literarische Gattung des Zukunftsromans (Utopie; *Science-Fiction*) gibt, ist es ein Leitmotiv, technische Artefakte mit geistigen Eigenschaften des Menschen auszustatten, vor allem mit Intelligenz und/oder Gefühlen. Dabei werden insbesondere dystopische Szenarien heraufbeschworen, bei denen Maschinen sich die Menschheit unterwerfen oder zumindest zu unterwerfen drohen.

Der Künstlichen Intelligenz kommt hierbei neuerdings die Rolle zu, die Implementierung von Gefühlen und Intelligenz in Maschinen nicht mehr als reine Zukunftsmusik erscheinen

¹ Carl Friedrich von Weizsäcker-Zentrum, Universität Tübingen, Döblerstr. 33, 72074 Tübingen, Deutschland, und CMA, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal, reinhard.kahle@uni-tuebingen.de

zu lassen. Im Gegenteil, zum Teil wird das als explizite Option, ja sogar als das Ziel der KI angesehen. Ein aktuelles Argument sind die beeindruckenden Erfolge des KI-getriebenen, textbasierten Sprachkommunikationssystems ChatGPT. Die Möglichkeit dieses Systems, in Dialogsituationen Antworten zu geben, die sowohl sprachlich als auch inhaltlich als adäquat betrachtet werden, befeuert den Eindruck, Computersysteme könnten die menschliche Sprache effektiv beherrschen, einschließlich der Gefühle und Intelligenz. Auf die Frage, ob letzteres wirklich schon in der Reichweite der KI ist, werden wir unten noch eingehen (s. § 6).

Es bleibt aber die Frage, warum ein mögliches Erreichen des benannten Ziels eigentlich Anlaß zu dystopischen Szenarien geben sollte. Ganz konkret wird das so unter dem Schlagwort *Superintelligenz* propagiert [LS23]:

Superintelligence will be the most impactful technology humanity has ever invented, and could help us solve many of the world's most important problems. But the vast power of superintelligence could also be very dangerous, and could lead to the disempowerment of humanity or even human extinction.

Die Autoren spezifizieren Superintelligenz als „AI systems much smarter than humans“, und fragen sich, „[how to] prevent[. . .] it from going rogue“.

Zur Zeit werden die Systeme der künstlichen Intelligenz von Menschen „überwacht“ – eine Überwachung, die im *supervised learning* durch die Kontrolle der von der KI gegebenen Ergebnisse durch den Menschen erfolgt. Hier wird das konzeptionelle Problem gesehen: „humans won't be able to reliably supervise AI systems much smarter than us“.

Aber was sollte, egal wie schlau eine KI-Software[!] auch werden mag, uns Menschen daran hindern, einfach den Stecker zu ziehen? Die dystopische Gefahr wird wohl vor allem dadurch geschürt, daß wir *Menschen* fürchten, die – in vollem Bewußtsein der Bösartigkeit einer Superintelligenz – diese *nicht* stoppen würden. Diese Befürchtung ist möglicherweise berechtigt, beläßt aber die Verantwortung bei den Menschen (s. unten § 5).

3 Herausforderungen für die KI und für den Menschen

Vor dem Hintergrund der skizzierten dystopischen Szenarien ist es in erster Linie eine Herausforderung *für den Menschen*, bei den möglichen Entwicklungen von KI-Systemen – und bei ihrem Einsatz – die Weichen so zu stellen, daß wir auf die beste der (noch) möglichen Welten zusteuern. Was das konkret bedeuten kann, werden wir im folgenden Abschnitt diskutieren.

Für die KI lassen sich noch keine Herausforderungen in dem Sinne formulieren, daß es der KI selbst überlassen wäre, sich diesen Herausforderungen zu stellen. Das liegt technisch in erster Linie daran, daß sich die KI (noch?) nicht selbst kontrolliert – weder in seiner

Lernphase, noch in seiner Anwendung. Im Sinne der Abschätzung möglicher Zukünfte ist es aber doch eine berechnete Frage, was passieren würde, wenn sich die KI selbst kontrollieren könnte. Diese Frage ist aber insofern hinfällig, da sie erfordern würde, die konkreten Kontrollverfahren in Augenschein zu nehmen. Da diese noch nicht vorliegen, fällt doch dem Menschen die Aufgabe zu, diese Kontrollverfahren (wenn sie denn überhaupt implementiert werden sollten) in verantwortungsvoller Weise zu gestalten.

Wenn wir die Eigenkontrolle der KI noch als „Zukunftsmusik“ betrachten, so sind es aber die menschlichen Entwickler, Regulierer und Benutzer der Künstlichen Intelligenz, die sich schon heute für die unterschiedlichen Zukunftsszenarien wappnen müssen.

4 Umweltherausforderungen

Zur Zeit sollten es weniger die informationstechnologischen Endzeitphantasien sein, die der Menschheit Sorgen bereiten müssen, als vielmehr die drohende – oder schon einsetzende – Klimakatastrophe.

Daraus ergeben sich *Anwendungsfelder* für die KI, allein schon deshalb, weil unsere Klimamodelle mit vielen, sehr vielen Daten arbeiten (*Big Data*). Doch auch viele verantwortlich gemachten Ursachen für die menschliche (negative) Klimabeeinflussung beinhalten Komponenten, die einer Kontrolle mittels KI zugänglich sein sollten. Entsprechend wird im Bereich der *Nachhaltigkeit* verstärkt nach Anwendungsmöglichkeiten von KI gesucht – und es werden auch solche gefunden (siehe z.B. [Ka22]).

In welchem Maß sich hier spezifische konzeptionelle Herausforderungen für die KI stellen, wurde an anderer Stelle diskutiert [Ma21]. Den konkreten Anwendungen der KI im Nachhaltigkeitsbereich, die unsere Zukunft bestimmen können, stehen eigentlich keine besonderen konzeptionellen Herausforderung für die KI mehr im Weg – im Gegenteil: die informatische Technik liegt im Prinzip bereits vor, es müssen nur Anwendungsgebiete identifiziert werden, für die die passenden Daten vorliegen (oder gesammelt werden können), und Anwender müssen geschult werden, mit den Ergebnissen von KI-Programmen umgehen zu können. Es gibt aber eine andere zentrale konzeptionelle Fragestellung, die unseren Umgang mit KI in der Zukunft bestimmen dürfte.

5 Verantwortungsübertragung

Am 1. November 2011 veröffentlichten 62 Informatiker (vorrangig Professoren) einen offenen Brief in einer führenden deutschen Tageszeitung unter dem Titel: *Initiative für ein internationales Abkommen zu Autonomie in Waffensystemen* [Fe21]. Dabei handelte es sich um einen „Appell von ForscherInnen in den Gebieten Künstliche Intelligenz und Robotik mit der Forderung, in den Koalitionsvertrag der neuen Bundesregierung eine Passage aufzunehmen, die ein internationales Abkommen zu Autonomie in Waffensystemen

anstrebt“.² Die Unterzeichner brachten dabei ihre „tiefe Besorgnis über Waffensysteme zum Ausdruck, die Ziele ohne echte menschliche Kontrolle auswählen und angreifen“.

Es ist offensichtlich, daß es kein technisches Verfahren geben wird, das die Anwendung von Künstlicher Intelligenz in Militärtechnologie verhindern würde. Wer ein selbstfahrendes Auto bauen kann, wird auch einen selbstfahrende Panzer bauen können. Und eine Drohne, die durch KI-Steuerung selbständig einem Objekt ausweichen kann, kann auch genauso gut dieses Objekt gezielt treffen.

Nun tobt in Europa gerade ein Krieg, der im Rückblick möglicherweise einmal als der erste „Drohnenkrieg“ in die Geschichte eingehen könnte. Zwar wurden ferngesteuerte Drohnen auch schon vorher zur Bekämpfung von Terroristen eingesetzt, doch der Umfang und die Bedeutung von Drohneneinsätzen dürfte in dem akutell stattfindenden Krieg über alles Bekannte hinausgehen.

Als Hochtechnologiewaffe sind moderne Drohnen erst einmal ferngesteuerte Flugobjekte, die im Prinzip mit „konventioneller Technik“, d.h. ohne spezifischen KI-Einsatz, konzipiert werden können. Wenn nun KI zur Verbesserung von aerodynamischen Eigenschaften dieser Drohnen zum Einsatz kommen sollte, so fällt das für sich noch nicht unter die Anwendung von KI, gegen die sich die Unterzeichner des offenen Briefes gewandt hatten. Solange die *Entscheidung* darüber, welches Ziel zu treffen ist, letztlich von Menschen getroffen wird, handelt es sich lediglich um technische Verbesserungen von Waffen, denen nicht das Prädikat *autonom* anhaftet.

Die in dem offenen Brief geäußerten Bedenken entstehen erst, wenn „Ziele ohne echte menschliche Kontrolle aus[ge]wähl[t] und an[gegriffen]“ werden. Und hier liegt tatsächlich ein grundlegender Unterschied vor, der sich schon bei der Frage nach Verantwortlichkeit zeigt. Wird versehentlich ein falsches Ziel getroffen, so läßt ich, solange ein Mensch die Entscheidung zum Angriff dieses Ziels gegeben hat, diese Person zur Verantwortung ziehen. Doch wenn eine autonome Drohne eine falsche Entscheidung trifft, bleibt die Verantwortungsfrage offen. Man könnte versucht sein, die Verantwortung in diesem Fall auf diejenigen Personen, die die Drohne auf den Weg geschickt haben, oder diejenigen Personen, die den Einsatz der KI-Software genehmigt haben, oder diejenigen Personen, die die KI-Software entwickelt und/oder trainiert haben, zu projizieren. Doch es scheint ein

² Der offene Briefe richtete sich an die – seinerzeit – neue deutsche Bundesregierung. Tatsächlich findet sich im Koalitionsvertrag der folgende Passus [SBF21, S. 115]:

Bewaffnete Drohnen wollen wir verstärkt in internationale Kontrollregime einbeziehen. Letale Autonome Waffensysteme, die vollständig der Verfügung des Menschen entzogen sind, lehnen wir ab. Deren internationale Ächtung treiben wir aktiv voran.

Es gilt hier nicht zu bewerten, in welchem Maße damit dem Anliegen der Initiative entsprochen wurde; sie forderte als konkret formuliertes Ziel im Koalitionsvertrag explizit eine „führende Rolle [Deutschlands] bei der Entwicklung eines neuen, völkerrechtlich verbindlichen Vertrages zur Regulierung von Waffensystemen, die in ihren kritischen Funktionen autonom sind“.

intrinsisches Problem der KI zu sein, daß man sich zu keiner Verantwortungszuschreibung durchringen kann.

In der technischen Umsetzung von KI gibt es, aufgrund der *Black Box*-Arbeitsweise, insbesondere keine Möglichkeit, Entscheidungsgründe für KI-Ergebnisse auch nur nachzuvollziehen. Die Versuche, die *Black Box* im Rahmen einer *Explainable AI* zu „öffnen“ sind bisher nicht besonders erfolgversprechend; sollte es tatsächlich gelingen, die algorithmischen Zusammenhänge zwischen Input- und Output-Daten einer KI-Software sichtbar zu machen, könnte man die KI in den Rahmen der bekannten algorithmischen Informatik zurückversetzen (und z.B. traditionelle, formale Spezifikations- und Verifikationsmethoden auf sie anwenden). In dem folgenden Beitrag [Ma23] zeigt Klaus Mainzer, wie dieses Problem im Rahmen einer *hybriden KI* in der Zukunft angegangen werden sollte.

Die KI liefert auch immer nur statistische Ergebnisse, die eine „richtige Entscheidung“ *wahrscheinlich* machen, aber eben nicht sicherstellen. Dies ist einer der Faktoren, warum im Hinblick auf autonome Entscheidungen über das „richtige Ziel“ einer Drohne, die intendiert mit der Bedrohung menschlichen Lebens einhergeht, auf eine Ächtung autonomer Waffen gedrängt wird.

Ein weiterer Faktor kann durchaus auch darin gesehen werden, daß, selbst wenn sich eine (statistische) Unzuverlässigkeit bei autonom ausgewählten Ziele dieses Drohnentypes festgestellt werden sollte, verantwortungslose Militärs den Einsatz nicht stoppen würden, sondern *Kollateralschäden* billigend in Kauf nähmen. Damit nähert man sich wieder den dystopischen Szenarien an, wobei es allerdings nicht realistisch erscheint, daß diese Militärs die Kontrolle soweit an autonome Drohnen abgeben könnten, das diese sich unkontrollierbar auch gegen die eigenen Truppen wenden würden.

Die Quintessenz unserer Überlegungen hier ist, daß man sich bei der Weiterentwicklung von Künstlicher Intelligenz bewußt sein muß, an welchen Stellen Verantwortung in welcher Form verankert wird. Insbesondere ist die *juristische Verantwortung* heute noch immer (und sinnvollerweise) an einen Menschen gebunden und kann nicht naiv auf ein technisches Produkt abgewälzt werden. Entsprechend wird die Frage, wie wir mit der Verantwortung für (Fehlentscheidungen einer) KI umgehen, unsere Zukunft mitbestimmen.

6 Intelligenz?

Für einen verantwortungsvollen Umgang mit Künstlicher Intelligenz gibt es augenblicklich noch eine besondere Herausforderung. Der mediale Erfolg von ChatGPT ist dazu geeignet in der breiten Bevölkerung den Eindruck zu erwecken, die KI sei ihrem Ziel, menschliche Intelligenz zu simulieren, entscheidend näher gekommen. Und tatsächlich muß man der neuesten Generation von *Chatbots* beeindruckende Ergebnisse zugestehen. Zum Beispiel sind sie in der Lage, praktisch fehlerfrei deutsche Sätze zu bilden (Grammatik), die weitestgehend sinnvoll sind (Semantik) und auch auf die in Rede stehenden Themen Bezug nehmen (Kontext).

Allerdings sind auch gleich konzeptionelle Schwächen der Systeme sichtbar geworden, besonders diskutiert bei der „Erfindung“ von Literaturreferenzen. Zudem antworten diese Systeme praktisch durchgängig auf reflexive und bewertende Fragen – „Woher wissen Sie das?“, „Sind Sie sich sicher?“, „Halten Sie das für gut?“, etc. – enttarnend mit einer Ausrede, beginnend mit „Als KI-Modell habe ich . . .“. Spätestens hier sollte es einem dämmern, daß es mit der „Intelligenz“ dieser *Chatbots* noch nicht sehr gut bestellt ist. (Siehe auch [MK22, MK23].)

Ein eingängiges Beispiel ist der vermeintliche Erfolg von GPT-4 bei der Lösung von Problemen, die z.B. bei Zulassungstests für Hochschulen gestellt werden. Man fand schnell heraus, daß die Erfolge nicht über den Stichtag der verfügbaren Lerndaten hinaus anhielten [Mi23]³:

The same critics noted that for one of the coding benchmarks, GPT-4’s performance on problems published before 2021 was substantially better than on problems published after 2021—GPT-4’s training cutoff. This is a strong indication that the earlier problems were in GPT-4’s training data.

Damit erscheint der Vorwurf, bei den aktuellen *Chatbots* handle es sich konzeptionell eher um „stochastic parrots“ als um intelligente Kommunikationspartner, nicht völlig unbegründet.

Tatsächlich deutet ja auch das prinzipielle Verfahren der *Chatbots*, Sätze dadurch zu bilden, daß auf der Grundlage erlernter Beispieldaten (ungeheuren Umfangs), immer einfach nur eine Wahrscheinlichkeit für das nächste Wort bestimmt wird, darauf hin, daß die Ergebnisse – aufgrund der erlernten Beispiele und nur derer – eine gewisse Plausibilität haben. Wie gesagt, grammatikalisch erscheinen die Sätze sogar praktisch immer korrekt zu sein – doch für eine inhaltliche Korrektheit kann sich niemand verbürgen; und innovatives Denken kommt so sicherlich nicht zustande. Und es ist auch nicht zu sehen, warum diese erkannten Schwierigkeiten durch eine vergrößerte Lernbasis *konzeptionell* gelöst werden sollten.

Unter diesem Aspekt ist es eine Zukunftsaufgabe der KI, das naive Vertrauen auf die Korrektheit von Ergebnisse einer KI – gerade bei Personen, die mit der Materie nicht vertraut sind, einschließlich Entscheidungsträgern in Politik und Wirtschaft – als gefährliches Mißverständnis offen zu bekennen. Ideen, *Chatbots* in Schulen und Universitäten einzusetzen, kommen einem bildungspolitischen Selbstmord gleich, wenn den Beteiligten nicht klar ist, daß deren Antworten nur unter der Maßgabe von Plausibilitätsstandards gewonnen wurden, aber intrinsisch keinen Anspruch auf Korrektheit erheben. Um Vertrauen in KI setzen zu können, müssen den Anwendern die Grenzen dieser Technologie bewußt sein. Die Verfechter, Entwickler und Vermarkter der KI stehen also in der Verantwortung, den Anwendern bei der Deutung von KI-Ergebnissen zusätzliches Wissen über die Funktionsweise an die Hand zu

³ Der Artikel [Mi23] wurde am gleichen Tag veröffentlicht, an dem dieser Artikel seine Ausgestaltung bekam. Er enthält alle wesentlichen Kritikpunkte, die wir für den vorliegenden Absatz konzipiert hatten, und die deshalb jetzt nicht mehr alle explizit wiederholt werden müssen.

geben. Nur in diesem Sinne verantwortliche KI dürfte die Weichen in die richtige Richtung einer Zukunft stellen, die keine dystopischen Ängste mehr begründet.

Danksagung

Diese Arbeit wurde u.a. von der Udo Keller-Stiftung und die Portugiesische Forschungsgemeinschaft FCT über das *Centro de Matemática e Aplicações*, UID/MAT/00297/2020, gefördert.

Literaturverzeichnis

- [Fe21] Federrath, H. et.al.: Offener Brief: Initiative für ein internationales Abkommen zu Autonomie in Waffensystemen. <https://autonomewaffen.org/>, 2021. Stand: 13.7.2023.
- [Ka22] Kahle, R.: Künstliche Intelligenz und Nachhaltigkeit. Informatik Aktuell, 2022. <https://www.informatik-aktuell.de/betrieb/kuenstliche-intelligenz/kuenstliche-intelligenz-und-nachhaltigkeit.html>. Stand: 14.7.2023.
- [LS23] Leike, J.; Sutskever, I.: Introducing Superalignment. <https://openai.com/blog/introducing-superalignment>, 2023. Version vom 5.7.2023, nachgeschlagen am 13.7.2023.
- [Ma21] Mainzer, K.: Künstliche Intelligenz und Nachhaltigkeit. In: INFORMATIK 2021, S. 1153–1162. Gesellschaft für Informatik, Bonn, 2021.
- [Ma23] Mainzer, K.: Zukunft durch hybride IT- und KI-Innovationen. In: INFORMATIK 2023. Gesellschaft für Informatik, Bonn, 2023.
- [Mi23] Mitchell, M.: How do we know how smart AI systems are? *Science*, 381(6654), 2023. DOI: 10.1126/science.adj5957.
- [MK22] Mainzer, K.; Kahle, R.: Grenzen der KI – theoretisch, praktisch, ethisch. Springer, 2022.
- [MK23] Mainzer, K.; Kahle, R.: Limits of AI – theoretical, practical, ethical. Springer, 2023. Englische Übersetzung von [MK22] mit einem zusätzlichen Kapitel über ChatGPT.
- [SBF21] Sozialdemokratische Partei Deutschlands; Bündnis 90/Die Grünen; Freie Demokratische Partei, Hrsg.: Mehr Fortschritt wagen. 2021. Koalitionsvertrag 2021–2025 zwischen der Sozialdemokratischen Partei Deutschlands (SPD), Bündnis 90/Die Grünen und den Freien Demokraten (FDP).

Kritisches Denken -
Zukunftswerkstatt Informatik
Gesellschaft (IUG 2023)

Materialität und Immaterialität – Begriffsbestimmungen und Herausforderungen für die Informatik

Patrick Liedtke¹ und Hans-Knud Arndt²

Abstract: Informatik ist im Laufe der Zeit immer wichtiger geworden und nimmt eine große Rolle im täglichen Leben ein. Aufgrund dieser Entwicklung strebt die Informatik nach ständiger Verbesserung, unter anderem bei den Produkten der Informations- und Kommunikationstechnik, die den Bedürfnissen der Nutzer gerecht werden sollen. Daraus ergeben sich Herausforderungen insbesondere im Umgang mit der Informatik. Die Informatik muss sich unter anderem mit den materiellen und immateriellen Erzeugnissen sowie deren resultierenden Komplexität beschäftigen, die das Verständnis erschweren. Für die Informatiker sollte eine bessere Vermittlung dieser Themenbereiche erreicht werden, sodass sie bei der Erarbeitung neuer Ideen unterstützt werden. Aus diesem Grund wird die Materialität und die Immaterialität genauer betrachtet und in einem Verhältnis mit der Informatik und ihren Erzeugnissen gesetzt. Daraus resultierende Herausforderungen und Probleme können somit gezielt angegangen werden.

Keywords: Materialität, Immaterialität, Informatik, Herausforderungen

1 Einleitung

Die Informatik spielt in der heutigen Zeit eine große Rolle, begünstigt durch die fortschreitende Digitalisierung und der ständigen Entwicklung neuer Technologien. Dabei muss natürlich die Frage aufgeworfen werden, inwieweit sich das Leben dadurch verändert und wie der Umgang mit der Technik in der Gesellschaft erfolgt. Es müssen Anreize geschaffen werden, die Informatik begreifbarer zu machen. Im Zusammenhang dessen müssen die Herausforderungen im Umgang mit der Informatik definiert werden. Aus diesem Grund ist es umso wichtiger, sich ebenfalls mit der Materialität und der Immaterialität im Bereich der Informatik zu beschäftigen. Doch wie unterscheiden sich die beiden Begriffe voneinander und was sind deren Eigenschaften? Dieser Artikel wird zunächst die Begriffe der Materialität und der Immaterialität genauer betrachten. Danach wird eine kurze Einführung zum Begriff Informatik und ihren Hauptkomponenten Hardware und Software gegeben. Anhand dessen werden abschließend Probleme und Herausforderungen für die Informatik abgeleitet.

¹ Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, Arbeitsgruppe Wirtschaftsinformatik - Managementinformationssysteme -, Universitätsplatz 2, 39106 Magdeburg, patrick.liedtke@ovgu.de

² Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, Arbeitsgruppe Wirtschaftsinformatik - Managementinformationssysteme -, Universitätsplatz 2, 39106 Magdeburg, hans-knud.arndt@iti.cs.uni-magdeburg.de

2 Materialität

Beim Begriff Materialität gibt es verschiedene Bedeutungen, die eng mit dem Begriff des Materials zusammenhängen [MFO15], sodass in diesem Abschnitt beide Begriffe betrachtet werden. Als Materialität werden körperlich präsenste Objekte beschrieben, welche ebenfalls die Eigenschaft der Greifbarkeit besitzen [KST15]. Die Materialität beinhaltet unter anderem auch den Umgang mit den Materialien, also beispielweise das Verständnis dieser [Dö16]. Daran lässt sich das Verhältnis der beiden Begriffe ebenfalls erkennen.

Allerdings ist auch der Begriff des Materials in der Wissenschaft nicht genau geklärt [VR07]. Die Literatur beschreibt, dass dieses Wort, abgeleitet vom mittellateinischen Begriff „materiale“, in den deutschen Sprachgebrauch aufgenommen wurde [MFO15]. Abgeleitet wird dieser Begriff wiederum von „materia“, welcher ebenfalls den Ursprung im Lateinischen hat und den „Stoff, woraus etwas verfertigt wird“ [SB42] beschreibt. Auch Aristoteles beschäftigte sich bereits mit Materie, wobei er darin das sieht, aus dem etwas Neues geschaffen werden kann [He15; Bu22]. Somit lässt sich erkennen, dass „materia“ und das Material gewissermaßen die Grundlage für neue Produkte oder etwas Neuartiges darstellen [MFO15]. Dies spiegelt sich auch in der Abstammung des Begriffs „materia“ vom Begriff „mater“ wider, welches die Bedeutung als „Mutter, Ursprung, Quelle“ [WSL13] innehat.

3 Immaterialität

Wie sich schon vermuten lässt, stellt die Immaterialität das Gegenteil von der Materialität dar und verweist somit auf körperlich nicht präsenste Objekte [KST15]. Des Weiteren stellt es „das vom Betrachter im Moment der Begegnung Erlebte“ [Sc13] dar. Im Dienstleistungsbereich wird die Immaterialität als eine Eigenschaft angesehen, welche „sowohl für die physische Unantastbarkeit von Dienstleistung, als auch für die fehlende psychische Erfassbarkeit von Dienstleistungen“ [BGZ13] steht. Die Entwicklung der Technologien sorgt dafür, dass beispielweise manuelle Arbeitsvorgänge automatisiert werden können, sodass sich mit dem Thema der Immaterialität beschäftigt werden muss.

Allerdings existieren unterschiedliche Meinungen in Bezug auf die Immaterialität. Nach Stiegler existiert das Immaterielle nicht, er spricht vielmehr von einer Hypermaterie [St10]. Seiner Meinung nach sind es „flüchtige Zustände der Materie“ [St10]. Darüber hinaus sagt er, dass Hardware, welche materiell ist, verwendet wird, um die flüchtigen Zustände zu produzieren [St10]. Als Problem sieht er die Unsichtbarkeit der Materie an und nicht die Immaterialität [St10]. Unabhängig von der beschriebenen Sichtweise, soll in dieser Arbeit trotzdem der Begriff der Immaterialität weiter betrachtet werden, auch wenn es sich dabei um einen Zustand der Materie handelt.

4 Informatik und ihre Hauptkomponenten Software und Hardware

Die Informatik kann als Wissenschaft angesehen werden, die dazu dient, Neues zu schaffen [K114]. Informatik wird ebenfalls als „eine anwendbare und nach Anwendungen verlangende Wissenschaft“ [Ba09] beschrieben. Nach dem Gabler Wirtschaftslexikon bezeichnet die Informatik die „Wissenschaft von der systematischen Verarbeitung von Informationen, insbes. der automatischen Verarbeitung mithilfe von Computern“ [HA00]. Dieser Aspekt verdeutlicht die Relevanz der Informatik in der Gesellschaft, da mithilfe der Informatik Prozesse automatisiert beziehungsweise unterstützt werden können.

In der Informatik werden oft ihre Hauptkomponenten Hardware und Software betrachtet. Als Hardware werden die Computerteile angesehen, die prinzipiell anfassbar sind, während Software, beispielsweise Programme und Daten, nicht anfassbar sind [PP14]. Aus den Definitionen lässt sich bereits erkennen, dass es sich bei der Hardware um materielle und bei der Software um immaterielle Bestandteile handelt. Eine Software kann ihre Funktionen nur mit Hardware ausführen, sodass Software und Hardware als „Komplementär-güter“ [Hi19] angesehen werden. Daraus kann geschlussfolgert werden, dass keine klare Grenze zwischen Software und Hardware gezogen werden kann, sondern dass dieser Übergang fließend ist.

5 Materialität und Immaterialität in der Informatik - Herausforderungen

Ein Architekt muss sich mit den verfügbaren Materialien auseinandersetzen und verstehen, wie diese Materialien eingesetzt werden können. Der Architekt wird als „Vermittler (...) zwischen Ideen und Materialien“ angesehen [Vi11]. Dasselbe sollte auch für einen Informatiker gelten, der sich ebenfalls mit seinen Materialien auseinandersetzen muss. Sowohl der Architekt als auch der Informatiker haben eine gesellschaftliche Verantwortung, indem sie die Zukunft mitgestalten und sich neuen Aufgaben annehmen. Durch das Erlernen der Materialien sowie der individuellen Weiterentwicklung kann aus den Architekten und den Informatiker ein guter Architekt beziehungsweise ein guter Informatiker werden. Der Architekt muss mit materiellen Stoffen, wie beispielweise Holz und Glas, umgehen können. Der Informatiker hingegen ist umgeben von immateriellen Objekten, die das Erlernen erschweren.

Das Erlernen der Eigenschaften von materiellen Stoffen erfolgt bereits in den jungen Jahren, sodass ein Bewusstsein darüber entsteht, was mit bestimmten Materialien gemacht werden kann und was nicht. Für den Architekten entsteht durch dieses Bewusstsein dann nochmal ein anderer Zugang zur Materialität. Der Architekt hat bereits das Vorwissen zu den Materialien, welches für die Ausübung seiner Tätigkeiten neu durchdacht werden muss. Dabei lässt sich eine Aufteilung der Materialien anhand ihrer Eigenschaften vornehmen, sodass der Einsatzbereich besser definiert werden kann. Eine solche Aufteilung gestaltet sich in der Immaterialität schwierig und stellt somit eine Herausforderung dar.

Um den Zugang für den Informatiker zur Immaterialität zu finden, könnte ein erster Schritt die bessere Vermittlung der Immaterialität darstellen. Das oberste Ziel bei der Vermittlung der Immaterialität, aber auch in der allgemeinen Ausbildung, ist, zu verstehen, was alles gemacht werden kann. Der Informatiker soll somit unterstützt werden, die richtigen Entscheidungen hinsichtlich aktueller Entwicklungen und dem gesellschaftlichen Bedarf zu treffen und neue Ideen zu entwickeln. Es ist wichtig, neue Ideen zu entwickeln, um die Qualität der Produkte der Informatik, unter anderem in der Informations- und Kommunikationstechnik, zu verbessern. Im folgenden Verlauf wird näher beschrieben, warum die Vermittlung der Immaterialität sich als schwierig erweist und somit eine große Herausforderung für die Informatik darstellt.

Wie bereits beschrieben, beschäftigt sich die Informatik mit der „Verarbeitung von Informationen“ [HA00]. Wenn wir nun auf die Definition der Software zurückgehen und diese mit den Definitionen zur Materialität und Immaterialität verbinden, können wir erkennen, dass Software als Produkt der Informatik immateriell ist. Gleiches gilt allerdings auch für die Informationen, die mithilfe einer Software verarbeitet werden sollen [We05]. Die Schlussfolgerung aus der Immaterialität der Software ist laut Weicker [We05], dass die Informatik an der Benutzerschnittstelle gemessen wird. Wie sich bereits in den Definitionen zur Software und Hardware zeigt, existiert eine Abhängigkeit zwischen diesen beiden Produkten. Die Immaterialität (in dem Fall die Software) kann nur mithilfe von Materialität (in dem Fall die Hardware) veranschaulicht werden. Die Veranschaulichung mittels der Benutzeroberfläche wird in der Literatur als wahrnehmbare Materialität bezeichnet. Diese steht im Gegensatz zu der technischen Materialität, welche die binären Daten darstellt [BMW22]. Durch die Wahrnehmung der Benutzerschnittstelle ist dem Anwender die Komplexität der Software nicht bewusst [We05]. Allerdings muss der Anwender für das Benutzen der Software kein vollständiges Verständnis für die Komplexität entwickeln, sondern nur wissen, welche Funktionen mit dieser Software ausführbar sind. Der Anwender erlernt in dem Fall auch, welche Software sich für welchen Einsatzbereich eignet.

Um allerdings bestimmte Aktionen auszuführen, müssen die Anwender in Interaktion mit den digitalen Geräten treten. Eine Interaktion kann das Drücken eines Knopfes darstellen, wobei in einem Praxisbeispiel aus der Literatur dazu beschrieben wird, dass der Nutzer nicht weiß, welche Prozesse durch den Knopfdruck in digitalen Geräten ausgelöst werden [Un23]. Durch die fortschreitende Entwicklung wird dieses Problem verschärft, wie sich an diesem Beispiel erkennen lässt. Oftmals werden die physisch anfassbaren Knöpfe bereits durch Aktionen wie „(s)treichen, streicheln, tippen, ansprechen, rufen, berühren“ [Un23] ersetzt. Das Problem, welches dabei existieren kann, ist, dass die Möglichkeiten zur Interaktion überlagert werden und der Nutzer somit versehentlich Aktionen ausführt, die nicht ausgeführt werden sollen. Wenn unerwünschte Funktionen ausgeführt werden, dann kann dies dazu führen, dass die Anwender unzufrieden werden und sich zukünftig für andere Produkte entscheiden. Auch hier muss der Nutzer nicht alle im Hintergrund ablaufenden Prozesse verstehen, aber es soll verstanden werden, welche Möglichkeiten sich zur Interaktion bieten. Das bedeutet, dass für den Nutzer eine verständliche und benutzerfreundliche Schnittstelle geschaffen werden muss. Um das zu erreichen, müssen wir zunächst auf den Informatiker zurückgehen, der neue Ideen aufgrund des Verständnisses

bezüglich der Immaterialität entwickeln kann. Bei der Umsetzung von neuen Ideen sollten auch immer die potentiellen Anwender einbezogen werden, um mögliche Probleme im Vorfeld zu minimieren und Vorstellungen seitens der Anwender einzubeziehen. Der technologische Wandel kann somit gemeinsam mitgestaltet werden.

6 Fazit und Ausblick

Aus den beschriebenen Erkenntnissen ergeben sich Herausforderungen, die den Umgang mit der Informatik betreffen. Dabei müssen in Zukunft Methoden und Ansätze zur Vermittlung entworfen werden, die sich primär an die Gestalter in der Informatik richten. Es ist wichtig, sich dieser Herausforderung anzunehmen, da Verbesserungen zur Zufriedenheit der Anwender beitragen und somit einen Vorteil gegenüber Mitwettbewerbern schaffen. Daher ist es wichtig, bei der Umsetzung von neuen Ideen die Anwender miteinzubeziehen, sodass die Produkte nutzerfreundlich gestaltet werden können und auch ein gemeinsames Verständnis zum Umgang dieser geschaffen wird. Das gemeinsame Verständnis kann dazu dienen, dass der Anwender leichter Erkenntnisse zum Einsatz dieser Produkte für die täglichen Aufgaben sammeln kann. Für die Anwender wird dadurch ein anderer Zugang für die Informatik geschaffen. Wenn diese Herausforderungen in Zukunft bewältigt werden können, dann könnte die Disziplin der Informatik in der Gesellschaft attraktiver und begreifbarer dargestellt werden. Der Nutzen, der daraus entstehen kann, ist, dass viel mehr Personen sich der Informatik widmen und neugierig sind, diesen Bereich mitzugestalten.

Literaturverzeichnis

- [Ba09] Bauer, F. L.: Informatik – Geburt einer Wissenschaft. In: (Bauer, F. L., Hrsg.) Historische Notizen zur Informatik. Springer-Verlag, Berlin u.a., S. 21-29, 2009.
- [BMW22] Bender, M.; Mell, R. M.; Wildfeuer, J.: Zur Spezifik digitaler Medien als Diskursraum: Materialität, Daten, Affordanzen. In: (Gredel, E., Hrsg.) Diskurse – digital. Theorien, Methoden, Anwendungen. De Gruyter, Berlin u.a., S. 27-45, 2022.
- [BGZ13] Biege, S.; Gotsch, M.; Zanker, C.: Anforderungen von Produktivitätsmessungen bei innovativen und wissensintensiven Unternehmensdienstleistungen. In: (Thomas, O.; Nüttgens, M., Hrsg.) Dienstleistungsmodellierung 2012, Product-Service Systems und Produktivität. Springer Fachmedien, Wiesbaden, S. 142-168, 2013.
- [Bu22] Buhr, L.: Materialität. In: Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen. Veröffentlicht am 07.04.2022, <https://diskursmonitor.de/glossar/materialitaet>, Stand: 02.05.2023.
- [Dö16] Döring, T.: A Materials Perspective on Human-Computer Interaction. Case Studies on Tangible, Gestural, and Ephemeral User Interfaces. Dissertation, Bremen, 2016.

- [HA00] Hadelner, T.; Arentzen, U.: Gabler Wirtschaftslexikon, E – J. 15., vollständig überarbeitete und aktualisierte Auflage, Gabler Verlag, Wiesbaden, 2000.
- [He15] Heuer, P.: Aristoteles' Konzept der Materie. *Rivista di Filosofia Neo-Scolastica* 4/15, S. 827-844, 2015.
- [Hi19] Hilty, L.: Software und Nachhaltigkeit – Von der informationellen zur materiellen Selbstbestimmung. In: (Göpel, M.; Leitschuh, H.; Brunnengräber, A.; Ibisch, P.; Loske, R.; Müller, M.; Sommer, J.; von Weizsäcker, E. U., Hrsg.) *Die Ökologie der digitalen Gesellschaft*. Hirzel Verlag, Stuttgart, S. 1-14, 2019.
- [K114] Kleiner, P.: Was ist Informatik?. Schriftenreihe, Hasler Stiftung, Bern, 2014.
- [KST15] Karagianni, A.; Schwindt, J. P.; Tsouparopoulou, C.: Materialität. In: (Meier, T.; Ott, M. R.; Sauer, R., Hrsg.) *Materiale Textkulturen, Konzepte – Materialien – Praktiken*. De Gruyter, Berlin u.a., S. 33-46, 2015.
- [MFO15] Meier, T.; Focken, F.-E.; Ott, M. R.: Material. In: (Meier, T.; Ott, M. R.; Sauer, R., Hrsg.) *Materiale Textkulturen, Konzepte – Materialien – Praktiken*. De Gruyter, Berlin u.a., S. 19-31, 2015.
- [PP14] Platzner, M.; Plessl, C.: Verschiebungen an der Grenze zwischen Hardware und Software. In: (Eke, N. O.; Foit, L.; Kaerlein, T.; Künsemöller, J., Hrsg.) *Logiken strukturbildender Prozesse: Automatismen*. Wilhelm Fink, Paderborn, S. 123-144, 2014.
- [SB42] Schulz, H.; Basler, O.: *Deutsches Fremdwörterbuch, Zweiter Band L - P*. De Gruyter, Berlin, 1942.
- [Sc13] Schmitt, U.: Der Doppelaspekt von Materialität und Immaterialität in den Werken der ZERO-Künstler 1957-67. Dissertation, Köln, 2013.
- [St10] Stiegler, B.: *Hypermaterialität und Psychomacht*. Diaphanes, Zürich, 2010.
- [Un23] Unterberg, L.: You Press the Button, We Do the Rest. Bildung und Knöpfe. In: (Leineweber, C.; Waldmann, M.; Wunder, M., Hrsg.) *Materialität – Digitalisierung - Bildung*. Verlag Julius Klinkhardt, Bad Heilbrunn, S. 30-43, 2023.
- [Vi11] Viray, E.: Warum Material Design? In: (Schröpfer, T., Hrsg.) *Material Design, Materialität in der Architektur*. Birkhäuser, Basel, S. 7-8, 2011.
- [VR07] Vallgård, A.; Redström, J.: Computational composites. In (Begole, B.; Payne, S.; Churchill, E.; St. Amant, R.; Gilmore, D.; Rosson, M. B., Hrsg.): *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, S. 513-522, 2007.
- [We05] Weicker, N.: Informatik – didaktische Weiterbildung von Lehrenden. In: (Friedrich, S., Hrsg.) *Unterrichtskonzepte für informatische Bildung*. Gesellschaft für Informatik e.V., Bonn, S. 101-110, 2005.
- [WSL13] Welsch, N.; Schwab, J.; Liebmann, C. C.: *Materie. Erde, Wasser, Luft und Feuer*. Springer Spektrum, Berlin u.a., 2013.

Aufnahme läuft! Partizipative, sozio-technische Gestaltung digitaler Kommunikationssysteme

Andrea Kienle ¹, Gabriele Kunau ²

Abstract: Der Beitrag beschäftigt sich mit einem Vorgehen zur partizipativen, sozio-technischen Gestaltung digitaler Kommunikationssysteme. Dieses Vorgehen wurde für ein konkretes Beispiel der Adoption eines Videokonferenztools in unserer Hochschule entwickelt und erprobt. Aus diesem Beispiel werden anschließend Best Practices für die partizipative, sozio-technische Systemgestaltung in Hochschulen abgeleitet.


Keywords: Partizipation, sozio-technische Systemgestaltung, organisatorische Wahlfreiheit, Kommunikationsunterstützung

1 Einleitung: Digitalisierung in Hochschulen

In den letzten Jahren hat sich digitale Kommunikation zur Unterstützung gemeinsamer Lern- und Arbeitsprozesse immer weiter verbreitet. Diese Entwicklung hat auch vor der Einführung in öffentlichen Einrichtungen wie Schulen, Hochschulen und Behörden keinen Halt gemacht. Einen ausführlichen Blick auf die Digitalisierung in Hochschulen gibt [HG23]. Charakteristisch für Hochschulen sind insbesondere eine große Heterogenität der Hochschulmitglieder sowie eine erhöhte Autonomie einiger Statusgruppen. Dabei wird immer wieder auch auf die Verschiedenheit der Verwaltungseinheiten und der Fachbereiche, die für Forschung und Lehre stehen, hingewiesen [Gr21].

Vielfach werden Standardwerkzeuge eingesetzt, deren Funktionalität sehr umfangreich ist. Hier stehen die Einrichtungen vor der Herausforderung, eine Auswahl der bereitgestellten Funktionalitäten so zu treffen, dass sie zu den Prozessen und Eigenheiten der jeweiligen Organisation passen.

In diesem Beitrag möchten wir nach einem kurzen Überblick über verwandte Arbeiten zunächst einen konkreten Fall der sozio-technischen Systemgestaltung unter Zuhilfenahme partizipativer Elemente vorstellen. Als Videokonferenztool wurde in unserer Hochschule Cisco Webex³ eingeführt. Zunächst wurden nur die Funktionen zur

¹ Fachhochschule Dortmund, Fachbereich Informatik, Emil-Figge Str. 42, 44227 Dortmund, andrea.kienle@fh-dortmund.de,  <https://orcid.org/0000-0002-3236-3418>

² Fachhochschule Dortmund, Fachbereich Informatik, Emil-Figge Str. 42, 44227 Dortmund, gabriele.kunau@fh-dortmund.de

³ www.webex.com Stand 14.07.2023

Unterstützung der synchronen Kommunikation zur Verfügung gestellt. Die Funktionen zur Aufzeichnung und Speicherung von Sitzungen wurden generell deaktiviert, weil dies mit Hinweis auf mögliche Überwachung und Kontrolle der Arbeitsleistungen problematisch erschien. Wie in Prozessen der Technikaneignung typisch, ergaben sich im Laufe der Zeit immer neue Anwendungsfälle, für die die Aufzeichnung von Sitzungen eine sehr sinnvolle Unterstützung darstellt. In einem partizipativen Format haben wir einen Empfehlungskatalog sozialer Regeln definiert, der für bestimmte Szenarien die Aufzeichnung ermöglicht und sie für andere explizit ausschließt. Aus diesem Beispiel werden wir anschließend Best Practices für die partizipative, sozio-technische Systemgestaltung in Hochschulen ableiten.

2 Verwandte Arbeiten: Sozio-technische Systemgestaltung

Methoden zur Gestaltung sozio-technischer Systeme beziehen immer die künftigen Nutzer*innen der Systeme aktiv ein. Seit den 1980er Jahren gibt es innerhalb der Informatik die Forschungsrichtung participatory design [KB98], die sich mit verschiedenen Optionen der Beteiligung auseinandersetzt. Ein wichtiges Ergebnis der frühen Forschungen zu sozio-technischen Systemen (zusammenfassend in [Sy85]) ist das Konzept der organisatorischen Wahlfreiheit: Es gibt immer organisatorische Optionen bei der Techniknutzung, die es zu gestalten gilt. Dementsprechend wurden im Laufe der Zeit Vorgehensmodelle wie STEPS [Fl89] oder STWT [He02] entwickelt, die die partizipative Gestaltung der Nutzungsprozesse von IT-Systemen zum Gegenstand haben.

Auch im Hochschulkontext hat sich die Partizipation der von Veränderung Betroffenen im Sinne des Prinzips „Betroffene zu Beteiligten machen“ [vL22, S. 20] etabliert. Speziell in der Hochschuldigitalisierung, die eher als Kommunikations- denn als Bereitstellungsproblem betrachtet werden muss, sind partizipative Prozesse notwendig, in denen dann „kulturelle Unterschiede – zentral dabei jene zwischen Wissenschaft und Verwaltung“ aufeinander treffen [Pa20, S. 66].

Beteiligt man zukünftige Nutzer*innen, so kann diese Beteiligung hinsichtlich ihres Grades sehr unterschiedlich ausfallen. Während eine Information an die oder eine Datenerhebung bei den Nutzer*innen eher eine kommunikative Einbahnstraße von den Nutzer*innen zu den Projektverantwortlichen darstellt, sind die Beratung oder Mitbestimmung dialogorientiert. Bezüglich der Form der Partizipation lassen sich die direkte und die repräsentative Beteiligung unterscheiden [KK14]. Das hier präsentierte Fallbeispiel stellt eine dialogorientierte Beratung und Mitbestimmung in Form einer repräsentativen Beteiligung dar. Inhaltlich geht es um die Erarbeitung von Regeln zur rechtmäßigen und sinnvollen Nutzung der Aufzeichnungsfunktion in Webex. Damit handelt es sich um die Ausgestaltung der organisatorischen Wahlfreiheit im Sinne der sozio-technischen Systemgestaltung.

3 Fallbeispiel: Sozio-technische Gestaltung von Aufnahmen in Webex

In unserer Hochschule wurde in einem ebenfalls partizipativen Prozess der Hochschulentwicklungsplan (HEP) 2025⁴ neu gefasst. Zur Digitalisierung als eines von vier Fokusthemen tauschen sich auf freiwilliger Basis Interessierte aus den Statusgruppen der Professorenschaft sowie der wissenschaftlich und nicht-wissenschaftlichen Beschäftigten in regelmäßig stattfindenden Fokusgruppen über aktuelle Entwicklungen in der Hochschule aus. Hier kam die Anforderung auf, dass Videokonferenzen per Webex aufgezeichnet und anschließend einem ausgewählten Kreis zur Verfügung gestellt werden sollten. Der Organisation und dem partizipativen Vorgehen im HEP Prozess entsprechend wurde eine Experimentiergruppe „Aufnahmen in Webex“ gegründet, für die es gelang, Repräsentanten der verschiedenen Zielgruppen der Professorenschaft sowie der wissenschaftlichen und nicht-wissenschaftlichen Beschäftigten als Fachseite zur Mitarbeit zu motivieren. Darüber hinaus war das IT-Team der Hochschule, Mitglieder des Justitiariats sowie der Personalräte vertreten. Die Experimentiergruppe nähert sich in einem iterativen Prozess dem Thema und vollzog im Wesentlichen diese Schritte:

- Erarbeitung von Anwendungsfällen: Im ersten Schritt wurden die Anforderungen an die Verwendung der Aufnahmefunktionalität konkretisiert, indem Anwendungsfälle gesammelt wurden. Durch die Konzentration auf konkrete Fälle konnte die Diskussion zwischen Befürworter*innen einer allgemeinen Freigabe der Aufnahmefunktionalität und Vertreter*innen einer generellen Deaktivierung sachlich moderiert werden. Beispiele, bei denen sich die Gruppe mehrheitlich für eine Ermöglichung der Aufnahme ausgesprochen hat, sind: Experteninterviews im Rahmen von Forschungsprojekten, Vorträge im Rahmen von Ringvorlesungen, Vorträge im Rahmen von Schulungsveranstaltungen der Personalentwicklung. Dagegen sind Beispiele, für die eine Aktivierung der Aufnahme prinzipiell abgelehnt wurde: Interaktive Lehrveranstaltungen (Vorlesungen und Seminare) in regulären Studiengängen, mündliche Prüfungen, Bewerbungsgespräche.
- Erstellung eines Empfehlungskataloges: Für jeden Anwendungsfall wurde weiter konkretisiert, welche Zielgruppen beteiligt sind und welche Handlungsempfehlungen bei der Nutzung gegeben werden. Ein Beispiel ist das Einholen einer Einverständniserklärung von Vortragenden in einer Ringvorlesung. Für die Anwendungsfälle, bei denen die Aufnahmefunktionalität abgelehnt wurde, wurden die Argumente in einer Begründung zusammengefasst. So sind Aufzeichnungen von regulären Lehrveranstaltungen, in denen Studierende identifizierbar sind, nicht zulässig. Der gewünschte interaktive Charakter verbietet eine Anonymisierung der Teilnehmenden. Um sie persönlich ansprechen zu können, sind die Nennung des Namens oder das Zeigen von Gesichtern erforderlich.

⁴ <https://www.fh-dortmund.de/hochschule/profil/hochschulentwicklung.php> Stand 14.07.2023

Da der Besuch einer Lehrveranstaltung im Rahmen eines Studiums datenschutzrechtlich als Zwangssituation zu werten ist, hülfe auch eine Einwilligung der Teilnehmer*innen in die Aufzeichnung nicht. Im Sinne einer konstruktiven Gestaltung wurden auch Alternativen wie bspw. die asynchrone Bereitstellung vorproduzierter Vorlesungen in den Katalog aufgenommen. In diesem Schritt waren die Einschätzungen der Juristen und Personalvertretungen besonders wertvoll für die spätere praktische Umsetzung der sozialen Regeln.

- **Experimentierstatus:** Während der Sammlung und Beurteilung der Anwendungsfälle kam mehr und mehr die Unsicherheit auf, wie die Aufnahmefunktionalität zu nutzen wäre. Aus diesem Grunde wurde den Mitgliedern der Gruppe in einem Experimentierstatus die Verwendung der Funktionalität ermöglicht. Besonders lehrreich war die gemeinsame Nutzung der Aufzeichnung in den Gruppensitzungen. So konnten die Mitglieder mit den Einstellungen sowie der Verwendung der entstandenen Dateien experimentieren und diskutieren. Dabei gingen die technische Umsetzung (Zugriff auf die Funktionalität) und die soziale Regelung (Beurteilung der Beteiligten) Hand in Hand. Mit diesem Schritt wurde erreicht, dass die Mitglieder die Nutzung der Funktionalität besser einschätzen konnten. Zudem wurden weitere Themen für den Empfehlungskatalog deutlich.
- **Weitere Spezifizierung des Empfehlungskatalogs:** Als Resultat der Experimentierphase wurde der Empfehlungskatalog um die Themen Ort, Dauer und Zugriffsmöglichkeiten der Aufnahmen pro Anwendungsfall weiter spezifiziert

Als Ergebnis lag schließlich ein differenzierter Empfehlungskatalog entlang der Anwendungsfälle vor, der den Personalräten zugeschrieben und dem in der Form zugestimmt wurde (vgl. Abbildung 1). Nach erfolgter Zustimmung durch die Personalräte wurde die Bereitstellung der Funktionalität einerseits in einer breiten Kommunikation an die Multiplikator*innen in der Hochschule wie zum Beispiel Senator*innen, Fachbereichs- und Dezernatsleitungen sowie die Mitglieder der HEP Fokusgruppe bekanntgegeben.

Experimentiergruppe "Aufnahmen in Webex"			
lfd. Nr.	Anwendungsfälle	Zielgruppe	Beurteilung der Aufnahmefunktionalität
3	Schulungs-Veranstaltungen (PEM, H-IT)	Lehrende & Beschäftigte	ja
5	Vorträge in Ringvorlesungen	Studierende & Lehrende	ja
6	Vorträge im Rahmen von Veranstaltungen (Konferenzen, internationale Woche, ...)	Alle	ja
7	Experteninterviews in Forschungsprojekten	Forschende & Lehrende	ja
4	Schulungs-Veranstaltungen (Career Service...)	Studierende	unter bestimmten Umständen
8	Großveranstaltungen (bspw. Akademische Jahresfeier)	Alle	unter bestimmten Umständen
9	Lerngruppen	Studierende	unter bestimmten Umständen
1	Interaktive Lehrveranstaltung: Vorlesung	Lehrende	nein
2	Interaktive Lehrveranstaltung: Seminare	Lehrende	nein
10	Bewerbungsgespräche	Beschäftigte	nein
11	Mitarbeiterjahresgespräche	Beschäftigte	nein

Abb. 1: Überblick über Anwendungsfälle, Zielgruppe und Beurteilung (Ausschnitt)

Andererseits wird auf die Möglichkeit der Aufzeichnung im Ausnahmefall beim Start einer jeden Webex Sitzung hingewiesen (vgl. Abbildung 2). Dabei wurden die Anwendungsfälle als Ausnahme von der Regel, dass prinzipiell nicht aufgezeichnet wird, verdeutlicht. Zudem werden explizit auch Teile der Handlungsempfehlungen wie beispielsweise Information und Benennung des Zwecks der Aufzeichnung thematisiert.

Ausschluss

AKTIVE TEILNAHME
 Online-Lehre lebt wie Präsenzlehre von einer aktiven Teilnahme. Bitte tragen Sie sich daher mit Vor- und Nachnamen ein und schalten Sie Kamera an, soweit die/der Dozent*In darum bittet. Nur so können die Dozent*Innen die Teilnehmer*Innen einbeziehen. Die Dozent*Innen als Gastgeber moderieren Diskussionen und können das Wort erteilen, aber bei Störungen auch durch Stummschaltung entziehen oder Teilnehmer*Innen notfalls ausschließen.

AUFZEICHNUNG
 Veranstaltung der Fachhochschule Dortmund, insbesondere Vorlesungen, werden grundsätzlich nicht aufgezeichnet. Lediglich in Ausnahmefällen, bei denen die Teilnahme freiwillig ist, können die Veranstalter*Innen eine Veranstaltung aufzeichnen. Zwecke der Aufzeichnung werden die Veranstalter*Innen den Teilnehmer*Innen darstellen und auf die Aufzeichnung hinweisen. Eine Aufzeichnung der Veranstaltungen durch die Teilnehmer*Innen ist in jeglicher Form untersagt und stellt eine Rechtsverletzung dar. Soweit die Studierenden eine Aufzeichnung tätigen, sind sie im Zweifel persönlich verantwortlich - auch bei Schadensersatzforderungen Dritter.

Ablehnen
Annehmen

Abb. 2: Regelung der Aufnahmen als Information in Webex

4 Best Practices für sozio-technische Systemgestaltung in Hochschulen

Das berichtete Fallbeispiel hat Best Practices zur partizipativen Gestaltung der organisatorischen Wahlfreiheit in sozio-technischen Systemen an einer Hochschule hervorgebracht, die sich auf andere Hochschulen übertragen lassen (vgl. Abbildung 3).

- Versachlichung der Diskussion durch Fokussierung auf konkrete Anwendungsfälle
- Beteiligung aller Zielgruppen für die Spezifizierung von Anwendungsfällen
- Beteiligung der Fach- und IT-Seite zur Förderung gegenseitigen Verständnisses
- Einbindung des Justiziariats zur rechtssicheren Umsetzung einer sozio-technischen Regelung
- Prozessbegleitende Mitbestimmung zur Beschleunigung der Zustimmung und anschließenden Umsetzung
- Experimentierstatus für die Erlernung der Nutzung und Identifikation weiterer Regelungsbedarfe
- Ergebnisse der Gestaltung als Informationen in das technische System einbetten für die Transparenz hinsichtlich der getroffenen Regelungen

Abb. 3: Best Practices zur partizipativen Gestaltung in Hochschulen

5 Zusammenfassung und Ausblick

Ausgehend von dem konkreten Fallbeispiel der sozio-technischen Gestaltung des Videokonferenztools Webex wurden Best Practices für die partizipative, sozio-technische Systemgestaltung in Hochschulen abgeleitet. Nach den sehr guten Erfahrungen mit dem Vorgehen kann es nun in weiteren Fällen wie beispielsweise des Umgangs mit ChatGPT an Hochschulen erneut angewendet werden. Als Erweiterung des partizipativen Vorgehens sollten zusätzlich die Studierenden in den Prozess eingebunden werden. Zudem ist in weiteren Schritten zu analysieren, wie die getroffenen Maßnahmen wirken.

Danksagung: Wir bedanken uns bei allen Teilnehmer*innen der Experimentiergruppe für ihren Beitrag.

Literaturverzeichnis

- [HG23] Hense, J.; Goertz, L.: Monitor Digitalisierung 360° Wo stehen die deutschen Hochschulen? Hochschulforum Digitalisierung, Arbeitspapier Nr. 68, 2023, Online, zuletzt abgerufen am 14.07.2023
- [He02] Herrmann, Th.; Kunau, G.; Loser, K.-U.: Sociotechnical Walkthrough – ein methodischer Beitrag zur Gestaltung soziotechnischer Systeme. In (M. Herczeg; Prinz, W.; H. Oberquelle, H. Hrsg.): Mensch & Computer 2002. Teubner, Stuttgart, S. 323-332, 2002.
- [Fl89] Floyd, Ch.; Reisin, F.-M.; Schmidt, G.: STEPS to software development with users. In: (Ghezzi, C.; McDermid, J.A. Hrsg.): 2nd European Software Engineering Conference, University of Warwick, Coventry, UK, September 11-15. Proceedings. Springer, Berlin, Heidelberg, S. 48-64, 1989.
- [Gr21] Graf-Schlattmann, M.: Hochschulorganisation und Digitalisierung. Springer Fachmedien, Wiesbaden, 2021.
- [KB98] Kensing, F.; Blomberg, J.: Participatory Design: Issues and Concerns. In: Computer Supported Cooperative Work (CSCW), Vol. 7, S. 167-185, 1989.
- [KK14] Kienle, A.; Kunau, G.: Informatik und Gesellschaft – eine sozio-technische Perspektive. de Gruyter, München, 2014.
- [Pa20] Pasternack, P.; Schneider, S.; Zierold, St.: Wie kann Partizipation an der Hochschuldigitalisierung gestaltet werden? In: Henke, J.; Pasternack, P. (Hrsg.): Wie die Hochschulen durch das Zeitalter des Frühdigitalismus kommen. Wiesbaden: Springer Fachmedien Wiesbaden GmbH, S. 66-68, 2020.
- [Sy85] Sydow, J.: Der soziotechnische Ansatz der Arbeits- und Organisationsgestaltung - Darstellung, Kritik, Weiterentwicklung. Campus Verlag, Frankfurt, New York, 1985.
- [vL22] von Lojewski, U.: Transparenz als Gelingensbedingung guter Partizipation. In: Strategie Digital. Magazin für Hochschulstrategien im digitalen Zeitalter, Ausgabe 3, 2022, S. 18-23.

„Ich denke, dass ethische Probleme nicht in der Entwicklung, sondern in der Anwendung auftreten“

Einstellungen von Studierenden zu ethischen Aspekten in der Informatik

Nils Pancratz¹

Abstract: In dem Beitrag wird eine Untersuchung der Einstellungen von Studierenden zu ethischen Aspekten in der Informatik vorgestellt. Eine entsprechende Erfassung der Studierendenperspektive ist gemäß des Forschungsrahmensmodells der Didaktischen Rekonstruktion nötig, um sie im Sinne konstruktivistischer Lehre bei der Planung von Vorlesungen und Seminaren zum Thema „Informatik und Gesellschaft“ nutzen zu können. An der Untersuchung nahmen 70 Studierende teil, die schwerpunktmäßig in verschiedenen Informatikstudiengängen eingeschrieben waren. Die Daten wurden mithilfe eines Online-Fragebogens erhoben und sowohl quantitativ als auch qualitativ ausgewertet. Die Ergebnisse deuten darauf hin, dass männliche Studenten ihre Fähigkeiten zur Analyse ethischer Aspekte in der Informatik tendenziell höher einschätzen als ihre weiblichen Kommilitoninnen. Informatik-Fach-Bachelor-Studierende haben dabei weniger Optimismus bzgl. zukünftiger gesellschaftlicher Auswirkungen von Informatiksystemen als Studierende anderer Studiengänge. Zusätzlich ergeben sich Hinweise darauf, dass Informatik-Fach-Bachelor-Studierende im Vergleich zu Studierenden von Bindestrich-Informatikstudiengängen ein geringeres Interesse an ethischen Fragestellungen in der Informatik haben könnten. Dies deutet darauf hin, dass ihre Teilnahme an Lehrveranstaltungen zu diesem Themenbereich eher extrinsisch motiviert ist. Zuletzt offenbaren die Ergebnisse zwar grundsätzlich eine Vielfalt an individuellen Vorstellungen zu ethischen Problemen in der Informatik, die sich bei genauerer Betrachtung jedoch vorrangig auf Themen wie Datenschutz, Überwachung und den Umgang mit personenbezogenen Daten konzentrieren. Die Ergebnisse der qualitativen Inhaltsanalyse der Vorstellungen zu ethischen Problemen in der Informatik können zukünftig zur Erweiterung des eingesetzten Messinstruments genutzt werden.

Keywords: Vorstellungen, Ethik, Informatik und Gesellschaft, Hochschuldidaktik

1 Einleitung

Eine verantwortungsvolle Gestaltung digitaler Technologien ist für zukünftige gesellschaftliche, kulturelle, soziale, ökologische und wirtschaftliche Entwicklungen von großer Bedeutung; die hochschuldidaktisch fundierte Vermittlung entsprechender Inhalte im Informatikstudium ist daher unverzichtbar. Aus konstruktivistischer Sicht ist es dabei unerlässlich, in Lehrveranstaltungen zu „Informatik & Gesellschaft“ auch Studierendenperspektiven einzubeziehen; Einstellungen und Alltagsvorstellungen sind jedoch bislang in diesem Bereich kaum erforscht. Das im vorliegenden Beitrag vorgestellte Forschungsvorhaben untersucht

¹ Universität Hildesheim, Universitätsplatz 1, 31141 Hildesheim, pancratz@uni-hildesheim.de

daher Einstellungen und Vorstellungen von Studierenden zu ethischen Aspekten in der Informatik. Im Beitrag werden erste Ergebnisse zur Untersuchung der ethischen Position in der Informatiksystementwicklung, zum Interesse an Fragestellungen zu ethischen Aspekten in der Informatik sowie zur Selbstwirksamkeitserwartung in diesem Kontext vorgestellt. Dabei wird u. a. der Frage nachgegangen, ob sich die genannten Einstellungen der Studierenden in Abhängigkeit von Geschlecht oder Studiengang zeigen. Außerdem werden Vorstellungen zum Auftreten ethischer Probleme in der Informatik und den damit einhergehenden Auswirkungen qualitativ-inhaltsanalytisch ausgewertet. Ziel des Forschungsvorhabens ist damit neben einer Untersuchung der oben genannten Fragestellungen insbesondere auch die Modellierung der Vorstellungen von ethischen Aspekten in der Informatik.

2 Methodik

Bei der **Datenerhebung** wird zur Erfassung der Selbsteinschätzung der ethischen Position bezüglich der Entwicklung von Informatiksystemen das validierte und im Rahmen einer Voruntersuchung modifizierte Fragebogeninstrument **EPOS** (*Ethische Position bzgl. der Entwicklung von Informatiksystemen*) [BSM18; PM22] eingesetzt. EPOS umfasst 17 Likert-Skalen-Items, die zu fünf Subskalen zusammengefasst sind. Zur tiefergehenden Untersuchung der Studierendenperspektiven wurde das Fragebogeninstrument für diese Untersuchung um Items zur Erfassung des Interesses an, der Einstellung ggü. Ethik und Informatik sowie der Selbstwirksamkeitserwartung bzgl. entsprechender Aspekte erweitert. Dazu wurde auf das sechs Items umfassende validierte Fragebogeninstrument **EASE** (*Ethics Attitudes and Self-Efficacy*) [Ho22] zurückgegriffen, wobei die Likert-Skalierung zur Passung an EPOS von sieben auf fünf reduziert wurde und die Items ins Deutsche übersetzt sowie auf Informatiksysteme im Speziellen (statt Technologie im Allgemeinen) bezogen wurden. Angelehnt an [Ho22] wurde zusätzlich ein Item zum Optimismus über den zukünftigen Einfluss von Informatiksystemen auf die Gesellschaft mit aufgenommen. Zur Untersuchung von *Vorstellungen zu ethischen Problemen in der Informatik* (**VEPI**) komplettiert die folgende Frage das in dieser Untersuchung eingesetzte Erhebungsinstrument: *„Wo und wie können deiner Ansicht nach ethische Probleme der Informatik auftreten? Du kannst gerne einen kurzen erläuternden Text verfassen oder alternativ Stichpunkte formulieren (am besten mit Kommas oder Zeilenumbrüchen getrennt).“* VEPI wird über ein mehrzeiliges Freitextantwortfeld erhoben. Biografische Daten zu Alter, Geschlecht, Semesteranzahl und Studiengang wurden am Ende des Online-Fragebogens erhoben. Das **Sample** setzt sich überwiegend aus Teilnehmenden einer Lehrveranstaltung zu „Informatik und Gesellschaft“ an der Universität Oldenburg zusammen, die in der ersten Vorlesungswoche des Wintersemesters 2022/2023 um die Teilnahme an der Untersuchung über ein Ausfüllen des entwickelten Online-Fragebogens gebeten wurden. Zur Erweiterung des Samples wurde der Link zum Fragebogen über Lehrende auch an zwei weiteren Universitäten (Hildesheim und Potsdam) geteilt, an denen die Datenerhebung unabhängig² von einer entsprechenden

² Um für alle Teilnehmenden Aussagen über ihre bisherige Ausbildung in „Informatik und Gesellschaft“ treffen zu können, wurde im Fragebogen, der außerhalb der Veranstaltung „Informatik und Gesellschaft“ eingesetzt

Lehrveranstaltung ebenfalls zu Semesterbeginn stattfand. Zur Auswertung von EPOS und EASE kann auf insgesamt 70 vollständig ausgefüllte Datensätze zurückgegriffen werden (Tab. 1). Da bei der Erhebung von VEPI durch das Freitextfeld von einer erhöhten Drop-Out-Rate auszugehen war, war die entsprechende Angabe optional. Zur Auswertung von VEPI kann daher lediglich auf die Angaben von 36 Teilnehmenden zurückgegriffen werden.

Geschlecht	EPOS/EASE			VEPI	
	männlich (m)	weiblich (w)	divers	m	w
Informatik (Bachelor)	20	4	-	11	2
Wirtschaftsinformatik (Bachelor)	8	4	-	4	2
Informatik (Zwei-Fächer-Bachelor)	7	3	1	5	1
Wirtschaftswissensch. (Bachelor)	7	4	-	4	1
Informatik (Master of Education)	4	2	-	2	1
Sonstige	5	1	-	3	0

Tab. 1: Anzahl an Teilnehmenden, die Angaben zu EPOS/EASE bzw. VEPI gemacht haben

Zur **Auswertung** von EPOS wurden die Subskalenwerte ermittelt und über Varianzanalysen (unabhängige *t*-Tests beim Zwischensubjektfaktor Geschlecht³, ANOVA beim Zwischensubjektfaktor Studiengang) untersucht. Zur Auswertung von EASE wurden vier der Items über Mittelwertbestimmungen zur Subskala Selbstwirksamkeitserwartung zusammengefasst; die Reliabilität dieser Zusammenfassung wurde über den Index Cronbachs Alpha bestimmt. Es ergibt sich ein Wert, der auf eine sehr gute interne Konsistenz der EASE-Subskala Selbstwirksamkeitserwartung ($\alpha = 0.815$, 4 Items) hindeutet. Die verbleibenden EASE-Items sind nicht sinnvoll zu weiteren Subskalen zusammenfassbar. Auch bei der EASE-Auswertung wurden Varianzanalysen durchgeführt (s.o.). Zur Untersuchung der VEPI wurden die vorliegenden Freitextantworten zunächst unter Verwendung der Software MAXQDA kodiert. Die Frage enthält bewusst zwei Teilfragen nach dem Auftreten ethischer Probleme in der Informatik („wo?“) und den damit einhergehenden Auswirkungen („wie?“). Entsprechend bilden beide Teilfragenbereiche die oberste Kategorie (deduktiv entwickelt) des Kategoriensystems. Unterhalb dieser Ebene wurde das Material induktiv kodiert; die zunächst über In-Vivo-Codes angelegten Kategorien wurden in mehreren Durchläufen subsumiert, bevor ein finaler Kategorienleitfaden mit Kategoriendefinitionen und Ankerbeispielen entwickelt wurde. Eine Intra-Rater-Reliabilitätsprüfung im Rahmen einer erneuten vollständigen Zuordnung der Antworten zu den definierten Kategorien führte in MAXQDA zu einem *Cohen's Kappa Wert* von $\kappa = 0.81$.

3 Ergebnisse

Die **Auswertung von EPOS** hat ergeben, dass in den einzelnen Subskalen bis auf eine Ausnahme keine signifikanten Unterschiede zwischen männlichen und weiblichen Studie-

wurde, neben den weiteren biografischen Daten auch abgefragt, ob eine entsprechende Veranstaltung in der Vergangenheit bereits besucht wurde. Dies bejahten lediglich zwei Teilnehmende, sodass bei den elaborierten Perspektiven überwiegend von präinstruktionalen Einstellungen ausgegangen werden kann.

³ Der Datensatz der Person, die divers als Geschlecht angegeben hat, blieb bei den *t*-Tests unberücksichtigt.

renden vorliegen. Lediglich in der Subskala „Wissen“ unterscheidet sich EPOS zwischen den Geschlechtern männlich und weiblich signifikant: Studentinnen bewerten die Relevanz, dass Informatikerinnen und Informatiker über Wissen zu ethischen Rahmenbedingungen im Zusammenhang mit der Entwicklung von Informatiksystemen verfügen sollten, signifikant ($t(39.645) = -2.169, p = 0.036$) höher ($M = 4.222, SD = 0.572$) als ihre männlichen Kommilitonen ($M = 3.850, SD = 0.761$). Die **Auswertung von EASE** ergab keine signifikanten Unterschiede zwischen den Geschlechtern. Bezüglich der EASE-Komponenten zur Selbstwirksamkeit ist lediglich ein Trend in Richtung statistischer Signifikanz ($t(67) = 1.737, p = 0.087$) zu erkennen: Männliche Studenten schätzen ihre Fähigkeiten zur Analyse ethischer Aspekte in der Informatik tendenziell höher ein ($M = 2.946, SD = 0.806$) als ihre weiblichen Kommilitoninnen ($M = 2.569, SD = 0.742$). Einfaktorielle Varianzanalysen (ANOVA) von EPOS, EASE und ihren jeweiligen Subskalen bzw. Items mit Zwischensubjektfaktor Studiengang ergaben zunächst keine signifikanten Unterschiede zwischen den sechs betrachteten Gruppen (vgl. Tab. 1). Zur weiteren Untersuchung wurden die Studierenden daraufhin den beiden Gruppen „Informatikstudiengang“ (Informatik, Wirtschaftsinformatik, Informatiklehramt) und „Andere“ (Wirtschaftswissenschaften und Sonstige zugeordnet, wobei sich ebenfalls keine Hinweise auf signifikante Unterschiede in EPOS und EASE ergaben. Bei der verbleibenden (letzten sinnvollen) Eingruppierung in „Reine Informatik“ (Informatik Fach-Bachelor) und „Bindestrich-Informatik und Andere“ (Rest) zeigte sich ein signifikanter Unterschied in der Einstellung gegenüber ethischen Aspekten in der Informatik: So zeigen sich Informatik-Fach-Bachelor-Studierende im Mittel ($M = 2.88, SD = 1.484, N = 24$) signifikant ($F(1) = 5.012, p = 0.028$) weniger optimistisch bzgl. zukünftiger Auswirkungen von Informatiksystemen auf die Gesellschaft als Studierende anderer Studiengänge ($M = 3.57, SD = 1.068, N = 46$). Ein weiterer Trend in Richtung statistischer Signifikanz ($F(1) = 3.806, p = 0.055$) zeigt sich bei der Frage „Ich würde gerne mehr über ethische Probleme in der Informatik erfahren“, die Informatik-Fach-Bachelor-Studierende ($M = 3.54, SD = 1.103, N = 24$) im Mittel deutlich niedriger bewerten als die anderen Studierenden ($M = 4.02, SD = 0.906, N = 46$).

Die Ergebnisse der **Auswertung von VEPI** werden im Folgenden über eine Darstellung des entwickelten Kategoriensystems vorgestellt, wobei stellenweise Ankerbeispiele (Zitate) zur Verdeutlichung herangezogen werden. Die Frage nach dem **Auftritt ethischer Probleme in der Informatik (Kategorie 1, insgesamt 54 Kodierungen)** wird am häufigsten im Zusammenhang mit der *Diskussion von Werten, Normen und rechtlichen Aspekten* (K1.1, 18 Kod.) beantwortet, wobei in den Textpassagen insbesondere auf *datenschutzrechtliche Aspekte* (K1.1.1, 8 Kod.) eingegangen und auf die *Vereinbarkeit verschiedener Moralsysteme* (bspw. kulturell oder intergenerationell bedingt) hingedeutet wird (K1.1.2, 3 Kod.). Auch ein Bezug auf den *Einsatz von Informatiksystemen* (K1.2, 16 Kod.) ist häufig, wobei insbesondere auf *militärische Anwendungen* (K1.2.1, 3 Kod.) wie IT-gestützte Waffensysteme oder Spionagesoftware und den *Einsatz von KI im Straßenverkehr* (K1.2.2, 2 Kod.) eingegangen wird. Weiterhin treten ethische Probleme in der Informatik nach Ansicht der untersuchten Studierenden im Rahmen von *Überwachungsmaßnahmen* (K1.3, 7 Kod.) und beim *Umgang mit (personenbezogenen) Daten* (K1.4, 7 Kod.) auf. Im Zusammenhang mit *Bildung und*

Forschung (K1.4, 4 Kod.) wird bspw. die Frage nach im Unterricht zu behandelnden Inhalten („Welche informatischen Inhalte an Schulen unterrichtet werden sollen (z.B. Hacking).“) als ethisches Problem genannt. In dieser Kategorie ist insbesondere folgende Angabe hervorzuheben: „Ich denke das ethische Probleme nicht in der Entwicklung sondern in der Anwendung auftreten. So wie zum Beispiel PenTester Tools sowohl dazu verwendet werden können Systeme sicherer zu machen als auch sie anzugreifen. Demnach sollten eher Anwender kontrolliert und geschult werden als Entwickler [sic!]“. Weiterhin können ethische Probleme in der Informatik den Textpassagen zufolge auch *im Zusammenhang mit wirtschaftlichen Prozessen* (K1.6, 2 Kod.) auftreten. An **Auswirkungen ethischer Probleme in der Informatik** (K2, insg. 25 Kod.) werden primär *Auswirkungen auf Gesundheit und Leben* genannt (K2.1, 8 Kod.). Hierunter fallen vorrangig Nennungen zu informatischen Anwendungen, die mit einer *Gefährdung menschlichen Lebens* (K2.1.1, 4 Kod.) oder *Auswirkungen auf die (geistige) Gesundheit* (K2.1.2, 2 Kod.) einhergehen können. Des Weiteren werden die *Beeinflussung des Selbstwertgefühls* (K2.1.3, 1 Kod.) sowie *subtile Zwänge* (K2.1.4, 1 Kod.), die konkret mit „panoptische[n] Überwachungsstrukturen [...] einhergehen“, genannt. Auch auf *ökonomische und ökologische Auswirkungen* (K2.2, 8 Kod.) wird häufig Bezug genommen. Hierunter fallen insbesondere Bezüge auf den *Wandel des Arbeitsmarkts durch IT* (K2.2.1, 2 Kod., „Ersetzen von menschlicher Arbeitskraft durch Informatiksysteme“, „Entfallen einiger Jobs durch ersetzen [sic!] von effizienteren Robotern/digitalen Systemen“) oder durch *Werbemaßnahmen* (K2.2.2, 2 Kod.). In einem Fall wird die *Informationsasymmetrie* (K2.2.3, 1 Kod.) als Folge ethischer Probleme der Informatik genannt. Weiterhin werden die *Spaltung zwischen Arm und Reich* (K2.2.4, 1 Kod., „Informatik fördert möglicherweise die Lücke zwischen Arm und Reich“) und (ohne weitere Spezifizierung) *finanzielle Umstände* (K2.2.5, 1 Kod.) genannt. Hier löse „Informatik [...] einfache Jobs ab“, sei dabei aber „nicht für jeden erschwinglich zugänglich“. Auf *ökologische Auswirkungen* (K2.2.6, 1 Kod.) wird nur in einer Textstelle verwiesen („allgemein alle Bezugspunkte von IT zum Wohl von Lebewesen/Umwelt“). Unter *gesellschaftliche/politische Auswirkungen* (K2.3, 5 Kod.) fallen Äußerungen zur „Entstehung von Echo Räumen [sic!] für radikale politische Ideologien“. Nach vergleichbaren Perspektiven erleichtern Informatiksysteme die Bildung von Gruppen „mit nicht legalen oder gesellschaftskritischen Themen“ oder auch die „Manipulation von Personen durch (absichtlich) falsche Informationen“. Insbesondere im Zusammenhang mit KI deuten einige Textpassagen auf die Vorstellung von *Diskriminierung/Stereotypisierung* (K2.4, 4 Kod.) als Auswirkung ethischer Probleme in der Informatik hin. So seien bestimmte Funktionen wie bspw. die Gesichtserkennung bei Smartphones „für manche Hauttypen besser geeignet [...] als für [...] andere.[.]“.

4 Fazit und Ausblick

Die Ergebnisse lassen die Hypothese zu, dass männliche Studenten ihre Fähigkeiten zur Analyse ethischer Aspekte in der Informatik tendenziell höher einschätzen als ihre weiblichen Kommilitoninnen; gleichzeitig bewerten weibliche Studentinnen die Bedeutung von Wissen zu ethischen Rahmenbedingungen im Zusammenhang mit der Entwicklung von




Informatiksystemen höher. Außerdem zeigen sich Informatik-Fach-Bachelor-Studierende weniger optimistisch bzgl. zukünftiger Auswirkungen von Informatiksystemen auf die Gesellschaft als Studierende anderer Studiengänge. Zugleich lassen die Ergebnisse die Hypothese zu, dass Informatik-Fach-Bachelor-Studierende eher weniger interessiert an ethischen Fragestellungen in der Informatik sind als Studierende anderer Studiengänge, insbesondere Bindestrich-Informatiken. Daraus lässt sich die Hypothese ableiten, dass sie eher extrinsisch motiviert an Lehrveranstaltungen zu diesem Themenbereich teilnehmen.

In diesem Ergebnisbericht konnte das entwickelte Kategoriensystem zu VEPI überblicksartig dargestellt werden. Die Ergebnisse zeigen, dass die herausgearbeiteten Vorstellungen zwar grundsätzlich vielfältig sind, sie sich in der Gesamtbetrachtung aber schwerpunktmäßig auf Themen wie Datenschutz, Überwachung und den Umgang mit personenbezogenen Daten beziehen. Das Kategoriensystem kann zukünftig genutzt werden, um das bislang eingesetzte Testinstrument um Skalenitems zu VEPI zu erweitern, wodurch wiederum eine tiefergehende Untersuchung bspw. zur Korrelation von VEPI mit EPOS und EASE ermöglicht wird. Im Kontext von Vorstellungen zu ethischen Problemen von „Fehlvorstellungen“ zu sprechen erscheint unangebracht. Das titelgebende Zitat (vgl. Abs. 3) zeugt aber dennoch von einer tendenziell folgenschweren Alltagsvorstellung: So werden bereits bei der Entwicklung von Informatiksystemen Entscheidungen getroffen (bewusst sowie unbewusst), die auf bestimmten Werten und Annahmen basieren, was wiederum Ungerechtigkeiten oder Schäden zur Folge haben kann. Erfolgt im Rahmen entsprechender Lehrveranstaltungen keine konzeptuelle Rekonstruktion entsprechender Vorstellungen, so könnten sich Absolventinnen und Absolventen von Informatikstudiengängen mit entsprechenden Einstellungen den Konsequenzen ihrer Arbeit für die Benutzerinnen und Benutzer ihrer Entwicklungen insgesamt entziehen und möglicherweise nur geringe Bereitschaft zu partizipativer Technikgestaltung zeigen. Des Weiteren ist zukünftig zu diskutieren, inwieweit die hohe Drop-Out-Rate bei der Erhebung von VEPI zum Teil auf die Abwesenheit konkreter Vorstellungen zurückzuführen ist, wie sie durch die Antwort einer Teilnehmerin auf die Frage geäußert wird: „Ich habe mir über diese Frage noch vorher keine Gedanken gemacht“.

Literatur

- [BSM18] Brandenburg, S.; Schott, R.; Minge, M.: Zur Erfassung der ethischen Position in der Softwareentwicklung. In: Mensch und Computer 2018 - Workshopband. GI, Bonn, 2018.
- [Ho22] Horton, D.; McIlraith, S. A.; Wang, N.; Majedi, M.; McClure, E.; Wald, B.: Embedding Ethics in Computer Science Courses: Does It Work? In: Proc. of the 53rd ACM Techn. Symposium on Computer Science Ed. ACM, Providence, S. 481–487, 2022.
- [PM22] Pancratz, N.; Matzner, M.: Die ethische Position von Informatikstudierenden bezüglich der Entwicklung von Informatiksystemen. In: INFORMATIK 2022. GI, Bonn, S. 711–716, 2022.

Transdisziplinäre Entwicklung eines WebGIS für die Atommüll-Endlagerstandortsuche in Deutschland

Paula Bräuer ¹, Lucas Schwarz ² und Isabella Peters ³

Abstract: Die Suche nach einem Standort für ein Endlager für hochradioaktive Abfälle in Deutschland ist herausfordernd und informationsintensiv. Trotz unterschiedlicher Informationsangebote der beteiligten Akteure besteht der Bedarf nach einer zentralen und interaktiven Plattform. Da viele der Informationen, die für die Endlagersuche relevant sind, geographisch verortbar sind, bietet sich ein webbasiertes geographisches Informationssystem (WebGIS) an. WebGIS-Anwendungen werden bereits eingesetzt, um Bürger:innen in politischen Entscheidungsprozessen zu informieren oder deren lokales Wissen in der Projektplanung zu berücksichtigen. In einem mehrstufigen transdisziplinären Forschungsansatz wurde gemeinsam mit nicht-akademischen Akteuren ein WebGIS im Endlagerkontext erarbeitet und evaluiert. Die Ergebnisse zeigen, dass durch ein solches System sowohl Expert:innen als auch Neueinsteiger:innen niedrigschwellig angesprochen werden können. Der Beitrag erläutert die transdisziplinären Methoden und reflektiert den transdisziplinären Arbeitsprozess.


Keywords: Transdisziplinäre Forschung, WebGIS, Partizipation, Evaluation, Standortauswahl.

1 Einleitung


Die Standortsuche für ein Endlager in Deutschland ist eine räumliche Herausforderung: Gesucht wird der Standort mit der bestmöglichen Sicherheit für die Endlagerung hochradioaktiver Abfälle aus den deutschen Kernkraftwerken (siehe § 2 StandAG). Aufgrund der konfliktreichen Geschichte des Umgangs mit Kernenergie und radioaktiven Abfällen in Deutschland [BL97], sind gute Kommunikation und Transparenz wichtige Faktoren beim Standortsuchprozess für ein Endlager. Die Endlager-Kommission hat daher Elemente eines transparenten, lernenden und partizipativen Verfahrens vorgeschlagen, die in das 2017 novellierte StandAG eingeflossen sind [EK16].

Mehrere Plattformen wollen diesen Anforderungen nachkommen (z.B. Endlagersuche-Infoplattform oder WeChange-Plattform). Informationen auf diesen Plattformen sind allerdings häufig schwer zugänglich und können nur gefunden werden, wenn Nutzer:innen mit der Plattform vertraut sind. Gleichzeitig wird aber eine niedrigschwellige


¹ Christian-Albrechts-Universität zu Kiel, Web Science, Christian-Albrechts-Platz 4, 24118 Kiel,

pbra@informatik.uni-kiel.de,  <https://orcid.org/0000-0001-5903-8829>

² Freie Universität Berlin, Otto-Suhr-Institut für Politikwissenschaft, Ihnestr. 22, 14195 Berlin,

lucas.schwarz@fu-berlin.de,  <https://orcid.org/0000-0003-1520-1182>

³ ZBW – Leibniz-Informationszentrum Wirtschaft, Web Science, Düsternbrooker Weg 120, 24105 Kiel,

ipe@informatik.uni-kiel.de,  <https://orcid.org/0000-0001-5840-0806>

Kommunikation vorausgesetzt, um Vertrauen in Lösungen für Themen mit unterschiedlichen Risikowahrnehmungen zu erhöhen [WBF21]. Hier besteht Forschungsbedarf, der mehrere Ansprüche vereint: Es bedarf einer niedrighschwelligen Plattform, die alle Informationen organisationsunabhängig bündelt und verknüpft, und gleichzeitig einer Möglichkeit für Personen, ohne bisherige Berührungspunkte mit dem Thema der Endlagersuche, sich selbstständig in das Thema einzuarbeiten. Adressiert wurde dieser Forschungsbedarf mit einem transdisziplinären Ansatz, welcher ein WebGIS als niedrighschwellige und vielseitig einsetzbare Kommunikationsplattform entwickelt und analysiert hat. Die Abkürzung WebGIS (auch Public Participation GIS, PPGIS [SS05]) steht für ein webbasiertes geographisches Informationssystem, das räumliche Daten auf verschiedenen Ebenen darstellt und deren Beziehung sichtbar macht. Dabei können Informationsansichten von Nutzer:innen individuell zu- und abgeschaltet werden.

Personen aus allen Bevölkerungsteilen einzubinden, ist eine Grundprämisse eines WebGIS [BK18]. Gerade Laien können ihre Erfahrungen und ihr Wissen einbringen, was im Kontrast zu rein technokratischen Verfahren steht [Nu18]. Insbesondere an der Schnittstelle zwischen gesellschaftlichen und technischen Fragen kann ein WebGIS daher sein Potenzial entfalten [RK19]. Durch klare und leicht verständliche Kommunikation kann das Vertrauen in die Anwendung erhöht und ein Beitrag zur Konfliktminimierung geleistet werden [RHP19]. Der monodirektionale Informations- und Kommunikationsfluss kann durch ein WebGIS aufgebrochen werden, sodass eine Ermächtigung der Prozessteilnehmer:innen entstehen kann, z.B. über ein Forum, das Austausch ermöglicht. Trotz der vielfältigen Einsatzmöglichkeiten besteht im Bereich der partizipativen WebGISe Forschungsbedarf, z.B. sparen bisherige Studien relevante Details zur Dateneingabe aus [DHL22]. Bisher gibt es kaum Studien, in denen WebGISe transdisziplinär unter Beteiligung nicht-akademischer Akteure entwickelt wurden.

Dieser Beitrag basiert auf einem Arbeitsbericht des Forschungsprojektes TRANSENS [Sc23] und konzentriert sich auf die Beschreibung der Umsetzung und Erfahrungen mit der transdisziplinären Entwicklung eines WebGIS für die Endlagerstandortsuche.

2 Transdisziplinärer Forschungsansatz

Transdisziplinarität wird allgemein als kritische und selbstreflexive Herangehensweise verstanden, die gesellschaftliche Herausforderungen und wissenschaftlichen Fragestellungen verbindet [JBK12]. In TRANSENS werden nicht-akademische Akteure im Sinne der *school of engagement* [La22] in den Forschungsprozess integriert, d.h. sie werden aktiv in die Konzeption von Forschungsarbeiten (Co-Design) sowie in die Produktion von Ergebnissen (Co-Produktion) eingebunden. Bei der Endlagerung müssen Entscheidungen getroffen werden, die nicht unbedingt Sicherheit bis ins letzte Detail garantieren können, da sie aufgrund der langen Zeiträume, der geologischen Veränderungen im Wirtsgestein und der gesellschaftlichen Transformationen inhärenten Unsicherheiten unterliegen. Diese Entscheidungsspielräume erfordern daher eine

Erweiterung über das akademische Wissen hinaus, auf alltägliches, disperses Wissen, wie es bspw. Bürger:innen beitragen können. Denn es geht hier nicht nur um eine wissenschaftsbasierte Entscheidung für einen Entsorgungspfad, sondern auch um ethische Abwägungen, wie mit dem öffentlichen Ungut der hochradioaktiven Abfälle umzugehen ist [Th22]. [He21] beschreibt einen Wissenszuwachs durch den Einbezug nicht-akademischer Akteure in Form von neuen Problemverständnissen, Zielen und Handlungsimpulsen. Grundvoraussetzung für einen erfolgreichen Wissenszuwachs in transdisziplinärer Forschung ist daher die aktive Einbindung nicht-akademischer Akteure [HPH17]. Transdisziplinäre Forschung geht damit über *Citizen Science* [Si09] hinaus, indem die Einbindung aktiv im gesamten Forschungsprozess erfolgt und nicht nur auf wissenschaftliche Datenerhebung abzielt. Für diese Studie wurde dazu kontinuierlich mit einer Begleitgruppe von fünf Bürger:innen zusammengearbeitet, die anlassbezogen durch weitere nicht-akademische Akteure aus dem Kontext der Standortsuche erweitert wurde.

Um die gesellschaftliche Relevanz der Forschungsergebnisse zu gewährleisten und sie mit einer wissenschaftlichen Fragestellung zu verbinden, folgte der hier beschriebene Forschungsansatz den *Ten Steps* von [PKS17]. Diese regen die kontinuierliche Reflexion darüber an, wie sich der transdisziplinäre Einbezug der nicht-akademischen Akteure im Forschungsprozess und wie sich das Verhältnis zu den akademischen Akteuren gestaltet.

3 Studiendesign und transdisziplinäres praktisches Arbeiten

Workshop 1 Thema/Ziel <ul style="list-style-type: none"> • Forschungsfragen • Studiendesign 	Teilnehmer:innen <ul style="list-style-type: none"> • Akademische Akteure (n=2) • Begleitgruppe (n=4) • Beobachter:innen (n=2)
Workshop 2 Thema/Ziel <ul style="list-style-type: none"> • Arbeit mit dem WebGIS (explorativ) 	Teilnehmer:innen <ul style="list-style-type: none"> • Akademische Akteure (n=2) • Begleitgruppe (n=6) • Weitere Akteure (n=39)
Workshop 3 Thema/Ziel <ul style="list-style-type: none"> • Ergebnisreflexion & Co-Produktion • Fragen Formulieren 	Teilnehmer:innen <ul style="list-style-type: none"> • Akademische Akteure (n=2) • Begleitgruppe (n=6) • Beobachter:innen (n=1)
Workshop 4 Thema/Ziel <ul style="list-style-type: none"> • Ergebnispräsentation • Abschlussdiskussion 	Teilnehmer:innen <ul style="list-style-type: none"> • Akademische Akteure (n=2) • Begleitgruppe (n=6) • Weitere Akteure (n=19) • Beobachter:innen (n=1)

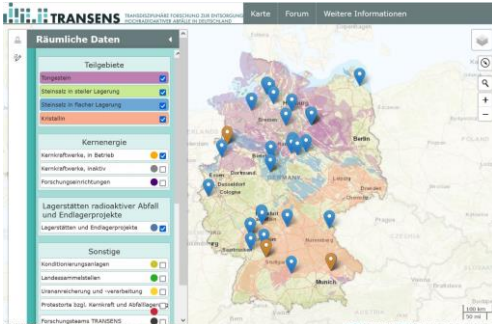


Abbildung 1. Studienablauf (links). Screenshot des explorativen WebGIS (rechts).

In der transdisziplinären Forschung werden unterschiedliche Methoden und Werkzeuge eingesetzt, um das Einbinden verschiedenster Akteure in verschiedenen Phasen der Forschung zu ermöglichen [DD19]. Ziel ist es, allen Akteuren die Möglichkeit zu geben, auf Augenhöhe zu interagieren. Da viele der Methoden nicht exklusiv auf transdisziplinäre Forschung anwendbar sind (z.B. *Design Thinking*), müssen diese Werkzeuge an die projektspezifischen Herausforderungen angepasst und ggf. modifiziert werden [Be10].

Die WebGIS-Studie wurde gemeinsam mit der Begleitgruppe in fünf verschiedenen Online-Workshops entworfen und evaluiert, in denen u.a. die Forschungsfrage sowie

Studienergebnisse diskutiert wurden (siehe Abb. 1 links). Die Initiierung des Projektes ging von den akademischen Akteuren aus, wurde aber mit der Begleitgruppe nachträglich abgestimmt und reflektiert. Auch Eingaben zur wissenschaftlichen Analyse und Methodik wurden weitestgehend von den akademischen Akteuren geleistet, jedoch jeweils in enger Abstimmung und unter Beteiligung der Begleitgruppe. Weitere nicht-akademische Akteure wurden einbezogen, wenn ihre Praxisnähe und die damit verbundenen Erfahrungen als bereichernd für die Co-Produktion angesehen wurden. Eine Besonderheit stellte die Erstellung der Ergebniszusammenfassung dar: Der Abschlussbericht wurde von den akademischen Akteuren verfasst, jedoch wurden alle Kapitel von der Begleitgruppe gelesen und überarbeitet. Zudem wurde jeder Abschnitt von der Begleitgruppe schriftlich kommentiert [Sc23], sodass das generierte Wissen bzw. die praktischen Erfahrungen der nicht-akademischen Akteure direkt in den Bericht einfließen. Durch diese Herangehensweise konnte auch in der Ergebniserzeugung ein flach hierarchisches Verhältnis zwischen den akademischen Akteuren und der Begleitgruppe gewahrt werden.

4 WebGIS und Evaluation durch nicht-akademische Akteure

Abb. 1 (rechts) zeigt das explorative WebGIS. Links befinden sich die verfügbaren räumlichen Daten, die durch Anklicken der Checkboxes einzeln ein- und ausgeblendet werden können. Somit ist eine individuelle Kombination der räumlichen Daten durch die Nutzer:innen möglich. Rechts befindet sich das Kartenfenster. Den Nutzer:innen war es möglich, andere Hintergrundkarten einzuschalten (Google Straßen, OSM Straßen, neutral, topographisch). Über die Suchfunktion konnten Orte/Adressen gesucht werden. Zu jedem Datensatz konnten durch Anklicken im Kartenfenster weiterführende Informationen abgerufen werden. Das WebGIS-Forum verfügt über eine Eingabefunktion, um ortsbezogene Hinweise oder Fragen im WebGIS einzutragen und ggf. durch einen Datenupload zu ergänzen. Alle Fragen und Hinweise wurden zudem auf einer Übersichtsseite zusammengefasst und konnten kommentiert werden. Das WebGIS basiert auf der JavaScript-Bibliothek Leaflet. Die Daten für die unterschiedlichen Kartenebenen werden über Esri gehostet und die Interaktion in einer MySQL Datenbank geloggt.

Das WebGIS wurde in einer Online-Studie evaluiert. Mit vier Aufgaben (z.B. „Blenden Sie die Teilgebiete im kristallinen Wirtsgestein ein.“) wurden die Funktionalitäten getestet und durch offene Fragen zu Vor- und Nachteilen eines WebGIS, Hindernisse und Chancen sowie wertvolle und fehlende Funktionen für den Standortsuchprozess erfasst. Außerdem wurden demografische Daten sowie die technische Affinität [WAF19] abgefragt.

5 Reflexion transdisziplinärer Arbeitsprozess

In der transdisziplinären Studie konnten vielfältige Aspekte eines WebGIS, von der Rolle im Verfahren über die Nutzung und Dialogmöglichkeiten bis hin zu technischen Aspekten und Funktionen, thematisiert werden. Die Evaluation zeigte, dass ein WebGIS Potenzial

bietet, bisher unbeteiligte Personen niedrigschwellig, zeit- und ortsunabhängig in den Standortauswahlprozess einzubinden. Neben inhaltlichen Ergebnissen ist vor allem die Reflexion der transdisziplinären Methodik zentral für diese explorative Studie. Zunächst lässt sich feststellen, dass sich transdisziplinäre Forschung (im hier aufgeführten Verständnis) durch einen hohen Zeitaufwand auszeichnet und aktiver Absprachen mit nicht-akademischen Akteuren bedarf. Gleichzeitig bietet sich aber die Chance, die eigenen akademischen Grenzen zu erweitern und neue Perspektiven einzubeziehen. Ferner ist vor allem das Mäandern zwischen trans- über inter- zu disziplinärem Arbeiten im Verlauf der WebGIS-Studie bemerkenswert. In den drei Phasen (Problem-Framing, Analyse und Auswirkungen erforschen) variierten die Aktivitäten stark und damit auch die Beteiligung der verschiedenen (nicht-)akademischen Akteure. Beim Problem-Framing und der Betrachtung der Auswirkungen waren Co-Produktion und interdisziplinäre Arbeit am stärksten ausgeprägt. Insbesondere die Erarbeitung der Forschungsfragen und die Ergebnisworkshops stechen hier hervor. In den weiteren Phasen der Studie waren zunächst die akademischen Akteure stärker in die Planung und Umsetzung eingebunden, während bei der Arbeit mit dem WebGIS der konsolidierende Beitrag der nicht-akademischen Akteure hervortritt. Zukünftige Projekte sollten sich bewusst sein, dass der Rahmen für transdisziplinäres Arbeiten fließend ist und von allen beteiligten Akteuren gestaltet wird.

6 Danksagung

TRANSENS wird vom Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz aufgrund eines Beschlusses des Deutschen Bundestages und im Niedersächsischen Vorab der Volkswagenstiftung vom Niedersächsischen Ministerium für Wissenschaft und Kultur von 2019 bis 2024 gefördert (FKZ 02E11849A-J).

7 Literaturverzeichnis

- [Be10] Bergmann, M.: Methoden transdisziplinärer Forschung. Ein Überblick mit Anwendungsbeispielen. Campus Verlag, Frankfurt, 2010.
- [BL97] Blowers, A.; Lowry, D.: Nuclear conflict in Germany: The wider context. *Environmental Politics* 3/97, S. 148–155, 1997.
- [BK18] Brown, G.; Kytä, M.: Key issues and priorities in participatory mapping: Toward integration or increased specialization? *Applied Geography* 95, S. 1–8, 2018.
- [EK16] Endlager-Kommission: Abschlussbericht der Kommission Lagerung hoch radioaktiver Abfallstoffe. Verantwortung für die Zukunft - Ein faires und transparentes Verfahren für die Auswahl eines nationalen Endlagerstandortes. Berlin. https://www.bundestag.de/resource/blob/434430/bb37b21b8e1e7e049ace5db6b2f949b2/drs_268-data.pdf, Stand: 20.07.2023
- [DD19] Defila, R.; Di Guilio, A.: Transdisziplinär und transformativ forschen, Band 2. Eine Methodensammlung. Springer Nature, Wiesbaden, 2019.

- [DHL22] Denwood, T.; Huck, J.; Lindley, S.: Participatory Mapping: A Systematic Review and Open Science Framework for Future Research. *Annals of the American Association of Geographers* 8/22, S. 2324–2343, 2022.
- [He21] Henze, J.: Zur Wissenschaftlichkeit transdisziplinärer Forschung. *GAIA -Ecological Perspectives for Science and Society* 1/21, S. 35–43, 2021.
- [HPH17] Hoffmann, S.; Pohl, C.; Hering, J.: Methods and procedures of transdisciplinary knowledge integration: empirical insights from four thematic synthesis processes. *Ecology and Society* 1/17, 2017.
- [JBK12] Jahn, T.; Bergmann, M.; Keil, F.: Transdisciplinarity: Between mainstreaming and marginalization. *Ecological Economics* 79, S. 1–10, 2012.
- [La22] Lawrence, M.; Williams, S.; Nanz, P.; Renn, O.: Characteristics, potentials, and challenges of transdisciplinary research. *One Earth* 1/22, S. 44–61, 2022.
- [Nu18] Nummi, P.: Crowdsourcing Local Knowledge with PPGIS and Social Media for Urban Planning to Reveal Intangible Cultural Heritage. *Urban Planning* 1/18, S. 100–115, 2018.
- [PKS17] Pohl, C.; Krütli, P.; Stauffacher, M.: Ten Reflective Steps for Rendering Research Societally Relevant. *GAIA - Ecological Perspectives for Science and Society* 1/17, S. 43–51, 2017.
- [RHP19] Rall, E.; Hansen, R.; Pauleit, S.: The added value of public participation GIS (PPGIS) for urban green infrastructure planning. *Urban Forestry & Urban Greening* 40, S. 264–274, 2019.
- [RK19] Rzeszewski, M.; Kotus, J.: Usability and usefulness of internet mapping platforms in participatory spatial planning. *Applied Geography* 103, S. 56–69, 2019.
- [Sc23] Schwarz, L.; Bräuer, P.; Consten, W.; Schöler, M.; Guth, A.; Nissen, N.; Gaebel, B.: Ein partizipatives WebGIS für die Endlagerstandortsuche: Eine transdisziplinäre Studie zu einer inklusiven Kommunikationsplattform [TRANSENS-Bericht-008], DOI: 10.21268/20230327-0, 2023.
- [SS05] Schlossberg, M.; Shuford, E.: Delineating "Public" and "Participation" in PGIS. *URISA Journal* 2/05, S. 15–26, 2005.
- [Si09] Silvertown, J.: A new dawn for citizen science. *Trends in ecology & evolution* 9/09, S. 467–471, 2009.
- [Th22] Themann, D.: Commoning in der Standortsuche für ein Endlager?: Neue Wege kollektiven Handelns. *TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 3/22, S. 51–57, 2022.
- [WAF19] Wessel, D.; Attig, C.; Franke, T.: ATI-S. An Ultra-Short Scale for Assessing Affinity for Technology Interaction in User Studies. In *Proceedings of Mensch und Computer 2019, Hamburg, Germany*, S. 147-154, ACM, 2019.
- [WBF21] Wiedemann, P.; Boerner, F.; Freudenstein, F.: Effects of communicating uncertainty descriptions in hazard identification, risk characterization, and risk protection. *PloS one* 7/21, e0253762, 2021.

Kultur Design - Digital Cultures
Cultural Analytics (InfDH 2023)

FMP Notebooks

Interaktives Lehren und Lernen der Digitalen Musikverarbeitung

Meinard Müller,¹ Frank Zalkow²

Abstract: Die meisten Menschen haben einen intuitiven Zugang zur Musik. Daher sind viele Aufgabenstellungen der digitalen Musikverarbeitung hervorragend geeignet, um abstrakte Konzepte in Gebieten wie der Signalverarbeitung, Mustererkennung oder Zeitreihenanalyse zu motivieren und in einer konkreten Anwendung einzusetzen. In diesem Beitrag stellen wir eine neuartige Sammlung von webbasiertem Multimediamaterial für das interaktive Studium und die Umsetzung eigener Projekte vor. Das als FMP Notebooks bezeichnete Multimediapaket basiert auf dem Lehrbuch *Fundamentals of Music Processing* (FMP) und erweitert traditionelles Lehrmaterial durch die Verknüpfung mit multimedialen Inhalten und Möglichkeiten der Online-Interaktion. Insbesondere kommen hierbei auf Python basierende interaktive Jupyter Notebooks zum Einsatz, die den Studierenden helfen, theoretische Konzepte nicht nur passiv nachzuvollziehen, sondern die gelernte Theorie auch aktiv in der Praxis einzusetzen. Die FMP Notebooks sind für Lehre und Forschung auf der Webseite <https://www.audiolabs-erlangen.de/FMP> frei verfügbar (CC BY-NC-SA 4.0).

Keywords: Musikverarbeitung; Music Information Retrieval; Audioverarbeitung; Python; Jupyter Notebook

Frei verfügbar Materialien zu den FMP Notebooks:

- FMP Notebooks: <https://www.audiolabs-erlangen.de/FMP>
- Educational Guide: [Mü21a]: <https://www.mdpi.com/2624-6120/2/2/18>
- Lehrbuch *Fundamentals of Music Processing* (FMP) [Mü21b]: www.music-processing.de
- Python-Paket libfmp [MZ21]: <https://github.com/meinardmueller/libfmp>

1 Musik trifft Informatik

Musik ist ein allgegenwärtiger und wichtiger Bestandteil unseres Lebens. Aufgrund des Angebots digitaler Musikdienste wie Spotify, Pandora und iTunes sind wir heute mehr denn je von Musik umgeben und interagieren mit ihr auf vielfältige Weise. Angesichts der explosionsartigen Zunahme digitaler Inhalte werden verstärkt computerbasierte Werkzeuge und Methoden zum Auffinden, Organisieren, Analysieren und Modifizieren von Musikdaten benötigt. Aus dieser Notwendigkeit heraus hat sich das neuartige Forschungsgebiet der digitalen Musikverarbeitung entwickelt, das im Englischen auch als *Music Information*

¹ International Audio Laboratories Erlangen, meinard.mueller@audiolabs-erlangen.de

² International Audio Laboratories Erlangen, frank.zalkow@audiolabs-erlangen.de

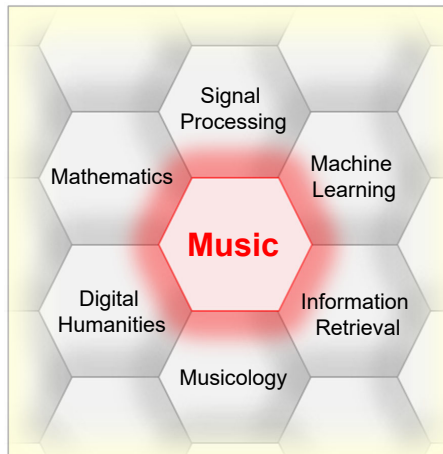


Abb. 1: Interdisziplinärer Bezug der Digitalen Musikverarbeitung.

*Retrieval*³ (MIR) bekannt ist. In ihren Anfängen hat die MIR-Forschung viele Ideen und Konzepte aus etablierten Disziplinen wie der Sprachverarbeitung oder der Computerlinguistik übernommen. Nach mehr als zwanzig Jahren ist MIR zu einem eigenständigen und interdisziplinären Forschungsgebiet gereift, das in engem Bezug mit unterschiedlichen Disziplinen wie der Informatik, den Ingenieurwissenschaften, der Signalverarbeitung, den Musikwissenschaften und den Digital Humanities steht (siehe Abbildung 1).

Dank der Vielfältigkeit und Komplexität von Musik gibt es viele Aufgabenstellungen der digitalen Musikverarbeitung, die als motivierende Szenarien für die Einführung, Erläuterung und das Studium von Techniken der Digitalen Signalverarbeitung, des Information Retrieval oder des Maschinellen Lernens dienen können. Zum Beispiel sind bei der Suche in heterogenen Musikdatenbeständen ganz unterschiedliche Modalitäten und Datentypen wie Partituren (Bilddaten) und Musikaufnahmen (Audiodateien) zu berücksichtigen. Zur Verarbeitung solcher Daten kommen Methoden der Bildverarbeitung, der Audioverarbeitung, und der Mustererkennung zum Einsatz. Für die Tempobestimmung benötigt man Verfahren zur Bestimmung von zeitlich, sich periodisch wiederholenden rhythmischen Mustern. Die Zerlegung von Musikaufnahmen in isolierte Instrumental- und Gesangsspuren (eine Problemstellung die allgemein hin als Quellentrennung bekannt ist) wird in den letzten Jahren insbesondere mittels Techniken des Deep Learning bewerkstelligt. Im Bereich der Audioidentifikation kommen effiziente Indexierungsverfahren zum Einsatz. Das Lehrbuch *Fundamentals of Music Processing* (FMP) [Mü21b] gibt einen umfangreichen Einblick in das Gebiet der digitalen Musikverarbeitung und vermittelt dabei allgemeine Grundlagen der Informatik und Signalverarbeitung in Theorie und Praxis. Die FMP Notebooks verknüpft

³ Hinter der Forschungsrichtung steht eine gemeinnützige Organisation, die International Society for Music Information Retrieval (ISMIR), siehe <https://ismir.net/>.

The screenshot displays a web browser window showing a Jupyter Notebook titled 'Fundamentals of Music Processing'. The notebook content includes a 'Definition' section with mathematical text and a formula for the Fourier transform, an 'Example: Shostakovich' section with musical notation and an audio player, and a 'Python Code' section with code snippets. The notebook is surrounded by a yellow border with a sidebar on the left containing navigation buttons: Teaching, Understanding, Programming, Baselines, Research, and Multimedia. At the bottom, there is a navigation bar with numbered tabs (0-8) and various icons. Red boxes highlight specific elements: 'Mathematics' (definition text), 'Theory' (mathematical formula), 'Music Example' (musical notation), 'Annotations' (text next to notation), 'Audio' (audio player), 'Links' (text below notation), 'Python Code' (code block), 'Functions' (code block), 'Visualization' (spectrogram plot), 'Results' (text below plot), 'Algorithms' (code block), 'Evaluation' (text below plot), and 'Sonification' (audio player at the bottom right).

Abb. 2: Komponenten und didaktische Aspekte der FMP Notebooks.

und erweitert das Lehrbuch mit Online-Interaktion und E-Learning durch den Einsatz innovativer und neuartiger multimedialer Werkzeuge [Mü21a; MZ19]. Hierauf gehen wir in den folgenden Abschnitten näher ein.

2 Jupyter Notebooks

In den letzten Jahren haben frei zugängliche Softwarepakete und webbasierte Frameworks die Ausbildung in der Informatik und den Ingenieurwissenschaften interaktiver gemacht. Der Einsatz solch neuartiger Technologien ermöglicht die Gestaltung von Kursen, die den Studierenden helfen, theoretische Konzepte nicht nur passiv nachzuvollziehen, sondern die gelernte Theorie auch aktiv in der Praxis einzusetzen. Die Motivation für die Entwicklung der FMP Notebooks ist, ein solches interaktives Lehren und Lernen der digitalen Musikverarbeitung zu unterstützen. Die FMP Notebooks basieren auf dem Open-Source Jupyter-Framework, das im Bildungsbereich vieler technischer Disziplinen zum Standard geworden ist [K116]. Mittels dieser umfassenden Webanwendung können Benutzer Dokumente erstellen, die lauffähigen Code, textbasierte Informationen, mathematische Formeln, Bilder, Klangbeispiele und Videos enthalten. Durch die Nutzung des Jupyter-Frameworks schließen die FMP Notebooks die Lücke zwischen Theorie und Praxis. Hierbei werden technische Konzepte und mathematische Details mit Codebeispielen, Illustrationen und Musikbeispielen unter Berücksichtigung didaktischer Aspekte verschachtelt (siehe Abbildung 2).



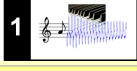



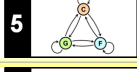

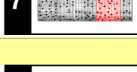
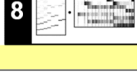
Part	Title	Notions, Techniques & Algorithms	HTML	IPYNB
	Basics	Basic information on Python, Jupyter notebooks, Anaconda package management system, Python environments, visualizations, and other topics	[html]	[ipynb]
	Overview	Overview of the notebooks (https://www.audiolabs-erlangen.de/FMP)	[html]	[ipynb]
	Music Representations	Music notation, MIDI, audio signal, waveform, pitch, loudness, timbre	[html]	[ipynb]
	Fourier Analysis of Signals	Discrete/analog signal, sinusoid, exponential, Fourier transform, Fourier representation, DFT, FFT, STFT	[html]	[ipynb]
	Music Synchronization	Chroma feature, dynamic programming, dynamic time warping (DTW), alignment, user interface	[html]	[ipynb]
	Music Structure Analysis	Similarity matrix, repetition, thumbnail, homogeneity, novelty, evaluation, precision, recall, F-measure, visualization, scape plot	[html]	[ipynb]
	Chord Recognition	Harmony, music theory, chords, scales, templates, hidden Markov model (HMM), evaluation	[html]	[ipynb]
	Tempo and Beat Tracking	Onset, novelty, tempo, tempogram, beat, periodicity, Fourier analysis, autocorrelation	[html]	[ipynb]
	Content-Based Audio Retrieval	Identification, fingerprint, indexing, inverted list, matching, version, cover song	[html]	[ipynb]
	Musically Informed Audio Decomposition	Harmonic/percussive separation, signal reconstruction, instantaneous frequency, fundamental frequency (F0), trajectory, nonnegative matrix factorization (NMF)	[html]	[ipynb]

Abb. 3: Inhaltlicher Überblick über die FMP Notebooks.

3 Struktur und Inhalte der FMP Notebooks

Die FMP Notebooks sind eine Sammlung bestehend aus mehr als 100 individuellen Notebooks, die in zehn Teile gegliedert sind (siehe Tabelle in Abbildung 3 für einen Überblick). **Teil 0** ist die Startseite der FMP Notebooks. Neben einem Überblick, der auch obige Tabelle als navigierbares Inhaltsverzeichnis enthält, spezifiziert diese Startseite auch die Lizenz (Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License) und stellt die wichtigsten Mitwirkenden bei der Erstellung der FMP Notebooks vor.

Die Notebooks von **Teil B** dienen unterschiedlichen Zwecken. Erstens beschreiben diese Notebooks die wichtigsten Werkzeuge, die für die Entwicklung der FMP Notebooks verwendet wurden, und dienen damit auch der Dokumentation. Zweitens geben sie kurze Einführungen in grundlegende technische Konzepte, die für die Installation und Verwendung von Jupyter und Python notwendig sind. Drittens findet man in **Teil B** zahlreiche Beispiele für die Erzeugung und Integration von Multimedia-Objekten (Text, Bilder, Audio, Video). Hierbei werden auch allgemeine Aspekte der Programmierung diskutiert, die in den anderen Teilen relevant werden.

Die restlichen acht Teile der FMP Notebooks orientieren sich eng an den acht Kapiteln des Lehrbuchs von Müller [Mü21b], das zentrale MIR-Themen wie Musikdarstellungen, Fourieranalyse, Musiksynchrisation, Strukturanalyse, Akkorderkennung, Beat-Tracking, Musiksuche und Audiozerlegung behandelt. Jeder der acht Teile besteht wiederum aus einer Übersichtsseite, die auf circa 10 bis 15 Notebooks verweist. Diese Notebooks enthalten Einführungen zu den jeweiligen MIR-Fragestellungen, geben zentrale mathematische Definitionen und beschreiben die algorithmischen Verfahren im Detail. Ein Hauptzweck der FMP Notebooks besteht darin, für diese Verfahren leicht verständliche Referenzimplementationen bereitzustellen. Diese werden dann anschließend auf konkrete und gut nachvollziehbare Musikbeispiele angewendet. Weiterhin kann die Benutzerin oder der Benutzer die Verfahren auf eigene Beispiele anwenden, mit Parametern experimentieren und die berechneten Ergebnisse durch geeignete Visualisierungen und Sonifikationen erfahrbar machen und damit besser verstehen. Diese Vorgehensweise ermöglicht auch die Generierung von Abbildungen und Illustrationen, die in Vorträgen und wissenschaftlichen Artikeln verwendet werden können. Auf diese Weise ergänzen die FMP Notebooks das eher theoretisch ausgerichtete Lehrbuch [Mü21b] und gehen darüber hinaus. Im folgenden gehen wir kurz auf die wesentlichen Inhalte der acht Teile ein.

In **Teil 1** (Music Representations) behandeln wir drei gängige Musikdarstellungen basierend auf Noten (graphische Darstellung), symbolischen Formaten und Wellenformen (Audio). Neben der Einführung der grundlegenden Terminologie, die in den folgenden FMP Notebooks verwendet wird, stellen wir Python-Code zur Verfügung, um musikalische und akustische Eigenschaften von Audiosignalen, einschließlich Aspekten wie Frequenz, Tonhöhe, Dynamik und Klangfarbe, erfahrbar und nachvollziehbar zu machen. So gibt es beispielsweise Python-Funktionen zum Vergleich verschiedener Stimmsysteme (gleichtemperierte, pythagoräische, harmonische Reihen) und zur Erzeugung von Shepard-Tönen.

In **Teil 2** (Fourier Analysis of Signals) nähern wir uns der Fourier-Transformation aus verschiedenen Perspektiven an. Wir stellen Code bereit, um komplexe Zahlen und Exponentialfunktionen und deren mathematische Eigenschaften besser zu verstehen. Diese bilden dann die Grundlage für die diskrete Fourier-Transformation (DFT) und die schnelle Fourier-Transformation (FFT) – ein Algorithmus von großer mathematischer Eleganz und praktischer Relevanz. Als weiteres wichtiges Thema diskutieren wir die Kurzzeit-Fourier-Transformation (STFT). In diesem Zusammenhang behandeln wir auch Aspekte wie

Sampling, Padding und Achsenkonventionen – Themen, die in der Theorie oft vernachlässigt, in der Praxis aber sehr wichtig werden.

In **Teil 3** (Music Synchronization) diskutieren wir Verfahren zum zeitlichen Abgleich von verschiedenen Versionen desselben zugrunde liegenden Musikstücks (z. B. unterschiedliche Interpretationen von Beethovens Fünfter Symphonie). Hierbei werden die Versionen in Merkmalsfolgen (sogenannte Chromagramme) transformiert, die harmonische und melodische Eigenschaften über den zeitlichen Verlauf hinweg erfassen, aber robust gegen klangfarbliche Änderungen sind. In diesem Zusammenhang befassen wir uns auch mit Fragen von hoher praktischer Relevanz wie Stimmung von Instrumenten, logarithmische Kompression sowie spektrale und zeitliche Auflösung – Aspekte, die einen wesentlichen Einfluss auf die Eigenschaften der Merkmale haben. Weiterhin führen wir die Technik des *Dynamic Time Warping* (DTW) ein, die für die Analyse und Alignierung allgemeiner Zeitreihen anwendbar ist.

In **Teil 4** (Music Structure Analysis) behandeln wir ein zentrales MIR-Themengebiet, bei dem es um die Segmentierung von Musikaufnahme hinsichtlich charakteristischer Strukturelemente geht. In diesem Kontext diskutieren wir grundlegende Segmentierungsprinzipien auf der Grundlage von Wiederholungen, Homogenität und Neuigkeit (Novelty). In den Notebooks stellen wir nicht nur Implementierungen unterschiedlicher Segmentierungsverfahren bereit (z. B. Novelty-Detection und Audio-Thumbnailing), sondern auch Code zur hierarchischen Visualisierung von Musikstrukturen (z. B. mittels Scape-Plots). Schließlich berühren wir kurz das Thema der Evaluation von Algorithmen und führen die Begriffe wie Precision, Recall und F-Measure ein. Diese Maßzahlen werden verwendet, um die berechneten Ergebnisse, die durch ein automatisiertes Verfahren erhalten werden, mit Referenzannotationen (Ground-Truth) zu vergleichen, die typischerweise von einem Domänenexperten manuell generiert werden.

In **Teil 5** (Chord Recognition) geht es um die Analyse harmonischer Eigenschaften eines Musikstücks, insbesondere der Bestimmung einer expliziten Akkordfolge für eine gegebene Audioaufnahme. Die Notebooks erklären zunächst musikalische Konzepte wie Intervalle, Akkorde und Tonleitern anhand audiovisueller Darstellungen. Nach der Einführung eines einfachen Akkorderkennungsverfahrens basierend auf einem Template-basierten Vergleichsverfahren, wird ein erweitertes Verfahren basierend auf Hidden-Markov-Modellen (HMMs) vorgestellt. Neben algorithmischen Aspekten und deren Implementierung dient das Akkorderkennungsszenario dazu, die Bedeutung von Entscheidungen beim Entwurf von Audiomerkmale und die Wechselwirkung zeitlicher und spektraler Glättungs- und Verarbeitungsstrategien herauszuarbeiten.

In **Teil 6** (Tempo and Beat Tracking) werden Tempo und Beat als grundlegende Eigenschaften von Musikaufnahmen analysiert. Die Notebooks behandeln zunächst energie- und spektralbasierte Verfahren zur Detektion von Noteneinsatzzeiten (Onsets) in Musiksignalen. Um Tempo- und Beat-Informationen abzuleiten, werden Noten-Einsatzkandidaten hinsichtlich quasiperiodischer Muster analysiert. Dieser zweite Schritt führt uns zum Studium

allgemeiner Methoden zur lokalen Periodizitätsanalyse von Zeitreihen. Insbesondere stellen wir zwei konzeptionell unterschiedliche Methoden – basierend auf Fourieranalyse und Autokorrelation – vor.

In **Teil 7** (Content-Based Audio Retrieval) besprechen wir unterschiedliche Suchstrategien für Musikaufnahme, die auf dem Paradigma *Query-by-Example* basieren. Bei der Audio-identifikation soll eine Musikaufnahme anhand eines kurzen, nur aus wenigen Sekunden bestehenden Audioclips identifiziert werden. Wir stellen hierzu ein auf spektralen Peaks basierendes Suchverfahren vor, das auch in kommerziellen Anwendungen wie Shazam⁴ zum Einsatz kommt. In weiterführenden Suchaufgaben geht es dann um Versionsidentifizierung, bei der auch andere Versionen wie z. B. Cover-Songs oder Arrangements gefunden werden sollen. Neben der Bereitstellung von Code für Basissysteme sollen die Notebooks ein grundlegendes Verständnis der praktischen Anforderungen für die verschiedenen Retrieval-Aufgaben vermitteln.

In **Teil 8** (Musically Informed Audio Decomposition) widmen wir uns drei Aufgabenstellungen, die dem allgemeinen Themengebiet der Quellentrennung zugeordnet werden können: harmonisch-perkussive Zerlegung, Extraktion der Hauptmelodie und Notentext-informierte Audiozerlegung. Innerhalb dieser Szenarien bieten die Notebooks detaillierte Erklärungen und Implementierungen grundlegender Techniken wie Momentanfrequenzschätzung, Grundfrequenzschätzung (F_0), Inversion von Spektrogrammdarstellungen und nicht-negative Matrixfaktorisierung (NMF). Diese Techniken sind für eine Vielzahl von Verarbeitungsschritten von allgemeinen Multimediadaten nützlich, die über die Quellentrennung und Musikverarbeitung hinausgehen.

4 Weitere Hinweise

Die meisten Menschen haben einen intuitiven Zugang zur Musik. Daher sind viele MIR-Aufgabenstellungen, wie bereits zuvor erwähnt, hervorragend geeignet, um abstrakte Konzepte der Signalverarbeitung, Mustererkennung oder Zeitreihenanalyse in einer konkreten Anwendung einzusetzen und damit besser zu verstehen [MMK21]. Weiterhin hat die zunehmende Verfügbarkeit von geeigneten Software-Toolboxen und webbasierten Frameworks nicht nur die Reproduzierbarkeit von Forschungsergebnissen vereinfacht, sondern auch die Ausbildung in der Informatik und anderen technischen Disziplinen erheblich verbessert [Mc19]. Seit ihren Anfängen vor mehr als zwanzig Jahren hat die MIR-Community zu solchen Open-Source-Entwicklungen mit unterschiedlichen Toolboxen beigetragen, welche modularen Quellcode für die Verarbeitung und Analyse von Musiksignalen bereitstellen.

⁴ <https://www.shazam.com/>

Prominente Beispiele sind `essentia`⁵ [Bo13], `madmom`⁶ [Bö16], `Marsyas`⁷ [Tz09], oder die `MIRtoolbox`⁸ [LT07].

Diese Toolboxen sind hauptsächlich für einen forschungsorientierten Zugang zur Audio- und Musikverarbeitung konzipiert und stellen eine breite Palette von Funktionalitäten für die Extraktion von semantischen Merkmalen zur Verfügung. Das Python-Paket `librosa` [Mc15] geht darüber hinaus, indem es modularen und didaktisch aufgearbeiteten Code für ganze MIR-Pipelines anbietet. Die FMP Notebooks greifen diesen Ansatz auf und erweitern `librosa`, indem Code und Implementierungen mit Lehrbuchelementen verquickt werden, um dadurch das Verständnis von MIR-Konzepten zu fördern. Anstelle von kompaktem und möglichst effizientem Code ist der Programmierstil der FMP Notebooks daher eher einfach und explizit gehalten – mit einer flachen funktionalen Hierarchie (auch auf Kosten einer gewissen Redundanz). Die mathematische Notation und die Namenskonventionen der FMP Notebooks sind sorgfältig aufeinander abgestimmt, um einen engen Bezug zwischen Theorie und Praxis herzustellen. Zusammenfassend sind pädagogische und didaktische Überlegungen der Hauptleitfaden bei der Entwicklung der FMP Notebooks. Daher hoffen wir, dass diese Notebooks bestehende Open-Source-Toolboxen gut ergänzen und Bildung sowie Forschung im Bereich der digitalen Musikverarbeitung und Informatik fördern.

Bei der Entwicklung von Lernenden zu Forschenden sollten Studierende über die Anwendung der FMP Notebooks hinauswachsen und ihre eigenen Methoden für die Musikverarbeitung entwickeln [MMK21]. Das Python-Paket `librosa` [Mc15] ist geeignet, um Studierende bei diesem Übergang zu unterstützen, da es die richtige Softwareunterstützung für das Erstellen und Experimentieren mit komplexen MIR-Pipelines bietet. Hiervon inspiriert wurde in einem ähnlichen Stil das Python-Paket `libfmp` [MZ21] entwickelt. Während die FMP Notebooks schrittweise in grundlegende MIR-Konzepte einführen, bündelt das `libfmp`-Paket die vorgestellten Kernkonzepte in Form einer gut dokumentierten und einfach zu bedienenden Python-Bibliothek. Zum einen integriert `libfmp` das `librosa`-Paket und baut auf diesem auf. Zum anderen stellt `libfmp` auch zum ersten Mal verschiedene Referenzimplementierungen von zuvor in der Forschungsliteratur veröffentlichten MIR-Methoden bereit, die von `librosa` und anderen öffentlich verfügbarer Python-Paketen noch nicht abgedeckt wurden.

Wir wollen abschließend erwähnen, dass der Code von `libfmp` in einem GitHub-Repository⁹ unter einer MIT-Lizenz öffentlich verfügbar ist. Neben den als Jupyter Notebooks und HTML-Export¹⁰ verfügbaren FMP Notebooks, in denen man didaktisch aufgearbeitete Erklärungen fast aller `libfmp`-Funktionen findet, wurde auch eine für Entwickler übliche API-Dokumentation¹¹ erstellt. Weiterhin wurde das `libfmp`-Paket in dem Python-Paketindex

⁵ <https://essentia.upf.edu>

⁶ <https://madmom.readthedocs.io/>

⁷ <http://marsyas.info/>

⁸ <https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox>

⁹ <https://github.com/meinardmueller/libfmp>

¹⁰ <https://www.audiolabs-erlangen.de/FMP>

¹¹ <https://meinardmueller.github.io/libfmp>

PyPi aufgenommen, sodass `libfmp` mit dem Standard-Python-Paketmanager `pip` installiert werden kann.

5 Danksagung

Die FMP Notebooks bauen auf Ergebnissen, Materialien und Erkenntnissen auf, die in enger Zusammenarbeit mit verschiedenen Personen gewonnen wurden. Wir möchten unseren Dank an ehemalige und aktuelle Studierende, Kolleginnen und Kollegen aussprechen, die uns bei der Erstellung der FMP Notebooks unterstützt haben, einschließlich Vlora Arifi-Müller, Stefan Balke, Michael Krause, Brian McFee, Sebastian Rosenzweig, Fabian-Robert Stöter, Steve Tjoa, Angel Villar-Corrales, Christof Weiß und Tim Zunner. Außerdem danken wir der Deutschen Forschungsgemeinschaft (DFG) für die finanzielle Unterstützung über die letzten zehn Jahre, ohne die eine Grundlagenforschung in der digitalen Musikverarbeitung kaum möglich wäre. Schließlich danken wir für die Unterstützung durch die International Audio Laboratories Erlangen – einer gemeinsamen Einrichtung der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und des Fraunhofer Instituts für Integrierte Schaltungen IIS.

Literatur

- [Bo13] Bogdanov, D.; Wack, N.; Gómez, E.; Gulati, S.; Herrera, P.; Mayor, O.; Roma, G.; Salamon, J.; Zapata, J. R.; Serra, X.: *Essentia: An Audio Analysis Library for Music Information Retrieval*. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Curitiba, Brazil, S. 493–498, 2013.
- [Bö16] Böck, S.; Korzeniowski, F.; Schlüter, J.; Krebs, F.; Widmer, G.: *madmom: A New Python Audio and Music Signal Processing Library*. In: *Proceedings of the ACM International Conference on Multimedia (ACM-MM)*. Amsterdam, The Netherlands, S. 1174–1178, 2016.
- [K116] Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, J.; Corlay, S.; Ivanov, P.; Avila, D.; Abdalla, S.; Willing, C.; development team, J.: *Jupyter Notebooks—a publishing format for reproducible computational workflows*. In: *Proceedings of the International Conference on Electronic Publishing*. Göttingen, Germany, S. 87–90, 2016.
- [LT07] Lartillot, O.; Toiviainen, P.: *MIR in MATLAB (II): A Toolbox for Musical Feature Extraction from Audio*. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Vienna, Austria, S. 127–130, 2007.
- [Mc15] McFee, B.; Raffel, C.; Liang, D.; Ellis, D. P.; McVicar, M.; Battenberg, E.; Nieto, O.: *Librosa: Audio and Music Signal Analysis in Python*. In: *Proceedings the Python Science Conference*. Austin, Texas, USA, S. 18–25, 2015.

- [Mc19] McFee, B.; Kim, J. W.; Cartwright, M.; Salamon, J.; Bittner, R. M.; Bello, J. P.: Open-Source Practices for Music Signal Processing Research: Recommendations for Transparent, Sustainable, and Reproducible Audio Research. *IEEE Signal Processing Magazine* 36/1, S. 128–137, 2019.
- [MMK21] Müller, M.; McFee, B.; Kinnaird, K.: Interactive Learning of Signal Processing Through Music: Making Fourier Analysis Concrete for Students. *IEEE Signal Processing Magazine* 38/3, S. 73–84, 2021.
- [Mü21a] Müller, M.: An Educational Guide Through the FMP Notebooks for Teaching and Learning Fundamentals of Music Processing. *Signals* 2/2, S. 245–285, 2021.
- [Mü21b] Müller, M.: *Fundamentals of Music Processing – Using Python and Jupyter Notebooks*. Springer Verlag, 2021.
- [MZ19] Müller, M.; Zalkow, F.: FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Delft, The Netherlands, S. 573–580, 2019.
- [MZ21] Müller, M.; Zalkow, F.: libfmp: A Python Package for Fundamentals of Music Processing. *Journal of Open Source Software (JOSS)* 6/63, 3326:1–5, 2021.
- [Tz09] Tzanetakis, G.: Music analysis, retrieval and synthesis of audio signals MAR-SYAS. In: *Proceedings of the ACM International Conference on Multimedia (ACM-MM)*. Vancouver, British Columbia, Canada, S. 931–932, 2009.

A Corpus of Memes from Reddit: Acquisition, Preparation and First Case Studies

Thomas Schmidt,¹ Fabian Schiller,² Mathias Götz³ and Christian Wolff⁴

Abstract: We present a corpus of memes and their textual components that were acquired from the popular meme platform r/memes, a subreddit of Reddit and one of the major outlets of online meme culture. The corpus consists of the most popular memes from 2013-2021 on the platform and we acquired 11,701 memes and 280,351 text tokens. We conduct several case studies focused on diachronic analysis to highlight the possibilities of the corpus for research in internet studies and online culture. We examine the general activity on the platform throughout the years and identify a significant increase in meme production beginning 2017. Results of sentiment analysis show a tendency towards memes with positively classified texts. The analysis of most frequent words per half-year spotlights the importance of certain cultural events for meme culture (e.g. the 2016 US election). Using the LIWC to analyze swear and sexual words shows an overall decrease in the usage of these words pointing to an increased moderation of the platform. The corpus is publicly available for the research community for further studies.

Keywords: memes; internet studies; corpus; natural language processing; sentiment analysis; Reddit

1 Introduction

Memos have become a popular media type in today's internet culture. Bauckhage [Ba11] defines a meme as "content or concepts that spreads rapidly among internet users" and Davison [Da12] extends this definition to memes as "a piece of culture, typically a joke, which gains influence through online transmission". One of the most popular meme types is the usage of an image with a text, usually consisting of one or two lines that formulate a joke or punch line (see figure 1 for a conceptual template, figure 2 for an example). The image is referred to as image macro or meme template and they are usually reused with different text changing based on context, goal, and situation of the overall meme. We refer to the different forms based on differing texts of the image macros as meme incarnations.

Due to the large popularity and widespread usage of memes in social media, memes have become a topic of interest for various humanities branches like internet and cultural studies. From a methodological point of view, research is focused on qualitative / hermeneutic methods with a rather low number of memes in qualitative case studies [Sh12, Sh14, Sh13,

¹ Media Informatics Group, University of Regensburg, thomas.schmidt@ur.de

² Media Informatics Group, University of Regensburg, fabian.schiller@ur.de

³ Media Informatics Group, University of Regensburg, mathias-christian.goetz@ur.de

⁴ Media Informatics Group, University of Regensburg, christian.wolff@ur.de

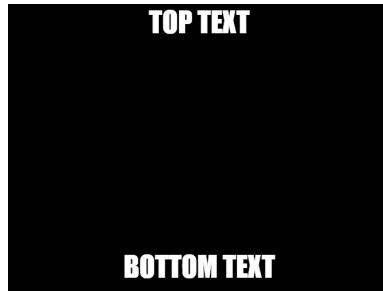


Fig. 1: Typical format of an image macro meme.

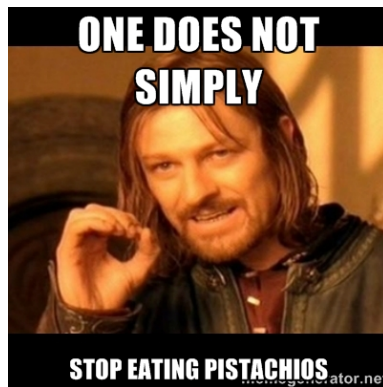


Fig. 2: Example of a meme of the corpus (posted in the second half-year of 2015).

Os15, HL16, Mc19]. The application of computational methods for meme analysis in digital humanities (DH) is rather rare [Sc20b]. In natural language processing (NLP), memes often serve as rich data source in the context of other primary tasks e.g. to evaluate algorithms for hate language detection [Ki21], AI-based meme generation [AT18, Sa20] or sentiment analysis [Sh20]. Sherratt [Sh22] proposes the focus on multimodal methods of online memes but also argues that there is still a lack of public well-structured corpora for the creation of necessary knowledge bases. First developments towards that goal can be seen in the data science and machine learning community with the publication of datasets on open data platforms⁵ but they often lack structured metadata, text captions or point to deleted online content. Schmidt et al. [Sc20b] acquired and analyzed around 8,000 meme incarnations of 16 image macros from the platform Know Your Meme. Vyalla and Udandarao [VU20] created a corpus of over 1 million meme text captions but limited to 128 image macro classes and use this to develop a meme generation system.

In this paper, we present the methodological approach and the results of a project for the acquisition and analysis of a corpus of memes. Our goal is to publish a multimodal

⁵ <https://www.kaggle.com/datasets/electron0zero/memegenerator-dataset>

(images and text) corpus with rich metadata of this unique media type to support research in media, internet, and cultural studies. Furthermore, we present first case studies with various methods of computational text analysis to showcase the potential of this corpus for further studies. We primarily examine the development of textual meme content across time to explore influence and interactions with general culture.

2 Corpus

We refer to the subreddit `r/memes`⁶ as our source of acquisition. The platform, which started in 2008, has more than 26 million subscribers, 22 million posts and is the 10th largest subreddit on Reddit.⁷ `r/memes` is one of the major curated outlets for memes of the image macro type in the English language. Users can post memes as submissions on the platform and receive comments, upvotes and downvotes. As one of the largest Reddit communities a strict moderation is in place deleting non-meme content or content violating the terms of use (e.g. racist content).

Table 1 illustrates general corpus statistics. Token and sentence-based analysis was performed via Spacy.⁸ The 11,701 memes sum up to 280,351 tokens of text. On average a meme consists of 24.5 tokens and 1.8 sentences. Figure 3 illustrates the distribution among memes in the corpus per half-year.

Type / Metric	absolute	AVG	SD	Min	Max
Tokens	280,351	24.5	37.7	1	2,120
Sentences	20,340	1.8	4.0	1	397
Punctuation	23,773	2.0	10.4	0	800

Tab. 1: General corpus statistics.

3 Case Studies

To highlight the possibilities of this corpus, we have performed a first exploratory analysis with a focus on diachronic analysis of the textual content of the memes based on half-year sub-corpora of this corpus. Additional data relating to these analyses as well as the corpus itself are publicly available for the research community.⁹

3.1 Most Frequent Words

To gain an overview of the topics that are discussed, and the terms used on memes, we analyzed most frequent words per half-year and most frequent words in the corpus as a

⁶ <https://www.reddit.com/r/memes/>

⁷ All data as of May 5th 2023.

⁸ <https://spacy.io/>

⁹ https://github.com/lauchblatt/reddit_memes

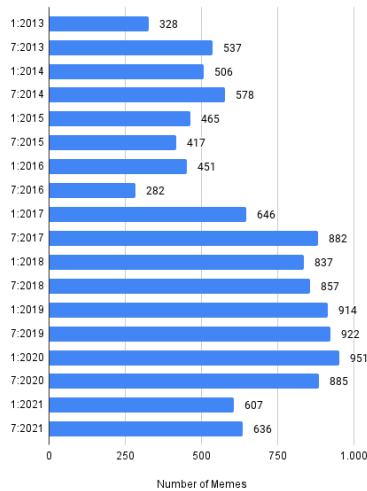


Fig. 3: Distribution of memes among each half-year.

whole using word clouds after stop word removal. The overall word cloud of all memes points to swear words and general terms as being very frequent (figure 4).

However, word clouds for specific half-years do mirror important topics in society (see figure 5 and 6 for examples). More word clouds can be examined on our GitHub repository.



Fig. 4: Word cloud of the overall meme corpus.

3.2 Sentiment Analysis

We also explored sentiment analysis using the Python toolkit Vader¹⁰ [HG14] which produces an average sentiment value for each text of each meme and is a recommended lexicon-based sentiment analysis often used for the analysis of social media content [SKW20, Mo20]. Following the recommendations of Vader, we regard an average sentiment value higher

¹⁰ <https://github.com/cjhutto/vaderSentiment>

content of the memes sums up to a rather neutral value (between -0.05 and 0.05). However, as figure 8 illustrates, there is a development towards more positive sentiment beginning in 2016. This is in line with the general rise of popularity of r\memes but also a more serious moderation and professionalization of the platform including a more controlled removal of racist and trolling content that is likely to be classified as negative.

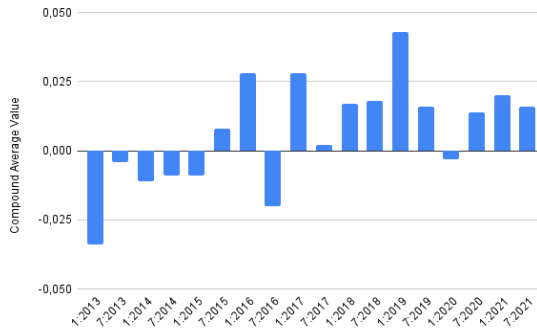


Fig. 8: The average sentiment score per half-year across time.

3.3 LIWC Analysis

Finally, we used the English version of the Linguistic Inquiry Word Count lexical resource¹¹ (LIWC; [Bo22]) to investigate the development of certain word categories across time. The LIWC is a psychologically validated lexical resource that offers lists of words for certain topics and while it is primarily used in psychology, it has also shown to be beneficial in the context of language analysis in social media by analyzing the overall proportions of word categories [Mo20]. The word lists consist of a fixed list of words including their inflections and words extended with regular expressions. We focus on the word categories swear (130 words, regex-forms not included, like bullshit, sucks, dumb and other slurs) and sexual (131 words like erotic, naked and sexual slurs) as those pose interesting insights in the context of the previously performed sentiment analysis. Please note however that this dictionary approach is limited due to the increased moderation of the subreddit (deleting and adjusting memes with explicit content) and also approaches by the community to avoid swear words by using placeholders (e.g. b*llsh*t for bullshit). Nevertheless, calculating the average proportion of these word categories among all memes results in an average of 0.19 for swear words and 0.07 for sexual words.

Looking at the proportion of these words across the time confirms previous findings of the sentiment analysis and shows the importance of swear words for this media type with average proportion values up to 0.42 in the second half-year of 2014 (figure 9). We also identified a moderate decrease from the peaks in 2014 to 2021 that is in line with the sentiment

¹¹ <https://www.liwc.app/>.

development and might be a sign of the increased professionalization and moderation of the platform deleting and managing memes with content like racial slurs that are part of the swear word category. Our results show that words with sexual connotations are not as important and frequent compared to swear words and rather rare in certain years (e.g. 2018, 2019). However, they do peak just recently in 2021.

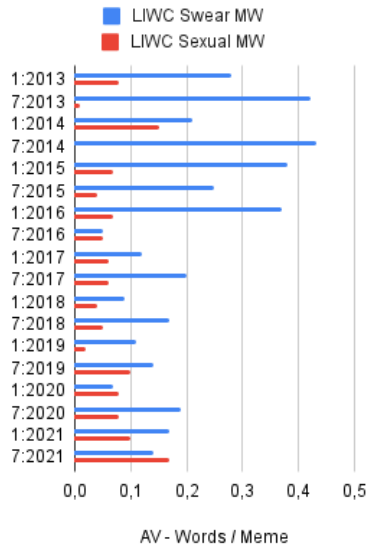


Fig. 9: Distribution of swear and sexual terms across the time. X-axis is the average proportion of word categories across all memes.

4 Limitations & Future Work

While these analyses are exploratory at the moment, they show the potential of this corpus and possible research ideas. We want to extend these cases studies by addressing several limitations:

We used Spacy, a standard NLP tool and general purpose solution to create tokens and sentences. However, we noticed that the special vocabulary and syntax used on memes pose challenges to this standard approach. We evaluated the functionality of the token and sentence segmentation with some random samples and while we got the impression that it works well most of the time, we also found problems dealing with slang, complex punctuation and URLs. In many cases, specific sentence segmentation decisions are open to interpretation. We plan to further analyze this problem by defining syntactical and token based units for our meme corpus and examining results of multiple segmentation tools in more detail.

The applied sentiment analysis as lexicon-based approach is rather limited and we did not evaluate the approach. We intend to annotate sub-corpora of the meme corpus with sentiment to evaluate methods and apply more sophisticated machine learning methods including more multifaceted emotion prediction [SDW21]. Furthermore, we intend to perform more fine-grained sentiment analysis considering the time spans e.g. sentiment progressions per month to get a more in-depth view of the data. We also argue that we can gain further insights of meme culture by analyzing other word categories of the LIWC. The analysis of the gender-based categories female and male offer potential for in-depth analysis in the context of gender studies [Sc20a]. We plan to investigate the method of topic modeling to examine the diachronic development of important topics discussed on memes in more depth. Additionally, we did not perform yet any major analysis into the interaction of certain metadata with sentiment or LIWC analysis e.g. the correlation of sentiment and popularity. We intend to uncover various analysis ideas of that branch in future work.

Furthermore, we agree with Sherratt [Sh22] that memes are a multimodal medium and we want to include methods of computer vision (e.g. image classification, object detection) to further analyze our corpora. This is further supported by the increased usage of video snippets, GIFs and similar content in recent meme culture. Currently, we are in the process to increase the overall corpus with more memes from r/memes but also other subreddits to get a broader view of the phenomenon and extend the corpus also with video based memes. Recent applications of computer vision methods applied in DH fields like theatre studies [SBW19, SW21], film studies [Sc21, Ho19, SK22, EKSW22] and internet studies [Sc20c] offer interesting possibilities in this context of meme analysis.

To our knowledge, only few structured meme corpora are publicly available [Sc20b] and we were able to gain interesting insights about the Reddit meme culture in regard to sentiment, topics and platform moderation that have implication for internet, media and social media studies.

Bibliography

- [AT18] AbelL.Peirson, V.; Tolunay, E. M.: Dank Learning: Generating Memes Using Deep Neural Networks. ArXiv, June 2018.
- [Ba11] Bauckhage, Christian: Insights into Internet Memes. Proceedings of the International AAAI Conference on Web and Social Media, 5(1):42–49, 2011. Number: 1.
- [Bo22] Boyd, Ryan; Ashokkumar, Ashwini; Seraj, Sarah; Pennebaker, James: The Development and Psychometric Properties of LIWC-22. February 2022.
- [Da12] Davison, Patrick: 9. The Language of Internet Memes. In: 9. The Language of Internet Memes, pp. 120–134. New York University Press, March 2012.
- [EKSW22] El-Keilany, Alina; Schmidt, Thomas; Wolff, Christian: Distant Viewing of the Harry Potter Movies via Computer Vision. In: Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries Conference (DHNB 2022). Uppsala, Sweden, pp. 33–49, 2022.

- [HG14] Hutto, C.; Gilbert, Eric: VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):216–225, May 2014. Number: 1.
- [HL16] Highfield, Tim; Leaver, Tama: Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, 2(1):47–62, January 2016. Publisher: Routledge _eprint: <https://doi.org/10.1080/22041451.2016.1155332>.
- [Ho19] Howanitz, Gernot; Bermeiteinger, Bernhard; Radisch, Erik; Gassner, Sebastian; Rehbein, Malte; Handschuh, Siegfried: Deep Watching - Towards New Methods of Analyzing Visual Media in Cultural Studies. July 2019.
- [Ki21] Kiela, Douwe; Firooz, Hamed; Mohan, Aravind; Goswami, Vedanuj; Singh, Amanpreet; Fitzpatrick, Casey A.; et al.: The Hateful Memes Challenge: Competition Report. In (Escalante, Hugo Jair; Hofmann, Katja, eds): *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*. volume 133 of *Proceedings of Machine Learning Research*. PMLR, pp. 344–360, December 2021.
- [Mc19] McCulloch, Gretchen: *Because Internet: Understanding the New Rules of Language*. Penguin, July 2019. Google-Books-ID: Tfo5DwAAQBAJ.
- [Mo20] Moßburger, Luis; Wende, Felix; Brinkmann, Kay; Schmidt, Thomas: Exploring Online Depression Forums via Text Mining: A Comparison of Reddit and a Curated Online Forum. In: *Proceedings of the Fifth Social Media Mining for Health Applications Workshop & Shared Task*. Association for Computational Linguistics, Barcelona, Spain (Online), pp. 70–81, December 2020.
- [Os15] Osterroth, Andreas: Das Internet-Meme als Sprache-Bild-Text. *IMAGE*, 22:26–46, 2015.
- [Sa20] Sadasivam, Aadhavan; Gunasekar, Kausic; Davulcu, Hasan; Yang, Yezhou: , memeBot: Towards Automatic Image Meme Generation, April 2020. arXiv:2004.14571 [cs].
- [SBW19] Schmidt, Thomas; Burghardt, Manuel; Wolff, Christian: Toward Multimodal Sentiment Analysis of Historic Plays: A Case Study with Text and Audio for Lessing’s Emilia Galotti. In (Navarretta, Costanza; Agirrezabal, Manex; Maegaard, Bente, eds): *Proceedings of the Digital Humanities in the Nordic Countries 4th Conference (DHN 2019)*. volume 2364 of *CEUR Workshop Proceedings*, CEUR-WS.org, Copenhagen, Denmark, pp. 405–414, March 2019.
- [Sc20a] Schmidt, Thomas; Engl, Isabella; Herzog, Juliane; Judisch, Lisa: Towards an Analysis of Gender in Video Game Culture: Exploring Gender specific Vocabulary in Video Game Magazines. In: *Proceedings of the Digital Humanities in the Nordic Countries 5th Conference (DHN 2020)*. Riga, Latvia, pp. 333–341, 2020.
- [Sc20b] Schmidt, Thomas; Hartl, Philip; Ramsauer, Dominik; Fischer, Thomas; Hilzenthaler, Andreas; Wolff, Christian: Acquisition and Analysis of a Meme Corpus to Investigate Web Culture. In (Estill, Laura; Guiliano, Jennifer, eds): *15th Annual International Conference of the Alliance of Digital Humanities Organizations (DH 2020)*, Conference Abstracts. Ottawa, Canada, July 2020.
- [Sc20c] Schmidt, Thomas; Mosiienko, Anastasiia; Faber, Raffaella; Herzog, Juliane; Wolff, Christian: Utilizing HTML-analysis and computer vision on a corpus of website screenshots to investigate design developments on the web. *Proceedings of the Association for Information Science and Technology*, 57(1):e392, 2020.

- [Sc21] Schmidt, Thomas; El-Keilany, Alina; Eger, Johannes; Kurek, Sarah: Exploring Computer Vision for Film Analysis: A Case Study for Five Canonical Movies. In: 2nd International Conference of the European Association for Digital Humanities (EADH 2021). Krasnoyarsk, Russia, September 2021.
- [SDW21] Schmidt, Thomas; Dennerlein, Katrin; Wolff, Christian: Emotion Classification in German Plays with Transformer-based Language Models Pretrained on Historical and Contemporary Language. In: Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature. Association for Computational Linguistics, Punta Cana, Dominican Republic (online), pp. 67–79, November 2021.
- [Sh12] Shifman, Limor: An anatomy of a YouTube meme. *New Media & Society*, 14(2):187–203, March 2012. Publisher: SAGE Publications.
- [Sh13] Shifman, Limor: *Memes in Digital Culture*. MIT Press, October 2013.
- [Sh14] Shifman, Limor: The Cultural Logic of Photo-Based Meme Genres. *Journal of Visual Culture*, 13(3):340–358, December 2014. Publisher: SAGE Publications.
- [Sh20] Sharma, Chhavi; Bhageria, Deepesh; Scott, William; PYKL, Srinivas; Das, Amitava; Chakraborty, Tanmoy; Pulabaigari, Viswanath; Gambäck, Björn: SemEval-2020 Task 8: Memotion Analysis- the Visuo-Lingual Metaphor! In: Proceedings of the Fourteenth Workshop on Semantic Evaluation. International Committee for Computational Linguistics, Barcelona (online), pp. 759–773, December 2020.
- [Sh22] Sherratt, Victoria: Towards Contextually Sensitive Analysis of Memes: Meme Genealogy and Knowledge Base. In (Raedt, Lud De, ed.): Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22. International Joint Conferences on Artificial Intelligence Organization, pp. 5871–5872, July 2022.
- [SK22] Schmidt, Thomas; Kurek, Sarah: Der Einsatz von Computer Vision-Methoden für Filme - Eine Fallanalyse für die Kriminalfilm-Reihe Tatort. In: DHd 2022 Kulturen des digitalen Gedächtnisses. 8. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum"(DHd 2022). Potsdam, Germany, March 2022.
- [SKW20] Schmidt, Thomas; Kaindl, Florian; Wolff, Christian: Distant Reading of Religious Online Communities: A Case Study for Three Religious Forums on Reddit. In: Proceedings of the Digital Humanities in the Nordic Countries 5th Conference (DHN 2020). Riga, Latvia, pp. 157–172, 2020.
- [SW21] Schmidt, Thomas; Wolff, Christian: Exploring Multimodal Sentiment Analysis in Plays: A Case Study for a Theater Recording of Emilia Galotti. In: Proceedings of the Conference on Computational Humanities Research 2021 (CHR 2021). Amsterdam, The Netherlands, pp. 392–404, 2021.
- [VU20] Vyalla, Suryatej Reddy; Udandaraao, Vishaal: Memeify: A Large-Scale Meme Generation System. In: Proceedings of the 7th ACM IKDD CoDS and 25th COMAD. CoDS COMAD 2020, Association for Computing Machinery, New York, NY, USA, pp. 307–311, January 2020.

Neue Wege für die Musikforschung

Computergestützte Analyse harmonischer Strukturen

Christof Weiß¹, Meinard Müller², Stephanie Klauk³ und Rainer Kleinertz⁴

Abstract: In dem von der DFG geförderten interdisziplinären Projekt „Computergestützte Analyse harmonischer Strukturen“ geht es um die Entwicklung computergestützter Konzepte, mit denen große Datenbestände von Musikaufnahmen verarbeitet und hinsichtlich harmonischer Strukturen analysiert werden können. Die musikwissenschaftliche Relevanz der zu entwickelnden Konzepte wird dabei paradigmatisch anhand der Analyse großer Werkzyklen wie zum Beispiel der Klaviersonaten von Ludwig van Beethoven verifiziert. In diesem Beitrag geben wir einen kleinen Einblick in ein solches computergestütztes Analyseverfahren, das die Visualisierung von tonalen Verläufen auf Basis von Musikaufnahmen ermöglicht. Hierbei werden unter Verwendung von Methoden der digitalen Signalverarbeitung und Mustererkennung geeignete Tonklasseninformationen aus den Audiodaten extrahiert und diatonische Tonalitätsverläufe abgeleitet. Die resultierenden Zeit–Diatonik-Darstellungen können schließlich visuell aufbereitet und für weitere Analysen verfügbar gemacht werden. Anhand zweier konkreter Musikbeispiele deuten wir das Potential solcher computergestützter Visualisierungsmethoden für die musikalische Korpusanalyse an.

Keywords: Musikverarbeitung, Harmonieanalyse, Signalverarbeitung, Visualisierung.

1 Einleitung

Die Analyse musikalischer Werke hinsichtlich verschiedener Aspekte wie Melodik, Harmonik oder Form stellt eine wichtige Grundlage des Erkenntnisgewinns in der Musikwissenschaft dar. Typischerweise wird eine Komposition dabei von Expert:innen unter Anwendung musiktheoretischer Konzepte beschrieben und bewertet. Für eine Systematisierung dieses Prozesses und dessen Ausweitung auf größere Werkbestände können (teil-)automatisierte Methoden von Nutzen sein. Auf Seite der Informatik hat es in den letzten zwanzig Jahren erhebliche Fortschritte bei der Entwicklung von computergestützten Methoden und Werkzeugen zur Erschließung und Analyse von Musikdaten, gegeben. Hierbei hat sich das neuartige

¹ Center for AI and Data Science, Universität Würzburg, Emil-Hilb-Weg 23, Hubland Nord (ZPD), 97074 Würzburg, christof.weiss@uni-wuerzburg.de

² International Audio Laboratories Erlangen, Am Wolfsmantel 33, 91058 Erlangen, meinard.mueller@audiolabs-erlangen.de

³ Institut für Musikwissenschaft, Universität des Saarlandes, Campus Geb. C6-3, 66123 Saarbrücken, s.klauk@mx.uni-saarland.de

⁴ Institut für Musikwissenschaft, Universität des Saarlandes, Campus Geb. C6-3, 66123 Saarbrücken, rainer.kleinertz@mx.uni-saarland.de

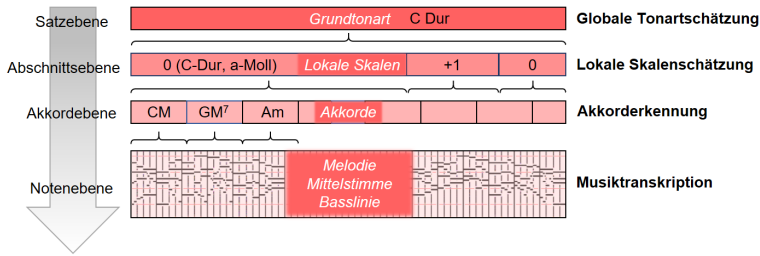


Abb. 1: Unterschiedliche Aufgabenstellungen der harmonischen Analyse bezüglich verschiedener Zeitskalen.

Forschungsgebiet der digitalen Musikverarbeitung etabliert, welches im Englischen auch unter dem Namen „Music Information Retrieval“ (MIR) bekannt ist. Vor dem Hintergrund der explosionsartigen Zunahme digitaler Inhalte, welche in unterschiedlichen Formaten und Ausprägungen gegeben sein können, hat dieses Forschungsgebiet enorm an Bedeutung gewonnen und eröffnet damit auch neue Wege für die Musikwissenschaft. Auf einen solchen Weg wollen wir in diesem Beitrag im Kontext der Analyse harmonischer Strukturen näher eingehen.

Im Bereich des MIR nimmt die harmonische Analyse von als Audiodaten gegebenen Musikaufnahmen eine zentrale Rolle ein. Dies ist nicht verwunderlich, stellt doch die Harmonik ein grundlegendes Prinzip westlicher Musik dar und spielt eine Schlüsselrolle für die formale Gestaltung von Musikwerken. Abhängig von der Zeitskala, also dem betrachteten zeitlichen Kontext, führt die harmonische Analyse zu unterschiedlichen MIR-Aufgabenstellungen, angefangen von der Bestimmung der Grundtonart eines gesamten Stückes oder eines Satzes [NS06], über die Schätzung lokaler Tonarten [WSM20] und Skalen [WH14], bis hin zur Extraktion von Akkordfolgen [CB14; KMK13; Pa19]. Auf noch feineren zeitlichen Skalen geht die harmonische oder tonale Analyse nahtlos in Fragestellungen der Musiktranskription [Be19] über. Für einen schematischen Überblick verweisen wir auf Abbildung 1. Die Analyse eines Musikstücks hinsichtlich auftretender tonaler Strukturen kann durchaus problematisch sein. Für komplexe Musikwerke sind sich Musikwissenschaftler:innen oft nicht in der Deutung harmonischer Verläufe einig. Unterschiedliche musiktheoretische Konzepte (Tonarten-, Funktionsbezeichnungen oder reduktionistische Notationen im Stile Heinrich Schenkers) liefern mehr oder weniger komplexe Beschreibungen, welche eine aufwändige Analyse erfordern und oft auch dem Spezialisten keinen intuitiven Überblick ermöglichen. Weiterhin hängt die Analyse von der jeweiligen Fragestellung, der musikalischen Stilistik sowie der betrachteten Zeitskala ab. Automatisierte Methoden zur Harmonieanalyse stoßen hier schnell an ihre Grenzen.

Anstatt eine explizite harmonische Analyse durchzuführen, geht es in diesem Beitrag um Visualisierungskonzepte, die automatisch gemessene harmonische Strukturen grafisch darstellen. Grundlage für unseren Beitrag sind die in [We17; WM21] vorgestellten Methoden zur Visualisierung von Akkorden und Skalen. Hierbei werden in einem gewissen mathema-

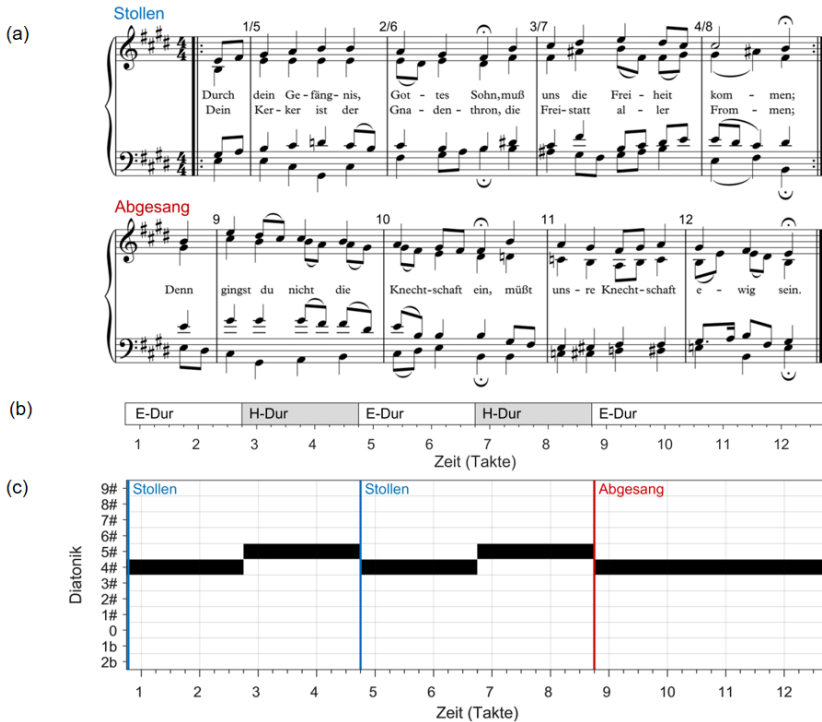


Abb. 2: Beispielhaftes Analyseszenario anhand des Chorals „Durch Dein Gefängnis“ aus der Johannespassion (BWV 245) von Johann Sebastian Bach. (a) Klavierauszug mit Formteilen, nach Walter Heinz Bernstein, Bärenreiter 1981, 18. Auflage 2012. (b) Manuelle Tonartanalyse (Beispiel). (c) Visualisierung des Tonalitätsverlaufs in Form einer Zeit–Diatonik-Darstellung.

tischen Sinne Wahrscheinlichkeiten für das Auftreten lokaler harmonischer Strukturen über die Zeit hinweg dargestellt. Eine solche Visualisierung hinsichtlich diatonischer Skalen findet man zum Beispiel in Abbildung 4 für einen Bach-Choral und in Abbildung 5 für den Kopfsatz einer Beethoven-Sonate. In Abschnitt 2 gehen wir näher auf die Visualisierungsmethode ein und zeigen dann in Abschnitt 3 anhand eines konkreten Beispiels – dem Kopfsatz einer Klaviersonate Beethovens – wie dieses Konzept zur Exploration größerer Werkzyklen eingesetzt werden kann.

2 Visualisierungskonzept

Im Folgenden beschreiben wir unser Konzept zur Visualisierung harmonischer Strukturen und fassen die wesentlichen algorithmischen Schritte zur Berechnung solcher Visualisierungen zusammen. In diesem Abschnitt folgen wir eng der Darstellung von [WM21]. Zur Illustration des Verfahrens betrachten wir beispielhaft den Choral Nr. 22 „Durch Dein



Abb. 3: Diatonische Skalen als sieben-tönige Ausschnitte der Quintenreihe, nach [WH14] und [GN02]. Für die absolute Bezeichnung diatonischer Skalen (unabhängig von der Grundtonart) orientieren wir uns an den für die Notation erforderlichen Vorzeichen.

Gefängnis“ aus der *Johannespassion* (BWV 245) von Johann Sebastian Bach. Abbildung 2a zeigt einen Klavierauszug dieses Chorals. Eine typische Aufgabenstellung liegt in der Beschreibung lokaler Tonarten und Modulationen, wie in Abbildung 2b angedeutet. Hierbei spielt die Erfassung der Tonhöhen und entsprechender musiktheoretischer Konzepte wie Akkorde, Figuration und Stimmführung eine Rolle. Eine grafische Darstellung der Analyseergebnisse kann hierbei die tonalen und formalen Verhältnisse übersichtlich veranschaulichen, was besonders für Werke mit komplexer Modulationsstruktur von Vorteil sein kann.

Als konkretes Beispiel zeigt Abbildung 2c eine Visualisierung des Tonalitätsverlaufs in diesem Choral. Die horizontale Achse dieser zweidimensionalen Darstellung entspricht einer Taktachse, die den musikalischen Zeitverlauf wiedergibt. Die vertikale Achse entspricht den zwölf diatonischen Skalen, wie sie der westlichen diatonischen Dur-Moll-Tonalität zugrunde liegen (siehe Abbildung 3). Hierbei bezieht sich zum Beispiel die Skala 1# (oder +1) auf eine Diatonik mit einem # (entspricht dem Tonvorrat von G-Dur oder e-Moll), 3b (oder -3) auf eine Diatonik mit drei b (Es-Dur oder c-Moll) und 0 auf die Diatonik ohne Vorzeichen (C-Dur oder a-Moll), welche den weißen Tasten des Klaviers entspricht. Ein Schwarzwert in der zweidimensionalen Darstellung gibt nun die jeweilige Diatonik zu einem gegebenen musikalischen Zeitpunkt an. Zum Beispiel visualisiert Abbildung 2c, dass sich der Choral in den ersten beiden Takten in der Diatonik 4# (hier E-Dur) bewegt und dann einen Schlag vor dem dritten Takt in die Diatonik 5# (H-Dur) wechselt.

Während die Visualisierung in Abbildung 2c aus der manuell erstellten Annotation in Abbildung 2b abgeleitet wurde, zeigen wir nun, wie sich eine solche Zeit-Diatonik-Darstellung automatisiert auf Basis von Musikaufnahmen berechnen lässt. Eine Alternative für eine solche Berechnung könnten auch Notentexte sein, die allerdings in einem computerlesbaren Format vorliegen müssten. Die Erstellung solcher digitaler Notentexte aus gedruckten Partituren muss dann entweder manuell erfolgen, was sehr zeitaufwändig ist, oder unter Einsatz von Software zur optischen Notenerkennung (*Optical Music Recognition*, kurz OMR), was jedoch in vielen Fällen zu stark fehlerbehafteten Ergebnissen führt [Ed21]. Ein Vorteil von Audiodaten liegt in der vergleichsweise einfachen Verfügbarkeit von Einspielungen für ein sehr breites Repertoire westlicher klassischer Musik. In den letzten Jahren hat sich der Zugang zu Musikaufnahmen durch Download- und Streamingdienste mit umfangreichen Beständen weiter vereinfacht.

Ausgangspunkt unseres Verfahrens ist also nun eine akustische Wellenform (im folgenden auch als *Signal* bezeichnet), wie in Abbildung 4a für eine Aufnahme des Bach-Chorals dargestellt. In einem ersten Schritt wird das Signal einer Spektralanalyse unterzogen [Mü21]. Dazu wird das Signal zunächst in lokale Zeitfenster unterteilt. Die Breite des Zeitfensters (gegeben in Sekunden) ist ein kritischer Parameter, der variabel gewählt und an die jeweiligen Anforderungserfordernisse angepasst werden kann (für unser Bach-Beispiel haben wir eine Breite von 4 Sekunden gewählt). Innerhalb eines Fensters werden nun die Anteile verschiedener Frequenzen berechnet, was unter anderem durch die Fouriertransformation realisiert werden kann. Dieses Zeitfenster wird nun über das Signal geschoben, sodass man für jeden Zeitpunkt eine lokale Frequenzverteilung erhält. Dies resultiert in einer Zeit–Frequenz-Darstellung, einem sogenannten *Spektrogramm*. Im Hinblick auf eine Harmonieanalyse westlicher Musik treffen wir die Annahme, dass der Tonhöhenvorrat hinreichend gut durch die zwölf chromatischen Tonhöhenklassen C, C♯, D, D♯, ..., H beschrieben werden kann. Auf Basis von Spektrogrammen ist eine enharmonische Unterscheidung von Tonhöhen wie z. B. C♯ (*cis*) und D♭ (*des*) nicht ohne Weiteres möglich. Ausgehend von einem Spektrogramm werden nun für jeden Zeitpunkt die Frequenzanteile auf die zwölf Tonhöhenklassen abgebildet und zu jeweils zwölf Koeffizienten pro Zeitpunkt zusammengefasst. Dies resultiert in einer Zeit–Chroma-Darstellung, einem sogenannten *Chromagramm*. Diese Darstellung erfasst die Energieverteilung des Musiksignals über die zwölf chromatischen Tonhöhenklassen im Verlauf der Zeit. Die Überführung von Musikaufnahmen in Chromagramme (als eine Zwischenstufe der Verarbeitungskette) ist eine grundlegende Vorgehensweise für unterschiedliche MIR-Aufgabenstellungen wie Tonartenbestimmung, Skalenanalyse und Akkorderkennung (siehe Abbildung 1). Für technische Details der zugrundeliegenden Signalverarbeitung verweisen wir auf [Mü21].

Die Betrachtung und Messung von Energieverteilungen in Tonhöhenklassen hat jedoch nur begrenzte Aussagekraft für harmonisch komplexere Sachverhalte. Sinnvollere Kategorien für diesen Zweck stellen Intervalle, Akkorde oder Skalen dar, die es durch eine geeignete Weiterverarbeitung zu erfassen gilt. Wir betrachten im folgenden die Skalenmessung bezogen auf die zwölf Diatoniken. Hierzu wird ein Chromagramm für jeden Zeitpunkt mit binären Prototypen verglichen, welche den Skalen entsprechen. Die 0-Diatonik (C-Dur, a-Moll) wird beispielsweise durch einen Prototyp modelliert, bei dem die Werte für die sieben Tonhöhenklassen C, D, E, F, G, A, H auf 1 und für die verbleibenden fünf Tonhöhenklassen C♯, D♯, F♯, G♯, A♯ auf 0 gesetzt werden. Über den lokalen Abgleich mit den zwölf unterschiedlichen Prototypen können Wahrscheinlichkeiten für das Auftreten dieser harmonischen Strukturen über die Zeit ermittelt werden. Dies führt auf eine verallgemeinerte Zeit–Diatonik-Darstellung, bei der die Wahrscheinlichkeiten über ein Graustufen-Schema visualisiert werden können (wobei Schwarz zur Wahrscheinlichkeit 1 und Weiß zur Wahrscheinlichkeit 0 korrespondiert). Abbildung 4b zeigt eine solche Visualisierung für eine Aufnahme des Bach-Chorals. Zur Hervorhebung der wichtigen Strukturen können weitere Verarbeitungsschritte erforderlich sein. Beispielsweise kann eine zeitliche Glättung oder eine Anhebung der höheren Energiewerte sowie eine Unterdrückung der niedrigen, rauschartigen Werte zu deutlicheren visuellen Strukturen führen. Letzteres wird im vorliegenden

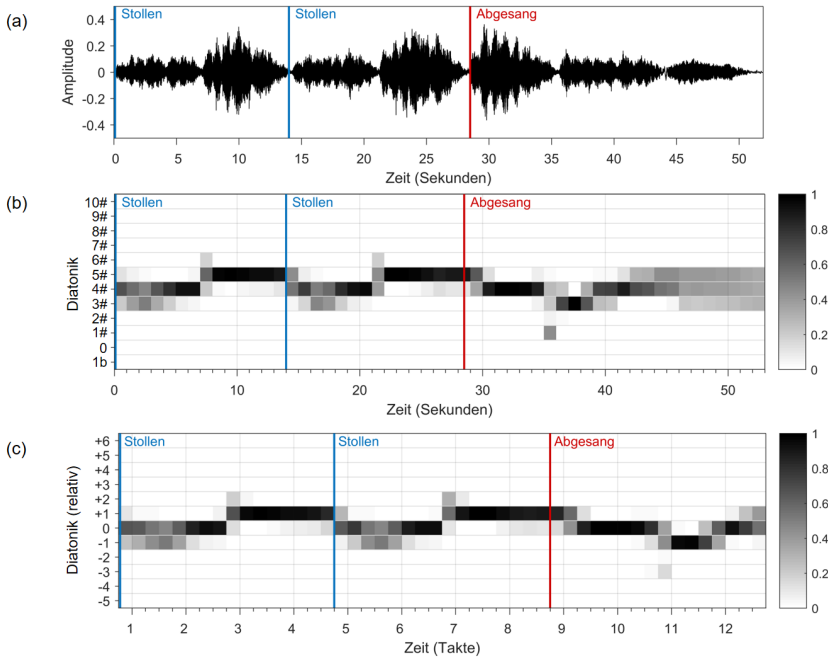


Abb. 4: Fortsetzung des Beispiels aus Abbildung 2. (a) Wellenformdarstellung einer Aufnahme unter Stephen Cleobury (Brilliant Classics 1996), mit manuell markierten Formteilen. (b) Visualisierung einer automatisch berechneten Zeit–Diatonik-Darstellung unter Verwendung einer Fensterbreite von vier Sekunden. (c) Analyseergebnis nach Konvertierung der physikalischen Zeitachse (in Sekunden) in eine musikalische Zeitachse (in Takten) und einer Ausrichtung und Bezeichnung der vertikalen Achse relativ zur Grundtonart.

Beitrag mittels einer exponentiellen Umskalierung der Energiewerte realisiert (ähnlich der *Softmax*-Funktion).

Die von einer Musikaufnahme abgeleitete Zeit–Diatonik-Darstellung (wie die Wellenform selbst) ist zunächst in physikalische Zeiteinheiten (Sekunden) gegliedert. Für manche Analysezwecke wie zum Beispiel die Bewertung von formalen Verhältnissen langer Werke kann dies eine sinnvolle Maßeinheit sein. Zum Vergleich mit Partiturdarstellungen bietet sich jedoch die Einbeziehung musikalischer Zeitinformationen an. Diese können sich auf die Form beziehen wie in Abbildung 2a oder auf die metrische Unterteilung in Takte und Schläge. Sind solche Informationen vorhanden⁵, so können die zeitlichen Bestandteile der Zeit–Diatonik-Darstellung zwischen musikalischen Zeitpunkten zusammengefasst werden und man erhält beispielsweise eine Darstellung mit Viertelnoten-Auflösung wie in Abbildung 4c. Beim Vergleich der beiden Darstellungen (b) und (c) ist die starke zeitliche Ausdehnung

⁵ Diese können entweder manuell erzeugt werden oder (mit Einschränkungen) auch unter Einbeziehung von Algorithmen, siehe hierzu [GMS10; PP11].

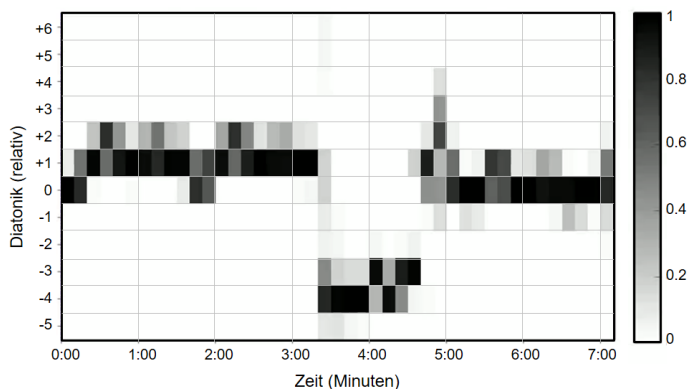


Abb. 5: Zeit–Diatonik-Darstellung von Beethovens Sonate Op. 14 No. 2, 1. Satz in einer Einspielung von Daniel Barenboim (EMI Classics 1998). Die relative 0-Diatonik entspricht der Grundtonart G-Dur.

des Schlussakkords in (b), der sich über mehrere Sekunden hinzieht, besonders deutlich sichtbar.

Zum Abschluss unserer Beschreibung stellen wir noch eine Variante der Zeit–Diatonik-Darstellung vor, die beim Vergleich verschiedener Werke hilfreich sein kann. Die Idee besteht darin, den Tonalitätsverlauf relativ zur Grundtonart zu betrachten und nach der Diatonik der Grundtonart auszurichten (relative Quintenmessung, siehe [GN02]). Dazu wird die Zeit–Diatonik-Darstellung zunächst durch zyklische Verschiebung in vertikaler Richtung auf die jeweilige Grundtonart zentriert. Die Grundtonart – im Fall des Bach-Chorals ist das die $4\sharp$ -Diatonik – wird dann als neue 0-Ebene definiert. Der Tonvorrat der Oberquinttonart ergibt sich entsprechend als +1-Ebene, die Unterquinttonart als –1-Ebene. Dieses Prinzip wird auf weiter entfernte Tonarten entlang des Quintenzirkels ausgeweitet. Die vertikale Achse wird entsprechend dieser Ebenenbezeichnungen neu beschriftet (siehe Abbildung 4c).

3 Beispiel: Beethovens Klaviersonate Op. 14 Nr. 2

Die automatisch berechneten Visualisierungen von Zeit–Diatonik-Darstellungen besitzen ein besonders Potential zur Analyse längerer Werke, was wir nun anhand einer Aufnahme eines Sonatensatzes von etwa sieben Minuten Spieldauer andeuten. Hierzu berechnen wir eine Visualisierung für den Kopfsatz aus Ludwig van Beethovens Sonate Op. 14 Nr. 2 (G-Dur) basierend auf einer Einspielung von Daniel Barenboim (EMI Classics 1998). Abbildung 5 zeigt das Ergebnis der Visualisierung, die sich in diesem Fall auf die physikalische Zeitachse der Aufnahme gemessen in Minuten und Sekunden bezieht. Um den groben harmonischen Verlauf zu verdeutlichen, ist die Fensterbreite mit etwa 20 Sekunden (und einer Auflösung von 10 Sekunden) deutlich größer gewählt als beim Bach-Choral.

In Abbildung 5 lassen sich formale Entwicklungen auf unterschiedlichen Granularitätsstufen beobachten. Zunächst fällt eine charakteristische Grobstruktur ins Auge, welche sich auf das gängige Modell der Sonatenhauptsatzform (Exposition, Durchführung, Reprise) bezieht. Die Exposition bewegt sich von der Grundtonart G-Dur (hier die relative 0-Diatonik) in die Oberquinttonart (D-Dur, +1) und wird ab Minute 1:40 wiederholt. Die Durchführung zeigt einen deutlichen Bruch und bewegt sich – als tonales Gegengewicht zur Exposition – in den relativen Minusbereich. Eine detaillierte Analyse des Notentextes zeigt, dass zunächst das erste Thema in der Varianttonart g-Moll (–3) begonnen wird, welches dann aber abbricht und sich zum zweiten Thema in B-Dur (ebenfalls –3) entwickelt. Über mehrere Modulationsvorgänge wird schließlich eine Scheinreprise in Es-Dur (–4-Diatonik) erreicht. Das Thema führt zurück nach g-Moll, worauf der Moll-Dur-Umschlag zurück in die Grundtonart (G-Dur) über den gemeinsamen Dominantseptakkord *d-fis-a-c* als „Scharnierakkord“ führt. Dieser Umschlag ist in der Visualisierung um Minute 4:50 deutlich zu erkennen. Die Reprise (ab 5:10) hält sich dann weitgehend in der Grundtonart auf, bestätigt damit also das typische Sonatensatz-Formmodell.

Neben dieser Grobstruktur ist jedoch auch die harmonisch–formale Gestaltung innerhalb der Exposition von Interesse, welche eine zentrale Fragestellung der gemeinsamen Arbeit [K121] ist. Anhand von Beethovens frühen Klaviersonaten wird hier das traditionelle Modell einer Sonatenexposition hinterfragt und der Sonatentheorie Francesco Galeazzis [K116] gegenübergestellt. Zur Unterstützung der Hypothesen werden hierbei die computergestützten Visualisierungen unterschiedlicher Granularität (durch Modifikation der Fensterbreite) herangezogen. Als eine wesentliche Erkenntnis aus dieser Arbeit wird gezeigt, dass die zeitliche Balance innerhalb der Exposition dem postulierten Themendualismus des traditionellen Sonatenmodells häufig entgegensteht. Wie in Abbildung 5 gut erkennbar ist, bewegt sich die Exposition nach dem ersten Thema zeitnah (ab Minute 0:20) von der Grundtonart weg und erreicht innerhalb kürzester Zeit das zweite Thema in der +1-Ebene, welche fortan die Exposition in harmonischer Hinsicht dominiert und in diskursiver Art weitere thematische Gedanken (welche bei Galeazzi mit Begriffen wie „passo caratteristico“, „passo di mezzo“ oder „Codetta“ bezeichnet werden) hervorbringt.

Das Beispiel der Sonate Op. 14 Nr. 2 zeigt, wie computergestützte Verfahren zur harmonischen Analyse angewendet, getestet und in Bezug zu aktuellen Forschungsfragen (z. B. hinsichtlich des Auftretens verschiedener Sonatensatz-Formmodelle) gesetzt werden. Durch das Wechselspiel zwischen Informatik und Musikwissenschaft können hierdurch einerseits bestehende Hypothesen unterstützt oder hinterfragt und andererseits aus visuellen Beobachtungen neue Hypothesen abgeleitet werden.

4 Resümee

In diesem Beitrag haben wir am Beispiel von automatisiert berechneten Visualisierungen harmonischer Strukturen aufgezeigt, inwieweit Methoden der digitalen Signalverarbeitung und Mustererkennung gewinnbringend im Bereich der Musikwissenschaft eingesetzt werden

können. An dieser Stelle wollen wir betonen, dass derartige computergestützte Konzepte in keiner Weise eine musikwissenschaftlich fundierte Analyse ersetzen können und sollen. Vielmehr sind sie als ein effizientes Werkzeug zur Exploration größerer Werkzyklen zu verstehen, mittels dessen großformatige harmonische Bezüge aufgedeckt werden können, die es dann durch eine verfeinerte manuelle Analyse zu verifizieren und zu vertiefen gilt. Eine schon in Teilen automatisierte Auswertung größerer Musikkorpora könnte zu einem Paradigmenwechsel innerhalb der historischen Musikwissenschaft führen: Bisherige Methoden würden nicht überflüssig, könnten aber auf Grund des ‚objektiv‘ vorgegebenen Befundes – der als solcher kritisch zu überprüfen wäre – mit neuen Fragestellungen konfrontiert werden. Ziel soll dabei nicht die Ablösung der historischen Musikwissenschaft durch eine wie auch immer zu definierende „Musikinformatik“ sein, sondern vielmehr ein Dialog zwischen beiden Disziplinen auf der Basis ihrer unterschiedlichen Voraussetzungen und Methoden.

Danksagung

Wir danken der Deutschen Forschungsgemeinschaft (DFG) für die finanzielle Unterstützung (MU 2686/7-2, KL 864/4-2). Die International Audio Laboratories Erlangen sind eine gemeinsame Einrichtung der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und des Fraunhofer Instituts für Integrierte Schaltungen IIS.

Literatur

- [Be19] Benetos, E.; Dixon, S.; Duan, Z.; Ewert, S.: Automatic Music Transcription: An Overview. *IEEE Signal Processing Magazine* 36/1, S. 20–30, 2019.
- [CB14] Cho, T.; Bello, J. P.: On the Relative Importance of Individual Components of Chord Recognition Systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22/2, S. 477–492, 2014.
- [Ed21] Edirisooriya, S.; Dong, H.; McAuley, J. J.; Berg-Kirkpatrick, T.: An Empirical Evaluation of End-to-End Polyphonic Optical Music Recognition. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. S. 167–173, 2021.
- [GMS10] Grosche, P.; Müller, M.; Sapp, C. S.: What makes beat tracking difficult? A case study on Chopin Mazurkas. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Utrecht, The Netherlands, S. 649–654, 2010.
- [GN02] Gárdonyi, Z.; Nordhoff, H.: *Harmonik*. Mösel, Wolfenbüttel, 2002, ISBN: 9783787730353.

- [K116] Kleinertz, R.: Streichquartette von Joseph Haydn und Wolfgang Amadeus Mozart im Spiegel der Sonatentheorie Francesco Galeazzis. In (Speck, C., Hrsg.): *The String Quartet: From the Private to the Public Sphere*. Brepols, Turnhout, S. 297–316, 2016.
- [K121] Klauk, S.; Kleinertz, R.; Weiß, C.; Müller, M.: ‚Seitensatz‘ versus ‚Mittelsatz‘: Expositionen in Beethovens frühen Klaviersonaten zwischen zeitgenössischer Theorie und computergestützter Analyse. In (Hohmaier, S., Hrsg.): *Jahrbuch des Staatlichen Instituts für Musikforschung – Preußischer Kulturbesitz 2017*. Schott Music, Mainz, S. 271–300, 2021.
- [KMK13] Konz, V.; Müller, M.; Kleinertz, R.: A Cross-Version Chord Labelling Approach for Exploring Harmonic Structures—A Case Study on Beethoven’s *Appassionata*. *Journal of New Music Research* 42/1, S. 61–77, 2013.
- [Mü21] Müller, M.: *Fundamentals of Music Processing – Using Python and Jupyter Notebooks*. Springer Verlag, 2021, ISBN: 978-3-030-69807-2.
- [NS06] Noland, K.; Sandler, M. B.: Key Estimation Using a Hidden Markov Model. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Victoria, Canada, S. 121–126, 2006.
- [Pa19] Pauwels, J.; O’Hanlon, K.; Gómez, E.; Sandler, M. B.: 20 Years of Automatic Chord Recognition from Audio. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Delft, The Netherlands, S. 54–63, 2019.
- [PP11] Papadopoulos, H.; Peeters, G.: Joint Estimation of Chords and Downbeats From an Audio Signal. *IEEE Transactions on Audio, Speech, and Language Processing* 19/1, S. 138–152, 2011.
- [We17] Weiß, C.; Zalkow, F.; Müller, M.; Klauk, S.; Kleinertz, R.: Versionsübergreifende Visualisierung harmonischer Verläufe: Eine Fallstudie zu Wagners Ring-Zyklus. In: *Proceedings of the Jahrestagung der Gesellschaft für Informatik (GI)*. Chemnitz, Germany, S. 205–217, 2017.
- [WH14] Weiß, C.; Habryka, J.: Chroma-Based Scale Matching for Audio Tonality Analysis. In: *Proceedings of the Conference on Interdisciplinary Musicology (CIM)*. Berlin, Germany, S. 168–173, 2014.
- [WM21] Weiß, C.; Müller, M.: Computergestützte Visualisierung von Tonalitätsverläufen in Musikaufnahmen. Möglichkeiten für die Korpusanalyse. In (Klauk, S., Hrsg.): *Instrumentalmusik neben Haydn und Mozart. Analyse, Aufführungspraxis und Edition*. Königshausen & Neumann, S. 107–130, 2021.
- [WSM20] Weiß, C.; Schreiber, H.; Müller, M.: Local Key Estimation in Music Recordings: A Case Study Across Songs, Versions, and Annotators. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 28/, S. 2919–2932, 2020.

An Analysis of Automatically Generated Music

Andrew McLeod¹

Abstract: In recent years, there has been an explosion of research into the automatic generation of music, both audio and symbolic. Countless deep learning approaches in particular have been proposed, using a wide range of methods and producing an equally wide range of outputs. However, the evaluation of such generations is very difficult, as the gold standard method of evaluation (listening experiments with musically-trained test participants) is expensive, in terms of both time and money (assuming the participants are fairly compensated), particularly when an extensive comparative evaluation is desired. Recent work [Yi23] has undertaken such a procedure, releasing human expert ratings and generated examples comparing human compositions to automatic compositions by several methods. We take the same generations (MIDI files of classical string quartets and piano improvisations), and analyze them instead statistically, comparing properties such as rhythmic density and pitch range across each of the methods and styles. We make no claim that our analysis represents an evaluation of the selected methods, but present our findings as an exploratory look at musically-relevant statistical properties of the outputs of each method, and draw conclusions based on that.

Keywords: Music Generation, Analysis

1 Introduction

The automatic generation of music has been subject to research for many decades. However, in recent years, as deep learning methods became the norm, larger datasets were produced (e.g. MAESTRO [Ha19]), and greater memory became available for training, focus on the task has increased exponentially [BP20]. It should be noted that music generation is a very broad topic, containing within it a wide variety of tasks, including different output formats, and various levels of control and conditioning. In this work, we focus specifically on non-constrained generation of symbolic music, where the style and form of each output is dependent only on the data used to train each model.

Regardless of the type of generation, evaluation including comparison against existing work is not a simple task, due to both the huge number of proposed models and the inherent difficulty of evaluating generated content [YL20]. The best evaluation method is a listening experiment with trained human participants rating the quality of generations, but such an experiment is extremely time consuming and costly (if the participants are compensated).

One recent paper performed such an evaluation, comparing five generation models across two musical styles [Yi23]. The musical styles are Western classical quartets (abbreviated

¹ Fraunhofer IDMT, Ehrenbergstraße 31, 98693 Ilmenau, Germany andrew.mcleod@idmt.fraunhofer.de

as CSQ here and in [Yi23]), lacking microtiming and tempo variation; and classical piano improvisations (abbreviated CPI here and in [Yi23]), performances taken from the MAESTRO dataset [Ha19] which represent live performances of piano compositions, containing expressive timing. We refer the reader to the original publication for further details about the two sets of data. The authors first generated 25 excerpts of 20–30s in each style by each relevant model (some were unable to generate in one of the styles). They then conducted listening experiments with trained human participants. The stimuli were all released publicly for future research.

Here, we take those stimuli and perform a different type of analysis. Rather than trying to estimate the quality of each generation, we measure musically-significant statistical properties (diatonicity, rhythmic texture, etc.). It is important to be clear that this is not a qualitative evaluation, nor an evaluation of any kind. We simply present statistics and discuss similarities and differences between the models rather than remarking on model performance or generation quality directly. Due to lack of space, we do not describe the models (MaMa [CL17], CoRe [Th18], MVAE [Ro18], MuTr [Hu18], LiTr²) here, and rather refer readers to the original paper for an overview.

2 Analysis

This section presents our analyses for each of the models in each style. CSQ includes two sets of 25 generations from human composers: one set from Haydn, Mozart, and Beethoven (Orig) matching the training style, and one set from Vivaldi and Brahms (BeAf) from before and after the era of the training data. CPI includes a set of 25 human compositions (Orig) matching the training style. We divide our analyses into two categories: rhythmic, which concern features related to note onset and offset positions; and pitch-based, which involve the pitch of each generated note. With the exception of LiTr (which was used pre-trained on a large corpus of scraped data), all models were trained from scratch on data in each style.

We visualize the analyses in violin plots. The white dots represent the median, while the thicker black bars range from the first to the third quartile of the data. The thickness of the colorful shape represents the distribution of values, smoothed using kernel density estimation. For each analysis, a generation is a single point in the distribution (e.g., for average duration, each point is the average duration of all notes in a single generation).

2.1 Rhythmic

First, we measure the average polyphony level of each generation (i.e., the average number of simultaneous notes at each point in time in each generation). The results are plotted in Figure 1, where it is clear that, for CSQ, MaMa, MVAE, and MuTr model the original data

² <https://magenta.github.io/listen-to-transformer>

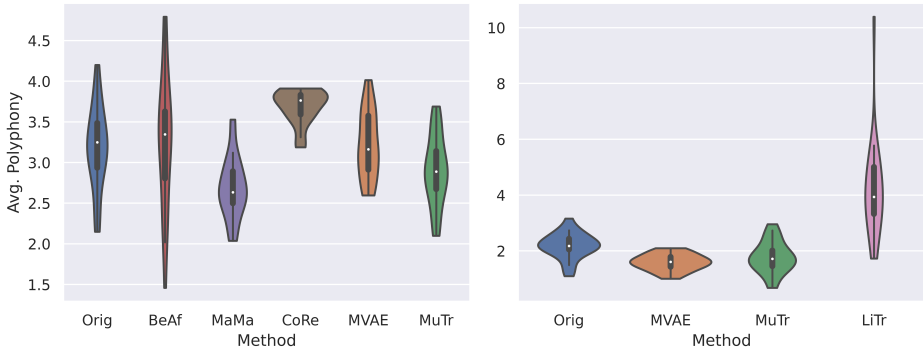


Fig. 1: The average polyphony level of each model on CSQ (left) and CPI (right).

most closely in this regard, with CoRe generating more music towards the high end of the training data’s polyphonic. CoRe’s explicit modeling of each voice may be the cause of this, as it is less likely to generate silences within each voice than a model without explicit voices. For CPI, both MVAE and MuTr modeled the training data’s polyphony quite well, with MuTr again slightly closer to the original distribution. LiTr is a clear outlier, generating pieces with much greater polyphony due to its training data.

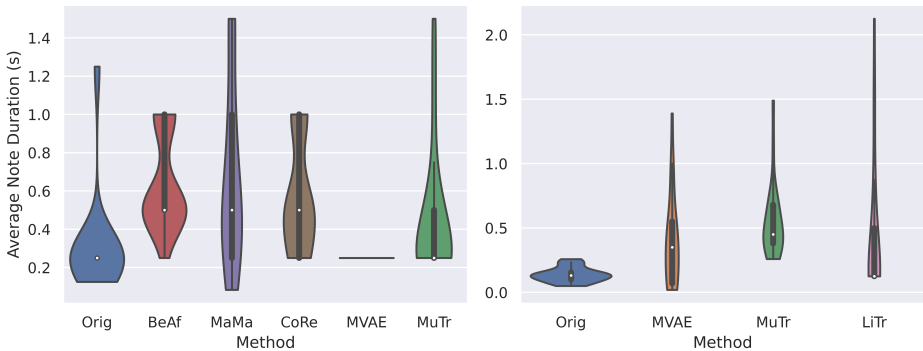


Fig. 2: The note duration of each model on CSQ (left) and CPI (right).

We next present the average note duration for each method in Figure 2. For CSQ, a quarter note is 0.5 seconds, and it can be seen that by far the most common average duration in the original style is around an eighth note (and MVAE produced only music with exactly that average duration), although there is a long tail towards higher durations. MaMa produced the widest range of average durations, roughly evenly distributed across the spectrum from 16th notes to dotted half notes, while none of the other methods generated a single piece whose average duration was less than a quarter note. For CPI, the original dataset contains

much shorter average durations (with a maximum of around 0.25 seconds) than all of the models, which each produced a wide range of average durations, up to over 1 second.

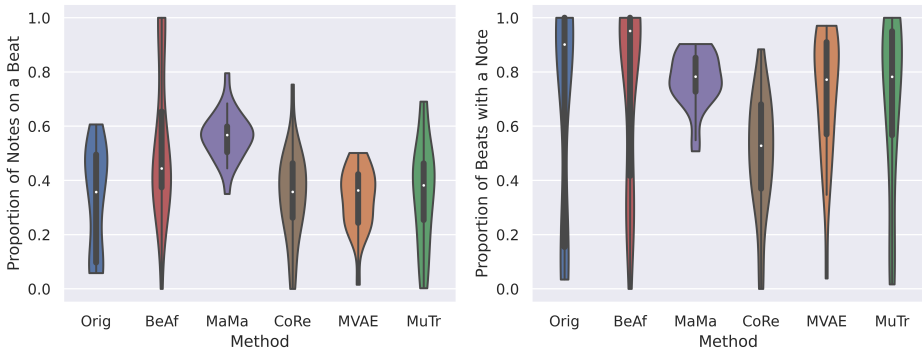


Fig. 3: The average proportion of notes that lie on a beat (left) and the average number of beats which have at least one note onset (right) for each model for the CSQ data.

Finally, in Figure 3, we present two features which are only informative for outputs which don't include expressive timing (in our case, CSQ), and generally describe the metrical density and regularity of the generation. In these plots, MaMa is the clear outlier, generating a much greater proportion of its notes on beat, and tending to produce only generations where around 80% of beats have at least one note. It's possible that, similar to CoRe with polyphony, the explicit modeling of beat position in MaMa's state space induces this property, whereas the other models all exhibit distributions closer to the original style.

2.2 Pitch-based

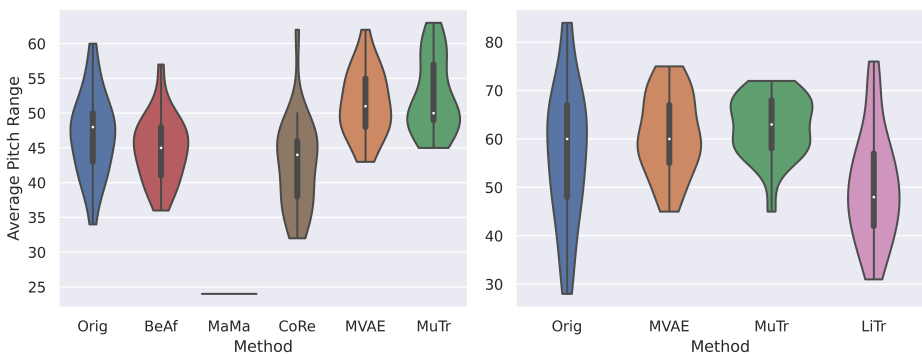


Fig. 4: The average pitch range of each model on CSQ (left) and CPI (right).

We first measure the average pitch range of each method in Figure 4. For CSQ, we see that MaMa is again an outlier, generating only music with a pitch range of 25. Again, this is likely due to a limitation of the model itself. For the deep learning methods, MVAE and MuTr both tend to generate music with slightly wider pitch ranges than the original data, while CoRe (which explicitly models the relationship between the notes in each voice), matches the training distribution more closely. It seems that CoRe’s modeling strategy has helped it to implicitly capture a pitch range dependency that the less constrained models missed. For CPI, the generations all lie within the range of the original data, although MVAE and MuTr are again more towards the upper-end of the distribution. In the case of pitch range, unconstrained models do not yet capture the dependence between high and low pitches that CoRe explicitly models.

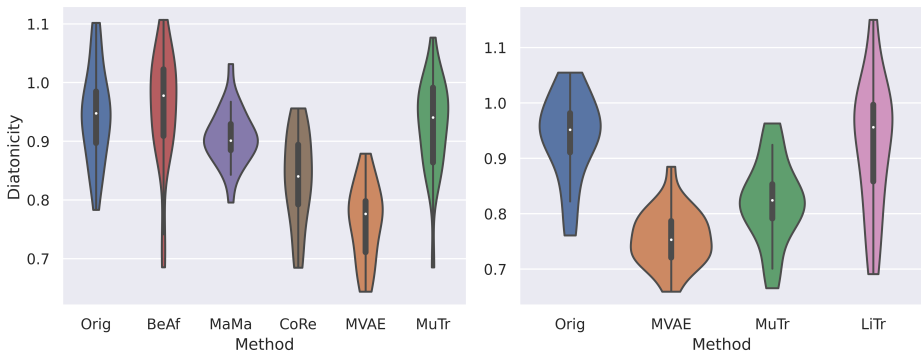


Fig. 5: The estimated diatonicity (based on overlap with pitch-class-profile templates from [TM08]) of each model on CSQ (left) and CPI (right).

Finally, we measure the level of diatonicity of each method, shown in Figure 5, by first measuring each piece’s normalized pitch class profile (PCP). That PCP is then multiplied by the major and minor template PCPs from [TM08], rotated from 0 to 11 times to model different tonics. The maximum of these 24 products is taken as a piece’s diatonicity. A PCP that exactly matches one of the rotations of a template would have a value of 1, while larger values represent pieces that tend to use more of the most common pitches of a template. Here, MVAE (and to a lesser extent CoRe) is a clear outlier for both CSQ and CPI, generating music that is less diatonic than the original style. The ability to stay within a tonal center relies on long-term modeling of pitch, something that the RNNs of CoRe and MVAE tend to struggle with compared to Transformers. MaMa models pitches explicitly relative to an estimated tonic, and thus needs no long-term model to remain in the same key.

3 Conclusion

In this paper, we have presented a musically-informed statistical analysis of some of the state-of-the-art systems for music generation. We used the generations produced and

released by [Yi23], covering both quantized outputs (i.e., compositions) in the style of Classical string quartets, and expressive performances generated in the style of classical piano improvisations. We measured properties of the resulting generations, producing both rhythmic and pitch-based analyses.

In general, we found that the generated examples tend to match the original data in the measured features. However, there were some interesting exceptions. Average duration was the most different for all methods compared to the original distribution, showing that perhaps rhythmic properties of the music are more difficult for the current models to learn. We also showed that for MVAE and CoRe (both forms of RNN for which long-term dependence is difficult to model), remaining in the same key for the entire generation was uncommon, which points to a potential drawback of such architectures. Finally, in two cases where a method explicitly models one aspect of the music (CoRe for polyphony and MaMa for metric regularity), they produced outputs that were skewed towards one end of the training distribution, suggesting that such explicit modeling, though helpful in some cases (e.g., MaMa’s diatonicity), can also lead to an over-regularity of that feature in the generations.

Bibliography

- [BP20] Briot, Jean-Pierre; Pachet, François: Deep learning for music generation: challenges and directions. *Neural Computing and Applications*, 32(4):981–993, 2020.
- [CL17] Collins, Tom; Laney, Robin: Computer-generated stylistic compositions with long-term repetitive and phrasal structure. *Journal of Creative Music Systems*, 1(2), 2017.
- [Ha19] Hawthorne, Curtis; Stasyuk, Andriy; Roberts, Adam; Simon, Ian; Huang, Cheng-Zhi Anna; Dieleman, Sander; Elsen, Erich; Engel, Jesse; Eck, Douglas: Enabling Factorized Piano Music Modeling and Generation with the MAESTRO Dataset. In: *ICLR*. 2019.
- [Hu18] Huang, Cheng-Zhi Anna; Vaswani, Ashish; Uszkoreit, Jakob; Shazeer, Noam; Simon, Ian; Hawthorne, Curtis; Dai, Andrew M.; Hoffman, Matthew D.; Dinculescu, Monica; Eck, Douglas: Music Transformer. In: *ICLR*. 2018.
- [Ro18] Roberts, Adam; Engel, Jesse; Raffel, Colin; Hawthorne, Curtis; Eck, Douglas: A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music. In: *ICML*. pp. 6939–6954, 2018.
- [Th18] Thickstun, John; Harchaoui, Zaid; Foster, Dean P; Kakade, Sham M: Coupled recurrent models for polyphonic music composition. In: *ISMIR*. 2018.
- [TM08] Temperley, David; Marvin, Elizabeth West: Pitch-class distribution and the identification of key. *Music Perception*, 25(3):193–212, 2008.
- [Yi23] Yin, Zongyu; Reuben, Federico; Stepney, Susan; Collins, Tom: Deep learning’s shallow gains: a comparative evaluation of algorithms for automatic music generation. *Machine Learning*, pp. 1–38, 3 2023.
- [YL20] Yang, Li-Chia; Lerch, Alexander: On the evaluation of generative models in music. *Neural Computing and Applications*, 32(9):4773–4784, 2020.

Computergestützte Bildtypenanalyse durch Zero-Shot Klassifikation mit CLIP

Michael Achmann,¹ Christian Wolff¹

Abstract: Wir stellen unsere computergestützte Bildtypenanalyse mittels CLIP am Beispiel einer Analyse visueller *Social Media*-Inhalte vor und evaluieren diese. Dabei betrachten wir 2208 Instagram *Stories* aus dem Bundestagswahlkampf 2021, die in den letzten zwei Wochen des Wahlkampfs auf den Kanälen von acht Parteien und 14 Spitzenkandidierenden veröffentlicht wurden. Durch die Bildtypenanalyse konnten wir feststellen, dass der Großteil der *Stories* Wahlkampfveranstaltungen dokumentiert und in einem kleineren Teil Inhalte anderer Plattformen und Formate geteilt werden. Bei ca. einem Fünftel der *Stories* stehen sachpolitische Themen im Mittelpunkt. Die automatisierte Klassifikation hat insgesamt mittelmäßig funktioniert, ein Teil der Bildtypen konnte aber mit guter Performance klassifiziert werden. Deshalb sehen wir den Bedarf für systematisches *Prompt Engineering* und schlagen eine *Few-Shot / Ensemble* Klassifikation in zukünftigen Vorhaben zu testen, um die *Performance* über alle Bildtypen zu erhöhen.

Keywords: Visual Social Media; Zero-Shot; CLIP; Political Communication; Instagram Stories; Bundestagswahl 2021

1 Einleitung

Seit der Veröffentlichung von ChatGPT im November 2022 erregen große Sprachmodelle wie GPT-4, das am 14. März als Weiterentwicklung von GPT-3 erschien, öffentliches Interesse². OpenAI, die Entwicklerfirma hinter ChatGPT, kündigte mit diesem Modell auch multimodale Fähigkeiten, insbesondere die Aufnahme visueller *Inputs*, an. Obwohl diese Fähigkeiten aktuell nur einem kleinen Kreis zugänglich sind, hat OpenAI bereits im Januar 2021 das neuronale Netz CLIP veröffentlicht [Ra21], das Zero-Shot-Klassifikationen von Bildern durch natürliche Sprache ermöglicht. Unsere Arbeit präsentiert Experimente zur Vorbereitung auf künftig verfügbare multimodale Modelle und erste Ansätze des *Prompt Engineerings* ([Yo22]) als Ergänzung zu bestehenden *Machine Learning*-Verfahren im Bereich *Computer Vision* für die Geistes- und Sozialwissenschaften.

Wir testen den CLIP-Ansatz zur automatisierten Klassifikation von Bildtypen für Instagram *Stories* der Parteien und Spitzenkandidierenden der deutschen Bundestagswahl 2021. Bildtypen kombinieren qualitative mit quantitativen Ansätzen [GA11] und wurden schon zur Untersuchung der politischen Kommunikation in Instagram Posts verwendet [HKK21; LB17].

¹ Universität Regensburg, Lehrstuhl f. Medieninformatik, Universitätsstr. 31, 93053 Regensburg, Deutschland
vorname.nachname@informatik.uni-regensburg.de

² <https://openai.com/research/gpt-4>, Aufgerufen am 25.04.2023.

1.1 Politische Kommunikation auf Instagram & Vergängliche Stories

Die politische Kommunikation auf Instagram wurde in den letzten Jahren mit diversen Ansätzen und Fragestellungen untersucht. Bast analysierte für ihre systematische Literaturübersicht 37 Studien und ordnete sie entlang des methodologischen Ansatzes. Dabei fiel ihr auf, dass die meisten Studien versuchen zu beantworten wer Instagram, wie und mit welchem Effekt nutzt [Ba21]. Die Studien zeigen einen unterschiedlichen Aufbau, einige vergleichen Kommunikationsstrategien zwischen Plattformen, andere zwischen Accounts, einzelne fokussieren sich auf einen einzigen Account. Insgesamt zeigt sich, dass Politikerinnen und Politiker Instagram nutzen, um ein überzeugendes Bild zu schaffen, indem sie die Wahlkampforganisation oder das Treffen mit Wählerinnen und Wählern dokumentieren statt sich mit politischen Themen auseinanderzusetzen. In einer jüngeren, europaweiten Studie zur Selbstdarstellung der Europaparlaments-Kandidierenden wurden computergestützte Verfahren unter anderem zur Auswertung des Aufnahmewinkels der Bilder genutzt [HJ21]. Darüber hinaus haben Towner & Muñoz 2022 eine erste Studie zur Nutzung von Instagram *Stories* im US-Wahlkampf 2020 veröffentlicht. *Stories* sind ein besonderes Format der Instagram-Plattform, das ursprünglich von Snapchat etabliert wurde, sie sind vergänglich und laufen nach 24 Stunden ab. Die Autoren sahen aus der Marketingperspektive mehrere Schwachstellen der Kampagne, wie verpasste Chancen, nutzergenerierte Inhalte zu teilen und eine inkonsistente Nutzung der Instagramkommunikationsnormen. Sie beobachteten, dass Wahlkampfveranstaltungen und Kundgebung die *Stories* dominierten [TM22].

In Anlehnung an die Arbeiten von Towner & Muñoz [TM22] und Haßler et. al. [HKK21] sehen wir den Bedarf, die vergänglichen Instagram-Inhalte aus dem Bundestagswahlkampf 2021 näher zu analysieren, um zu verstehen, welche Inhalte von den Akteuren geteilt werden und welche Kommunikationsnormen und Nutzungsmuster im deutschen *Social Media*-Wahlkampf genutzt wurden.

1.2 Computergestützte Bildanalyse & Distant Viewing

Um die wachsende Menge an visuellen *Social Media* Inhalten verarbeiten zu können, wollen wir computergestützte Methoden zur Analyse nutzen. Arnold & Tilton [AT19] haben 2019 ein grundsätzliches Framework für das *Distant Viewing* zur Arbeit mit großen visuellen Korpora vorgestellt. Ein Anwendungsfall für das *Distant Viewing* und damit auch ein Anwendungsfall für die Übertragung der bestehenden Werkzeuge und Code-Systeme ist die Analyse ikonischer Fotos. In einer systematischen Literaturübersicht zu ikonischen Bildern und computergestützten Verfahren beschreibt van Noord [No22] den *Semantic Gap* als eine der größten Herausforderungen für die *Computer Vision* (CV), besonders auch im Bereich der Geistes- und Kulturwissenschaften. Die kulturelle Lücke beschreibt dabei die fehlende Übereinstimmung zwischen den Informationen, die man aus visuellen Daten extrahieren kann, und den Interpretationen, die dieselben Daten für kulturelle Gruppen im Laufe der Zeit haben [No22]. Um die kulturelle Lücke zu schließen, bzw. zu minimieren, schlägt van Noord die Miteinbeziehung von Kontextinformationen vor, um Unsicherheiten zu überwinden. Während die einzelnen Bildtypen im vorgestellten

Vorhaben zunächst objektiv beschreibbar erscheinen und damit zur Analyse der politischen Kommunikation computergestützt klassifizierbar wirken, gibt es auch hier den Bedarf, auf Kontextinformationen zurückzugreifen, beispielsweise die im Bild eingebetteten Texte, die mit CV-Verfahren zunächst keine Beachtung finden. Wir sehen hier Parallelen zwischen den Herausforderungen der Geistes- und Kulturwissenschaften bei der computergestützten Bildanalyse und den Sozialwissenschaften, weshalb beide Disziplinen von methodischen Weiterentwicklungen profitieren können.

1.3 Zero-Shot Klassifikation mit CLIP & Prompt Engineering

Konkret wollen wir die Bildtypen mit Hilfe des CLIP-Modells automatisiert klassifizieren. CLIP wurde im Januar 2021 vorgestellt [Ra21], seitdem wurde es in unterschiedlichen Domänen getestet und evaluiert. Yong et. al. haben sich für ihre Arbeit im Bauingenieurwesen intensiv mit dem *Prompt Engineering* und der Performanz der *Zero-*, bzw. *Few-Shot*-Klassifikation unter Nutzung des CLIP-Netztes beschäftigt. Bei der systematischen Optimierung ihrer *Prompts* für die Klassifikation von Bildern mit Gebäudedefekten erzielten sie die besten Ergebnisse durch die Nutzung ganzer Sätze, die Nutzung von Domänenwissen und die Eingabe von Beispielbildern [Yo22]. In den Digital Humanities haben Smits und Kestemont mit der binären Klassifikation von Laterna-Magica-Glasplatten aus dem späten 19. Jahrhundert begonnen. Während die *Zero-Shot*-Klassifikation für die Kategorien innen / außen vergleichsweise gut funktionierte, erzielte ein *transfer learning*-Modell bessere Ergebnisse [SK21]. In einer weiteren Arbeit klassifizieren Smits und Kestemont Bilder aus einem niederländischen Kinderbuchkorpus. Die Autoren sehen einen *multimodalen Turn* in den Digital Humanities bevorstehen, und nennen auch das *Few-Shot*-Konzept, wie im obigen Beispiel durch Eingabe der Beispielbilder in das Modell umgesetzt, als Chance zur Klassifikationsverbesserung.

Zusammenfassend wollen wir mit Hilfe der Bildtypenanalyse die Inhalte der Instagram-*Stories* aus dem Bundestagswahlkampf 2021 untersuchen. Die Analyse bietet uns einen besseren Einblick in die Nutzung vergänglicher Medien durch politische Akteure und tragen zur Untersuchung der politischen Kommunikation in den sozialen Medien bei. Um die Bilder auszuwerten, nutzen wir ein computergestütztes Verfahren, die *Zero-Shot*-Klassifikation mit dem CLIP-Modell. Wir wollen in dieser Arbeit folgende Fragen beantworten: (1) Welche Bildtypen dominieren die Instagram *Stories* der Parteien und Spitzenkandidierenden im Bundestagswahlkampf 2021? (2) Wie gut können die Bildtypen mit Hilfe des CLIP-Modells im *Zero-Shot*-Verfahren computergestützt klassifiziert werden?

2 Methoden

Wir haben vom 13. bis zum 26. September 2021 insgesamt 2208 Instagram *Stories* von acht Partei-Accounts und 14 Spitzenkandidierenden gesammelt. Neben allen im Bundestag sitzenden sieben Parteien wurden noch die Freien Wähler als achte Partei aufgenommen. Darüber hinaus wurden für alle Parteien zwei Spitzenkandidierende ausgewählt, die Union

wurde bei der Auswahl als eine Gruppe angesehen. Die *Stories* wurden täglich um 00:00 (MESZ) mit Hilfe von Selenium gesammelt, das die menschliche Betrachtung der *Stories* simulierte. Im folgenden stellen wir einen Ansatz zu computergestützten Auswertung visueller Social Media-Inhalte mit Hilfe des OpenAI CLIP-Modells vor. Zunächst gehen wir auf die Bildtypenanalyse als Methode der quantitativen Inhaltsanalyse ein und stellen unsere aus der Literatur gewonnenen Bildtypen vor. Danach gehen wir auf das CLIP-Modell und den Klassifikationsprozess ein.

2.1 Bildtypen

Die quantitative Bildtypenanalyse kombiniert das qualitativ ikonografisch-ikonologische Vorgehen der Kunstgeschichte mit der quantitativen Inhaltsanalyse aus der Sozialwissenschaft. Ziel ist es, Bildelemente und Symbole sowie Bildinhalte und Bildmotive zu erfassen und daraus Bildtypen abzuleiten, die zur Interpretation des Korpus genutzt werden können. Grittmann und Ammann stellen hierfür einen dreistufigen Ablauf vor: Zunächst findet die ikonografische Analyse statt, darauf aufbauend werden die Bildtypen gebildet und schließlich folgt die ikonologische Interpretation der Ergebnisse [GA11]. In der Kommunikationswissenschaft wurde diese Methode bereits zur Untersuchung von Instagraminhalten genutzt. Liebhart & Bernhardt nutzten den Ansatz, um das politische *Storytelling* auf Instagram am Beispiel des heutigen Bundespräsidenten Österreichs, Alexander Van der Bellen, zu untersuchen [LB17]. Haßler et al. [HKK21] bauen für Ihre Untersuchung der Bundestagswahl 2017 auf Instagram auf den Bildtypen der österreichischen Arbeit auf.

Diese bilden die Grundlage für unseren Ansatz. In einem ersten Schritt haben wir zur Vorbereitung die gesammelten Instagram *Stories* exploriert und 1104 Bilder auf Grundlage der Annotationsanleitung von Haßler et al. [HKK21] kodiert. Während die meisten Typen übertragbar sind, haben wir keine Bilder aus der Kategorie Umfrageergebnisse gesichtet und diesen Bildtyp deshalb nicht übernommen. Einzelne Typen, z. B. *Negative Campaigning* oder *Call for Action*, konnten wir nicht übernehmen, da die korrekte Klassifikation von (Text-)inhalten abhängig ist. Mit der Beschreibung der Bildinhalte für unsere Vergleichsphrasen konnten wir keine zufriedenstellende Unterscheidung vom Bildtyp *Positioning* erreichen. Außerdem haben wir die Bildtypen um *Social Media Posts* erweitert, da *Stories* auch genutzt werden, um Beiträge und Inhalte aus anderen Plattformen (z. B. Twitter) oder Medientypen (Posts) zu übernehmen. Die ausgewählten Bildtypen und zugehörigen Vergleichsphrasen sind in Tabelle 1 dargestellt.

Bei Haßler et al. [HKK21] konnte ein Bild mit mehreren Bildtypen kodiert werden. Bei der menschlichen Kodierung konnten wir diese Entscheidung gut nachvollziehen. Um die Komplexität unseres Ansatzes zu reduzieren, gehen wir allerdings im Folgenden davon aus, dass jedem Bild genau ein Bildtyp zugeordnet werden kann. Dieser Schritt scheint durchaus kompatibel mit Grittmann und Ammanns Erwartung, „dass der einzelne Bildtyp intern homogen und extern heterogen ist.“ [GA11].

2.2 CLIP & Evaluation

Ziel unseres Projekts ist die computergestützte Klassifikation von Bildtypen. Zur maschinellen Klassifikation von visuellen Medien stehen verschiedene Verfahren zur Verfügung, vor allem aus dem *Deep Learning* Bereich. In der Regel setzen diese Ansätze die Existenz eines genügend großen Trainingskorpus voraus, also einer genügend großen Zahl von annotierten Daten. Daraus ergeben sich mehrere Probleme, beispielsweise der hohe Zeit- und Kostenaufwand um visuelle Medien in ausreichender Qualität zu annotieren, sowie eine begrenzte Übertragbarkeit von Trainingsdaten für spezifische Domänen.

Wir haben den CLIP-Ansatzes als möglichen Kandidaten für die Bildklassifikation für unseren Anwendungsfall identifiziert, da das Modell sogenannte *Zero-Shot* Klassifikation erlaubt. Das CLIP-Netz wurde auf einer Vielzahl von Bild-Text-Paaren trainiert, damit erlaubt es den Vergleich zwischen einem Eingabetext und einem Eingabebild [Ra21]. Das Vergleichsergebnis mündet in der *confidence*, einer metrischen Variable, die den Grad der Übereinstimmung des Bildes mit dem Text ausdrückt. Um nun eine Klassifikation mit Hilfe des CLIP-Modells vorzunehmen, wird eine Sammlung von Vergleichsphrasen angelegt. In einer Schleife werden für jedes Bild alle Phrasen mit dem Bild verglichen, zur Klassifikation wählen wir die Phrase aus, die die höchste *confidence* für alle Bild-Phrasen-Paare aufweist. Konkret wird eine *softmax*-Funktion angewandt und die Klassifikation mit der höchsten Wahrscheinlichkeit ausgewählt. In Anlehnung an [LTG22] wurden im Anschluss an die erste Kodierung insgesamt 21 Phrasen erstellt, worin auch die Erfahrungen aus der manuellen Kodierung einfließen. Die Phrasen wurden dann zur Klassifikation genutzt.

Zur Evaluation wurden die Klassifikationsergebnisse einem geschulten Annotator in LabelStudio zur Kontrolle vorgelegt. Er hat alle vorgeschlagenen Bildtypenklassifikationen korrigiert. Wie oben erwähnt, haben vergangene Arbeiten die Kodierung mit einem oder mehreren Bildtypen vorgenommen, in unserem Fall wurde die *softmax*-Funktion genutzt, um die Phrase mit der größten Ähnlichkeit zum Bild zu identifizieren, und davon abgeleitet genau ein Bildtyp klassifiziert. Entsprechend wurde bei der Korrektur der Entscheidung des CLIP-Ansatzes Vorrang gewährt.

3 Ergebnisse & Evaluation

In unserer Studie klassifizierten wir die visuellen Inhalte von 2208 Instagram *Stories* im Bundestagswahlkampf 2021. Die Zahl der Übereinstimmungen mit den Vergleichsphrasen der CLIP-Klassifikation und der daraus abgeleitete Bildtyp sind in Tabelle 1 angegeben und den Frequenzen der manuellen Korrektur gegenübergestellt. Für die inhaltliche Auswertung wird Bezug auf die korrigierten Ergebnisse genommen.

Zur Evaluation der computergestützten Bildklassifikation wurden die menschlichen Korrekturen erfasst und *Precision*, *Recall* und *F1-Score* als Performance-Metriken berechnet. Insgesamt betrachtet erreicht das CLIP-Verfahren über alle Vergleichsphrasen lediglich

Bildtyp	Vergleichsphrase	n-Phrase C	n-Typ C	n-Typ M
Campaign events	The main subject of the photo is on stage giving a speech	328	765	888
	A photo of a crowd gathered at a political rally	41		
	A photo of an event location without a focus on people	300		
	A photo of one or more people watching TV.	21		
	A photo documenting a political campaign event	55		
	A photo showing one or more people on stage during a discussion	9		
	A travel photo	9		
Individual Voter Contact	A roadtrip photo	2		
	A photo of a politician talking face to face with constituents	225	227	106
Media work	A Selfie	2		
	The main subject of the photo is being interviewed	80	124	146
Positioning	A photo of a TV studio	44		
	A person speaking to the camera	111	519	507
Call for action	A digital collage for an election campaign	408		
	A poster or image calling to vote or an invitation to an event	0		(116)
Campaign material	A photo of flyers or other campaign material for an election campaign	0	32	49
	A photo of posters or flyers	32		
Everyday political work	A photo documenting the parliamentary work of a politician	65	90	71
	An image documenting a party convention	25		
	A screenshot of a social media post	451	451	395
NA				49
Private background story	A photo showing the private life of a politician	0	0	(3)

Tab. 1: Übersicht der genutzt Bildtypen nach Haßler et. al. mit den dazugehörigen Vergleichsphrasen. In den hinteren Spalten sind die Vorkommen der Phrasen bei Klassifikation und die Zahl der daraus abgeleiteten Bildtypen für das CLIP-Modell (C), bzw. die menschliche Kodierung (M) angegeben.

Der Bildtyp *Call for Action* wurde durch CLIP nur bei der manuelle Korrektur erfasst und entsprechend dem Bildtypen *Positioning* zugeordnet. Die wenigen *Private background stories* wurden unter NA aussortiert, vgl. Diskussion.

Bildtyp	Precision	Recall	F1
Social Media Posts	0.88	0.79	0.84
Campaign events	0.77	0.90	0.83
Positioning	0.73	0.73	0.73
Media work	0.57	0.69	0.62
Individual Voter Contact	0.89	0.42	0.57
Campaign Material	0.37	0.60	0.46
Everyday Political Work	0.39	0.33	0.36
Overall	0.66	0.64	0.63

Tab. 2: Performance-Metriken für die Bildtypenklassifikation mittels CLIP (Makro-F1) im Vergleich zum menschlich korrigierten Datensatz.

einen *Makro-F1-Score* von 0.60. Ein genauerer Blick in die Ergebnisse zeigt allerdings auch, dass sich der Klassifikationserfolg zwischen den Vergleichsphrasen stark unterscheidet: Die Phrasen *A photo of an event location without a focus on people* und *The main subject of the photo is on stage giving a speech* schnitten mit einem F1-Wert *Accuracy* 0.92 und 0.91 sehr gut ab, die Phrasen *A photo documenting the parliamentary work of a politician* (F1=0.37) und *An image documenting a party convention* (F1=0.29) hingegen sehr schlecht ab. Da die Performance für die Bildtypenklassifikation direkt von der Klassifikation anhand der Vergleichsphrasen abhängig ist, zeichnet sich ein ähnliches Ergebnis ab, siehe Tabelle 2.

Im Instagram *Story*-Wahlkampf ist somit der größte Teil der veröffentlichten Inhalte eine Dokumentation von Wahlkampfveranstaltungen. Diese Bilddokumentation zeigt sehr häufig Kandidierende auf der Bühne bei Reden und Blicke in das Publikum und Überblicke der Veranstaltungsorte. *Social Media Posts* war der zweithäufigste Bildtyp, dabei handelt es sich um *Stories*, die der Interaktion mit dem Publikum dienen, durch Teilen von Beiträgen anderer Accounts, oder Teilen von Ergebnissen über Instagram-Sticker. Darüber hinaus wurden in dieser Kategorie viele Screenshots von eigenen Beiträgen auf anderen Plattformen wie Twitter beobachtet, zum Teil überschneiden sich diese Inhalte mit dem Bildtyp *Positioning*. Dieser war der dritthäufigste, hier wurden *Stories* kategorisiert, die sachpolitische Inhalte aufgriffen und über Share-Pics, Text-Integrierte *Stories* oder Videos vermittelt wurden. Bei der *Media Work*-Kategorie konnten viele Interviews und Ausschnitte aus dem Triell als besonderer Kommunikationssituation zwischen drei Kanzlerkandidat:innen 2021 beobachtet werden, *Individual Voter Contact* fand primär bei Wahlkampfveranstaltungen statt. Aufgrund der Beobachtung in den letzten Tagen des Wahlkampfs wurden vergleichsweise wenige *Stories* des Typs *Everyday Political Work* gesichtet, hier stach die FDP mit ihrem Bundesparteitag heraus. Am seltensten wurden *Campaign Materials* geteilt, oft wurden hier Bilder von Wahlkampfständen kleinerer Ortsgruppen gezeigt.

Zusammenfassend hat die automatisierte Klassifikation der Bildtypen für die Typen am besten funktioniert, für die die detailliertesten³ Vergleichsphrasen entwickelt wurden. Durch die Klassifikation der *Stories* zeigt sich ein starker Fokus der Parteien und Kandidierenden auf der Dokumentation des Wahlkampfs und dem Teilen von Inhalten anderer Plattformen oder Kanäle, der durch die manuelle Korrektur bestätigt werden konnte. Darüber hinaus gab es (nur) bei rund einem Fünftel der *Stories* einen sachpolitischen Fokus.

4 Diskussion

Unsere Auswertung zeigt vergleichbare Ergebnisse zu den Instagram-*Stories* im U.S. Präsidentschaftswahlkampf, auch dort dominierten die Wahlkampfveranstaltungen und Kundgebungen. Towner & Muñoz [TM22] haben darüber hinaus aus der Marketing-Perspektive verpasste Chancen durch Einbindung von nutzergenerierten Inhalten gesehen,

³ Detailliert ist hier einerseits im Sinne einer möglichst präzisen Beschreibung der Bilder zu verstehen, andererseits als die Abdeckung eines Bildtyps durch mehrere Phrasen.

hier ist die große Zahl der *Social Media Posts* in unserem Korpus näher zu betrachten: Dieser Bildtyp wurde abweichend von allen anderen Bildtypen nicht aus der Literatur herausgearbeitet. Bei der Evaluation hat sich gezeigt, dass CLIP hier sehr präzise geteilte *Stories*, *Posts* und *Screenshots* von eigenen oder anderen Profilen sowie Instagram und anderen Plattformen klassifiziert hat. Bilder dieses Typs sollten in zukünftigen Arbeiten tiefergehender analysiert werden, um herauszuarbeiten, in welchem Umfang es sich um nutzergenerierten Inhalt handelt. Dies könnte teils automatisiert durch den Blick in die Metadaten erfolgen. Die *Stories* dieses Bildtyps zeigen eine weitere Limitierung des Verfahrens auf: Die Bildtypen der Literatur waren darauf ausgelegt, dass ein oder mehr Bildtypen einem Post zugeordnet werden können, ein Teil der geteilten *Social Media*-Inhalte hätte inhaltlich auch einem weiteren Bildtyp zugeordnet werden können. Ein anderer Teil der Bilder hätte allerdings auch durch den Menschen nicht spezifischer eingeordnet werden können, beispielsweise Mitmachspiele auf dem CDU-Account.

Bei der Korrektur der Ergebnisse zeichneten sich auch Muster für besonders herausfordernde Klassifikationen ab: Die Differenzierung zwischen *Call for Action* und *Positionierung* ist mit unseren Vergleichsphrasen nicht möglich gewesen. Die meisten Bilder, die eine Aufforderung zur Wahl oder Veranstaltungseinladung darstellten, wurden naheliegenderweise als „A digital collage for an election campaign“ klassifiziert. Gespräche mit Wählerinnen und Wählern wurden darüber hinaus zu oft erkannt. Außerdem konnten wir bei der Annotation der *Stories* noch Bilder finden, die nicht optimal unter den bestehenden Bildtypen eingeordnet werden können: Es gibt mehrere *Behind the Scenes*-Aufnahmen, bei denen Unterstützer:innen nach einem erfolgreichen Wahlkampftag noch etwas gemeinsam trinken. Weiterhin wurden mehrere Events von einzelnen Parteimitgliedern moderiert. Da die Personen dabei in die Kamera gesprochen haben, wurden diese Szenen als *Positionierung* klassifiziert, allerdings ist der sachpolitische Kontext hier nicht immer gegeben. Liebhart & Bernhardt [LB17] hatten auch noch den Bildtyp *Site Visits* genutzt, auch in unserem Korpus gab es mehrere Dokumentation von Besuchen bei Unternehmen und Interessenvertretungen, die durch den Bildtyp *Campaign Events* nicht exakt repräsentiert werden können – wobei die Klassifikation der Besuche und Moderation mit CLIP eine weitere Herausforderung darstellt. Im Vergleich zu früheren Arbeiten konnten wir keine Einblicke in das Privatleben der Politiker:innen beobachten, Robert Habecks *Stories* zum Besuch seines ehemaligen WG-Hauses bei der Wahlkampftour wären die einzigen Kandidaten gewesen, aufgrund der Uneindeutigkeit haben wir die Bilder nicht in die Analyse einbezogen.

Einige Bildtypen konnten nicht zufriedenstellend klassifiziert werden und gerade die Unterscheidung zwischen Bildern mit sachpolitischen Inhalten und *Call for Action* sowie die korrekte Einordnung des individuellen Wählerkontakts hat schlecht funktioniert. Darüber hinaus wurde in der vorliegenden Studie primär das Bild betrachtet, ein Großteil der gesichteten *Stories* beinhaltet Bildtext, dessen Inhalte weitere wichtige Anhaltspunkte für eine korrekte Bewertung der *Stories* geben – interessanterweise haben die Ergebnisse zum Teil den Eindruck erweckt, als ob bestimmte Stichwörter (z. B. "Plenar-") die CLIP-Klassifikationen beeinflusst haben. Wir haben lediglich einen Zwei-Wochen-Zeitraum

betrachtet, unsere Aussagen lassen sich deshalb nicht auf den gesamten Wahlkampf übertragen. Darüber hinaus haben wir nur *Stories* betrachtet und auch hier nur Bilder oder den ersten *frame* aus Videos, wobei Videos rund 56% unsere *Story*-Korpus ausmachen. Durch die Bevorzugung der CLIP-Ergebnisse bei mehrdeutigen Bildern hat unsere Evaluation einen Bias. Um die Mehrdeutigkeiten zu reduzieren könnten künftige Arbeiten Bildtexte und Bildinhalte klarer getrennt betrachtet, z. B. durch Schwärzung der Texte für CLIP und menschlichen Annotator.

Aufbauend auf unsere Ergebnisse sehen wir den Bedarf, das *Prompt Engineering* weiter zu systematisieren und für die Bildtypenklassifikation *Ensemble*- und *Few-Shot*-Ansätze zu testen und nutzen. Dabei werden neben den Suchphrasen noch einzelne Vergleichsbilder als *Input* in das Model gegeben, in der Literatur wurde die Klassifikationsperformance dadurch stark erhöht. Für eine gesamtheitliche Betrachtung der multimodalen *Social Media* Inhalte sollten die Bildtexte und Metadaten sowie Audioinhalte bei Videos für eine valide Klassifikation zusammen mit den Bildern (bzw. *frames*) genutzt werden. Abschließend sollten zukünftige Arbeiten auch die *Posts* der Akteure beinhalten.

Zusammenfassend dominiert der Bildtyp *Campaign Events* die *Stories* aus dem Instagramwahlkampf deutlich. Außerdem gibt es viele geteilte Inhalte aus anderen Medien und Formaten und ca. ein fünftel der *Stories* mit sachpolitischen Bezügen (*Positioning*). Die automatisierte Bildklassifikation mittels CLIP hat insgesamt mittelmäßig funktioniert, bei detaillierter ausgearbeiteten Bildtypen aber durchaus eine gute *Performance* erzielt. Um Bildtypen zukünftig zuverlässig mit CLIP zu klassifizieren, ist eine systematische Erarbeitung der Vergleichsphrasen notwendig, außerdem sollten *Few-Shot* & *Ensemble* Lösungen näher betrachtet werden.

Literatur

- [AT19] Arnold, T.; Tilton, L.: Distant viewing: analyzing large visual corpora. en, *Digital scholarship in the humanities* 34/Supplement_1, S. i3–i16, 2019, ISSN: 2055-7671, 2055-768X, URL: https://academic.oup.com/dsh/article/34/Supplement_1/i3/5694340.
- [Ba21] Bast, J.: Politicians, Parties, and Government Representatives on Instagram: A Review of Research Approaches, Usage Patterns, and Effects. en, *Review of Communication Research* 9/, Juli 2021, ISSN: 2255-4165, 2255-4165, URL: <https://www.rcommunicationr.org/index.php/rcr/article/view/108>.
- [GA11] Grittmann, E.; Ammann, I.: Quantitative Bildtypenanalyse. In (Petersen, T.; Schwender, C., Hrsg.): *Die Entschlüsselung der Bilder: Methoden zur Erforschung visueller Kommunikation : ein Handbuch*. von Halem, S. 163–178, 2011, ISBN: 9783869620435.

- [HJ21] Haim, M.; Jungblut, M.: Politicians' Self-depiction and Their News Portrayal: Evidence from 28 Countries Using Visual Computational Analysis. *Political Communication* 38/1-2, S. 55–74, 2021, ISSN: 1058-4609, URL: <https://doi.org/10.1080/10584609.2020.1753869>.
- [HKK21] Haßler, J.; Kümpel, A. S.; Keller, J.: Instagram and political campaigning in the 2017 German federal election. A quantitative content analysis of German top politicians' and parliamentary parties' posts. *Information, Communication and Society*/, S. 1–21, Juli 2021, ISSN: 1369-118X, URL: <https://doi.org/10.1080/1369118X.2021.1954974>.
- [LB17] Liebhart, K.; Bernhardt, P.: Political storytelling on Instagram: Key aspects of Alexander Van der Bellen's successful 2016 presidential election campaign. *Media and communication* 5/4, S. 15–25, 2017, ISSN: 2183-2439, 2183-2439, URL: <https://www.cogitatiopress.com/mediaandcommunication/article/view/1062>.
- [LTG22] Lucas, L.; Tomás, D.; Garcia-Rodriguez, J.: Exploiting the Relationship Between Visual and Textual Features in Social Networks for Image Classification with Zero-Shot Deep Learning. In: 16th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2021). Springer International Publishing, S. 369–378, 2022, URL: http://dx.doi.org/10.1007/978-3-030-87869-6_35.
- [No22] van Noord, N.: A survey of computational methods for iconic image analysis. *Digital Scholarship in the Humanities* 37/4, S. 1316–1338, 2022, ISSN: 2055-7671, URL: <https://academic.oup.com/dsh/article-pdf/37/4/1316/46607811/fqac003.pdf>.
- [Ra21] Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; Krueger, G.; Sutskever, I.: Learning Transferable Visual Models From Natural Language Supervision./, 2021, arXiv: 2103.00020 [cs.CV], URL: <http://arxiv.org/abs/2103.00020>.
- [SK21] Smits, T.; Kestemont, M.: Towards multimodal computational humanities. Using CLIP to analyze late-nineteenth century magic lantern slides. In: CEUR Workshop Proceedings. Amsterdam, Netherlands, 2021, URL: http://ceur-ws.org/Vol-2989/short_paper23.pdf.
- [TM22] Towner, T. L.; Muñoz, C. L.: A Long Story Short: An Analysis of Instagram Stories during the 2020 Campaigns. *Journal of Political Marketing*/, S. 1–14, 2022, ISSN: 1537-7857, URL: <https://doi.org/10.1080/15377857.2022.2099579>.
- [Yo22] Yong, G.; Jeon, K.; Gil, D.; Lee, G.: Prompt engineering for zero-shot and few-shot defect detection and classification using a visual-language pretrained model. en, *Computer-aided civil and infrastructure engineering*/, 2022, ISSN: 1093-9687, 1467-8667, URL: <https://onlinelibrary.wiley.com/doi/10.1111/mice.12954>.

Designing a Prototype for Visual Exploration of Narrative Patterns in News Videos

Bernhard Liebl¹, Manuel Burghardt²

Abstract: News videos play an important rule in shaping our everyday communication. At the same time, news videos use narrative patterns to keep people entertained. Understanding how these patterns work and are being applied in news videos is crucial for understanding how they may affect a videos ideological message, which is an important dimension in times of fake news and disinformation campaigns. We present Zoetrope, a web-based tool that supports the discovery of narrative patterns in news videos by means of a visual exploration approach. Zoetrope integrates a number of multimodal information extraction frameworks into an interactive visualization, to allow for an efficient exploratory access to large collections of news videos.

Keywords: visual analytics, video analytics

1 Introduction

Audiovisual media play an important role in shaping cultural communication [Mi01; Mi99]. In the realm of news media, the proliferation of video news formats on social media platforms has drastically influenced the way news are consumed [Ne22]. Given an increase in news video formats, recent events such as the COVID-19 pandemic or the war against Ukraine have also highlighted the important role news videos play in shaping the public opinion in one direction or the other. We also find a number of examples where videos are deliberately used for disinformation purposes [BL18]. In a current research project called “FakeNarratives”³, we aim to examine how exactly video news work, and how – in particular – the applied audiovisual narrative strategies may affect its ideological message. Examples for such narrative patterns would be emotionalization or individualization [Ts23]. Although news reporting as narrative is well-established, analyzing its audiovisual properties is limited to small datasets and interpretive methods so far [Du05; MKW06; Yt01]. In the light of vast amounts of different news videos via social media platforms, more scalable methods are desperately needed. Similar developments toward automatized and computational video analytics can be observed in the digital humanities field, where we find conceptions such as “distant viewin” [AT19] and “deep watching” [Be19], and even the proclamation of a “visual

¹ Leipzig University, Computational Humanities Group, Augustusplatz 10, 04109 Leipzig, Germany
bernhard.liebl@uni-leipzig.de

² Leipzig University, Computational Humanities Group, Augustusplatz 10, 04109 Leipzig, Germany
burghardt@informatik.uni-leipzig.de

³ <https://fakenarratives.github.io/>

digital turn” [WS20]. These recent developments heavily build on the availability of neural networks, which are able to automatically extract all kinds of multimodal information from videos with great success [Sh21].

In this paper we present ongoing work that is trying to harness the advantages of state-of-the-art frameworks for video analysis to investigate the use of narrative strategies in news videos. While there is some work on narrative strategies that are being used in films [Wu18], it is not entirely clear how they are implemented for the specific case of news videos. This is why we are currently working on an exploratory visualization of news videos. Exploration is generally understood as a form of searching, but in new, unknown information spaces [He09]. [TS20] summarize this form of undirected search as “I-know-it-when-I-see-it” approach. While Exploratory Data Analysis (EDA) originates in statistics and goes back to the late seventies [Tu77], the objective of using visual exploration is to provide an overview of the data and to enable users to navigate various sections of the data interactively [To06]. For the case of news videos, the task of visual exploration is particularly challenging, as we are dealing with highly complex data objects. Nevertheless, visualizations constitute a valuable instrument to facilitate the examination of both quantitative and qualitative dimensions of videos. In particular, visualizations can assist scholars in deciphering and dissecting films and videos and aid them in uncovering noteworthy patterns and trends on a broader scale [Ma13]. Based on this premise, we introduce the Zoetrope prototype⁴, which serves as an exploratory visualization tool for news videos. Zoetrope incorporates state-of-the-art frameworks for multimodal information extraction and displays these features in an integrated, interactive visual interface.

2 Related work

Zoetrope can be contextualized in the area of what has been called visual movie analytics [Ku16]. At the same time, it joins a number of existing video analysis tools, such as the ones surveyed by [Pu20], who discuss a total of eight software tools for quantitative film studies. The same article also discusses the potential of more recent neural architectures for multimodal information extraction, which are part of the Zoetrope conception.

Although a large scale analysis of narrative patterns in news videos is still missing, there is indeed a small body of related work on more general computational approaches for news video analysis. [Qi00] present an early news video browser that combines image and text detection (OCR) as features. With NewsLab, [Gh07] present an exploratory visualization for the presentation of emerging topics in news videos. [Sa02] presents an approach for the location of identical shots in news videos, which can be useful to track topics by means of visual cues and their reuse. However, all of these related works are clearly from the

⁴ Zoetrope is currently still in development and not available to the public.

pre-deep learning era. We hope Zoetrope will be able to showcase what current state-of-the-art information extraction methods can achieve and how these features can be smoothly integrated into an interactive, exploratory visualization.

3 Zoetrope functionality

The overall task of this project is to be able to automatically detect the use of narrative strategies in news videos and to understand how they are used for the purpose of disinformation. However, before we can think about training a model that is able to perform this task automatically, we first of all have to understand what constitutes a narrative strategy in news videos, which may function rather differently than in film [Wu18]. This is where Zoetrope⁵ comes into play, as it allows media scholars to explore huge collections of news videos and to search for narrative patterns. Zoetrope currently provides a video collection overview as well as an in-depth single video exploration (see Figure 1).

3.1 Video collection overview

Upon opening the web-based application, users are provided with an overview of the whole video collection (see Figure 1, section 1). By means of a text search feature, videos can be searched and filtered by their title and tag as well as by their textual contents that have been indexed before, by means of optical character recognition (OCR) for written text and automatic speech recognition (ASR) for spoken language (sec. 2). Under each video thumbnail, a movie barcode [BKW18] summarizes the most dominant colors of each video and can also be used as a timeline to navigate the video by means of scrubbing, to get a first preview of its contents (sec. 3). Upon double-clicking on a specific video, users are directed to another screen, where they can explore the video in a more detailed view.

3.2 Single video exploration

This video analysis view provides an interactive video player (sec. 4) that can be navigated via a timeline (sec. 5). To facilitate the exploration and the discovery of narrative strategies, Zoetrope provides a number of search and navigational functions, which are organized in different tracks (sec. 6) and which can be adjusted and fine-tuned in the properties pane (sec. 7). The timeline can be zoomed in and out, to show different levels of detail in the various tracks. In the following, we will focus on the tracks that provide exploratory visualizations of various multimodal video features. These can be used to interactively browse a video from different perspectives.

⁵ A Zoetrope demo video is available here: <https://www.dropbox.com/s/0y0755zfk1bnvqt/demo-zoetrope-EuroVis-2023.mp4?dl=0>.

Text features First of all, Zoetrope provides various textual layers of a video. These include any written language, which we captured by means of OCRed scene text [Ch21], as well as spoken language, which was transcribed by means of the Mozilla DeepSpeech model⁶. Another textual layer provides captions for all of the video frames by using multimodal BLIP embeddings [Li22]. Each of the textual layers can be explored by either conducting an exact verbatim search for a term or by performing a semantically broader search using FastText embeddings. Each query will produce a new track, which contains potential results aligned with the video timeline, so users can easily see where a certain term appears in the video and navigate to this position manually. Another, rather experimental feature, builds on the relative difference between CLIP embeddings [Ra21] of adjacent frames, to calculate a score for what we call *dynamics*. The basic idea is to visualize how much a frame changes – according to its textual CLIP representation – from what was previously displayed.

Audio features For the extraction of some basic audio features, we used the librosa framework⁷. We are currently providing a track with amplitude visualization, to get an idea of the loudness, and another track with a spectrographic analysis of different frequencies in the audio signal.

Visual features For the visual features, we first of all provide a track with information on the mean color of each frame. Furthermore, we were able to extract facial information [De19; De20] from the videos, allowing us to search for and filter the video content based on specific individuals. Zoetrope provides information about the average face size, which can be used to infer the camera perspective, as well as the average face count per frame, which can be used to distinguish mass scenes from interview scenes for instance. As a further means for exploratory navigation, all of the faces are clustered in a separate track and are assigned a unique color code, to make clusters distinguishable. In the player pane we provide additional views that allow for video exploration based on face information, for instance a face network, where each node describes a unique face and each edge describes a co-appearance of two faces in one frame.

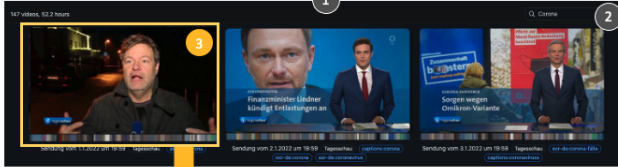
4 Implementation details

Given the complexity involved in incorporating the aforementioned features into a web-based application, this section is devoted to elucidating the technical intricacies of the Zoetrope tool. The architecture of Zoetrope consists of a backend that is fast enough to allow for real-time delivery of a zoom-level dependent, aggregated form of different portions of the data, and a web frontend for visualizing these different portions of the data interactively.

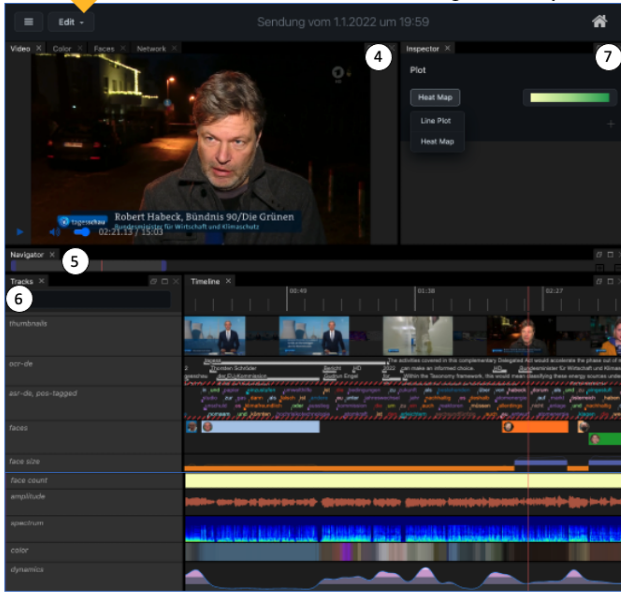
⁶ Mozilla DeepSpeech: <https://github.com/mozilla/DeepSpeech>

⁷ <https://librosa.org/doc/latest/index.html>

Video Collection Overview



Single Video Exploration



- Frame preview thumbnails
- OCR track
- ASR track
- Face cluster track
- Face size track
- Face count track
- Amplitude track
- Spectrography track
- Mean colors track
- Dynamics track

Abb. 1: Zoetrope UI – Video collection overview sections: (1) overview screen, (2) text search, (3) video scrubbing and movie barcode; Single video exploration sections: (4) player pane, (5) navigator and timeline pane, (6) tracks pane and (7) properties pane.

4.1 Aggregation

With aggregation of data we mean resampling data at a different time resolution and thereby delivering data in chunks that are suitable for a certain zoom level. For example, whereas our actual raw data might consist of one data point per video frame (e.g. 30 times per second), aggregation could resample this underlying data to produce one data point per minute only. This type of aggregation has two purposes: First, aggregation dramatically reduces the amount of data we need to transfer from the backend to the frontend. This allows very fast overview display of large amounts of underlying data, since we only load (and render) what is actually seen. Even with huge amounts of data, the memory footprint in the frontend stays small and constant. Second, and even more important, aggregation offers a similar function to that of mipmaps in classic computer graphics, which are used to reduce

aliasing effects due to sampling. They offer a solution to more faithfully display the actual underlying data. For example, in Zoetrope, a single pixel column in a color band is actually the average of all frames in the underlying time span represented by that column, and not a more or less randomly picked sample from that time span. In practice, aggregation is computed by running a series of range queries on indexed versions of the data.

4.2 Backend

One of the challenges of the backend is to provide data both fast and accurately at different zoom levels. In order to meet this goal, the following preprocessing steps were applied to our sample corpus of 274 news videos. First, numerical data (e.g. color data and scalar data such as e.g. audio amplitude or CLIP similarity) is kept in a Hierarchical Data Format (HDF) file of prefix sums. This allows for $O(1)$ range queries to obtain average values, which is the cornerstone of our implementation of the aforementioned aggregation. Second, all frames are extracted and compressed into small PNG thumbnails (stored in the file system) and color average values (stored in a HDF files). Color values are also kept as prefix sums.

Computation and presentation of data happens in a cascade of caches and pipelined modules. Various machine learning components – e.g. insightface for face recognition, different huggingface models such as CLIP, spaCy for NLP tasks, as well as low-level feature extractors such as librosa – are run either as preprocessing or on-demand on the videos on a dedicated server. Computed data is cached for each frame in an SQLite database. All data arrives fully incremental from the (cached) machine learning data algorithms through the backend database into the web frontend through a cascade of a low-latency server push technology (SSE) and web sockets, allowing the user to see partial results of data extraction or searches immediately as they are computed.

4.3 Frontend

The web frontend uses PixiJS to display all track data, including video thumbnails, plots of data, and any other timeline-based visualizations. Loading of data happens on-demand, using a tile-based system, similar to the tiled map approaches popular in the geographic information system (GIS) community⁸. In short, the navigable timeline space is split into zoom levels and tiles. Data gets loaded tile by tile, and is pre-aggregated in the backend (i.e. sized down as described before) to what the current zoom level needs for display. Our design transfers the underlying data, but not pre-rendered raster images, and is therefore most similar to GIS vector tiles. This vector tile approach enables Zoetrope to fully animate zooming for various types of tracks, giving the user the impression of a non-discrete navigation space - even though the underlying request architecture is based on discrete

⁸ See for example the OpenGIS® Web Map Tile Service Implementation Standard at http://portal.opengeospatial.org/files/?artifact_id=35326.

integer zoom levels. Also, other operations such as highlighting of elements can happen solely in the frontend without a round-trip to the backend.

Since Zoetrope displays a larger number of visually complex tracks at the same time, zooming and scrolling at interactive frame rates poses challenges. For example, our first implementation of densely populated views such as the text span view, which is used for OCR and BLIP, or the face span view, which is a mix of text spans and face images, exhibited severe lags that broke interactivity. Only after moving layout algorithms to vectorized numpy code into the backend and generating image bitmaps through a web worker that pre-renders tiles to an off-screen canvas could we achieve interactive frame rates. Similarly, horizon plots of scalar data are rendered using a custom WebGL fragment shader to drastically reduce the amount of vertex data that needs to be prepared in JavaScript for the GPU.

5 Showcase: Using Zoetrope to detect the example narrative pattern “individualization of elites”

As was described in the introduction, Zoetrope was primarily developed to support media scholars in exploring video collections in order to detect narrative patterns. In the aforementioned FakeNarratives project, we are working together with experts in multimodal analysis, who are responsible for detecting and describing narrative patterns in news videos. One of the patterns they have identified is that of “individualization of elites”. It describes how elites, such as politicians, are often introduced as antagonists of a story, typically to criticize their actions. This narrative pattern can be formalized in the following way:⁹: (1) Still image in the background of anchor’s intro, (2) elite’s name/identity in written and spoken texts, (3) after the intro, at least one shot of the specific elite, and (4) elite’s name/identity mentioned in voiceover

Zoetrope can help to identify this very pattern in a news video by searching for a specific elite’s name combined with a face track that has the elite and anchor co-appear in the same frames. A further indicator for the still image in the background might be an increased face size for this portion of the video. The values in the dynamics track are also expected to be rather stable, as we do not expect a lot of change in the CLIP embeddings of the above pattern. This way – having identified the pattern in the first place – Zoetrope provides a number of visual cues and search opportunities to identify scenes that are likely to contain individualized elites, which can then be analyzed in more detail, for instance via the integrated video player. Also note that Zoetrope offers an advanced function which allows the combination of parameter values from various tracks as a complex filter option. This means that for this specific pattern, users could directly search videos for a combination of an elite’s name, its close-up face together with the face of an anchor plus generally low dynamics.

⁹ Credits for this specific formalization go to Chiao-I Tseng.

6 Conclusions

In this paper we have presented ongoing work as part of the research project FakeNarratives. While Zoetrope was specifically developed for the scenario of news video analysis, we believe it could also be used for other types of videos, such as movies or series, and thus add to the digital humanities landscape of computational film studies and distant viewing. At this point, Zoetrope is mainly a technical feasibility study (see sec. 4), as especially the gradual zoom into different areas of the video involved some technical hurdles. At the same time, the working prototype has demonstrated that the tool can indeed help to find narrative patterns in a larger video corpus (see sec. 5). Zoetrope currently has its focus on the visualization of individual videos. Nonetheless, we aim to incorporate a macro-perspective that would enable us to visualize larger collections of videos, allowing gradual zooming in and out, as per Shneiderman's well-known design mantra [Sh96]. By doing so, we aspire to create an exploratory video analysis tool that can scale beyond the scope of a single video [BKW18; Hö15] and help to discover narrative patterns across multiple videos.

7 Acknowledgements

We are grateful to the Federal Ministry Education of Research for funding this research as part of the project "FakeNarratives – Understanding Narratives of Disinformation in Public and Alternative News Videos"(16KIS1516). We are also grateful to Chiao-I Tseng and John Bateman for providing valuable feedback on Zoetrope and for their definition of narrative patterns used in news video.

Literatur

- [AT19] Arnold, T.; Tilton, L.: Distant viewing: analyzing large visual corpora. Digital Scholarship in the Humanities 34/Supplement_1, S. i3–i16, 2019.
- [Be19] Bermeitinger, B.; Gassner, S.; Handschuh, S.; Howanitz, G.; Radisch, E.; Rehbein, M.: Deep Watching: Towards New Methods of Analyzing Visual Media in Cultural Studies. Book of Abstracts, DH Conference 2019, Utrecht/, 2019.
- [BKW18] Burghardt, M.; Kao, M.; Walkowski, N.-O.: Scalable MovieBarcodes—An Exploratory Interface for the Analysis of Movies. In: IEEE Vis4DH Workshop. 2018.
- [BL18] Bennett, W.L.; Livingston, S.: The disinformation order: Disruptive communication and the decline of democratic institutions. European journal of communication 33/2, S. 122–139, 2018.
- [Ch21] Chen, X.; Jin, L.; Zhu, Y.; Luo, C.; Wang, T.: Text recognition in the wild: A survey. ACM Computing Surveys (CSUR) 54/2, S. 1–35, 2021.

- [De19] Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S.: Arcface: Additive angular margin loss for deep face recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. S. 4690–4699, 2019.
- [De20] Deng, J.; Guo, J.; Ververas, E.; Kotsia, I.; Zafeiriou, S.: Retinaface: Single-shot multi-level face localisation in the wild. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. S. 5203–5212, 2020.
- [Du05] Dunn, A.: Television news as narrative. *Narrative and media/*, S. 140–152, 2005.
- [Gh07] Ghoniem, M.; Luo, D.; Yang, J.; Ribarsky, W.: Newslab: Exploratory broadcast news video analysis. In: 2007 IEEE Symposium on Visual Analytics Science and Technology. IEEE, S. 123–130, 2007.
- [He09] Herczeg, M.: Interaktionsdesign. In: *Interaktionsdesign*. Oldenbourg Wissenschaftsverlag, 2009.
- [Hö15] Höferlin, B.; Höferlin, M.; Heidemann, G.; Weiskopf, D.: Scalable video visual analytics. *Information Visualization* 14/1, S. 10–26, 2015.
- [Ku16] Kurzhals, K.; John, M.; Heimerl, F.; Kuznecov, P.; Weiskopf, D.: Visual movie analytics. *IEEE Transactions on Multimedia* 18/11, S. 2149–2160, 2016.
- [Li22] Li, J.; Li, D.; Xiong, C.; Hoi, S.: Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In: *International Conference on Machine Learning*. PMLR, S. 12888–12900, 2022.
- [Ma13] Manovich, L.: Visualizing Vertov. *Russian Journal of Communication* 5/1, S. 44–55, 2013.
- [Mi01] Mitchell, W.: “What is visual culture?” In: *Meaning in the visual arts: views from the outside*. Princeton: Institute for Advanced Study, S. 202–217, 2001.
- [Mi99] Mirzoeff, N.: *An Introduction to Visual Culture*. London & New York: Routledge, 1999, ISBN: 9780415158763.
- [MKW06] Machill, M.; Köhler, S.; Waldhauser, M.: Narrative Fernsehnachrichten: Ein Experiment zur Innovation journalistischer Darstellungsformen. *Publizistik* 51/4, S. 479–497, 2006.
- [Ne22] Newman, N.; Fletcher, R.; Robertson, C. T.; Eddy, K.; Kleis Nielsen, R.: *Reuters Institute Digital News Report 2022*. Reuters Institute & University of Oxford., 2022.
- [Pu20] Pustu-Iren, K.; Sittel, J.; Mauer, R.; Bulgakowa, O.; Ewerth, R.: Automated Visual Content Analysis for Film Studies: Current Status and Challenges. *DHQ: Digital Humanities Quarterly* 14/4, 2020.
- [Qi00] Qi, W.; Gu, L.; Jiang, H.; Chen, X.-R.; Zhang, H.-J.: Integrating visual, audio and text analysis for news video. In: *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*. Bd. 3, IEEE, S. 520–523, 2000.

- [Ra21] Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J. et al.: Learning transferable visual models from natural language supervision. In: International conference on machine learning. PMLR, S. 8748–8763, 2021.
- [Sa02] Satoh, S.: News video analysis based on identical shot detection. In: Proceedings. IEEE International Conference on Multimedia and Expo. Bd. 1, IEEE, S. 69–72, 2002.
- [Sh21] Sharma, V.; Gupta, M.; Kumar, A.; Mishra, D.: Video processing using deep learning techniques: A systematic literature review. IEEE Access 9/, S. 139489–139507, 2021.
- [Sh96] Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations. In: Proceedings 1996 IEEE symposium on visual languages. IEEE, S. 336–343, 1996.
- [To06] Tominski, C.: Event-Based Visualization for User-Centered Visual Analysis, Diss., Institute for Computer Science, Department of Computer Science und Electrical Engineering, University of Rostock, Germany, 2006.
- [TS20] Tominski, C.; Schumann, H.: Interactive visual data analysis. CRC Press, 2020.
- [Ts23] Tseng, C.-I.; Liebl, B.; Burghardt, M.; Bateman, J.: FakeNarratives – First Forays in Understanding Narratives of Disinformation in Public and Alternative News Videos. In: DHd2023. 2023.
- [Tu77] Tukey, J. W. et al.: Exploratory data analysis. Reading, MA, 1977.
- [WS20] Wevers, M.; Smits, T.: The visual digital turn: Using neural networks to study historical images. Digital Scholarship in the Humanities 35/1, S. 194–207, 2020.
- [Wu18] Wu, H.-Y.; Palù, F.; Ranon, R.; Christie, M.: Thinking Like a Director: Film Editing Patterns for Virtual Cinematographic Storytelling. ACM Trans. Multimedia Comput. Commun. Appl. 14/4, Okt. 2018, ISSN: 1551-6857, URL: <https://doi.org/10.1145/3241057>.
- [Yt01] Ytreberg, E.: Moving out of the Inverted Pyramid: narratives and descriptions in television news. Journalism Studies 2/3, S. 357–371, 2001.

Digitalisierung in der Amateurmusik

Anforderungen an ein digitales Notenarchiv für einen Musikverein

Stefan Balke,¹ Adrian Maiworm²

Abstract: Die Rolle der Notenverantwortlichen ist in einem Musikverein von großer Bedeutung. Neben der Bereitstellung der Noten für die regelmäßige Probenarbeit sind sie ebenfalls für die Archivierung des Notenbestandes verantwortlich. Ein wachsender Anteil dieses Bestandes wird heutzutage in digitaler Form erworben bzw. bei Eingang durch Einscannen digitalisiert. Die Ablage dieser digitalen Artefakte stellt viele Vereine vor eine besondere und oftmals ungewohnte Herausforderung. In diesem Beitrag wollen wir einen exemplarischen Digitalisierungsprozess im Amateurbereich am Beispiel eines Sinfonischen Blasorchesters beschreiben. Ferner legen wir typische Anforderungen aus dem Orchester an die Notenverantwortlichen dar und explorieren in einem ersten Anwendungsfall, wie diese durch algorithmische Unterstützung automatisiert werden können.

Keywords: Computergestützte Musikverarbeitung, Multimedia, Datenarchivierung

1 Einleitung

Die digitale Transformation von vormals analogen Prozessen und Abläufen schreitet in allen gesellschaftlichen Bereich voran, so auch in der Amateurmusik. Viele Musikverlage und Komponist*innen bieten ihre Werke heutzutage in digitaler Version an: Angefangen von den Notentexten als Scans³ bis hin zu Demoaufnahmen auf großen Streamingplattformen. Was sich im Endeffekt als eine Erleichterung für die Notenverantwortlichen herausstellen soll, erfordert vorab Hilfsmittel, die eine strukturierte Ablage der digitalen Artefakte (Notentexte, Audioaufnahmen, Metadaten) gewährleisten können.

Der Status Quo in den Musikvereinen lässt sich als eine Kombination von gedruckten und eingescannten Notentexten beschreiben. Im Normalfall werden lediglich Kopien an die Musizierenden herausgegeben, die Originale verbleiben im Archivschrank. Durch die lange Tradition mancher Musikvereine stammen manche Noten allerdings aus den 50er und 60er Jahren des letzten Jahrhunderts. Damals waren Fotokopierer noch nicht im heutigen Maßstab verbreitet. Aus diesem Grund wurden zwangsläufig die Originalnoten herausgegeben, was die heutige Vollständigkeit mancher Notensätze signifikant schmälert und die Noten im schlimmsten Fall für den Verein unbrauchbar bzw. für die Notenverantwortlichen extrem

¹ Hochschule Weserbergland, Am Stockhof 2, 31785 Hameln, Deutschland balke@hsw-hamelnde

² Musikzug Ennest, 57439 Attendorn, Deutschland

³ Exporte in digitalen Notentextformaten (z.B. MusicXML) sind heute noch die absolute Ausnahme.

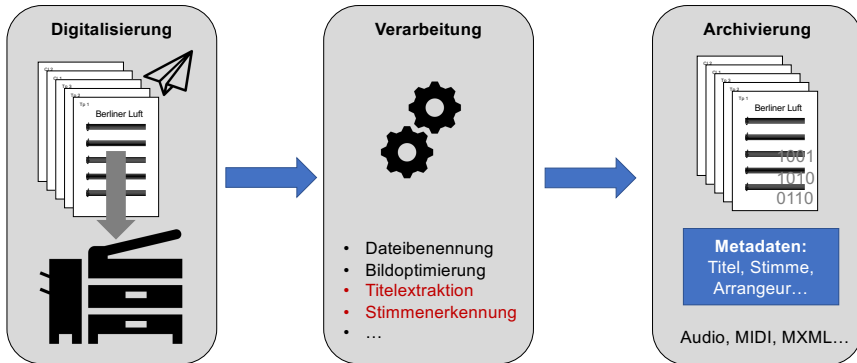


Abb. 1: Übersichtsdarstellung eines Digitalisierungsprozesses. Im ersten Schritt werden die Notentexte gescannt. In der Verarbeitung können neben der Optimierung der gescannten Bilder weitere Algorithmen, bspw. zur Metadatenextraktion angewendet werden. Die Archivierung erfolgt auf einem Speichermedium, das sowohl strukturierte Metadaten, als auch unstrukturierte Daten in Form von Scans oder Audioaufnahmen aufnehmen kann.

mühsam in der Wiederbeschaffung machen. Die vollständige Archivierung dieser existierenden Notenbestände, in Kombination mit neu erworbenen Werken, ist für viele Vereine ein Mammutprojekt und stellt die Notenverantwortlichen vor ungewohnte Herausforderungen, insbesondere wenn der Datenbestand eine gewisse Größe erreicht.

In diesem Beitrag beschreiben wir zunächst in Abschnitt 2 aus technischer Sicht ein digitales Notenarchiv eines Bläserorchesters, als beispielhafter Repräsentant eines Musikvereins. Neben dem initialen Digitalisierungsprozess ziehen wir in diesem Abschnitt einen Vergleich zur existierenden Literatur, insbesondere den Forschungsarbeiten aus dem Bereich der „Digitalen Bibliotheken.“ Abgeschlossen wird dieser Abschnitt mit einer kurzen Beschreibung der typischen Anforderungen an ein digitales Notenarchiv eines Bläserorchesters. Dies stellt den Rahmen für erste Aufgabenstellungen hinsichtlich algorithmischer Unterstützung in dieser Art von Archiven. Als exemplarische, erste Aufgabenstellung widmet sich Abschnitt 3 der automatisierten Extraktion von Metadaten wie dem Werktitel mittels Optical Character Recognition (OCR). Abschnitt 4 fasst den Beitrag zusammen und beschreibt die weiteren Schritte (z.B. Architekturauswahl, Datenstrukturierung etc.) zum Aufbau eines digitalen Notenarchives.

2 Digitales Notenarchiv eines Bläserorchesters

Wie eingangs erwähnt stellt die wachsende Digitalisierung der Notenarchive und auch die steigenden Anforderungen der Musiker*innen, Notenverantwortliche in Musikvereinen vor mitunter neue Herausforderungen. Im Folgenden beschreiben wir einen Digitalisierungsprozess, den die Autoren in der Praxis mehrfach beobachtet haben bzw. selbst praktizieren.

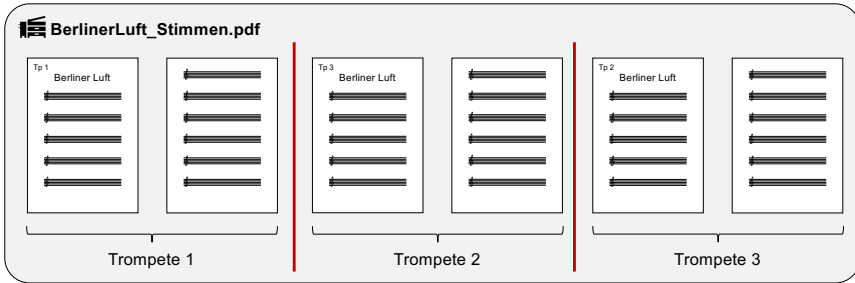


Abb. 2: Darstellung eines typischen Archivdokumentes. Alle Stimmen sind aneinandergehängt. Der Beginn einer neuen Stimmen wird durch ein Notenblatt mit Titelinformation begonnen und endet vor dem nächsten Titelblatt, dargestellt durch die roten Trennlinien.

2.1 Digitalisierungsprozess

Abb. 1 stellt einen typischen Digitalisierungsprozess eines in Papierform vorliegenden Notensatzes schematisch dar (in unserem Beispiel den Notensatz für das Stück „Berliner Luft“).⁴ Im ersten Schritt der *Digitalisierung* wird der Notentext konventionell eingescannt. Dies erfolgt im Normalfall mit Flachbettscannern oder im Falle der Partitur mit speziellen Buchscannern (falls vorhanden). In vielen Musikvereinen erfolgt das Scannen über ein bestehendes Kopiersystem. Die Scans werden hierbei direkt auf ein Speichermedium (z.B. USB-Stick) gespeichert. In diesem Fall werden die Stimmen nacheinander eingescannt und in einer zusammenhängenden Datei abgelegt. Für die entstandene Datei vergibt das Kopiersystem mit Abschluss des Scanvorgangs einen eindeutigen Namen (meistens eine Kombination aus Datum und Uhrzeit).

Im zweiten Schritt werden die Dateien vom Speichermedium an einen Computer zur weiteren Verarbeitung übertragen. Darauf folgt die Benamung der Datei. Abb. 2 zeigt den Aufbau der Datei *BerlinerLuft_Stimmen.pdf*. Wie eingangs beschrieben sind in dieser Datei alle Stimmen hintereinander abgelegt. Eine Stimme startet mit einer Titelseite und endet vor der nächsten Titelseite (Anfang der nächsten Stimme). Weitere Operationen zur Bildoptimierung können optional folgen, bspw. das Drehen der Noten oder das Beschneiden der Ränder.

Im finalen Schritt der Archivierung werden die Scans auf einem geeigneten Speichermedium archiviert. Viele Vereine verwenden dazu Cloudanbieter, damit der Zugriff später für alle Musiker*innen einfach gewährleistet werden kann. Im Sinne der Archivierung und Übersicht werden für alle gescannten Werke weitere Metadaten in strukturierter Form erfasst (z.B. in einem Tabellenkalkulationsprogramm). Zu den Metadaten gehören der Stücktitel, Komponist*in bzw. Arrangeur*in, ggf. das Genre sowie weitere archivarisches Anmerkungen (z.B. ob Stimmen fehlen). Neben den eigentlichen Scans und den dazugehörigen Metadaten

⁴ Bei digital vorliegenden Notentexten entfällt der erste Schritt des Scannens sowie Teile des zweiten Schrittes.

werden hier unter Umständen weitere Daten erfasst. Dazu zählen Demoaufnahmen der Stücke, Konzertmitschnitte, falls vorhanden MIDI-Dateien oder digitale Notentexte (z.B. MusicXML).

2.2 Literatureinordnung

Das Thema der digitalen Notenarchive ist mitnichten neu und wurde in der Vergangenheit bereits in vielen wissenschaftlichen Arbeiten und Projekten dokumentiert. Der nun folgende Literaturvergleich stellt eher einen Auszug dar und erhebt keinen Anspruch auf Vollständigkeit. Im Bereich der musikwissenschaftlich motivierten Editionsforchung sind in den letzten Jahren bereits viele Arbeiten entstanden, die sich (nicht ausschließlich) mit der Erstellung von kritischen Editionen beschäftigt. Beispielhaft sind hier die Projekte *Edirom* (<https://www.edirom.de/edirom-projekt>), *Beethovens Werkstatt* (<https://beethovens-werkstatt.de>) oder *FreischützDigital* [Rö15] genannt. Bibliotheken digitalisieren ihre Datenbestände und machen diese für die Öffentlichkeit zugänglich [Di12; Ku15; PKK18]. Von äußerst hoher Relevanz sind die Ergebnisse und Ausführungen aus dem *PROBADO* Projekt, welches sich zum Ziel gesetzt hatte, eine komplette Systemlandschaft, von der Digitalisierung, bis hin zu multimodalen Benutzerschnittstellen zu konzeptionieren und prototypisch zu entwickeln [Da12]. Weitere Literatur ist beispielsweise in den Konferenzbänden der seit 2014 jährlich stattfindenden International Conference on Digital Libraries for Musicology (DLfM) zu finden.⁵

Dieser Beitrag siedelt sich am ehesten an den Ausführungen des *PROBADO* Projektes an. Der größte Unterschied ist, dass die Perspektive eine andere ist: Wendet sich die Ausführungen des *PROBADO* Projektes (und auch der genannten Editionsprojekte) eher an Musikinteressierte oder Musikwissenschaftler*innen, so richten sich die hier präsentierten Darstellungen eher an die Notenverantwortlichen sowie die Musiker*innen und deren Anforderungen.

2.3 Anforderungen an ein digitales Notenarchiv

Die Anforderungen an ein digitales Notenarchiv in einem Musikverein sind vielfältig und wir betrachten im Folgenden nur die gängigsten. Eine der größten Herausforderungen für die Notenverantwortlichen ist es, die Musiker*innen von dem unregulierten Zugriff auf die Originalnoten abzuhalten. Ein Grund dafür ist, dass die Stimmen in Notensätzen häufig in einer bestimmten Reihenfolge vorsortiert sind (z.B. Partitur, Holz, hohes Blech, tiefes Blech, Schlagwerk, etc.). Diese Vorsortierung/Konvention ermöglicht ein schnelles Auffinden einzelner Register und Stimmen. Ist der Zugang zu den Originalnoten für alle frei, lässt sich diese Sortierung nicht lange halten, da Noten aus der Erfahrung heraus nicht

⁵ Link zu den gesammelten Proceedings: https://dl.acm.org/conference/dl_fm/proceedings.

wieder korrekt einsortiert werden nach dem Gebrauch. Im schlimmsten Fall werden für die Probe Originale entwendet und nicht wieder zurückgelegt. Die Notensätze zu erhalten ist eine Grundvoraussetzung für die Orchesterarbeit. Nicht immer sind alle Werke im aktiven Repertoire des Orchesters, werden aber zu bestimmten Anlässen immer wieder benötigt (z.B. Weihnachtslieder). Insbesondere dann erzeugen unvollständige Sätze weitere Arbeit für die Notenverantwortlichen.

Um den oben genannten unkontrollierten Zugriff auf die Originalnoten zu unterbinden, bedarf es einer einfach zu handhabenden Alternative. Der Zugang zu einem digitalen Notenarchiv muss demnach so simpel wie möglich gestaltet werden und von allen nutzbar sein. In der Praxis haben sich hierfür Cloudlösungen bewährt. Der Zugang zum Archiv kann mit Hilfe dieser Lösungen an das aktive Orchester erfolgen bzw. auch ausscheidenden Mitgliedern wieder entzogen werden.

Eine typische Probensituation verdeutlicht die nächste Anforderung: Ein/e Musiker*in ist erkrankt. Die Stimme soll aber in der Probe besetzt sein. Im Normalfall sind dann die Notenverantwortlichen gezwungen ins Archiv zu gehen, die Noten aufzusuchen und ggf. direkt eine Kopie anzufertigen. Dieser manuelle Prozess erzeugt Unruhe in der Probe und schmälert die effektive Zeit für die musikalische Probenarbeit. Insbesondere in der frühen Phase einer Konzertvorbereitung treten diese Situationen in den Proben mehrmals auf. Ein digitales Notenarchiv sollte in der Lage sein, Noten schnell auffindbar zu machen, entweder mit bereitgestellten Suchfunktionen oder einer intuitiven Ablagestruktur (gängig ist hier die Sortierung nach Werken und anschließend Stimmen).

Darüberhinaus gibt es für die musikalisch Verantwortlichen im Musikverein immer wieder Situationen, wo eine Gesamtsicht auf das Archiv von Nutzen sein kann (z.B. bei der Zusammenstellung einer Werkreihenfolge für ein Konzert). Hierzu sind weiterführende Metadaten wie Komponist*in, Arrangeur*in, Schwierigkeitsgrad, Genre oder spezielle Spielanweisungen oder benötigte Instrumente (oftmals in den Percussioninstrumenten zu finden) hilfreich in der Planung. Auch bereits gespielte Konzertreihenfolgen können Inspiration sein bzw. eine Hilfe sein, Wiederholungen zu vermeiden.

An dieser Stelle halten wir die grundlegenden Anforderungen fest:

- Komplette Notensätze vorhalten bzw. den Bestand daraufhin prüfen.
- Einfacher Zugang von allen Musiker*innen von möglichst überall.
- Schnelle Auffindbarkeit einzelner Stimmen um die Probenarbeit nicht unnötig zu verzögern.
- Alle Metadaten im Überblick bereitstellen, ggf. auch Informationen darüber, wann ein Stück gespielt wurde.

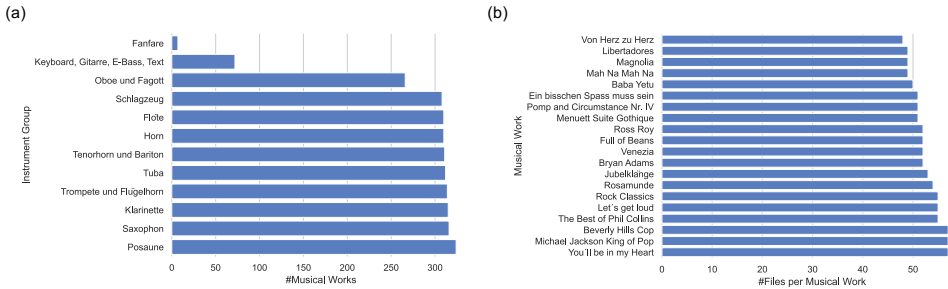


Abb. 3: Übersichtsstatistiken des verwendeten Datensatzes. (a) Zeigt die Anzahl der musikalischen Werke pro Register. (b) Stellt die Anzahl der Dateien/Stimmen pro musikalischem Werk dar.

3 Proof of Concept: Metadatenextraktion mittels OCR

Viele der oben genannten Anforderungen können durch ein geeignetes, relationales Datenbankschema und einer sortierten Ablage für die nicht-relationalen Daten realisiert werden. Problematisch ist, dass insbesondere die alten Notenbestände nur als Scans vorliegen und ggf. auch in völlig unsortierter Form. Im Folgenden beschreiben wir ein erstes Experiment zur Metadatenextraktion mittels Optical Character Recognition (OCR), das wir auf einem digitalen Notenarchiv eines Musikvereins durchgeführt haben. Es handelt sich bei dem Musikverein um ein Sinfonisches Blasorchester, das sowohl konzertant auftritt, als auch im Zuge der Brauchtumpflege (z.B. Volksfeste oder kirchliche Anlässe). Im Wesentlichen bedeutet das für den Datensatz, dass sowohl Noten im DIN A4, als auch im Format DIN A5 vorliegen. Die Forschungsfrage ist, inwiefern existierende, modellbasierte OCR-Programme in der Lage sind, die textuellen Daten in einem gemischten Dokument aus Text und Notentext, zu extrahieren. Als OCR-Programm verwenden wir für diese Experimente das Pythonpaket EasyOCR, da es sehr einfach in der Nutzung ist und in Vorexperimenten gute Ergebnisse erzielt hat.⁶ EasyOCR verwendet zwei hintereinander geschaltete neuronale Netze. Das erste Modell erkennt die relevanten, textuellen Felder im Bild (Detection). Das zweite Modell wandelt diese in Text um (Recognition). Wir nehmen an, dass der Notentext als solcher eher störend für das Modell wirken wird, da diese Art von Dokumenten typischerweise nicht im Training verwendet wird.⁷

Unser betrachteter Datensatz umfasst insgesamt 351 musikalische Werke, die sich auf 12.403 Einzelstimmen (z.B. 1. Trompete, 2. Klarinette, 4. Posaune, etc.) verteilen.⁸ Die Einzelstimmen bilden zusammengefügt ein Dokument mit 17.372 Seiten Notentext. Spezialstimmen wie Partituren oder Einzelstimmen mit mehreren Werken pro Seite wurden im Vorfeld bestmöglich aussortiert. In Abb. 3 sind zwei exemplarische Übersichtsstatistiken

⁶ <https://github.com/JaidedAI/EasyOCR>

⁷ Die Implementierung der Experimente kann unter <https://github.com/stefan-balke/sheet-music-archive-GI> eingesehen werden.

⁸ In diesem Datensatz sind die Einzelstimmen in separaten PDFs abgelegt. Die Anzahl der Einzelstimmen entspricht demnach der Anzahl der Dateien.

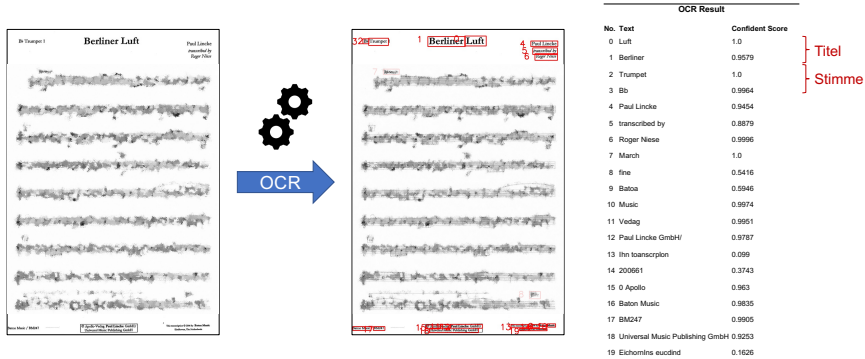


Abb. 4: Darstellung der Verarbeitungspipeline. Das Deckblatt der Einzelstimme wird geöffnet und mittels eines OCR-Programmes verarbeitet. Das Ergebnis des Detection Modells besteht aus Bounding Boxes (rote Rechtecke) und den dazugehörigen, extrahierten Texten als Ergebnis des Recognition Modells (Tabelle). (Aus lizenzrechtlichen Gründen ist der eigentliche Notentext für die Publikation unkenntlich gemacht worden.)

dargestellt. Abb. 3a zeigt die Anzahl der musikalischen Werke pro Register.⁹ Die typisch anzutreffenden Register eines Sinfonischen Blasorchesters weisen alle eine Werkanzahl >300 auf. Bei dem Register der Doppelrohrbläser (Oboe und Fagott) sinkt die Zahl auf ca. 260 ab. Dies hängt damit zusammen, dass diese Instrumente insbesondere in der traditionellen Blasmusik nicht besetzt sind. Ähnlich verhält es sich mit dem Register um das Keyboard herum, auch diese Instrumente sind oftmals nur in konzertanten Stücken besetzt. Eine Sonderstellung haben die Fanfaren, diese sind lediglich im Zusammenhang eines Sinfonischen Blasorchesters als Spezialinstrumente zu betrachten. Beispielsweise können die drei Posten des „Der große Zapfenstreichs“ auf diesen geblasen werden und ergänzen dann die Trompeten. Abb. 3b zeigt die Anzahl der Einzelstimmen pro musikalischem Werk (Darstellung wurde reduziert auf die Top 20 des Datensatzes). Bei den dargestellten Stücken handelt es sich fast ausschließlich um konzertante Werke. Wie eben erwähnt werden zum einen in diesen Werken Spezialinstrumente eingesetzt, zum anderen sind die Arrangements komplexer gestaltet, meistens mit 4-stimmigen Posaunen- und Trompetensätzen, 2 Oboen plus einer eigenen Stimme für Englischhorn etc.

Der eigentliche Ansatz zur Metadatenextraktion ist in Abb. 4 dargestellt. Für den ersten Schritt sind wir ausschließlich an dem Titel des Werkes interessiert (im Beispiel „Berliner Luft“). Mittels eines OCR-Programmes wird die erste Seite einer Einzelstimme (hier der ersten Trompete) analysiert und die Texte extrahiert. Als Ergebnis liefert EasyOCR die Bounding Boxes aus dem Detection Modell (rote Rechtecke) und die extrahierten Texte aus dem Recognition Modell, mitsamt einer Konfidenz (Tabelle). In unserem Beispiel hat EasyOCR alle relevanten Textfelder extrahiert, was in 19 Bounding Boxes resultiert. Nach qualitativer Begutachtung macht es dies auch für andere Beispiele sehr zuverlässig. Darüberhinaus

⁹ Ein sog. Register fasst bspw. alle Hörner oder alle Posaunen zu einer Gruppe zusammen.

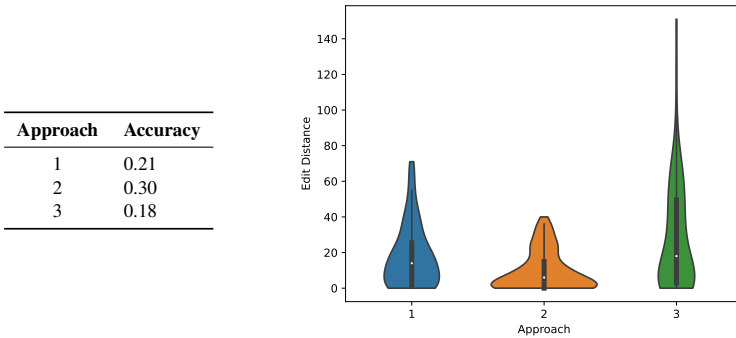


Abb. 5: Evaluationsergebnisse der drei Ansätze zur Titelerkennung. Die Tabelle zeigt die mittlere Accuracy auf den 99 Sample. Die Graphik stellt die Verteilung der Edit Distance der drei Ansätze als sog. Violin-Plots dar.

werden auch oftmals Dynamikzeichen (z.B. *p*, *mf*, etc.), sowie Spielanweisungen mit detektiert und extrahiert. Die Extraktion des Textes ist ebenfalls von sehr hoher Qualität, ebenso bei den anderen Beispielen. Beim dargestellten Beispiel zeigt sich, dass Texte, die eigentlich zusammenhängend sind (z.B. der Titel), in zwei Bounding Boxes aufgeteilt wird. Darüberhinaus ist auffällig, dass bei der Stimmenangabe die „1“ außerhalb der erkannten Bounding Box liegt, weshalb der extrahierte Text dies auch nicht anzeigt.

Im folgenden evaluieren wir drei regelbasierte Ansätze, um die Titelinformation zu extrahieren. Wir betrachten im Rahmen des Proof of Concepts einen zufällig gezogenen Unterdatenbestand von 99 Samples (bleibt für alle drei Ansätze identisch). Im ersten Ansatz nehmen wir die größte Bounding Box (Höhe x Breite in Pixeln) und nehmen den extrahierten Text als Titelinformation. Im zweiten Ansatz filtern wir die Bounding Boxes vor. Es werden nur die behalten, die sich in den oberen 20% des Dokuments befinden und wählen davon die größte als Titelkandidat aus. Im dritten Ansatz filtern wir wie im zweiten Ansatz vor, zusätzlich führen wir die potenziellen Bounding Boxes noch algorithmisch zusammen. Das Zusammenführen der Bounding Boxes ist in EasyOCR vorimplementiert und führt im Wesentlichen eine Gruppierung auf Grundlage der Abstände der Boxes durch.¹⁰

Die Evaluation führen wir auf zwei Arten durch: Zum Einen mit einer Accuracy, die prüft, ob der extrahierte Text exakt der Ground Truth aus dem Datensatz entspricht. Durch OCR Fehler ist dies allerdings ein sehr hartes Maß, weshalb wir zusätzlich noch die Edit Distance berechnen, um ein Gefühl dafür zu bekommen, wie weit der extrahierte Text vom eigentlichen Text entfernt ist. Abb. 5 zeigt die Ergebnisse der drei Ansätze. Die Accuracy des ersten Ansatzes liegt bei 0.21. Das Filtern auf die obersten 20% des Dokuments hebt die Accuracy auf 0.30. Das zusätzliche Zusammenführen der Bounding Boxes ergibt eine Accuracy von 0.18. Qualitativ betrachtet waren die Ergebnisse des dritten

¹⁰ Für das Experiment verwenden wir folgenden Aufruf: `easyocr.utils.get_paragraph(input_boxes, x_ths=0.5, y_ths=0)`. Die beiden Parameter `x_ths` und `y_ths` wurden manuell ermittelt.

Ansatzes am vielversprechendsten, es passiert hier allerdings häufig, dass zu viele Boxen zusammengeführt werden und beispielsweise der Titel mit dem Arrangeur vermischt wird. Ebenfalls in Abb. 5 zeigt rechts die Verteilung der Edit Distances. Hier geht es nicht um absolute Zahlen, sondern rein um die relative Entwicklung (die Edit Distance variiert mit der Länge der zu vergleichenden Zeichenketten). Im relativen Vergleich der Ansätze wird deutlich, dass von Ansatz 1 zu 2 die Verteilung etwas schmaler wird, hingegen zum 3. Ansatz wieder sehr in die Breite geht. Letzteres ist vor allen dem Phänomen geschuldet, dass zu viele Boxen zusammengeführt werden, die Zeichenkette länger wird und somit die Edit Distance proportional anwächst.

4 Zusammenfassung und Ausblick

In dieser Arbeit wurden digitale Notenarchive in Musikvereinen betrachtet. Ausgehend vom Digitalisierungsprozess, über die Literatureinordnung bis hin zur Anforderungsanalyse wurden grundlegende Aspekte und Anforderungen beschrieben und erörtert. Mit einem ersten Proof of Concept wurde gezeigt, in wie fern sich OCR-Verfahren, die für Textdokumente ausgelegt und trainiert sind, eignen, um den Titel aus einem Notentext zu extrahieren. Die Ergebnisse der Experimente zeigen, dass dies als ein erster explorativer Schritt in diese Forschungsrichtung zu betrachten ist. Das verwendete Pythonpaket EasyOCR bietet die Möglichkeit, eigene Modelle zu trainieren, dazu muss allerdings der Datensatz mit weiteren Annotationen angereichert werden. Von der reinen Größe her hat dieser wahrscheinlich das Potenzial nach erfolgter Annotation, die Modelle mit ausreichend Daten zu versorgen.

Im Vergleich zu ähnlichen Experimenten, wie beispielsweise [BAM15], zeigt sich, dass insbesondere die Erkennung der relevanten Bounding Boxes (Detection) ein Hauptproblem für die existierenden OCR Modelle ist. Dieser Schritte wurde in den Experimenten von Balke et al. [BAM15] durch regelbasierte Verfahren übersprungen und es wurde lediglich die eigentliche Text Recognition durch ein OCR-Programm übernommen. Die damaligen Recognition Modellen hatten allerdings qualitativ betrachtet eine schlechtere Qualität.

Die präsentierten Arbeiten eröffnen eine Reihe weiterer Fragestellungen. Die Qualität der Titelextraktionen noch nicht in der Praxis einsetzbar und bedarf daher weiterer Anstrengungen. Insbesondere der Detection-Schritt, also das Auffinden ggf. der relevanten Bounding Boxes weist noch erhebliches Verbesserungspotenzial auf. Zudem könnten weitere Informationsquellen eingebunden werden, um die OCR-Ergebnisse zu verbessern (z.B. externe Datenbank wie <https://de.musicainfo.net>). Ein weiterer Anwendungsfall, der oben kurz angerissen wurde, ist das automatische Trennen der Scans in die Einzelstimmen (vergleichbar zu einer Boundary Detection aus der computergestützten Musikverarbeitung). In einer weiteren Ausbaustufe wird auch die Aufgabenstellung der Optical Music Recognition (OMR) ein Thema werden [CJP20]. Der derzeitige Fokus von OMR liegt vor allen Dingen auf Klaviermusik, die durch ihre Systemanordnung und den regelmäßigen auftretenden Anomalien, ein sehr schwieriges Feld darstellt. Die monophone Einzelstimmen der Instrumente der Sinfonischen Blasmusik könnten hier ein interessantes Zwischenszenario bieten.

Aktuelle Forschung in der Sinfonischen Blasmusik zeigt (z.B. [Ba23]), dass die Musik als solche und die zugehörigen Orchesterstrukturen als interessantes Versuchsfeld im Bereich der computergestützten Musikverarbeitung dienen kann. Durch die ehrenamtlich getragenen Strukturen und vor allen Dingen durch die Möglichkeit kleine Laborumgebungen einzurichten, entstehen viele neue praxisbezogene Forschungsfragen, die allerdings schnell Anschluss zu den derzeit weltweit diskutierten Fragestellungen des angewandten Maschinellen Lernen finden können.

Danksagung

Wir danken Boris Wortmann vom Musikzug Ennest, der in seiner unermüdlichen, ehrenamtlichen Arbeit als Notenwart den Datensatz für unsere Experimente ermöglicht hat. Diese Arbeit wurde durch die Deutsche Stiftung für Engagement und Ehrenamt (DSEE) unterstützt.

Literatur

- [Ba23] Balke, S.: Sinfonische Blasmusik als neuer Anwendungsfall in der automatisierten Musikverarbeitung. In: Proceedings der Deutschen Jahrestagung für Akustik (DAGA). Hamburg, Germany, S. 1366–1369, 2023.
- [BAM15] Balke, S.; Achankunju, S. P.; Müller, M.: Matching Musical Themes based on Noisy OCR and OMR Input. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Brisbane, Australia, S. 703–707, 2015.
- [CJP20] Calvo-Zaragoza, J.; Jr., J. H.; Pacha, A.: Understanding Optical Music Recognition. *ACM Computing Surveys* 53/4, 77:1–77:35, 2020, ISSN: 0360-0300, URL: <https://doi.org/10.1145/3397499>.
- [Da12] Damm, D.; Fremerey, C.; Thomas, V.; Clausen, M.; Kurth, F.; Müller, M.: A digital library framework for heterogeneous music collections: from document acquisition to cross-modal interaction. *International Journal on Digital Libraries: Special Issue on Music Digital Libraries* 12/2-3, S. 53–71, 2012.
- [Di12] Diet, J.: Digitalisierung und Bereitstellung von Musikedokumenten in der Bayerischen Staatsbibliothek. In: Proceedings der Jahrestagung der Gesellschaft für Musikforschung. Göttingen, Germany, S. 29–35, 2012.
- [Ku15] Kurth, S.: Das Schott-Archiv in der Bayerischen Staatsbibliothek. In: Proceedings des Bibliotheksforums Bayern. S. 297–300, 2015.
- [PKK18] Paakkonen, T.; Kervinen, J.; Kettunen, K.: Digitisation and Digital Library Presentation System – Sheet Music to the Mix. In: Proceedings of the Workshop on Music Reading Systems. Paris, S. 1–2, 2018.
- [Rö15] Röwenstrunk, D.; Prätzlich, T.; Betzwieser, T.; Müller, M.; Szwillus, G.; Veit, J.: Das Gesamtkunstwerk Oper aus Datensicht – Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt “Freischütz Digital”. *Datenbank-Spektrum* 15/1, S. 65–72, 2015.

Diskettenmagazine als Frühes Kulturerbe: Erschließung und Re-Digitalisierung

Dr. Torsten Roeder ¹

Abstract: Diskettenmagazine sind frühe, ausschließlich digital publizierte Periodika, die in der Zeit der frühen Heimcomputer (1980er) entstanden und den Höhepunkt ihrer Verbreitung in den 1990er Jahren erlebten. Die „Diskmags“ waren häufig nicht-kommerzielle Community-Produkte und geben weitreichende Einblicke in die frühen Szenen digitaler Kultur. Sie schließen insofern potenziell eine Überlieferungslücke, allerdings ist ihre Bewahrung und Zugänglichkeit für die Forschung nicht gesichert, da Bibliotheken in den entscheidenden Jahrzehnten lediglich Printerzeugnisse sammelten und katalogisierten. Das vorliegende Paper erörtert Strategien und gangbare Lösungswege, um Diskettenmagazine langfristig als erforschbares Kulturerbe verfügbar zu halten.


Keywords: Digitales Kulturerbe, Heimcomputer, Disketten, Sammlung

1 Einleitung

1.1 Die Ära der Heimcomputer

Wer erinnert sich noch an die Zeit der ersten Heimcomputer? In den 1980er Jahren begannen Geräte von Herstellern wie Acorn, Atari, Commodore, Olivetti, Schneider oder Sinclair sowohl Wohnzimmer als auch Schreibtische zu bevölkern. In jener Zeit etablierten sich Heimcomputer nach und nach als Unterhaltungskonsolen und Arbeitsgeräte gleichermaßen. Die Basis, die diese Entwicklung vorantrieb, bestand aus einer Subkultur von Computer-Enthusiasten, in der Anwender, Computerspieler, Programmierer, Spieleentwickler, Democoder sowie Hacker und Cracker in Kreisen zusammenfanden.

Das authentische Erlebnis jener frühen digitalen Kultur ist uns heute nicht mehr präsent. Definiert wurde es zunächst durch die Hardware, die einerseits durch ein typisches „look and feel“, andererseits durch eine spezifische Ausstattung (Geschwindigkeit, Speicher, Soundkanäle, Farben, Bildauflösung, Peripherieanschlüsse) oft sehr charakteristisch war. Zudem waren die Geräte waren längst nicht so hermetisch versiegelt wie heute, so dass man einige Dinge noch mit dem Lötkolben selbst richten konnte. Des Weiteren war die

¹ Universität Würzburg, Zentrum für Philologie und Digitalität, Emil-Hilb-Weg 23, 97074 Würzburg,
torsten.roeder@uni-wuerzburg.de,  <https://orcid.org/0000-0001-7043-7820>

Bedienung von Programmen mehr hybrid als „born digital“:² Computern lagen dicke Handbücher bei, Anwendungsprogramme wurden oft mit umfangreichen Bedienungsanleitungen ausgeliefert, und auch in Spielen war das Druckbeiwerk manchmal unerlässlich. Programme wurden manchmal nicht nur digital kopiert, sondern auch händisch aus Printmagazinen abgetippt. Software erforderte Einarbeitungszeit und eine hohe Frustrationsschwelle, bei vergleichsweise geringer Stabilität der Betriebssysteme.

Entgegen dieser scheinbaren Widrigkeiten des damals neuen Mediums „Heimcomputer“ fand sich eine Community bereit, sich mit den Geräten und den Programmen intensiv auseinanderzusetzen und deren begrenzte Möglichkeiten immer wieder aufs Neue kreativ auszuloten. Wie lässt sich die Kultur jener frühen digitalen Bewegung heute noch vermitteln?

1.2 Heimcomputer-Journale: Dokumente der Kulturgeschichte

Die intensive Selbstdokumentation, die von der heutigen digitalen Massenkultur durch zahlreiche Medienkanäle geleistet wird, steht für die damalige Zeit nicht in derselben Dichte zur Verfügung. Eine besondere Rolle fällt deshalb den zahlreichen Periodika zu, die in jener Zeit produziert wurden und Einblicke in die Vielfalt der Heimcomputerszene geben. Einen großen Teil decken hier zunächst die professionellen Printmagazine ab, die allerdings das Geschehen innerhalb der diversen Subszene nur selten fokussierten. Daher lohnt die nähere Betrachtung einer besonderen Art von multimedialen Born-Digital-Journalen, die schon Anfangs jener Periode früher digitaler Kultur entstanden ist und nicht nur die charakteristischen multimedialen Ansätze jener Zeit spiegelt, sondern auch Dokumente der Vielfältigkeit und Lebendigkeit der Heimcomputer-Community darstellen.

2 Diskettenmagazine

2.1 Begriff, Technik, Verbreitung

Diskettenmagazine, im Deutschen auch „Diskmags“ abgekürzt,³ waren digitale Zeitschriften, die am Bildschirm gelesen und auf zeitgenössischen Trägermedien wie 3,5- oder 5,25-Zoll-Disketten, in der Frühzeit auch auf Datensetten⁴ herausgegeben wurden. Der Höhepunkt ihrer Verbreitung ist zwischen den späten 1980er und den frühen 1990er

² Zur Diskussion um den Begriff „born digital“ im Kontext hybrider Überlieferungsgegenständen siehe Ruan/McDonough [RM09].

³ Im englischsprachigen Raum wurden Diskettenmagazine auch ‚diskzines‘ oder ‚magazette‘ genannt.

⁴ Magazine auf Datensette erschienen in den späten 1970ern, darunter CURSOR für den Radio Shack TRS-80 und CLOAD für den Commodore PET.

Jahren anzusetzen. Sie wurden primär in Gegenden mit hoher Heimcomputerdichte hergestellt, also in den USA, in Großbritannien, Westdeutschland, Italien, Schweden, Finnland und anderen westlichen Ländern. Diskettenmagazine waren fast immer an bestimmte Computersysteme gebunden. Ihr Verbreitungsgrad wurde daher nicht vorrangig durch sprachliche Grenzen limitiert, sondern primär durch technische Hürden bestimmt; diese wieder hingen von der marktbedingten, teils landesspezifischen Verbreitung der Geräte ab.⁵ Diskettenmagazine richteten sich deshalb oft spezifisch an die Communities der jeweiligen Systeme und beförderten damit eine teils einschlägige Identitätsbildung.⁶ Insgesamt kursierten vermutlich mehrere tausend unterschiedliche Titel. Die Mehrzahl der Diskettenmagazine wurden auf Englisch publiziert, aber es gab eine relevante Community von deutschsprachigen Diskettenmagazinen: Ein Hobbyprojekt listet allein 30 deutschsprachige Titel, die für den Commodore 64 erschienen sind [GaoD]; sie existierten aber z. B. auch auf Französisch, Russisch, Polnisch, Schwedisch, Spanisch oder Niederländisch.⁷

Verbreitet wurden Diskettenmagazine stets auf ihrem namengebenden Medium, aber auf ganz unterschiedlichen Vertriebswegen. Einige konnte man regulär am Kiosk beziehen, mit einer relevanten sekundären Leserschaft, denn es wurde viel getauscht und auch kopiert. Andere wurden gegen Einsendung von Retourporto und Leerdiskette oder einen angemessenen Unkostenbeitrag per Post versendet, darunter z. B. „Killer Drive“ und das seit 1993 (mit Unterbrechungen bis heute) existierende „Digital Talk“; manche verbreiteten sich ausschließlich über Privatkopien.

Diskettenmagazine ließen sich durch mitgelieferte Reader-Software am Bildschirm lesen, die individuell für die Magazine programmiert waren.⁸ Für die Navigation wurde meistens ein Menüsystem angeboten. Verlinkungen (innerhalb eines Magazins) wurden nur selten implementiert, allerdings findet sich in manchen Readern ein videotext-ähnliches Chiffren-System, so dass Seiten direkt aufgerufen werden konnten. Einige Magazine boten scrollbaren Text, andere besaßen eine feste Seiteneinteilung. Bei Diskettenmagazinen aus der Demoszene ist hin und wieder auch Lauftext zu finden. Es ist noch zu erwähnen, dass Diskettenmagazine ein „langsames“ Medium waren, denn die

⁵ In Großbritannien beispielsweise waren Acorn BBC Micro (1981) und Sinclair ZX Spectrum (1982) weit verbreitete Geräte, während in Westdeutschland amerikanische Produkte von Commodore dominierten, darunter vor allem der VC-20 (1981) und der C64 (1982). In Frankreich existierte indessen der Onlinedienst „Minitel“, der in der Breite erfolgreich war und die Nachfrage nach Diskettenmagazinen möglicherweise limitierte.

⁶ In der „Enzyklopädie der Diskmags“ beschreibt Volko [Vo12] allein 80 Diskettenmagazine für IBM-PC im Detail. Aus den Beschreibungen wird ersichtlich, dass bereits seit den 1990er Jahren vermehrt über ‚ältere‘ Computersysteme diskutiert wird, sei es hinsichtlich der Migration von Inhalten oder bezüglich der weiterhin aktiven Subszene.

⁷ Ein niederländisches Diskmag ist beispielsweise RUN Flagazine, archiviert im Internet Archive.

⁸ Durch die Einbettung in die Reader-Software unterscheiden sich Diskettenmagazine grundlegend von ASCII-Magazinen („Zines“), die im Plaintext-Format ausgeliefert wurden und für die Verbreitung durch die damaligen Datenfernübertragungskonzepte optimiert waren. Dennoch wird die Bezeichnung „Diskmag“ manchmal auch für ASCII-Magazine verwendet, vor allem wenn ausführbare Software mitgeliefert wurde.

Ladezeit – gerade auf älteren Systemen – konnte durchaus einige Minuten benötigen [Re20].⁹

Einige Titel seien hier exemplarisch erwähnt. Als allererstes Diskettenmagazin kann „Softdisk“ angesehen werden, das ab 1981 für Apple II erschien (zunächst auf Datasette) und später für andere Systeme portiert wurde. Das international bekannteste Diskmag für Commodore 64 war „Loadstar“, dessen Name auf den berühmten Ladebefehl „LOAD“^{8,1} zurückgeht; es stellte einen Rekord auf, indem es ununterbrochen von 1984 bis 2007 existierte. Populäre Titel waren außerdem „Diskmag“ für IBM-PC, „Generation“ für Commodore Amiga und „ST News“ für Atari ST.¹⁰

Diskettenmagazine konkurrierten allgemein zunächst wenig mit den textbasierten Bulletin Board Systems (BBS) der späten 1980er, jedoch übernahm das World Wide Web im Verlauf der 1990er die Funktion der multimedialen Informationsverteilung und viele der Diskettenmagazine wurden zusammengelegt oder beendet. Nur wenige wurden online weitergeführt, manche führten dabei die Bezeichnung „Diskmag“ im Titel weiter; in der Demoszene ist dies bis heute üblich.

Die derzeitige Dokumentationslage ist erbärmlich. Bibliotheken sammelten Diskettenmagazine nicht, da sie kein Printprodukt darstellten. Nur falls ein wesentlicher Printanteil vorlag, wurden manchmal Ausgaben katalogisiert und ggf. auch eine ISSN vergeben, z. B. für „Loadstar“ (0886-4144); oder für „INPUT 64“ (0177-3771). Verlässliche Informationen und digitale Dokumente (sowohl Scans als auch Binärdateien) finden sich am ehesten auf einschlägigen Fansites. Der rechtliche Status solcher Archive ist allerdings nur teilweise transparent.

2.2 Inhalte

Diskettenmagazine boten zwei verschiedene Inhaltsarten. Auf der einen Seite waren das Texte zu verschiedenen Rubriken, wie etwa Nachrichten zu digitalen Technologien, Peripheriegeräten, Spielen, Anwendungsprogrammen bis hin zu Bastelanleitungen, manchmal auch Literatur. Auf der anderen Seite¹¹ gab es Software: Spiele, Anwendungsprogramme und Demos, manchmal auch Systemtools und kleine Hacks. Oft

⁹ Dem Ladezeiten-Problem wurde auf unterschiedliche Weise behandelt, beispielsweise durch Schnellladeverfahren oder Datenkomprimierung; außerdem war bei einigen Systemen das Abspielen von Musik während des Ladevorgangs möglich.

¹⁰ Ein zeitgenössisches Review zu Diskettenmagazinen präsentiert eine Auswahl einiger populärer Titel [BB86].

¹¹ Bei Diskettenmagazinen auf 5,25-Zoll-Disketten war „Seite“ wörtlich zu nehmen, da diese Disketten beidseitig beschrieben wurden und für die Benutzung manuell umgedreht werden mussten. Auf der Vorderseite waren in der Regel das Magazin und ggf. kleinere Programme untergebracht, auf der Rückseite war Platz für Software größeren Umfangs.

waren dies freie Programme, damals unter dem Label „PD“ (Public Domain)¹² verbreitet wurden, manchmal eigens lizenzierte Versionen.

Diskettenmagazine setzten ihre inhaltlichen Schwerpunkte sehr unterschiedlich. Beispielsweise war INPUT 64, herausgegeben vom bis heute operativen Heise-Verlag, vor allem ein Software-Magazin, das neben vielen Programmen lediglich deren Beschreibungen und Bedienungsanleitungen mitlieferte. Auch „Magic Disk 64“, „Starlight“ und „Loadstar“ wiesen einen Software-Schwerpunkt auf, enthielten aber z. B. auch Nachrichten aus der Computerszene, Elektronik-Bastelanleitungen sowie Rezensionen zu aktueller Hardware und Software. Diese Magazine wurden wie Printzeitschriften am Kiosk verkauft, denn sie boten durch die Software entsprechenden Mehrwert.

Eine ganz andere Kategorie bildeten die privat vertriebenen Magazine, bei denen die Textinhalte im Vordergrund standen. Letztere bestanden oft aus Leserbeiträgen und formten diese Diskettenmagazine insofern zu einer Kommunikationsplattform der Community. Dazu gehörten auch Literaturmagazine wie z. B. „Intel Outside“ (1999–2000). Das ausschließlich über Privatkopien kursierende Diskettenmagazin „Sex'n'Crime“ (der Titel spricht sehr für die Selbstzuordnung zur Untergrundszene) ermöglichte sogar das sogar die Hinzufügung von eigenen Nachrichten und Kommentaren; äußerlich glich es mehr einem Demo. Diese und vergleichbare Magazine enthielten Produkte der frühen digitalen Kreativszene – handgezeichnete Pixelgrafiken, Musik und Demos, für die es bis heute eigene Wettbewerbe gibt. Die mediale Gestaltung der Diskettenmagazine ist somit nicht zwingend auf Text fokussiert. Sie sind ein Zusammenspiel aus Text, Grafik, Animation und Klang, die die multimedialen Fähigkeiten der jeweiligen Computersysteme gezielt ausreizten und charakteristische Erzeugnisse schufen.

Man sollte an dieser Stelle erwähnen, dass die Autoren der Diskettenmagazine oft zugleich deren Programmierer waren. Es schrieben also Personen aus der Community, anders als bei Printmagazinen, bei denen ein professionalisierter Computer-Journalismus existierte, der jedoch vom Firmenmarkt, nicht vom Heimcomputermarkt herrührte. Die Autoren waren vor allem Laienjournalisten, bei denen nach und nach eine Professionalisierung stattfand. Viele Magazine erhielten sich das Flair eines authentischen Grassroots-Produkts. Insofern können die Artikel der Diskettenmagazine eine Lücke in der Überlieferungslage zum damaligen Erleben von Digitalität füllen. In ihnen spiegeln sich Jargon, Selbstwahrnehmung, Habitus und interne Strukturen der frühen Heimcomputerszene.

¹² Der Begriff stammt aus dem angelsächsischen Common Law und bezeichnet copyright-freie Werke. Programmierer deklarierten ihre Software als Public Domain, um weite Verbreitung zu ermöglichen. In Kontinentaleuropa hingegen, wo das personengebundene Urheberrecht gilt, bezieht sich der Ausdruck (juristisch) nur auf Werke, die durch abgelaufene Urheberrechte (70 Jahre nach dem Tod eines Urhebers) gemeinfrei geworden sind. In der Heimcomputerszene galt die angelsächsische Bedeutung.

2.3 Fallbeispiel

Als ein Fallbeispiel möge hier das im deutschsprachigen Raum populäre Diskettenmagazin „Magic Disk 64“ detaillierter vorgestellt werden, das von 1987 bis 1993 existierte. Herausgegeben wurde es von „CP Computer Publications“ in Nürnberg durch Christian Geltenpoth, der später Mitgründer der Firma CompuTec werden sollte. Das Diskettenmagazin konnte man zum Preis von 9,80 DM oder 80 Schilling erwerben (heute etwa 10 Euro). Es wurde außerdem in der Schweiz, in den Niederlanden und in Italien vertrieben.

Ausgeliefert wurde es auf einem bunten Trägerkarton, der als Titelblatt fungierte und auf die wesentlichen Inhalte auf der Diskette hinwies – meist bezog sich dies auf die mitgelieferte Software. Der Reader der Magic Disk wurde mehrmals überarbeitet: Das erste Interface (1987) bot ein grafisch nüchternes Hauptmenü mit einem steuerbarem Pointer zur Auswahl der Rubriken; innerhalb der Rubriken waren einzelne Seiten anwählbar. Die zweite Version war grafisch deutlich aufwändiger gestaltet und bot eine Hintergrundmusik, zudem wurde das Interface vereinfacht bzw. an die zeittypischen Navigationsmechanismen von Computerspiel-Menüs angepasst (begrenzt auf die Funktionen auf/ab/auswählen).

Die Inhalte der Magic Disk 64 waren nach mehreren Rubriken aufgeteilt. In den „News“ erfuhr man von aktuellen Veranstaltungen und Ereignissen aus der Heimcomputerszene, während unter „Hardware“ und „Software“ neue Geräte und Programme (hier vor allem Spiele) vorgestellt und kritisch beleuchtet wurden. Die Rubrik „Bastelware“ beinhaltete elektrotechnische Anleitungen, während unter „Tips und Tricks“ einige Programmiertricks zu finden waren. Die mitgelieferte Software wurde schließlich in den Rubriken „Utilities“ und „Games“ präsentiert. Die Kategorien wurden in späteren Ausgaben teilweise umbenannt, blieben inhaltlich jedoch erhalten. Außerdem kamen mit „Hack Meck“ noch Kommentare zur Hackerszene hinzu, und einige Ausgaben beinhalteten mehrteilige Programmierkurse. Daneben gab es ein Editorial, Leserbriefe, Impressumsangaben und eine Vorschau auf die nächste Ausgabe. Die erste Ausgabe umfasste etwa 42.000 Anschläge¹³ und 14 Abbildungen; der Satz des C64 wurde angepasst, um Umlaute und Scharfes S abbilden zu können.

Nach 1993 wurde die Magic Disk 64 in das Printmagazin „Play Time“ integriert; 1995 erschien sie dann kurzzeitig als Printausgabe mit Diskettenbeilage unter dem Titel „Magic Disk 64 Classic“. Genau diese beiden Ausgaben wurden dann auch tatsächlich von der Nationalbibliothek katalogisiert [ZeoD]. Indessen liegt eine intensive Dokumentation durch die Community vor: Floppy-Binaries für Emulatoren sind beispielsweise auf einer aufwändig gestalteten Fan-Website erhältlich [ZaoD]. Auf der Website „Kultboy.com“ finden sich außerdem Scans der farbig bedruckten Trägerpappen oder Titelseiten [ScoD]. Daneben existieren YouTube-Kanäle, bei denen Ausgaben der Magic Disk 64 am

¹³ Zum Vergleich: Ein Printmagazin aus derselben Zeit (z. B. das 64er Magazin aus dem Verlag Markt&Technik) enthielt grob geschätzt 3.000–6.000 Anschläge pro Seite.

Bildschirm durchgeblättert und kommentiert werden [PaoD]. Das C64-Wiki dokumentiert die Inhaltsverzeichnisse sämtlicher Ausgaben [CW15] und eine weitere Seite in einem eher schattigen Bereich des Internets enthält den vollständig extrahierten Plaintext und die Grafiken aller Ausgaben [Ne08]. Insgesamt ist dies für den Erhalt des überlieferten Materials eine günstige Situation, die aber auf die besondere Popularität der Magic Disk 64 zurückzuführen ist, während bei weniger weit verbreiteten Diskettenmagazinen die Überlieferungssituation deutlich prekärer ist. Die Nutzbarkeit der überlieferten Materialien wird zudem durch rechtliche Gegebenheiten (Urheberrecht) und teilweise Unklarheiten erschwert (Verwendung von Pseudonymen, Kollektivautorschaft, undeutliche Rechteangaben).

3 Diskettenmagazine als Digital Heritage

3.1 Perspektiven für den Erhalt

Seit die UNESCO im Jahr 2003 ihre „Charter on the Preservation of the Digital Heritage“ verabschiedet hat und damit Veränderungen der politischen Rahmenbedingungen des Erhalts von Born-Digital-Heritage angestoßen hat [UN03], gelangt das digitale Kulturerbe verstärkt ins öffentliche Bewusstsein. 2021 wurde die Demoszene, die dem Medium Diskettenmagazin eng verbunden ist und ihre Wurzeln ebenfalls in den 1980ern hat, in die deutsche UNESCO-Liste des immateriellen¹⁴ Kulturerbes aufgenommen [La19] [De21]. Ob etwas vergleichbares für Diskettenmagazine geschehen wird, hängt maßgeblich von der Community ab – und auch vom Interesse der Forschung. Zudem erfordert die Bindung der Software an historische Geräte einerseits und die teilweise ausgeprägte Hybridität des Materials andererseits besondere Umsicht, wenn es darum geht, einen ausgewogenen Gesamteindruck von dem historischen Medienerlebnis zu bekommen.

Hinsichtlich des materiellen Erhalts von Diskettenmagazinen beginnen die Probleme bei den Disketten selbst, die nach etwa 10–30 Jahren infolge von Demagnetisierung unlesbar werden [Na89]. Selbst wenn sie lesbar sind, bedarf es funktionstüchtiger Hardware, die ebenfalls der Korrosion ausgesetzt ist. Früher oder später wird man auch mit Rekonstruktionen arbeiten müssen, während der Einsatz von Originalmedien und -geräten auf besondere Anlässe zu reduzieren sein wird. Die materielle Gebundenheit des Digitalen [Sh16] ist daher nicht zu unterschätzen.

Eine Methode, die sich für einen breitenorientierten Zugriff eignet, ist die Emulation. Die originale Umgebung wird dabei in einer anderen, aktuellen Umgebung gekapselt. Voraussetzung dafür ist lediglich, dass ein digitales Abbild der originalen Dateien existiert (binary copy), die von der Emulation dann wie ein virtueller Datenträger behandelt wird.¹⁵

¹⁴ Immateriell bezieht sich hier nicht auf Digitalität, sondern auf „lebendige kulturelle Traditionen und Ausdrucksformen“ [DeoD].

¹⁵ Verfahren zur Herstellung von Binaries sind bereits hinreichend standardisiert [BE13].

Emulation ist auch im Browser möglich. Dies kann die Funktion von Faksimilia erfüllen, wie sie aus dem Gebiet der Manuskripte bekannt ist. Im Internet Archive finden sich über in verschiedenen Collections mehrere Tausend emulierte Ausgaben von Diskettenmagazinen, allerdings unsystematisch zusammengestellt, so dass man leider nicht von einer repräsentativen Zusammenstellung sprechen kann. Immerhin ist Emulation ein möglicher Weg, die Funktionalität von Diskettenmagazinen zu bewahren, wenngleich dabei zum einen das originale ‚Look-and-Feel‘ der Hardware sowie die analogen Bestandteile schnell aus der Berücksichtigung fallen und sie somit lediglich Teilaspekte abdeckt.

Auch eine Portierung oder Migration der Inhalte von Diskettenmagazinen auf moderne Geräte, also eine Re-Medialisierung, ist denkbar, beinhaltet aber mehrere Schwierigkeiten. Bei einer solchen Übertragung werden die Inhalte und Funktionen für ein anderes Mediengerät adaptiert und dabei üblicherweise an dessen Möglichkeiten angepasst. Der Kontakt zur Materialität des Originalmaterials ginge hier ebenso wie schon bei der Emulation verloren, bei gleichzeitiger Verringerung der historischen Distanz zum Medium. Die spezifischen Bedingungen der Erzeugung – im vorliegenden Fall vorwiegend die charakteristischen Limitierungen der Originalgeräte – wären hier nicht mehr nachvollziehbar. Dabei ist beispielsweise zu berücksichtigen, dass die historischen CRT-Bildschirme durch ihre runde, weiche Pixelform digitale Bilder anders wiedergeben als die späteren LCDs, bei Pixel rechteckig erschienen [Sm95].¹⁶ Letztlich kann Portierung kaum in Sinne einer Erhaltung genutzt werden, da es – bei den mittlerweile extrem großen Unterschieden – mehr auf eine Karikatur des Originals in neuem Gewand hinausliefe.

Diesem Umstand könnte indessen Rechnung getragen werden, wenn das Ziel nicht die Ersetzung des Originals, sondern dessen digitalphilologische Edition wäre. Die Texte und auch die Bild- und Soundmedien würden einer textkritischen Betrachtung und Kommentierung unterzogen werden und die Wiedergabe müsste nicht auf die faksimilierte Reproduktion des Originals fokussieren, sondern auf die Präsentation des verfügbaren Materials in seinem historischen Kontext. Dabei könnte die oben beschriebene Emulation die Rolle von Faksimilia übernehmen, bei der sogar der direkte Sprung aus einem edierten Artikel hinüber an die genau entsprechende Stelle in einer Emulation; dies ist ohne weiteres im Browser möglich [KO21].

Die digitale Edition von Born-Digital-Material scheint zunächst trivial zu sein, da die Inhalte ja bereits in digitaler Form vorliegen. Allerdings sind Disketten-Magazine nicht vorwärts-kompatibel zu aktuellen Standards. Das betrifft fundamentale Aspekte wie beispielsweise die Zeichenkodierung: Ältere Geräte wie z. B. der Commodore 64 beschränkten sich auf einen Zeichensatz lateinischer Buchstaben, typische Interpunktionszeichen und einige Blockgrafiken, die teilweise nicht mit ASCII und damit auch nicht ohne den Custom-Bereich von Unicode abbildbar sind. Es bedarf also eines interpretativen Mappings der Zeichensätze (welche zudem häufig modifiziert wurden) und

¹⁶ Neben der grundlegend unterschiedlichen Pixelform sind auch die historischen, fernsehtechnik-basierten Farbschemata nicht eindeutig auf den heutigen Standard RGB(a) abbildbar [Wi22].

die Konservierung des binären Originalcodes parallel zu einem re-digitalisierten Text, einschließlich der Berücksichtigung besonderer typographischer Merkmale des Originalsystems oder der modifizierten Schriftsätze. Die Notwendigkeit der Re-Digitalisierung betrifft außerdem, wie schon angedeutet, auch Grafik sowie Interfaces und Sound mit jeweils geräte- und softwarespezifischen Transformationsverfahren.

3.2 Digitale Edition als Re-Digitalisierung

Die Erstausgabe der oben bereits allgemein beschriebenen Magic Disk 64 dient hier als Vorlage für eine exemplarische digitale Edition. Dabei sind mehrere technische und interpretative Schritte notwendig [Ro22]. Es beginnt mit der Extrahierung der originalen Textdaten, die hier komprimiert vorliegen und somit erst nach Ermittlung und Anwendung des richtigen Kompressionsalgorithmus korrekt ausgelesen werden können. Anschließend sind die Textzeichen entsprechend des modifizierten Zeichensatzes soweit möglich nach UTF-8 zu mappen, einschließlich einer Dokumentation der nicht präzise abbildbaren Zeichen. Anschließend wird die im Interface sichtbar abgebildete Struktur des Diskettenmagazins angelegt, ggf. mit dem Dateisystem verglichen und der Text in die Struktur eingefügt. Anschließend sind außerdem Paratexte wie Splash Screens, Dateiverzeichnisse und ggf. das bedruckte Trägermaterial einzubeziehen. Alles zusammengenommen lässt sich dies im Standardformat TEI/XML beschreiben und abbilden.¹⁷ Bilddaten sind unter Berücksichtigung der historischen Farbschemata einzubeziehen.

Anschließend ist zu entscheiden, wie das re-digitalisierte Material präsentiert wird. Die vorliegende Beispiel-Edition wird mittels eines Webbrowsers aufgerufen, wie es bei digitalen Editionen üblich ist. Im Interface und im Layout werden einige besondere Aspekte konserviert, während andere gegenüber dem Original verändert wurden. Das betrifft zunächst die Seitenaufteilung: Die gesamte Ausgabe ist auf einer einzigen Webseite dargestellt und funktioniert somit nach dem Schriftrollen-Prinzip anstatt nach dem ursprünglichen Seiten-Prinzip. Die Anordnung der Artikel erfolgt nach der Reihenfolge im Hauptmenü, die auch der Gliederung im Dateisystem entspricht. Anstelle des originalen Navigations-Interfaces treten hier Hyperlinks. Wo es notwendig für das Layout ist, werden die originalen Blockgrafiken verwendet, an anderen Stellen wurde zugunsten des Leseflusses eine Schriftart mit variabler Laufweite und ohne die originalen Zeilenumbrüche gewählt. Die Interpretation der Kodierung (z. B. Leerzeichen als Sperrung) ist jeweils im TEI-Code dokumentiert.

¹⁷ Beispieledition unter: Torsten Roeder: „Magic Disk 64. Ausgabe November 1987“. GitHub Pages, 26.02.2020. https://diskmags.github.io/md_87-11.html (letzter Abruf: 23.01.2023); XML-Dateien unter: Torsten Roeder: „diskmags.github.io“. GitHub, 01.06.2022. <https://github.com/diskmags/diskmags.github.io> (letzter Abruf: 23.01.2023). Die exemplarische Edition wurde diskutiert auf der TEI-Konferenz 2022, siehe Newcastle University: TEI2022 Conference, Session 7B. <https://campus.recap.ncl.ac.uk/Panopto/Pages/Viewer.aspx?id=35e3495b-713b-49d5-b3e4-af1000a08f58> (letzter Abruf: 23.01.2023).

Zusätzlich wurden Register angebunden und einige Stellen textkritisch oder inhaltlich kommentiert. Die Artikel enthalten viele Eigennamen von Firmen, Entwicklerteams, Demo- und Hackergruppen, außerdem von Softwaretiteln und Gerätenamen sowie von Programmiersprachen und zeitgenössischen Konferenzen. Normdateneinträge bei der DNB fanden sich dazu jedoch nur in Einzelfällen.

Die Differenz zwischen dem Originalausgabe und der digitalen Edition ist gravierend und die hier vorgestellte Re-Digitalisierung geht mit erheblichen Eingriffen in das überlieferte Material einher. Durch Editionsrichtlinien kann Transparenz gewahrt werden, allerdings ist es unerlässlich, dem vor allem für die Forschung aufbereiteten Material eine dem Original nahe Emulation anzubieten.

4 Offene Desiderate

Digitales Kulturerbe hängt an originalen Datenträgern und an originalen Geräten. Aufgrund des Verfalls sind wir außerhalb von Ausstellungs- und Laborkontexten weitestgehend auf Emulation angewiesen, für die Forschung wird aber auch eine datenorientierte Erschließung benötigt, die auch systematische Zugänge ermöglicht. Die Herstellung des Datenmaterials ist nicht trivial und stellt nicht nur die Digital Humanities vor eine größere Herausforderung, sondern auch Archive und Museen.

Diskettenmagazine bilden ein breites Spektrum von kommerziellen Softwaremagazinen bis hin zu Untergrund-Zines ab. Auch die Zielgruppen reichten von Spiele-Konsumenten bis hin zu involvierten Szenevertretern. Die Zukunft der teils bis heute aktiven Szenen über ihre eigene Generation hinweg bleibt im Ungewissen. Es erscheint daher sinnvoll, die noch aktiven Communities bei der Erschließung von Diskettenmagazine als Citizen Scientists einzubeziehen.

Für die eingehende wissenschaftliche Bearbeitung von Diskettenmagazinen ist als Grundlage zunächst ein umfassender Katalog der überlieferten Titel notwendig.¹⁸ Anschließend wären einzelne Ausgaben zu katalogisieren, möglichst mit Inhaltsverzeichnissen. Schließlich wäre eine umfangreiche Erschließung der Texte und Medien wünschenswert. Dies eröffnet nicht nur die Möglichkeit der Volltextsuche, sondern erlaubt auch analytische Verfahren wie beispielsweise Topic Modeling. Es bedarf zudem systematisch gepflegter Register für die Recherche, die wiederum mit Normdaten anzureichern wären. Die Binaries sind zu archivieren und entsprechend der Rechtslage mit geregelten Zugängen zu versehen bzw. in emulierten Systemen verfügbar zu machen. In besonderen Fällen – etwa bei besonders breitenwirksamen Titeln wie die Magic Disk 64 – kann auch der Aufwand einer digitalen Edition gerechtfertigt sein. Als Zeugnisse einer hoch kreativen Ära, die uns medial inzwischen nicht mehr zugänglich ist, in der aber die

¹⁸ Ein Kooperationsprojekt des NFDI-Konsortiums „Text+“ befasst sich derzeit mit der systematischen Katalogisierung von Diskettenmagazinen [Ze23].

Grundlagen der heutigen digitalen Kultur gelegt wurden, bieten Diskettenmagazine nicht nur einen vielschichtigen Anwendungsfall für den Umgang mit dem frühem digitalem Kulturerbe, sondern ermöglichen auch Einblicke in das Denken und Schreiben in multimedialen Kontexten.

Literaturverzeichnis

- [BE13] Barrera-Gomez, J.; Erway, R.: [Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house](#). OCLC Research, Dublin, 2013.
- [BB86] Bertuca, D. J.; Bertuca, C. A.: [Is There a Disk Magazine in Your Future?](#) In: *Serials Review* 12, No. 2/3 (1986), S. 41–45.
- [CW15] C64 Wiki: [Magic Disk 64](#). https://www.c64-wiki.com/wiki/Magic_Disk_64, 25.02.2015.
- [DeoD] Deutsche UNESCO-Kommission: [Bundesweites Verzeichnis des Immateriellen Kulturerbes](#). <https://www.unesco.de/kultur-und-natur/immaterielles-kulturerbe/immaterielles-kulturerbe-deutschland/verzeichnis-ike>, o.D.
- [De21] Deutsche UNESCO-Kommission: [Demoszene – Kultur der digitalen Echtzeit-Animationen](#). <https://www.unesco.de/kultur-und-natur/immaterielles-kulturerbe/immaterielles-kulturerbe-deutschland/demoszene>, 2021.
- [GaoD] Gansberger, F. (Hg.): [www.c64.at: Das Verzeichnis von deutschsprachigen Magazinen für den C64](#). <https://www.c64.at>, o.D.
- [KO21] Kaltman, E.; Osborn, J.; Wardrip-Fruin, N.: [From the Presupposition of Doom to the Manifestation of Code: Using Emulated Citation in the Study of Games and Cultural Software](#). *DHQ* 15,1 (2021).
- [La19] Lange, A.: [Weltkulturerbe wird auch digital werden](#). In: *Telepolis*, 11.12.2019.
- [ScoD] Schmitzer, M. (Hg.): [Magic Disk 64](#). Kultboy, <https://www.kultboy.com/index.php?site=kult/kultmags&id=294>, o.D.
- [Na89] National Semiconductor Corporation: [Section 5: Floppy Disk Controller](#). In: *Mass Storage Handbook*, Santa Clara, 1989.
- [Ne08] Nerdherrschaft: [Magic Disk 64](#). <http://magicdisk.untergrund.net/>, 15.08.2008.
- [PaoD] Patul, P.: [The World of Diskmag](#). <https://www.youtube.com/@Patrickpatul>, o.D.
- [Re20] Rettinghaus, K.: [Magic Disk 64](#). YouTube, 25. Februar 2020.
- [Ro22] Roeder, T.: [Rescuing Diskmags: Towards Scholarly \[Re-\]Digitisation of an Early Born-Digital Heritage](#). In: *Magazén* 1,3, 2022, S. 139–58.
- [RM09] Ruan, J.; McDonough, J. P.: [Preserving born-digital cultural heritage in virtual world](#). In: *IEEE International Symposium on IT in Medicine & Education*, 2009, S. 745–48.
- [Sh16] Shep, S. J.: [Digital Materiality](#). In: Susan Schreibman, Ray Siemens und John Unsworth (Hg.): *A New Companion to Digital Humanities*, Chichester, 2016, S. 322–330.

- [Sm95] Smith, Alvy Ray: [A Pixel Is Not A Little Square. A Pixel Is Not A Little Square. A Pixel Is Not A Little Square! \(And a Voxel is Not a Little Cube\)](#). Microsoft Technical Memo 6, 1995.
- [UN03] UNESCO: [Charter on the Preservation of the Digital Heritage](https://unesdoc.unesco.org/ark:/48223/pf0000179529.locale=en). <https://unesdoc.unesco.org/ark:/48223/pf0000179529.locale=en>, 17. Oktober 2003.
- [Vo12] Volko, C.: Enzyklopädie der Diskmags. Norderstedt 2012.
- [Wi22] Wikipedia: [YUV: Conversion to/from RGB](#). 03.05.2022.
- [ZaoD] Zauner, C. (Hg.): [Images](#). Magic Disk 64, <http://www.magicdisk64.com/images-magic-disk-64.html>, o.D.
- [ZeoD] Zeitschriftendatenbank: [Magic disk 64: das C64-Magazin auf Diskette](#). Nürnberg: CP-Verlag, 1992–1993 nachgewiesen, ZDB-ID: 1275979-X.
- [Ze23] Zentrum für Digitalität und Philologie: [NFDI Text+ Projekt ‚Diskmags‘ bewilligt](#). Universität Würzburg, 11.01.2023.

Annotation, Simulation und Analyse eines historischen Datenbanksystems

Moritz Feichtinger¹

Abstract: Seit den 1960er Jahren verwenden Regierungen und Behörden der westlichen Welt digitale Kommunikations- und Informationstechnologien. Damit die Geschichtswissenschaft und damit auch die Bürgerinnen und Bürger ihrer demokratischen Rolle der Aufarbeitung und Kontrolle nachkommen können, ist es wichtig, Standards zur Bewahrung und Analyse digitaler Artefakte des Regierungshandelns zu entwickeln. Dieses Paper zeigt die Schwierigkeiten eines solchen Unterfangens anhand der Überlieferung von Datenbanksystemen aus dem Vietnamkrieg exemplarisch auf. Seit 1966 setzte die US Armee in Vietnam Systeme zur automatisierten Sammlung und Auswertung von Daten für ihre strategische und politische Planung ein. Zwar sind einzelne dieser Datensätze überliefert, jedoch teilweise in einer obsoleten Formatierung, teilweise mit nur fragmentarischer Dokumentation. Der eigentliche Quellcode des Database Management Systems, der Rückschlüsse auf die Funktionsweise des historischen Systems erlauben würde, ist nicht überliefert. Das Paper erläutert diese Situation und diskutiert die Möglichkeit, aus den überlieferten Datensätzen und der Dokumentation eine kritische Simulation der Funktionsweise des historischen Systems zu entwickeln.

Keywords: Geschichtswissenschaft; Informationswissenschaft; Archivwissenschaft; Computergeschichte; Medienarchäologie; Forensik

1 Die Herausforderung durch *digital born* Objekte

Die Geschichte der zweiten Hälfte des 20. Jahrhunderts und des Beginns des 21. Jahrhunderts läßt sich ohne Einbezug digitaler Quellen und Artefakte ("born digitals"), insbesondere digitaler Datenbanken, nicht mehr verstehen und beschreiben. Datenbanken sind zentrale Infrastrukturen digitaler Informations- und Kommunikationstechnologien (ICT). Ihre Struktur bestimmt, welche Informationen erhoben, gespeichert und verarbeitet werden, und wie sie zueinander in Beziehung gesetzt werden können. Seit den 1960er Jahren bilden die jeweiligen Architekturen und Funktionsweisen von Datenbankmanagement Systemen (DBMS) die Grundlage für die Produktion, Ordnung und Weiterverarbeitung von Wissen in digitalen ICTs.[Ha09]. Datenbanken sind weit mehr als bloße Container von Informationen. Sie sind komplexe soziotechnische Systeme, deren Architektur die Generierung und Repräsentation von Wissen im digitalen Zeitalter maßgeblich beeinflussen.[Ki14] Um die vielfältigen Elemente und Auswirkungen der Entwicklung und Nutzung von Datenbank-Technologie historisch fassen zu können, bietet es sich daher an, über das technische Objekt hinaus zu gehen und epistemologischen Vorbedingungen, die Praktiken der Interaktion mit dieser

¹ Universität Basel, Departement Geschichte, Hirschgässlein 21, 4051 Basel, Schweiz moritz.feichtinger@unibas.ch

Technologie, ihre materiellen sowie informationelle Umgebung und die Rückwirkungen auf die Welt mit in die Analyse einzubeziehen. Das Verb “*databasing*” drückt dieses holistische Verständnis aus, da es den Umgang mit Datenbank-Technologie als Praxis betont.[Do17] *Databasing* kann und sollte als soziale und kulturelle – und damit historisch verortete – Praxis verstanden werden. Diese Perspektive entspricht den neueren Ansätzen der Geschichte von Informationstechnologie, in denen nicht mehr die Technik, “der Computer” als Ideal im Singular, sondern “Computing”, also die vielfältigen Praktiken der Interaktion und Nutzung im Fokus stehen. Diese Perspektive hat es ermöglicht, den Umgang mit Informationstechnik als soziale Praktik zu verstehen und somit die beteiligten Personen und Gruppen genauer zu analysieren und den Fokus der “klassischen” Computergeschichte auf männliche Erfinder und Geschäftsleute in westlichen Ländern zu überwinden. [Hi18], [Mc20], [Ch21] Die etablierten geschichtswissenschaftlichen Zugänge müssen auf ihre Übertragbarkeit auf born digitals überprüft und gegebenenfalls erweitert werden. Die Grundlage der wissenschaftlichen Untersuchung von Zeugnissen der Vergangenheit bildet die Quellenkritik, also die Überprüfung der Echtheit, Vollständigkeit, Herkunft und Glaubwürdigkeit von Überlieferungen und Hinterlassenschaften. Aber sind hergebrachte Konzepte etwa der Originalität und Authentizität sinnvoll anwendbar auf digitale Artefakte angesichts ihrer einfachen Kopierbarkeit?[Fi21] Wie könnte in diesem Zusammenhang der Nachvollzug der Provenienz (also Herkunft) von digitalen Quellen aussehen? Kann im Falle von kollaborativ oder interaktiv entstandenen Quellen noch von Autorschaft gesprochen werden? Um diese neue Quellengattung bearbeiten zu können, müssen Methoden aus der Medianarchäologie und der Computer-Forensik in den Werkzeugkasten der Geschichtswissenschaft aufgenommen werden.[Ri22] Auch das analytische Instrumentarium der Geschichtswissenschaft, das verstehende Erschließen (Hermeneutik) muss an die spezifischen Eigenschaften von born digitals angepasst werden. Die Frage nach dem Zusammenhang zwischen der Medialität und dem (semantischen) Inhalt erhält bei digitalen Objekten ein noch größeres Gewicht als bei nicht-digitalen, da digitale Objekte nicht einfach so gelesen werden können, sondern eine passende Softwareumgebung und Computer für den Zugriff auf sie notwendig sind. Klassische Fragen der historischen Wissenschaften etwa nach der Perspektivität und Repräsentativität, dem Kontext und der kulturellen Gebundenheit von Quellen müssen in der Analyse von digital born Objekten ebenfalls bearbeitet werden, erschließen sich aber nicht durch heuristische Methoden aus den Quellen selbst, sondern setzen das Vorhandensein von Kontextinformationen und Dokumentation voraus.[B117] Aus archivalischer Perspektive stellt sich die Frage der Auswahl des zu bewahrenden digitalen Erbes auf zweifacher Ebene: Zunächst ist die Frage der Selektion zu beantworten, also zu entscheiden, welche Quellen und Artefakte der stetig und exponentiell anwachsenden Menge an digitaler Kommunikation und Information für die Nachwelt potentiell von Interesse sein könnten. Dann ist zu klären, wie genau die ausgewählten Zeugnisse zu bewahren sind. Wie ist der technischen Obsoleszenz zu begegnen, also dem Problem, dass sich historische Formate von heutiger (und zukünftiger) Software nicht mehr auslesen lassen? Was genau gehört zu einem digitalen Artefakt, nur seine Repräsentation in Binärcode und die entsprechende Codierung, oder auch die Umgebung im weitesten Sinne, oder sogar Hinweise auf die historische Nutzung? Wie soll der Zugriff auf digitale Objekte in veralteten Formaten gewährleistet werden? Indem sie

in modernere Formate migriert werden oder indem ihre ursprüngliche Funktionsweise in modernen Betriebssystemen emuliert wird?[Lo14], [Bo21]

Das laufende Forschungsprojekt “Ctrl + All – Computing the Social” untersucht ein Datenverarbeitungssystem aus dem Vietnamkrieg entlang der hier vorgestellten analytischen Linien als historisches Fallbeispiel von databasing. Bevor auf die einzelnen Schritte des Forschungsprojektes, erste Ergebnisse und Schwierigkeiten eingegangen werden kann, muss der Entstehungszusammenhang der Datenverarbeitung und die Besonderheiten der archivierten Überreste dargestellt werden.

2 Fallstudie: Databasing im Vietnamkrieg

Die US Armee setzte während ihres Eingreifens in den Krieg in Vietnam (1965-1975) Computer für verschiedene Zwecke ein, für ihre Logistik, Verwaltung, Sammlung von Geheimdienst Informationen, strategische Planung und Evaluation. Auch für die Counterinsurgency, die integrierte politisch-militärische Bekämpfung der Guerilla, setzten die Amerikaner und ihre Verbündeten Computer und automatisierte Datenverarbeitung ein. Die 1967 gegründete Zentralbehörde zur Koordination der zivilen und militärischen Counterinsurgency Massnahmen, das CORDS (*Civil Operation and Revolutionary Development Support*), verwendete Computer und Datenbanken um ihre Projekte zu koordinieren und zu evaluieren. Zwei davon sollten ein monatliches umfassendes Bild der demographischen, sozioökonomischen Situation der ländlichen Bevölkerung produzieren und ihre politischen Ansichten und Haltungen erheben, das *Hamlet Evaluation System* (HES) und das *Pacification Attitudes Analysis System* (PAAS). Das HES war ein Verzeichnis aller 12.000 Weiler und Dörfer in Südvietnam, das die geografische Lage der Siedlungen, soziodemografische Daten über ihre Bewohner und eine fünfstufige Bewertung enthielt, in die der Grad der Kontrolle der jeweiligen Gemeinschaft durch die Regierung oder die Kommunisten angegeben wurde. Die Daten für das HES wurden von 250 US-amerikanischen Militärberatern in allen 44 Provinzen Südvietnams monatlich erhoben und in einer zentralen Datei gesammelt. Das HES wurde seit seiner Inbetriebnahme im Jahr 1967 bis zur Übergabe an die vietnamesische Regierung im Jahr 1973 dreimal überarbeitet und gilt als die umfassendste Datenbank über die Bevölkerungsverteilung und -dichte im Südvietnam der Kriegszeit. Aufgrund seiner Geoinformationen und Bevölkerungsregister diente es als Grundlage für die Verknüpfung vieler anderer von CORDS betriebenen Datenbanken.[Fe22] Das PAAS kann als qualitatives Korrelat zum eher quantitativen HES betrachtet werden. Es bestand aus behavioristischen Umfragen, die von speziell geschulten vietnamesischen Fragestellern monatlich in bestimmten Regionen durchgeführt wurden. Der Fragebogen umfasste Themen wie Lebensmittelpreise, Hoffnungen und Befürchtungen für die Zukunft sowie allgemeine politische Einstellungen und Ansichten. Obwohl PAAS nur auf einer recht kleinen Stichprobe der Bevölkerung basierte, wurden die Erhebungen unter ethnischen und religiösen Minderheiten sowie unter Frauen und damit einem erstaunlich breiten sozialen Spektrum durchgeführt. Die Ergebnisse wurden kodiert und in einer Datenbank zusammengefasst, die

mit dem HES-System kompatibel war. Beide Systeme, HES und PAAS, produzierten monatliche umfassende Ergebnisberichte, deren wichtigsten Erkenntnisse zudem auf Diagrammen, Landkarten und Broschüren visuell dargestellt wurden. Diese Befunde dienten Politikern und Militärs zur Einschätzung der Lage in Südvietnam, zur taktischen und strategischen Planung, aber auch zur Legitimation solcher Manöver sowie des amerikanischen Engagements in Vietnam insgesamt.

2.1 Provenienz, Formate, Dokumentation

Das US-Nationalarchiv (*National Archive and Records Administration*, NARA) erhielt die HES und PAAS Datenbanken (und einige andere) auf 9-Spur Magnetbändern in den späten 1970ern. Da die Abteilung für elektronische Dokumente (*Electronic Records Archive*, ERA) damals nicht über das benötigte Equipment verfügte, extrahierte das National Institute for Health die Daten von den (schnell degradierenden) Magnetbändern. Die Daten waren zum Teil im damals gängigen, IBM-eigenen, EBCDIC Format abgelegt, zum Teil in einem Format namens NIPS: Das *National Information Processing System Formatted File System* war ein aus dem *1400 Formatted File System* (FFS) hervorgegangenes Datenbankformat, das speziell für das US Militär entwickelt worden war. Es lässt sich der Familie der selbstbeschreibenden Datenformaten zuordnen, das heißt, Struktur und Schema der Datensätze war in den vorderen Segmenten der Datensätze selbst beschrieben.[FS76] Um Forschenden den Zugang zu den Datensätzen des Vietnamkrieges zu erleichtern, begann das ERA in den 1980ern damit, die Daten in das ASCII Format zu übersetzen. Dafür wurde eine Software entwickelt (NIPStran), mit der sich grosse Teile der Konversion automatisieren ließ. Die Datensätze liegen also in mehreren Formaten und Speichermedien vor: einige Reports (die Endprodukte beider Systeme) sind als Ausdrucke auf Endlospapier erhalten; das NARA verfügt über Bitstream-Images der kompletten Datensätze in ihrer originären Formatierung (NIPS) und stellt auf seiner Homepage die ASCII-konvertierten Dateien zum Download zur Verfügung. HES und PAAS wurden von spezialisierten Analyse-Abteilungen in mehreren Regierungsbehörden benutzt und analysiert: von der *Research and Analysis Division* des CORDS in Saigon, vom *Combined Intelligence Data Center* der Armee in Saigon, vom *Systems Analysis Office* im *Department of Defense* in Washington, einigen Sozial- wissenschaftlichen Forschungsinstitutionen, sowie von privaten Computerfirmen, die mit Regierungsverträgen die Datenbanken betreuten und weiterentwickelten. Alle diese Institutionen hatten Nutzerhandbücher und haben ihren Umgang mit den Systemen mehr oder weniger umfassend dokumentiert. Das NARA bewahrt zudem einen um fangreichen Bestand der *Research and Analysis Division* des CORDS auf, aus dem sich die Geschichte der Entstehung der Systeme, die Praxis der Datensammlung, sowie mehrere Revisionen der Systeme und die jeweiligen Prozessierungsverfahren zumindest teilweise rekonstruieren lassen. Weitere Quellen und Dokumente im NARA und anderen Archiven lassen Rückschlüsse auf die Verwendung der Daten in der strategischen und taktischen Einsatzplanung sowie der politischen Entscheidungsfindung und Öffentlichkeitskommunikation entscheidender Regierungsbehörden zu.

3 Forschungsdesign

Das Forschungsprojekt zielt darauf ab, ein in Fragmenten überliefertes Datenverarbeitungssystem durch eine Untersuchung der entsprechenden Dokumentationen möglichst breit zu kontextualisieren, mit (quellenkritischen) Metadaten zu annotieren und durch eine (experimentelle) Rekonstruktion und Simulation zu analysieren. Das von den Archivarinnen und Archivaren des ERA im amerikanischen Nationalarchiv entwickelte Programm NIPStran produziert ASCII-formatierte Datensätze als .csv Files und HTML Files, in denen Tabellen-, Spalten- und Attributnamen abgelegt sind. Grundlage für diese Umwandlung sind das *fixed set* der NIPS Dateien, in denen nicht-repetitive Einträge gespeichert sind (also die Kategorien) und jeweils mehrere zugehörige *periodic sets*, also dynamisch Datensätze, in denen die zugehörigen Werte abgelegt waren und monatlich oder vierteljährlich erneuert wurden. Mit diesen Daten lassen sich sehr einfach Tabellen und Data-Arrays erstellen und auswerten. Allerdings enthielten die originalen NIPS-Daten (und auch die in EBCDIC gespeicherten Daten) darüber hinaus in den vorderen Segmenten weitere Informationen, namentlich ein *control set*, das einen Unique Identifier enthielt und zur Verknüpfung der Datensätze genutzt werden konnte, ein *variable set* für Kommentare, File-Maintenance Statements und Log-files, sowie weitere Strukturdaten, die der Interaktionen mit verschiedenen Prozessoren und Verarbeitungsverfahren dienen.

3.1 Extraktion von Strukturinformationen

Der erste Schritt der Analyse besteht darin, ein Verfahren zu entwickeln, das technisch-funktionale Metadaten aus den ersten Segmenten des Bitstreams extrahiert. Grundlage dafür ist der Source-Code von NIPStran und frei verfügbare Programme zur Migration von EBCDIC Dateien. Dabei geht es darum zu klären, ob sich aus dem archivierten Bitstream Aussagen über Funktionsweise und Architektur des Systems ziehen lassen. Bestenfalls lassen sich Verfahren entwickeln, mit denen sich Strukturinformationen und Dateischemata nicht nur aus verschiedenen Datensätzen im NIPS-Format auslesen, sondern das Verfahren so konzipiert werden, dass sich auch aus Daten in verwandten Formaten der FFS Familie Schemata extrahieren lassen. Das angestrebte Ziel dieses Schrittes besteht im Entwurf eines RDF Knowledge-graphs, in dem die Datensätze in ihrer Gesamtheit gemeinsam mit den technischen Strukturinformationen abgelegt sind. Damit soll in einem späteren Schritt die Funktionsweise der Systeme nachvollzogen (und kritisiert) werden können.

3.2 Annotation mit Metadaten

Die im ersten Schritt gewonnenen Strukturinformationen müssen nach einem einheitlichen Schema abgelegt werden und als Metadaten zu den (konvertierten) Daten in den RDF Graph hinzugefügt werden. Zur Beschreibung und Ablage der technisch-strukturellen Informationen wird der Metadatenstandard PREMIS verwendet. Das Vokabular von PREMIS

(PREservation Metadata: Implementation Strategies), dessen Data Dictionary in der letzten Version 3.0 2016 publiziert wurde, hat sich in den letzten Jahren als de facto Standard für Metadaten zur Beschreibung digitaler Objekte etabliert. Mit PREMIS können die strukturellen und technischen Eigenschaften von historischen Datenbanksystemen erfasst werden. Zudem kann PREMIS, das vier große Klassen von Einheiten (entities) kennt - *Object*, *Event*, *Agent*, *Rights* - die Interaktion sowohl von Personen als auch von technischen Systemen mit Datenbanken abbilden. Die Einheit *Object* hat drei Unterkategorien *file*, *bitstream* und *representation*, die geeignet sind, den Prozess der Datenverarbeitung und die jeweiligen Soft- und Hardware Umgebungen zu beschreiben. Die Einheitsklassen *Event* und *Agent* können jegliche Interaktion mit dem System ausweisen, da *Agent* sowohl Maschinen oder Software als auch Personen bezeichnen kann. Die Daten des HES und PAAS durchliefen jeweils einen mehrstufigen Lebenszyklus, in dessen Verlauf sich ihr Inhalt, ihr Ablage-Format, die entsprechenden Speichermedien und ihr Verarbeitungs- und Verfeinerungsgrad veränderte. Zudem konnten sich die Datensätze in untereinander in den vorgenannten Aspekten unterscheiden. In den Archivwissenschaften hat sich die Erstellung von Data-Lifecycle-Modellen zur Beschreibung von Daten etabliert. Um den Lebenszyklus der Datensätze zu ermöglichen, muss ein formalisiertes Schema entwickelt werden, in dem sich die jeweilige Provenienz-Information zu den Datensätzen für die jeweiligen Stufen in ihrem Lebenszyklus beschreiben lassen. Damit soll zum Beispiel festgehalten werden, wo, von wem, mit welcher Methode und auf Grundlage welchen konzeptionellen Modells Daten erhoben wurden. Auch soll sich damit festhalten lassen, welchen Grad der Weiterverarbeitung, Säuberung, Aggregation oder Verknüpfung mit anderen Daten die jeweiligen Datensätze aufweisen.

Die Data Documentation Initiative (DDI) hat mit dem Discovery Vocabulary einen Metadaten-Standard entwickelt, mit dem sich sozialwissenschaftliche Studien im Hinblick auf die vorgenannten Schritte in ihrem Lebenszyklus beschreiben lassen. Das Ziel des zweiten Schrittes besteht darin, eine quellen-kritische Beschreibung für die Daten aus dem Vietnamkrieg aus dem Vokabular von DDI-Discovery zu entwickeln und in den PREMIS-basierten Knowledge-graph aus Schritt 1 zu integrieren. Damit wird dieser nicht nur um die Dimension einer (kritischen) Feststellung der Provenienz Informationen erweitert, sondern es lassen sich erste Korrelationen zwischen technischen und Provenienz-Metadaten erheben. So lassen sich, wenn sämtliche überlieferte Daten einheitlich annotiert und erfasst sind, die Kohärenz zwischen verschiedenen Versionen überprüfen und die Auswirkungen verschiedener Aggregationsstufen feststellen. Die Anwendung von Metadaten-Standards wie DDI-Discovery und PREMIS wird auch deren Grenzen und entsprechenden Bedarf an deren Ausbau für historische Datenbanken und Software aufzeigen. Beide Standards kommen aus der Archivwissenschaft und dienen vornehmlich der Dokumentation, dem Austausch und der Zugänglichmachung von digitalen Objekten, nicht deren Analyse. Für eine kritische Erforschung nach den Standards der Geschichtswissenschaft ist gerade die Ausweisung von Lücken und Unklarheiten in Provenienz und Integrität von Bedeutung und ein entsprechende Auszeichnungs-Vokabular muss entwickelt und erprobt werden. Bestenfalls kann das Forschungsprojekt dafür Anhaltspunkte und Erfahrungen generieren.

3.3 Simulation der Funktionsweise

Der finale und wichtigste Schritt des Forschungsprojekts besteht aus einer kritischen Simulation der Funktionsweise der historischen Systeme. Damit soll überprüft werden, ob sich das experimentelle Reenactment, das in den Science and Technology Studies ein erprobtes Verfahren ist, in den Methodenkanon der reflexiven Computergeschichte und Digital History integrieren lässt.[Du20], [He20], [FPS16]

Die Simulation umfasst auf der Ebene des Gesamtsystems eine formale Modellierung der Datenverarbeitungsprozesse, inklusive der zugrundeliegenden Problemstellungen, der beteiligten Personen, Maschinen, Medien und Materialien sowie der einzelnen Systemkomponenten und ihr Ineinandergreifen auf Softwareebene. In Kombination mit den technischen Metadaten soll somit die Funktionsweise des Gesamtsystems nachvollziehbar und Wechselwirkungen zwischen Materialien, Formaten, und Inhalten überprüfbar gemacht werden. Die zugrundeliegende Idee ist, einen Graph zu erstellen, der den gesamten Lebenszyklus der Daten im System zusammen mit der Funktionsweise und Interaktion mit den einzelnen Komponenten repräsentieren kann. Damit soll auch ermöglicht werden, die jeweiligen Effekte der Veränderung einzelner Parameter und Komponenten auf die Performance des Systems und die Validität der durch das System generierten Informationen zu testen. So lässt sich nachvollziehen, ob sich statistische und konzeptuelle Verzerrungen und Vorannahmen im Prozess der Datenverarbeitung verschärften, oder ob sie durch Aggregationsverfahren nivelliert wurden. Nicht zuletzt sollte die kritische Simulation überprüfbar machen, wie sich die Ergebnisse des Systems zu Lageberichten und Einschätzungen zum Kriegsverlauf aus anderen Quellen verhielten. Das Ziel der Simulation auf Softwareebene besteht darin, die Abfrage- und Systemroutinen in moderner Software nachzubauen. Einerseits sollen verschiedene Query-Prozesse in SPARQL nachgebaut werden, um die historischen Abfrage-Routinen nachvollziehbar zu machen, andererseits soll eine Grundlage für die Entwicklung weiterer Abfrage-Kombinationen ermöglicht werden um zusätzliche Mustererkennung, Varianz-Analyse, etc. zu ermöglichen. Die HES und PAAS Systemprogramme bestanden aus jeweils etwa 80 Routinen und Subroutinen, die in COBOL, FORTRAN 66 und Assembler Language geschrieben waren. Neben Update- und File-Maintenance Routinen hatten die Entwickler Programme zur Abfrage unter bestimmten Aspekten, zur Erstellung von Reports und geographischen Plots, sowie zur Analyse ausgewählter und frei kombinierbarer Attribute geschrieben. Die meisten dieser Programme sind in den Dokumenten umfassend beschrieben, durch Flussdiagramme abgebildet, und in Nutzerhandbüchern detailliert beschrieben.

Die Anwendung von experimentellem Reenactment zur Analyse von historischen Systemen der Informationsverarbeitung wirft einige praktische und ontologische Probleme auf, welche die Aussagekraft der Simulation für die Interpretation der historischen Rolle und Wirkung der Systeme beeinträchtigen. Augenscheinlich handelt es sich bei der kritischen Simulation um eine weitere Stufe der Modellierung und Abstraktion. Versuchten schon die Akteure, komplexe soziale Dynamiken in vereinfachte Modelle zu fassen, um sie im Medium der Datenbanken kalkulierbar und prozessierbar zu machen, so repliziert eine formale Modellierung des Systems diese Abstrahierung von den konkreten, historischen

Gegebenheiten ein weiteres mal. Es ist daher stets zu bedenken, dass durch die Simulation keine Aussagen über die Validität des Systems in der Repräsentation der Realwelt getroffen werden können, sondern nur über die Kohärenz der Informationen in ihrer Sammlung und Verarbeitung innerhalb des Systems. Zudem Abstrahiert eine Simulation und Replikation einzelner Prozessierungsschritte stark von den tatsächlichen materiellen und technischen Charakteristika des historischen Systems. So fand die Kommunikation mit dem System hauptsächlich mit Lochkarten statt, die sowohl für den Input von Daten als auch für das Laden von Programmen genutzt wurden. Wird diese materielle Dimension dieses historischen Mediums nur abstrakt modelliert oder gar ganz unterschlagen, so geraten maßgebliche Systemeigenschaften und Probleme, wie etwa der auf jeweils 80 Spalten beschränkte Speicherplatz oder die aufwändige und arbeitsintensive Normalisierung und Validierung der Daten, aus dem Blick. Gleiches gilt für Simulation der Verarbeitungsgeschwindigkeit und die Materialität der Maschinen und Prozessoren. Die IBM S/360 Systeme der späten 1960er Jahre waren raumfüllenden und extrem wartungsintensive Maschinen, die oftmals nicht physisch integriert waren, sondern aus Einzelkomponenten und Peripheriegeräten bestanden, zwischen denen Input und Output in Form von Lochkarten, die von Menschen hin- und her getragen wurden, ausgetauscht wurde. Die Berechnungszeiten selbst für einfache Operationen innerhalb solcher Systeme konnten Tage und Wochen beanspruchen, was ebenfalls bei einer Simulation in moderner Software nicht vergessen gehen darf. Schließlich ist die Dokumentation in zweifacher Hinsicht lückenhaft: Zum einen sind zwar die meisten der eingesetzten Programme und Subroutinen einigermaßen umfassend dokumentiert, aber das gilt eben nicht für alle und der Source-Code selbst ist nicht erhalten. Zum anderen handelt es sich bei Handbüchern um Darstellungen von einfachen und optimierten Nutzungsszenarien. Sie geben naturgemäß wenig Auskunft über Probleme und Fehler einzelner Anwendungen oder des Gesamtsystems. Zwar kann eine formale Modellierung Probleme mit der Datenkohärenz sichtbar machen, aber Probleme, die sich im Betrieb und in konkreten Situationen ergaben, bleiben bei dieser Analyse unentdeckt und lassen sich nur durch den Abgleich mit der historischen Dokumentation und eines möglichst breiten Quellenkorpus erfassen. Alle Simulationen können daher genau genommen nur als experimentelle Annäherungen gelten und müssen dringend umfassend dokumentiert, kritisch evaluiert und mit anderen historischen Zeugnissen abgeglichen werden.

4 Conclusion

Der Vietnamkrieg ist in die Geschichte eingegangen als krasser Fall der Mißinformation der Öffentlichkeit durch eine Regierung. Weder die politischen Motive, noch das militärische Vorgehen oder die tatsächliche strategisch-politische Situation wurden der amerikanischen und internationalen Öffentlichkeit durch die US-Regierung mitgeteilt, sondern kamen erst sukzessive, dank Whistleblowern, Leaks, mutige journalistische Arbeit und einer kritischen historischen Aufarbeitung ans Licht. Dabei ist das Fallbeispiel des Vietnamkriegs kein isolierter historischer Vorfall, denn sowohl die Konstellation - ein Bündnis aus westlichen Alliierten versucht in einer Gesellschaft der Dritten Welt in einen politischen Konflikt

einzugreifen - als auch die angewandten Mittel, nämlich eine Kombination aus militärischer Repression und politischer Manipulation auf Grundlage automatisiert erhobener und verarbeiteter Daten, sind auch in der Gegenwart und der Zukunft zu erwarten. Umso dringlicher ist die Frage, inwieweit die Fehlinformation der Öffentlichkeit eine bewusste Täuschung war und inwieweit sie auf einer fehlerhaften Interpretation der Fakten oder gar einer verzerrten oder gänzlich falschen Faktenlage beruhte. Um dies zu klären ist es unerlässlich, die Systeme der Informationsverarbeitung selbst einer historischen Analyse und Kritik zu unterziehen. Die Geschichtswissenschaft hat für den Umgang mit digitalen Objekten, seien es Datenbestände oder Systeme, bislang keine allgemein anerkannten analytischen Verfahren und Standards der Dokumentation entwickelt. Das in diesem Paper vorgestellte Experiment soll dazu beitragen, solche Standards und Verfahren zu entwickeln und zu prüfen. Es gründet auf der Annahme, dass ein Austausch und ein gegenseitiges Lernen zwischen den Geistes- und Sozialwissenschaften auf der einen Seite und den Informations- und Computerwissenschaften auf der anderen Seite für beide Seiten bereichernd sein kann und letztlich dazu beitragen wird, die Gesellschaft als Ganzes in die Lage zu versetzen, digitale Informationsverarbeitung und darauf beruhendes staatliches Handeln besser verstehen und kritisch bewerten zu können.

Literaturverzeichnis

- [Bl17] Blewer, Ashley: The PREFORMA handbook: validating formats, a prerequisite for preserving digital objects. PREFORMA, 2017. OCLC: 1263230909.
- [Bo21] Digital Preservation Coalition: Born digital archive cataloguing and description. Bericht, Digital Preservation Coalition, Mai 2021.
- [Ch21] Chakraborty, Indranil: Invisible labour: support service workers in India's information technology industry. Routledge, Abingdon, Oxon ; New York, NY, 2021.
- [Do17] Dourish, Paul: The stuff of bits: an essay on the materialities of information. The MIT Press, Cambridge, Massachusetts, 2017.
- [Du20] Dupré, Sven; Harris, Anna; Kursell, Julia; Lulof, Patricia; Stols-Witlox, Maartje, Hrsg. Reconstruction, Replication and Re-enactment in the Humanities and Social Sciences. Amsterdam University Press, November 2020.
- [Fe22] Feichtinger, Moritz J.: Computing Counterinsurgency: The Hamlet Evaluation System (HES) and Databasing During the Vietnam War. IEEE Annals of the History of Computing, 44(4): S. 28–43, Oktober 2022.
- [Fi21] Fickers, Andreas: Authenticity: Historical Data Integrity and the Layered Materiality of Digital Objects. In (Balbi, Gabriele; Ribeiro, Nelson; Schafer, Valérie; Schwarzenegger, Christian, Hrsg.): Digital Roots, S. 299–312. De Gruyter, August 2021.
- [FPS16] Fors, Hjalmar; Principe, Lawrence M.; Sibum, H. Otto: From the Library to the Laboratory and Back Again: Experiment as a Tool for Historians of Science. Ambix, 63(2): S. 85–97, April 2016.

- [FS76] Fry, James; Sibley, Edgar: Evolution of Data-Base Management Systems. *ACM Comput. Surv.*, 8: S. 7–42, März 1976.
- [Ha09] Haigh, Thomas: How Data Got its Base: Information Storage Software in the 1950s and 1960s. *IEEE Annals of the History of Computing*, 31(4): S. 6–25, Oktober 2009.
- [He20] Hendriksen, Marieke M. A.: Rethinking Performative Methods in the History of Science. *Berichte zur Wissenschaftsgeschichte*, 43(3): S. 313–322, September 2020.
- [Hi18] Hicks, Marie: Programmed inequality: how Britain discarded women technologists and lost its edge in computing. *History of computing*. MIT Press, Cambridge, MA London, UK, 2018. OCLC: on1004074349.
- [Ki14] Kitchin, Rob: The data revolution: big data, open data, data infrastructures & their consequences. SAGE Publications, Los Angeles, California, 2014. OCLC: ocn871211376.
- [Lo14] Loebel, Jens-Martin: Lost in translation: Leistungsfähigkeit, Einsatz und Grenzen von Emulatoren bei der Langzeitbewahrung digitaler multimedialer Objekte am Beispiel von Computerspielen. *Game studies*. Hülsbusch, Glückstadt, 2014. OCLC: 912044941.
- [Mc20] McIlwain, Charlton D.: Black software: the Internet and racial justice, from the AfroNet to Black Lives Matter. Oxford University Press, New York, NY, 2020.
- [Ri22] Ries, Thorsten: Digital history and born-digital archives: the importance of forensic methods. *Journal of the British Academy*, 10: S. 157–185, 2022.

Evaluating Real-Time Pitch Estimation Algorithms for Creative Music Game Interaction

Peter Meier^{1,2}, Simon Schwär², Gerhard Krump¹, Meinard Müller²

Abstract: Music-based games are an important genre in the gaming community and have become increasingly popular with games like SingStar and Guitar Hero. These types of games are usually based on reactive game mechanics, where the player must hit a certain note at a certain time in order to score points. In this contribution, we present a game prototype that goes beyond purely music-reactive game mechanics and focuses more on the creative aspect of making music in games. In particular, we developed a jump-and-run game that can be controlled with a gaming controller but also uses the player's singing voice to interact with the game world. To this end, we estimate the pitch of a microphone signal in real time and use it as a creative input to the game. This input can be used to control parts of the game world, for instance by singing and adding stair-like elements that allow the player to overcome obstacles and reach the end of a game level. With our game prototype, we demonstrate how game designers can incorporate musical challenges into a well-known game environment while motivating musicians to creatively explore and practice their musical skills. Furthermore, motivated by our game prototype, we evaluate different real-time pitch estimation algorithms using common MIR metrics on a publicly available dataset to analyse what works best for our gaming scenario.

Keywords: Evaluation, Real-Time, Pitch, Estimation, Algorithm, Creative, Music, Game, Interaction

1 Creative Music Game Interaction

This paper is part of a series on “Real-Time Signal Processing Algorithms for Interactive Music Analysis Applications.” In previous work, we presented a music-reactive game with beat tracking, where the game world is generated in real time from ambient music (e.g., radio, live music) of the player [Me22]. In this article, we consider a different type of music game that uses pitch estimation as a control input, see Fig. 1 for an illustration. In developing this game, we want to encourage the creative use of melody and singing for players and game designers and concentrate on exploring and teaching real-time aspects of pitch estimation. In particular, we focus on two pitch estimation algorithms, YIN [CK02] and SWIPE [CH08], which we adapted from the Python library libf0 [RSM22] and modified to work in real time for our gaming context. In this paper, we closely follow [Me23], where we already described the conceptual idea of our game prototype, and now complement it with experiments to evaluate the developed real-time pitch estimation algorithms.

¹ Deggendorf Institute of Technology, 94469 Deggendorf, Germany, peter.meier@th-deg.de, gerhard.krump@th-deg.de

² International Audio Laboratories Erlangen, 91058 Erlangen, Germany, peter.meier@audiolabs-erlangen.de, simon.schwaer@audiolabs-erlangen.de, meinard.mueller@audiolabs-erlangen.de

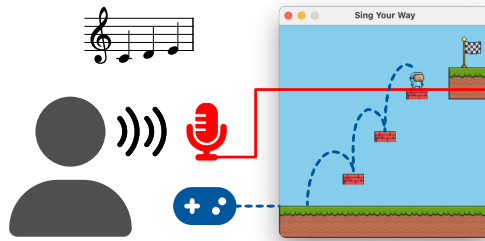


Fig. 1: Illustration of our “Sing Your Way” music game prototype. It uses a combination of traditional gaming controller and singing voice input to add note blocks to the game.

In our music game prototype called “Sing Your Way”, the player can interact with the game world by using their singing voices. Fig. 2 shows what this might look like for a simple use case. In the first step, the player can sing a note. The pitch of this note is displayed as a red line in the game. The higher the note is sung, the higher the red pitch line is displayed. While singing a note, the player can interact with this red pitch line by jumping on it. By doing this, a fixed note block is created in the game for the player to stand on. The same step can be repeated several times at different pitches. As a result, players can add stair-like elements to the environment, using their voices to overcome obstacles, avoid enemies, or reach the end of a level.

One of the main goals of our game prototype is to create a platform that connects scientists, level designers, and musicians, especially in an educational context. For this reason, we have integrated the ability to design levels using a map editor³ and a tileset⁴. This allows for creative level design without any technical background in programming and gives researchers new and creative ways to test real-time algorithms in the context of a music game. Furthermore, this platform could also be interesting for musicians to explore and practice their musical skills in a creative and entertaining setting.

2 Real-Time Pitch Estimation

In order to use the player’s singing voice as an input controller for the game, we need to estimate the pitch of the singing voice, which is a common task in Music Information Retrieval (MIR). The pitch estimation algorithm used for this task has to meet the following requirements to create a smooth and enjoyable gameplay experience. First, the algorithm must be computed in real time with low latency. Second, the pitch estimation must be accurate. Third, one requires a robust voicing detection that can distinguish between singing and non-singing, as we only want to display a red pitch line when the user is actually

³ <https://www.mapeditor.org>

⁴ <https://pixelfrog-assets.itch.io/pixel-adventure-1>

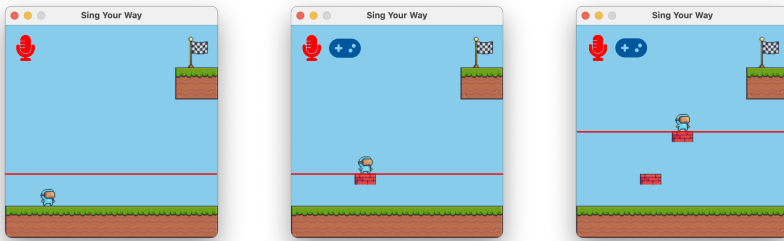


Fig. 2: Gameplay sequence showing how to add note blocks and stairs to the game: Sung notes appear as a red line that can be jumped on. This creates a fixed note block on which the player can stand.

singing⁵. In order to find suitable solutions for these requirements, we will first take a closer look at the general challenges of “Real-Time Music Information Retrieval” (Sect. 2.1) and then focus on two concrete examples of “Pitch Estimation Algorithms” (Sect. 2.2).

2.1 Real-Time Music Information Retrieval

Music Information Retrieval (MIR) is an area of research that focuses mainly on the analysis of large music datasets [ST22]. The algorithms used for these analyses are often not optimized for execution time and assume full availability of all data at all times. Compared to this *offline* processing, *online* algorithms used for real-time applications differ in three important aspects:

- (a) **Causality of information:** For offline analysis, the entire audio signal is already available, and you can move forward or backward in the signal and select analysis windows of any size. For online analysis, only past and present information is available.
- (b) **Trade-off between accuracy and latency:** In general, the more data is available, the more robust an analysis can be. For instance, a DFT with more audio samples results in a higher frequency resolution. However, for online analysis, this also has a direct impact on latency, as more audio samples also mean longer waiting times, especially for causal systems.
- (c) **Signal processing with frames:** Online analysis is done in frames, processing a fixed number of samples as they are provided, for example, by a soundcard. There is usually a requirement to update the analysis result at the same frame rate, so the algorithm has limited time to complete before processing the next frame. As a result, online analysis methods may not be as complex as in the offline case and need to be optimized for execution time.

⁵ General-purpose algorithms for voicing detection may not be able to differentiate between singing and speech since there is no clear boundary between the two. For our user interface, however, speech is not a suitable input as it typically does not allow for fine-grained control of pitch.

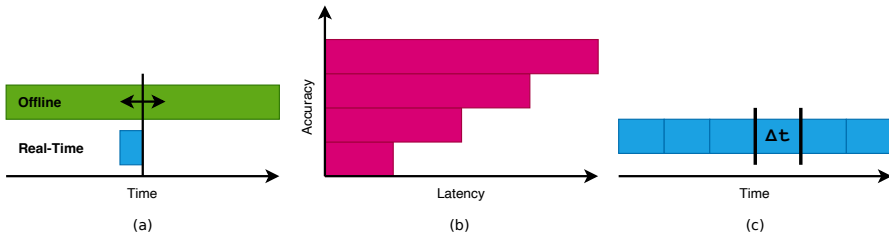


Fig. 3: Illustration of three main differences between offline and real-time Music Information Retrieval. (a) Causality of information. (b) Trade-off between accuracy and latency. (c) Signal processing with frames.

2.2 Pitch Estimation Algorithms

In addition to the general challenges of real-time analysis discussed in the previous section, we now want to take a closer look at the specific challenges of real-time analysis with respect to different pitch estimation algorithms. When we talk about pitch estimation in this article, we actually mean fundamental frequency estimation (F0 estimation), which is often used synonymously with pitch estimation. The basic idea of F0 estimation is to find the signal's fundamental frequency that correlates with the perceived pitch for each time position [CH08; CK02; SG12]. The input to the algorithm is the waveform, spectrogram, or other representation of the signal. The output of the algorithm is the frequency estimate of the input signal and an indication of how confident the algorithm is that the signal is tonal and periodic, called the “confidence value.” This confidence value is often used in combination with a threshold to provide a form of voicing detection, for example, to distinguish singing from non-singing signals. Two methods that compute the values on a frame-by-frame basis, and are therefore in principle suitable for real-time applications, are shown in Fig. 4 and will be discussed in the following section.

One of the most well-known F0 estimation approaches is called YIN [CK02]. The YIN algorithm is based on an autocorrelation method and is calculated entirely in the time domain, using the signal's waveform as input, see also Fig. 4. The algorithm provides accurate pitch estimates for signals with singing, but also incorrectly detects pitches for non-singing parts of the signal. With respect to real-time applications, YIN is relatively simple and can be implemented in an efficient way with low latency. In fact, the offline algorithm can be directly transferred into an online variant, since only one input frame of audio data is required for each output frame of pitch estimation. In addition, YIN can be extended with post-processing steps [MD14], which we would like to investigate more closely for real-time applications in the future. For voicing detection, YIN calculates the so-called “aperiodicity” of the input signal. The lower the aperiodicity, the higher the probability that the signal is periodic and tonal. This is opposite to the confidence, where lower values indicate a lower probability that the signal is periodic and tonal. To allow a

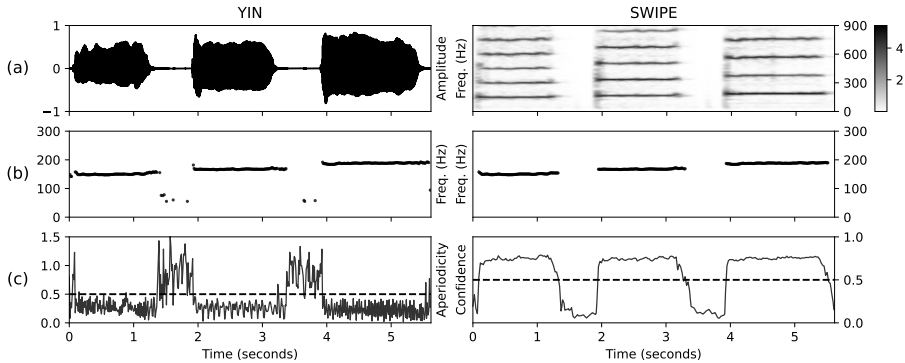


Fig. 4: An overview of two pitch estimation algorithms used for our music game prototype. YIN: (a) Waveform input. (b) Pitch estimation trajectory. (c) Aperiodicity for voicing detection. SWIPE: (a) Spectrogram input. (b) Pitch estimation trajectory. (c) Confidence for voicing detection.

more direct comparison with SWIPE, in this paper we define a confidence = 1 – aperiodicity for YIN, which will be used in the following.

Unlike the YIN time-domain algorithm, SWIPE is based on spectral correlations in the frequency domain [CH08]. To this end, SWIPE uses multiple pitch candidates with a unique fundamental frequency, each represented by a single pre-computed kernel inspired by a sawtooth waveform. The original offline version of the algorithm (SWIPE-Off in the following) transforms the input signal into a time–frequency representation using Short-Time Fourier Transforms (STFTs). Each pitch candidate requires a different optimal window size for the spectral input, which results in more accurate pitch estimates, but also in a higher computational complexity. Pitch estimates are calculated by correlating the spectrum of each input frame with the set of pitch kernels, finding the kernel with the highest correlation. The amount of correlation is also used as a confidence value for voicing detection. In the online version of SWIPE, we use a rolling buffer that stores the previous input frames needed to compute all the different window sizes for each frame. This concept allows for three different online variants of SWIPE (SWIPE-On-A/B/C), which are illustrated in Fig. 5 and described below:

SWIPE-On-A: The most obvious arrangement of windows follows the offline algorithm by keeping the spectra perfectly centered and aligned for each point in time. For a causal real-time system, this results in a delayed output signal because the rolling buffer must be filled completely before computation, see (A) of Fig. 5.

SWIPE-On-B: To compensate for this delay, the windows can be filled in such a way that the most recent frames are always in the center of each window, while keeping the alignment between the different window sizes as in SWIPE-On-A. Any samples in the future are set to zero. As illustrated in (B) of Fig. 5, this causes a reduction in the confidence values.

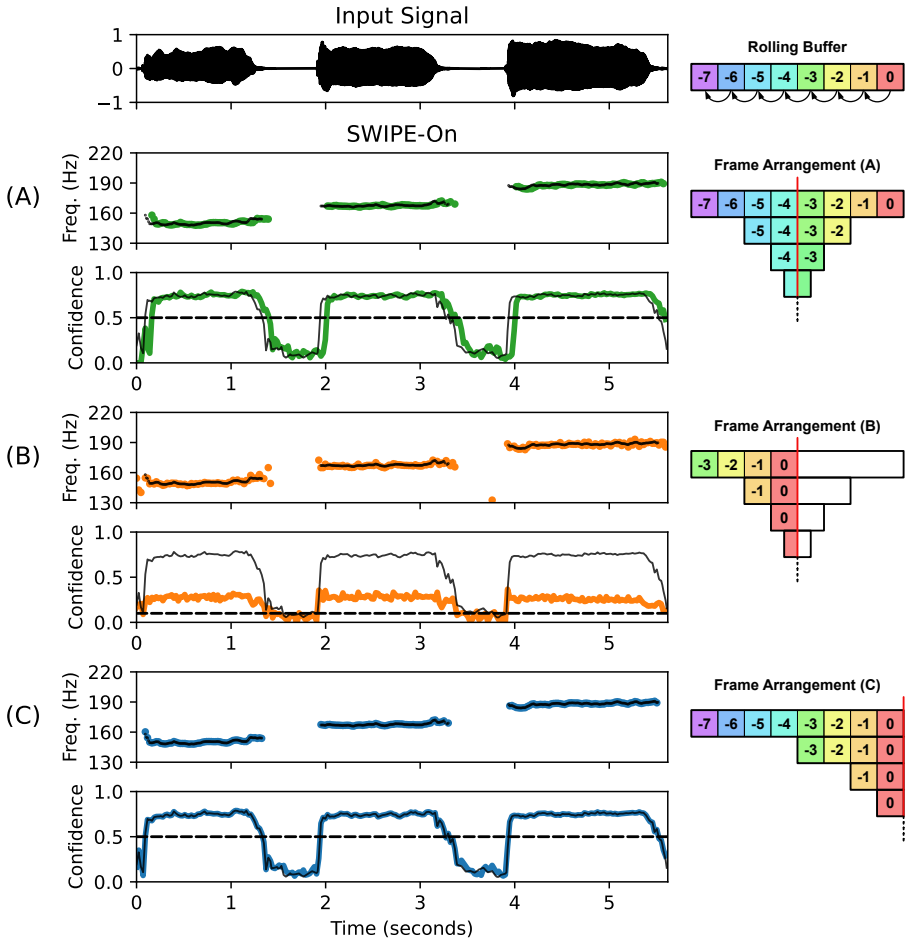


Fig. 5: **Left:** Frequency estimates and confidence values of the SWIPE-On variants (denoted by A, B, C, and colored in green, orange, blue). Values for SWIPE-Off are displayed in black. Confidence thresholds for each variant are indicated by black dashed lines. **Right:** The rolling buffer holds several input frames, symbolized by squares. The numbers indicate the frame index, where 0 is the current frame, -1 is the previous frame, and so on. The frame arrangements for each variant visualize the used signal segment for different window sizes. We only show four rows per frame arrangement, although there are usually much more windows sizes to calculate for SWIPE.

This is because only half-filled windows are available for the spectral correlation with the frequency kernels, lowering the score.

SWIPE-On-C: In the third and final online variant, all the different windows are filled with the most recent audio data. This discards a perfectly centered time alignment of frames between

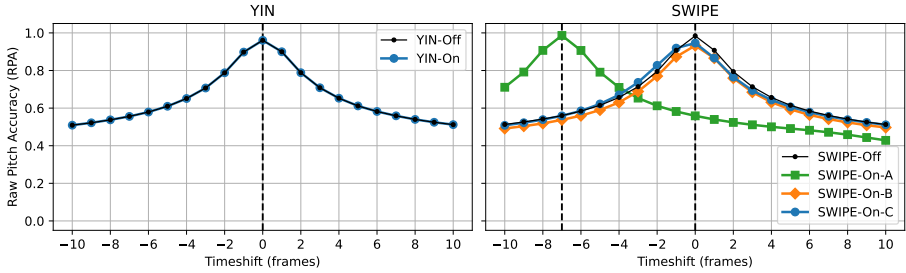


Fig. 6: Average raw pitch accuracy versus the timeshift (frames) of the estimated frequency values, tested with a pitch tolerance of 50 cents.

Metric	RPA		RCA	
	-7	0	-7	0
YIN-Off	0.535±0.172	0.926±0.093	0.556±0.166	0.961±0.056
YIN-On	0.535±0.172	0.926±0.093	0.556±0.166	0.961±0.056
SWIPE-Off	0.554±0.171	0.976±0.092	0.560±0.166	0.984±0.070
SWIPE-On-A	0.978±0.091	0.553±0.171	0.986±0.069	0.559±0.166
SWIPE-On-B	0.521±0.168	0.909±0.116	0.538±0.161	0.933±0.089
SWIPE-On-C	0.556±0.170	0.938±0.101	0.559±0.169	0.946±0.087

Tab. 1: Average raw pitch/chroma accuracies and their standard deviations between different tracks of the dataset for different timeshifts (frames) of the estimated frequency values, tested with a pitch tolerance of 50 cents.

different window sizes and thus between F0 estimates at different frequencies. However, it avoids the problems of (SWIPE-On-A) delayed output and (SWIPE-On-B) reduced confidence, see (C) of Fig. 5.

3 Experiments

For an objective evaluation of the different algorithms, we use the MDB-melody-synth dataset [Sa18]. It consists of 65 different tracks with mainly singing voice melodies derived from the MedleyDB dataset [Bi14]. The MDB tracks were re-synthesized by the method described in [Sa17] to provide perfect ground truth annotations. Thus, MDB-melody-synth is well suited for evaluating real-time algorithms, as the annotations can be easily converted to a frame size of 512 samples and a sampling rate of 44100 Hz, which are also typical real-time audio settings that we use for our game.

In our experiments, we compare a total of six different pitch estimation algorithms: YIN-Off/On, SWIPE-Off, and SWIPE-On-A/B/C. To find a setting that is compatible with all the pitch estimation algorithms as well as the dataset, we run each algorithm with a

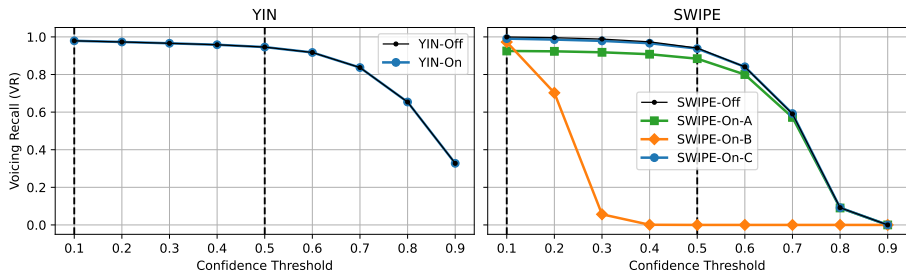


Fig. 7: Average voicing recall rates versus the confidence threshold.

Metric	VR		VFA	
	0.1	0.5	0.1	0.5
YIN-Off	0.979±0.020	0.946±0.070	0.017±0.021	0.013±0.018
YIN-On	0.979±0.020	0.946±0.070	0.017±0.021	0.013±0.018
SWIPE-Off	1.000±0.003	0.941±0.088	0.012±0.010	0.000±0.000
SWIPE-On-A	0.925±0.054	0.884±0.096	0.131±0.106	0.105±0.089
SWIPE-On-B	0.973±0.075	0.000±0.000	0.009±0.010	0.000±0.000
SWIPE-On-C	0.991±0.008	0.938±0.087	0.009±0.010	0.003±0.007

Tab. 2: Average voicing recall/false alarm rates and their standard deviations between different tracks of the dataset for different confidence thresholds.

sampling rate of 44100 Hz, a frame size of 512 samples, a minimum F0 estimate of 45 Hz, and a maximum F0 estimate of 1760 Hz. For the evaluation, we measure the quality of the pitch estimation in Raw Pitch Accuracy (RPA) and Raw Chroma Accuracy (RCA), both with 50 cents pitch tolerance. Furthermore, we measure the quality of the voicing detection in Voicing Recall (VR) and Voicing False Alarm (VFA), as described in [SG12]. For all evaluation metrics, we use the reference implementation of `mir_eval` [Ra14]. In addition, we run all evaluations for different timeshifts (from -10 to 10 frames) and confidence thresholds (from 0.1 to 0.9).

In Fig. 6 we show the Raw Pitch Accuracy (RPA) for estimated frequency values shifted in time between -10 and 10 frames. With Tab. 1 we report both RPA and RCA for timeshifts of both -7 and 0 frames. We see that YIN-Off and YIN-On produce exactly the same accuracy values. SWIPE-Off outperforms YIN-Off by five percentage points with regard to RPA. When shifted -7 frames, SWIPE-On-A achieves similar accuracy as SWIPE-Off. For a zero frames timeshift, SWIPE-On-C comes closest to SWIPE-Off with a drop of 0.038, followed by SWIPE-On-B with a drop of 0.067.

In Fig. 7 we show the Voicing Recall (VR) for different confidence thresholds, and with Tab. 2 we report the values for Voicing Recall (VR) and Voicing False Alarm (VFA). YIN-Off and YIN-On provide exactly the same Voicing Recall (VR). Furthermore, SWIPE-Off

has a perfect Voicing Recall (VR) of 1.0 compared to `YIN-Off` with a drop of 0.021, both with similar Voicing False Alarm (VFA) rates close to zero. `SWIPE-On-B` works only with lower confidence thresholds, where it achieves a similar Voicing Recall (VR) as the other algorithms. The best overall online variant is `SWIPE-On-C` with a Voicing Recall (VR) similar to `SWIPE-Off` and a Voicing False Alarm (VFA) rate close to zero.

4 Conclusion and Discussion

With our music game prototype, we provide a platform that uses real-time pitch estimation as an input to an interactive game. Our general aim is to bring together researchers, level designers, and musicians to explore algorithms, design interesting levels and practice musical skills in a creative context. In this paper, we compared and evaluated several real-time pitch estimation algorithms to be used in our interactive gaming environment, where both the accuracy of pitch estimation and voicing detection are equally important. We have shown that `YIN` transfers well from offline to online and runs very efficiently on the waveform directly, but does not reach the pitch estimation accuracy of `SWIPE`. Compared to `YIN`, `SWIPE` is computationally more complex and requires the additional step of performing multiple STFT operations per frame. Nevertheless, these spectrogram computations are efficient enough to not add noticeable latency to our Python-based test implementation. We presented three different variants for `SWIPE-On`, where (`SWIPE-On-A`) problems with delayed outputs can be solved with timeshifts, and (`SWIPE-On-B`) problems with reduced confidence values can be handled with lower thresholds. We also showed that `SWIPE-On-C` produces the highest overall accuracy for both pitch estimation quality and voicing detection, and is therefore the best option for our gaming scenario. For future research, we would like to look at other pitch estimation algorithms based on both classical engineering and more recent machine learning techniques and extend our case study to other datasets.

5 Acknowledgements

This work was supported by the BayWISS Joint Academic Partnership “Digitalisation.” The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institut für Integrierte Schaltungen IIS.

References

- [Bi14] Bittner, R. M.; Salamon, J.; Tierney, M.; Mauch, M.; Cannam, C.; Bello, J. P.: MedleyDB: A Multitrack Dataset for Annotation-Intensive MIR Research. In: Proceedings of the International Society for Music Information Retrieval Conference (ISMIR). Taipei, Taiwan, pp. 155–160, 2014.

- [CH08] Camacho, A.; Harris, J. G.: A sawtooth waveform inspired pitch estimator for speech and music. *The Journal of the Acoustical Society of America* 124/3, pp. 1638–1652, 2008.
- [CK02] de Cheveigné, A.; Kawahara, H.: YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America (JASA)* 111/4, pp. 1917–1930, 2002.
- [MD14] Mauch, M.; Dixon, S.: pYIN: A Fundamental Frequency Estimator Using Probabilistic Threshold Distributions. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Florence, Italy, pp. 659–663, 2014.
- [Me22] Meier, P.; Schwär, S.; Rosenzweig, S.; Müller, M.: Real-Time MIR Algorithms for Music-Reactive Game World Generation. In: *Mensch und Computer 2022 - Workshopband*. Bonn, 2022.
- [Me23] Meier, P.; Schwär, S.; Krump, G.; Müller, M.: Real-Time Pitch Estimation for Creative Music Game Interaction. In: *Fortschritte der Akustik - DAGA 2023 Hamburg*. Deutsche Gesellschaft für Akustik e.V. (DEGA), Berlin, pp. 1346–1349, 2023, ISBN: 978-3-939296-21-8.
- [Ra14] Raffel, C.; McFee, B.; Humphrey, E. J.; Salamon, J.; Nieto, O.; Liang, D.; Ellis, D. P. W.: MIR_EVAL: A Transparent Implementation of Common MIR Metrics. In: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Taipei, Taiwan, pp. 367–372, 2014.
- [RSM22] Rosenzweig, S.; Schwär, S.; Müller, M.: libf0: A Python Library for Fundamental Frequency Estimation. In: *Late Breaking Demos of the International Society for Music Information Retrieval Conference (ISMIR)*. Bengaluru, India, 2022.
- [Sa17] Salamon, J.; Bittner, R. M.; Bonada, J.; Vicente, J. J. B.; Gómez, E.; Bello, J. P.: An Analysis/Synthesis Framework for Automatic F0 Annotation of Multitrack Datasets. In: *Proceedings of International Society for Music Information Retrieval Conference (ISMIR)*. Suzhou, China, pp. 71–78, 2017.
- [Sa18] Salamon, J.; Bittner, R.; Bonada, J.; Bosch, J. J.; Gómez, E.; Bello, J. P.: MDB-melody-synth, version 1.0.0, Zenodo, Nov. 2018, URL: <https://doi.org/10.5281/zenodo.1481168>.
- [SG12] Salamon, J.; Gómez, E.: Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics. *IEEE Transactions on Audio, Speech, and Language Processing* 20/6, pp. 1759–1770, 2012.
- [ST22] Stefani, D.; Turchet, L.: On the challenges of embedded real-time music information retrieval. In: *Proceedings of the 25-th Int. Conf. on Digital Audio Effects (DAFx20in22)* 3/, pp. 177–184, Sept. 2022.

Perspektiven computergestützter harmonischer Analyse: Beethovens op. 14 Nr. 1 als Gegenstand gattungsübergreifender Korpusanalyse

Stephanie Klauk¹, Rainer Kleinertz², Pascal Schmolenzky³, Christof Weiß⁴, Meinard Müller⁵

Abstract: Mit der zunehmenden Digitalisierung von Notentexten und Einspielungen vermehren sich die Möglichkeiten computergestützter Darstellungen und Analysen. Derartige harmonische Analyseverfahren sollen im Beitrag einem Fassungsvergleich des Kopfsatzes von Beethovens op. 14/1 zugrunde gelegt und mit traditionellen musikwissenschaftlichen Methoden abgeglichen werden. Die daraus abgeleiteten Hypothesen erweitern einerseits den Blick auf Beethovens Bearbeitungspraxis und lassen sich andererseits vom konkreten Fallbeispiel – Klaviersonate versus Streichquartettfassung – auf die entsprechenden Gattungen im Sinne von Korpusanalysen übertragen.

Keywords: Computergestützte harmonische Analyse, Korpusanalyse, Musikverarbeitung.

1 Einleitung

Beethovens Bearbeitungen eigener Werke nehmen eine Sonderstellung in seinem Œuvre ein. Abhängig vom Veränderungsgrad können sie alternative, gar ‚verbesserte‘ Fassungen des Originals darstellen oder mehr oder weniger mechanische Bearbeitungen und Arrangements für andere Besetzungen: Abhängig von der kammermusikalischen Instrumentation unterteilt Unverricht in [Un78] Beethovens Bearbeitungen in drei Gruppen, während Lühning in [Lü92] die sinfonische Musik mit einbezieht. Die kompositorische Herausforderung einer solchen Übertragung in eine andere Gattung ist laut Beethovens eigener Aussage nicht hoch genug einzuschätzen. 1802 schrieb er an den Verlag Breitkopf & Härtel: „Ich habe eine einzige Sonate von mir in ein Quartett von Geigeninstrumenten verwandelt, worum man mich so sehr bat und ich weiß gewiß, das macht mir nicht so leicht ein anderer nach“ [Fi81, 12].

¹ Universität des Saarlandes, Institut für Musikwissenschaft, Campus Geb. C6-3, 66123 Saarbrücken, s.klauk@mx.uni-saarland.de

² Universität des Saarlandes, Institut für Musikwissenschaft, Campus Geb. C6-3, 66123 Saarbrücken, rainer.kleinertz@mx.uni-saarland.de

³ Universität des Saarlandes, Institut für Musikwissenschaft, Campus Geb. C6-3, 66123 Saarbrücken, s8plschr@stud.uni-saarland.de

⁴ Universität Würzburg, Center for AI and Data Science, Emil-Hilb-Weg 23, Hubland Nord (ZPD), 97074 Würzburg, christof.weiss@uni-wuerzburg.de

⁵ Friedrich-Alexander-Universität, International Audio Laboratories Erlangen, Am Wolfsmantel 33, 91058 Erlangen, meinard.mueller@audiolabs-erlangen.de

Die Unterschiede, die sich bei Beethovens Bearbeitung der E-Dur-Sonate op. 14 Nr. 1 (1799) zum Streichquartett in F-Dur (1802) ergaben, waren bereits Gegenstand musikwissenschaftlicher Beiträge: Broyles diskutiert in [Br70] die These, dass op. 14 Nr. 1 ursprünglich als Streichquartett konzipiert wurde, Finscher legt in [Fi81] den Fokus auf die Rezeption der Bearbeitung als „vollgültiges Streichquartett“. Schwager sieht in [Sc73] Änderungen unabhängig von der Idiomatik der Instrumente, Lockwood argumentiert dagegen [Lo98]. Aspekte wie Dynamik, Phrasierung, Satztechnik, Kontrapunktik stehen im Mittelpunkt der Diskussionen, die Harmonik spielt selbst in der Monografie von Sturgis-Everett [St86] für den Vergleich keine Rolle, da dort die geringsten Unterschiede verortet werden.

Im vorliegenden Beitrag sollen audiobasierte Analysen des Kopfsatzes beider Fassungen miteinander verglichen werden. Dabei wird sich zeigen, dass das in [WM21] vorgestellte Verfahren einer auf Skalen basierenden Audioanalyse über die grundverschiedenen Instrumente hinweg stabil bleibt und zugleich eine sinnvolle Ergänzung zu ‚manuellen‘ Annotationen von Experten und auf dem Notentext basierenden automatisierten Analysen darstellt.

2 Computergestützte harmonische Analyse

Computergestützte harmonische Analysen können anhand zweier Arten von Daten durchgeführt werden: ‚symbolisch‘ auf der Grundlage von computerlesbaren Notentexten oder audiobasiert auf der Grundlage von Musikaufnahmen. Im ersten Fall wird der Notentext in digital bearbeitbare Formate übertragen. Auf der Basis solcher symbolischer Daten wurde im Kontext des Akademieprojekts *Beethovens Werkstatt* eine automatisierte Harmonieanalyse sowohl des Kopfsatzes der Klaviersonate op. 14 Nr. 1 als auch des entsprechenden Streichquartettsatzes durchgeführt.⁶ Diese Akkord-Analysen sind online abrufbar in einem von sechs Darstellungsmodi („Harmonievergleich“) eines Fassungsvergleichs von Original und Bearbeitung des Kopfsatzes. In [CS20] werden die Darstellungsverfahren des im Projekt entwickelten Anwenderprogramms VideApp und die vier Perspektiven „Fassungssynopse“, „Bearbeitungsmaßnahmen“, „Einzelnotenvergleich“ und „Stimmenkontur“ besprochen; die weiteren Kategorien „Ereignisdichte“ und „Harmonievergleich“ sind auf der Internetseite des Projekts auswählbar und näher beschrieben [Be23].

Abbildung 1 zeigt die Takte 65 bis 68 des Fassungsvergleichs in der Perspektive „Harmonievergleich“. Die relativen harmonischen Unterschiede sollen automatisiert rot eingefärbt werden. Insbesondere in diesen Beispieltakten liegen sowohl die computergestützte Unterschieds- als auch die Harmonieerkennung falsch: Das Ergebnis

⁶ In [Ne18] sind Beethovens Streichquartette (ohne die Bearbeitung von op. 14 Nr. 1) von Musikwissenschaftlern annotiert und in symbolischer Form zur digitalen Weiterverarbeitung zur Verfügung gestellt. Es handelt sich also nicht um eine computergestützte Harmonieanalyse, sondern die Ergebnisse lassen sich lediglich automatisiert auswerten, siehe [Mo19].

Abb. 1: Screenshot der Benutzeroberfläche der Videapp [Be23], Perspektive „Harmonievergleich“, Takte 65-68

der lokalen Erfassung der Noten der ersten Zählzeit (*c*, *e*) ergibt C-Dur statt a-Moll, und nur a-Moll entspricht dem richtig erkannten b-Moll der transponierten Streichquartettfassung.

Im Gegensatz dazu gehören automatisierte Harmonieanalysen auf der Basis von Audiodaten zum Gebiet der digitalen Audiosignalverarbeitung. Damit beschäftigt sich seit 2014 das zwischen Musikwissenschaft und Informatik angelegte interdisziplinäre DFG-Projekt *Computergestützte Analyse harmonischer Strukturen*, das an der Universität des Saarlandes und den International Audio Laboratories Erlangen durchgeführt wird. Eines der im Projekt entwickelten Analyseverfahren der Kopfsätze von Beethovens Klaviersonaten sowie die entsprechende Visualisierung [KI21] werden auf die Streichquartettbearbeitung übertragen und in Abschnitt 3 diskutiert. Beim gewählten Analyseverfahren handelt es sich um eine aus dem Audio abgeleitete Zeit–Diatonik-Darstellung [WM21]. Dabei werden zunächst die Anteile verschiedener Frequenzen der akustischen Wellenformen innerhalb bestimmter Zeitfenster mit Hilfe einer Spektralanalyse berechnet. Die resultierende Zeit–Frequenz-Darstellung (*Spektrogramm*) wird in eine Zeit–Chroma-Darstellung (*Chromagramm*) überführt, in dem die Energieverteilung des Musiksignals über die zwölf chromatischen Tonhöhenklassen (ohne Enharmonik und Oktavierung zu unterscheiden) und Zeit erfasst wird. Um daraus Intervalle, Akkorde oder, wie in diesem Fall, Skalen ableiten zu können, wird das Chromagramm für jeden Zeitpunkt mit binären Prototypen verglichen, die den Skalen entsprechen. So korrespondiert das E-Dur der Klaviersonate op. 14 Nr. 1 mit der „+4(Kreuz)-Diatonik“, deren Prototyp den Wert 1 für die sieben Tonhöhenklassen E, F#, G#, A, H, C#, D# aufweist, den Wert 0 für die fünf verbleibenden Tonhöhenklassen

F, G, A#, C, D. Mit den 12 Prototypen werden Wahrscheinlichkeiten für die 12 Skalen bzw. Dur- (oder entsprechende Moll-)Tonarten ermittelt. Dies wird in der Zeit–Diatonik-Darstellung visualisiert, bei der die Wahrscheinlichkeiten über ein Graustufen-Schema angezeigt werden: Schwarz entspricht der Wahrscheinlichkeit 1, Weiß entspricht 0; je dunkler der Farbwert ist, umso größer ist die Wahrscheinlichkeit, dass der Tonvorrat an der jeweiligen Stelle der angegebenen Diatonik/Tonart zugeordnet werden kann. Zur besseren Vergleichbarkeit von Klaviersonate und Streichquartett wird der Tonalitätsverlauf in den Abbildungen 2 und 3 relativ zur Grundtonart betrachtet und nach der Diatonik der Grundtonart ausgerichtet. Die Grundtonart wird dabei als „0-Ebene“ definiert, der Tonvorrat der Oberquinttonart etwa als „+1-Ebene“ bezeichnet, der der Unterquinttonart als „-1-Ebene“.

3 Audioanalyse op. 14 Nr. 1: Klaviersonate versus Streichquartett

Die Abbildungen 2 und 3 zeigen jeweils Visualisierungen des in Abschnitt 2 vorgestellten Verfahrens einer automatisierten harmonischen Audioanalyse der Klaviersonate op. 14 Nr. 1, 1. Satz (Einspielung von Daniel Barenboim, EMI Classics 1998; Abb. 2) sowie der Streichquartett-Fassung (Einspielung des Gewandhaus-Quartetts, Nca 2012; Abb. 3). Beethovens Transposition um einen Halbton nach oben ist durch die Darstellung der relativen Diatonik angeglichen, sodass die Visualisierungen unmittelbar miteinander vergleichbar sind.

Die Länge des sich kontinuierlich verschiebenden Zeitfensters beträgt 12 Sekunden, siehe [WM21]. Zunächst fällt auf, dass die x-Achse der Streichquartett-Einspielung (SQ) 376 Sek. (6:16 Min.), die der Sonaten-Einspielung (KS) 408 Sek. (6:48 Min.) umfasst. Dies besagt, dass sich Barenboim eines langsameren Grundtempos bedient als das Gewandhaus-Quartett, und das, obwohl Beethoven die Satzüberschrift „Allegro“ der Klaviersonate in das langsamere „Allegro moderato“ des Streichquartetts abgeändert hat. Die formale Grobstruktur beider Visualisierungen ist jedoch identisch und kann mit der Sonatenhauptsatzform in Verbindung gebracht werden: Deutlich erkennbar ist das sich wiederholende ‚Muster‘ der Exposition (KS: 0-108, 108-216 Sek.; SQ: 0-102, 102-203 Sek.) mit dem jeweiligen Beginn in der Grundtonart („0-Ebene“) sowie der Wechsel zur Oberquinttonart („+1-Ebene“) mit eingeschobenen Abschnitten in der „+2-Ebene“. Dies entspricht den Expositionen der vorangehenden frühen Klaviersonaten Beethovens [KI21]. Die Durchführung sticht durch ihren Beginn in der „-4-Ebene hervor (KS: 216-272 Sek.; SQ: 203-254 Sek.), während für die Reprise mit abschließender Coda der überwiegende Verbleib in der Grundtonart charakteristisch ist (KS 272-408 Sekunden; SQ: 254-376 Sek.).

Kleinere Abweichungen in Form von unterschiedlich hellen Grauwerten lassen sich selbst bei der Wiederholung der Exposition innerhalb derselben Einspielung beobachten.

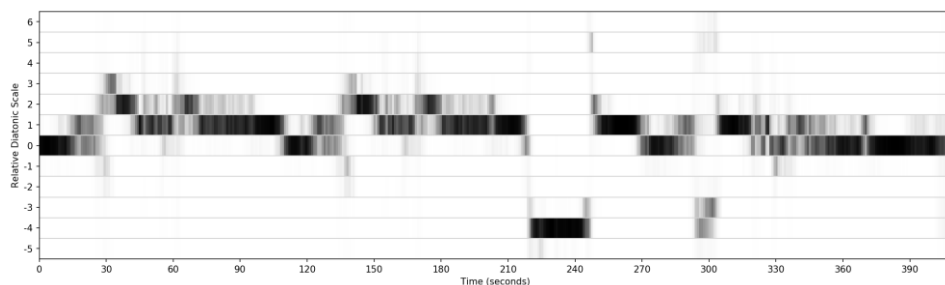


Abb. 2: Zeit-Diatonik-Darstellung einer Einspielung von Daniel Barenboim (EMI Classics 1998) von Beethovens Klaviersonate op. 14 Nr. 1, 1. Satz (relative 0-Diatonik entspricht E-Dur)

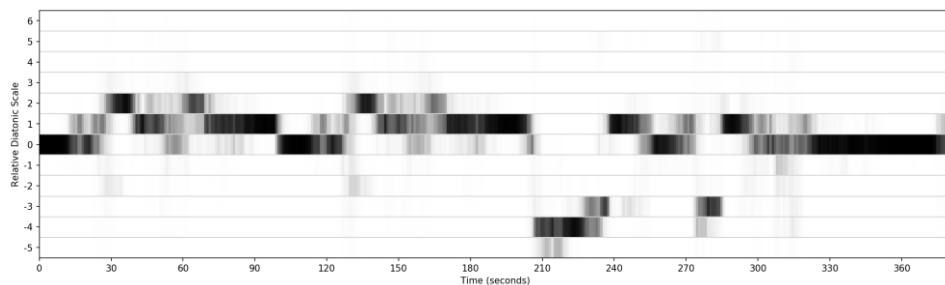


Abb. 3: Zeit-Diatonik-Darstellung einer Einspielung des Gewandhaus-Quartetts (Nca 2012) von Beethovens Bearbeitung der Klaviersonate op. 14 Nr. 1, 1. Satz, zum Streichquartett (relative 0-Diatonik entspricht F-Dur)

Dies kann auf leichte Unterschiede in der Dynamik oder Spieltechnik zurückgeführt werden. Eine der wenigen markanteren Abweichungen wird am Anfang der Durchführung deutlich: Während für die Klaviersonate zwischen Sek. 220-246 mit der überwiegenden Schwarzfärbung eine sehr hohe, durchgehende Wahrscheinlichkeit für die „-4-Ebene“ (C-Dur/a-Moll) angezeigt wird, ist die Parallelstelle der Streichquartett-Einspielung (Sek. 206-232) differenzierter. Die „-4-Ebene“ (Des-Dur/b-Moll) wird mit höchster Wahrscheinlichkeit lediglich zwischen Sek. 219-228 angegeben, davor und danach wird eine ähnliche Wahrscheinlichkeit auch für die „-5-“ und die „-3-Ebene“ abgebildet. Hat Beethoven in diesem Abschnitt einschneidende harmonische Veränderungen für die Streichquartettfassung vorgenommen, die aus dem Notentext hervorgehen? Tatsächlich weist die Streichquartettfassung einige kleinere harmonische Abweichungen gegenüber der Klaviersonate auf. So antizipiert in T. 75 in der zweiten Takthälfte die Viola bereits den Grundton des folgenden Des-dur-Akkords in T. 76, der hier als dissonierende Quarte erklingt, und in T. 81 wechselt im Streichquartett die Harmonik von Des-dur zum verminderten Dominantseptnonakkord (Dv) von C-dur mit

tiefalterierter Quinte *des*, während die Klavierfassung hier jeweils eine konstante Sechzehntel-Begleitung aufweist. Der Grund für die Abweichungen bei der Audio-Analyse scheint aber vor allem in den gebrochenen Dreiklängen im Cello mit anschließender Pause (T. 66-71) und des Cellos und der Viola in T. 76-79 sowie den repetierenden Sechzehnteln in den Mittelstimmen zu liegen. Während die Harmonien in der Klavierfassung hier konstant in arpeggierten Sechzehnteln erklingen, werden in den Streichern die einzelnen, über zwei bis drei Oktaven verteilten Töne unterschiedlich deutlich wahrgenommen und durch die Wiederholungen eine größere Anzahl an für die jeweilige Skala relevanten Tönen erfasst. So sind *ges* und *es* in T. 67-68 zentral und für Grauwerte in der „-5-Ebene“ verantwortlich, der Ton *as* in T. 73-74 für die „-3-Ebene“.

4 Fazit und Ausblick

Das vorgestellte automatisierte Verfahren zur harmonischen Audioanalyse hat sich unabhängig von der Gattung bzw. von der Instrumentierung als robust erwiesen. Wenige markante Abweichungen in den entsprechenden Visualisierungen der Klaviersonate und der Streichquartett-Fassung sind offenbar nicht instrumentenspezifisch, sodass das Verfahren zur gattungsübergreifenden Korpusanalyse geeignet ist. Während Musikwissenschaftler (teilweise auch automatisierte harmonische Analysen, die auf symbolischen Daten basieren) zu dem Schluss kommen, dass Beethoven die Harmonik im Kopfsatz der beiden Fassungen von op. 14 Nr. 1 nicht grundlegend verändert hat, erkennt das Verfahren der harmonischen Audioanalyse durchaus einige Unterschiede. Dies korrespondiert mit der auditiven Wahrnehmung: Ohne sich auf konkrete Harmonien beziehen zu können, wirkt die Durchführung des Streichquartetts auf das Ohr deutlich strukturierter. Die Audioanalyse bietet eine Perspektive musikalischer Analyse, die sich nicht aus dem Notentext, aber dennoch aus statistischen Werten und objektiver Messung ergibt. In zukünftigen Arbeiten sollen symbolische und audiobasierte Harmonieanalysen bereits im automatisierten Verarbeitungsschritt kombiniert werden, um eine Optimierung und Stabilisierung der computergestützten Harmonieanalyse insgesamt zu erzielen.

Danksagung. Wir danken der Deutschen Forschungsgemeinschaft (DFG) für die finanzielle Unterstützung (MU 2686/7-2, KL 864/4-2). Die International Audio Laboratories Erlangen sind eine gemeinsame Einrichtung der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und des Fraunhofer Instituts für Integrierte Schaltungen IIS.

Literatur

- [Be23] Beethovens Werkstatt, VideAppArrangements, <https://videapp-arr.beethovens-werkstatt.de>, Stand: 02.05.2023.
- [Br70] Broyles, M. E.: Beethoven's Sonata Op. 14, Nr. 1. Originally for Strings?. *Journal of the American Musicological Society* 23/3, S. 405-419, 1970.

- [CS20] Cox, S.; Sanger, R.: Digitale Fassungsvergleiche am Beispiel von Beethovens Eigenbearbeitungen. In (Acquavella-Rauch, S.; Munzmay, A.; Veit, J. Hrsg.): Bruckenschlage zwischen Musikwissenschaft und Informatik. Theoretische und praktische Aspekte der Kooperation. Musikwissenschaftliches Seminar der Universitat Paderborn und der Hochschule fur Musik Detmold, Detmold, S. 97-104, 2020.
- [Fi81] Finscher, L.: „Das macht mir nicht so leicht ein anderer nach“. Beethovens Streichquartettbearbeitung der Klaviersonate op. 14 Nr. 1. In (Staehelin, M. Hrsg.): Divertimento fur Hermann J. Abs. Beethoven-Haus, Bonn, S. 11-23, 1981.
- [KI21] Klauk, S. et.al.: ‘Seitensatz’ versus ‘Mittelsatz’: Expositionen in Beethovens fruhem Klaviersonaten zwischen zeitgenossischer Theorie und computergestutzter Analyse. In (Hohmaier, S. Hrsg.): Jahrbuch des Staatlichen Instituts fur Musikforschung Preuischer Kulturbesitz 2017. Schott Music, Mainz, S. 271-300, 2021.
- [Lo98] Lockwood, L.: Beethoven as Colourist: Another Look at his String Quartet Arrangement of the Piano Sonata, Op. 14 No. 1. In (Brandenburg, S. Hrsg.): Haydn, Mozart, and Beethoven. Studies in the Music of the Classical Period. Essays in Honour of Alan Tyson. Clarendon Press, Oxford, S. 175-180, 1998.
- [Lu92] Luhning, H.: Beethoven als Bearbeiter eigener Werke. In (Fischer, J. Hrsg.): Munchener Beethoven-Studien. Katzbichler, Munchen u.a., S. 117-127, 1992.
- [Mo19] Moss, F. C. et.al.: Statistical Characteristics of Tonal Harmony: A Corpus Study of Beethoven’s String Quartets. PloS ONE 14/6, 2019. <https://doi.org/10.1371/journal.pone.0217242>
- [Ne18] Neuwirth, M. et.al.: The Annotated Beethoven Corpus (ABC): A Dataset of Harmonic Analyses of All Beethoven String Quartets. Frontiers in Digital Humanities. Sec. Digital Musicology 5, 2018. <https://doi.org/10.3389/fgdh.2018.00016>
- [Sc73] Schwager, M.: A Fresh Look at Beethoven’s Arrangements. Music & Letters 54, S. 142-160, 1973.
- [St86] Sturgis-Everett, B. A.: The First Movements of Beethoven’s „Piano Sonata in E major, op. 14, no. 1“ and his „String Quartet in F major, op. 14, no. 1“. A Critical Comparison. University of Cincinnati, Cincinnati, 1986.
- [Un78] Unverricht, H.: Original und Bearbeitung. Ein Beitrag zu Beethovens eigenen Bearbeitungen seiner Kammermusikwerke. In (Klein, R. Hrsg.): Beethoven-Kolloquium 1977: Dokumentation und Auffuhrungspraxis. Barenreiter, Kassel u.a., S. 190-196, 1978.
- [WM21] Weiss, C.; Muller, M.: Computergestutzte Visualisierung von Tonalitatsverlaufen in Musikaufnahmen. Moglichkeiten fur die Korpusanalyse. In (Klauk, S. Hrsg.): Instrumentalmusik neben Haydn und Mozart. Analyse, Auffuhrungspraxis und Edition. Konigshausen & Neumann, Wurzburg, S. 107-130, 2021.

Kultur Design - Designing
alternative future home stories

Designing alternative future home stories


Benedikt Haupt ¹, Christian Pentzold ², Alexa Becker ³ and Arne Berger ⁴


Abstract: Usually the smart home imaginaries portray nuclear families living in detached houses with a focus on efficiency and energy. In this workshop, we follow the co-design approach to imagine alternative future homes beyond this stereotype. First, we brainstorm the concept of *home* addressing idiosyncratic needs and expectations. Second, we try three co-design tool adaptations in small groups which allow us to collaboratively imagine speculative futures. *LittleBoxes* is a cultural probe variation with which participants can configure a speculative future by using inspirational and technological materials. The *Tiles IoT Inventor Toolkit* adaptation helps to tell detailed stories of how to possibly live with technology in the future. With the *IoT Service Kit* adaption participants are able to build speculative home scenarios and playfully explore imaginary future technologies within them. There will be two rounds of co-designing, so every participant will explore two of the toolkits. In the last part, we reflect on the experiences as well as discuss if and how participants could imagine adapting the introduced tools beyond the workshop.


Keywords: Smart Home; Co-Design; Participatory Design; Internet of Things; toolkit; design fiction


1 Background

Current smart home innovations often advertise solutions for comfort, security, and efficiency [Li21], [LPH21], [Mc23]. The underlying image of the inhabitants of such homes is very limited and they are often portrayed as a family living in a detached house [De19], [Mc23]. By treating smart home inhabitants only as users and consumers we perpetuate the underlying heteronormative and capitalist values already inscribed in prevailing smart home visions [Be22]. In doing so, the actual needs and wishes of people as well as the personal idiosyncrasies of homes are not reflected in smart home

¹ Leipzig University, Department for Communication and Media Studies, Nikolaistraße 27-29, Leipzig, Germany, 04109, benedikt.haupt@uni-leipzig.de,  <https://orcid.org/0000-0002-2344-3491>

² Leipzig University, Department for Communication and Media Studies, Nikolaistraße 27-29, Leipzig, Germany, 04109, christian.pentzold@uni-leipzig.de,  <https://orcid.org/0000-0002-6355-3150>

³ Leipzig University, Department for Communication and Media Studies, Nikolaistraße 27-29, Leipzig, Germany, 04109, alexa.becker@uni-leipzig.de,  <https://orcid.org/0000-0002-9761-3211>

⁴ University of Applied Science Anhalt, Department 5 - Computer Science and Languages, Lohmannstr. 23, Koethen, Germany, 06366, arne.berger@hs-anhalt.de,  <https://orcid.org/0000-0002-6398-839X>

technologies. But how can future smart homes be envisioned and materialized so as to consider alternative forms of living as well as the individual needs of people in their home?

In recent years HCI and design researchers have developed a multitude of approaches and methods in attempting to broaden the view and gain a deeper understanding of how to move beyond narrow market driven smart home schemes [DWO15], [LPH21], [SAO19]. These approaches draw on collaborative creative processes [SS14] with the aim to create new technological setups for what they consider alternative homes.

In this workshop, we will build on the approaches of speculative design and co-design ventures [LC15], [SS08] as well as on work of human-computer interaction and research-through-design. The twist of this workshop however is, rather than demonstrating a newly developed method of co-creatively addressing the aforementioned issues regarding smart home visions, we adapt existing co-design methods and shift the focus to a narrative perspective which enables participants to speculate and tell stories about their individual visions of future smart home technologies.

By scaffolding a storytelling process participants are empowered to concretize their imaginaries and detail the diegetic contexts in which speculative technologies can be used without flattening the rich field of possible future modes of IoT-enhanced dwelling [Be22]. Furthermore, by demonstrating our narrative adaptations of co-design approaches we want to show the versatility such methods can yield if they are appropriated and tweaked for a certain research interest.

2 Workshop

2.1 Aim of Workshop

Legen In the workshop, we challenge conceptions of home. By sharing and reflecting individual perspectives, we attempt to uncover the underlying assumptions in order to gain new insights and further develop everyone's individual concept of home.

The workshop will address the following questions a) What/where is home? b) What is a standard home? c) What is an alternative home? d) What makes home home? e) How can we imagine and narrate future homes? f) How can alternative imaginaries of smart homes be materialized? g) What roles does/can technology play in the home? h) What makes the smart home 'smart'? i) What does this implicate for the interactions with technologies in the home?

To think beyond contemporary imaginaries of smart home technologies, three co-design tools and methods will be provided. Participants have the opportunity to explore and try out two of the methods and speculate on their future homes in a playful manner. The tools enable the participants to express their wishes and needs regarding how they want to live with technology in a home of the future. The methodical scaffolding supports participants

in imagining and exploring smart home scenarios in novel ways to create stories about the speculative everyday life in them.

In detail, the three tools that will be available in the workshop are: *LittleBoxes* – a cultural probe variation with which participants can configure a speculative future by using inspirational and technological materials [GDP99]. The *Tiles IoT Inventor Toolkit* adaptation which enables people to tell detailed stories of how to possibly live with technology in the future [MGD17], [PI23]. With the *IoT Service Kit* adaption participants are able to build speculative home scenarios and playfully explore imaginary future technologies within them [Io23].

2.2 Brief Schedule of Workshop

The workshop will be held in-person on Friday, 29.09.23, 10 am to 5 pm. The workshop will be in German, as are all the tools and methods.

- Morning 10 am to 11 am
 - Welcome and introduction to the workshop, hosts & participants
 - Brief and general overview on co-design methods and toolkits for the IoT
 - Brainstorming about the concept of the home, guided by four questions
- Coffeekick 11 am to 11.30 am
- Noon 11.30 am to 12.30 am
 - Participants will be split in small groups
 - First round of try-outs with one of the tools and methods
 - Participants can choose between an adaption of The Tiles IoT Inventor Toolkit, an adaption of the IoT Service Kit and a cultural probe called LittleBoxes
 - Each method will be facilitated by one of the hosts
- Lunchbreak 12.30 am to 2 pm
- Afternoon 2 pm to 3.30 pm
 - Participants will be split in small groups
 - Second round of try-outs with one of the tools and methods
 - Participants can choose between an adaption of The Tiles IoT Inventor Toolkit, an adaption of the IoT Service Kit and a cultural probe called LittleBoxes
 - Each method will be facilitated by one of the hosts

- Coffeebreak 3.30 pm to 4.30 pm
- Reflection & Wrap Up 4.30 pm to 5 pm

We encourage participants to reflect with us about their experiences of the workshop & to connect with each other beyond the workshop

3 Organizers

Benedikt Haupt: I am a research associate at the Department for Communication and Media Studies at the University of Leipzig. Coming from a qualitative social science background my interests became greatly linked during my studies to human machine interactions as well as the role technology plays in our everyday life. As a relatively new member of a transdisciplinary team working with co-design, I am fascinated by the variety and diversity of methods and approaches in this field. Therefore, my motivation mostly derives from the tension between how methods are perceived in the field of co-design and co-creation in contrast to the perception of methods in the qualitative media and communication science and how these perspectives could potentially benefit from each other.

Christian Pentzold: I am Chair and Professor of Media and Communication in the Department for Communication and Media Studies at the University of Leipzig. At Leipzig, I'm one of the co-directors of the Center for Digital Participation. I'm broadly interested in the construction and appropriation of digital media and the roles information and communication technologies play in modern society. As part of my work on the interaction between people and smart digital technologies, I am maintaining a living lab that involves citizens in workshops and lectures, creating tinkering, and co-design events. Next, as a researcher and scholar, I have been involved in peer production since 2005 where people from all walks of life come to create information goods as well as tangible technologies beyond the governance mechanisms and legal constraints of markets or formal organizations.

Alexa Becker: I am a research associate at the Department for Communication and Media Studies at the University of Leipzig with an interdisciplinary background in literature, art, media and computer science. My research interests are co-design, tangibles and micro-interactions especially in the context of the smart home. During the last years I conducted multiple co-design workshops for the smart home in different settings with a variety of people. As a relatively new member in academic research my motivation is to gain more experience and to get in touch with other researchers.

Arne Berger: I am a Professor of Human-Computer Interaction at Hochschule Anhalt. HeI am fascinated by the complex, idiosyncratic and unintended interactions between humans and digital technology. My work is influenced by the Scandinavian tradition of Participatory Design, which recognizes that those who will be affected by a future technology should have an active say in its creation. My research focuses on early phases

of design and development processes and he is particularly interested in how errors, failures, blips, and oversights shape how we think about future technology.

4 Post-Workshop Plans

During the workshop the process will be audio recorded and photographed. The results will be documented on post-its and canvas as well as on photos. The photographed results will be shared with the workshop participants. A data privacy statement will be provided as well as a table to fill in the email address if the participant would like to receive the documentation of the workshop.

References

- [Be22] Becker, A.; Haupt, B.; Berger, A.; Pentzold, C.: Future home stories: participatory predicaments and methodological scaffolding in narrative speculation on alternative domestic lives, *Digital Creativity*, 2022. <https://doi.org/10.1080/14626268.2022.2082488>
- [De19] Desjardins, A.; Viny, J.; Key, C.; Johnston, N.: Alternative Avenues for IoT: Designing with Non-Stereotypical Homes. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, pp. 1-13, 2019. <https://doi.org/10.1145/3290605.3300581>
- [DWO15] Desjardins, A.; Wakkary, R.; Odom, W.: Investigating Genres and Perspectives in HCI Research on the Home. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 3073-3082, 2015. <https://doi.org/10.1145/2702123.2702540>
- [GDP99] Gaver, B.; Dunne, T.; Pacenti, E.: Design: Cultural probes. *interactions* 6/1, pp. 21-29, 1999. <https://doi.org/10.1145/291224.291235>
- [Io23] IoT Service Kit, Creating IoT Services has never been easier, <https://iotservicekit.com> accessed 22/02/2023.
- [LC15] Lindley, J.; Coulton, P.: Back to the Future: 10 Years of Design Fiction. Proceedings of the 2015 British HCI Conference, pp. 210-211, 2015. <https://doi.org/10.1145/2783446.2783592>
- [Li21] Li, W.; Yigitcanlar, T.; Erol, I.; Liu, A.: Motivations, Barriers and Risks of Smart Home Adoption: From Systematic Literature Review to Conceptual Framework. *Energy Research & Social Science* 80, 2021. <https://doi.org/10.1016/j.erss.2021.102211>
- [LPH21] Lupton, D.; Pink, S.; Horst, H.: Living in, With and Beyond the 'Smart Home'. *Convergence* 27/5, pp. 1147-1154, 2021. doi:10.17/13548565211052736
- [Mc23] McKinsey & Company, The Connected Smart Home Market, https://www.mckinsey.com/spcontent/connected_homes/pdf/mckinsey_connectedhome.pdf, accessed 22/02/23

- [MGD17] Mora, S.; Gianni, F.; Divitini, M.: Tiles: A Card-based Ideation Toolkit for the Internet of Things. Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17), pp. 587-598, 2017. <https://doi.org/10.1145/3064663.3064699>
- [PI23] Play with Tiles IoT Inventor Toolkit. Change the world!, <https://www.tilestoolkit.io> accessed 22/02/2023
- [SAO19] Shin, J.; Aceves Sepúlveda, G.; Odom, W.: Collective Wisdom: Inquiring Into Collective Homes as a Site for HCI Design. Proceedings of the 2019 CHI Conference on Human Factors in Computing systems, pp. 1-14, 2019. <https://doi.org/10.1145/3290605.3300546>
- [SS08] Sanders, E. B.-N.; Stappers, P. J.: Co-creation and the new landscapes of design. *CoDesign*, 4/1, pp. 5-18, 2008. <https://doi.org/10.1080/15710880701875068>
- [SS14] Sanders, E. B.-N.; Stappers, P. J.: Probes, toolkits and prototypes: three approaches to making in codesigning, *CoDesign*, 10/1, pp. 5-14, 2014. <https://doi.org/10.1080/15710882.2014.888183>

**Offentliche Infrastruktur - Research
Data Infrastructures for Data
Science and AI (RDI4DataScience)**

NFDI4DS at a Glance

Sonja Schimmler,¹ Christine Hennig¹

Abstract: The consortium NFDI4DS supports researchers along all stages of the research data lifecycle to conduct their research in line with the FAIR principles. By conducting interviews and surveys, NFDI4DS continuously identifies the needs and challenges of researchers from various disciplines regarding data science and artificial intelligence, keeping ethical, legal, and social aspects in mind. Those identified needs and challenges are continuously addressed by picking up existing services, developing new ones and integrating them into the NFDI4DS infrastructure. By systematically adding digital objects (articles, data, machine learning models, workflows, scripts/code, etc.) to the NFDI4DS research knowledge graph within the infrastructure, transparency, reproducibility, and fairness are steadily improved. The process is continuously accompanied by providing resources such as educational videos and organizing events such as community challenges.

In this short paper, we give an overview of NFDI4DS, and provide details about our approach to address the current challenges. We will report on how we plan to utilize FAIR digital objects (FDOs) and Research Knowledge Graphs (RKGs) as a basis for the infrastructure envisioned. We will also give an overview of the services planned, and how they are meant to interact.

Keywords: NFDI; NFDI4DS; Data Science; Artificial Intelligence; Research Data Infrastructures

1 Introduction

The past years have seen a paradigm shift, with computational methods increasingly relying on data-driven and often deep learning-based approaches, leading to the establishment of data science as a discipline driven by advances in the field of computer science. Transparency, reproducibility and fairness have become crucial challenges for data science (DS) and artificial intelligence (AI) due to the complexity of contemporary DS methods, often relying on a combination of scripts/code, models, workflows and data used for training. NFDI4DS will promote fair and open research data infrastructures supporting all involved resources such as scripts/code, workflows, models, data, or articles through an integrated approach.

The vision of NFDI4DS is to support all steps of the complex and interdisciplinary research data lifecycle, including collecting/creating, processing, analyzing, publishing, archiving, and reusing resources in DS and AI.

The overarching objective of NFDI4DS is the development, establishment, and sustainment of a national research data infrastructure (NFDI) for the DS and AI community in Germany.

¹ Fraunhofer FOKUS, Berlin, Germany

sonja.schimmler@fokus.fraunhofer.de, christine.hennig@fokus.fraunhofer.de

This will also deliver benefits for a wider community requiring data analytics solutions, within the NFDI and beyond. The key idea is to work towards increasing the transparency, reproducibility and fairness of DS and AI projects, by making all digital artifacts available, interlinking them, and offering innovative tools and services. Based on the reuse of these digital objects, this enables new and innovative research.

NFDI4DS intends to represent the DS and AI community in academia, which is an interdisciplinary field rooted in computer science. We aim to reuse existing solutions and to collaborate closely with other NFDI consortia and beyond. In the beginning, the consortium will focus on four DS intense application areas: language technology and natural language processing, biomedical research and clinical decision-making, information sciences and social sciences. The expertise available in NFDI4DS ensures that metadata standards are interoperable across domains and that new ways of dealing with digital objects arise.

2 Challenges and Approach to Address Challenges

Sharing scientific knowledge is not just about publishing articles. Instead, it involves documenting the entire *research data lifecycle* and providing a multitude of digital objects in compliance with the FAIR principles by making them *findable, accessible, interoperable* and *reusable*. As DS and AI are continuously evolving, the methods used become more complex, and it is difficult to maintain transparency, reproducibility, and fairness in research. Challenges related to ethical, legal, or social aspects further limit the willingness and/or ability of researchers to conduct, archive, or publish their research in line with the FAIR principles.

NFDI4DS is part of the NFDI initiative to build a German National Research Data Infrastructure. It supports all stages of the complex and interdisciplinary research data lifecycle to enable the efficient and effective reuse of research data and other digital objects. These range from humanities and social sciences, life sciences, and natural sciences, to engineering sciences. The research data lifecycle includes six stages: collection/creation of data, its processing, analysis, preservation, access, and reuse. Additionally, NFDI4DS steadily contributes to establishing best practices in research, fostering Open Science to enable researchers to make full use of valuable resources.

NFDI4DS is organized around six task areas: (1) Community and Training, (2) Research Knowledge Graphs, (3) Infrastructure and Services, (4) Transfer and Application, (5) Interoperability and Cooperation, and (6) Management. In addition, working groups are temporarily set up in case they are needed.

To achieve its goals, the consortium is continuously striving for an integrated approach to research data management. The infrastructure is composed of already existing tools and services as well as new ones developed within the scope of this initiative and beyond serving specific purposes to enable all stages of the research data lifecycle. These are embedded in a

framework making the individual tools freely available, and easily usable without extensive prior knowledge.

By regularly conducting interviews, insights are gathered on the needs and challenges, especially about ethical, legal, and social aspects. In addition, yearly surveys identify gaps for new implementations, as well as tools and services that already exist and are useful for the initiative. As DS and AI are important in many disciplines, with often contradicting requirements, building an infrastructure is a collaborative effort including the scientific communities. By regularly providing benchmark datasets and fostering joint work on shared tasks interdisciplinary as well as domain-specific services and solutions are achieved. Each shared task focuses on a specific aspect of the research data lifecycle and has the goal to initiate a concrete service that is being integrated into the NFDI4DS infrastructure later on. Besides the development and integration of tools, another important aspect of the consortium are learning materials, best practice guidelines, tutorials, workshops, and hackathons.

The NFDI4DS core services focus on different steps within the research data lifecycle utilizing digital objects. These digital objects include artifacts such as articles, data, machine learning models, workflows, and scripts/code, building upon the emerging concept of FAIR Digital Objects. The NFDI4DS Research Knowledge Graph will form the basis of the infrastructure, providing details about digital objects and their interrelations. Key elements of the infrastructure will be the NFDI4DS gateway and portal as well as the NFDI4DS registries and repositories. Digital objects will be harmonized, aggregated, and preserved via the repositories and exposed via the registries and the gateway and portal. The gateway and portal will offer different functionalities to search and explore the knowledge base as well as provide recommendation and evaluation services e. g. for benchmarking or assessing quality, bias, harm and FAIRness. They will utilize facets allowing browsing datasets, methods, or tools complying with certain community-driven requirements.

3 Conclusion

In this short paper, we gave an overview of NFDI4DS, and provided details about our approach to address the current challenges. We reported on how we plan to utilize FAIR Digital Objects (FDOs) and Research Knowledge Graphs (RKGs) as a basis for the infrastructure envisioned. We also gave an overview of the services planned, and how they are meant to interact.

Acknowledgements

This work has received funding through the German Research Foundation (DFG) project NFDI4DS (no. 460234259).

Community and Training in NFDI4DS

Anna-Lena Lorenz¹, Maria Christoforaki², Christine Hennig³, Angelie Kraft⁴, Stephanie von Maltzan⁵, Sonja Schimmler³

Abstract: Key to NFDI4DS's success is an active and vibrant community as establishing a common data culture and practice relies on the community's participation and acceptance. We address this challenge by leveraging the network of NFDI4DS partners to raise awareness for topics around FAIR data and establish international standards. By identifying requirements, we improve our services and develop new strategies for building and finding user communities.

Keywords: NFDI; NFDI4DS; Community; Training; Requirements Engineering; ELSA

1 Introduction

Due to the broad scope spanning Data Science (DS) and Artificial Intelligence (AI) in different fields, the NFDI4DS community is rather diverse compared to other NFDI consortia. Skillsets and prior knowledge vary a lot, even between people of the same career level. A crucial aspect is thus to get a thorough understanding of who the community consists of and what their needs are. Therefore we continuously elicit, gather and analyze requirements in interviews and surveys via our various communication channels and in-person events.

2 Outreach

We regularly host events targeting the different stakeholder groups - from traditional Lecture Series to more dynamic approaches like our Science Slam that already took part twice in the context of the Berlin Science Week. This event was open to the general public and invited perspectives from all NFDI consortia.

To connect to our community online, we also run several communication channels. We started with a Twitter account¹ and are now also on Mastodon², as part of our community moved there due to the recent developments. We can also be found on LinkedIn³ and for

¹ TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany, anna.lorenz@tib.eu

² University Hospital of Cologne, Germany, maria.christoforaki@uk-koeln.de

³ Fraunhofer FOKUS, Berlin, Germany

sonja.schimmler@fokus.fraunhofer.de, christine.hennig@fokus.fraunhofer.de

⁴ Universität Hamburg, Germany, angelie.kraft@uni-hamburg.de

⁵ FIZ Karlsruhe, Germany, stephanie.maltzan@fiz-karlsruhe.de

¹ <https://twitter.com/nfdi4ds>

² <https://nfdi.social/@nfdi4ds>

³ <https://www.linkedin.com/company/nfdi4ds>

hosting our lecture series and training videos, we have a YouTube channel⁴. We are also on Zenodo⁵, GitLab⁶ and GitHub⁷, since our community is active there.

While building a community is a crucial first step, training researchers and practitioners is equally important. We thus also develop interactive training material facilitating well established platforms and services within the community and curricula. We are currently in the process of creating a video series to introduce our consortium and our services. Not only do we develop material, but we also organize and participate in community events on different topics regarding data management, data ethics and data protection.

All of our created material is given a common branding. Since our beginning we have a clear corporate design consisting of an orange-pink color scheme and a stylized tree with nodes as a logo. This design emphasises the modernity of the topic via its fresh, young colors and symbolizes the interconnection between services, data and communities which is crucial for our work. Since our material is as broadly diverse as our communities, this is a unique way of creating a uniform framework, while allowing for formats, services and training material to be as manifold and specific as necessary. And of course, the design can not only be found in our digital material, but also on our merchandise that we hand out at in-person events to spread the word. Currently we have stickers, pencils, blocks, sticky notes, and t-shirts. Our slogan is "We are the missing link", as we believe that NFDI4DS truly provides a unique unified approach for all research data of our community.

3 Personas

To not think about our community as an abstract concept, but of a heterogeneous group of people with different skill sets and requirements, we developed personas - fictional individuals - with their distinct wants and needs. The first input for those personas was developed during our first in-person consortium meeting with several people from different scientific backgrounds and career levels present. In this first session, we introduced three fictional people. A computer science PhD student, a post-doc from biomedical research and a professor in social sciences. The participants were asked to take the perspective of these people based on their own experiences in the respective fields or career stages. We then collected answers to three broad questions: (1) What are this person's tasks? (2) What are their problems and obstacles in executing these tasks? (3) What do they wish for? What could help them with the aforementioned problems?

As a first step, we encouraged the participants to not only answer these questions with specific services in mind but rather generally. After this brainstorming session, we clustered

⁴ <https://www.youtube.com/@nfdi4ds>

⁵ <https://zenodo.org/communities/nfdi4ds>

⁶ <https://gitlab.com/nfdi4ds>

⁷ <https://github.com/nfdi4ds>

the answers into different categories. We excluded the points that can not be addressed by our consortium, such as the wish for more money, and identified the main issues.

In total we determined six personas for our core stakeholder groups from these inputs. At our next meeting we then presented the personas and aligned them with our services to ensure that we can offer added value to these fictional people and also for our services to keep in mind. During that process, we identified missing information, unclear descriptions and redundancies, which then left us with our refined version of the personas:

- **Alex** is a PhD student in computer science. He spends a lot of time searching for and cleaning datasets. For some of his tasks he wants to re-use sensitive data from another group. Sometimes he struggles to access this data and often he wishes there would be a system to help him find, access and semi-automatically clean data. Especially the last point would reduce his workload significantly. Moreover, better search options supporting different content types would help him.
- **Ben** does his PhD in biomedicine and searches for papers and methods for his task. He often gets lost in the flood of publications and does not have time to read all papers. Additionally, often methods and data have incomplete descriptions. He wishes for a better documentation as well as an overview of the current state of his research area.
- **Cassie** just started her PhD in social sciences and thus still has a steep learning curve. She wishes for better training material tailored to her needs.
- **David** is in the early postdoc phase of his career in life sciences. He wants to compare his results with some experimental data in the field, but often is missing standards for sharing data. He also wishes for more computational time and power.
- **Emma** plans to run a journal and thus has to deal with a lot of editorial tasks. She needs to find a publisher, acquire suitable reviewers and keep up with the latest community news in terms of topics, papers and people. The ongoing reviewer fatigue does not make her situation easier. A system that informs her about current developments and also suggests possible reviewers would be useful for her.
- **Finn** is a professor in social sciences and responsible for his group's data management plans. In his proposals he wants to fulfill the expectations, but sometimes doing open science and protecting sensitive data seem mutually exclusive. He needs clearer guidelines and a safe and secure storage for his group's sensitive data that should still be linked to executable code.

For these fictional people we develop pathways of how they can navigate our website and services. We keep these people's needs in mind when designing and optimizing services. Moreover, these personas are not only used for development. We also create marketing material such as videos tailored to their use cases. We have a set of secondary stakeholders as well, including e.g. developers who use our services as a back-end to their own product.

4 Ethical, Legal and Social Aspects

When working with data, it is crucial to keep ethical, legal and social aspects (ELSA) in mind. Especially when scraping data or training models on sensitive or personal data, e.g. patients' medical diagnoses or survey answers, researchers need to make sure to comply with data protection laws and copyright restrictions. Furthermore, other European regulations, particularly the forthcoming AI Act, must be taken into account. Nevertheless, open science and the use of personal data do not mutually exclude each other. We want to support open data by giving out guidelines and developing curricula to train researchers on these issues. A key step is to understand the legal and ethical challenges our community faces. Therefore, we started conducting interviews with researchers at all career levels and from all our disciplines. Based on the results of the interviews we will be able to make recommendations on how to address these challenges.

Storing and accessing personal data is a challenge faced by many NFDI consortia, but in the domain of AI, another topic arises: In times of generative AI like ChatGPT and StableDiffusion researchers have not only to think about copyright or privacy issues, but also about the societal implications of their work. Choices and jobs that have traditionally been done by people are gradually assigned to algorithms. But as these algorithms are only as fair as their training data, there is a danger of teaching the model an implicit bias and increasing the risk of discriminating people treated by the algorithms. NFDI4DS aims to identify and address these issues. For that we regularly host events such as the Weizenbaum Forum on 'Chat GPT, Stable Diffusion and Co', which recently took place.

In addition, we are currently producing a video series about important concepts and questions regarding AI ethics. This series will be a knowledge resource for practitioners in DS- and AI-related fields and will facilitate critical reflections of our present and future with AI. The videos are comprised of interviews with renowned experts in the fields of ethics, sociology, computer science, design, and law. A wide range of topics are covered, such as the EU AI Act, trustworthiness, explainability, and algorithmic bias.

5 Conclusion

NFDI4DS builds and trains a community of DS and AI researchers and practitioners to establish a common culture around FAIR data, data ethics and data management. The focus of our outreach measures is on our services. By providing easy entry points and showing pathways, we simplify data management as much as possible. To train researchers, we develop curricula and participate in community events regularly.

Acknowledgements

This work has received funding through the German Research Foundation (DFG) project NFDI4DS (no. 460234259).

Research Knowledge Graphs in NFDI4DS

Saurav Karmakar,¹ Matthäus Zloch,¹ Fidan Limani,² Benjamin Zopilko,¹ Sharmila Upadhyaya,¹ Jennifer D'Souza,³ Leyla Jael Castro,⁴ Georg Rehm,⁵ Marcel R. Ackermann,⁶ Harald Sack,⁷ Zeyd Boukhers,⁸ Sonja Schimmler,⁹ Danilo Dessì,¹ Peter Mutschke,¹ Stefan Dietze^{1 10}

Abstract: The ever-increasing amount of research output through scientific articles requires means to enable transparency and a better understanding of key entities of the research lifecycle, referred to as research artifacts, such as methods, software, datasets, etc. Research Knowledge Graphs (RKG) make research artifacts findable, accessible, interoperable, and reusable (FAIR) and facilitate their interpretability. In this article, we describe the role of RKGs, from their construction to the expected benefits, including an overview and a vision of their use within the German National Research Data Infrastructure (NFDI) consortium NFDI4DataScience (NFDI4DS). This paper includes insights about the existing RKGs, how to formally represent research artifacts, and how this supports better transparency and reproducibility in data science and artificial intelligence. We also discuss key challenges, such as RKG construction, and integration, and give an outlook on future work.

Keywords: NFDI; NFDI4DS; Data Science; Research Knowledge Graph; Scholarly Data; Knowledge Graph Integration; Knowledge Graph Federation

1 Introduction

Scientific results are typically disseminated through unstructured research publications that describe problems, outline methodologies, present contributions, and discuss experimental results. However, processing information from unstructured documents is costly and requires substantial efforts, making it challenging to actually understand the state-of-the-art in a respective field and the relations and dependencies between tasks, datasets, methods, machine learning models, or overall performance. This problem has been elevated by the shift towards increasingly data-driven and deep learning-based systems, which are known to suffer from dramatic reproducibility issues and often do not outperform simple baselines despite being considered state-of-the-art approaches [DCJ19]. Behind these significant challenges in the field of data science and AI is a lack of transparency about the inherent

¹ GESIS – Leibniz Institute for the Social Sciences, Cologne, Germany saurav.karmakar@gesis.org

² ZBW – Leibniz Information Centre for Economics, Kiel, Germany

⁵ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Berlin, Germany

³ TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany

⁴ ZB MED Information Centre for Life Sciences, Cologne, Germany

⁶ DBLP computer science bibliography, Schloss Dagstuhl - Leibniz Center for Informatics, Trier, Germany

⁷ FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, Karlsruhe, Germany

⁸ Fraunhofer Institute for Applied Information Technology -FIT-, Sankt Augustin, Germany

⁹ Fraunhofer Institute for Open Communication Systems -FOKUS-, Berlin, Germany

¹⁰ Heinrich Heine University Düsseldorf, Germany

dependencies of data, models and code as being used in increasingly complex data processing pipelines.

The current practices as part of the scientific process still involve substantial efforts in searching, browsing and parsing unstructured documents to find relevant resources and insights, where the increasing amount of scholarly literature only increases required efforts. This has further increased the need to represent, store, and document research outcomes, including intricate and complex relationships among research questions, models, data, results, and all-produced research artifacts. Making research outcomes available implies making the entire research lifecycle transparent and all relevant research artifacts findable, accessible, interoperable, and re-useable in the best sense of the FAIR Data Principles [Wi16, Ba22]. While some solutions suggest prompting authors to annotate certain research entities (e.g., contributions) during the creation of their articles [BBK23], the variable nature of manuscripts pre-publication leaves no assurance that authors will consistently and diligently annotate these entities. Therefore, novel and scalable means for extracting and representing digital research artifacts are crucial for the preservation, exploration, use, and reuse of research.

Knowledge graphs (KGs) are being widely used to represent data in a machine-actionable fashion following established principles, such as the FAIR principles and W3C standards, and are particularly focused on capturing relations between entities, usually expressed as RDF triples. KGs have been used already to represent and manage information in scientific contexts, where examples include Nanopublications¹ [GGV10], SoftwareKG [SZK20], Springer SciGraph², Computer Science KG [De22b] and the Open Research Knowledge Graph (ORKG) [Au20, Ja19]. We do refer to such applications of KGs for research data and information management as *Research Knowledge Graphs* (RKG).

Building on these foundations, the NFDI4DataScience (NFDI4DS)³ consortium, part of the German National Research Data Infrastructure (NFDI) initiative, aims at addressing the aforementioned challenges with respect to reproducibility and transparency in data science and AI through an integrated RKG infrastructure. The goal is to establish an integrated infrastructure that provides efficient access to DS and AI resources, most notable, data, code, and machine learning models, and to provide transparent information about their interdependencies. This will be achieved through both building on established RKGs and constructing novel RKGs, e.g. through scholarly information extraction models applied to large corpora of data science and AI literature. This work gives an overview of current work and future directions within this context.

¹ <https://nanopub.org/>

² <https://www.springernature.com/gp/researchers/scigraph>

³ <https://www.nfdi4datascience.de/>

2 Research Knowledge Graphs in NFDI4DS

RKGs allow for structuring relations between research artifacts and other entities (e.g., authors, organizations, etc.), making them machine-understandable and searchable. They are also a means toward re-usability and further machine-actionability of the structured information. RKGs are typically generated from different kinds of (meta)data sources. They can serve various use cases, such as search across a specific knowledge domain, understanding of entities dependencies, use of data within computer-based approaches, etc. Examples include searching for research datasets, understanding research methods, their use, and dependencies on other research artifacts (code, used datasets, etc.), as well as understanding scientific discourse. Thus, RKGs enable researchers to explore a research field from various facets by exploiting their inherent linkage structures.

Category	RKG	Scope
quality-controlled ground truths datasets	SoMeSci [Sc21]	software mentions in scientific articles, currently consisting of about 4.4 million triples, describing metadata and context of 3,756 mentions in 1,367 articles
scholarly resource metadata	GESIS KG Infrastructure ⁴	tools for constructing RKGs of research information, metadata and primary research data at GESIS, i.e., the GESIS Search KG [Hi19]
scholarly resource metadata	DBLP KG [db23]	metadata about 6.7 million computer science publication entities and 3.2 million author entities, 362 million RDF triples in total (Jun. 2023)
primary research data	TweetsKB [Fa18]	metadata about 1.5 billion tweets (collected Feb. 2013 - Mar. 2018)
primary research data	TweetsCOVID19 [Di20]	subset of TweetsKB containing COVID-related tweets reflecting the societal discourse about COVID-19 on Twitter (Oct 2019 - Apr 2020)
primary research data	ClaimsKG [Te19, Ga23]	claims and their evaluation from fact-checking websites, currently holding about 28 thousand claims from six English-language websites
automatically (NLP-)generated resource relations	SoftwareKG [Sc22]	information about software mention statements from more than 51,000 scientific articles from the social sciences
automatically (NLP-)generated resource relations	CS-KG [De22a, De22c]	large-scale automatically generated knowledge graph composed by over 350 million RDF triples describing 41 million statements from 6.7 million articles about 10 million entities (tasks, methods, materials, metrics, etc.) linked by 179 semantic relations
community expression of scholarly papers	The Open Research Knowledge Graph [Au20, Ja19]	structured semantic descriptions of articles, crowd-sourced from authors and researchers, incl comparisons of contributions, visualizations, leaderboards

Tab. 1: Overview of RKGs in NFDI4DS

The NFDI4DS RKGs aim at providing knowledge about research artifacts and their relationships in a well-structured representation. Next to providing individual research knowledge graphs, NFDI4DS will focus on integrating individual research artifacts (e.g., datasets) within each respective category and into a joint research knowledge graph infrastructure that preserves their identity and provenance. A particular focus is on sharing data as well as methods and infrastructure for extracting and curating scholarly knowledge.

The RKG infrastructure of NFDI4DS will be based upon existing RKGs and resource collections targeting different types of entities and purposes, which typically can be classified into the following categories: (i) RKGs built on quality-controlled ground truth datasets, (ii) RKGs built on scholarly resource metadata, (iii) RKGs built on primary research data, (iv) RKGs built on automatically (NLP-)generated resource relations, and (v) RKGs built on community expression of scholarly papers. For an overview of the RKGs in NFDI4DS and their categorization, we refer you to Table 1. Additionally, efforts will also be made to establish and include new RKGs that will be developed, e.g., by the methods generated in the shared tasks (see Section 3).

3 Scholarly Information Extraction and RKG Construction

There are different ways RKG construction takes place depending on the category the RKG belongs to and the purpose it serves. For example, RKGs can be generated by lifting a structured database (e.g., DBLP), manually annotating high-quality-controlled corpora (e.g., SoMeSci), by information extraction (IE) using a supervised mechanism (e.g., SoftwareKG), and also by community efforts as the case of ORKG. Depending on the mechanism an RKG is generated, it bears pros and cons with respect to costs, quality, and scalability. An important effort in NFDI4DS is about scholarly IE methods. In fact, a shared task is in place within NFDI4DS for software mention, MLModel, and dataset mention detection which aims at (i) collecting insights and reproducible methodologies from the ML and DS community, and (ii) further promoting the automatic extraction of research artifacts. A major task in NFDI4DS remains the enrichment of existing RKGs with entities (representing research artifacts) and relations extracted from scholarly papers and the provision of ground truth datasets based on scholarly IE. Recent IE tasks on scholarly publications are:

- **Classification of Scholarly Publications:** This IE task aims at fostering discoverability by multi-class single-label research field classifiers trained on a manually built benchmark dataset based on the ORKG research field taxonomy. For the area of DS and AI research, a specific dataset will be provided to further extend the classifier to solve a multi-class multi-label classification on a more fine-grained taxonomy of research fields.
- **Software Mention Detection:** To advance transparency of the software used in research initial efforts are made to provide an RKG containing articles that link to the software cited, e.g., as done in SoftwareKG [Sc22]. However, as software mentions are usually

informal and incomplete we aim at advancing the field of software mention detection and disambiguation through novel methods.

- **Leaderboard Extraction:** *Leaderboards* aim to aggregate a few related research artifacts, such as tasks, models, datasets, evaluation metrics, and scores [Ho19]. The goal is to automatically mine leaderboards for empirical AI research [KDA21, KDA23a, KDA23b] and provide them in a machine-readable way to enable scientists to better keep track of research progress.
- **Machine Learning and Dataset Mention Detection:** Today, two of the most important entities in scientific articles are *Machine Learning (ML) models* and *datasets*. In an era where DS has an ever-increasing contribution of data to scholarly communication, data is processed and produced by various ML paradigms in different experimental scenarios. This paves the way for intricate relationships between these two entity types such as “related to”, “derived from”, “cited by”, etc.
- **Metadata Extraction:** Linking methodologies, datasets, and results to their respective authors and research papers is crucial. However, the task is challenging due to the variety of paper templates [Bo21]. Hence, the objective is to leverage multimodal techniques for extracting author metadata and for parsing and extracting references [BAS19].

In addition to IE activities, there exist other initiatives for constructing RKGs from existing data. An example of this is the GESIS Data Search⁵ platform, consisting of research and survey studies data provided to users through a search portal. Very recently, an effort has been made to thoroughly exploit these data via metadata mining and structuring them as an RKG. Finally, a further activity is lifting the GESIS Methods Hub into a RKG. The Methods Hub is being developed within the Digital Behavioral Data⁶ project about reproducible and reusable methods, that aims to support social scientists to reuse state-of-the-art methods on newly collected data. The Methods Hub RKG will provide methods as first-class research output, and enable their search and citability.

4 NFDI4DS Registry and RKGs Integration

Given the variety of RKGs and their use cases, we envision an NFDI4DS Registry to make them searchable and available to the community. More precisely, we plan to launch an NFDI4DS Registry platform that will present the RKGs along different categories (e.g., scholarly resource metadata, primary research data, etc. as presented in Section 2.), including their documentation, schemas, example queries, etc. Additionally, the NFDI4DS Registry will provide links to the SPARQL endpoints that can be used to explore the RKGs. However, just providing the RKGs is not sufficient to meet the demand for using RKGs to address

⁵ <https://datasearch.gesis.org/start>

⁶ <https://www.gesis.org/en/institute/digital-behavioral-data>

cross-domain research questions; therefore, a particular focus is set on RKG integration and related topics such as schema integration, linking, federation, use of PIDs, among others. In this regard, we pursue the following activities:

Schema Integration. One of the main goals of the NFDI4DS initiative is to make research digital objects FAIR. To achieve this, we are investigating schemas and vocabularies to formally represent research entities, their semantics across disciplines (i.e., an entity can have different meanings in different disciplines), and RKGs' intended uses (e.g., find machine learning models applied on a dataset, describe research trends, etc.). In detail, this activity involves the alignment of current classes and properties among the schemas and vocabularies that are already in use in the existing RKGs (e.g., SoftwareKG, GESIS Search KG, etc.). As an outcome, we plan to develop a unified schema that describes existing RKGs that can be reused for other RKGs (e.g., new RKGs that are being built within the NFDI4DS activities). We will deliver mappings among the existing schemas and provide core schema elements that can be reused in a variety of scenarios. The schema alignment will enable a common representation and search possibilities (e.g., the same query can be used to explore more than one RKG), making the exploration and use of research digital objects easier. In addition, one more expected benefit is that services that are tailored to a specific RKG can be adopted with less or zero costs to other RKGs.

Interlinking and Enrichment. Different RKGs can describe the same entity from different perspectives. For instance, the conceptual entity *Random Forest* is described by both SoftwareKG and CS-KG. In SoftwareKG, *Random Forest* is described based on its mentions in research papers and the kind of mention i.e., use, creation, deposition, and allusion; in CS-KG the same entity is described by its connection to other entities e.g., <cskg:random_forest, skos:broader, cskg:machine_learning_method>, <cskg:random_forest, cskg-onto:methodUsedBy, cskg:classification>, etc. Linking various RKGs is an essential activity within NFDI to take advantage of all the current RKGs developments. This will be done based on the type of content RKGs describe (e.g., by exploiting connections to Semantic Web hubs like DBpedia⁷ and Wikidata⁸, identifiers for authors e.g., ORCIDs, DOIs for papers, etc.). The linking methodology will be based on the content of the various RKGs and their intended use. The linking will enable a more complete representation of the entities, thus allowing RKG users to get a better understanding of the entities and their scope. To achieve this, we plan to explore automatic methods based on semantic similarity and relatedness in order to scale to the millions of entities that the various RKGs listed in Table 1 contain. As a deliverable, we plan to provide links among RKGs to enable federation and enrich the overall knowledge about the described entities.

Persistent Identifiers. When it comes to linking entities and resources from different RKGs, persistent identifiers (PIDs) play an important role. A PID inventory and assessment at GESIS revealed that different RKGs involve heterogeneous PID systems, namespaces, and URI schemas, ranging from third-party ensured DOIs to informal, self-hosted URIs. The

⁷ <https://www.dbpedia.org/>

⁸ https://www.wikidata.org/wiki/Wikidata:Main_Page

situation in NFDI4DS is similar, where formal PIDs and DOIs (e.g. used to refer to scientific literature and datasets as e.g. in DBLP KG [db23]) and informal URIs (like e.g. in the SoftwareKG [Sc22]) coexist. To overcome these inconsistencies, we will generate a PID strategy that serves as a guideline and gives recommendations on the assignment of PIDs to entities and resources, e.g. which PID system (externally hosted like DOIs or internally hosted like URIs) to use for which entity types/resource types, and how to organize the joint existence of PIDs like DOIs and URIs for the same entities.

5 Conclusion and Outlook

This paper has summarised current work and future directions for building up an integrated RKG infrastructure for scientific resources in data science and AI, most notably, data, code, and machine learning models. Building on a rich set of established RKGs, datasets, and corpora, involving key resources such as the DBLP bibliographic repository, NFDI4DS aims at enabling a more integrated and machine-interpretable view of research information and resources. The RKG infrastructure will form the foundation for searching and using resources, understanding the dependencies, and more generally, provide a transparent view on the very dynamic research landscape in the field of artificial intelligence.

Another challenge is the need for a community-based agreement on shared PIDs and vocabularies used to identify and describe the research artifacts within the RKG infrastructure. While there is a wealth of vocabularies currently in use, consensus-finding processes are required to establish jointly shared and understand vocabularies and standards. An activity to be taken into account here is the initiative on a common NFDI core ontology⁹ aiming to facilitate FAIR access to research data across all scientific domains within the NFDI program [Ti23]. As an outcome for 2024, we plan to provide a web-based registry on which all NFDI4DS RKGs are presented along the different categories and are accessible, e.g., via their SPARQL endpoints. This registry will also provide the elaborated schema mappings between different RKGs, exemplary federated SPARQL queries, computed links between resources/entities of different RKGs as well as a joint NFDI4DS PID strategy to ensure share practices for naming and identifying resources within DS and AI.

Acknowledgements

This work has received funding through the German Research Foundation (DFG) project NFDI4DS (no. 460234259).

Bibliography

[Au20] Auer, Sören; Oelen, Allard; Haris, Muhammad; Stocker, Markus; D'Souza, Jennifer; Farfar, Kheir Eddine; Vogt, Lars; Prinz, Manuel; Wiens, Vitalis; Jaradeh, Mohamad Yaser;

⁹ NFDI core ontology: <https://ise-fizkarlsruhe.github.io/nfdicore/>

Improving access to scientific literature with knowledge graphs. *Bibliothek Forschung und Praxis*, 44(3):516–529, 2020.

- [Ba22] Barker, Michelle; Chue Hong, Neil P; Katz, Daniel S; Lamprecht, Anna-Lena; Martinez-Ortiz, Carlos; Psomopoulos, Fotis; Harrow, Jennifer; Castro, Leyla Jael; Gruenpeter, Morane; Martinez, Paula Andrea et al.: Introducing the FAIR Principles for research software. *Scientific Data*, 9(1):622, 2022.
- [BAS19] Boukhers, Zeyd; Ambhore, Shriharsh; Staab, Steffen: An end-to-end approach for extracting and segmenting high-variance references from pdf documents. In: 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL). IEEE, pp. 186–195, 2019.
- [BBK23] Bless, Christof; Baimuratov, Ildar; Karras, Oliver: SciKGT_X—A LaTeX Package to Semantically Annotate Contributions in Scientific Publications. In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2023*. 2023.
- [Bo21] Boukhers, Zeyd; Beili, Nada; Hartmann, Timo; Goswami, Prantik; Zafar, Muhammad Arslan: MexPub: Deep Transfer Learning for Metadata Extraction from German Publications. In: 2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL). IEEE, pp. 250–253, 2021.
- [db23] dblp team: , dblp Knowledge Graph. Download, June 2023. Monthly RDF release of June 2023, <https://dblp.org/rdf/release/dblp-2023-06-01.nt.gz>.
- [DCJ19] Dacrema, Maurizio Ferrari; Cremonesi, Paolo; Jannach, Dietmar: Are we really making much progress? A worrying analysis of recent neural recommendation approaches. In (Bogers, Toine; Said, Alan; Brusilovsky, Peter; Tikk, Domonkos, eds): *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019*. ACM, pp. 101–109, 2019.
- [De22a] Dessì, Danilo; Osborne, Francesco; Recupero, Diego Reforgiato; Buscaldi, Davide; Motta, Enrico: CS-KG: A Large-Scale Knowledge Graph of Research Entities and Claims in Computer Science. In: *The Semantic Web - ISWC 2022 - 21st International Semantic Web Conference*. volume 13489 of *Lecture Notes in Computer Science*. Springer, pp. 678–696, 2022.
- [De22b] Dessì, Danilo; Osborne, Francesco; Recupero, Diego Reforgiato; Buscaldi, Davide; Motta, Enrico: SCICERO: A deep learning and NLP approach for generating scientific knowledge graphs in the computer science domain. *Knowledge-Based Systems*, 258:109945, 2022.
- [De22c] Dessì, Danilo; Osborne, Francesco; Recupero, Diego Reforgiato; Buscaldi, Davide; Motta, Enrico: SCICERO: A deep learning and NLP approach for generating scientific knowledge graphs in the computer science domain. *Knowl. Based Syst.*, 258:109945, 2022.
- [Di20] Dimitrov, Dimitar; Baran, Erdal; Fafalios, Pavlos; Yu, Ran; Zhu, Xiaofei; Zloch, Matthäus; Dietze, Stefan: TweetsCOV19 - A Knowledge Base of Semantically Annotated Tweets about the COVID-19 Pandemic. In: *Proceedings of the 29th ACM International Conference on Information and Knowledge Management. CIKM '20*, Association for Computing Machinery, New York, NY, USA, p. 2991–2998, 2020.
- [Fa18] Fafalios, Pavlos; Iosifidis, Vasileios; Ntoutsis, Eirini; Dietze, Stefan: TweetsKB: A Public and Large-Scale RDF Corpus of Annotated Tweets. In (Gangemi, Aldo; Navigli, Roberto; Vidal, Maria-Esther; Hitzler, Pascal; Troncy, Raphaël; Hollink, Laura; Tordai, Anna; Alam, Mehwish, eds): *The Semantic Web*. Springer International Publishing, Cham, pp. 177–190, 2018.

- [Ga23] Gangopadhyay, Susmita; Boland, Katarina; Dessí, Danilo; Dietze, Stefan; Fafalios, Pavlos; Tchechmedjiev, Andon; Todorov, Konstantin; Jabeen, Hajira: Truth or Dare: Investigating Claims Truthfulness with ClaimsKG. In: *Linked Data-driven Resilience Research 2023. CEUR Workshop Proceedings*. volume 3401, 2023.
- [GGV10] Groth, Paul; Gibson, Andrew; Velterop, Jan: The anatomy of a nanopublication. *Information Services & Use*, 30(1-2):51–56, 2010.
- [Hi19] Hienert, D.; Kern, D.; Boland, K.; Zapilko, B.; Mutschke, P.: A digital library for research data and related information in the social sciences. In: *Proceedings of 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. pp. 148–157, 2019.
- [Ho19] Hou, Yufang; Jochim, Charles; Gleize, Martin; Bonin, Francesca; Ganguly, Debasis: Identification of Tasks, Datasets, Evaluation Metrics, and Numeric Scores for Scientific Leaderboards Construction. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, pp. 5203–5213, July 2019.
- [Ja19] Jaradeh, Mohamad Yaser; Oelen, Allard; Farfar, Kheir Eddine; Prinz, Manuel; D’Souza, Jennifer; Kismihók, Gábor; Stocker, Markus; Auer, Sören: Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge. In: *Proceedings of the 10th International Conference on Knowledge Capture. K-CAP ’19*, Association for Computing Machinery, New York, NY, USA, p. 243–246, 2019.
- [KDA21] Kabongo, Salomon; D’Souza, Jennifer; Auer, Sören: Automated Mining of Leaderboards for Empirical AI Research. In: *Towards Open and Trustworthy Digital Societies: 23rd International Conference on Asia-Pacific Digital Libraries, ICADL 2021, Virtual Event, December 1–3, 2021, Proceedings 23*. Springer, pp. 453–470, 2021.
- [KDA23a] Kabongo, Salomon; D’Souza, Jennifer; Auer, Sören: ORKG-Leaderboards: A Systematic Workflow for Mining Leaderboards as a Knowledge Graph. *arXiv preprint arXiv:2305.11068*, 2023.
- [KDA23b] Kabongo, Salomon; D’Souza, Jennifer; Auer, Sören: Zero-shot Entailment of Leaderboards for Empirical AI Research. In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2023*. 2023.
- [Sc21] Schindler, David; Bensmann, Felix; Dietze, Stefan; Krüger, Frank: SoMeSci- A 5 Star Open Data Gold Standard Knowledge Graph of Software Mentions in Scientific Articles. In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management. CIKM ’21*, Association for Computing Machinery, New York, NY, USA, p. 4574–4583, 2021.
- [Sc22] Schindler, David; Bensmann, Felix; Dietze, Stefan; Krüger, Frank: The role of software in science: a knowledge graph-based analysis of software mentions in PubMed Central. *PeerJ Computer Science*, 8, 2022.
- [SZK20] Schindler, David; Zapilko, Benjamin; Krüger, Frank: Investigating Software Usage in the Social Sciences: A Knowledge Graph Approach. In (Harth, Andreas; Kirrane, Sabrina; Ngonga Ngomo, Axel-Cyrille; Paulheim, Heiko; Rula, Anisa; Gentile, Anna Lisa; Haase, Peter; Cochez, Michael, eds): *The Semantic Web. Lecture Notes in Computer Science*, Springer International Publishing, Cham, pp. 271–286, 2020.

- [Tc19] Tchechmedjiev, A.; Fafalios, P.; Boland, K.; Gasquet, M.; Zloch, M.; Zapilko, B.; Dietze, S.; Todorov, K.: ClaimsKG: A Live Knowledge Graph of Fact-Checked Claims. In: 18th International Semantic Web Conference (ISWC19). 2019.
- [Ti23] Tietz, Tabea; Bruns, Oleksandra; Söhn, Linnaea; Tolksdorf, Julia; Posthumus, Etienne; Steller, Jonatan Jalle; Fliegl, Heike; Norouzi, Ebrahim; Waitelonis, Jörg; Schrade, Torsten; Sack, Harald: From Floppy Disks to 5-Star LOD: FAIR Research Infrastructure for NFDI4Culture. In: Proc. of the 3rd Workshop on Metadata and Research (objects) Management for Linked Open Science, DaMaLOS 2023. PUBLISSO - Fachrepositorium Lebenswissenschaften (FRL), May 2023.
- [Wi16] Wilkinson, Mark D; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

NFDI4DS Infrastructure and Services

Sonja Schimmler,¹ Bianca Wentzel,¹ Arnim Bleier,² Stefan Dietze,² Saurav Karmakar,² Peter Mutschke,² Angelie Kraft,³ Tilahun A. Taffa,³ Ricardo Usbeck,³ Zeyd Boukhers,⁴ Sören Auer,⁵ Leyla Jael Castro,⁶ Marcel R. Ackermann,⁷ Thomas Neumuth,⁸ Daniel Schneider,⁸ Ziawasch Abedjan,⁹ Atif Latif,¹⁰ Fidan Limani,¹⁰ Raia Abu Ahmad,¹¹ Georg Rehm,¹¹ Sima Attar Khorasani,¹² Matthias Lieber¹²

Abstract: NFDI4DataScience (NFDI4DS) is a consortium founded to support researchers in all stages of the research data lifecycle in order to conduct their research in line with the FAIR principles. The infrastructure developed targets researchers from a wide range of disciplines working in the field of data science and artificial intelligence. NFDI4DS contributes to systematically understanding the needs and challenges of researchers in various disciplines regarding data science and artificial intelligence, keeping in mind ethical, legal and social aspects. The identified needs will be addressed by support structures such as educational videos and challenges. Transparency, reproducibility and FAIRness will be improved by integrating existing and newly developed services into the NFDI4DS infrastructure, and by systematically adding all digital objects (articles, data, machine learning models, workflows, scripts/code, etc.) to the NFDI4DS research knowledge graph. This paper presents the goals of NFDI4DS, and gives an overview on what the consortium is going to contribute to the data science and artificial intelligence communities. It focuses on existing and newly developed services and their integration.

Keywords: NFDI; NFDI4DS; Research Data Infrastructures

1 Introduction

Making research available to other researchers does not only mean to publish articles. It means making the research lifecycle and all digital objects *findable, accessible, interoperable and reusable* (FAIR) [Wi16]. As data science (DS) and artificial intelligence (AI) continue to evolve and the methods used for data collection/creation, processing and analysis become

¹ Fraunhofer Institute for Open Communication Systems, Berlin, Germany

² GESIS Leibniz Institute for the Social Sciences, Cologne, Germany

³ Universität Hamburg, Germany

⁴ Fraunhofer Institute for Applied Information Technology, Sankt Augustin, Germany

⁵ TIB Leibniz Information Centre for Science and Technology and L3S, University of Hannover, Germany

⁶ ZB MED Information Centre for Life Sciences, Cologne, Germany

⁷ Schloss Dagstuhl, Leibniz Center for Informatics, Trier, Germany

⁸ Leipzig University, Germany

⁹ Leibniz Universität Hannover and L3S, Hannover, Germany

¹⁰ Leibniz Information Centre for Economics, Kiel, Germany

¹¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Berlin, Germany

¹² Technical University Dresden, Germany

more complex, it can be difficult to maintain transparency, reproducibility, and fairness in research.

The NFDI4DS consortium, part of the NFDI initiative to build a German national research data infrastructure, aims to provide comprehensive support for all phases of the complex and interdisciplinary *research data lifecycle* within the DS and AI communities. The research data lifecycle includes collection/creation of data, its processing, analysis, preservation, access, and reuse of digital objects.

2 Approach

To achieve its goals, the consortium plans to adopt an integrated approach towards research management. This involves making all relevant digital objects available and interlinking them to provide a comprehensive view of research (improving transparency). Additionally, tools and services will be offered to facilitate the management of these resources. These efforts will ensure that valuable resources are available and accessible for reuse (improving reproducibility), ultimately contributing to the growth and development of these fields.

NFDI4DS aims to tackle the needs of various scientific fields involved in DS and AI by offering interdisciplinary as well as domain-specific services, accompanied with shared tasks and associated benchmarking datasets. Furthermore, support structures are being set up, including the circulation of learning materials and best practice guidelines, and the conduction of tutorials, workshops and challenges.

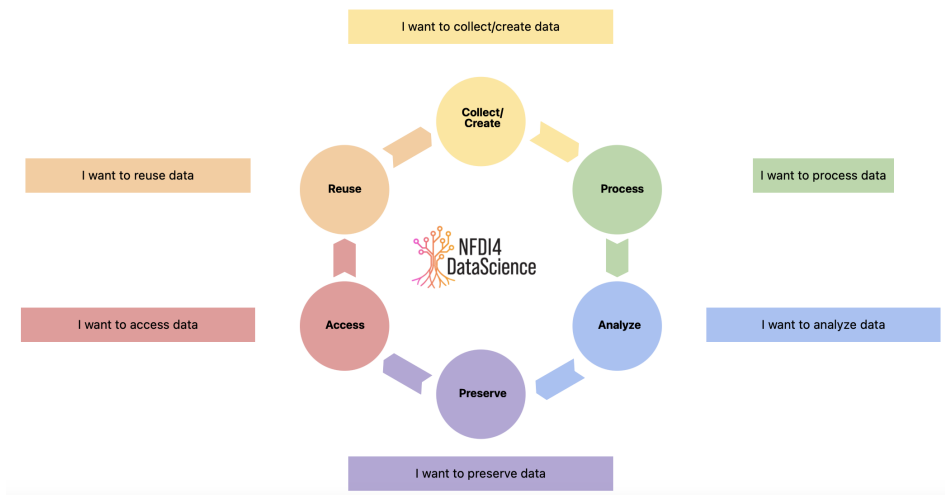


Fig. 1: Core Services Dashboard

3 Core Services

The core services focus on six main areas around digital objects: collection/creation, processing, analysis, preserving, access, reuse, and evaluation (see Figure 1). Digital objects include artifacts beyond articles, such as data, machine learning models, workflows, and scripts/code. Collection/creation, processing and analysis mainly deal with data, including data profiling [AGN15], cleaning [Ma19] and transformation[Ab16]. Preservation, access, reuse and evaluation covers a broader spectrum of digital objects as they refer not only to data. The NFDI4DS research knowledge graph forms the basis of the infrastructure, providing details about digital objects and their interrelation.

The NFDI4DS infrastructure is based on a number of already existing software components and already well-established tools and services, which target different phases of the research data lifecycle (see Figure 1). Some of these are detailed below:

- *The Open Research Knowledge Graph (ORKG)* [Au20, Ja19] is a service for semantically describing research contributions in a knowledge graph. The semantic descriptions of articles are crowd-sourced from authors and researchers leveraging NLP of articles.
- *The GESIS Knowledge Graph Infrastructure* consists of tools and pipelines for constructing actual research knowledge graphs of research information, metadata and primary research data, such as SoMeSci [Sc21], SoftwareKG [Sc22] or TWetsKB [Di20].
- *The DBLP Computer Science Bibliography* is an open bibliographic data base, search engine, and knowledge graph on computer science publications. All of DBLP's curated data is available via its website, APIs, and as a data dump download [dt23].
- *The Data Management Platform Piveau* is developed within the scope of *data.europa.eu* [Ki19, Ki20]. Piveau provides services and pipelines for harvesting data and metadata from various sources saving it into a knowledge graph utilizing semantic linking. The knowledge graph can be explored and accessed via a Web interface, APIs and a SPARQL endpoint.
- *The General Entity Annotator Benchmark (GERBIL)* [RUN18] is a FAIR benchmarking platform for entity annotation and disambiguation, as well as question-answering systems [Us19].
- *GESIS Notebooks* [B122, Ra18] is an online reproducibility service for FAIR digital objects. Its main components are a BinderHub, a JupyterHub, and a place to publish, explore, try out, and learn about DS and AI methods. GESIS Notebooks is part of the BinderHub Federation powering mybinder.org.

4 Service Integration

Our service integration strategy is inspired by the EOSC Interoperability Framework [Co21], which is based on: (1) persistent identification using PIDs, such as DOI, ORCID

or authoritative URIs, (2) authentication and authorisation (AAI) adhering to common standards, (3) semantic interoperability using RDF, vocabularies and ontologies, and (4) API integration based on REST principles. Below we describe some exemplary integrations between the NFDI4DS core services, currently under development:

- *NFDI4DS Research Knowledge Graph*. The NFDI4DS research knowledge graph will entail automatically extracted metadata about resource relations, e. g. software mentions in scholarly publications [Sc22], highly quality-controlled manual annotations of scholarly resources [Sc21] and community-annotations of scholarly publications from the ORKG [Au20, Ja19]. A particular focus is on sharing data as well as methods and infrastructure for extracting and curating scholarly knowledge.
- *NFDI4DS Registries*. The consortium aims at providing registries for different digital objects, one of which is the DBLP computer science bibliography. DBLP and the ORKG started linking author and publication entities within their respective knowledge graphs. The integration makes use of WikiData [Vr12] as a central, community-curated ID hub; matching results will be mirrored there. More sources and digital objects will be integrated as needed.
- *NFDI4DS Gateway and Portal*. The portal will include a federated search engine and recommendation system. Through a unified and intuitive search interface, users are enabled to query a wide range of scientific databases such as DBLP, Zenodo, and OpenAlex. While the gateway queries APIs in an ad-hoc fashion, the portal, which will be based on Piveau, provides a harvesting-based service to account for larger research data dumps. The overall aim is to design an entry point that categorises and summarises multiple search results (such as researchers, articles, machine learning models, benchmarking results etc.) such that practitioners and researchers gain a swift overview of existing contributions.
- *NFDI4DS Notebooks*. The consortium aims at integrating different tools and services, including a Jupyter Hub instance. To facilitate the further analysis and visualisation of digital objects contributed by the participating services, they can be directly loaded into a new MyBinder Jupyter notebook data frame for further programmatic processing, e. g. for generating visualisations.
- *Compute Infrastructure*. While most of the NFDI4DS tools and services will be hosted in a cloud infrastructure located at various sites of the partners, some DS and AI tasks require access to specialized high-performance computing (HPC) resources such as GPU accelerators and large memory capacities. The goal is to make the transition from cloud to HPC as easy as possible for both service providers and users. On a Web interface users can specify their compute resource needs, which will be handed over to the batch system SLURM to launch JupyterHub on the HPC system.

5 Conclusion

Although the idea of portals around DS and AI is not new, there is still a lot of work to do in this area of such portals, especially when it comes to the inclusion of different digital

objects and the support of multiple interdisciplinary communities. One difference between the NFDI4DS portal and other related efforts (e. g., HuggingFace¹³) is the emphasis on FAIR metadata for digital objects. Proper and rich metadata enables not only FAIR but also further empowers *Linked Open Science* to improve best practices for conducting research including aspects such as transparency, reproducibility and fairness.

Acknowledgements

This work has received funding through the German Research Foundation (DFG) project NFDI4DS (no. 460234259).

Bibliography

- [Ab16] Abedjan, Ziawasch; Morcos, John; Ilyas, Ihab F.; Ouzzani, Mourad; Papotti, Paolo; Stonebraker, Michael: DataXFormer: A robust transformation discovery system. In: 32nd IEEE International Conference on Data Engineering, ICDE 2016, Helsinki, Finland, May 16-20, 2016. IEEE Computer Society, Helsinki, Finland, pp. 1134–1145, 2016.
- [AGN15] Abedjan, Ziawasch; Golab, Lukasz; Naumann, Felix: Profiling relational data: a survey. VLDB J., 24(4):557–581, 2015.
- [Au20] Auer, Sören; Oelen, Allard; Haris, Muhammad; Stocker, Markus; D’Souza, Jennifer; Farfar, Kheir Eddine; Vogt, Lars; Prinz, Manuel; Wiens, Vitalis; Jaradeh, Mohamad Yaser: Improving Access to Scientific Literature with Knowledge Graphs. Bibliothek Forschung und Praxis, 44(3):516–529, 2020.
- [BI22] Bleier, Arnim; Erdogan, Kenan; Kahmann, Christian; Posch, Lisa: , GESIS Notebooks: Online Reproducible Computational Analysis for the Social Sciences, 2022.
- [Co21] Commission, European; for Research, Directorate-General; Innovation; Corcho, O; Eriksson, M; Kurowski, K; Ojsteršek, M; Choirat, C; Sanden, M; Coppens, F: EOSC interoperability framework : report from the EOSC Executive Board Working Groups FAIR and Architecture. Publications Office, Brussels, Belgium, 2021.
- [Di20] Dimitrov, Dimitar; Baran, Erdal; Fafalios, Pavlos; Yu, Ran; Zhu, Xiaofei; Zloch, Matthäus; Dietze, Stefan: TweetsCOVID-19 - A Knowledge Base of Semantically Annotated Tweets about the COVID-19 Pandemic. In: Proceedings of the 29th ACM International Conference on Information and Knowledge Management. CIKM ’20, Association for Computing Machinery, New York, NY, USA, p. 2991–2998, 2020.
- [dt23] dblp team: , dblp computer science bibliography. Download, January 2023. Monthly RDF release of January 2023.
- [Ja19] Jaradeh, Mohamad Yaser; Oelen, Allard; Farfar, Kheir Eddine; Prinz, Manuel; D’Souza, Jennifer; Kismihók, Gábor; Stocker, Markus; Auer, Sören: Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge. In: Proceedings of the 10th International Conference on Knowledge Capture. K-CAP ’19, Association for Computing Machinery, New York, NY, USA, p. 243–246, 2019.

¹³ <https://huggingface.co/>

- [Ki19] KIRSTEIN, Fabian; DITTWALD, Benjamin; DUTKOWSKI, Simon; GLIKMAN, Yury; SCHIMMLER, Sonja; HAUSWIRTH, Manfred: Linked Data in the European Data Portal: A Comprehensive Platform for Applying DCAT-AP. In (Lindgren, Ida; Janssen, Marijn; Lee, Habin; Polini, Andrea; Rodríguez Bolívar, Manuel Pedro; Scholl, Hans Jochen; Tambouris, Efthimios, eds): *Electronic Government*. Springer International Publishing, Cham, pp. 192–204, 2019.
- [Ki20] KIRSTEIN, Fabian; STEFANIDIS, Kyriakos; DITTWALD, Benjamin; DUTKOWSKI, Simon; URBANEK, Sebastian; HAUSWIRTH, Manfred: Piveau: A Large-Scale Open Data Management Platform Based on Semantic Web Technologies. In (Harth, Andreas; Kirrane, Sabrina; Ngonga Ngomo, Axel-Cyrille; Paulheim, Heiko; Rula, Anisa; Gentile, Anna Lisa; Haase, Peter; Cochez, Michael, eds): *The Semantic Web*. Springer International Publishing, Cham, pp. 648–664, 2020.
- [Ma19] MAHDAMI, Mohammad; ABEDJAN, Ziawasch; FERNANDEZ, Raul Castro; MADDEN, Samuel; OUZZANI, Mourad; STONEBRAKER, Michael; TANG, Nan: Raha: A Configuration-Free Error Detection System. In (Boncz, Peter A.; Manegold, Stefan; Ailamaki, Anastasia; Deshpande, Amol; Kraska, Tim, eds): *Proceedings of the 2019 International Conference on Management of Data, SIGMOD Conference 2019, Amsterdam, The Netherlands, June 30 - July 5, 2019*. ACM, Amsterdam, The Netherlands, pp. 865–882, 2019.
- [Ra18] RAGAN-KELLEY, Benjamin; WILLING, Carol; AKICI, F; LIPPA, D; NIEDERHUT, D; PACER, M; BINDER 2.0-Replicable, interactive, sharable environments for science at scale. In: *Proceedings of the 17th python in science conference*. F. Akici, D. Lippa, D. Niederhut, and M. Pacer, eds., scipy.org, Austin, TX, USA, pp. 113–120, 2018.
- [RUN18] RÖDER, Michael; USBECK, Ricardo; NGOMO, Axel-Cyrille Ngonga: GERBIL - Benchmarking Named Entity Recognition and Linking consistently. *Semantic Web*, 9(5):605–625, 2018.
- [Sc21] SCHINDLER, David; BENSCHMANN, Felix; DIETZE, Stefan; KRÜGER, Frank: SoMeSci- A 5 Star Open Data Gold Standard Knowledge Graph of Software Mentions in Scientific Articles. In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management. CIKM '21, Association for Computing Machinery, New York, NY, USA, p. 4574–4583, 2021*.
- [Sc22] SCHINDLER, David; BENSCHMANN, Felix; DIETZE, Stefan; KRÜGER, Frank: The role of software in science: a knowledge graph-based analysis of software mentions in PubMed Central. *PeerJ Computer Science*, 8:e835, January 2022.
- [Us19] USBECK, Ricardo; RÖDER, Michael; HOFFMANN, Michael; CONRADS, Felix; HUTHMANN, Jonathan; NGOMO, Axel-Cyrille Ngonga; DEMMLER, Christian; UNGER, Christina: Benchmarking question answering systems. *Semantic Web*, 10(2):293–304, 2019.
- [Vr12] VRANDIĆ, Denny: Wikidata: a new platform for collaborative data collection. In (Mille, Alain; Gandon, Fabien; Misselis, Jacques; Rabinovich, Michael; Staab, Steffen, eds): *Proceedings of the 21st World Wide Web Conference, WWW 2012, Lyon, France, April 16–20, 2012 (Companion Volume)*. ACM, Lyon, France, pp. 1063–1064, 2012.
- [Wi16] WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, IJbrand Jan; et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018, December 2016. Number: 1.

NFDI4DS Transfer and Application

Ekaterina Borisova,¹ Raia Abu Ahmad,¹ Georg Rehm,¹ Ricardo Usbeck,² Jennifer D'Souza,³ Markus Stocker,³ Sören Auer,³ Judith Gilsbach,⁴ Anastasia Wolschewski,⁵ Johannes Keller,⁵ Daniel Schneider,⁵ Thomas Neumuth,⁵ Sonja Schimmler⁶

Abstract: Due to the ever increasing importance of Data Science and Artificial Intelligence methods for a wide range of scientific disciplines, ensuring *transparency* and *reproducibility* of DS and AI methods and research findings have become essential. The NFDI4DS project promotes the *findability*, *accessibility*, *interoperability*, and *reusability* in DS and AI by developing an open integrated research data infrastructure in which all artefacts (e. g., papers, code, models, datasets) will be interlinked in a FAIR and transparent way. One of the key aspects is to build a bridge between NFDI4DS and other research communities which actively apply DS and AI methods. This paper describes the main actions taken to engage with the relevant (sub)communities.

Keywords: NFDI; NFDI4DS; Research Data Infrastructure; Data Science; Artificial Intelligence

1 Introduction

Most research artefacts nowadays are spread across repositories, digital libraries, and institutional databases. Such unsystematic and decentralised storage complicates the *findability*, *accessibility*, *interoperability*, and *reusability* (FAIR) [WDA16] of scientific data. The German National Research Data Infrastructure (NFDI)⁷ initiative aims to interconnect and preserve interdisciplinary scholarly data in a FAIR way. NFDI for Data Science and Artificial Intelligence (NFDI4DS)⁸ is one of the 26 NFDI consortia which promotes the idea of FAIRness, transparency and reproducibility in DS and AI. NFDI4DS's goal is to support all phases of the research data life cycle, including collecting, creating, comparing, processing, analysing, publishing, archiving, and reusing resources in DS and AI through innovative tools and services.

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Berlin, Germany
ekaterina.borisova@dfki.de, raia.abu_ahmad@dfki.de, georg.rehm@dfki.de

² Universität Hamburg, Hamburg, Germany, ricardo.usbeck@uni-hamburg.de

³ TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany
jennifer.dsouza@tib.eu, markus.stocker@tib.eu, auer@tib.eu

⁴ GESIS Leibniz Institute for the Social Sciences, Cologne, Germany, judith.gilsbach@gesis.org

⁵ Innovation Center Computer Assisted Surgery (ICCAS), Leipzig University, Germany
anastasia.wolschewski@iccas.de, johannes.keller@iccas.de, daniel.schneider@iccas.de,
thomas.neumuth@iccas.de

⁶ Fraunhofer FOKUS, sonja.schimmler@fokus.fraunhofer.de

⁷ <https://www.nfdi.de>

⁸ <https://www.nfdi4datascience.de>

To achieve this objective, it is important to engage with the relevant communities and to transform NFDI4DS's visions into concrete applications. We collaborate with four scientific fields, i. e., Natural Language Processing (NLP) and Language Technology (LT) including Semantic Web, Biomedical Research, Information Sciences and Social Sciences.

2 Applications in Natural Language Processing

Like the other domains in scope of the project, researchers from the LT, NLP, and Semantic Web communities apply a wide variety of Research Data Management (RDM) methodologies. Although these communities follow similar research avenues, they utilise heterogeneous RDM techniques. This leads to inconsistency in documentation and sharing of digital artifacts (e. g., data, code, models) contributing to low transparency and reusability. Our goal is to bridge the gap between LT, NLP and Semantic Web by hosting different events (e.g., workshops, shared tasks, challenges) and by spreading knowledge about RDM best practices. We currently work on the following activities:

- Three shared tasks⁹ which address problems in the area of scholarly information processing, i. e., software mention detection, leaderboard mining [KDA21, KDA23a, KDA23b], and research field classification. By organising these shared tasks we enable and encourage the sharing of RDM techniques, digital artifacts and evaluation measures across LT, NLP and Semantic Web communities.
- The ISWC 2023 Scholarly Question Answering challenge¹⁰ allows researchers to compare their Knowledge Graph Question Answering Systems.
- We also work on compiling Open Science Best Practices in DS and AI, i. e., recommendations for ensuring a FAIR life cycle of digital artifacts. The guidelines will be published on the NFDI4DS website.

3 Applications in Biomedical Research and Clinical Decision Making

In the biomedical domain, Data Science faces obstacles due to 1. the safety-critical and ethically relevant nature of clinical applications, 2. complex heterogeneous and incomplete datasets, and 3. a lack of standardisation and data-privacy regulations. We plan to leverage the NFDI4DS infrastructure to foster data and knowledge transfer between the biomedical research community and DS. This ensures that the requirements of biomedical stakeholders can be adequately met within DS research. Our effort until 2026 includes:

- Host challenges dealing with the requirements of biomedical applications, starting with a competition related to reliable uncertainty estimation in 2024. The competitions will raise awareness for open research questions relevant to biomedical DS, flatten

⁹ <https://www.nfdi4datascience.de/docs/community-engagement/shared-tasks/>

¹⁰ <https://kgqa.github.io/scholarly-QALD-challenge/>

- learning curves through tutorials and workshops, present overviews of the state of the art, provide benchmarks, and obtain novel solutions driving AI adoption in medicine.
- Promote the NFDI4DS infrastructure in the biomedical engineering, medical informatics, and DS research communities by giving talks and hosting focus sessions at relevant conferences (e. g., BMT¹¹, GMDS¹², and within ScaDS.AI¹³). These events will also help gather community feedback and requirements.
 - Design tutorials and best practices for use cases from clinical prognosis and decision support using heterogeneous medical datasets. The tutorials will be deployed using the binder notebook service and involve topics including data preparation, synthetic data generation, dealing with missing data, and uncertainty estimation.

4 Application for Information Sciences

The Open Research Knowledge Graph (ORKG)¹⁴ [St23, Au20, Ja19] is an NFDI4DS service for semantically describing research contributions published in scientific articles in a knowledge graph. Research contributions are semantic (i. e., machine-actionable) descriptions of published research findings together with the employed materials and methods and the addressed research problem. Semantic descriptions are crowd-sourced from authors and researchers. The ORKG also leverages NLP information extraction from articles to automate knowledge graph construction. ORKG supports numerous downstream services, including comparisons of research contributions addressing a common research problem, visualisations of compared information, leaderboards, etc. The ORKG is central for the NFDI4DS Application for Information Sciences, which has three goals:

- Deploy the most promising NLP models for information extraction from scholarly articles developed in NFDI4DS Shared Tasks in an infrastructure to further advance automation of scholarly knowledge graph construction and curation.
- Catalyse the adoption of the ORKG and, more generally, Scholarly Knowledge Graph infrastructure, in relevant conference series and journals. First experiments have been conducted in 2022 with ISWC and SEMANTiCS by actively guiding authors on how to include ORKG Comparisons in their submitted papers.
- Develop approaches to ensure expressions of research findings are produced machine actionable, and automatically flow into digital scholarly communication infrastructure such as ORKG. We currently test approaches for model performance evaluation conducted using Python or R computing environments to ensure that TDMS (Task, Dataset, Metric, Score) data published in articles automatically flows into ORKG benchmarks and leaderboards.

¹¹ <https://bmt2023.de>

¹² <https://www.gmds2023.de>

¹³ <https://scads.ai>

¹⁴ <https://orkg.org>

5 Applications in Social Sciences

Researchers from the Social Sciences handle very heterogeneous types of data, documented and shared in very diverse ways. The same applies to code written for data wrangling and analysis. Our goal is to encourage Social Science researchers to practice FAIR science and provide them with the knowledge and infrastructure. We currently compile the status quo regarding documentation, quality assessment and sharing behaviour of data and code. Subsequently, we will contribute to the implementation of our goals in the Social Sciences with an emphasis on making data and code available to the research community in a reusable way. Until 2026 we will be engaged in the following activities:

- Conduct a survey among researchers, investigating current data and code-sharing practices as well as quality assessment of digital research objects with a focus on relational social network data. For sharing behaviour a similar survey exists in the field of computational biology [CH22].
- Define use cases from an online access web-tracking panel under development at GESIS¹⁵ to define requirements for data quality and sharing practices.
- Host a workshop for social scientists on open science practices, present NFDI4DS services and gather feedback from the Social Science community.

6 Call for Speedboat Projects

In addition to the four scientific areas mentioned in the last paragraphs, we are issuing a call for speed boat projects. These projects are meant to kick off collaboration with further domains, and to complement our overall service portfolio.

7 Conclusion

This paper presents NFDI4DS's main strategies to promote FAIR research data management across four scientific areas: 1. NLP and LT as well as Semantic Web, 2. Biomedical Research and Clinical Decision Making, 3. Information Sciences, and 4. Social Sciences. The current community engagement tools range from conducting various surveys and proposing best practices to organising shared tasks.

Acknowledgements

This work has received funding through the German Research Foundation (DFG) project NFDI4DS (no. 460234259).

¹⁵ <https://www.gesis.org>

Bibliography

- [Au20] Auer, Sören; Oelen, Allard; Haris, Muhammad; Stocker, Markus; D'Souza, Jennifer; Farfar, Kheir Eddine; Vogt, Lars; Prinz, Manuel; Wiens, Vitalis; Jaradeh, Mohamad Yaser: Improving Access to Scientific Literature with Knowledge Graphs. *Bibliothek Forschung und Praxis*, 44(3):516–529, 2020.
- [CH22] Cadwallader, Lauren; Hrynaszkiewicz, Iain: A Survey of Researchers' Code Sharing and Code Reuse Practices, and Assessment of Interactive Notebook Prototypes. *PeerJ*, 10:e13933, 2022.
- [Ja19] Jaradeh, Mohamad Yaser; Oelen, Allard; Farfar, Kheir Eddine; Prinz, Manuel; D'Souza, Jennifer; Kismihók, Gábor; Stocker, Markus; Auer, Sören: Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge. In: *Proceedings of the 10th International Conference on Knowledge Capture. K-CAP '19*, Association for Computing Machinery, New York, NY, USA, pp. 243–246, 2019.
- [KDA21] Kabongo, Salomon; D'Souza, Jennifer; Auer, Sören: Automated Mining of Leaderboards for Empirical AI Research. In: *Towards Open and Trustworthy Digital Societies: 23rd International Conference on Asia-Pacific Digital Libraries, ICADL 2021, Virtual Event, December 1–3, 2021, Proceedings 23*. Springer, pp. 453–470, 2021.
- [KDA23a] Kabongo, Salomon; D'Souza, Jennifer; Auer, Sören: ORKG-Leaderboards: A Systematic Workflow for Mining Leaderboards as a Knowledge Graph. *arXiv preprint arXiv:2305.11068*, 2023.
- [KDA23b] Kabongo, Salomon; D'Souza, Jennifer; Auer, Sören: Zero-shot Entailment of Leaderboards for Empirical AI Research. In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2023*. 2023.
- [St23] Stocker, Markus; Oelen, Allard; Jaradeh, Mohamad Yaser; Haris, Muhammad; Oghli, Omar Arab; Heidari, Golsa; Hussein, Hassan; Lorenz, Anna-Lena; Kabenamualu, Salomon; Farfar, Kheir Eddine; Prinz, Manuel; Karras, Oliver; D'Souza, Jennifer; Vogt, Lars; Auer, Sören: FAIR Scientific Information with the Open Research Knowledge Graph. *FAIR Connect*, 1(1):19–21, Jan 2023.
- [WDA16] Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, et al.: The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3(160018), 3 2016.

NFDI4DS Shared Tasks

Raia Abu Ahmad,¹ Ekaterina Borisova,¹ Georg Rehm,¹ Stefan Dietze,² Saurav Karmakar,² Wolfgang Otto,² Jennifer D’Souza,³ Fidan Limani,⁴ Ricardo Usbeck⁵

Abstract: Shared tasks have proven to be successful in proposing innovative solutions for challenging research problems. The NFDI4DS consortium plans to host various shared tasks to tackle problems under the umbrella of scholarly information processing. We discuss three shared tasks in detail: Field of Research Classification, Software Mention Detection, and Tracking State-of-the-Art in Empirical AI. We also briefly mention other shared tasks planned to be released in the future.

Keywords: NFDI; NFDI4DS; Shared Tasks

1 Introduction

Shared tasks are scientific competitions in which teams attempt to find efficient solutions to a specific problem using shared data and evaluation measures [Pa17]. The goal is to objectively and directly compare different methods for tackling the same problem by using gold-standard data and common performance measures. Shared tasks are usually organised either at conferences or workshops such as the Conference on Natural Language Learning⁶ or the International Workshop on Semantic Evaluation⁷, or by companies (e. g., Kaggle⁸).

In recent years, shared tasks have been very successful in advancing state-of-the-art methods and standards to solve challenging problems [FU18]. In most shared tasks, the gold-standard dataset is made publicly available after the competition, thus providing valuable resources for the research community [Pa17]. Recent work has also introduced guidelines to ensure transparency and reproducibility of shared tasks in order to benefit scientific progress [Es21]. By utilising competitiveness among participants, shared tasks have proven to significantly encourage the development of novel and innovative solutions [Pa17].

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Berlin, Germany
ekaterina.borisova@dfki.de, raia.abu_ahmad@dfki.de, georg.rehm@dfki.de

² GESIS Leibniz Institut für Sozialwissenschaften, Germany
stefan.dietze@gesis.org, saurav.karmakar@gesis.org, wolfgang.otto@gesis.org

³ TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany, jennifer.dsouza@tib.eu

⁴ ZBW – Leibniz-Informationszentrum Wirtschaft, Kiel, Germany, f.limani@zbw-online.eu

⁵ Universität Hamburg, Germany, ricardo.usbeck@uni-hamburg.de

⁶ <https://conll.org>

⁷ <https://semeval.github.io>

⁸ <https://www.kaggle.com>

This paper presents three shared tasks that are part of the NFDI for Data Science and AI (NFDI4DS) project.⁹ We also briefly introduce tasks currently under development. All shared tasks mentioned in Section 2, 3 and 4 are planned to start in 2024.

2 FoRC: Field of Research Classification for Scholarly Publications

The fast-growing pace of research has given rise to repositories and digital libraries that capture and manage scientific knowledge. For automated processes, the information of which scientific field a publication belongs to can assist downstream tasks. However, existing classification systems are limited either in terms of the used taxonomy, or in terms of the classification model itself.

We propose the Field of Research Classification (FoRC) shared task, which consists of two sub-tasks. **Subtask I** is a multi-class problem for general fields of research (FoR). This subtask uses an existing FoR taxonomy from the Open Research Knowledge Graph (ORKG) [Au20]. The dataset was constructed by fetching scientific publications from ORKG, arXiv, Crossref¹⁰ API, and S2AG¹¹. The subtask I dataset consists of 59,500 scholarly papers in English, classified into 123 FoR over five high-level fields and four hierarchical levels. **Subtask II** is a fine-grained multi-label classification problem that deals with multiple sub-fields within one specific FoR, i. e., computational linguistics (CL). This subtask will use a hierarchical taxonomy of ca. 170 CL sub-fields structured into three hierarchical levels that was constructed using a topic modeling approach. The dataset is currently being developed by a team of six annotators and will contain 1,500 manually labeled articles from the ACL Anthology.

3 SOMD: Software Mention Detection in Scholarly Publications

Across all disciplines science has become increasingly data-driven, leading to additional needs regarding software for collecting, processing and analysing data. Transparency about software used as part of the scientific process is crucial to ensure reproducibility and to understand provenance of research data and insights. Furthermore, understanding software usage, citation habits, and their evolution over time within and across disciplines can shape the understanding of the evolution of disciplines, the influence of software on scientific impact, and the emerging needs for computational support.

Given the scale and heterogeneity of software citations, robust methods are required, able to detect and disambiguate mentions of software and related metadata. However, despite the existence of software citation principles [SKN16], software mentions in scientific articles

⁹ <https://www.nfdi4datascience.de>

¹⁰ <https://www.crossref.org>

¹¹ <https://www.semanticscholar.org/product/api>

are usually informal and incomplete. Initial pipelines for the automated extraction through deep learning-based NLP pipelines [Is22] usually rely on fine-tuned or otherwise supervised models and sufficient ground truth data [Sc21].

The SOMD shared task will build on a corpus and annotation framework and existing baselines (SoMeSci [Sc21]). In particular, we consider three different subtasks. **Subtask I** will deal with recognising Software from individual sentences. At the same time software mentions shall be classified according to its mention type, e.g., mention, usage, or creation and the software type, e.g., application, programming environment, and plugin or package. **Subtask II** will deal with extracting additional information for each software according to the SoMeSci schema. And **Subtask III** will deal with classifying relations to other recognised entities. This includes versions and developers, but also URLs or Host applications for plugins. Evaluation will be based on exact matches rather than partial matches.

4 SOTA? Tracking the State-of-the-Art in AI Scholarly Publications

The growing number of publications poses a major challenge [Ji10]. How can we stay updated with the fast-paced research progress? The current format of scholarly communication, with results buried in unstructured PDFs, hinders comprehension and limits machine usability. One potential solution is to represent research results in structured and semantic formats within knowledge graphs (KG) of scientific knowledge [Sh09] such as the ORKG¹². When it comes to AI research, the leaderboards construct of information organisation provides an overview of the state-of-the-art (SOTA) by aggregating results from multiple studies addressing the same research challenge. A leaderboard typically comprises a task (T), a dataset (D), evaluation metrics (M), and scores obtained by the model (S).

The SOTA shared task would facilitate reaping diverse machine learning observations on a relatively non-trivial task as the automated mining of leaderboards for empirical AI research. **Subtask I** will deal with TDM extraction, **Subtask II** with TDMS extraction, and **Subtask III** with extracting URLs of source-code from publications. The dataset for each subtask is extracted from the community-annotated Leaderboards on PapersWithCode (PwC).

5 Future Work

Four additional shared tasks are currently under development:

1. Question Answering (QA): There is currently no natural language interface that allows users to access scholarly data, e. g., papers, authors, models, or datasets. We will introduce a shared task at ISWC 2023¹³ with two subtasks. SciQA [Au23] focuses

¹² <https://orkg.org>

¹³ <https://kgqa.github.io/scholarly-QALD-challenge/2023/>

on QA over scholarly data using the ORKG. DBLP-QUAD involves KG question answering over the DBLP Knowledge Graph using DBLP-QUAD [Ba23].

2. **Machine Learning Model Detection:** Machine learning model (MLModel) mentions in scholarly articles are currently unable to be detected. We are preparing a manually annotated gold standard dataset as part of the GESIS Scholarly Annotation Project (GSAP), comprised of annotating MLModel family, dataset model family, and their relationships. The GSAP data will be used in two subtasks: HowMLMod will deal with detecting how an MLModel is mentioned in an article. AncestorMLMod will figure out which MLModel architecture is used in specific MLModel mentions.
3. **Dataset Mention Detection:** In Machine Learning research, ensuring the traceability of datasets is vital due to issues like biases, data leakage, train-test contamination, and ethical concerns. Reuse and overlap of datasets further complicate matters. We will propose two shared tasks. The first focuses on detecting dataset mentions that are aligned with mentions of ML models. The second task emphasizes traceability by detecting relationships between datasets. It aims to uncover how datasets are reused or transformed, unraveling the dataset's history.
4. **Standards, FAIR principles, and FAIR Digital Objects (FDO):** The project targets different types of research artefacts, including the ones generated from the shared tasks, meant to benefit the NFDI4DS research community. Our goal is to evaluate the adoption of the FDO model as means of elevating these artefacts into “actionable knowledge units” [DSKW20], able to support existing or enable new use cases. FAIR principles [Wi16] are a key component for this model, and this “dependency” implies that evaluating their adoption for the NFDI4DS resources remains an important aspect for this shared task. Finally, there are tasks across the project, such as the adoption of metadata standards, Knowledge Organization Systems, machine-readable representation, to name just a few, that provide a valuable input to and will be considered in their supporting role to addressing the challenges of FDO adoption.

6 Conclusion

This paper presents three of NFDI4DS's shared tasks, as well as four additional ones currently under development. The shared tasks tackle challenging problems in the field of scholarly information processing. All tasks plan to release publicly available datasets and follow guidelines of transparency and reproducibility.








Acknowledgements

This work has received funding through the German Research Foundation (DFG) project NFDI4DS (no. 460234259).

Bibliography


- [Au20] Auer, Sören; Oelen, Allard; Haris, Muhammad; Stocker, Markus; D'Souza, Jennifer; Farfar, Kheir Eddine; Vogt, Lars; Prinz, Manuel; Wiens, Vitalis; Jaradeh, Mohamad Yaser: Improving Access to Scientific Literature with Knowledge Graphs. *Bibliothek Forschung und Praxis*, 44(3):516–529, 2020.
- [Au23] Auer, Sören; Barone, Dante A.C.; Bartz, Cassiano; Cortes, Eduardo G.; Yaser, Mohamad Jaradeh; Karras, Oliver; Koubarakis, Manolis; Mourontsev, Dmitry; Pliukhin, Dmitrii; Radyush, Daniil; Shilin, Ivan; Stocker, Markus; Tsalapati, Eleni: SciQA Benchmark: Dataset and RDF Dump. 2023.
- [Ba23] Banerjee, Debayan; Awale, Sushil; Usbeck, Ricardo; Biemann, Chris: DBLP-QuAD: A Question Answering Dataset over the DBLP Scholarly Knowledge Graph. *CoRR*, abs/2303.13351, 2023.
- [DSKW20] De Smedt, Koenraad; Koureas, Dimitris; Wittenburg, Peter: FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units. *Publications*, 8(2):21, 2020.
- [Es21] Escartín, Carla Parra; Lynn, Teresa; Moorkens, Joss; Dunne, Jane: Towards Transparency in NLP Shared Tasks. *arXiv preprint arXiv:2105.05020*, 2021.
- [FU18] Filannino, Michele; Uzuner, Özlem: Advancing the State of the Art in Clinical Natural Language Processing through Shared Tasks. *Yearbook of medical informatics*, 27(01):184–192, 2018.
- [Is22] Istrate, Ana-Maria; Li, Donghui; Taraborelli, Dario; Torcar, Michaela; Veytsman, Boris; Williams, Ivana: A Large Dataset of Software Mentions in the Biomedical Literature. 2022.
- [Ji10] Jinha, Arif E.: Article 50 Million: An Estimate of the Number of Scholarly Articles in Existence. *Learned Publishing*, 23(3):258–263, 2010.
- [Pa17] Parra Escartín, Carla; Reijers, Wessel; Lynn, Teresa; Moorkens, Joss; Way, Andy; Liu, Chao-Hong: Ethical Considerations in NLP Shared Tasks. In: *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*. Association for Computational Linguistics, Valencia, Spain, pp. 66–73, April 2017.
- [Sc21] Schindler, David; Bensmann, Felix; Dietze, Stefan; Krüger, Frank: SoMeSci—A 5 Star Open Data Gold Standard Knowledge Graph of Software Mentions in Scientific Articles. In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM '21)*. Association for Computing Machinery, Virtual Event, QLD, Australia, November 2021.
- [Sh09] Shotton, David: Semantic Publishing: The Coming Revolution in Scientific Journal Publishing. *Learned Publishing*, 22(2):85–94, 2009.
- [SKN16] Smith, Arfon M; Katz, Daniel S; Niemeyer, Kyle E: Software Citation Principles. *PeerJ Computer Science*, 2:e86, 2016.
- [Wi16] Wilkinson, Mark D; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E et al.: The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific data*, 3(1):1–9, 2016.

NFDI4Energy Task Area 4: FAIR Data for Energy System Research


Amanda Wein ¹, Jan Reinkensmeier², Anke Weidlich ³, Johan Lilliestam ⁴,
Veit Hagenmeyer ⁵, Mascha Richter⁶, Sören Auer ⁷, Astrid Nieße ⁸ and
Sebastian Lehnhoff ⁹


Abstract: The NFDI4Energy consortium will create a research data infrastructure for energy system research, emphasizing the openness and FAIRness of data and models in this research domain. Within the consortium, Task Area 4 focuses on the development of resources and services that will provide a semantic layer for the overall platform built by NFDI4Energy. The team of this Task Area will produce artifacts including a domain ontology, metadata standards, a knowledge graph, a Persistent Identifier service, and integration infrastructure to join these artifacts to the NFDI4Energy platform..


Keywords: FAIR Data, Metadata, Ontologies, Energy System Research, Research Data Infrastructure

¹ OFFIS e.V., R&D Division Energy, Escherweg 2, Oldenburg, 26121, amanda.wein@offis.de, 
<https://orcid.org/0009-0009-2960-3474>


² OFFIS e.V., R&D Division Energy, Escherweg 2, Oldenburg, 26121, jan.reinkensmeier@offis.de


³ Albert-Ludwigs-Universität Freiburg., Institut für Nachhaltige Technische Systeme (INATECH), Friedrichstr. 39, Freiburg, 79110, anke.weidlich@inatech.uni-freiburg.de, 
<https://orcid.org/0000-0003-2361-0912>


⁴ Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Sustainability Transition Policy, Lange Gasse 20, Nürnberg, 90403, Johan.Lilliestam@iass-potsdam.de, 
<https://orcid.org/0000-0001-6913-5956>

⁵ Karlsruhe Institute of Technology, Institute for Automation and Applied Informatics, Kaiserstr. 89, Karlsruhe, 76131, veit.hagenmeyer@kit.edu, 
<https://orcid.org/0000-0002-3572-9083>

⁶ Reiner Lemoine Institut gGmbH, Research Unit Transformation of Energy Systems, Rudower Chaussee 12, Berlin, 12489, Mascha.Richter@rl-institut.de

⁷ Technische Informationsbibliothek, Data Science & Digital Libraries research group, Welfengarten 1 B, Hannover, 30167, auer@tib.eu, 
<https://orcid.org/0000-0002-0698-2864>

⁸ Carl von Ossietzky Universität Oldenburg, Department für Informatik, Ammerländer Heerstr. 114-118, Oldenburg, 26129, astrid.niesse@uni-oldenburg.de, 
<https://orcid.org/0000-0003-1881-9172>

⁹ OFFIS e.V., R&D Division Energy, Escherweg 2, Oldenburg, 26121, sebastian.lehnhoff@offis.de, 
<https://orcid.org/0000-0003-2340-6807>

1 Introduction

The National Research Data Infrastructure (Nationale Forschungsdateninfrastruktur, or NFDI) initiative in Germany aims to establish research data infrastructures for many scientific disciplines. Various consortia have been established to pursue this goal, each focusing on a specific research area. Among these consortia is the NFDI4Energy consortium, started in 2023, which is building research data infrastructure for energy system research. The envisioned final product will support researchers in obtaining, sharing, and understanding FAIR research data and models. [Ni22]

NFDI4Energy has divided its work into seven Task Areas (TAs) that together will create services for researchers focusing on the energy system domain (see Figure 1). TAs 1, 2, and 3 focus on the involvement of the energy research community, society, and the energy industry, respectively, to ensure that the services provided by the consortium are able to meet the needs of these diverse stakeholders. TAs 4, 5, and 6 will tackle the creation of services and infrastructure, while TA7 provides organizational support across the consortium. [Ni22]

This paper describes the objectives of TA4: FAIR Data for Energy System Research, as well as the methodology that the TA4 team will use to achieve these objectives. Section 2 provides background and motivation for the consortium's work, with an emphasis on TA4's role. Section 3 describes the artifacts that TA4 has been tasked with creating, as well as the proposed methods to create them. Finally, Section 4 concludes the paper with an overview of the TA4 stakeholders and participants.

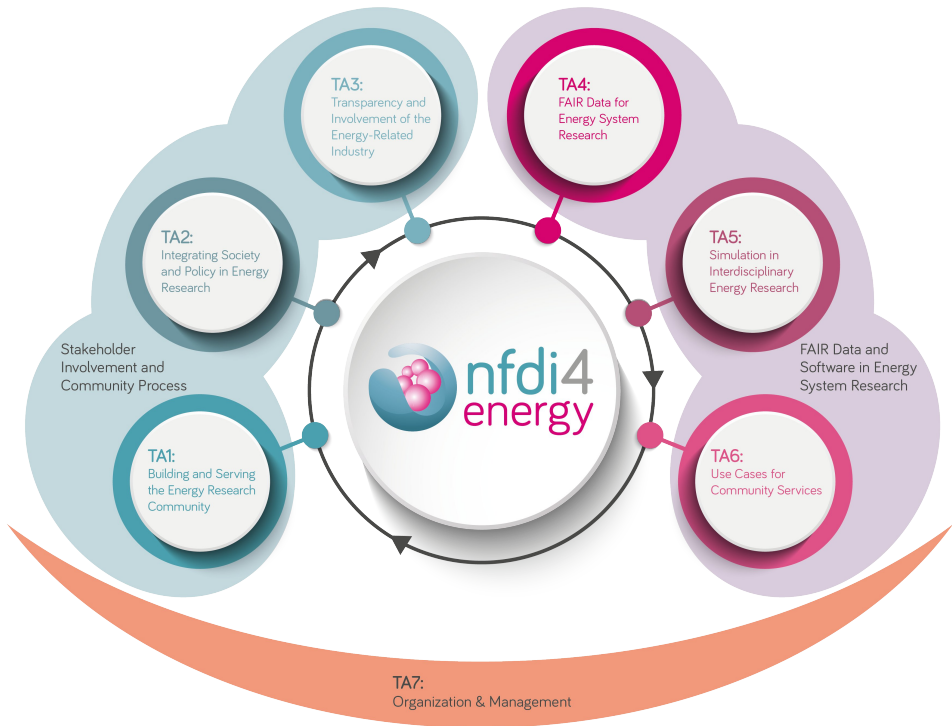


Fig. 1: The seven Task Areas (TAs) of NFDI4Energy. Figure from [Ni22].

2 Background

The FAIR Principles, first published in 2016 [Wi16], are a set of data management guidelines which the NFDI4Energy consortium aims to follow in its service development. To conform to these principles, research data and/or metadata must be:

- Findable (clearly indexed such that it can be searched for and located),
- Accessible (retrievable by standardized methods and usable in multiple contexts),
- Interoperable (using standardized terminology so that the data may be used as broadly as possible), and
- Reusable (described in detail, including provenance and usage license information, according to relevant standards).

In the domain of energy system research, FAIR data can be difficult to obtain and use. Energy systems impact and are impacted by many diverse factors, ranging from new technology adoption to governmental climate change policies, to global weather patterns, to financial structures of energy markets, and more. Studies in this domain therefore often require information from many disciplines, necessitating substantial time and effort on the part of researchers to not only track down relevant datasets, but to clean and map the data appropriately. [Ni22]

The NFDI4Energy consortium aims to ease this burden by building a data infrastructure that promotes and supports the use of FAIR data in energy system research. The consortium's platform will include services for researchers to discover best practices and methods to be used in their projects, obtain standardized datasets and models, find and collaborate with like-minded research teams, and share their work with the scientific community. [Ni22]

Within the consortium, TA4 will develop the semantic layer of this platform. Ontologies will define terminology and relationships for relevant concepts in energy system research, and metadata standards will promote the findability and reusability of research data. Various services, including a Knowledge Graph, a Terminology Service, and a Persistent Identifier (PID) service, will enable the storage and retrieval of research artifacts. These resources, when integrated into the NFDI4Energy platform, will be key components of the new data infrastructure, streamlining the work of researchers in the energy domain. [Ni22]

3 Objectives & methodology

The deliverables that TA4 will produce are summarized in Figure 2. As shown, the artifacts can be divided into two categories: semantic resources, and services & infrastructure. Each of the software services will be designed according to the terminologies and relationships defined by the semantic resources. Input from several other TAs will be incorporated into TA4's work, and TA4's work will in turn be incorporated into the tasks of other TAs, ensuring a cohesive NFDI4Energy platform.

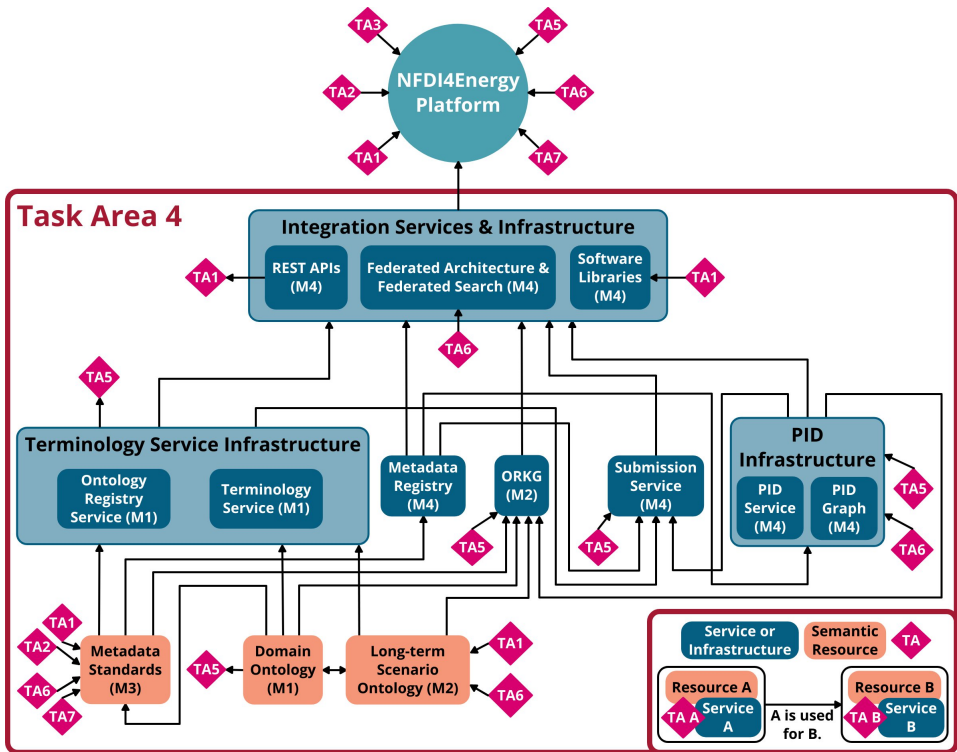


Fig. 2: The resources and services to be created by TA4. Arrows indicate how the artifacts are linked together, and also how the work of other TAs will influence the work of TA4. The tasks of TA4 are referred to internally as Measures, abbreviated here as M1 – M4.

3.1 Semantic resources

Two related ontologies will be fashioned by TA4.

A domain ontology will provide the semantic backbone for NFDI4Energy’s research data infrastructure, defining a standardized vocabulary for energy system research. This ontology will lay out concept definitions and relationships in a hierarchical arrangement. It will form a basis for research data interoperability not just within the energy domain, but also in interdisciplinary applications. [Ni22]

Building from the domain ontology, a scenario ontology will provide a vocabulary for specification of energy system scenarios. Various long-term scenarios are often used in studies examining what the power grids of the future may look like and how entities can best meet their carbon reduction goals; however, there is generally wide variation in scenario criteria (such as time frames, energy demand predictions, energy price changes,

greenhouse gas emission forecasts, etc.) from study to study. The scenario ontology aims to normalize the definitions of these criteria and simplify comparisons between scenarios, making the resulting research data more accessible to policy-makers. [Ni22]

Multiple ontologies already exist for the energy domain, such as the Open Energy Ontology [Bo21] and the Common Information Model [Mc07]. Rather than reinvent the wheel, one or more of these existing ontologies will be used as a starting point for TA4's work. The starting ontologies will be selected based on a review of their features and a comparison to features requested by energy system researchers. The most appropriate ontologies will then be combined and extended to create the NFDI4Energy domain and scenario ontologies.

Metadata standards will be developed alongside the ontologies to promote data FAIRness. These standards will utilize the terminology of the domain ontology, but will also be designed for interoperability with other research domains, as far as is possible. To that end, TA4 will collaborate closely with other TAs and other NFDI consortia in the creation of the metadata standards. As with the ontologies, existing metadata standards will be investigated and those that are suitable will be integrated into the new NFDI4Energy standards. [Ni22]

Two working groups shall be established as part of the ontology and standards development process. The Metadata & Ontologies working group will provide a forum for energy system researchers and other NFDI consortia to participate in the development of NFDI4Energy's semantic resources, aligning the ontologies and metadata standards and providing training to support the adoption of these resources. The Scenario Description Standardization working group will focus on harmonizing descriptions of energy system scenarios for interdisciplinary research, ensuring alignment with the NFDI4Energy ontologies and metadata standards. Both working groups will allow the TA4 team to receive continuous feedback from stakeholders throughout the ontology and standards development, thus making certain that the new artifacts will be well-suited to serve the needs of the scientific community. [Ni22]

3.2 Services & infrastructure

In support of FAIR research data and to enable the use of the semantic resources, TA4 will develop several software services that will be integrated into the NFDI4Energy platform. Various registries and submission services will promote data findability and accessibility by providing researchers with infrastructure for sharing data and models.

The primary means of accessing the ontologies and standards will be through a Terminology Service. This service will also support the registry of ontologies from other domains that are considered relevant for research in the energy domain. It will follow W3C standards and allow users to search for and retrieve any terminology information that is stored in the platform. [Ni22]

TA4 will also prepare an Open Research Knowledge Graph (ORKG) [Au20] for energy research. ORKGs are platforms designed to link knowledge from multiple research sources and make this knowledge machine-readable, thus simplifying the location and comparison of data. TA4's ORKG will be focused specifically on energy system scenarios, easing comparisons between scenarios to give users a clearer picture of the potential future pathways for energy systems that have been studied. [Ni22]

Two additional services will focus specifically on improving the findability and reusability of research data. A PID service, using either existing concepts such as DOI or using a new concept if needed, will assign PIDs to digital objects and link the identified objects through a PID graph. This service will be tied to a metadata registry, based on the Leibniz Data Manager [Le23], which will register digital repositories of research data from across the internet. The metadata registry will allow users to visualize and compare data sets, thus helping researchers to locate the ideal data for their work. [Ni22]

4 Conclusion

TA4's work will be closely tied to the work of several other TAs. [Ni22]

- TA1 (Building and Serving the Energy Research Community) will provide feedback from researchers to inform the development of TA4's semantic resources. Infrastructure assembled by each of these TAs will be linked together.
- TA2 (Integrating Society and Policy in Energy Research), similar to TA1, will provide stakeholder input for the semantic resources.
- TA5 (Simulation in Interdisciplinary Research) will produce artifacts for energy system simulations, which must be compatible with TA4's resources and services.
- TA6 (Use Cases for Community Services) will identify use cases for the NFDI4Energy platform that TA4 will take into account during resource development.
- TA7 (Organization & Management) will support TA4 in connecting with other NFDI consortia, to pursue the creation of resources with interdisciplinary uses.

Several organizations will be participating in the efforts of TA4, which will be lead by OFFIS e.V. Additional contributors include Albert-Ludwigs-Universität Freiburg, Friedrich-Alexander-Universität Erlangen-Nürnberg, Karlsruhe Institute of Technology – Institute for Automation and Applied Informatics, Reiner Lemoine Institut, Technische Informationsbibliothek, and Carl von Ossietzky Universität Oldenburg. [Ni22] Taken together with those who choose to participate in the working groups, TA4 will benefit from a wealth of diverse experience, viewpoints, and suggestions throughout the crafting of its artifacts.

Bibliography

- [Au20] Auer, S. et.al.: Improving Access to Scientific Literature with Knowledge Graphs. *Bibliothek Forschung und Praxis*, 3/20, pp. 516-529, 2020
- [Bo21] Booshehri, M. et.al.: Introducing the Open Energy Ontology: Enhancing data interpretation and interfacing in energy systems analysis. *Energy and AI* 5/21, 2021.
- [Le23] Leibniz Data Manager, <https://labs.tib.eu/info/en/project/leibniz-data-manager/>, accessed: 11/05/2023.
- [Mc07] McMorran, A.W.: An Introduction to IEC 61970-301 & 61968-11: The Common Information Model. University of Strathclyde 124/07, 2007.
- [Ni22] Nieße, A. et.al.: nfdi4energy – National Research Data Infrastructure for the Interdisciplinary Energy System Research, <https://doi.org/10.5281/zenodo.6772013>, accessed: 03/05/2023.
- [Wi16] Wilkinson, M. D. et.al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3/16, 2016.

Öffentliche Infrastruktur -
IT-Governance und Strategisches
Informationsmanagement
(ITG-SIM)

Operationalizing Enterprise Service Management Capability – the OATIP Model

Daria Goscinska¹, Till J. Winkler²

Abstract: In times of digital transformation, service has evolved to a key concept in the Information Systems research. With a service lens, we focus on Enterprise Service Management (ESM) as an innovative value proposition for organizations. In this paper we attempt to operationalize an organization's ESM capability through a scale-based measure. Based on a service-dominant logic and existing literature we develop the OATIP model for ESM capability consisting of five resources: **organizational, architectural, technological, shared information and people resources**. We qualitatively validate the model through expert interviews and card sorting. Our results can guide organizations in their endeavour to build an ESM capability.

Keywords: Enterprise Service Management, IT Service Management, Organizational Transformation, Construct Development, Service Systems.

1 Introduction

In times of digital transformation, it has become very important for organizations to modernize their service understanding, value creation mechanisms and relations between IT and business. Particularly service has evolved to a key concept in the transdisciplinary Information Systems (IS) research, which is well equipped to understand and develop the complex models of service systems [BLM14]. Bardhan et al. [Ba10] recommend future research to focus on studying the business value of service in organizations.

With a service system lens based on the service-dominant logic of Vargo and Lusch [VL06] we want to focus on Enterprise Service Management (ESM) as an innovative value proposition for organizations. ESM is an approach which aims at transforming and managing organizational workflows as services. The idea of ESM is based on transferring best practices from IT Service Management (ITSM) frameworks like IT Infrastructure Library (ITIL) and COBIT to the (non-IT) business areas of a company [GW22]. Previous research based on a systematic literature review provided the foundation for the ESM construct through a thorough cross-field literature analysis and a unified definition. It also encouraged further research on ESM towards a conceptualized, operationalized model [GW22].

In this paper we want to undertake a theoretically grounded operationalization and

¹ University of Hagen, B*IMA Faculty, Hagen, Germany, daria.goscinska@fernuni-hagen.de

² University of Hagen, B*IMA Faculty, Hagen, Germany; & Copenhagen Business School, Department of Digitalization, Frederiksberg, Denmark, till.winkler@fernuni-hagen.de

qualitative validation by asking the research question: *how can we operationalize an organization's ESM capability through a scale-based measure?* Our contribution is twofold: on the one hand it is conceptual, as we develop the ESM capability as a measurable construct by designing the OATIP—organization, architecture, technology, shared information and people resource—model; on the other hand it is practical, as our results provide organizations with guidance on the necessary resources for developing an ESM capability. We first present the theoretical and research background, followed by the methodology. We then present the qualitatively validated ESM capability model and conclude with a discussion and outlook.

2 Theoretical background

We base our understanding of service on the foundations of Vargo and Lusch's service-dominant logic, where service is understood as the application of resources and competences for the benefit of others. Based on this logic, we apply the service system lens, which is a useful perspective for a service-based concept like ESM. A service system is the basic unit of analysis of service [MS13, VL06]. Service involves at least two actors co-creating value, called service systems, a provider and a client side, which can be individuals, firms or organizations [Sp07, Ma09]. Service systems are evolutionary, adaptive systems, meaning that their operand and operand resources can compose and re-compose dynamically [Ma09]. This recombination of resources creates service innovation which leads to new value propositions [MS13].

Service systems are defined as “a dynamic value-cocreation configuration of resources, including people, organizations, shared information (language, laws, measures, methods), and technology, all connected internally and externally to other service systems by value propositions” [Ma09: 399]. Next to the four main resources people, organization, shared information and technology, other authors emphasize further resources like architecture, customers, other service systems or physical artifacts [Sp07, MP10, BLM14]. Likewise, in Bardhan's et al. [Ba10] listing of useful perspectives on (IT) Service Management we find elements from architectural (modularity, loose coupling, Service-Oriented Architecture (SOA)), technological (interoperability, service-oriented technologies), human (culture, IT alignment) and organizational (dynamic capabilities, strategy) perspectives.

To transfer the theoretical concepts our paper, we view an organization as the service system, in which ESM is a higher-level capability consisting of five main resources: organization, architecture, technology, people, and shared information. The organization reflects the internal and external service structures; architecture decomposes the organization into functional components and its interactions; technology evolves around a tool system and infrastructure; shared information comprises a form of encoded language for coordination; and people refer to human labour but also human capabilities [Sp07, BPLM14]. Each of the resources consists of particular characteristics, which we

put into items. The process of adapting resources towards ESM is viewed as a service innovation creating a new value proposition for the organization.

The ground for the conceptualization was laid in the literature review conducted by Goscinska and Winkler [GW22]. The paper offers a detailed analysis of ESM characteristics across two main research fields: the customer and process-centric ESM related to the IT organization thinking and system-centric ESM related to IT architecture thinking [GW22]. Moreover, the paper provides a common ESM definition of the concept, unifying the two perspectives. The conceptual theme of the ESM construct, understood as a “set of fundamental attributes/characteristics that are necessary and sufficient for something to be an exemplar of the construct” [MPP11: 300], includes the service-oriented organizational structure, the modular IT architecture, the application of service management technology, organization-wide shared information and the adapted people factor. Through the interplay of those five resources, the conceptual domain of ESM is best characterized as an (IT) capability on the entity level of an organization. Tab. 1 summarizes the definition, the conceptual domain of the construct, the entity to which it applies and the conceptual theme of ESM [MPP11]. Having conceptualized the ESM construct, it is still missing an operationalization, which would help organizations to measure their ESM capability and to determine which resources to enable. For this, we develop and qualitatively validate the ESM capability with our OATIP model.

Construct element	Value used for ESM construct	Main reference
Definition	ESM is an approach which aims at transforming and managing organizational workflows as services, which can be used by internal or external customers to fulfill business or IT requests. ESM is created as an interplay of service-orientation, architectural modularization, and technology agnosticism of service management tools.	GW22
Conceptual domain	(IT) capability is a construct that refers to a firm’s abilities or competencies in the IT area.	Bh00
Entity	Organization as a service system	Sp07
Conceptual theme	Interplay of five resources: a service-oriented organization, a modular IT architecture, an agnostic service management technology application, organization-wide shared information and an adapted people factor	GW22, MA09, BPLM14

Tab. 1: Conceptualization elements of ESM according to MacKenzie et al. [MPP11]

3 Methodology

Our construct development research followed the approach suggested by MacKenzie et al. [MPP11], which gives a detailed methodological description particularly tailored to the IS research area. Due to space and time limitations, this paper focuses on the first four steps of the procedure, which include: the conceptualization, development of items and model specification. Fig. 1 summarizes the construct development approach.

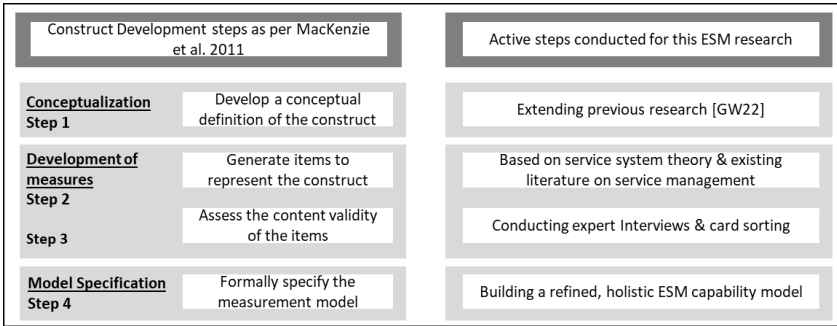


Fig. 1: Own methodological approach based on MacKenzie et al. [MPP11]

Steps two and four develop the ESM capability model, while step three ensures qualitative validity. The goal of the item generation process is to capture all essential parts of the construct and at the same time demarcate it from other concepts outside of its core domain [MPP11]. We used various sources as suggested by MacKenzie et al. [MPP11], including the above theoretical basis, the referenced literature review and existing research on established items. To ensure qualitative content validity of the ESM capability model we conducted expert interviews and a card sorting exercise.

3.1 ESM capability – qualitative model validation

Expert Interviews. The ESM capability model was discussed with seven experts in the service management area for content validation purposes. The semi-structured interviews included experts from academia and practice with 10-25 years of experience (Tab. 2). Involving both expert types ensures that feedback is given holistically from both perspectives. As a result, refinement of the ESM model included main changes, i.e., the modification of resource dimensions; and medium changes like deleting, adding, merging, or shifting resource items; lastly, minor changes in wording increased clarity. Overall, we put more emphasis on the service value chain as a key organizational feature [EI2, EI4], deleted the focus on internal customers [EI2, EI4] and on the IT department for service provision [EI1, EI5], and merged the process-orientation into the workflow logic [EI3, EI7].

Interview coding	EI1	EI2	EI3	EI4	EI5	EI6	EI7
Years of experience	15+	25+	20+	20	10+	10+	15+
Type of expert	academia	practice	practice	practice	academia	practice	academia

Tab. 2: Summary of expert interviews

Card Sorting. As a next step of qualitative validity, we conducted a card sorting of the refined ESM model, in order to advance the rigor of item purification [YG22]. We opted for a randomized, closed card sorting with 39 question item cards and 10 categories according to the ESM capability model. The card sorting was administered via maze.co

among a sample of fellow researchers and affiliated IT professionals. The results were assessed twofold: with the interrater reliability (calculating Cohen's Kappa κ) and with the item placement ratio (percentage of card-category mappings of participants) [YG22]. Subsequently, two raters below the reliability threshold of $\kappa=0.4$ were excluded from the sample, as the result suggests a poor level of agreement [FLP03]. This also demonstrates the necessity to conduct a pre-screening of future survey participants to select the right respondent group. Furthermore, three cards were below the item placement threshold of 50% [YG22] and respectively rephrased for more clarity. In sum, the collective feedback from the experts and fellow researchers clarified and complemented the ESM capability, leading to the final OATIP model as shown in Fig. 2.

4 Results: the OATIP model – an ESM capability operationalized

In this step we elaborate in dept on the ESM capability OATIP model (Fig. 2), structured into five construct resource dimensions, 10 sub-dimensions and 29 construct items.

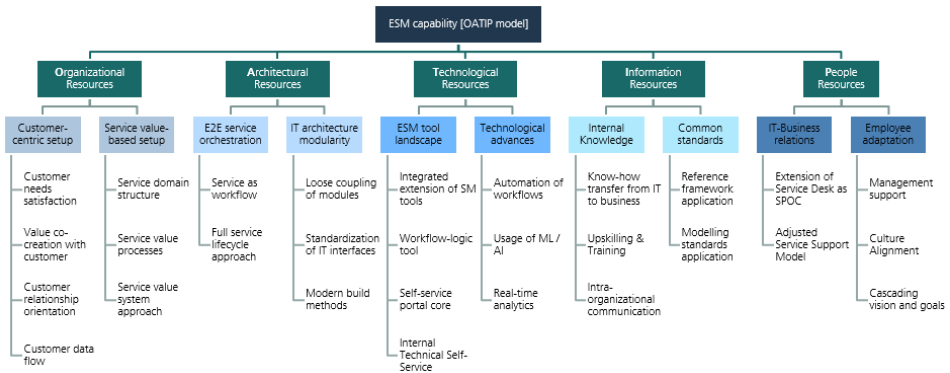


Fig. 3: Validated ESM capability model

Organizational Resources. The construct dimension organizational resources has been compiled into two sub-dimensions: the organization has a *customer-centric* and *service value-based setup*. The design of service systems centres around enhancing service value for the customer through a customer centric focus [Ba10]. Many authors see the customers, their needs and satisfaction, as main drivers of ESM implementation [Mi19, OD19]. As customer-centricity is already an established concept within IS, we base our understanding on Lamberti's [La13] and Spottke's et al. [SBW15] work. Customer-centricity ensures *customer' needs satisfaction* by providing services which address those needs [SBW15, OD19]. An organization with the objective of customer-centricity focuses on *value co-creation with the customer*, characterized by a collaborative dialogue during service creation, usage, and improvement [SBW15]. Customer data and related value opportunities are gathered through a continuous interactive exchange build upon a *customer relationship orientation* [La13, SBW15]. The internal structures of an

organization should be coordinated to ensure that the collected customer information is shared and managed in an integrated *customer data flow* [La13].

The second sub-dimension of organizational resources is a *service value-based setup*. Service orientation can be broadly understood as “a basic set of relatively enduring organizational policies, practices, and procedures intended to support and reward service-giving behaviours that create and deliver ‘service excellence’” [LHM98: 459]. The organization has a *service domain structure*, which views departments as service domains providing internal services to each other [He17a]. A service value-based setup requires the creation of *service value processes*, which are a set of activities in a service value chain from demand to value delivery [Eb21, At23]. The processes are based on the request fulfilment logic, since processing requests for information or incident notifications in the business area—including the allocation of tasks, approvals, or escalations—is not any different than similar requests in the IT area [La17]. Lastly, to create value, an organization should follow the *service value system approach*, a holistic approach for service organizations based on the ITIL v4 framework build around the service value chain. It includes the ITSM components: guiding principles, governance, practices, and continual improvement [Eb21, At23].

Architectural Resources. The construct dimension IT architecture resources “refers to the overarching structure and properties of the relationships among the systems and applications in an organization's IT portfolio” [TK10: 288]. This dimension has been compiled into two sub-dimensions: the IT architecture of an organization follows the ideas of *End-to-End (E2E) service orchestration* as well as *IT modularity*. From IT architecture perspective, “[...] enterprises have redefined their business and IT activities using a component-based approach where a business or IT function is decomposed into a set of standardized business or IT services that have well-defined basic service offerings and standard-based interfaces and permit flexible re-compositions” [HKC05: 203]. To achieve E2E service orchestration, the *services follow a workflow logic*, where “a set of tasks [is] coordinated according to predefined procedural rules to deliver a service and each workflow task is performed by one or multiple services” [Vi21: 758]. The IT architecture of orchestrated services characterizes a *full service-lifecycle approach*, i.e., managing an enterprise service in various phases of a service's life from planning through execution and analysis to decommissioning [En21, Wi08]. This design idea is in accordance with established ITSM practices based on the ITIL domains Service Strategy, Design, Transition and Operations.

The second sub-dimension of IT architecture resources is the *modularity of the IT architecture*, defined as “the degree to which an organization's IT portfolio is decomposed into relatively autonomous subsystems [...] that communicate through standardized interfaces [TK10: 288-290]. Since IT architecture modularity is a well-established IS construct, we base our understanding on Tiwana and Konsynski [TK10]. The *loose coupling of modules* refers to the degree of independence between an organization's IT components, i.e., how much internal changes to one component (infrastructure or application) do not affect the behaviour of others. The *standardization*

of *IT interfaces* refers to the application of organization-wide standards, policies, or guidelines for the IT portfolio to allow for high interoperability between the modules. Lastly, IT modularity is achieved through *modern build methods* like container or cloud solutions for micro-services [TK10, EI3, EI4].

Technological Resources. The construct dimension technological resources has been compiled into two sub-dimensions: an *ESM tool landscape* and the *use of technological advances*. For a tool landscape, a starting point is the *extension of existing SM tools* towards the application areas of ESM, at best based on commercial or open standard software [OD19, Ps21, SK09]. A holistic landscape uses complementary instead of replacement tools and shall therefore ensure technical integration-ness of the diverse toolsets [EI2]. A needed feature and crucial element from other ESM dimensions is a *tool with a workflow logic*, to process and keep track of support requests [En21]. A strong enabler of ESM is a tool landscape with a *self-service portal at its core* [He17a, EI2]. The portal shall also act as an *internal technical self-service* for the IT department and related teams (e.g., DevOps) to support fast delivery of IT outputs [EI3, EI4].

The second sub-dimension is the *usage of technological advances* to support innovation in service system interactions [BLM14]. A first step is the *automation of whole workflows or particular tasks*. The level of automation, i.e., the degree of mixture between human and technological execution of the process, can range from a nearly fully automated process execution by the system, via a co-operating execution between the system and manual work to a mere monitoring system of an otherwise manually conducted process [CMJ06]. O’Connell and Drogseth [OD19: 5] found that “advanced levels of automation favor more mature levels of ESM adoption”. Another technological advance is the *usage of machine learning or artificial intelligence* in the service management area. The involved technology takes over activities beyond mere automation, as in the case of an automated ticket routing or a chatbot guiding the user through the service options [En21]. With the help of *real-time or proactive analytics* modern service management systems allow to react immediately towards occurring problems or to solve them even before the customer notices any inconveniences [Ps21].

Information Resources. The construct dimension information resources has been compiled into two sub-dimensions: the collection and application of *internal knowledge* as well as the organization-wide adaption of *common standards*. *Internal know-how on service management should be transferred from IT to the business organization*, as the IT department has gained years of experience in establishing and improving ITSM processes, tools, and practices [Eb21, He17a]. Moreover, establishing an ESM capability necessitates *upskilling and training*, be it for the service executing staff or the service using customers [EI1, EI3]. Lastly, knowledge is enabled through *intra-organizational communication*, which includes the removal of silos and the use of multichannel, available platforms for inter-team communication [EI4, Mi19, OD19].

The second sub-dimension on *common standards* creates a collective understanding of norms and policies. To implement or enhance the ESM capability, an organization

applies best practices from reference frameworks like ITSM or ITIL, with broad organizational implications for its market-, service-, lifecycle- and process-orientation [Ps21, Mi19]. ITSM follows the idea that IT is delivered as an E2E service [At22], whereas ESM would consider the delivery of non-IT as a service. With regards to IT architecture, many authors refer to *modelling standards application* for the design of unified workflows or services. A key reference for ESM are SOA principles [Wi08, Bh10], which create a common language for the analytical unit of a service [SK09].

People Resources. The construct dimension people resources has been compiled into two sub-dimensions: the organization's transformed *IT-Business relations* and the needed *employee adaptation* for ESM. To process ESM workflows, an integrated centre of competence should be established as a *Single Point of Contact* for customers, meaning to *extend the previously IT-based Service Desk* to include experts from various business units beyond IT [EI1, EI2, EI5]. Further, committing towards ESM as a holistic transformational approach should include an *adapted service support model*, i.e., a Global Business Service or organization-wide shared service model with IT and business stakeholder [Ps21].

The second sub-dimension is the *employee adaptation* during and after the ESM implementation. An organization-wide ESM initiative requires *top-down management support*. Managers guide the affiliated change management measures among the staff [Ps21]. An organizational transformation requires organizational *culture alignment* towards the new structures and working modes [He17a, Wi08, EI6]. Lastly, to commit to ESM, it is important to include the whole organization in the *strategic vision and goals, and cascade them* across all working levels, both on IT and business side [EI4, EI6].

5 Discussion and Conclusion

In this paper we attempted to operationalize an organization's ESM capability through a scale-based measure. We developed the holistic OATIP model as an ESM capability construct based on the methodology of MacKenzie et al. [MPP11]. The applied lens of service systems has proven useful for a construct like ESM, which is a complex model of various resources within an organization. Based on service theory and existing literature the OATIP model consists of five resources: organizational, architectural, technological, shared information and people resources. We have further used two measures to account for qualitative validity. The expert interviews and card sorting refined the model by adjusting dimensions, items, and overall wording. With the OATIP model, our aim was to establish a holistic picture of a complex set of organizational abilities. Our results can guide organizations in their endeavour to build an ESM capability. Adapting resources towards ESM is a service innovation which creates a new value proposition in an organization. This paper fits well into the current service research, as it sheds lights on the key concept of service in times of digital transformation.

The following limitations of this research merit consideration. With a different

theoretical lens, the dimension or item setup of the model could have been depicted alternatively [EI2, EI5, EI7]. However, even so, we believe that the core ideas of the identified capability items would remain similar. Moreover, experts [EI6, EI7] commented that some resources might be more crucial than others, unbalancing the weight of the dimensions and items in the model. We agree and emphasize that as a (IT) capability, the ESM capability is an evolving, maturing type of organizational ability. Organizations interested in establishing ESM should use an iterative, gradual approach towards the resource development. Lastly, in accordance with the scale development methodology, our future research will further analyse the ESM capability model, i.e., collect data, validate the model statistically and evaluate mediating and core dimensions.

References

- [At23] Atlassian: A guide to ITIL and its place in modern ITSM, Blog post, online at: <https://www.atlassian.com/itsm/itil>, 2023 [accessed: 26.04.2023].
- [Ba10] Bardhan, I.; Demirkan, H.; Kannan, P.; Kauffman R.; Sougstad, R.: An Interdisciplinary Perspective on IT Services Management and Service Science. *Journal of Management Information Systems* 26:4, pp.13-64, 2010.
- [Bh00] Bharadwaj, A.: A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation, In: *MIS Quarterly* 24/1, pp. 169–196, 2000.
- [BLM14] Böhmman, T.; Leimeister, J.; Moeslein, Kathrin M.: Service Systems Engineering: A Field for Future Information Systems Research. In: *Business & Information Systems Engineering (BISE)*, 2014.
- [CMJ06] Cuellar, M.; McLean, E.; Johnson, R: The measurement of information system use: preliminary considerations. In: *Conference Proceedings of the ACM SIGMIS CPR*, pp. 164-168, 2006.
- [Eb21] Ebel, N.: Basiswissen ITIL 4: Grundlagen und Know-how für das IT Service Management und die ITIL-4-Foundation-Prüfung. dpunkt.verlag, 2021.
- [En21] Engelbrecht, A.; Pumplun, L.; Bauer, C.; Vida, K.: Künstliche Intelligenz im Enterprise Service Management. In: *Künstliche Intelligenz*, Buxmann, P., Schmidt, H. (eds.), Springer Verlag, pp.149-163, 2021.
- [FLP03] Fleiss, J., Levin, B.; Paik, M.: The Measurement of Interrater Agreement. In: *Statistical Methods for Rates and Proportions* (eds W.A. Shewart et al.), 2003.
- [GW22] Goscinska, D.; Winkler, T.: Enterprise Service Management between IT Organization and IT Architecture Thinking: A Clarifying Literature Review. In: H. Koç, J. Stirna, K. Sandkuhl, et al. (Eds.), *CEUR Workshop Proceedings* 3223, pp.44-55, 2022.
- [He17a] Hennhöfer, F.: Enterprise Service Management als Treiber der digitalen Transformation. *IT Service Management* 42, itVerlag, pp. 6-9, 2017a.
- [HKC05] Huang, Y.; Kumaran, S.; Chung, J.: A model-driven framework for enterprise service management, *Information Systems & e-Business Management* 3/2, pp.201-217, 2005.
- [La13] Lamberti, L: Customer centricity: the construct and the operational antecedents. In: *Journal of Strategic Marketing* 21/7, pp. 588-612, 2013.

- [La17] Landis, M.: Enterprise Service Management im Mittelstand - Der Schlüssel zur Kostenreduktion in allen Servicebereichen. In: *IT Service Management: Fitnesskur für IT und Organisation*. itVerlag, pp. 9-12, 2017.
- [LV06] Lusch, R.; Vargo, S.: Service-Dominant Logic: Reactions, Reflections and Refinements. In: *Marketing Theory*. 6, 2006.
- [LHM98] Lytle, R.; Hom, P.; Mokwa, M.: SERV*OR: A managerial measure of organizational service-orientation. In: *Journal of Retailing* 74/4, pp.455-489, 1998.
- [MPP11] MacKenzie, S.; Podsakoff, P.; Podsakoff, N.: Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *MIS Quarterly* 35/2, pp. 293–334, 2011.
- [Ma09] Maglio, P.; Vargo, S.; Caswell, N.; Spohrer, J.: The Service System Is the Basic Abstraction of Service Science. In: *Information Systems and e-Business Management*, 7, pp. 395-406, 2009.
- [MS13] Maglio, P.; Spohrer, J.: A service science perspective on business model innovation. In: *Industrial Marketing Management* 42, pp. 665-670, 2013.
- [MP10] Mele, C.; Polese, F.: Key Dimensions of Service Systems in Value-Creating Networks. In: *The Science of Service Systems*, 2010.
- [Mi19] Mitrakis, N.: *Die Ausrichtung des IT-Service-Managements auf die Digitalisierung*. Springer Vieweg Research, 2019.
- [OD19] O’Connell, V.; Drogseth, D.: Automation, AI, and Analytics Reinventing ITSM. Research Summary Report. Enterprise Management Association (EMA), 2019.
- [Ps21] Pscheidl, M.: Vom Kunden-Service-Management zum Enterprise-Service-Management durch angepasste Organisationsstrukturen und Digitalisierungsvorteile. In: *Services Management und digitale Transformation*, Altenfelder K., Schönfeld D., Krenkler W. (eds.), Springer Gabler, Wiesbaden, 2021.
- [Sp07] Spohrer, J.; Maglio, P.; J. Bailey; Gruhl, D.: Steps toward a science of service systems. In: *Computer* 40/1, pp. 71-77, 2007.
- [SK09] Spohrer, J.; Kwan, S.: Service Science, Management, Engineering, and Design: An Emerging Discipline-Outline & References. In: *IJISSS* 1, pp.1-31, 2009.
- [SWB15] Spottke, B.; Wulf, J.; Brenner, W.: Consumer-Centric Information Systems: A Literature Review and Avenues for Further Research, at 36th International Conference on Information Systems (ICIS), 2015.
- [TK10] Tiwana, A.; Konsynski, B.: Complementarities Between Organizational IT Architecture and Governance Structure. In: *Information Systems Research* 21/2 pp. 288–304, 2010.
- [Vi21] Viriyasitavat, W.; Da Xu, L.; Dhiman, G.; Sapsomboon, A.; Pungpapong, V.; Bi, Z.: Service Workflow: State-of-the-Art and Future Trends. In: *IEEE Transactions on Services Computing* 16/1, pp. 757-772, 2021.
- [Wi08] Wisnosky, D.; Feldshteyn, D.; Mancuso, W.; Gough, A.; Riutort, E.; Strassman, P.: DoD Business Mission Area Service-Oriented Architecture to Support Business Transformation. *The Journal of Defense Software Engineering*, pp. 25-29, 2008.
- [YG22] Yan, R.; Gong, X.: Peer-to-peer accommodation platform affordance: Scale development and empirical investigation. In: *Journal of Business Research* 144(C), Elsevier, pp. 922-938, 2022.

Ein vereinfachter Ansatz zur Entwicklung einer Digitalstrategie in Unternehmen

Warum das Assessment technischer Schulden ein wesentliches Element der digitalen Transformation darstellt

Mona Möstl¹, Patrick Stoll²

Abstract: Digitalisierung und digitale Transformationen beschäftigen heutzutage praktisch jedes Unternehmen. In dieser Ausarbeitung geht es um die Entwicklung einer Digitalstrategie in Unternehmen. Dabei werden sowohl Treiber und Herausforderungen für digitale Transformationsprojekte berücksichtigt als auch der Bereich technischer Schulden näher beleuchtet, der in vielen Unternehmen in den Hintergrund geraten ist und bei digitalen Transformationsprojekten schnell zu einem unüberwindbaren Problem werden kann. Das ausgearbeitete Konzept stellt eine Handlungsempfehlung basierend auf Erfahrungen aus der Praxis sowie auf wissenschaftlicher Evidenz dar.

Keywords: Digitalisierung, Digitale Transformation, Technische Schulden, IT-Strategie, IT-Management, Strategieentwicklung, Praxisbericht

1 Einleitung

Digitalisierung und digitale Transformation spielen bei Unternehmen eine immer wichtiger werdende Rolle, denn die disruptive Kraft der digitalen Transformation erlaubt es vor allen Dingen branchenfremden Wettbewerbern, neue Märkte zu erschließen und Marktanteile etablierter Unternehmen zu gefährden. Nach [OSK22] stehen Unternehmen mehr denn je vor der Herausforderung, diesen Veränderungen und den wachsenden Kundenanforderungen zu begegnen und ihre eigene Innovationskraft zu beweisen. Die Herangehensweisen bei der Umsetzung von Digitalisierung und digitaler Transformation sind hierbei sehr verschieden; während sich manche Unternehmen auf die Weiter- und Neuentwicklung erprobter Produkte und Dienstleistungen oder die Digitalisierung analoger Prozesse beschränken, können andere Unternehmen innovative, digitale Geschäftsmodelle entwickeln, um neue Möglichkeiten der Wertschöpfung entdecken und nutzen zu können. Laut [OSK22] „[...] ist nicht jedes neu eingeführte Tool oder die Erneuerung oder Optimierung von IT mit einer digitalen Transformation gleichzusetzen.“ Immer mehr Unternehmen denken in dem Zusammenhang über die Entwicklung und Einführung einer Digitalisierungs- oder einer digitalen Transformationsstrategie nach. Da viele Unternehmen jedoch nicht wissen, wo sie technisch genau stehen und wie ihre IT-

¹ FH Kufstein/Tirol, Andreas-Hofer-Str. 7, 6330 Kufstein, mona.moestl@hotmail.de,

² FH Kufstein/Tirol, Andreas-Hofer-Str. 7, 6330 Kufstein, patrick.stoll@lektor.fh-kufstein.ac.at

Architektur aufgebaut ist, besteht die Gefahr, dass die Umsetzung einer entwickelten Strategie die Dauer mehrerer Technologiezyklen übersteigt. Dies zeigt sich nach [Fu21] darin, dass im Schnitt nur circa 43 % aller digitalen Transformationsprojekte als überwiegend erfolgreich bezeichnet werden. Ein Grund hierfür ist, dass Unternehmen den Grad ihrer technischen Verschuldung nicht kennen und neue Digitalisierungsprojekte zu ambitioniert planen. Ziel dieser Arbeit ist es herauszuarbeiten, wie Unternehmen eine Digitalstrategie im Unternehmen entwickeln und anwenden können. Dabei werden u. a. Treiber und Herausforderungen berücksichtigt sowie der Bereich von technischen Schulden näher beleuchtet, da diese laut [El23] oft zum Scheitern von Transformationsprojekten führen können.

2 Definitionen

In der Literatur existieren verschiedene Definitionen der Begriffe Digitalisierung, digitale Transformation und technische Schulden. [Sc17] fasst die gängigsten Definitionen zusammen. Das Bundesministerium für Wirtschaft und Klimaschutz, [Bm15], definiert den Begriff der Digitalisierung wie folgt: „Die Digitalisierung steht für die umfassende Vernetzung aller Bereiche von Wirtschaft und Gesellschaft sowie die Fähigkeit, relevante Informationen zu sammeln, zu analysieren und in Handlungen umzusetzen.“ Der Begriff der digitalen Transformation wird vom Technikum Wien, [Te23], als „[...] der Wandel der Unternehmenswelt durch neue Internet Technologien mit Auswirkungen auf die gesamte Gesellschaft [bezeichnet]. Dabei werden zwecks Performancesteigerung digitale Informations- und Kommunikationstechnologien eingesetzt, indem Unternehmensprozesse, Kundenerlebnisse und Geschäftsmodelle transformiert bzw. weiterentwickelt werden.“ Ähnlich definiert auch PwC in [Sc17] die digitale Transformation. Diese Definition wurde gewählt, da zum einen der konkrete Bezug auf die Transformation von Geschäftsmodellen hergestellt wird und da zum anderen diese Definition nicht nur das System „Unternehmen“ einschließt, sondern den Kreis auf das gesamte soziotechnische Umfeld ausweitet. Als Hauptziel der digitalen Transformation wird die Performancesteigerung durch den Einsatz digitaler Technologien angeführt. Für den Begriff der Digitalstrategie gibt es bisher keine klare einheitliche Definition. [Sc20] definieren basierend auf den gängigsten Definitionen der aktuellen Literatur eine Digitalstrategie als „[...] ganzheitliche Ausrichtung von Digitalisierungsvorhaben in Unternehmen und Organisationen, um den digitalen Wandel zu antizipieren und mitzugestalten. Das mittel- und langfristige Ziel ist es, Wettbewerbsvorteile zu erhalten oder neu zu schaffen.“ In dieser Definition lässt sich die Ausrichtung an den Unternehmenszielen sowie die Unterstützung der Zielerreichung erkennen. Im weiteren Verlauf gehen [Sc20] auch auf den Einsatz von Technologien und Methoden ein. Grundsätzlich lässt sich sagen, dass die Digitalisierung eine Voraussetzung für die digitale Transformation ist. Dabei hat die Digitalisierung bereits mit Einsetzen der Industrie 3.0 gegen Ende der 1960er Jahre begonnen.

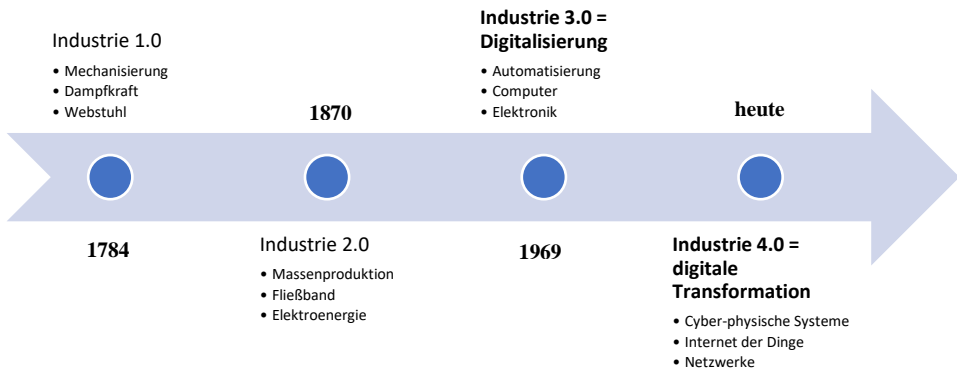


Abbildung 1: Entwicklung der industriellen Revolution (Eigene Darstellung nach Stoll, 2022 und Vario, 2023)

Auch der Begriff der technischen Schulden wird in der Literatur aus verschiedenen Blickwinkeln beleuchtet. Nach [Mi23] sind „Technische Schulden bzw. technische Schuld [ist] in der Informatik der zusammenfassende Begriff für alle Konsequenzen, die aufgrund mangelhaft umgesetzter Software entstehen sowie Kompromisse schon zu Beginn der Entwicklung. Darunter fällt auch der Mehraufwand der betrieben werden muss, schlecht programmierte Software zu ändern bzw. deren Funktionen zu ergänzen.“ Diese Definition zeigt die verschiedenen Ausprägungen, in denen technische Schulden anfallen können; diese können bereits bei der Softwareentwicklung entstehen und nicht ausschließlich im Rahmen von Maintenance und Servicing. Dies ist ein Aspekt, der in der Praxis oftmals nicht berücksichtigt wird und dadurch zu hohen Aufwänden führen kann.

3 Diskussion der Ergebnisse

[Wh20] bezeichnet die digitale Transformation nicht als ein einzelnes Ziel im Unternehmen, sondern als „wesentliche[n] Teil der Geschäftsstrategie, um die Wettbewerbsfähigkeit zu erhalten, Produkte und Dienstleistungen zu differenzieren und die Effizienz zu steigern“. Im folgenden Abschnitt soll eine mögliche Vorgehensweise zur Entwicklung einer Digitalstrategie dargestellt werden. Dabei werden zunächst die häufigsten Treiber für digitale Transformationsvorhaben in Unternehmen näher beleuchtet. Im weiteren Verlauf werden Herausforderungen, die bei der Umsetzung von digitalen Transformationsvorhaben auftreten können, erläutert und abschließend ein Lösungsansatz für die Entwicklung einer Digitalstrategie erörtert.

3.1 Treiber für Digitalisierung und digitale Transformation

Mögliche Haupttreiber für Unternehmen zur Durchführung von Digitalisierung und digitaler Transformation sind nach [OSK22] die sich stetig verändernden Marktbedingungen und Kundenerwartungen. [Ib23] fanden heraus, dass immer mehr Kunden heutzutage die Erwartung haben, ihre Geschäfte möglichst vollständig digital abwickeln zu können. Diese Entwicklungen sind nicht branchenspezifisch zu verorten, was die generische Auswahl der folgenden Beispiele für einen ersten Überblick darstellen soll. In [Sc16] wird die Einführung eines digitalen Showrooms bei der Firma Mazda zur Gewinnung von Kunden über soziale Medien dargestellt. Demnach wurden bereits 2016 etwa 57 % der Kaufprozesse online durchgeführt, bevor der erste persönliche Kundenkontakt entstand. Nicht nur Mazda bedient sich moderner Technologien, um Kundendaten zu gewinnen und effektiv einzusetzen. [SC16] beschreibt darüber hinaus, dass immer mehr Unternehmen Funktionen wie Programmatic Advertising nutzen, um zielgerichtete Werbung zu schalten und Kunden zu gewinnen. Ein weiterer treibender Faktor für die Bestrebungen von Unternehmen, eine digitale Transformation zu beginnen, sind externe Kräfte wie zum Beispiel Wettbewerbsdruck von neuen, innovativen Unternehmen am Markt. Nach [OSK22] müssen bereits etablierte Unternehmen ihre bisherigen Geschäftsmodelle überdenken, um „mit der Zeit mitgehen“ zu können und ihre Marktmacht sowie Kundenstämme nicht zu verlieren. Ein Beispiel hierfür ist die Zunahme an Insolvenzen und Filialschließungen im stationären Modehandel durch eine Zunahme des Onlinehandels, unter anderem getrieben durch temporäre Schließungen im Rahmen der COVID 19-Pandemie. Das angepasste Konsumentenverhalten führte dazu, dass etablierte Modehäuser wie zum Beispiel Peek & Cloppenburg, Appellath & Cüpper oder Galeria Kaufhof mit ihren Geschäftsmodellen des klassischen Ladenverkaufs dem Preis- und Kundenwunsdruck nicht mehr standhalten konnten. Auch nach der Pandemie setzt sich dieser Trend fort, wie z. B. [Rn21] oder [Wi23] berichten. Dabei haben bereits vor einigen Jahren Pioniere – sogenannte radikale Erneuerer – wie Amazon und Netflix, durch die Verwendung neuer Technologien und durch die Entwicklung digitaler Geschäftsmodelle, große Marktanteile gewinnen können. Laut [Ib23] setzen sie diese Entwicklungen unter anderem zur Neuerfindung von Geschäftsmodellen (E-Commerce), zur Prozessoptimierung und zur Verbesserung der Kundenerlebnisse (personalisierte Empfehlungen basierend auf dem Kundenverhalten) ein.

3.2 Die größten Herausforderungen bei der digitalen Transformation in Unternehmen

Wie bereits in Kapitel 1 erwähnt, sind nur ca. 43 % der digitalen Transformationsprojekte erfolgreich. Das viele Projekte scheitern, geht auch aus einer Studie von [Mc20] hervor, die sich mit dem Erfolg von Transformationen auseinandersetzt. Für das Scheitern von Vorhaben zur Umsetzung der digitalen Transformation kann es mehrere Ursachen geben. So werden zum Beispiel von [Hi22] die Komplexität der IT, fehlendes Fachwissen oder mangelndes Interesse der Mitarbeiter:innen angeführt. Nach [OSK22] liegt „Die große Herausforderung einer digitalen Transformation [...] jedoch [...] auch in der Konzeption

einer Digitalisierungsstrategie und [...] in der erfolgreichen Initiierung und Steuerung des Transformationsprozesses.“ Um eine Digitalstrategie zu entwickeln, bedarf es zunächst einer Unternehmens- und einer IT-Strategie, aus der die Digitalisierungsziele abgeleitet werden können. Selbiges gilt für eine digitale Transformationsstrategie. Ist die Strategie entwickelt, erfolgt der Umsetzungsprozess, der sich in der Praxis oft als langwierig erweist. Hier entstehen Schwierigkeiten, sobald der Umsetzungsprozess länger als zwei Technologiezyklen (in etwa drei Jahre) in Anspruch nimmt, da die neuen Technologien bei Einführung unter Umständen bereits veraltet sind. An dieser Stelle kommt das Mooresche Gesetz zum Tragen, was die Notwendigkeit der Eingrenzung des zeitlichen Horizonts unterstreicht. Moores Gesetz besagt, dass sich die Rechenleistung neuer Computerchips alle 12 bis 24 Monate in etwa verdoppelt. Aus Sicht des Endverbrauchers verdoppelt sich nach [Wa16] der Nutzen für den Käufer alle zwei Jahre. Dies sollte bei der Strategieplanung berücksichtigt werden. Die wesentliche Herausforderung für Unternehmen sind vorhandene technische Schulden, die geplante Digitalisierungsprojekte in die Länge ziehen können, wenn sie nicht erkannt und entsprechend behandelt werden. Unternehmen sollten sich bewusst sein, dass mit der Existenz von substantiellen technischen Schulden im Unternehmen die Voraussetzungen für die Umsetzung einer erfolgreichen digitalen Transformation nicht gegeben sind. Zum Abbau der technischen Schulden werden sowohl finanzielle als auch personelle Ressourcen benötigt, die bei digitalen Transformationsprojekten fehlen, diese in die Länge ziehen und unter Umständen sogar zum Scheitern bringen.

3.3 Mögliche Lösungsansätze zur Entwicklung einer Digitalstrategie

Wichtig ist, dass für jegliche Art von digitaler Transformation eine grundlegende Digitalisierung erforderlich ist. Laut [Wi23] starten viele Unternehmen mit der digitalen Transformation, ohne vorher zu prüfen, welchen Grad die Digitalisierung in ihrem Unternehmen bereits erreicht hat. Erst wenn die technischen Schulden weitestgehend abgebaut sind, kann mit einem sinnvollen und nachhaltigen Strategieentwicklungsprozess für die Digitalstrategie begonnen werden. Eine der Hauptherausforderungen ist es nach [Co20], zunächst eine einheitliche Definition im Unternehmen zu finden und diese nachhaltig zu etablieren. Dabei ist nach [Sc20] zu beachten, dass viele Begriffe wie z. B. „Digitalstrategie“, „Digitalisierungsstrategie“ oder „digitale Transformationsstrategie“ sowohl im englischen als auch im deutschen Sprachgebrauch häufig synonym verwendet werden. In der Literatur gibt es verschiedene Ansätze zur Entwicklung einer Digitalstrategie. Hierzu zählt das folgende Beispiel von SAP. Die Firma SAP – laut [St23] mit 33,2 Mrd. US-Dollar Umsatz im Jahr 2021 eines der größten Softwareunternehmen weltweit - empfiehlt eine vierschrittige Vorgehensweise zur Entwicklung einer Digitalstrategie.

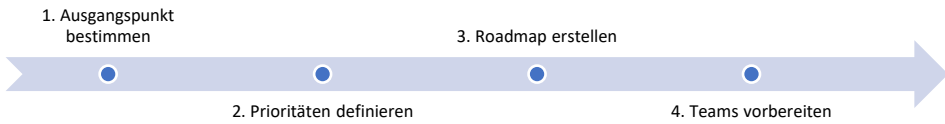


Abbildung 1: Vier Schritte der Transformationsstrategieentwicklung (Eigene Darstellung nach SAP, 2023)

Im ersten Schritt wird der aktuelle Ausgangspunkt bestimmt, um zu ermitteln, wo das Unternehmen (vor allem aus technologischer Sicht) steht (IST-Analyse). Hierzu zählt unter anderem die Prüfung vorhandener Systeme und Anlagen sowie deren Beziehungen zueinander. Im zweiten Schritt werden klare Prioritäten definiert. Ziel ist, die digitale Transformation in mehreren Teilprojekten abzuwickeln. So können Risiken minimiert und durch Teilerfolge positive Effekte bei Mitarbeiter:innen erzielt werden. Im dritten Schritt ist die Erstellung einer Transformations-Roadmap vorgesehen, die laut [Sa23] „[...] konkrete und erreichbare Ziele“ verfolgt und gleichzeitig „[...] Flexibilität und Wachstum“ ermöglicht. Dabei werden nicht nur technologische sondern auch menschliche Aspekte wie das Change Management berücksichtigt, denn die digitale Transformation bringt auch Veränderungen für Mitarbeiter:innen mit sich. Im letzten Schritt geht es darum, die Mitarbeiter:innen für die bevorstehenden Neuerungen zu sensibilisieren und „ins Boot“ zu holen, denn von den Vorteilen, die die digitale Transformation mit sich bringt, kann das Unternehmen nach [Sa23] nur profitieren, wenn alle Mitarbeiter:innen auch offen sind und sich den Veränderungen annehmen. Daher ist eine frühe Einbindung der Mitarbeiter:innen in den Transformationsprozess empfehlenswert.

Das wesentliche Problem bei dieser Herangehensweise ist, dass bei der Entwicklung einer Digitalstrategie in Unternehmen oftmals nur Geschäftsprozesse und / oder Unternehmensziele als Ausgangspunkt gesehen werden. Der technologische Stand wird in der Regel nur unvollständig oder gar nicht ermittelt. Die reine Ausrichtung der Digitalstrategie an den Unternehmenszielen oder der IT-Strategie, ohne den aktuellen technologischen Stand des Unternehmens zu kennen, kann jedoch dazu führen, dass technische Schulden im Unternehmen nicht erkannt und bei der Strategieentwicklung auch nicht berücksichtigt werden. Gerade im Hinblick auf eine digitale Transformation, die zumeist eine Transformation des gesamten Geschäftsmodells zur Konsequenz hat, ist es nach [Go18] wichtig, dass bei der Entwicklung einer Digitalstrategie alle Faktoren berücksichtigt werden und nicht nur einzelne Elemente eine Rolle spielen. Der Zeithorizont für einen klassischen Strategieentwicklungsprozess beträgt in der Regel mindestens ein Jahr. Wäre bereits im Vorfeld bekannt, dass eine digitale Transformation aufgrund zu hoher technischer Schulden nicht umgesetzt werden kann, dann sollte die IT-Strategie primär auf den Abbau der technischen Schulden ausgerichtet werden. Dies schützt nicht nur vor finanziellen oder ressourcenbezogenen Engpässen, sondern hilft auch, den Fokus von Projekten wegzulenken, die aufgrund der technischen Schulden ohnehin nicht umgesetzt werden können. Bei einem geringen Grad der Verschuldung sollte nicht außer Acht gelassen werden, dass die Gefahr eines exponentiellen Wachstums

der technischen Schuld besteht; auch hier sollte der Abbau der technischen Schuld mit hoher Priorität behandelt werden. Hat ein Unternehmen keine technischen Schulden, können Ressourcen und finanzielle Mittel direkt in die Umsetzung der digitalen Transformation investiert werden.

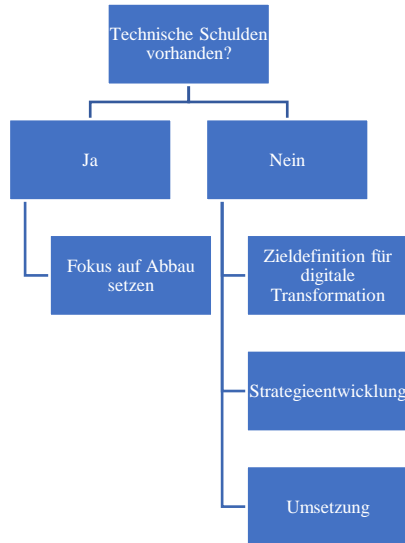


Abbildung 2: Prozess zur Entwicklung einer Digitalstrategie (Eigene Darstellung)

Als Folge aus diesen Überlegungen sollte der Fokus bei der Entwicklung einer Digitalstrategie zunächst auf die Durchführung einer digitalen oder technologischen Inventur gelegt werden. Ziel der Ermittlung ist es, alle vorhandenen (IT-)Systeme des Unternehmens und deren Nutzen/Nutzung zu identifizieren und den aktuellen „Status Quo“ der IT-Infrastruktur zu ermitteln. Dazu zählen auch Bestandteile wie deren Benutzer sowie Beziehungen und Schnittstellen. Des Weiteren spielt die strategische Bedeutung der einzelnen (IT-)Systeme zur Erreichung der Unternehmensziele eine Rolle. Werden an dieser Stelle bereits technische Schulden im Unternehmen aufgedeckt, sollte festgestellt werden, in welchem Umfang diese vorhanden sind und wie sich diese perspektivisch entwickeln. Dies ist erforderlich, um die Ausrichtung der IT-Strategie entsprechend anpassen und steuern zu können. Besonders in großen Unternehmen ist zu überlegen, ob die Einführung eines Enterprise Architecture Managements (EAM) langfristig einen Mehrwert stiften kann, um nachhaltig einen aktuellen Überblick über die IT-Infrastruktur im Unternehmen zu erhalten und entsprechenden Wartungs- / Aktualisierungsbedarf sowie Schwachstellen und Risiken frühzeitig erkennen und beseitigen zu können. Des Weiteren können System-Redundanzen oder veraltete Systeme entdeckt sowie möglicher Optimierungsbedarf ermittelt werden.

4 Limitationen

Eine Limitation der Ausarbeitung ist, dass keine konkreten Zahlen und Daten vorliegen und sich die Ausführungen auf Erfahrungen aus der Praxis stützen. Es gibt keine genauen Informationen darüber, wie viele Unternehmen tatsächlich von technischen Schulden und deren Auswirkungen betroffen sind und wie groß die Ausmaße sind. Des Weiteren sprechen Unternehmen grundsätzlich zwar gerne über erzielte Erfolge, nicht aber über das Scheitern. Eine weitere Limitation, die es zu berücksichtigen gilt, ist, dass technologische Entwicklungen Unternehmen zwar dabei helfen können, neue Anforderungen des Marktes zu erfüllen und damit einhergehenden Veränderungen gerecht zu werden. Wichtig ist nach [OSK22] jedoch, dass diese Technologien nicht willkürlich, sondern zielgerichtet und mehrwertbringend eingesetzt werden.

5 Fazit und zukünftige Forschung

Ziel der Ausarbeitung ist es, eine Handlungsempfehlung basierend auf Praxiserfahrungen zur Entwicklung einer Digitalstrategie in Unternehmen zu geben. Dabei wurden vor allen Dingen Herausforderungen im Hinblick auf technische Schulden beleuchtet, die in nahezu allen Unternehmen vorhanden sind, aber bei der Planung von digitalen Transformationsprojekten oft in den Hintergrund geraten oder gar nicht berücksichtigt werden. Dabei ist gerade die Berücksichtigung von technischen Schulden ein essentieller Aspekt, der viele Transformationsprojekte aus den aufgezeigten Gründen zum Scheitern bringt. Mit der verhältnismäßig simplen Vorgehensweise, vor der Entwicklung einer Digitalstrategie zunächst eine digitale Inventur durchzuführen, kann dieses Risiko eliminiert werden. Für zukünftige Ausarbeitungen wäre eine Großzahlenerhebung bei Unternehmen eine Möglichkeit, um Informationen zum Erfolg digitaler Transformationsprojekte sowie Gründe für das Scheitern von digitalen Transformationsprojekten zu erhalten (vgl. Limitationen). Darüber hinaus könnte auch die Berücksichtigung verschiedener Arten von technischen Schulden (z. B. Requirements Engineering debts) in zukünftigen Ausarbeitungen beachtet werden.

Literaturverzeichnis

- [Bm15] Bundesministerium für Wirtschaft und Klimaschutz, Industrie 4.0 und Digitale Wirtschaft – Impulse für Wachstum, Beschäftigung und Innovation, Berlin, 2015
- [Co20] Correani, A. et al.: Implementing a Digital Strategy: Learning from the Experience of Three Digital Transformation Projects, in California Review Management Summer 2020 Vol. 62 No. 4, 2020
- [El23] Eliav, R.: Priorisierung von technischen Schulden auf agile Art, <https://www.panaya.com/de/blog/modernes-alm/technische-schulden-agile-art/>, Abruf: 05.07.2023

- [Fu21] Fuhrmann, P.: Digitale Transformation: Warum viele Initiativen scheitern, URL: <https://business-user.de/digitalisierung/digitale-transformation-warum-viele-initiativen-scheitern/>, Abruf: 18.04.2023
- [Ib23] IBM: Was ist digitale Transformation?, URL: <https://www.ibm.com/de-de/topics/digital-transformation>, Abruf: 24.03.2023
- [Jo19] Johanning, V.: IT-Strategie. Die IT für die digitale Transformation in der Industrie fit machen, 2. Auflage, Springer Vieweg, Wiesbaden, 2019
- [Go18] Gobble, M.: Digital Strategy and Digital Transformation, in Research-Technology Management 09-10/2018, pg. 66-70, 2018
- [Hi22] Hill, J.: Digitales Dilemma. Daran scheitern Unternehmen bei der digitalen Transformation, URL: <https://www.computerwoche.de/a/daran-scheitern-unternehmen-bei-der-digitalen-transformation.3553761>, Abruf: 18.04.2023
- [Mc20] McKinsey & Company: The secrets to transformation success, <https://www.mckinsey.com/featured-insights/Mckinsey-live/webinars/the-secrets-to-transformation-success>, Abruf: 08.03.2023
- [Mi23] Mindworks, Softwareentwicklung: Technische Schulden, <https://www.mindworks.de/technische-schulden-in-der-softwareentwicklung/>, Abruf: 24.03.2023
- [OSK22] Oswald, G., Saueressig, T., Krcmar, H.: Digitale Transformation. Fallbeispiele und Branchenanalysen, 2. Auflage, Springer Gabler, Wiesbaden, 2022
- [Pa23] Palladio Consulting: Die IT-Strategie – erstellen, beschließen und erfolgreich verankern, <https://www.palladio-consulting.de/it-strategie/>, Abruf: 16.04.2023
- [Rn21] RedaktionsNetzwerkDeutschland: Adler, Esprit und Co.: Diese Modeketten hat Corona in Not gebracht, <https://www.rnd.de/wirtschaft/adler-esprit-und-co-insolvenzen-bei-modehaendlern-in-der-corona-krise-4A3LEUUMCRE4TOLLN4UHYXTCAU.html>, Abruf: 16.04.2023
- [Sa23] SAP: Was ist digitale Transformation?, <https://www.sap.com/germany/insights/what-is-digital-transformation.html>, Abruf: 08.03.2023
- [Sc20] Schallmo, D. R. A. und Lohse, J.: Digitalstrategien erfolgreich entwickeln. Grundlagen, Ansätze und Vorgehensweisen, Springer Gabler, Wiesbaden, 2020
- [Sc17] Schallmo, D. R. A. et al.: Digitale Transformation von Geschäftsmodellen. Grundlagen, Instrumente und Best Practices, Springer Gabler, Wiesbaden, 2017
- [Sc16] Schwarz, T.: Digitale Transformation. Beispiele aus der Praxis, 1. Auflage, Marketing Börse GmbH, Waghäusel, 2016
- [St23] Statista: Die größten Softwareunternehmen weltweit nach Umsatz im Jahr 2021, <https://de.statista.com/statistik/daten/studie/151056/umfrage/umsatz->

- [führender-software-hersteller-durch-software-in-europa/](#), Abruf: 16.04.2023
- [St22] Stoll, P.: IT-Management, Vortrag FH-Kufstein, Tirol, 2022
- [Te23] Technikum Wien Academy, Was ist digitale Transformation?, <https://academy.technikum-wien.at/ratgeber/digitale-transformation-was-ist-das/>, Abruf: 24.03.2023
- [Va23] Vario Lexikon: Industrie 4.0, URL: <https://lexikon.vario-software.de/industrie-4-0/>, Abruf: 18.04.2023
- [Wa16] Waldrop, M. M.: Mehr als Moore, URL: <https://www.spektrum.de/news/mehr-als-moore/1405206>, Abruf: 18.04.2023
- [Wh20] White, N.: Sieben Grundsätze einer effektiven digitalen Transformationsstrategie, <https://www.ptc.com/de/blogs/corporate/digital-transformation-strategy>, Abruf: 08.03.2023
- [Wil23] Wilhelm Büchner Hochschule: Digitalisierung vs. Digitale Transformation, URL: <https://www.wb-fernstudium.de/magazin/ratgeber/digitalisierung/digitalisierung-vs-digitale-transformation.html>, Abruf: 19.04.2023
- [Wi23] Wirtschaftswoche: Diese Modeunternehmen haben Insolvenz-Erfahrung, URL: <https://www.wiwo.de/unternehmen/handel/adler-galeria-esprit-und-co-diese-modeunternehmen-haben-insolvenz-erfahrung/29012944>, Abruf: 16.04.2023

Von Data Fabrics zu Smart Data Fabrics – eine Literaturanalyse

Lena Schlosser¹, Helmut Beckmann²

Abstract: Die zunehmende Digitalisierung der Welt sorgt für eine immer größer werdende Menge an Daten, die gespeichert, verarbeitet und geteilt werden müssen [BS16]. Zudem gehören Daten und die Datafizierung zu den aktuellen Trends [Fs22]. Dabei wird jedoch die Fragen nach Lösungsmöglichkeiten für Technologien für das Datenmanagement nicht beantwortet. Um Unternehmen bei der Strukturierung ihrer Daten und dem Datenmanagement zu unterstützen, helfen Architekturen wie Smart Data Fabrics. In dieser Arbeit soll Anhand einer Literaturanalyse nach [Fp06] und [WW02] dargelegt werden, was Data Fabrics grundlegend sind. Zudem wird dargestellt worin sich Smart Data Fabrics von Data Fabrics unterscheiden und welche Vorteile sie gegenüber diesen und gegenüber traditionellen Datenarchitekturen haben. Genauer werden dabei Bereiche wie Datenexploration, künstliche Intelligenz (KI) und Machine Learning (ML) betrachtet, die die zusätzliche Analysefunktion von Smart Data Fabrics unterstützen [Sj21]. Am Ende lässt sich das große Potential rund um die Smart Data Fabrics darlegen, welches sich vor allem in den Produkten der Datamanagementservicedienstleistern widerspiegelt.

Keywords: Data Fabric, Smart Data Fabric, Digitalisierung, Datenanalyse, künstliche Intelligenz, maschinelles Lernen

1 Einleitung

Durch den technologischen Fortschritt, dessen Geschwindigkeit und den vermehrten Einsatz von Informationstechnologie befindet sich die Gesellschaft in einem Wandel. Um diesen Wandel zu beschreiben können Begriffe wie Digitalisierung und digitale Transformation verwendet werden [Bt19]. Die zunehmende Digitalisierung sorgt dafür, dass vermehrt analoge Prozesse digitalisiert werden. Das führt zu großen Mengen an Daten [BS16].

Diese Daten und deren Analyse spielen in jedem Bereich der Gesellschaft eine Schlüsselrolle. Laut [Fs22] unterstützen die Daten dabei, Muster zu erkennen und diese sinnvoll zu nutzen. Durch die immer größer werdende Menge an Daten wird der Bedarf diese zu speichern, zu teilen und zu verarbeiten immer größer. Um den Menschen dabei zu unterstützen, werden Methoden oder Technologien benötigt, die Entscheidungen übernehmen oder Prozesse vereinfachen [Fs22]

Um diese zu entwickeln, müssen laut [Jv22] zwei Dimensionen beachtet werden: die Semantik von Daten und das Teilen der Daten zwischen unterschiedlichen Speicherstätten.

¹ Hochschule Heilbronn, Wirtschaftsinformatik, Bildungscampus, 74076 Heilbronn, slena@stud.hs-heilbronn.de

² Hochschule Heilbronn, Wirtschaftsinformatik, Bildungscampus, 74076 Heilbronn, helmut.beckmann@hs-heilbronn.de

Durch diese Komplexitäten sind bisherige Datenarchitekturen und -verwaltungssysteme immer mehr ausgelastet.

Die Aktualität des Themas sieht auch [Fs22], der die Trends des Software-Engineering im Jahr 2022 zusammenfasst. Er sieht Daten und deren Analyse als wichtigen Faktor jedes Unternehmens heutzutage. Durch Daten können Services und Angebote geformt werden. Er lässt jedoch auch am Ende des Artikels die Frage nach den Technologien und Methoden im Raum stehen, die notwendig sind, um die Datenverarbeitung zu optimieren.

Bei dieser Beantwortung der Frage trägt diese Arbeit einen Teil bei, da Data Fabrics und Smart Data Fabrics zu den Technologien gehören, die die Datenverarbeitung optimieren.

Um die aufgeführte Problemstellung bezüglich der Methoden und Technologien zu beheben, stellt diese Arbeit eine Technologie für die Datenverarbeitung vor. Dabei orientiert sich der Inhalt dieser Seminararbeit an folgenden Forschungsfragen:

- Welche Unterschiede bestehen zwischen Data Fabrics und Smart Data Fabrics?
- Welche Vorteile haben Smart Data Fabrics?

Das Ziel dieser Arbeit ist die Darstellung der Unterschiede von Data Fabrics und Smart Data Fabrics. Im Laufe der Forschungsarbeit sollen die Merkmale der beiden Begriffe dargelegt und die Vorteile der Smart Data Fabrics erläutert werden.

1.1 Aufbau der Arbeit

Zu Beginn der Arbeit soll in der Einleitung in das Thema eingeführt werden. Dafür werden Problemstellung, Forschungsfrage und Zielsetzung aufgeführt. Im Anschluss daran folgt das Grundlagenkapitel, welches in Data Fabrics, deren Aufbau und Abgrenzung zu anderen Datenspeicher und -architekturen einführt. Im Kapitel 3 wird die Forschungsmethodik dieser Arbeit festgelegt und erläutert. Daraufhin folgt die Präsentation der Ergebnisse dieser Arbeit. Dazu wird zuerst auf Smart Data Fabrics und im Anschluss auf deren Vorteile eingegangen. Im letzten Kapitel der Arbeit findet eine Zusammenfassung der Ergebnisse statt, ebenso wie die kritische Reflektion dieser. Das Ende des Kapitels und damit der Arbeit bildet der Ausblick.

2 Grundlagen und Begriffe

In diesem Kapitel soll eine Einordnung in und deutliche Abgrenzung zu anderen Datenverwaltungs- und -analysearchitekturen stattfinden. Dazu gehören unter anderem Data Lakes und Data Warehouses, die im Folgenden definiert und von Data Fabrics und Smart Data Fabrics differenziert werden sollen.

Außerdem wird zu Beginn der Aufbau von Data Fabrics dargestellt und deren Funktionen aufgezählt. Dies dient als Grundlagen für den Vergleich mit den Smart Data Fabrics, der in einem späteren Kapitel dieser Arbeit durchgeführt wird.

2.1 Abgrenzung

In diesem Kapitel soll eine eindeutige Abgrenzung von Data Fabrics zu anderen, häufig verwendeten Datenverwaltungs- und -analysearchitekturen gezogen werden. Dafür werden einige dieser Architekturen definiert und Unterschiede aufgezeigt.

Grundsätzlich kann zwischen Datenverwaltungssystemen und Datenarchitekturen unterscheiden werden. Zu den Datenverwaltungssystemen gehören unter anderem traditionelle Speichersysteme wie Data Lakes und Data Warehouses [Ic23].

Data Warehouse – Data Warehouses dienen als Speicher für Daten aus unterschiedlichsten Quellen von Unternehmen. Durch Datentransformationen werden die Daten in das korrekte Datenformat des gewünschten Modells gebracht. Die abgespeicherten Daten können im Anschluss für unterschiedliche Anwendungen genutzt werden [Ic23], [Em22].

Data Lake – Data Lakes speichern große Mengen an Rohdaten ab. Diese Daten können sowohl strukturiert als auch unstrukturiert vorkommen. Ihre Entwicklung wurde zur Bewältigung der immer größer werdenden Datenmengen von Big Data angestoßen, um Data Warehouses zu ersetzen. Data Lakes sind langsamer als Data Warehouse aber aufgrund der mangelnden Datenverarbeitung kostengünstiger [Ic23], [Em22].

Datenarchitekturen stellen dar, wie Datenverwaltungssysteme wechselwirken. Diese Architekturen sind unter anderem Datengewebe und Datennetze.

Data Mesh – In Data Meshs werden die Daten dezentral nach Geschäftsbereich sortiert. Das ermöglicht die Betrachtung der Daten als eigenständiges Produkt und den unternehmensweiten Austausch der Daten [Ic23], [Mi22]

Data Fabric – Data Fabrics lassen sich ebenfalls bei den Datenarchitekturen einordnen. Data Meshs und Data Fabrics können auch zusammenarbeiten, wobei Data Fabrics die Erstellung von Produkten beschleunigt [Ic23], [Sj21].

2.2 Data Fabric

Dieses Unterkapitel fasst den Aufbau von Data Fabrics und deren häufigste Anwendungsfelder zusammen.

Data Fabrics zählen, wie im vorherigen Unterkapitel bereits erwähnt, zu den Datenarchitekturen, die den Zugriff auf Daten im Unternehmen ermöglichen und beschleunigen und das rapide Wachstum der Datenmengen kontrollierbar machen sollen. Zudem kann mit ihnen die große Diversität der Daten, bestehend aus Texten, Bildern oder

Sensordaten reguliert werden [Aâ19]. Dafür greifen sie auf Daten unterschiedlicher Quellen zu, transformieren diese und ermöglichen eine sinnvolle Nutzung. [Sj21]

Data Fabrics können architektonisch in sieben Komponenten unterteilt werden (siehe Abb. 2). Komponente eins sind die Datenquellen, die Daten generieren. Diese Daten werden verarbeitet, abgespeichert und innerhalb der Data Fabric genutzt [Bc22]. Die Datenerhebung kann in einzelnen, großen Chargen oder mehreren kleineren Chargen passieren [Am20]. Zu den Quellen können Internet of Things (IoT) Sensoren, Daten aus Enterprise Resource Planning (ERP) Programmen oder von externen Quellen, wie Social Media [Bc22] zählen. Außerdem wird mit dieser Komponente die Priorisierung von Datenquellen festgelegt, die die Relevanz von Daten beeinflusst. [Am20]

Ein weiterer Bestandteil von Data Fabrics sind Systeme, die die Daten analysieren und strukturieren. Dies ist notwendig, da die Informationen meist nur teils oder gar nicht strukturiert gesammelt werden. Um die Daten zu verwerten, werden diese in ein kohärentes Format gebracht, welches eine Weiterverarbeitung ohne Probleme ermöglicht und ein Verstehen der Daten erleichtert [Bc22].

Element drei der Data Fabrics sind die Algorithmen, die bei der Gewinnung von Erkenntnissen unterstützen. Relevante Daten müssen an Use Cases angepasst werden und für Entscheidungsfindung notwendige Einblicke aufzeigen. Gleichzeitig finden Protokollierungen der Daten statt, die für Security und Compliance Themen relevant sind [Bc22]. Durch die Programmierung können Verbindungen erstellt und getrennt und CRUD (Erstellen, Lesen, Update und Löschen) Operationen ausgeführt werden [Am20].

Die vierte Komponente besteht aus den Programmierschnittstellen (Application Programming Interfaces, kurz: API) und den Softwareerstellungstools (Software Development Kits, kurz: SDK). Durch diese beiden Bestandteile kann die Verbindung zum Frontendnutzer hergestellt werden, der über das UI (User Interface) einen Überblick über die wichtigsten Erkenntnisse bekommen kann [Bc22].

Ein weiteres Element ist die Datenverbrauchsschicht. Diese bezieht sich auf das UI, welches Endnutzern zur Verfügung steht und dadurch die Nutzung der Daten ermöglicht. Diese Nutzung kann auf unterschiedliche Wege ermöglicht und erleichtert werden, unter anderem durch die anschauliche Darstellung oder Analysefunktionen [Bc22]. Diese Schicht ist vor allem für wirtschaftliche Betrachtung der Daten relevant und hilfreich, da sie Analysten bei der Arbeit unterstützt [Am20].

Komponente fünf ist die Datentransportschicht, welche bei der Verschiebung von Daten innerhalb der Data Fabric unterstützt. Mit dieser Schicht können Daten zudem verschlüsselt übermittelt und das Entstehen von Duplikaten bei Verschiebung verhindert werden, wodurch unnötige Verwendung von Speicherplatz vermieden wird [Bc22].

Der letzte Bestandteil ist die Hostingumgebung, in der die Data Fabric läuft. Diese beeinflusst den Kern der Data Fabrics, wird jedoch meist extern gehostet. Das kann on-

premises oder in der Cloud passieren [Bc22]. Hier werden zudem wichtige Themen wie Data Security und Data Governance gemanaged. [Aj22]

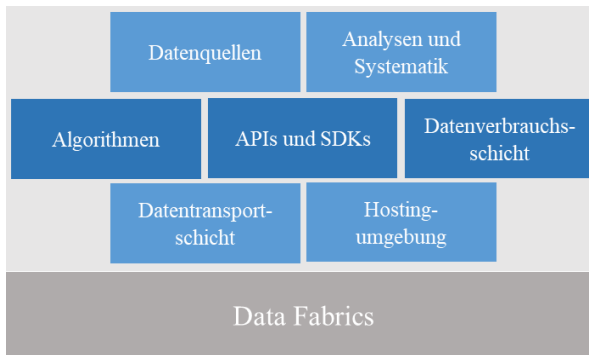


Abbildung 1: Architektur der Data Fabrics (in Anlehnung an [Bc22])

Die Data Fabrics haben mehrere Anwendungsfelder. Ihre Anwendung ist nicht nur technisch affinen Nutzern vorbehalten, sondern kann auch von Nutzern ohne technischen Hintergrund schnell verstanden werden, die mit einfach gehaltenen Dashboards schnell einen Überblick erhalten. Gleichzeitig können mit Data Fabrics aber auch Data Scientists tiefgreifende Analysen der Daten durchführen [Bc22].

Vor allem im IoT Bereich finden Data Fabrics rege Verwendung. Durch diese wird die immer größer werdende Menge an Daten verknüpft und zur Weiterverwendung vorstrukturiert, um einen guten Überblick zu bieten. Dies ermöglicht einen deutlichen Marktvorteil, da eine schnelle Einsicht auf die Daten möglich ist [JG20].

3 Forschungsmethodik

Das Vorgehen dieser Arbeit orientiert sich an der Forschungsmethodik nach [Fp06]. In Kapitel 1 wird die Problemstellung, die dieser Arbeit zugrunde liegt, formuliert und die Zielsetzung definiert.

Die Literatursuche wird mit einer Forward-/Backwardsuche nach [WW02] ergänzt. Für die Literatursuche werden die Beiträge zuerst nach Verfügbarkeit und Relevanz gefiltert und ein Aussortieren der Duplikate durchgeführt. Die Literaturquellen wurden mit Hilfe einer Konzeptmatrix nach [WW02] ausgewertet und im Anschluss interpretiert (siehe Abb. 2).

Die Präsentation der Ergebnisse findet in Form der Grundlagen in Kapitel 2, des Ergebniskapitels 4 und der Diskussion in Kapitel 5 statt.

Paper	Konzepte				
	Grundlagen		Smart Data Fabrics		
	Data Fabrics	Abgrenzung	Allgemeine Infos	Vorteile	Potentiale
Alpoim et al. (2019)	x				
Alvord et al. (2020)	x				
Anand (2022)	x				
Forti et al. (2022)	x				
IBM (2021)	x				
Jahns (2022)	x				
Jane und Ganesh (2020)			x		x
Kuftinova et al. (2022)	x				
Stegmann (2021)	x	x	x	x	x
Theodorou (2021)	x		x		x
Summe	9	1	3	1	3

Abbildung 2: Konzeptmatrix nach [WW02]

4 Ergebnis

In diesem Kapitel findet die Ergebnisdarstellung dieser Arbeit statt. Zu Beginn werden dafür Smart Data Fabrics, deren Aufbau und wichtige Funktionen dargelegt. Darauf folgt die Aufzählung der Vorteile der Smart Data Fabrics gegenüber traditionellen Datenarchitekturen und Data Fabrics.

4.1 Smart Data Fabrics

Um in die Ergebnisdarstellung einzuführen, wird zuerst ein Überblick über Smart Data Fabrics gegeben. Dafür werden Aufbau und Funktion erklärt. Um zu Beginn des Kapitels das Verständnis des Begriffs klarzustellen, wird darauf hingewiesen, dass sich folgende Ausführungen auf Smart ‚Data Fabrics‘ beziehen, womit intelligente Data Fabrics thematisiert werden, und nicht ‚Smart Data‘ Fabrics, was sich auf Strukturen beruft, die mit intelligenten Daten arbeiten.

Grundsätzlich ähnelt der Aufbau von Smart Data Fabrics dem der klassischen Data Fabrics. Sie führen deren Konzept fort und verfügen über zusätzliche Analysefunktionen, die integriert sind, um die aufbereiteten Daten direkt zu verwenden, statt diese an andere Systeme oder Anwendungen zu übertragen [Sj21] (siehe Abb. 3).

Eine dieser Analysefunktionen umfasst die Durchführung von Datenexplorationen [Sj21]. Dabei werden die Daten mittels eines iterativen Prozesses analysiert, um diese bestmöglich zu verstehen. So können immer mehr neue Informationen bestimmt werden, die eine spätere Erstellung von Modellen erleichtern [Fs22]. Als weitere Analysefunktion kann Business Intelligence (BI) genutzt werden [Sj21]. Unter BI-Anwendungen fallen

Prozesse und Verfahren, welche für Einsicht die systematische Analyse des Unternehmens unterstützen. Dabei integrierte Komponenten reichen von Datenextraktion, über Transformation bis hin zu der entscheidungsorientierten Aufbereitung der relevanten Daten und der Analyse dieser [Sb18]. Eine letzte zu nennende Analysefunktion ist das Natural Language Processing (NLP) [Sj21]. Dazu zählen eine Sammlung an Computertechnologien, die sich mit der automatisierten Erfassung und Analyse von menschlicher Sprache, ebenso wie mit deren Wiedergabe auseinandersetzen. Darunter versteht sich die Analyse von Texten ebenso die von gesprochenen Worten. [Ck20]

Diese Analysefunktionen werden unter anderem durch die Verwendung von künstlicher Intelligenz (KI) oder Machine Learning Modellen unterstützt [Sj21]. Beide sind allgemein Funktionen, die in der Data Science Pipeline Anwendung finden und auch bei der Erstellung und dem Training von Modellen genutzt werden [Ho19].

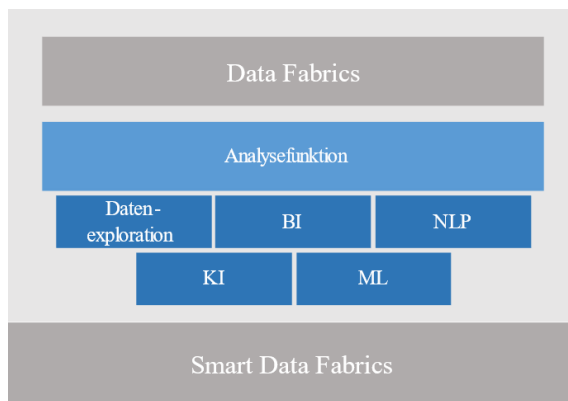


Abbildung 3: Architektur der Smart Data Fabrics (eigene Darstellung)

Smart Data Fabrics können Unternehmen in vielen Situationen unterstützen. Beginnend mit der Planung von Szenarien und der Unterstützung bei Entscheidungsfindungen bis hin zur Vermögensverwaltung des Unternehmens. Außerdem vereinfachen sie das Management von Unternehmensrisiken und das Einhalten von Vorschriften. Verwendung finden Smart Data Fabrics in vielen unterschiedlichen Industrien [IS23].

Ein explizites Beispiel kommt aus dem Bereich des Edge Computing mit dem KI-basierten Data Fabric Konzept MEDAL von [Tv21]. Mit MEDAL sollen Data Operations unterstützt und das Datenmanagement automatisiert werden. Dafür wird ein Dashboard angeboten, auf welchem die Visualisierung der Daten stattfindet. Das Cloud-to-Edge Management findet autonom statt, wobei Workflow Deployment und Qualitätskontrolle eine Rolle spielen. Diese Qualitätskontrolle wird kontinuierlich garantiert. Mittels der KI-Techniken werden Situation analysiert und eventuell notwendige Änderungen vorgenommen.

4.2 Vorteile von Smart Data Fabrics

In diesem Kapitel sollen die Vorteile von Smart Data Fabrics gegenüber anderen Datenarchitekturen und Datenverwaltungssystemen aufgeführt werden, ebenso wie gegenüber Data Fabrics.

Gegenüber traditionellen Datenarchitekturen:

Smart Data Fabrics sind in ihrer Leistungsfähigkeit weniger begrenzt als traditionelle Architekturen. Außerdem wird durch das einheitliche Format der Daten die Bereinigung und Zusammenführung dieser vereinfacht und gestaltet sich dadurch als weniger zeitintensiv. Durch diese Vereinheitlichung ist es Unternehmen möglich, das Zugreifen auf Daten zu beschleunigen, was wiederum das schnelle Schaffen eines Überblicks erlaubt. Mit Hilfe der Smart Data Fabrics ist zudem eine Synchronisation und Verfügbarkeit von Echtzeitdaten leichter möglich [Sj21]. Durch diese Vorteile kann der Arbeitsaufwand, welcher für das allgemeine Datenmanagement nötig ist, deutlich reduziert werden [Kn22].

Gegenüber Data Fabrics:

Durch die Verwendung der in Kapitel 4.1 aufgeführten Analysefunktionen der Smart Data Fabrics haben diese den Vorteil, historische und aktuelle Daten zu kombinieren, um schneller und einfacher neue Erkenntnisse zu gewinnen oder fundierte Entscheidungen aufgrund der Echtzeitdaten zu treffen. Zudem können durch Predictive Maintenance Anwendungen zukünftige Entwicklungen vorweggenommen werden. Weiterhin unterstützen Smart Data Fabrics dabei, die Komplexität von IT-Infrastrukturen zu reduzieren, sowie deren Wartung und Betrieb zu vereinfachen [Sj21].

5 Diskussion und Ausblick

Am Ende der Arbeit sollen die Ergebnisse zusammengefasst werden. Zudem findet eine kritische Diskussion der Ergebnisse statt und eine Aufzählung von Potentialen für die Zukunft. Zum Schluss wird ein Ausblick auf weitere Forschungsarbeit gegeben.

Smart Data Fabrics bestehen, wie auch Data Fabrics, aus den Komponenten Datenquellen, Analyse und Systematik, Algorithmen, APIs und SDKs, Datenverbrauchsschicht, Datentransportschicht und Hostingumgebung [Bc22]. Der Unterschied liegt in den zusätzlichen Analysefunktionen, die in Smart Data Fabrics integriert sind. Dabei werden Datenexploration, Business Intelligence und Natural Language Processing mit Unterstützung durch künstliche Intelligenz und Machine Learning Methoden genutzt [Sj21]. Das bringt einige Vorteile gegenüber traditionellen Datenanalyse-systeme mit sich, wie etwa die Vereinfachung und Zeiteinsparung durch ein einheitliches Datenformat. Zusätzlich dazu haben Smart Data Fabrics auch einige Vorteile gegenüber Data Fabrics. Diese sind unter anderem die Kombination von Echtzeitdaten mit historischen Daten oder

die Nutzung von Predictive Maintenance Anwendungen. Damit könne auch die Forschungsfragen des Papers beantwortet werden, die sich mit den Unterschieden zwischen Data Fabrics und Smart Data Fabrics und den Vorteilen der Smart Data Fabrics auseinandersetzen.

Zum Abschluss der Ergebniszusammenfassung wird noch ein kleiner Überblick über die Nachteile der Smart Data Fabrics gegeben, die trotz der überwiegenden Menge an Vorteilen nicht vernachlässigt werden dürfen. Ein Nachteil ist der hohe Arbeitsaufwand, der vor allem bei der Errichtung einer Data Fabric anfällt, aber auch bei der Maintenance der Architektur ansteht [Ss21]. Außerdem wird die Komplexität zu Beginn der Nutzung erst komplexer, bis alle Systeme korrekt integriert sind und das volle Verständnis für alle Komponenten und Funktionen gegeben ist.

Das Konzept der Smart Data Fabrics ist ein sehr neues Konzept. Aufgrund dessen finden sich nicht viele Artikel mit Informationen dazu, in den meisten Fällen ist von Konzepten die Rede. Die Beiträge, welche sich mit Smart Data Fabrics auseinandersetzen sind zudem meist nicht vor dem Jahr 2021 veröffentlicht worden. Die meisten Information zu diesem Thema sind auf Seiten von Technologie- und Datamanagementunternehmen zu finden. Diese bieten ihren Kunden seit kurzer Zeit Lösungen zum Datenmanagement auf Basis von Smart Data Fabrics an. Dazu führen sie Kunden auf ihren Webseiten in die Thematik ein [Ic21], [Is23].

Dadurch, dass die Thematik rund um Smart Data Fabrics so neu ist, besteht noch viel Potential für Entwicklung in der Zukunft. Durch die konstante Weiterentwicklung der vielen neuen Technologien, wie künstliche Intelligenz und Machine Learning, die für die Umsetzung genutzt werden, findet ein ständiger Ausbau des Konzepts statt. Zudem hat das Datenmanagement und die kommerzielle Nutzung von Daten im Unternehmensbereich einen immer höheren Stellenwert [Fs22]. Dieser wird ebenfalls zu einer konstanten Weiterentwicklung von Smart Data Fabrics führen, die vor allem auch von Anbietern von Datenmanagementsystemen vorangetrieben wird. Als Hypothese lässt sich deswegen die immer größer werdende Popularität von Smart Data Fabrics aufstellen.

Am Ende soll ein Blick auf zukünftige Forschungsarbeit in der Thematik geworfen werden. Da Smart Data Fabrics aktuell in der Industrie vermehrt aufzufinden sind, ist eine Analyse der Angebote und Produkte von Anbietern von Datenmanagementsystemen sinnvoll. Das dient dazu, einen tieferen Überblick darüber zu erhalten, was im Rahmen von Smart Data Fabrics angeboten wird. Zudem ist eine Forschung in Richtung der einzelnen Komponenten sinnvoll, die detailliert deren Aufbau darstellt. Aufgrund der Neuheit des Themas wird außerdem eine Beobachtung der Veränderungen im Laufe der nächsten fünf bis zehn Jahre empfohlen, die dann Mittels einer wissenschaftlichen Arbeit dokumentiert werden kann.

Literatur

- [Aâ19] Alpoim Â, Guimarães T, Portela F, Santos MF Evaluation Model for Big Data Integration Tools. In: ACIS, 2019. Springer,
- [Am20] Alvord MM, Lu F, Du B, Chen C-A (2020) Big Data Fabric Architecture: How Big Data and Data Management Frameworks Converge to Bring a New Generation of Competitive Advantage for Enterprises.
- [Aj22] Anand JV (2022) Digital Transformation by Data Fabric. IRO Journal on Sustainable Wireless Systems 4 (3):196-201
- [Bt19] Barton T, Müller C, Seel C (2019) Hochschulen in Zeiten der Digitalisierung. SpringerLink,
- [Bc22] BasuMallick C (2022) What Is Data Fabric? Definition, Architecture, and Best Practices. Spiceworks.
- [BS16] Bengler K, Schmauder M (2016) Digitalisierung. Springer-Verlag Berlin Heidelberg, DOI 10.1007/s41449-016-0021-z
- [Ck20] Chowdhary KR (2020) Natural Language Processing. In: Fundamentals of Artificial Intelligence. Springer,
- [Em22] El Aissi, M. et al. (2022). Data Lake Versus Data Warehouse Architecture: A Comparative Study. In: Bennani, S., Lakhrissi, Y., Khaissidi, G., Mansouri, A., Khamlichi, Y. (eds) WITS 2020. Lecture Notes in Electrical Engineering, vol 745. Springer, Singapore. https://doi.org/10.1007/978-981-33-6893-4_19
- [Fp06] Fettke P (2006) State-of-the-Art des State-of-the-Art. WIRTSCHAFTSINFORMATIK 48:257-266
- [Fs22] Forti S, Breitenbücher U, Soldani J (2022) Trending Topics in Software Engineering. ACM SIGSOFT Software Engineering 47 (20-21)
- [Ho19] Hein O (2019) Aufbau einer Data Science Pipeline. Informatik Aktuell.
- [Ic21] IBM C (2021) Data fabric architecture delivers instant benefits.
- [Ic23] IBM C (2023) Was ist eine Datenarchitektur?
- [Is23] Intersystems (2023) Smart Data Fabrics.
- [Jv22] Jahns V (2022) Data Fabric and Datafication. ACM SIGSOFT Software Engineering 47:30-31
- [JG20] Jane JB, Ganesh EN (2020) Big Data and Internet of Things for Smart Data Analytics Using Machine Learning Techniques.
- [Kn22] Kuftinova NG, Ostroukha AV, Filippovaa NA, Gaevskii VV, Podgorny AV (2022) Integration of Scalable IT Architectures on the Basis of Data Fabric Technology.
- [Mi22] Machado IA, Costa C, Santos MY (2022) Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures, In: Procedia Computer Science, Volume 196
- [Sb18] Scarff B (2018) So beeinflusst Machine Learning die Business Intelligence. Bigdata Insider,
- [Sj21] Stegmann J (2021) Smart Data Fabric ist das neue Ideal im Datenmanagement. Informatik Aktuell.
- [Ss21] Swoyer S (2021) Data Fabric Architecture: Advantages and Disadvantages. ITPro Today,
- [Tv21] Theodorou V, Gerostathopoulos I, Alshabani I, Abell'o A, Breitgand D (2021) MEDAL: An AI-Driven Data Fabric Concept for Elastic Cloud-to-Edge Intelligence.
- [WW02] Webster J, Watson RT (2002) Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26

Wertbeitrag von Big Data, KI und Maschinellem Lernen. Bezugsrahmen und Demonstration für Prüfungs- und Aufsichtsprozesse

Lisa Otten  und Matthias Goeken¹

Abstract: IT und Informationsverarbeitung wurde in Organisationen lange vor allem als Unterstützungsfunktion angesehen. Dies wandelte sich in jüngerer Zeit und sie sollen heute darüber hinaus Wertbeiträge hervorbringen. In der Informatik-Literatur wird jedoch überwiegend betrachtet, wie neue Technologien implementiert werden. In diesem Beitrag soll hingegen untersucht werden, welche Wertbeiträge sich durch IT ergeben können, wobei der Schwerpunkt auf Big Data, Künstlicher Intelligenz und Machine Learning (BD/KI/ML) liegt. Basierend auf einer Literaturrecherche werden Bezugsrahmen zur Kategorisierung möglicher Wertbeiträge erarbeitet. Diese werden anhand von Wertbeiträgen von BD/KI/ML in Prüfungs- und Aufsichtsprozessen illustriert, was methodisch als Fundierung und Demonstration der Bezugsrahmen angesehen werden kann.

Keywords: Wertbeitrag, Framework, Aufsichtsprozesse, Prüfungsprozesse, Machine Learning, KI

1 Einleitung

Aktuell investieren viele Organisationen in Big Data, KI und Machine Learning und BD/KI/ML werden als bedeutsame Ansätze angesehen, die in vielen Gebieten tiefgreifende Veränderungen und Innovationen hervorbringen werden. In der öffentlichen Debatte – zumal in Deutschland und Europa – scheinen zurzeit jedoch Risiken und Gefahren im Vordergrund zu stehen. Der Nutzen und ein möglicher Wert für Individuen, Betriebe und die Gesellschaft wird öffentlich weniger diskutiert. Allerdings ist es für Unternehmen und andere Organisationen von großem Interesse, neben Risiken und Gefahren auch Nutzen und Wertbeiträge in den Blick zu nehmen – nicht zuletzt, um Investitionen in entsprechende IT-Ressourcen zu beurteilen und zu begründen.

Im Folgenden wird – anknüpfend an die Literatur zum Wertbeitrag der IT im Allgemeinen (Business Value of IT) – der Wertbeitrag von BD/KI/ML beleuchtet. Nach einer Diskussion grundlegender Begriffe (Abschnitt 2) werden ein multidimensionaler Bezugsrahmen zum Wertbeitrag der IT vorgestellt (3.1) sowie ein spezialisierter Bezugsrahmen inhaltlicher Wertbeitragsdimensionen für BD/KI/ML entwickelt (3.2). Diese werden in (4) herangezogen, um mittels einer Fall- und Literaturstudie zu Prüfungs- und Aufsichtsprozessen seine Anwendbarkeit zu demonstrieren. Den Verfassern sind keine Arbeiten bekannt,

¹ Hochschule der Deutschen Bundesbank, Hachenburg, lisa.otten | matthias.goeken@bundesbank.de

die den Wertbeitrag und Nutzen von BD/KI/ML in diesem Bereich systematisch untersuchen. Daher wird mit diesem Beitrag auch der Versuch unternommen, diese Lücke zu füllen. Dabei wird auf vorhandener Literatur aufgebaut und die Ergebnisse sind somit als deskriptiv zu verstehen.

2 Grundlegendes zum Wert und Wertbeitrag von IT

Was unter „Wert“ zu verstehen ist, wird in verschiedenen Denkschulen und Wissenschaftsdisziplinen sehr unterschiedlich beschrieben. Ganz allgemein kann Wert als „ideelle, geistige Bedeutung und Wichtigkeit von etwas“ angesehen werden. In diesem Fall ist etwas „*ein Wert (an sich)*“, was auch als substanzieller oder intrinsischer Wert bezeichnet wird („Wert-Sein“) (vgl. [De14], S. 25 ff.).

In der Ökonomie (BWL/VWL) werden hingegen eher (1) ein Überschuss als Wert angesehen oder ein sogenannter „attributiver Wert“ betrachtet („Wert-Haben“/ „Wert für etwas“). In der attributiven Interpretation bezeichnet Wert (2) etwas, das zur Befriedigung von Bedarfen und Bedürfnissen beiträgt oder (3) den Grad der Brauchbarkeit eines Gegenstandes oder einer Leistung für die Erfüllung eines Ziels (vgl. [Fi04], S. 23).

Im Sinne von (1) kann einem einfachen ökonomischen Kalkül folgend als Wertbeitrag ein Nettoeffekt verstanden werden, der sich ergibt, wenn positive Wirkungen negativen gegenübergestellt werden [Tö20, S. 4]. Diese Sichtweise verdeutlicht, dass nicht nur in der IT und durch Einsatz von IT Kosten gesenkt werden können, sondern IT auch dazu beitragen kann, zusätzliche Umsätze zu generieren. Allerdings sind dieser Sichtweise enge Grenzen gesetzt, da IT-Einsatz aufgrund seiner Querschnittsfunktion oftmals keine direkt messbaren oder gar monetär bewertbaren Effekte mit sich bringt. Tallon et al. empfehlen daher eine breitere Sichtweise einzunehmen: „managers may want to consider more holistic measures than simply trying to link IT to financial performance“ [Ta20, S. 9].

Sieht man entsprechend und im Sinne von (2) einen Bedarf als eine Diskrepanz zwischen einem tatsächlichen und einem gewünschten Zustand, dann würde ein Wertbeitrag (beispielsweise eines Anwendungssystems) darin bestehen, diese Diskrepanz zu beseitigen oder zu verringern. Diese Sichtweise findet sich zwar in der Informatik, allerdings ist sie dort eher für das Anforderungsmanagement einschlägig.

Beschreibt man wie bei (3) ein Ziel durch seinen *Zielinhalt* (Beschreibung des Sachverhalts, auf den sich das Ziel bezieht – also „was“), ein *Zielausmaß* (Messung des Ziels und Festlegung, bei welchem Maß es als erreicht gelten kann – „wie viel“) sowie einen *zeitlichen* (wann) und *organisatorischen Bezug* (wer) [Th20, S. 46 ff.], lässt sich der Wertbeitrag eines Gegenstandes oder einer Leistung – beispielsweise der IT oder einer IT-Investition – als ihre Rolle bei der Zielerreichung bzw. ihre Unterstützung zur Zielerreichung auffassen. Dies bedeutet auch, dass beschriebene und definierte Unternehmensziele inhaltlich festlegen, was als Wert und Wertbeitrag der IT angesehen werden kann, und er sich dadurch ergibt, dass Unternehmensziele in einem höheren Ausmaß erreicht werden.

Ansätze, die den Wertbeitrag im Sinne von (3) attributiv und „holistisch“ betrachten, sind zum Teil direkt oder indirekt mit dem ‘Bundle of Benefits’-Ansatz verbunden [Tö20, S. 5]. Folgt man der ganzheitlichen Auffassung, dass IT-Einsatz eben nicht nur direkte, quantitativ messbare Effekte hat, sondern vielfältige Ziele verfolgt werden, wirkt die Sicht als „Nutzenbündel“ sinnvoll und geeignet, intangible, indirekte und qualitative Effekte zu erfassen und dabei verschiedene Zielinhalte und -dimensionen in den Blick zu nehmen.

3 Kategorisierung von Wertbeiträgen der IT und von BD/KI/ML

In diesem Kapitel werden inhaltliche Wertbeitragsdimensionen für den IT-Einsatz allgemein dargestellt (Abb. 1) und eine weitergehende multidimensionale Kategorisierung erarbeitet (Abb. 2); hierauf aufbauend wird ein konzeptioneller, allgemeiner Bezugsrahmen für Wertbeiträge von BD/KI/ML entwickelt (Abb.3 – Abschnitt 3.2).

3.1 Bezugsrahmen aus der Literatur zum „Business Value of IT“

Bereits vor mehr als 20 Jahren sind in der Literatur im Sinne des „Bundle of Benefits“-Ansatzes erste Aufstellungen von möglichen IT-Wirkungen erschienen [ML98, SS00], die einen sachlich-inhaltlichen Fokus haben und auch als „Wirkungskataloge“ bezeichnet werden [Sc22, S. 197 ff.]. Abbildung 1 zeigt eine Reihe solcher Aufstellungen und stellt sie einander gegenüber (zu einer ausführlicheren Herleitung siehe [KI23, S. 79 ff.]). Sie ist „high level“ da jeweils nur die oberste Ebene einbezogen wird (hier als „Wertbeitragsdimension“ bezeichnet). Darüber hinaus sind die Dimensionen der Tendenz nach – von links nach rechts – von effizienz-/technikorientiert/operativ zu effektivitätsorientiert/strategisch angeordnet.

	Wertbeitragsdimensionen				
Afflerbach [Af15]	Cost investment		Revenue investment		
Melville et al. [Me04]	Efficiency impacts		Competitive impacts		
Töhönen [Tö20]	Efficiency	Effectiveness		Business transformation	
Oh/Pinsonneault [OP07]	Cost-reduction	Quality-improvement		Revenue-growth	
Mirandi/Lederer [ML98]	Transactional Dimension		Informational Dimension		Strategic Dimension
Mooney [Mo96]	Automational		Informational		Transformational
Gregor et al. [Gr06]	Transactional	Informational		Strategic	Transformational
Shang/Saddon [SS00]	IT Infrastructure	Operational	Organizational	Managerial	Strategic
Bartsch [Ba15]	Transaktions-bezogene Wertbeiträge	Informations-bezogene Wertbeiträge	Strategische Wertbeiträge	Transformations-bezogene Wertbeiträge	Wertsichernde Wertbeiträge

Abb. 1 – Sachlich-inhaltliche Dimensionen des Wertbeitrags – „high level“

Es ist offensichtlich, dass die Aufstellungen verschiedener Autoren sich mit Blick auf die Anzahl der Wertbeitragsdimensionen unterscheiden (aufsteigend angeordnet). Des Weiteren ist zu erkennen, dass es zwischen den Aufstellungen deutliche Überschneidungen in

Bezug auf die gewählten Dimensionen gibt. Bspw. umfassen nahezu alle Aufstellungen *transaktionsorientierte* Wertbeiträge (auch als effizienz-/kostenorientierte bzw. operative Effekte erfasst). Wettbewerbsbezogene und umsatzsteigernde Effekte des IT-Einsatzes werden in anderen Aufstellungen unter den Begriff der „*strategische Wertbeiträge*“ subsumiert. Ebenfalls mehrfach finden sich informatorische Wertbeiträge. Andere Dimensionen hingegen sind lediglich bei einer Aufstellung enthalten.

Neben der sachlich-inhaltlichen Betrachtung werden in der Literatur alternative bzw. ergänzende Möglichkeiten der Kategorisierung von Wertbeiträgen vorgeschlagen. Diese werden im Folgenden betrachtet und zu einer multidimensionalen Strukturierung möglicher *Fokusse* auf Wertbeiträge (WBe) ausgebaut (siehe Abb. 2, in der in der ersten Zeile, basierend auf Abb. 1, eine Konsolidierung der Wertbeitragsdimensionen zu finden ist).

Seufert et al. [Se21] und Chau et al. [Ch07] betrachten – anknüpfend an Literatur zur IT-Effektivität – *Ebenen*, auf denen der Wertbeitrag wirksam wird, und unterscheiden in dieser Hinsicht Benutzerzufriedenheit sowie individuelle, organisatorische und gesellschaftliche Wirkung von IT-Einsatz (siehe Abb. 2, Zeile 2: F-E – steht für Fokus-Ebene)

Chau et al. [Ch07] beziehen des Weiteren die *Untersuchungseinheit* („Unit of Analysis“) (Individuell, Organisatorisch, Branche/Gesellschaft) mit ein (F-U). In dieselbe Richtung geht die Unterscheidung bei [Tö20], die allerdings noch die Untersuchungseinheit „Prozess“ vorsieht (dazu auch [Me04]). Auf den ersten Blick mag „Untersuchungseinheit“ redundant zu den aufgeführten „Ebenen“ erscheinen; allerdings sind beide Fokusse unabhängig voneinander zu betrachten: Bspw. könnte die organisatorische Wirkung des IT-Einsatzes in einer bestimmten Branche erhoben werden.

Fokus	Ausprägungen							
Sachlich-inhaltliche WBe - F-I	Infrastruktur-bezogene WBe	Transaktions-bezogene WBe	Organisations-bezogene WBe	Informations-bezogene WBe	Strategische WBe	Transformations-bezogene WBe	Wertsichernde WBe	
Ebenen - F-E	Benutzerzufriedenheit		Individuelle Wirkung		Organisatorische Wirkung		Gesellschaftliche Wirkung	
Untersuchungseinheit - F-U	Individuum		Prozess		Organisation		Branche/Gesellschaft	
Interessengruppen - F-IG	Unabhängiger Beobachter		Einzelpersonen		Gruppen		Management oder Eigentümer	
Systemtypen - F-S	ein Aspekt der IT-Nutzung		einzelne IT-Anwendung	Art von IT/IT-Anwendung	alle IT-Anwendungen einer Organisation		ein Aspekt einer Systementwicklungsmethodik	IT-Funktion einer Organisation
Art der Daten - F-D	objektiv gemessener Nutzen/Wertbeitrag				wahrgenommener Nutzen/Wertbeitrag			
Funktionsbereich - F-F	Cross Org. Activities	Logistic	Service	Procurement	HR	Marketing & Sales	Operations	...

Abb. 2 - Bezugsrahmen für Fokusse auf den Wertbeitrag der IT

Bei Seddon et al. [Se99] finden sich zum einen fünf Arten von *Interessengruppen* („Stakeholder“): (1) unabhängige Beobachter, (2) Einzelpersonen, (3) Gruppen, (4) Management/Eigentümer und (5) Land (F-IG). Zum anderen unterscheiden sie nach der Art des zu bewertenden Systems *Systemtypen*: (1) ein Aspekt der IT-Nutzung, (2) eine einzelne IT-Anwendung, (3) eine Art von IT oder IT-Anwendung, (4) alle von einer Organisation verwendeten IT-Anwendungen, (5) ein Aspekt einer Systementwicklungsmethodik und (6) eine IT-Funktion einer Organisation (F-S). Beim „*Type of data*“ (F-D) wird mit Blick

auf die Art der Messung bzw. die erhobenen Daten zwischen objektiv gemessenem Nutzen/Wertbeitrag und wahrgenommenem Nutzen/Wertbeitrag unterschieden [Tö20].

In ihrer sehr aktuellen und umfangreichen Aufstellung möglicher Wertbeiträge durch IT strukturieren Schütte et al. [Sc22, S. 314 ff] in ihrer Untersuchung entlang von Unternehmensbereichen und betrieblichen Funktionen (unterste Zeile in Abbildung 2 – F-F).

Die resultierende Kategorisierung in Abb. 2 kann verwendet werden, um Untersuchungen und Literaturstudien zum Wertbeitrag zu charakterisieren. Seine Anwendung wird in Abschnitt 4 demonstriert.

3.2 Spezifische Wertbeiträge von BD/KI/ML

Während die Wertbeitragsdimensionen aus Abb. 1 (bzw. Abb. 2, F-I) einen allgemeingültigen, generischen Anspruch bzgl. des IT-Einsatzes haben, finden sich in der Literatur vermehrt Arbeiten, die funktions-, branchen- oder anwendungssystemspezifische Betrachtungen vornehmen. Mit Blick auf BD/KI/ML wurden drei Aufsätze identifiziert, die hinsichtlich einer Erweiterung des obigen Bezugsrahmens geprüft werden.

Grover et al. [Gr18] entwerfen eine Strukturierung des Wertbeitrags aus „Big Data Analytics“-Anwendungen und betrachten v.a. einen sog. „Strategic Business Value“. Sie teilen den strategischen Wert von „Big Data Analytics“ (BDA) in einen „Symbolic Value“ und „Functional Value“. Letzterer entspricht weitestgehend dem, was bereits oben unter Wertbeitrag subsumiert wurde: Verbesserung der Performance und Produktivität durch die Verwendung von BDA, gemessen v.a. durch finanzielle Kennzahlen, d.h. Steigerung von Effizienz, bessere Koordination und Entscheidungsfindung. Unter den Symbolic Value, der sich vor allem auf die externe Umwelt einer Organisation und ihre Stakeholder bezieht, fallen indirekte Wirkungen wie Verbesserung von Reputation und Image. Theoretische Basis für diese „neue“ Art von Wertbeiträgen sind „Signaling“ und „Herding“ (Herdenverhalten). Im Sinne der Signaling-Theorie könnten Organisationen durch BD-Initiativen und Investitionen, die kommuniziert, d.h. der Umwelt „signalisiert“ werden, einen Wertbeitrag erzielen, da sie positive Reputationseffekte bewirken, die sich in steigenden Aktienkursen und Unternehmenswerten niederschlagen [Gr18, S. 395]. „Herding“ bzw. „herd behaviour“ – ein Konzept, das dem Marketing entstammt – meint an dieser Stelle, dass Unternehmen Informationstechnologie einführen und anwenden, um zu demonstrieren, dass sie dem „Mainstream“ angehören und innovativ sowie wettbewerbsfähig sind [Gr18, S. 396]. Während bei Grover et al. der Fokus auf der organisationsbezogenen Wirkung liegt, kann das Signaling darüber hinaus auch einen Effekt auf der Ebene der Branche und/oder der Gesellschaft haben (siehe unten).

Reis et al. [Re20] und Enholm et al. [En21] analysieren nicht BD, sondern adressieren den Wertbeitrag von ML und KI. [En21] entwickeln einen Bezugsrahmen für Effekte künstlicher Intelligenz und ihres Geschäftswerts. Darin werden neben Enablern und Hemmnissen sowie Möglichkeiten der Nutzung von AI („Automation“ und „Augmentation“) Wertbei-

trags-Effekte erster und zweiter Ordnung beschrieben. Effekte erster Ordnung sind Prozesseffizienz, Erkenntnisgewinnung sowie die Transformation von Geschäftsprozessen. Als Effekte zweiter Ordnung bezeichnen sie operationelle, finanzielle und marktbasiertere Performance sowie Nachhaltigkeitseffekte, aber auch unbeabsichtigte Folgen und negative Auswirkungen durch künstliche Intelligenz. Auch hier ist eine deutliche Übereinstimmung mit zu den allgemeinen Wertbeiträgen zu erkennen. Eine Erweiterung stellen die zuletzt genannten Wertbeiträge dar.

Aufbauend auf [En21, Gr18, Re20, ML98, SS00, Ba15, Gr06, Mo96, Tö20 und K123] wurde der Versuch unternommen, einen spezielleren Wirkungskatalog für BD-, KI- und ML-Anwendungen hinsichtlich des sachlich-inhaltlichen Fokus herzuleiten. Dieser ist in Abbildung 3 enthalten und aufgrund der Nichtrepräsentativität der Literaturrecherche noch als „research in progress“ anzusehen. Darüber hinaus soll er wegen der Dynamik des Gegenstandsbereichs als grundsätzlich erweiterbar aufgefasst werden.

Er ergänzt und erweitert den sachlich-inhaltlichen Fokus des Wertbeitrags (F-I) aus Abbildung 2: die informationsbezogenen Wertbeiträge sind um den Aspekt der „Analyse“ erweitert worden, weil dieser mit Blick auf BD, KI und ML besonders relevant ist (F-I-4); Aspekte der organisationsbezogenen Dimension sind in andere Dimensionen integriert worden, insbesondere in die transformatorische (F-I-6), da sich hier deutliche Überschneidungen ergeben. F-I-8 bis F-I-10 sind in Anlehnung an die zuvor dargestellten Arbeiten von [Gr18] und [En21] als Erweiterungen aufgenommen worden.

Sachlich-inhaltliche Wertbeitragsdimension	Wertbeiträge mit Blick auf BD/KI/ML [gemäß En21, Gr18, Re20, ML98, SS00, Ba15, Gr06, Mo96, Tö20 und K123]
F-I-1 Infrastrukturbezogene Wertbeiträge	<ul style="list-style-type: none"> • Erhöhte analytische und geschäftliche Flexibilität durch auswertungsneutrale Rohdaten (bspw. in Data Lakes) oder föderierte Datenarchitekturen und -infrastrukturen (Data Meshes) • Senkung von (IT-)Kosten in der Datenbereitstellung und der Bereitstellung von „Data Services“ sowie durch einen Ordnungsrahmen bezüglich Daten (Data Governance)
F-I-2 Transaktionsbezogene Wertbeiträge (Effizienz)	<ul style="list-style-type: none"> • Operative Effizienz, Prozesseffizienz sowie Effizienz der Kommunikation • Reduktion von Prozesskosten • Automatisierung von Aufgaben, die den Menschen ersetzen • Schnellere Aufgabenerledigung, Beschleunigung von Prozessen, weniger Zeitverzögerungen • Erhöhung der Prozessqualität, höhere Präzision, geringere Fehlerquoten
F-I-4 Informations- und analysebezogene Wertbeiträge	<ul style="list-style-type: none"> • Qualität und Flexibilität der bereitgestellten Information • Leichter und schnellerer Zugang zu Informationen/Reports (auch mittels Visualisierungen) • Neue Erkenntnisse aus und (systematisches und effektives) Erkennen von verborgenen Mustern in großen Datenmengen, die Menschen möglicherweise übersehen • Überwindung der kognitiven Einschränkungen menschlicher Analysten/Entscheidungsträger • Bessere bzw. besser informierte Entscheidungen • Qualität und Geschwindigkeit der Entscheidungsfindung • Verbesserung von Kontrolle und Steuerung • Verbessert Reaktionsfähigkeit und Agilität (z.B. wegen Repriorisierung) • Schnellere und präzisere Reaktion auf Signale und Muster
F-I-5 Strategische Wertbeiträge	<ul style="list-style-type: none"> • Wettbewerbsvorteile, strategische Vorteile • Effektivität des Marketings, der Beschaffung und anderer betrieblicher Funktionen • Verbesserte Kundenbeziehungen und Kundenzufriedenheit • Angebot neuer, innovativer Produkte und Dienstleistungen
F-I-6 Transformationsbezogene Wertbeiträge	<ul style="list-style-type: none"> • Transformation von Prozessen (radikale Transformation) • Re-Engineering und Neugestaltung der bestehenden Organisationsstrukturen • Veränderung des Einsatzes von Humanressourcen bzw. Neuzuweisung von Ressourcen und dadurch Umgestaltung von Organisationsstrukturen • Verbesserung von Geschäftsmodellen • Erweiterung organisatorischer Fähigkeiten (Capabilities)

F-I-7 Wertsichernde Wertbeiträge	<ul style="list-style-type: none"> • Minderung von Risiken wie zum Beispiel Betriebs-, Projekt-, Portfolio- und Outsourcingrisiken (sowohl IT-interne als auch IT-externe Risiken) • Erfüllung von Compliance-Anforderungen
F-I-8 Nachhaltigkeitsbezogene Wertbeiträge	<ul style="list-style-type: none"> • Geschäftsmodelle, die Nachhaltigkeit von Organisationen und/oder Gesellschaft fördern • Ökologische Nachhaltigkeit: Reduktion von Energiekosten, Energieverbrauch und so Verringerung negativer Auswirkungen auf die Umwelt; Verringerung/-meidung von Umweltverschmutzung und Abfall; Transformation in Richtung Kreislaufwirtschaft • Soziale Nachhaltigkeit: Objektive, rationale KI-gestützte Entscheidungen können menschliche Voreingenommenheit und Verzerrungen (Bias) reduzieren; Verbesserung (bspw. durch KI-Roboter) bzgl. Sicherheit und Arbeitsbedingungen der Mitarbeiter; Automatisierung von Routineaufgaben, sodass Mitarbeiter ihre Fähigkeiten anderweitig einsetzen können, was zu sinnstiftenderen und kreativeren Tätigkeiten führen kann
F-I-9 Unbeabsichtigte Konsequenzen/negative Wirkungen	<ul style="list-style-type: none"> • Versäumnis, Verzerrungen in den Daten oder KI-Algorithmen zu erkennen und zu beseitigen • Diskriminierung oder unerwünschten Ergebnissen mit Blick auf bestimmte ethnische Gruppen, Geschlechter oder Bevölkerungsgruppen • negative Auswirkungen auf Reputation und Image, finanzielle Schäden und Geldstrafen • mangelnde Transparenz, Erklärbarkeit und Rechenschaftspflicht (Compliancerisiken) • reduziertes Vertrauen/Misstrauen und Nichtnutzung
F-I-10 „Symbolic Value“	<ul style="list-style-type: none"> • Verbesserung von Reputation und Image durch Signaling und Herding, was zu direkter Wertsteigerung führt und/oder indirekte Effekte haben kann (bspw. Reputationsgewinne)

Abb. 3 Wirkungskatalog/Bezugsrahmen möglicher Wertbeiträge von BD/KI/ML

Dieser Bezugsrahmen hat das Ziel, – im Sinne des ‘Bundle of Benefits’-Ansatzes – Wertbeitragsdimensionen von BD/ML/KI und deren jeweilige Ausprägungen zu illustrieren. Damit ist es im Folgenden möglich, Wertbeiträge entlang der Dimensionen und mit Blick auf eine spezielle Untersuchungseinheit strukturiert zu erfassen.

4 Wertbeiträge in Aufsichts- und Prüfungsprozessen

Nachdem zuvor potenzielle Wertbeiträge von IT im Allgemeinen (3.1) und BD/KI und ML im Besonderen (3.2) dargestellt worden sind, werden in diesem Abschnitt auf der Grundlage einer Literaturstudie konkrete Wertbeiträge, die mit Blick auf Aufsichts- und Prüfungsprozessen genannt werden, identifiziert und mit Hilfe der genannten Wertbeitragsdimensionen strukturiert. Dies dient der Demonstration bzw. Illustration der Bezugsrahmen und damit ihrer Anwendung und Erprobung. Darüber hinaus sind nach unserem Wissen bislang keine systematischen Untersuchungen zum Einsatz von BD/KI/ ML in diesen Prozessen veröffentlicht worden, obwohl dieser für Zentralbanken sowie Prüfungs- und Aufsichtsinstitionen besondere Relevanz besitzt [EC19, Mc22].

4.1 Charakterisierung der Fall- und Literaturstudie

In Aufsichtsbehörden und Prüfungsgesellschaften sind IT sowie BD/KI und ML ein wichtiger Unterstützer und „Enabler“. Dabei sind deren Tätigkeiten hoch reglementiert. Dies führt dazu, dass sie teilweise hohe Barrieren überwinden müssen, sollen neue Technologien eingesetzt werden [Ma20, S. 1339]. Denn Aspekte wie Vertrauen der Bevölkerung [Br18, S. 17-18], algorithmische Fairness [De21, S. 12] und Dokumentation [Ma20, S. 1339] sind in diesem Bereich besonders sensibel und müssen gewährleistet werden, sodass die genannten Dimensionen F-I-7 und F-I-9 von besonderer Bedeutung sind. Gleichwohl

werden auch Potenziale betrachtet, die Tätigkeiten effizienter und effektiver zu gestalten sowie die Prüfungsergebnisse zu verbessern, und damit Wertbeiträge im Sinne der hergeleiteten Dimensionen zu erzielen.

Anhand des multidimensionalen Bezugsrahmens aus Abbildung 2 können wesentliche Merkmale der Fall- und Literaturstudie beschrieben werden:

- F-U: Untersuchungseinheit sind hier primär Prüfungs- und Aufsichtsprozesse, wobei jedoch auch die „benachbarten“ Untersuchungseinheiten (sekundär) in den Blick genommen werden (Individuum und Organisation).
- F-E: Die Veränderung von Prüfungs- und Aufsichtsprozessen mittels des (vermehrten) Einsatzes von IT und BD/KI/ML hat Auswirkungen auf verschiedenen Ebenen: sie hat vornehmlich organisatorische Wirkungen, darüber hinaus möglicherweise aber auch weitergehende, die es zu identifizieren gilt.
- Der Systemtyp (F-S) lässt sich nicht konkret und eng fassen: In den analysierten Quellen ist branchentypisch im aufsichtlichen Zusammenhang von „SupTech“ (Supervisory Technology) und „RegTech“ (Regulatory Technology) die Rede. Es werden jedoch auch generische Funktionalitäten zur Analyse und Datenaufbereitung sowie verschiedene Lernstile und -verfahren von BD/KI/ML adressiert.
- Hinsichtlich der erhobenen Daten (F-D) lässt sich feststellen, dass in den Quellen keine monetär bewerteten und/oder objektive gemessenen Wertbeiträge genannt oder belegt werden. Vielmehr geht es um wahrgenommenen (oder angenommenen) Nutzen aus der Verwendung von BD/KI/ML in Prüf- und Aufsichtsprozessen. Dieser wird lediglich qualitativ begründet.

4.2 Identifizierte Wertbeiträge und Diskussion

In Abbildung 4 werden die im Rahmen einer Literaturstudie identifizierten Nachweise dargestellt (zweite Spalte). Diese Literaturnachweise sind nach einer Funktionalität oder einer spezifischen Wirkung in Aufsichts- und Prüfungsprozessen „gelabelt“ (Spalte 1) dargestellt. Spalte 3 stellt einen Bezug zu den zuvor hergeleiteten Wertbeitragsdimensionen her. In der letzten Spalte ist eine kurze, zusammenfassende Beschreibung der Wirkungen für das jeweilige Label enthalten.

Label	Literaturnachweise für Wertbeiträge durch BD/KI/ML	F-I-	Zusammenfassende Kurzbeschreibung
Daten- erfassung	[Br18 S.16-17] “In the area of data collection, supotech supports a faster and more flexible data capture from supervised firms compared to the traditional template-based approach.”	2	Daten können schneller, flexibler, effizienter und in einem größeren Umfang erfasst werden. Zudem wird die Erfassung von detaillierteren und granulareren Daten möglich. Damit gehen Ressourceneffekte einher.
	[FS20 S. 8] “SupTech applications can improve the value of data collected by enriching its intelligibility and interoperability. (...) RegTech [3] can drive resources efficiencies in collecting data for regulatory filings.”	2 & 4 & 7	
	[FC18 S. 13] “[Machine-executable reporting regulations] provide [...] potential for regulators to collect more detailed, granular data.”	2	
Daten- kombi- nationen	[Br18, S. 8] “SupTech applications are capable of combining multiple data sources to support analytical work. Often this involves connecting structured data and unstructured data. A certain example is found at the Bank of Italy [...], which combines suspicious transactions reports (structured data) with press reviews (unstructured data) for anti-money laundering [...] detection.”	2 & 6	Kombination von strukturierten und unstrukturierten Daten sowie von Daten aus unterschiedlichen Quellen zum Zwecke der Datenanalyse.

	[FS20, S 7] "Certain SupTech tools also allow unstructured data to be integrated into existing data sets for analysis."	2 & 4	
	[Ko18, S. 300] „Im Kontext der Abschlussprüfung werden als Datenanalysen diejenigen Tätigkeiten des Abschlussprüfers verstanden, die dazu dienen, aus strukturierten oder unstrukturierten Daten Informationen zu gewinnen, die im Rahmen von Prüfungsplanungen entsprechend dem Prüfungsprozess –von der Prüfungsplanung bis zu den abschließenden Prüfungshandlungen- genutzt werden können.“	2 & 6	
	[Ma20, S. 1338] „Aufbereitung und Verknüpfung von Daten aus unterschiedlichen Quellen“	2	
Vollprüfung	[Br18, S. 17] "It is impossible for supervisors to review each one closely. Suptech applications can sift through all these regulatory filings and identify potential supervisory issues."	2 & 7	Durch BD/MI/ML können alle prüfungsrelevanten Informationen gesichtet werden. Das Stichprobenrisiko entfällt. Gleichzeitig kann die Transparenz der Prüfungshandlung erhöht werden. Prozesstransformation: Von Stichproben- zur Vollprüfung.
	[Ve18, S. 2582] „So könnte durch den Einsatz von Massendaten die Wesentlichkeitsschwelle abgesenkt und der bisherige stichprobenorientierte Prüfungsansatz zu einer Vollprüfung umgewandelt werden. Durch die Erhöhung der zu Verfügung stehenden Informationen könnte das Entscheidungsverhalten des Abschlussprüfers somit entscheidend beeinflusst werden.“	6 & 4	
	[Ma20, S. 1338] „Höhere Transparenz durch Prüfungen vollständiger Datensätze“	6	
Datenaufbereitung/-exploration	[FS20 S. 8] "SupTech that enables visualisation (such as risk dashboards and charts) can improve the density and complexity of data, helping to transform it into accessible indicators. In addition, SupTech can provide the necessary data inputs for ML and deep learning applications for trend and forecasting analysis."	2 & 4	BD/MI/ML kann Visualisierungen ermöglichen und somit die Komplexität der Daten verringern. Zudem können durch die Aufbereitung Eingaben für weiterführende Analysen geliefert werden.
Analysen	[Br18, S. 17] "In the area of data analytics, suptech drastically cuts the time needed for analysis."	2	Durch BD/MI/ML können Analysen schneller sowie mit mehr und komplexeren Daten durchgeführt werden. So kann die Qualität der Analysen und Prognosen erhöht werden.
	[FS20, S. 7] "[RegTech] can also enable the analysis of large or complex data pools, including customer and risk management data"	4	
	[Ko18, S. 300] „Datenanalysen vertiefen insbesondere in komplexen Prüfungsfeldern die erforderlichen Kenntnisse des Prüfers und erhöhen die Qualität der Identifikation von Fehlrisiken und der erforderlichen Prüfungshandlungen.“	6 & 7	
	[Ve18, S. 2583] „In einer Gesamtschau wird auf Basis der bisherigen empirischen Studien gezeigt, dass Big-Data-Technologien durch eine verbesserte Analyse- und Prognosequalität dazu beitragen, die künftige Ertragskraft von Unternehmen präziser einzuschätzen und Fehlverhalten des Managements besser aufzudecken.“	4 & 7	
Mustererkennung	[Br18, S. 10] "Smart technologies can detect unusual transactions, relations and networks that are not apparent to human supervisors."	4	Durch BD/MI/ML können Muster erkannt werden, welche durch den Menschen nicht entdeckt werden können. Dies erlaubt eine Repriorisierung von Aufgaben und ggf. die Anpassung von Prüfungsplänen.
	[FS20, S. 8] "[...] the use of some AI/ML applications may identify patterns in data that may not be apparent to human review."	4	
	[Ko18, S. 300] „Diese Analysen lassen Muster in für die Abschlussprüfung bedeutsamen Daten erkennen und können prüfungsrelevante Abweichungen bzw. Inkonsistenzen aufdecken.“	4	
	[Ma20, S. 1332] „Das Herausarbeiten aussagekräftiger Muster und Tendenzen aus großen Datenmengen ist für die Risikobeurteilung in der Abschlussprüfung von großem Nutzen.“	2 & 4	
Outlier-Entdeckung	[Hu17, S. 9] "And our Market Oversight team can clarify the normal behaviour of traders and detect deviations that might flag insider trading."	4	Durch BD/MI/ML können Outlier-Datenpunkte schneller erkannt werden.
	[Ma20, S. 1338] „Schnelle Erkennung von Outlier-Datenpunkten“	2	
Entscheidungsunterstützung	[Br18, S. 3] "Risk indicator dashboards, centralised data warehouse for supervisory reports and early warning systems are just a few examples of tools that are now entrenched in a number of supervisory agencies around the world. Suptech offers the potential to either radically improve these tools or develop considerably better ones. "	4 & 6	BD/MI/ML unterstützt die Prüfer in ihrer Entscheidungsfindung im Rahmen von aufsichtlichen Maßnahmen, gezielten Prüfungshandlungen, ebenso wie bei der Risikoeinschätzung. Ggf. können durch andere und bessere "Tools" Prüfungsprozesse transformiert werden.
	[FS20, S. 8] "The potential benefits of SupTech include: [...]. Enhanced decision making of supervisory measures; [...]."	4	
	[Ko18, S. 302] „Bei „Predictive Analytics“ geht es u.a. darum, Ermessensentscheidungen im Bereich von Schätzwerten in der Rechnungslegung durch den Einsatz von „Big Data“-basierten alternativen Assoziations- oder Regressionsanalysen durch den Prüfer zu hinterfragen.“	2 & 4	
	[Ko18, S. 302] „Die [...] Nutzung von „Big Data“ durch innovative Datenanalysen führt zu verbesserten Risikoeinschätzungen und gezielteren Prüfungshandlungen.“	2 & 6	

Echtzeitprüfung/Continuous Auditing (CA)	[FS20, S. 8] "SupTech applications that leverage AI/ML models, may improve surveillance and assessment of risk in real time while also proving predictive analysis. Such timely forward-looking monitoring may allow authorities early insights into risk factors that might threaten financial stability."	4 & 6	Durch BD/MI/ML können prüfungsrelevante Unternehmensdaten jederzeit überprüft werden. Dies kann zu einer frühzeitigen Erkennung von Risiken führen oder gar dazu, dass Anreize zu Regelkonformität verstärkt werden. Möglicherweise können Prüfer und Aufseher ihr Mandat besser erfüllen, was gesellschaftlich positive Effekte mit sich bringt.
	[Ma20, S. 1333] „Durch die vollständige Aufzeichnung aller Geschäftsvorfälle können Verstöße gegen Normen sowie unternehmensinterne Richtlinien systemseitig identifiziert und dem Prüfer gemeldet werden. [...] Aufgrund der kontinuierlichen Anwendung eines KI-basierten CA ist eine schnelle Reaktionszeit möglich. [...]“	2	
	[Ve18, S. 2582] „Durch die konsequente Verfolgung des Real Time Accounting könnte die Entscheidungsnützlichkeit der Finanzberichterstattung infolge der Bereitstellung aktueller Unternehmensinformationen signifikant erhöht werden. Es wird auch angenommen, dass die Anreize des Managements zu einer aggressiven Rechnungslegungspolitik und mithin das Prüfungsrisiko des Abschlussprüfers reduziert werden könnten.“	10	
Kostensenkung	[Br18, S. 16] "Enhanced effectiveness, reduced costs and improved capabilities are the most often cited motivations for developing suptech applications."	2 & 4	Durch BD/MI/ML können Kosten seitens der Aufseher und Wirtschaftsprüfungsgesellschaften sowie seitens der Institute und Unternehmen reduziert werden, was zu nachhaltiger Mandaterfüllung führt. Auch erzeugt dies einen öffentlichen Nutzen (gesellschaftliche Ebene des Wertbeitrags).
	[Br18, S. 17] "Suptech applications reduce costs by automating processes that typically used to involve several people. Cost efficiency certainly motivates solutions related to data collection. [...] Data analyses, particularly those that involve a significant amount of data, also benefit from cost reduction since these now require less staff time. At one supervisory agency, an AML solution is estimated to have reduced costs by 80%."	2 & 8	
	[FC18, S. 5] "As outlined in our Mission, we aim to regulate efficiently and cost-effectively, using our finite resources to deliver the greatest public value. Regulation that involves high costs and inefficient processes can inhibit competition and mean firms pass higher costs on the customers. By using technologies specifically designed to overcome the regulatory challenges faced by firms we can add the public value we deliver as a regulator."	2 & 8	
	[Ma20, S. 1338] „Zeit- und Kosteneinsparungen durch automatisierte Prüfungshandlungen“	2	

Abb. 4 Ergebnisse des Literaturreviews zu Wertbeiträge in Prüfungs- und Aufsichtsprozessen

Im Folgenden werden wesentliche Ergebnisse aus der Literatur mit Blick auf die Wertbeitragsdimensionen dargestellt und diskutiert. Allerdings ist anzumerken, dass nicht in jedem Fall eine überschneidungsfreie bzw. eindeutige Zuordnung von Wirkungen gelingt.

- **Transaktionsbezogene Wertbeiträge (F-I-2)** adressieren die Effizienz der Organisation bzw. von Prozessen. Dies beinhaltet unter anderen die Steigerung der Mitarbeiterproduktivität. Im Falle von Prüfungs- und Aufsichtsprozessen beziehen sich Wertbeiträge darauf, dass die herkömmlichen Tätigkeiten bei BD-/KI-/ML-Unterstützung schneller, kostengünstiger und/oder mit einer höheren Qualität und Verlässlichkeit ausgeführt werden. So können dieselben Daten schneller erfasst, verarbeitet und analysiert werden. Ebenfalls können mehr Daten – neben strukturierten auch unstrukturierte sowie Daten aus unterschiedlichen Quellen – (in der vorhandenen Zeit oder ebenfalls schneller) erfasst, kombiniert, verarbeitet und analysiert werden. Die Beschleunigung der Erfassung und Bereitstellung von Daten kann dabei Auswirkungen auf die Rechtzeitigkeit von Entscheidungen haben und so mittelbar einen Wertbeitrag bedeuten.
- **Informations- und analysebezogenen Wertbeiträge (F-I-4)** zielen auf die Verbesserung der Informationsversorgung, Erkenntnisgewinnung und Entscheidungsfindung; darüber hinaus auf die Verbesserung von Automatisierung, Qualität, Transparenz und Methodik von Analysen sowie eine verbesserte Dokumentation der Analyseergebnisse und Analyseverfahren. In Prüfungs- und Aufsichtsprozessen können durch BD/KI/ML Daten unterschiedlichen Typs und unterschiedlicher Herkunft für Analysen und Entscheidungen herangezogen werden; so werden Analysen auf komplexeren und heterogenen Datenbestän-

den ermöglicht und sie lassen sich zu präziseren und fundierteren Urteilen verdichten. Explizit erwähnt wird, dass Maschinen in großen Datenmengen Muster, Ausreißer und andere Auffälligkeiten erkennen können und dabei möglicherweise den Analysekapazitäten von menschlichen Bearbeitern überlegen sind. In den Literaturquellen wird des Weiteren deutlich, dass bessere Analysefähigkeiten flexiblere Reaktionen in Prüfungsprozessen erlauben und dass ggf. Prüfungsprozesse daher dauerhaft angepasst und geändert (transformiert) werden können (siehe auch unten F-I-6).

- Strategische Wertbeiträge (F-I-5) sind auf die Unterstützung bei der Umsetzung der langfristigen Unternehmensziele gerichtet. Hierbei spielen die (IT-)Strategien und die schnelle Reaktion auf Veränderungen des Umfeldes eine entscheidende Rolle. Solche Wertbeiträge, wie auch solche, die sich aus der Erzielung von Wettbewerbsvorteilen ergeben, sind mit Blick auf Prüfungs- und Aufsichtsprozesse in der herangezogenen Literatur nicht direkt zu identifizieren. Dies mag daran liegen, dass zumindest die öffentliche Aufsicht sich nicht im Wettbewerb befindet – entsprechend Strategien in diesem Bereich eine untergeordnete Rolle spielen.
- Transformationsbezogene Wertbeiträge (F-I-6) zielen auf die Ermöglichung von organisatorischem Wandel und Innovation ab. Dazu gehört im Allgemeinen die Entwicklung neuer Produkte, Geschäftsfelder und Prozesse. Mit Blick auf Prüfungs- und Aufsichtsprozesse bezieht sich dies darauf, dass sich die Prozesse infolge des Einsatzes von BD, KI und ML wandeln und/oder „transformiert“ werden können. In den Literaturquellen wird angeführt, dass Prüfungshandlungen automatisiert werden können, was eine Prozesstransformation bedeuten kann. Ebenfalls werden als Effekte bessere Risikoeinschätzung sowie frühere und flexiblere Reaktionen in Prüfungsprozessen genannt, was dazu führen kann, dass standardisierte Vorgehensweise und Prüfungsprozesse dauerhaft transformiert werden können, wenn entsprechende BD/KI/ML-Werkzeuge regelmäßig und zweckgerichtet zum Einsatz kommen. Als Prozesstransformation kann auch der mögliche Übergang von Stichprobenprüfungen zu Vollprüfungen angesehen werden. Auch durch die Möglichkeit zu Echtzeitprüfungen wandeln sich Prüfungs- und Aufsichtsprozesse (bzw. haben sich in der Vergangenheit schon gewandelt).
- Wertsichernde Wertbeiträge (F-I-7) zielen darauf ab, negative Zielbeiträge, also Risiken, einzuschränken, zu vermeiden oder besser beherrschbar zu machen. Dies gelingt durch traditionelle Risikoanalyse und die gezielte Steuerung der Risiken. Ebenso durch die Erfüllung von Compliance-Anforderungen ([Ba15], S. 203ff). Es ist davon auszugehen, dass die Möglichkeit zur Vollprüfung sowie der Einsatz von BD/KI/ML-Werkzeugen in verschiedenen Prozessschritten relevante Risiken – bspw. das Entdeckungsrisiko oder das Prüfungsrisiko – senken kann. Dies kann auch für Rechtzeitigkeit/frühere Informationsbereitstellung angenommen werden.
- Nachhaltigkeitsbezogene Wertbeiträge (F-I-8) beziehen sich u.a. auf die Reduktion von Ressourcenverbräuchen. Dies kann zum Beispiel bedeuten, dass weniger Speicherplatz benötigt wird, da doppelte Datenhaltung nicht mehr notwendig ist oder dass eine Tätigkeit, welche zuvor manuell oder papierhaft durchgeführt wurde, nun komplett elektronisch durchgeführt werden kann, wodurch sich in Prüfungs- und Aufsichtsprozessen

Ressourcen einsparen oder anderweitig wertstiftend verwenden lassen. Während anderweitige Verwendung zu einer Verbesserung der Prüfungsergebnisse führen kann, können eingesparte Ressourcen, dann, wenn es sich um die öffentliche Aufsicht handelt, bewirken, dass öffentliche Finanzen entlastet oder Abgaben und Steuern an anderer Stelle Gewinn bringend eingesetzt werden.

- Unbeabsichtigte Konsequenzen/negative Wirkungen (F-I-9) wurden in den herangezogenen Literaturquellen nicht thematisiert. Allerdings handelt es sich bei diesen auch nicht im eigentlichen Sinne um Wertbeiträge – vielmehr ist Vermeidung eine Wertsicherung im zuvor beschriebenen Sinne (F-I-7). Die genannten Herausforderungen sind insofern von Prüfungs- und Aufsichtsinstanzen zu beachten, als dass sie Prüfungsurteile verzerren und Reputationsschäden sowie sinkendes Vertrauen bewirken können.
- Symbolic Value wird in den identifizierten Literaturquellen an mehreren Stellen angesprochen. Dass direkte Reputations- und Imagesteigerung durch die Nutzung von BD/KI/ML angestrebt wird, ist zwar denkbar; in den herangezogenen Quellen wurden diesbezüglich jedoch keine Hinweise gefunden. In anderer Hinsicht ist das Potenzial jedoch erkennbar: Mit Blick auf Echtzeitprüfungen wird zum Beispiel angeführt, dass als Reaktion auf die Nutzung von BD/KI/ML durch Prüfungs- und Aufsichtsinstanzen Fehlverhalten erschwert und damit wahrscheinlich reduziert wird. In diesem Zusammenhang hat die Nutzung entsprechender Technologien (oder die Kommunikation der Nutzung) einen (abschreckenden) Signaling-Effekt mit Blick auf relevante Player in der Finanzindustrie, was als eigenständiger Wertbeitrag angesehen werden kann. Darüber hinaus wird mit „public value“ und der verbesserten Überwachung von Risiken im Finanzsystem (Finanzstabilität) auch die gesellschaftliche Dimension angesprochen.

Anzumerken ist, dass keine Wertbeiträge bezüglich der (Daten-)Infrastruktur (F-I-1) in den identifizierten Literaturquellen gefunden werden konnten.

5 Fazit

Ausgangspunkt des Beitrags war die Betrachtung von Wertbeiträgen, die durch IT erbracht werden und – hieran anknüpfend – welche Wertbeiträge sich durch den Einsatz von Big Data, Künstlicher Intelligenz und Machine Learning ergeben können. Dies ist bezüglich des Hypes sowie Fragestellungen, die Entscheidungsträger mit Blick auf den Einsatz dieser Technologien beantworten müssen, aktuell von besonderer Relevanz. Zurzeit sind viele Organisationen bestrebt, Einsatzszenarien und Anwendungsfälle für den Einsatz von BD/KI/ML zu identifizieren. Dabei können die vorgestellten Bezugsrahmen Hilfestellung bieten, indem sie die Identifikation möglicher Wertbeiträge unterstützen.

Dabei sind die Bezugsrahmen „research in progress“ und als vorläufig anzusehen. Weitere Forschung sollte sich darauf beziehen, sie zum einen zu erweitern und zu konkretisieren und sie – zum anderen – weitergehend zu evaluieren und zu erproben. Die hier durchge-

fürhte, kurze Fall- und Literaturstudie ist ein erster Schritt in dieser Hinsicht. Weitere Fallstudien und auch konzeptionelle Evaluierungen könnten hierauf aufbauen und so zu einer weiteren Fundierung – auch hinsichtlich der Nützlichkeit und Anwendbarkeit – oder Anpassung und Weiterentwicklung beitragen.

Ein weiteres Ziel bestand darin, den Wertbeitrag und Nutzen von BD/KI/ML in Prüfungs- und Aufsichtsprozessen systematisch zu untersuchen und diesbezüglich eine Forschungslücke zu füllen. Allerdings kann in Zweifel gezogen werden, dass die gemeinsame Behandlung von Aufsicht und Prüfung, wie hier vorgenommen, sachgerecht ist. Die gemeinsame Betrachtung erfolgte, um auf einer breiteren Literaturbasis aufbauen zu können. Allerdings könnten die institutionellen Unterschiede und jeweiligen Besonderheiten auch nahelegen, zwischen Aufsicht und Prüfung zu trennen. Dies sollte in zukünftiger Forschung adressiert werden und auch dabei müssten sich die Bezugsrahmen als sinnvoll anwendbar, nützlich und robust erweisen.

Literaturverzeichnis

- [Af15] Afflerbach, P.: The Business Value of IT in Light of Prospect Theory. *Business & Information Systems Engineering*, 5, 2015, S. 299 - 310.
- [Ba15] Bartsch, S.: Ein Referenzmodell zum Wertbeitrag der IT. Wiesbaden, 2015.
- [Br18] Broeders, Dirk; Prenio, Jermy: Innovative technology in financial supervision (suptech) - the experience of early users, 2018.
- [Ch07] Chau, P. Y. K.; Kuan, K. K. Y.; Liang, T.-P.: Research on IT value: what we have done in Asia and Europe. *EJIS* 3, 2007, S. 196–201.
- [De14] Deppert, W.: Individualistische Wirtschaftsethik (IWE). Wiesbaden, 2014.
- [EC19] ECB: Bringing artificial intelligence to banking supervision, Newsletter Article. https://www.bankingsupervision.europa.eu/press/publications/newsletter/2019/html/-ssm.nl191113_4.de.html
- [En21] Enholm, I. M. et al.: Artificial Intelligence and Business Value: a Literature Review, *Information Systems Frontiers*, 2021
- [FC18] FCA: Call for Input: Using technology to achieve smarter regulatory reporting, 2018.
- [Fi04] Fischer, T.; Rothe, A.: Wertbeitrag der Informationstechnologie. In: Moormann, J.; Fischer, T. (Hrsg.): *Handbuch IT in Banken*. Berlin, 2004, S. 19 - 41.
- [FS20] FSB: The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions, 2020.
- [Gr06] Gregor, S. et al.: The transformational dimension in the realization of business value from information technology. *The JSIS* 3/15, 2006, S. 249 - 270.
- [Gr18] Grover, V. et al.: Creating Strategic Business Value from Big Data Analytics: A Research Framework, *JMIS* 2018, Vol. 35, No. 2, S. 388 - 423.

- [Hu17] Hunt, S.: From Maps to Apps: the Power of Machine Learning and Artificial Intelligence for Regulators, 2017.
- [Kl23] Klotz, M.; Goeken, M.; Fröhlich, M.: IT-Governance. Ordnungsrahmen und Handlungsfelder für eine erfolgreiche Steuerung der Unternehmens-IT. Heidelberg 2023.
- [Ko18] Kompenhans, H.; Wermelt, A.: Big Data - Potenziale für innovative Abschlussprüfungen. In: *Betriebs-Berater*, H. 6, 2018, S. 299 - 303.
- [Ma20] Marten, K.-U.: Wirtschaftsprüfung und Künstliche Intelligenz, In: *Die Wirtschaftsprüfung*, H. 22, 2020, S. 1331 - 1340.
- [Mc22] McCaul, E.: The impact of suptech on European banking supervision. 14.9.2022. <https://www.bankingsupervision.europa.eu/press/speeches/date/2022/html/ssm.sp220914~d0201e42a9.en.html>
- [Me04] Melville; Kraemer; Gurbaxani: Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value. *MISQ* 2/28, S. 283, 2004.
- [Mi98] Mirani, R.; Lederer, A. L.: An Instrument for Assessing the Organizational Benefits of IS Projects. *Decision Sciences*, 29, 1998, S. 803 - 838.
- [Mo96] Mooney, J. G.; Gurbaxani, V.; Kraemer, K. L.: A process oriented framework for assessing the business value of information technology. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 2/27, 1996, S. 68 - 81.
- [OP07] Oh, W.; Pinsonneault, A.: On the Assessment of the Strategic Value of Information Technologies: Conceptual and Analytical Approaches. *MISQ* 2/31, 2007.
- [Sc22] Schütte, R. et al.: IT-Systeme wirtschaftlich verstehen und gestalten. Methoden – Paradoxien – Grundsätze. Wiesbaden, 2022.
- [Se21] Seufert, S. et al.: A Literature-based Derivation of a Meta-framework for IT Business Value: Proceedings of the 23rd ICEIS 2021, S. 291 - 302.
- [Se99] Seddon, P. B. et al.: Dimensions of Information Systems Success. *CAIS*, 2, 1999.
- [Sh00] Shang, S.; Seddon, P. B.: A Comprehensive Framework for Classifying the Benefits of ERP Systems. In: *AMCIS 2000 Proceedings*, S. 1005 - 1014.
- [SS00] Shang, S.; Seddon, P. B.: A Comprehensive Framework for Classifying the Benefits of ERP Systems. undefined, 2000.
- [Ta20] Tallon, P. P.; Mooney, J. G.; Duddek, M.: Measuring the Business Value of IT. In: Lynn, T. et al. (Hrsg.): *Measuring Business Value of Cloud Computing*. Cham 2020, S. 1 - 17.
- [Th20] Thommen, J.-P.; Achleitner, A.-K.; Gilbert, D. U.; Hachmeister, D.; Jarchow, S.; Kaiser, G.: *Allgemeine Betriebswirtschaftslehre*. Wiesbaden, 2020.
- [Tö20] Töhönen, H.; Kauppinen, M.; Männistö, T.; Itälä, T.: Evaluating the business value of information technology: Case study on game management system. *IJAIS*, 36, 2020.
- [Ve18] Velte, P.; Drews, P.: Herausforderungen in der Wirtschaftsprüfung durch den Einsatz von BD-Technologien, In: *Der Betrieb*, H. 43, 2018, S. 2581–2585.

Agile Strategieentwicklung

State-of-the-Art

Betül Demirözer¹, Helmut Beckmann²

Abstract: Die zunehmende Digitalisierung bewirkt eine Disruption ganzer Branchen, die im Strategieprozess nicht adäquat adressiert werden. Im Fokus steht daher die unternehmerische Perspektive der agilen Strategieentwicklung und die Notwendigkeit, nicht nur alltägliche operative Entscheidungen, sondern auch langfristige strategische Entscheidungen zu überdenken. Darüber hinaus hat die fortschreitende VUCA-Welt zwangsläufig Auswirkungen auf die Strategieentwicklung, da die meisten Organisationen nur selten auf eine VUCA-Welt vorbereitet sind. In diesem Zusammenhang setzt die agile Strategieentwicklung wichtige Impulse. Agile Strategieprozesse und agile Methoden unterstützen den Einsatz der agilen Strategieentwicklung. Von einer methodischen Perspektive aus betrachtet, ist SWOT3 ein wirksames Instrument, um mit Herausforderungen in der VUCA -Welt umzugehen und gleichzeitig die Anforderungen der agilen Strategie zu erfüllen.

Keywords: Agile Strategieentwicklung, Agile Strategie, State-of-the-Art, Agile Strategieprozesse, Agile Strategy, Agile Strategieentwicklungsmethoden

1 Einleitung

Die Welt wird komplexer und dynamischer. Organisationen müssen bestehende Strategien, Strukturen und Prozesse häufiger sowie schneller anpassen und modifizieren [Ba19]. In einer VUCA-Welt³ müssen Strategien und ganze Organisationen immer agiler werden. Darüber hinaus erfordert die zunehmende Digitalisierung eine strategische Anpassung für viele Unternehmen [DT19]. Eine solide entwickelte Strategie ist entscheidend für die erfolgreiche Positionierung einer Organisation. Allerdings stößt der klassische Strategieprozess in der VUCA-Welt an seine Grenzen. In der aktuellen Forschung wird vermehrt die Verbindung zwischen der agilen Strategieentwicklung und strategischen Prozessen in einem agilen Umfeld untersucht. Dabei liegt der Fokus darauf, wie strategische Prozesse an die dynamischen Anforderungen des Marktes angepasst werden können [PMP22]. In der Literatur scheint agiles Management besser geeignet zu sein, um die Lebensfähigkeit von Organisationen in einer Welt des Wandels, der Ungewissheit, Komplexität und Ambiguität sicherzustellen. Im Rahmen der Strategieentwicklung ist eine VUCA-kompatible und agile Methode von entscheidender

¹ Hochschule Heilbronn, WIN-MDT, Max-Planck-Str. 39, 74081 Heilbronn, bdemiroe@stud.hs-heilbronn.de,

² Hochschule Heilbronn, WIN-MDT, Max-Planck-Str. 39, 74081 Heilbronn, helmut.beckmann@hs-heilbronn.de,

³ Volatility (Volatilität), Uncertainty (Ungewissheit), Complexity (Komplexität), Ambiguity (Mehrdeutigkeit)

Bedeutung. Sie ermöglicht es, Expertise unabhängig von Hierarchie und inhaltlichem Status zu berücksichtigen und den Prozess aus prozessualer Sicht agiler zu gestalten. Diese Methode erlaubt es einer Organisation, schnell und flexibel auf unerwartete Faktoren zu reagieren, die kurzfristige Auswirkungen auf, die sich verändernde Umwelt haben könnten [DT19]. Bislang bietet die Literatur wenig Unterstützung für die Gestaltung agiler Strategieentwicklung. Der agile Strategieprozess und die Methode der agilen Strategie können Unternehmen dabei unterstützen, agile Strategieentwicklung agiler und strukturierter zu gestalten [DT19, PMP22]. In dieser Arbeit wird der agile Strategieprozess mit dem klassischen Strategieprozess gegenübergestellt und mithilfe einer Literaturanalyse werden die Gestaltung und Methoden zur agilen Strategieentwicklung identifiziert und dargelegt. Daraus ergibt sich die folgende Forschungsfrage: „Wie ist der State-of-the-Art im Forschungsbereich zur Gestaltung der agilen Strategieentwicklung?“

Ziel dieser Arbeit ist es, die Darstellung des aktuellen Forschungsstandes vorzubereiten und die Methodik für die agile Strategieentwicklung mithilfe einer Konzeptmatrix darzulegen. Diese Untersuchung ermöglicht die Identifizierung von Forschungsbedarfen und zukünftigen Trends. Um dieses Ziel zu erreichen, werden zunächst in Kapitel 2 die Begrifflichkeiten und Grundlagen erläutert. Anschließend wird in Kapitel 3 die Methodik zur systematischen Literaturanalyse beschrieben, welche zur Vorbereitung der Darstellung des aktuellen Forschungsstandes dient. In Kapitel 4 werden die Ergebnisse mithilfe einer Konzeptmatrix analysiert, eine Gegenüberstellung zwischen klassischer und agiler Strategieentwicklung gezogen und die agilen Methoden dargelegt. Abschließend werden in Kapitel 5 Schlussfolgerungen getroffen und mögliche zukünftige Forschungsbedarfe diskutiert.

2 Grundlagen und Begrifflichkeiten

In diesem Kapitel werden die Grundbegriffe Agilität, agile Strategieentwicklung, agile Strategieprozesse und VUCA-Welt erläutert. Dadurch soll ein einheitliches Verständnis und ein entsprechender theoretischer Kontext zur Beantwortung der Forschungsfrage geschaffen werden.

2.1 Agilität

Agilität bezeichnet die Geschicklichkeit, Wendigkeit und Mobilität, die sich in Organisationen, Menschen, Strukturen und Prozessen manifestieren können. Sie haben die Flexibilität, auf unvorhergesehene Ereignisse und neue Anforderungen zu reagieren. Wenn es beispielsweise um Veränderungen geht, ist eine Person nicht nur reaktiv, sondern auch proaktiv [Be22]. Im Kontext einer Organisation oder eines Unternehmens wird Agilität als die Fähigkeit einer Organisation beschrieben, die sich kontinuierlich an ihr komplexes, turbulentes und unsicheres Umfeld anpasst. Ähnlich definiert [Ha17] Agilität als die Fähigkeit, sich an alle Arten von Veränderungen anzupassen und Ziele, Inhalte,

Organisation sowie Prozesse zeitnah zusammenzustellen [Mü22]. [AU15] argumentieren, dass agile Unternehmen in der Lage sind, schnell und effizient zu reagieren, Unsicherheiten zu antizipieren und entsprechende proaktive Vorbereitungen zu treffen. In einer Hays-Studie wurden agile Unternehmen als anpassungsfähig, flexibel, selbstorganisierend und initiativ beschrieben [Mü22]. Weiterhin geht das Verständnis von Agilität über diese Merkmale hinaus. Es handelt sich um eine Haltung, ein Mindset und ein Managementansatz, bei dem die Kundenzufriedenheit ständig im Mittelpunkt steht und an dem Organisationen konsequent arbeiten sollten. Basierend auf den Erkenntnissen in diesem Abschnitt, beruht diese Ausarbeitung auf der folgenden Definition von Agilität, als ein Unternehmen, dass sich proaktiv auf Unsicherheiten im Geschäftsumfeld vorbereitet, sich schnell an Veränderungen anpasst und dadurch die Anforderungen der Kunden erfüllt und somit langfristige Zufriedenheit erlangt. Insofern ist Agilität nicht nur eine Methodik, sondern ein Mindset [Mü22].

2.2 Agile Strategieentwicklung

Die Strategie bestimmt grundsätzlich den Kurs, den ein Unternehmen erfolgreich einschlagen will. Bei der agilen Strategieentwicklung sucht das Unternehmen nach geeigneten Lösungsmöglichkeiten, um bestehende und identifizierte Herausforderungen zu bewältigen [F022]. Agile Strategieentwicklung ist ein Ansatz zur Erstellung eines Plans zur Erreichung der langfristigen Ziele eines Unternehmens angesichts des ständigen Wandels. Daher setzt agile Strategieentwicklung mehr auf Schnelligkeit, notwendige Anpassungen und Flexibilität. Weniger Streben nach einer einzigen vermeintlich optimalen Lösung. Der wesentliche Unterschied zwischen der agilen Strategieentwicklung und der klassischen Herangehensweise an die Strategieentwicklung ist das sogenannte „agile Mindset“, also eine agile Haltung und Perspektive sowie ein grundlegend anderes proaktives methodisches Vorgehen. Daher legt die agile Strategieentwicklung ein besonderes Augenmerk darauf, sich den Grundlagen und zukunftsweisenden Richtungen vollständig anzupassen und vor allem schnell handlungsfähig zu sein [F022]. Angesichts sich ändernder Rahmenbedingungen und Anforderungen ist es wichtig, die Aktualität der Strategie regelmäßig zu überprüfen und anzupassen. Darüber hinaus wird in der Regel zunächst nach einer internen und externen Analyse auf Basis des Leitbildes eine Weiterentwicklungsthese verfasst und einfache, pragmatische sowie strategische Eckpunkte definiert. Nachfolgend werden Entwicklungsoptionen verfeinert, konzipiert und in konkrete Schritte umgesetzt. Dies zeigt, welche Schritte erfolgreich waren, welche Veränderungen aufgetreten sind und gibt Aufschluss darüber, welche Bereiche überdacht werden müssen. Diesbezüglich wird die Strategie sukzessive angepasst und zur Überprüfung und Anpassung formuliert [Zo21]. Die entsprechenden Inhalte der agilen Strategieentwicklung lassen sich gemäß [Zo21] in die folgenden dargestellten 5 Grundelemente differenzieren.

- Ein strategisches Leitbild, das mit einigen Kernaussagen den Rahmen der agilen Strategieentwicklung gestaltet

- Der agile Strategieentwicklungs- und Umsetzungsprozess gliedert sich in kurze, rollierende Arbeitsphasen, die strukturiert organisiert und evaluiert werden
- Situative Strategieteams gestalten agile Strategien flexibel und weitgehend selbstbestimmt
- Dabei werden die Erfahrungen und Lerneffekte der agilen Strategie untersucht und die Strategie entsprechend iterativ weiterentwickelt
- Die agile Strategieentwicklung stellt eine Verknüpfung zwischen dem Analyse- und Umsetzungswissen sowie den Fähigkeiten und Erfahrungen der verschiedenen Beteiligten dar

2.3 Agile Strategieprozesse

Der agile Strategieprozess verwendet interdisziplinäre und Iterative Methoden, um Strategien zu formulieren und umzusetzen. Es wird durch kurze Planungszyklen, beständiges Feedback und eine enge Zusammenarbeit zwischen Verbraucher und Arbeitnehmer gekennzeichnet [SE19]. Eine weitere Definition des agilen Strategieprozesses besteht darin, basierend auf den Prinzipien der Agilität zu iterieren und zusammenzuarbeiten und sich schnell an sich ändernde Marktbedingungen anzupassen. Diese Strategie wird als dynamischer Prozess verstanden, der kontinuierlich angepasst und optimiert wird [Sc20]. In der aktuellen Literatur gibt es derzeit keine einheitliche Definition eines agilen Strategieprozesses. Deshalb basiert diese Definition auf kombinierten Beschreibungen und Anforderungen von Agilität und strategischem Management. Alle bekannten Beschreibungen betonen die Reaktionsfähigkeit der strategischen Planung auf Veränderungen der Informationen. Sie wird als strategische Entscheidung definiert, die in der Lage ist, auf Veränderungen der Informationslage zu reagieren. Sie bestehen aus aufeinanderfolgenden iterativen Kombinationen von Informationsanalyse und Entscheidungskonkretisierung [DT19].

2.4 VUCA-Welt als Rahmenbedingung für die Agile Strategieentwicklung

Strategien für alle Arten von Organisationen werden in einem zunehmend komplexen und dynamischen Umfeld entwickelt. Dies hat zwangsläufig Auswirkungen auf die Strategieentwicklung. In der aktuellen Literatur und Forschung wird in diesem Zusammenhang häufig das Akronym VUCA-Welt [DT19] verwendet. "VUCA" steht dabei für Volatilität (Schwankung), ein Begriff, der die Häufigkeit, Geschwindigkeit und das Ausmaß von Veränderungen im Laufe der Zeit beschreibt. Diese Schwankungen werden oft durch unerwartete Ereignisse ausgelöst [Au17]. Uncertainty (Ungewissheit) beschreibt die Entwicklung und den Ausgang zunehmend unvorhersehbarer Ereignisse. Darüber hinaus bedeutet Complexity (Komplexität), dass Organisationen in ihrem Umfeld mit einer nahezu unkontrollierbaren Anzahl von interagierenden sozialen, ökologischen, politischen, regulatorischen, technologischen und ökonomischen Faktoren (Variablen)

konfrontiert sind. Ambiguity (Mehrdeutigkeit) zeigt, dass die entstehende Komplexität erst nach Ablauf der Zeit, also nachträglich, verständlich wird. Die derzeitige Unklarheit führt dazu, dass die Akteure die Situation sehr unterschiedlich einschätzen und daher unterschiedliche Schlussfolgerungen ziehen [DT19].

3 Methodik zur Literaturanalyse

Die Durchführung der Literaturanalyse wird mithilfe zweier Methoden ausgeführt. Zunächst wird die Methode von [Fe06] angewandt, die eine Suchstring-basierte Suche umfasst, die unter Berücksichtigung von Suchzeiträumen durchgeführt wird. Um diese Methode zu ergänzen, wird auch das Verfahren von [WW02] verwendet, welches sich durch eine Vorwärts-Rückwärts-Suche auszeichnet. Anschließend werden alle gefundenen Quellen und Fachartikel in einer eigens erstellten Tabelle gesammelt und anhand ihrer Titel und Abstracts auf Eignung für die Anfertigung der Arbeit beurteilt, ob sie geeignet oder nicht geeignet sind. Für die Literaturlauswertung wird eine Konzeptmatrix von [WW02] verwendet, in der die Themenfelder der Arbeit übersichtlich nach verschiedenen Kriterien wie Forschungsschwerpunkt oder Methoden dargestellt werden. Die mit der Analyse in Zusammenhang stehenden Literaturdatenbanken sind in Tabelle 1 aufgeführt. Die verwendeten Suchbegriffe lauten wie folgt.

- („Agile“) AND („Strategie“) AND („Gestaltung“ OR „Methoden“ OR „Prozess“)
- („Agile Strategieentwicklung“) AND („Forschungsbedarf“ OR „zukünftige Forschung“ OR „Trends“)
- („Agil“) AND („Strategy“) AND („Methodology“ OR „Process“ OR „Development“) AND („State of the art“ OR „Review“)
- („Agile“) AND („Strategy“) AND („Development“ OR „Process“ OR „Design“) AND („Future Research“ OR „Trends“)

#	Name der Datenbank	Abkürzung
L1	AIS Electronic Library	AISel
L2	Google Scholar	GS
L3	IEEE Xplore Digital Library	IEEE
L4	ResearchGate	SD
L5	ScienceDirect	SD
L6	Springer Link	SL

Tab. 1: Literaturdatenbanken

Insgesamt wurden zunächst 423 Artikel als Suchergebnisse in bibliografischen Datenbanken identifiziert (siehe Tabelle 2). Aufgrund der Auswahl nach Relevanz und wissenschaftlicher Anwendbarkeit und Duplizierung beträgt die Anzahl der Beiträge zu dieser Arbeit 32. Der nächste Schritt besteht darin, eine Zusammenfassung dieser 32 Beiträge zu überprüfen. Nach Sichtung der Abstracts betrug die Anzahl der verbleibenden relevanten Artikel 9. Anschließend wurden 2 weitere relevante Artikel durch eine Vorwärts- und Rückwärtssuche nach [WW02] identifiziert. Durch dieses systematische Vorgehen wurden insgesamt 11 relevante Beiträge dokumentiert, die im nächsten Kapitel näher betrachtet und anhand einer Konzeptmatrix analysiert werden.

	Suchtreffer
Gesamtzahl der Suchtreffer innerhalb der Literaturdatenbanken	423
Basierend auf Relevanz und wissenschaftlicher Eignung und Entfernung von Duplikaten	32
Sichtung: Abstract	9
+ Vorwärts- und Rückwärtssuche	2
Summe Beiträge	11

Tab. 2: Suchtreffer

Die Kategorisierung der agilen Strategieentwicklung und Methoden basiert auf selbst definierten Themen und durchgeführter Literaturanalyse. Diese Kategorisierung wird im nächsten Kapitel näher beschrieben und anhand der Konzeptmatrix mit den aus der Literaturanalyse gewonnenen Beiträgen veranschaulicht.

4 Ergebnisse

In diesem Kapitel werden die Erkenntnisse aus der Literaturanalyse detailliert dargestellt und erläutert. Dazu wird eine Konzeptmatrix nach der Methode von [WW02] erstellt, welche die klassischen und agilen Strategieprozesse sowie die agilen Methoden kategorisiert, die für die Gestaltung der agilen Strategieentwicklung berücksichtigt werden. Werden in den untersuchten Publikationen klassische oder agile Strategieprozesse und agile Methoden zur Gestaltung der agilen Strategieentwicklung erwähnt, ist dies mit einem X in der Konzeptmatrix gekennzeichnet. Dabei wird eine systematische Übersicht erstellt, die den aktuellen State-of-the-Art abbildet. Anschließend kann die in der Einleitung vorgestellte Forschungsfrage beantwortet werden. Zu diesem Zweck ist die Konzeptmatrix in Tabelle 3 unten dargestellt.

Nr.	Referenzen	Agile Strategieentwicklung		Methoden		
		Klassische Strategieprozesse	Agile Strategieprozesse	SOWT-Analyse	TOWS-Matrix	SWOT3
1.	[Ba19]		x			
2.	[DT19]		x	x	x	x
3.	[F022]		x			x
4.	[HN10]			x	x	x
5.	[Jo16]	x	x			
6.	[Mü22]		x			
7.	[PMP22]	x	x	x	x	x
8.	[Uh19]		x			x
9.	[Wa19]		x			x
10.	[WTB17]		x			x
11.	[Zo21]		x			
Gesamt		2	10	3	3	7

Tab.3: Konzeptmatrix für die Gestaltung der agilen Strategieentwicklung (In Anlehnung an [WW02])

4.1 Entwicklung einer agilen Strategie durch den agilen Strategieentwicklung

Die Konzeptmatrix in Tabelle 3 zur agilen Strategieentwicklung zeigt, dass die agile Gestaltung zur agilen Strategieentwicklung einen agilen Strategieprozess erfordert. Außerdem wird ein kurzer Gegenüberstellung zwischen klassischen und agilen Strategieprozessen dargelegt.

Agilität und strategisches Management basieren auf zwei grundlegend unterschiedlichen Paradigmen. Während Agilität, Schnelligkeit und Flexibilität erfordert, um sich kontinuierlich an veränderte Gegebenheiten anzupassen, ist die Grundprämisse der strategischen Strategieentwicklung, einen logisch ableitbaren linearen Entscheidungsprozess für die langfristige Unternehmenspositionierung sicherzustellen. Diese beiden gegensätzlichen Paradigmen müssen im Strategieprozess konstruktiv kombiniert werden, um einen agilen Strategieentwicklungsprozess zu erreichen [Jo16]. Zusätzlich helfen die folgenden fünf Prinzipien, den agilen Strategieprozess zu optimieren.



Abb. 1: Optimierung der agilen Strategieprozesses (eigene Darstellung nach [PMP22])

Im ersten Prinzip „Alignment“ ist die Ausrichtung ein Richtungselement des iterativen Prozesses. Dynamische, unbekannte Entwicklungen erfordern Regeln als Leitplanken. Sie verhindern ziellose Versuche einer strategischen Vorgehensweise. Entscheidend für die strategische Handlungswahl sind beispielsweise die Vision und die strategische Ausrichtung einer Organisation. Zudem werden im zweiten Prinzip „Evidenzbasiert“, sorgfältige, explizite und rationale Anwendungen der besten Informationen bei der Entscheidungsfindung getroffen, klare Formulierung einer Strategiethese und Durchführung experimenteller Tests überprüft [PMP22]. Die „Iteration“ geht davon aus, dass der erste Strategieplan einen gewissen Fehler enthält, der durch wiederholte Anpassungen optimiert werden kann [Wa19]. Mehrere Iterationen der Informationsbeschaffung und Entscheidungsfindung tragen dazu bei, strategische Risiken zu reduzieren und die Kreativität zu optimieren. Das vierte Prinzip handelt von der "Kollaboration", die als Grundsatz der Agilität fungiert. Dieses Prinzip geht davon aus, dass Iterationen evidenzbasierter strategischer Entscheidungen nur im Dialog zwischen Wissens- und Entscheidungsträgerinnen und -trägern entstehen können. Das fünfte und letzte Prinzip ist die „Autonomie“, bei der die strategische Arbeit und Entscheidungsfindung von getrennten Teams durchgeführt werden. Die Kombination dieser fünf Prinzipien mit den üblichen Schritten des Strategieentwicklungsprozesses optimiert den agilen Strategieprozess (siehe Abbildung 2) und unterstützt die Gestaltung der agilen Strategieentwicklung [PMP22].

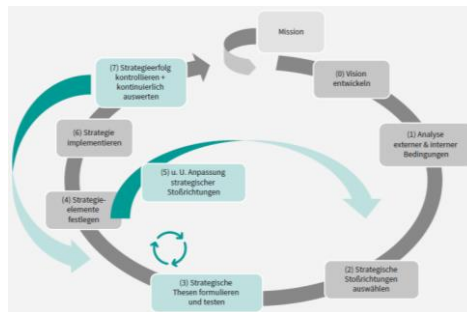


Abb. 2: Agiler Strategieprozess nach [PMP22]

Abbildung 2 skizziert die Schritte im agilen Strategieentwicklungsprozess. Während die Schritte 0-2 dem klassischen Strategieprozess entsprechen und vor allem die Prinzipien des Alignments beinhalten, beschreiben die Schritte 3, 5 und 7 ein agiles Vorgehen mit evidenzbasierter und iterativer Entscheidungskonkretisierung. Die Prinzipien der Kollaboration und Autonomie sind für Schritt 3 oder 4 relevant, je nach Entscheidungskultur des Unternehmens. Schritt 6 wiederum ist dem klassischen Strategieprozess sehr ähnlich, bringt aber bei ausreichendem Wissen aus den Schritten 3 bis 5 die Herausforderung mit sich, vor der Strategie Implementation eine Entscheidung zu treffen [PMP22]. Der agile Strategieprozess hebt sich vom klassischen Strategieprozess durch die zusätzlichen iterativen Phasen ab, in denen Hypothesen aus der strategischen Ausrichtung gezogen und getestet werden. Hier liegt die größte Stärke des agilen Strategieprozesses. Im besten Fall arrangiert ein agiler Strategieprozess, dass Unternehmen besser und schneller reagieren, weil fehlende Informationen schnell festgestellt und Fehleinschätzungen und strategische Ausrichtungen von Entscheidungsgremien revidiert werden können [PMP22].

4.2 Agile Methodik für die agile Strategieentwicklung

Die Konzeptmatrix in Tabelle 3 zeigt 3 Methoden, mit denen eine agile Strategieentwicklung gestaltet werden kann. Die Methoden SWOT-Analyse und TOWS-Matrix werden nicht vertieft, der Fokus liegt auf der Kombination der beiden Methoden, die schließlich eine agile Methode darstellt. SWOT3 erbt die Eigenschaften und die grundlegende Logik der klassischen SWOT-Analyse und der TOWS-Matrix. Aufgrund der Kombination beider Methoden, können die bestehenden Nachteile der einzelnen Methoden, teilweise durch SWOT3 behoben werden. Gründe dafür sind einerseits die Mehrdimensionalität, andererseits die Verwendung höherer, geordneter oder konsistenter Messskalen bei der Bewertung von Umweltfaktoren. Die Agile Methode SWOT3 bildet die vier klassischen SWOT-Elemente Stärken und Schwächen sowie Chancen und Risiken differenziert und verallgemeinert in drei strategische Dimensionen ab, Bedeutung (Impact), Eintrittszeit (Impact Time) und Polung (Polarity). Letzteres spiegelt eine Einordnung zwischen den Polen Chance und Risiko bzw. Stärke und Schwäche wider [DT19]. Dadurch erfüllt SWOT3 die Anforderungen von [HN10]: „Die Zukunft der

SWOT-Analyse scheint darin zu liegen, die Variablen zu ordnen und sie zu gewichten, um den Entscheidungsträgern eine bessere Orientierung zu ermöglichen.“ Um den Diskussions- und Entscheidungsprozess genauer und effektiver zu unterstützen, wird in Abbildung 3 ein auf einem Koordinatensystem basierendes Diagramm visualisiert. Diese hat folgende vier Attribute: Das erste Attribut ist die vertikale Achse, von +5 (Umweltfaktoren sind nur Chancen) bis -5 (Umweltfaktoren sind nur Risiken). Als zweite Achse folgt die horizontale Achse, die für die relative Einordnung aller identifizierten Umweltfaktoren in ein Zeitkontinuum mit fünf Orientierungspunkten von der „Gegenwart“ bis zur „fernen Zukunft“ zuständig ist. Das dritte Attribut ist die Bewertung, die die spezifische Bewertung des Umweltfaktors durch die Organisation auf einer fünfstufigen Skala beschreibt, die durch die Größe des Faktors (oder Kreises) dargestellt wird. Das vierte und letzte Attribut sind darüber hinaus die TOWS-Verknüpfungen, die gleichzeitige Verknüpfung wichtiger Umweltfaktoren mit faktorbezogenen Stärken und Schwächen der Organisation zu strategischen Handlungsfeldern ermöglicht [PMP22].

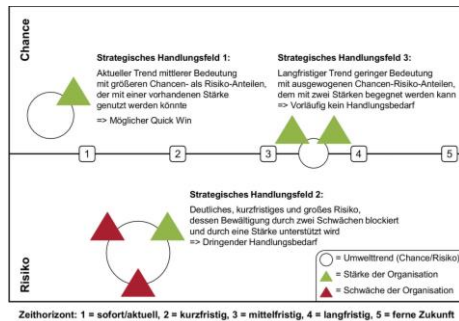


Abb. 3: SWOT3 als Methode der agilen Strategieentwicklung [DT19]

Die dreidimensional erfassten Umfeldfaktoren führen zu sogenannten strategischen Handlungsfeldern, denen organisatorische Attribute zugeordnet werden. Es geht um den Anteil an Chancen, durch den Vorteile genutzt werden können, bzw. welche Schwächen die Organisation überwinden muss, damit dieser Chancenanteil genutzt werden kann. Gleiches gilt für Risikoanteil [DT19].

[WTB17] kommt zu dem Schluss, dass dieser textuelle Zwischenschritt das Auffinden und Formulieren der folgenden Ziele erleichtert. Strategische Ziele kommen nicht zwangsläufig aus jedem strategischen Handlungsfeld. In manchen Fällen gingen mehrere Ziele aus einem einzigen Handlungsfeld hervor, in anderen Fällen wurde ein einziges Ziel in einer Kombination mehrerer Handlungsfelder präsentiert, in anderen Fällen wiesen die Handlungsfelder keinen vorrangigen strategischen Handlungsbedarf auf. Anstatt zu abstrakte und realistische Standardstrategien abzuleiten, wie es in TOWS-Matrizen üblich ist, wird eine für alle Beteiligten leicht verständliche Brücke von der Strategieanalyse zur Strategieformulierung geschlagen. Das strategische Handlungsfeld ist ein wichtiger Zwischenschritt, der die Komplexität der Analyse der VUCA-Welt widerspiegelt, die sich aus einer Vielzahl entsprechender Umfeldfaktoren (Chancen und Risiken) und Organisationsmerkmalen (Stärken und Schwächen) ergibt [DT19].

5 Diskussion und Ausblick

In dieser Arbeit wurden anhand einer systematischen Literaturanalyse von [Fe06] und einer Vorwärts- und Rückwärtssuche von [WW02] 11 relevante Beiträge zur Gestaltung agiler Strategieentwicklung identifiziert. Die 11 identifizierten Beiträge beinhalten Gestaltungsmöglichkeiten zur agilen Strategieentwicklung, die anhand verschiedener Themen kategorisiert wurden. Generell lässt sich sagen, dass agile Strategieentwicklung kein Werkzeug ist, das man nebenbei nutzen kann. Es erfordert eine entsprechend agile Haltung und ein offenes Verständnis für die Bereitschaft zur Veränderung. Agile Strategieentwicklung ist nur wirksam, wenn agile Prinzipien verinnerlicht, Prozesse entsprechend agil optimiert und konsequent umgesetzt werden. Grundlagen sind eine Unternehmenskultur, agile Strategieprozesse und agile Methoden. Bei näherer Betrachtung können agile Strategieprozesse manchmal komplexer erscheinen als klassische Strategieprozesse. Das Hin- und Herwechseln zwischen dem Erwerb von Wissen und der Verkörperung von Entscheidungen verzögert den Abschluss von Entscheidungen, und das Testen strategischer Hypothesen kann zeitaufwändig und mühsam sein. Andererseits hat es Vorteile, strategische Strategiefelder schrittweise zu erkunden, anstatt Entscheidungen nur auf der Grundlage von Hypothesen zu treffen [PMP22]. Aus methodischer Sicht stellt SWOT3 ein praktikables Werkzeug dar, welches die Herausforderungen der VUCA-Welt adressieren kann und gleichermaßen den Anforderungen einer Gestaltung der agilen Strategieentwicklung gerecht wird. Aufgrund seiner Vielseitigkeit scheint SWOT3 auch auf jede Art von Organisation und Strategie anwendbar zu sein [DT19]. Es ist denkbar, dass in der zukünftigen Forschung weitere methodische Anpassungen und Erweiterungen der Agilen SWOT3-Methodik vorgenommen werden. Darüber hinaus haben agile Prozesse in der agilen Strategieentwicklung noch Schwierigkeiten, klare Schlussfolgerungen für den Strategieentwicklungsprozess zu ziehen. Daher gilt es zu klären, in welche strategische Richtung agile Methoden passen. Vorrangig sind Anweisungen und Fragen zu stellen, die nicht ausreichend analysiert wurden.

Literaturverzeichnis

- [AU15] Anderson, K., Uhlig, J.: Das agile Unternehmen-Wie Organisationen sich neu erfinden. Campus Verlag GmbH, Frankfurt am Main, 2015.
- [Au17] Aulinger, A.: Die drei Säulen agiler Organisationen. Whitepaper IOM Steinbeis-Hochschule, Berlin: Steinbeis-Hochschule, 2017.
- [Ba19] Barthel (Hrsg.), C. (2019): Managementmoden in der Verwaltung: Sinn und Unsinn. Springer Fachmedien, Wiesbaden, 2019.
- [Be22] Bendel, Prof. Dr. O: Wirtschaftslexikon.gabler.de, Agilität Definition-Was ist "Agilität"? unter: <https://wirtschaftslexikon.gabler.de/definition/agilitaet-99882/version-368852> (abgerufen am 09.12.2022).

- [DT19] Dahm, M. H., Thode (Hrsg.). S. T.: Strategie und Transformation im digitalen Zeitalter: Inspirationen für Management und Leadership. Springer Fachmedien, Wiesbaden, 2019.
- [Fe06] Fettke, P.: “State-of-the-Art des State-of-the-Art: Eine Untersuchung Der Forschungsmethode ‘Review’ Innerhalb Der Wirtschaftsinformatik. *Wirtschaftsinformatik* 48 (4): 257–66. <https://doi.org/10.1007/s11576-006-0057-3>, 2006.
- [F022] Foerster, B.: marketinginstitut.biz, Agile Strategieentwicklung – flexibel und handlungsschnell zum Ziel, unter: <https://www.marketinginstitut.biz/blog/agile-strategieentwicklung/> (abgerufen am 12.12.2022).
- [Ha17] Hanschke, I.: Agilität - was bedeutet das? Agile Strategien leben und umsetzen. Springer Gabler, Wiesbaden, 2017.
- [HN10] Helms, M. M., & Nixon, J.: Exploring SWOT analysis – where are we now? A review of academic research from the last decade. *Journal of Strategy and Management*. <https://doi.org/10.1108/17554251011064837>, 2010.
- [Jo16] Johnson, G., Whittington, R., Scholes, K., Angwin, D., Regner, P.: Strategisches Management – Eine Einführung. 10. akt. Aufl., Pearson Deutschland GmbH: Hallbergmoos 2016.
- [Mü22] Müller, K.: Entwicklung eines High Performance Learning Journey Konzepts zur organisationalen Weiterbildung agiler Rollen im Rahmen von SAFE: Spezifiziert am Beispiel des Scrum Masters. Springer Fachmedien, Wiesbaden, 2022.
- [PMP22] Pichel, K., Müller, A., Pödlert, C.: Strategility: Vorschläge zur Gestaltung agiler Strategieprozesse. Schäffer-Poeschel. In: zfo. S. 42-50. <https://digitalcollection.zhaw.ch/handle/11475/24911>, 2022.
- [SE19] Schwab, P., & Eichhorn, M.: Agile Strategieentwicklung: Wie Führungskräfte agile Methoden nutzen können. Springer-Verlag, 2019.
- [Sc20] Schwuchow, K., Müller, J., Schumacher, T., & Sadowski, D.: Agilität in der Strategieentwicklung: Konzeptualisierung und empirische Überprüfung. *Zeitschrift für Betriebswirtschaft*, 90(1), 47-81, 2020.
- [Uh19] Uhl, A., Loretan (Hrsg.), S.: Digitalisierung in der Praxis: So schaffen KMU den Weg in die Zukunft. Springer Fachmedien, Wiesbaden, 2019.
- [Wa19] Walter, S. (2019). Strategie Design. Ein ganzheitliches Strategieverständnis für das digitale Zeitalter. Springer Fachmedien, Wiesbaden, 2019.
- [WW02] Webster J, Watson RT.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 26(2): xiii–xxiii, 2002.
- [WTB17] Wistuba, L., Thode, S., & Balzer, Y.: Die Strategie 2020 der Polizei Niedersachsen – Neue Wege aus der Ökonomisierungsklemme. In J. Stierle, D. Wehe & H. Siller (Hrsg.), *Handbuch Polizeimanagement* (S. 835–866). Springer Gabler, Wiesbaden, 2017.
- [Zo21] Zornek, W.: Agile Strategieumsetzung: Wirkungsvoll führen durch aktives Selbstmanagement. Haufe Group, Freiburg, 2021.

Digital Business Technology Platforms

Anforderungen und Kriterien

Vanessa Wegner¹, Helmut Beckmann²

Abstract: In der heutigen, sich schnell entwickelnden Zeit, in denen die verfügbaren Datenmengen exponentiell steigen, ist es für Unternehmen wichtiger denn je, diese Daten effizient zu nutzen. Um dies zu erreichen, wird eine Plattform benötigt, welche die Technologien geschickt auswählt, aufnimmt, aufeinander abstimmt und für den Business Prozess verfügbar macht. Eine Digital Business Technology Platform muss eine Reihe von Kriterien und Anforderungen erfüllen, um Unternehmen bei ihrer digitalen Transformation unterstützen zu können. In der geplanten Forschungsarbeit soll der State-of-the-Art der Digital Business Technology Platforms dargestellt werden. Eine umfangreiche Literaturrecherche nach Fettke, sowie eine Konzeptmatrix nach Webster und Watso wird für das Forschungsdesign ausgewählt. Diese Literaturanalyse untersucht die Kriterien und Anforderungen, die eine Digital Business Technology Platform erfüllen muss, um Unternehmen bei ihrer digitalen Transformation erfolgreich zu unterstützen. Die Ergebnisse zeigen, dass die Wahl einer geeigneten Plattform eine wichtige Entscheidung für den Erfolg des Unternehmens darstellt und sorgfältig überdacht werden sollte.

Keywords: Digital Business Technology Platform, Capabilities, Requirements, State of the Art, Literature Review

1 Einleitung

Die digitale Transformation hat in den letzten Jahren einen starken Einfluss auf die Art und Weise, wie Unternehmen ihre Geschäftsprozesse gestalten und durchführen [DT18, S. 11f.]. Mit der zunehmenden Verbreitung von digitalen Technologien werden Unternehmen gezwungen, ihre Arbeitsabläufe und Geschäftspraktiken anzupassen, um auf den sich ständig verändernden Markt zu reagieren und erfolgreich zu bleiben. Diese Herausforderungen können durch die Verwendung einer Digital Business Technology Platform überwunden werden, welche eine wertvolle Unterstützung für Unternehmen darstellt [HCB16, S. 7f.].

Eine Digital Business Technology Platform bietet eine Vielzahl von Vorteilen, darunter die Automatisierung von Geschäftsprozessen, die Optimierung von Arbeitsabläufen und

¹ Hochschule Heilbronn, Institut für Wirtschaftsinformatik, Max-Planck-Str. 39, 74081 Heilbronn, vwegner@stud.hs-heilbronn.de,

² Hochschule Heilbronn, Institut für Wirtschaftsinformatik, Max-Planck-Str. 39, 74081 Heilbronn, helmut.beckmann@hs-heilbronn.de

die Überwachung der Geschäftsleistung als Funktionalitäten innerhalb einer Plattform anstatt verteilt über eine Anzahl unterschiedlicher Informationssysteme. Dadurch können Unternehmen Zeit und Ressourcen sparen, ihre Effizienz erhöhen und sich auf andere wichtige Geschäftsthemen konzentrieren [Ca22]. Es ist jedoch wichtig, eine Plattform auszuwählen, die den Anforderungen und Kriterien des Unternehmens entspricht, um eine erfolgreiche digitale Transformation zu gewährleisten. Eine gründliche Überprüfung und Auswahl der richtigen Digital Business Technology Platform ist daher von entscheidender Bedeutung für die erfolgreiche Durchführung einer digitalen Transformation. Dieses Paper befasst sich mit den wichtigsten Anforderungen und Kriterien, die bei der Wahl einer geeigneten Digital Business Technology Platform zu berücksichtigen sind, um Unternehmen bei ihrer Entscheidungsfindung zu unterstützen und den aktuellen Stand des Wissens im Bereich der Digital Business Technology Platforms darzustellen.

1.1 Problemstellung

Neben dem Internet existieren mittlerweile viele weitere Datenquellen wie zum Beispiel das Internet of Things, welche permanent Daten generieren. Durch die Fülle an Daten ist der Bedarf entstanden weitere Informationstechnologien und Geschäftsinnovationen zu entwickeln, um ein besseres Verständnis für den Wert der hierbei generierten Daten herzustellen [Li19, S.15-24]. In der heutigen Zeit dienen Plattformen dazu, verschiedene Ressourcen an einem gemeinsamen Ort zu bündeln. Datenplattformen unterstützen die Speicherung, Verarbeitung und Analyse verschiedener Datenquellen und Datentypen für gemeinsame Zwecke. Mit der fortschreitenden Digitalisierung von Unternehmen ist die Digitalisierung durch Digital Business Technology Platforms in den Fokus gerückt [Ci14, S. 3ff.]. Aufgrund der steigenden Komplexität und Menge der zu analysierenden Daten ist es vorteilhaft, Digital Business Technology Platforms nach den jeweiligen Anforderungen der Nutzer zu entwickeln. Um die Möglichkeiten und Vorteile, die sich aus diesen Datenmengen ergeben zu nutzen, ist der Einsatz einer geeigneten Digital Business Technology Platform, welche als Grundlage für ein intelligentes und nachhaltiges Unternehmen dienen kann, notwendig. Folglich bietet diese Lösung eine einheitliche Umgebung zur Bündelung von Daten, Analysen, künstlicher Intelligenz, Anwendungsentwicklung und Automatisierung [SA22]. Obwohl es schon große Digital Business Technology Platforms gibt, ist es wichtig die aktuellen Anforderungen und Fähigkeiten einer solchen Plattform zu definieren.

1.2 Forschungsfrage

Ausgehend von der aktuellen Relevanz des Einsatzes von Digital Business Technology Platforms soll in der vorliegenden Forschungsarbeit der aktuelle Forschungsstand dargelegt und um zukünftige Forschungsmöglichkeiten ergänzt werden. Hierzu ergibt sich für den vorliegenden Beitrag die folgenden Fragestellungen:

„Wie ist der State-of-the-Art im Forschungsbereich der Digital Business Technology Platforms?“

Unterfragen:

- Q1: Welche Anforderungen und Kriterien sind für die Auswahl einer geeigneten Digital Business Technology Plattform von Bedeutung?
- Q2: Welche Anforderungen und Kriterien werden am häufigsten erwähnt?

1.3 Zielsetzung

Die Zielstellung dieser Forschung ist es, den aktuellen Stand der Forschung im Bereich der Digital Business Technology Platforms zu bestimmen und die wichtigsten Anforderungen und Kriterien zu identifizieren, die bei der Wahl einer geeigneten Plattform zu berücksichtigen sind.

2 Grundlagen und Begriffe

Gemäß den Empfehlungen von. [vSS09] werden in den nachfolgenden Abschnitten die wesentlichen Begriffe des Forschungsvorhabens erläutert und definiert. Dies soll dabei helfen ein gemeinsames Verständnis zu erlangen.

2.1 Digitale Plattform

Digitale Plattformen können als elektronische Systeme definiert werden, die es Benutzern ermöglichen, miteinander zu interagieren, Inhalte zu produzieren und zu konsumieren sowie Geschäftstransaktionen durchzuführen. Diese Plattformen stellen eine digitale Infrastruktur bereit, die es Benutzern ermöglicht, sich über ein Netzwerk zu vernetzen und Interaktionen und Transaktionen auf einer gemeinsamen digitalen Schnittstelle auszuführen. Einige digitale Plattformen sind als offene Plattformen konzipiert, bei denen jeder Zugang und Interaktion hat, während andere geschlossene Plattformen sind, bei denen der Zugang und die Interaktionen beschränkt sind [HBK18, S. 181ff.]. Zusätzlich können digitale Plattformen als Marktplätze verstanden werden, auf denen Anbieter und Kunden zusammenkommen, um ihre Bedürfnisse und Angebote auf einer gemeinsamen digitalen Schnittstelle abzustimmen. Auf diesen Plattformen können Anbieter ihre digitalen Produkte, Dienstleistungen und Anwendungen auf einfache und effiziente Weise bereitstellen, während Kunden Zugang zu einer breiten Palette digitaler Angebote haben [HBK18, S. 187f.]. Digitale Plattformen gibt es in vielen verschiedenen Formen und Funktionen [HBK18, S. 183f.]. Einige der bekanntesten Beispiele sind E-Commerce-Plattformen wie Amazon und eBay, soziale Netzwerke wie Facebook und Twitter, Videoplattformen wie YouTube und Vimeo, Musik-Streaming-Dienste wie Spotify und

Apple Music, Online-Marketingplattformen wie Google AdWords und DoubleClick sowie Unternehmenssoftware wie Salesforce und Microsoft Dynamics [Le21, S. 39ff.].

2.2 Digital Business Technology Platform

Eine Digital Business Technology Platform (DBTP) ist eine Sammlung von Technologiefähigkeiten und Komponenten, die zusammenarbeiten, um eine integrierte Plattform zu bilden [Iv19, S. 77]. Diese Plattform bietet eine zentralisierte und standardisierte Basis für Geschäftstransaktionen und Prozesse, Datenformatierung und Analyse sowie die Fähigkeit, Echtzeitreaktionen auf Geschäftsereignisse durchzuführen. Sie dienen als das "Nervensystem" eines modernen Unternehmens und ermöglichen eine optimierte Datenverarbeitung und Überwachung von Geschäftsereignissen [Iv19, S. 77]. Eine wichtige Komponente von DBTPs ist die Datenanalyse, die eine effektive Steuerung von Geschäftsabläufen ermöglicht [GXP19, S. 371ff.]. Darüber hinaus spielt die Datenformatierung eine wichtige Rolle bei der Verbesserung der Datenqualität und der Effizienz von Analysen [Fi22, S. 49].

3 Methodik zur Literaturanalyse

Die folgenden Abschnitte umfassen die Dokumentation und Planung der systematischen Literaturanalyse, um diese transparent zu beschreiben und offenzulegen [Fe06].

3.1 Verfahren

Zur Erstellung der Forschungsarbeit wird als Verfahren die Literaturrecherche nach [Fe06] in Kombination mit der Vorwärts- und Rückwärtssuche nach [WW02] verwendet. Die Auswertung der Literaturrecherche findet durch die Erstellung einer Konzeptmatrix ebenfalls, nach [WW02], statt. Diese dient einerseits der Unterstützung für die Analyse und Interpretation und andererseits bei der abschließenden Visualisierung der Ergebnisse. Die ausgewählten Verfahren eignen sich für die Strukturierung und dienen insbesondere der Nachvollziehbarkeit der Literaturanalyse.



Abbildung 1: Methodische Vorgehensweise (Eigene Darstellung in Anlehnung an [Fe06])

Gemäß dieser Vorgehensweise durchläuft die Literaturanalyse die nachfolgenden fünf Phasen:

1. **Problemformulierung:** Die Ausformulierung und Abgrenzung der zu beantwortenden Forschungsfrage (Kapitel 1.1 und Kapitel 1.2).
2. **Literatursuche:** Recherche geeigneter Literatur, welche in der Lage ist, die Forschungsfrage zu beantworten (Kapitel 3.2).
3. **Literaturauswertung:** Inhaltliche Relevanz der identifizierten Literatur in Bezug auf die Forschungsfrage prüfen (Kapitel 3.3).
4. **Analyse und Interpretation:** Ergebnisse der Literatursuche und Auswertung unter Berücksichtigung der Problemformulierung evaluieren (Kapitel 4).
5. **Präsentation:** Forschungsergebnisse aufbereiten und visualisieren.

3.2 Literatursuche

Für die Suche nach relevanten Publikationen wurden die Datenbanken AIS Electronic Library (AISEL), Association for Computing Machinery Digital Library (ACM), EmeraldInsight (Emerald), Google Scholar (GS), IEEE Xplore Digital Library (IEEE), ScienceDirect (SD) und Springer Link (SL) betrachtet. Außerdem wurden noch die folgenden Konferenzen: Americas Conference on Information Systems (AMCIS), European Conference on Information Systems (ECIS), International Conference on Information Systems (ICIS) und Internationale Tagung Wirtschaftsinformatik (WI) berücksichtigt. In einem nächsten Schritt wurden Suchbegriffe in deutscher als auch in englischer Sprache definiert. Um möglichst relevante Suchergebnisse zu erhalten, wurden die Suchbegriffe durch Operatoren logisch verknüpft:

ST1: („Digital Business Technology Platform“) AND („Anforderungen“ OR „Fähigkeiten“ OR „Kriterien“)

ST2: („Digital Business Technology Platform“) AND („Requirements“ OR „Capabilities“ OR „Criteria“)

ST3: („Anforderungen“ OR „Fähigkeiten“ OR „Kriterien“) AND („an digitale Geschäfts-Technologie Plattformen“)

ST4: („Digital Business Technology Platform“) OR („Digitale Geschäfts-Technologie Plattformen“)

Nachdem die Suchterme definiert wurden, werden im nächsten Schritt die Suchfilter festgelegt. Dieses Forschungsvorhaben konzentriert sich auf die Anforderungen und Kriterien für Digital Business Technology Platforms in dem Zeitrahmen von 2016 bis 2023, da die Plattformen seit 2016 zunehmend im Umlauf sind [Ob20]. Die Suchterme werden auf die Suchfelder Titel, Abstract und Schlüsselwörter begrenzt.

3.3 Ergebnis der Literaturanalyse

Die in aufgeführten Literaturdatenbanken und die aufgeführten Konferenzen wurden mit den festgelegten Suchtermen durchsucht. Eine gründliche Literaturrecherche wurde durchgeführt, um eine solide Datengrundlage zu schaffen. Alle 107 Ergebnisse der Literaturabfrage wurden sorgfältig bewertet, indem zuerst die Titel geprüft und schließlich 14 Abstracts auf inhaltliche Relevanz überprüft wurden. Mit einer Kombination aus Rück- und Vorwärtssuche [WW02] ergaben sich insgesamt sieben publizierte Arbeiten, die für das Thema von Bedeutung waren. Da es jedoch nur wenige Literaturquellen zu dem aktuellen Thema gab, wurden zusätzlich auch Internetquellen, wie Berichte renommierter Unternehmen wie Deloitte und Gartner, in die Überprüfung einbezogen.

4 Ergebnisse

Für die Darstellung der Ergebnisse der Literaturanalyse wird – wie in der Methodischen Vorgehensweise beschrieben – die Methode der Konzeptmatrix nach [WW02] angewandt, um einen Überblick über die relevante Literatur und deren Kernmerkmale zu erhalten. In der Konzeptmatrix werden bestimmte Anforderungen und Kriterien in Form von Zeilen und Literatur in Form von Spalten dargestellt. Wenn ein Thema in einem Artikel erwähnt wurde, wird ein Kreuz "x" in der entsprechenden Zeile und Spalte der Matrix gesetzt. Am Ende werden die Kreuze für jede Anforderung zusammengezählt, um einen Überblick darüber zu erhalten, welche Anforderungen eine Digital Business Technology Platform erfüllen sollte. Diese Methode dient dazu, die Häufigkeit von Anforderungen in einer Sammlung von Artikeln zu quantifizieren und damit Trends und Schwerpunkte in der wissenschaftlichen Literatur aufzuzeigen. Indem die Häufigkeit der Anforderungen in der Konzeptmatrix als Prozentangabe aufgezeigt wird, können Pflichtenanforderungen herausgearbeitet werden.

4.1 Konzeptmatrix

Die in Abbildung 2 dargestellte Konzeptmatrix zeigt die Anforderungen und Kriterien einer Digital Business Technology Platform, unterteilt in sieben Kategorien, auf. Darunter finden sich folgende Kategorien: Prozess, Effiziente Datenbehandlung, Analyse, Infrastruktur, Technologien, Preis und Benutzerfreundlichkeit. Die gesetzten Kreuze in der Konzeptmatrix geben einen Überblick darüber, in welcher Quelle welche Anforderungen und Kriterien genannt wurden.

Bereich	Kriterien	Quelle															Anzahl					
		Giacobbe, M.; Billa, M.; Polifino A. (2019)																				
		Cox, J (2020)																				
		Gartner (2017)																				
		Wumelic, J. (2022)																				
		Wanov, I. (2019)																				
		Ganesan, V. (2019)																				
		Bjorkback, A. (2022)																				
		Starin, L. (2021)																				
		Figueredo, M. (2022)																				
		McKinsey (2022)																				
		Burgardt, S. (2022)																				
		Bode, S. (2022)																				
		Chook, L. (2022)																				
		Heffner, R. (2019)																				
		Borggreve, C.; Wieke, R. (2022)																				
		Oppenheim (2022)																				
		Business Horizons (2021)																				
		Salesforce (2022)																				
		Microsoft (2022)																				
		Vantky, A. (2020)																				
		Oberlander, F. (2020)																				
Prozess	Prozessoptimierung																					25
	Prozessautomatisierung																					27
Effiziente Datenbehandlung	Datenformatierung																					3
	Datenkombination	x	x	x																		15
	Datenminimierung																					2
	Datenschutz	x	x																			9
Analyse	Echtzeitanalysen																					6
	Expressanalyse	x	x																			13
	Prognoseanalyse	x																				9
	Entscheidungshilfe / Empfehlung	x	x																			13
Infrastruktur	Skalierbarkeit	x																				7
	Integration von Technologien / Plattformen	x	x	x	x																	21
	Flexibilität	x	x																			6
	Anpassungsfähigkeit / Agilität																					8
	Cloudbasiert	x																				10
	Zentrale Verwaltung																					2
Technologien	Cybersicherheit																					6
	Blockchain																					3
	Internet of Things																					6
	Machine Learning	x																				6
	Künstliche Intelligenz	x	x																			11
Preis	Niedrige Einstiegskosten																					2
	Kosteneffiziente Infrastruktur	x																				5
	Reduzierte Entwicklerkosten																					4
Benutzerfreundlichkeit	Schnelle Anbindung																					8
	Einfacher Umstieg	x	x	x																		8
	Zugänglichkeit																					5
	Unkomplizierte Bereitstellung	x	x																			6
	Flexible Architektur	x																				5
	Internationalisierung																					4
	Support	x	x	x																		7

Abbildung 2: Konzeptmatrix

4.2 Anforderungen und Kriterien

Anhand der Konzeptmatrix konnten 31 Anforderungen und Kriterien an eine Digital Business Technology Platform identifiziert werden. Anschließend werden die einzelnen Anforderungen und Kriterien beschrieben. Diese werden in derselben Reihenfolge, wie in der Konzeptmatrix angeordnet.

Prozessoptimierung und **Prozessautomatisierung** sind zentrale Aspekte einer Digital Business Technology Platform. Die Optimierung von Geschäftsprozessen ist heutzutage ein wichtiger Faktor für den Erfolg eines Unternehmens. Durch die Automatisierung von Prozessen können Unternehmen ihre Effizienz verbessern, Kosten senken und Fehler minimieren. Eine moderne Digital Business Technology Platform ermöglicht es, bestehende IT-Systeme zu verbinden und durchgängige Prozesse abzubilden und zu

automatisieren, was eine wichtige Rolle bei der Optimierung von Geschäftsprozessen spielt. Die Funktionalitäten einer BPM (Business Process Management) Suite oder einer DBTP bieten dafür die nötigen Werkzeuge [Ob20, S. 73].

Digital Business Technology Platforms müssen eine einfache und flexible **Datenformatierung** bereitstellen, um Daten aus verschiedenen Quellen zu integrieren und zu kombinieren. Diese Datenkombination ermöglicht es Unternehmen, eine umfassende Sicht auf ihre Daten zu erhalten und diese effektiv zu nutzen. Eine **Datenminimierung** an der Quelle, bei der nur relevante Daten an die Plattform gesendet werden, trägt dazu bei, dass das System reibungslos läuft und Echtzeitmaßnahmen schnell ergriffen werden können [Va20]. Die DBTPs nutzen verschiedene Verschlüsselungsdienste, um Daten und Protokolle zu schützen, da der Datenschutz bei der Verwendung dieser Systeme sichergestellt werden sollte [Fi22].

Echtzeitanalysen sind ein wichtiger Bestandteil moderner Digital Business Technology Platforms. Durch die Nutzung von Datenvirtuelisierung können Daten aus verschiedenen Quellen vereinheitlicht werden, ohne dass sie ihren ursprünglichen Speicherort verlassen müssen. Dies ermöglicht es Unternehmen, in Echtzeit auf Daten zuzugreifen und wichtige Einsichten zu gewinnen [Co20, S. 5]. **Expressanalyse** bezieht sich auf tiefe Prozessanalyse für individuelle Empfehlungen. **Prognoseanalyse** nutzt historische Daten, um zukünftige Trends und Ereignisse vorherzusagen [Bo22]. **Entscheidungsunterstützung** nutzt Daten, Modelle und Analysen, um Entscheidungen zu informieren und zu verbessern. Dies kann durch Expressanalyse und Prognoseanalyse erreicht werden. Digitaltechnologie und KI ermöglichen Echtzeitanalyse, Expressanalysen, Prognoseanalysen und bessere Entscheidungen [Ch22].

Eine Digital Business Technology Platform muss bestimmte Anforderungen an die Infrastruktur erfüllen, um eine effektive Unterstützung für das Geschäft zu bieten. **Skalierbarkeit** ermöglicht es, die Kapazitäten der Plattform an die wachsenden Anforderungen des Geschäfts anzupassen [sa22]. **Integration von Technologien** ermöglicht es, unterschiedliche Systeme und Anwendungen in einer einzigen Plattform zu vereinen [Ga22]. **Flexibilität** und **Anpassungsfähigkeit** erlauben es, die Plattform an die spezifischen Anforderungen des Geschäfts anzupassen [BW22]. Eine **cloudbasierte Plattform** bietet zusätzliche Vorteile, wie z.B. die Verfügbarkeit von Ressourcen und Daten von jedem Ort aus, ohne dass lokale Infrastrukturen benötigt werden [Wu22]. **Zentrale Verwaltung** ermöglicht es, die Plattform effektiv zu überwachen und zu verwalten, um die optimale Leistung und Effizienz sicherzustellen [Co20].

Des Weiteren spielen ebenfalls verschiedene Technologien eine wichtige Rolle in Digital Business Technology Platforms. **Cybersicherheit** ist bei digitalen Geschäftsprozessen entscheidend, da sie sicherstellt, dass vertrauliche Daten und Systeme vor Hackerangriffen und Datenverlust geschützt sind [St21]. Cybersicherheitstechnologien, wie Verschlüsselung, Zwei-Faktor-Authentifizierung und Firewalls, werden in DBTPs verwendet, um eine sichere Umgebung für die Übertragung und Speicherung von Daten bereitzustellen. **Blockchain-Technologie** ist eine verteilte Datenbank, die eine

unveränderliche Aufzeichnung von Transaktionen ermöglicht [BW22]. In DBTPs wird Blockchain genutzt, um sichere und transparente Geschäftstransaktionen zu ermöglichen, ohne dass eine zentrale Instanz erforderlich ist. Das **Internet of Things (IoT)** ermöglicht es, Daten von Geräten und Sensoren in Echtzeit zu sammeln und zu analysieren. Diese Technologie wird genutzt, um eine bessere Überwachung und Steuerung von Geräten und Prozessen zu ermöglichen und zukünftige Trends vorherzusagen [Bö22, S. 26, Bu22]. **Machine Learning (ML)** und **künstliche Intelligenz (KI)** ermöglichen es, Daten zu analysieren und Mustererkennungsmodelle zu entwickeln. In DBTPs werden ML und KI eingesetzt, um Prozesse zu automatisieren, Entscheidungen zu treffen und Prognosen über zukünftige Trends und Ereignisse zu erstellen [Mc22].

Eine Digital Business Technology Platform kann eine kosteneffiziente Lösung für Unternehmen sein, die ihre Geschäftstätigkeiten digitalisieren möchten. Einer der Vorteile sind die **niedrigen Einstiegskosten**. Im Vergleich zu traditionellen Geschäftslösungen ist die Implementierung einer digitalen Plattform oft wesentlich günstiger. Dies liegt daran, dass keine teuren Hardware-Systeme oder physischen Räume erforderlich sind. Eine weitere Kosteneinsparung bei der Verwendung einer DBTP ist die **kosteneffiziente Infrastruktur**. Die Plattform nutzt Cloud-Computing, was bedeutet, dass Unternehmen keine teuren Systeme oder Ressourcen für den Betrieb und die Wartung ihrer Lösung bereitstellen müssen [Mc22]. Zusätzlich zu den Einsparungen bei den Infrastrukturkosten können Unternehmen auch ihre **Entwicklerkosten reduzieren**. DBTPs bieten bereits eine Vielzahl von Tools und Funktionen, die für die meisten Geschäftsanforderungen ausreichend sind. Dies bedeutet, dass Unternehmen weniger Zeit und Geld für die Entwicklung von maßgeschneiderten Lösungen ausgeben müssen [Ca22].

Benutzerfreundlichkeit ist ein wichtiger Faktor, der dazu beiträgt, dass eine Technologie für den Einsatz im täglichen Geschäft erfolgreich ist. Um eine hohe Benutzerfreundlichkeit zu gewährleisten, sollten verschiedene Faktoren berücksichtigt werden. Einer davon ist die **schnelle Anbindung**. Es sollte einfach sein, die Technologie in das bestehende Geschäftssystem zu integrieren. Ein weiterer wichtiger Faktor ist die **Einfachheit des Umstiegs**. Es sollte eine intuitive Benutzeroberfläche und einfache Anweisungen geben, um den Übergang zur neuen Technologie so unkompliziert wie möglich zu gestalten. **Zugänglichkeit** ist ebenfalls ein bedeutender Aspekt. Die Technologie sollte für eine breite Zielgruppe verfügbar sein, unabhängig von Alter, technischen Fähigkeiten oder Standort [Ca22]. Eine **unkomplizierte Bereitstellung** und **flexible Architektur** sind essenziell für die Benutzerfreundlichkeit. Eine einfache Installation und Wartung kann den Benutzern mehr Zeit für ihre täglichen Aufgaben geben [He19]. **Internationalisierung** ist ein wichtiger Faktor, wenn die Technologie in verschiedenen Teilen der Welt eingesetzt werden soll. Eine Technologie, die in mehreren Sprachen verfügbar ist, kann eine bessere Benutzerfreundlichkeit für Benutzer aus verschiedenen Teilen der Welt bieten. Schließlich ist ein guter Support von hoher Bedeutung. Benutzer sollten jederzeit Zugang zu qualifiziertem **Support** haben, um Fragen beantwortet und Probleme gelöst zu bekommen [Mi22].

5 Konklusion

Im Folgenden wird das Ergebnis diskutiert, das Vorgehen kritisch reflektiert und ein Ausblick über mögliche, auf diesem Paper aufbauende Forschungstätigkeiten gegeben.

5.1 Diskussion der Ergebnisse

Durch die Literaturanalyse von Digital Business Technology Platforms konnten 31 Kriterien identifiziert werden, die für die Bewertung von DBTP von Bedeutung sind. DBTP werden durch fünf am häufigsten identifizierten Kriterien beschrieben und deuten somit darauf hin, dass eine Plattform, diese Merkmale beinhalten sollte um als gut bewertet zu werden. Es ist zu beachten, dass diese Ergebnisse auf der ausgewählten, nicht repräsentativen Stichprobe und der Vorgehensweise der Studie basieren und möglicherweise in anderen Kontexten unterschiedlich ausfallen können.

Die aus der Literatur extrahierten Kriterien sind als Basis zu sehen, um eine unternehmensspezifische Plattform zugeschnitten auf die Unternehmensanforderungen am Markt auszuwählen. Somit die beste Lösung für ein Unternehmen gefunden werden. Es ist zu beachten, dass die Verwendung einer Liste von 31 Kriterien für die praktische Nutzung eine Herausforderung darstellen kann. Somit ist es denkbar, eine mehrstufige Betrachtung einzuführen, z.B. in Form eine gewichteten Nutzwertanalyse, um eine Vorsektion der Auswahlkriterien vorzunehmen.

5.2 Fazit

Abschließend lässt sich festhalten, dass das Ziel dieses Papers und die aus der Problemstellung entwickelte Forschungsfrage: „Wie ist der State-of-the-Art im Forschungsbereich der Digital Business Technology Platforms?“, sowie die dazugehörigen Unterfragen, im Rahmen des Papers erfolgreich beantwortet werden konnten. Hierfür wurden alle gewünschten Artefakte erfolgreich erstellt.

5.3 Kritischer Reflektion der Vorgehensweise

Die ausgewählte methodische Vorgehensweise der Literaturanalyse nach [Fe06] in Kombination mit [WW02] und der daraus entwickelte Vorgehensprozess hat sich, trotz geringer Anzahl an Literatur, als geeignete Methodik und Prozessdurchführung erwiesen. Aufgrund der Relevanz und der Aktualität des Themengebietes, könnte diese Forschung um andere Methodiken erweitert werden, um weitere Anforderungen und Kriterien an eine Digital Business Technology Platform zu identifizieren.

5.4 Zukünftiger Forschungsbedarf

Aufgrund der präsentierten Ergebnisse in diesem Papier ist eine Fortführung der Forschung angestrebt. Hierbei kann einerseits das entwickelte Kriterienraster als Zwischenergebnis in Richtung einer umfassenderen, mehrstufigen Bewertungsmatrix konzipiert und darauffolgend eine entsprechende Marktanalyse durchgeführt werden. Dadurch können DBTP-Anbieter anhand der Kriterien bewertet und miteinander verglichen werden. Andererseits sollte weitere Forschung in Richtung einer möglichst einfachen praktischen Handhabung des Kriterienrasters unternommen werden, insbesondere unter Berücksichtigung der Anforderungen kleiner und mittlerer Unternehmen.

Zudem wäre eine zukünftige Reproduktion dieses Forschungsvorgehens angebracht, um zeitnah Veränderungen und Trends zu identifizieren.

Literaturverzeichnis

- [Bo22] Bode, S.: Solutions. <https://www2.deloitte.com/de/de/pages/technology/solutions/sap-s4-hana.html>.
- [Bö22] Bögelsack, A. et al.: SAP S/4 HANA-Systeme in Hyperscaler Clouds. Amazon Web Services und Google Cloud. Springer Fachmedien Wiesbaden GmbH, Wiesbaden, 2022.
- [Bu22] Burghardt, S.: SAP Business Technology Platform. <https://www.gambit.de/wiki/sap-business-technology-platform/>
- [BW22] Borggreve, C.; Wickel, R.: Step up with SAP business technology platform. <https://www.accenture.com/us-en/insights/technology/sap-business-technology-platform>
- [Ca22] Capgemini: MICROSOFT DYNAMICS 365 AND POWER PLATFORM
- [Ch22] Cheok, L.: Digital-First: The Evolving Relationship of Business and Technology. <https://blogs.idc.com/digital-first-the-evolving-relationship-of-business-and-technology/>
- [Ci14] Ciobanu, R.-I. et al.: Big Data Platforms for the Internet of Things. Big Data and Internet of Things: A Roadmap for Smart Environments 546, 2014.
- [Co20] Cox, J.: SAP's Business Technology Platform is a major milestone in the Intelligent Enterprise story from SAP, 2020.
- [DT18] Dahm, M.; Thode, S. Hrsg.: STRATEGIE UND TRANSFORMATION IM DIGITALEN ZEITALTER. Inspirationen fr. GABLER, [Place of publication not identified], 2018.
- [Fe06] Fettke, P.: State-of-the-Art des State -of-the-Art. Wirtsch. Inform. (WIRTSCHAFTSINFORMATIK) 4/48, 2006.
- [Fi22] Figueiredo, M.: SAP HANA Cloud in a Nutshell. Apress, Berkeley, CA, 2022.
- [Ga22] Gartner: Plattform (Digital Business). <https://www.gartner.com/en/information-technology/glossary/platform-digital-business>.
- [GXP19] Giacobbo, M.; Xibilia, M.; Puliafito, A.: Building a Digital Business Technology Platform in the Industry 4.0 Era., Springer, [Cham], Switzerland, 396-375, 2019.

- [HBK18] Hein, A.; Böhm, M.; Krcmar, H.: Digitale Plattformen. In (Dahm, M.; Thode, S. Hrsg.): STRATEGIE UND TRANSFORMATION IM DIGITALEN ZEITALTER. Inspirationen fr. GABLER, [Place of publication not identified], S. 181–199, 2018.
- [HCB16] Hoelck, K.; Cremer, S.; Ballon P. Hrsg.: Cross-Platform Effects: Towards a Measure for Platform Integration Benefit, 2016.
- [He19] Heffner, R.: A Developer's Guide To Forrester's Strategies For Integration And Digital Business Platforms. <https://www.forrester.com/report/a-developers-guide-to-forresters-strategies-for-integration-and-digital-business-platforms/>
- [Iv19] Ivanov, I.: Chasing the Crowd: Digital Transformations and the Digital Driven System Design Paradigm, 9th International Symposium, BMSD 2019, Lisbon, Portugal, July 1-3, 2019, proceedings. Springer, Cham, Switzerland, S. 64–80, 2019.
- [Le21] Lexa, C.: Fit für die digitale Zukunft. Springer Fachmedien Wiesbaden GmbH; Springer Gabler, Wiesbaden, 2021.
- [Li19] Li, Q. et al.: Big Data Architecture and Reference Models, 2019.
- [Mc22] McKinsey: McKinsey Digital. <https://www.mckinsey.com/capabilities/mckinsey-digital/how-we-help-clients>.
- [Mi22] Microsoft: Erreichen Sie mit weniger mehr, indem Sie mithilfe codearmer Tools Anpassungen vornehmen. <https://powerplatform.microsoft.com/de-de/>.
- [Ob20] Oberländer, F.: Digital Business Platform: Die neue Basis für die Digitalisierung? <https://www.frox.ch/newsroom/blog-artikel/dbp-digital-business-platform/>
- [sa22] salesforce: Salesforce Platform, <https://www.salesforce.com/de/products/platform/>
- [SA22] SAP: Discover How SAP Business Technology Platform Driver Digital Transformation at Mohawk Industries, 2022.
- [St21] Starita, L.: How to Build a Digital Business Technology Platform, <https://www.gartner.com/smarterwithgartner/how-to-build-a-digital-business-technology-platform>
- [Va20] Vantiq, A.: What is a Digital Business Technology Platform? <https://vantiq.com/blog/what-is-a-digital-business-technology-platform/>
- [vSS09] vom Brocke, J.; Sonnenberg, C.; Simons, A.: Wertorientierte Gestaltung von Informationssystemen. Konzeption und Anwendung einer Potenzialmodellierung am Beispiel Serviceorientierter Architekturen. *Wirtsch. Inform.* 3, S. 261–272, 2009.
- [Wu22] Wunschick, J.: Business Technology Platform. <https://e-3.de/business-technology-platform/>
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 2/26, S. xiii–xxiii, 2002 <https://powerplatform.microsoft.com/de-de/>

Öffentliche Infrastruktur -
IT-Kompetenzen für die digitale
Verwaltung der Zukunft

The Public Official's Selection Parameters for E-Competence Continuous Education

Michael Koddebusch,¹ Paul Brützke,² Holger Koelmann³ Jörg Becker⁴

Abstract: Given the public officials' importance in public sector digitalization, e-government-competence education must be designed in accordance with their individual needs. Only by fostering e-government-competences, or short, e-competences, can they actively participate in the transformative process. Although the importance of e-competences is evident, the landscape of education and continuous education offers does not address the diversity of requirements of the public sector's workforce. Therefore, enabling public administrations to provide their workforce with appropriate continuous education is imperative. The present study sheds light on the parameters public officials apply to assess the appropriateness of a given continuous education offer for their individual situation. To achieve these insights, we conducted an in-depth interview study with a subsequent qualitative content analysis, utilizing eleven semi-structured interviews with members of the German public sector. Upon the interview results, we discovered the parameters and constructed an Entity-Relationship-Model reflecting them for future use in continuous education systems.

Keywords: E-Competence; Continuous Education; E-Government; Public Sector; Digital Transformation

1 Introduction

Public officials, i.e., public sector employees, play an essential part in implementing public e-services and e-government as a whole [GGP05, LL01, LMS21]. In a recent conceptualization of e-government, Lindgren et al. [LMS21] include public officials in their definition of stakeholders as "people who contribute to e-government [...] through direct or indirect participation in [...] processes and activities". However, this "participation", i.e., the possibility to act in a digitalized public sector, requires the public official to possess e-government-competences [DOB19]. In line with the steadily progressing digitalization of the public sector, the demand for e-government-competences is growing [Ko22]. Fulfilling this demand is a challenging endeavor as *digitalization* is a cross-cutting issue and no longer a matter of the IT department only – the necessity to obtain e-government-competences affects public officials of all hierarchy levels and units [Li19].

Despite the evident importance, the current workforce's education never revolved around digitalization. Public officials, especially in Germany, are trained to implement services

¹ University of Münster, European Research Center for Information Systems, koddebusch@ercis.de

² University of Münster, European Research Center for Information Systems, p.bruetzke@uni-muenster.de

³ University of Münster, European Research Center for Information Systems, koelmann@ercis.de

⁴ University of Münster, European Research Center for Information Systems, becker@ercis.de

with a particular focus on satisfying legal frameworks. Moreover, even though there is no scientific analysis regarding the composition of public administration degree programs or dual training, spot checks of education offers show that digitalization is merely a side note.⁵

Consequently, a particular focus must be on today's public officials' further education and professional development. Demographic trends, the shortage of skilled labor, and the competition for talent with the private sector permit us to rely on future generations of public officials to bring e-competence with them. However, competence acquisition is a complex process, and simply providing public officials with a set of continuous education offers will most likely be unsuccessful. Considering the variety of different roles and the individually differing life- and job circumstances, we must understand which parameters public officials apply when selecting a particular continuous education offer and which characterizations they appreciate. Based on this understanding, public administrations can design their continuous education portfolios addressing these parameters and, thus, pick the public officials up where they are. Therefore, our *research question* reads as follows:

According to which parameters do public officials select continuous education offers for e-government-competences?

To that end, we employ an in-depth interview study with eleven public officials. The remainder of the article entails the research background, research method, presentation of our results with a subsequent discussion, and a conclusion.

2 Research Background

While digital transformation generally benefits organizations to work more efficiently, the public sector of many developed economies struggles to use these new opportunities effectively [Eu22]. The competences of the public officials in dealing with all aspects of this digital transformation (often also referred to as *e-government* [LMS21, Le02]) are one cornerstone to succeed in these endeavors [DOB19, Ko22]. In contrast, these competences are often found to be lacking in the existing workforce and even identified as the *e-competence gap* within the public sector [Ko22]. In this context, we refer to *competences* as “*an individual's work-related knowledge, skills and abilities*” [DOB19], with competences needed for e-government-related activities to be called *e-government-competences* (or *e-competences* in short). Regarding the digitalization of the public sector, multiple works exist trying to understand and classify e-competences [DOB19, HS13, HPS15], such as the categorization by Hunnius et al. [HPS15] into the broad categories of “*technical*”, “*socio-technical*”, “*organizational*”, “*managerial*”, and “*political-administrative*” competences, each comprising distinct competences such as *process management competence*, *IS design competence*, or *competence in e-government policy*.

However, the pure conceptualization of these e-competences does not help to apply them. They need to be acquired and further developed by public officials to be useful in the digital

⁵ Example: HWR Berlin - Public Administration Bachelor of Arts

transformation of the public sector. Beyond improvements in the education of new public sector employees [Sa22], which might only help with the digitalization efforts in decades to come, *continuous education* in e-competences for existing public sector employees is a more promising way to close the e-competence-gap in the present day [Ha21, OH18]. For this to succeed, continuous education offers should facilitate further e-competence development and fit the learners' requirements. With recent developments in new and digital teaching formats, such as purely digital and asynchronous formats like Massive Open Online Courses (MOOCs) [Eg18], the number of options to choose from when deciding which continuous education offer to take, has increased substantially. When looking at the current reality for the public sector, we can see a need for more attractive and fitting offers, e.g., for the German market [St19]. In addition, each public official needs to find or be presented with the right continuous education offers for the best possible results regarding their specific competence needs. This study, therefore, takes a closer look into the relevant selection parameters public officials use when assessing a specific continuous education offer's usefulness to them.

3 Research Method

We conducted an in-depth interview study with public officials to answer our research question. We sent out inquiries to 23 mid-sized town administrations, outlining our research endeavor and requesting interview partners from their workforce. Mid-sized towns are characterized by a population of 20.000-100.000 inhabitants and generally deal with particular digitalization challenges because Smart City concepts are typically designed for large cities [ABD15] and cannot be easily transferred to the circumstances of mid-sized towns. The sampling happened based on project commitments and availability to the researchers. To achieve an inclusive, open environment of input perspectives, we did not apply a focus on particular responsibilities within the municipal administration. From 23 inquiries, five towns provided us with interview partners, which enabled us to conduct a total of eleven interviews (see Tab. 1). As we reached saturation, i.e., no new insights emerged, after the 10th interview, we refrained from conducting a second round of interviews. The 11th interview, from which we gained no new results, was still conducted, as it had already been scheduled.

Throughout the interviews, we employed a *semi-structured interview design* described by Rowley [Ro12] and Brinkmann [Br13], which enabled us to dynamically adapt the questions to the respondents' answers and ask follow-up questions at appropriate points. The *guideline* was developed partly based on current public sector education literature (e.g., [OH18, JEO12]) and partly on domain experience and expertise resulting from projects operating in the same domain. We divided our interview guideline into five sections, promoting each section to explore different aspects of public officials' viewpoints:

- 1. General introduction:** Introduction and personal experience with continuous education. Example: *What kind of continuous education have you had experience with yourself so far?*
- 2. Motivation to partake in continuous education:** Factors influencing the individual

perception of continuous education. Example: *What requirements do you have for continuous education in e-competences?*

3. Participation requirements: Requirements that affect participation in continuous education. Example: *What factors in your professional life influence your participation in continuous education?*

4. Performance requirements: Reflection on the current design and diversity of the continuous education landscape. Example: *How should continuous education opportunities be best integrated into your personal and professional life?*

5. Outlook: Reflection on general aspects the interviewees wanted to address, which we did not specifically ask about.

Tab. 1: Interview Partners

Interview	Job	Responsibilities			
		Management	Continuous Edu.	Digitalization	Other
I1	Head of Project Management & Organizational Development Unit	X		X	
I2	Digitalization Officer	X		X	
I3	Clerk for Organization		X		X
I4	Department Lead for Internal Services, Staff, Organization & IT	X	X	X	
I5	Clerk for Organization			X	X
I6	Clerk in the Administration Control				X
I7	Coordination for Digitalization & E-Government			X	
I8	Clerk for Staff and Organization		X		X
I9	City Planning Office & Digital Ambassador			X	X
I10	Teamlead Waste Management, Contributions & Traffic	X			X
I11	Teamlead Staff & Organization	X		X	

Overall, we collected 523 minutes of recordings throughout the 11 interviews, whereas the average interview length was 47 minutes. Furthermore, the shortest interview lasted 33 minutes, and the longest interview was 61 minutes. The interviews were transcribed and analyzed by employing qualitative content analysis [Ma00]; we applied an inductive-deductive coding scheme to categorize our results. We then decided on an *Entity-Relationship-Model (ERM)* to demonstrate the findings. The choice of using an ERM is founded in its ability to examine situations regarding the entities (=single objects) involved in this situation. In our context, the situation is the public official choosing a continuous education offer and each selection parameter is one entity. Each code category relates to one parameter category (course/professional/private) in the ERM, and each sub-code category relates to one or more specializations of the parameter categories. Thus, the ERM enables us to organize the trade-offs that public officials often make only implicitly in their decision-making process for continuous education offers. Additionally, the data model notation provides a sound basis for potential implementation into a working prototype in which the specialization can be realized as filter functions. Overall, the elaboration on participation parameters affecting the continuous education choice of public officials provides an overview of potential parameter combinations.

4 Results

Based on the outlined interview analysis, we understand that, when participating in a continuous education offer, each participant applies a varying set of parameters that determine the choice for a particular offer. The entirety of parameters emerging from our interview analysis can be divided into four categories: content-related and course-related parameters, as well as parameters emerging from a given participant's professional and private lives (Figure 1). Each of these categories entails 1-n specializations. Each generalization-specialization is of the type *non-disjoint, partial (N,P)*, meaning that each entity (=category) can belong to one or more of its specializations. At the same time, as our results originate from a limited set of interviews, we must provide for the possibility of our model to be non-exhaustive.

In the following, the course-, professional-, and private parameters are elaborated on. As the interviews were conducted focusing on continuous education for e-competences, elaborating on the content parameters is moot. To have the interview participants speak freely about their requirements, we did not further specify our understanding of e-competences. Some interviewees understood e-competences to be rather broad concepts such as process management, organizational management, technical understanding, and a general understanding of the digitalized world, which closely relates to scientific e-competence frameworks, e.g., by [HPS15]. In contrast, other interviewees were much more focused on specific applications, such as Microsoft Outlook or Microsoft Excel or information systems specifically required for their job.

As large parts of the following diagrams (Figure 2, Figure 3, Figure 4) should be self-explanatory, the elaboration will be limited to a sample of entities in need of more explanation.

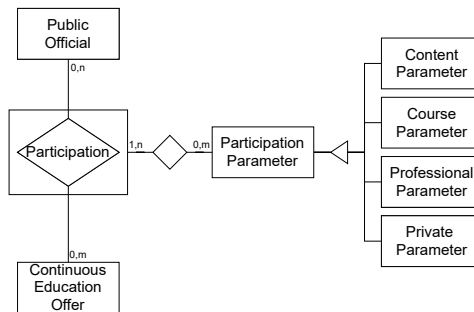


Fig. 1: Entity-Relationship-Model for E-Competence Continuous Education Participation Parameters

Course Parameters: The *Course Parameters*, Figure 2, show the structural characteristics of continuous education offers, and relate to the research problem because, based on individual circumstances, the requirements of *how* public officials partake in continuous education can differ significantly from each other. The course parameters are divided into four specializations. The *course focus* determines whether the offer is adapted to the public sector and whether its contents are more oriented towards acquiring (conceptual) or applying (practical) knowledge. *Acquisition* hereby refers to offers conveying abstract concepts and

theories of how something should be done, e.g., procedure models, management approaches, and frameworks. *Application*, in contrast, means learning by doing, e.g., a programming language or a certain software application.

“It is important that I can apply my knowledge. From my point of view, this has not only to do with the competence of having this knowledge but also with the practical experience.”

The *content delivery* specialization differentiates between conveying content in rather conventional classroom settings, in which the participants are required to be at a certain place at a certain time, or whether digital technologies are used to facilitate more innovative training formats. For *digital formats*, the interviewees again differentiated between synchronous communication, i.e., real-time exchange using a conferencing tool, or asynchronous settings, i.e., Massive-Open-Online-Courses with a supervised forum. Interestingly, the great majority of interviewees did generally prefer one of these delivery methods; all of them highlighted the dependence on content and envisioned outcomes of participation. “But I also notice that there are certain formats or topics that are simply much better in presence or digital. So it depends a bit on the topic.”

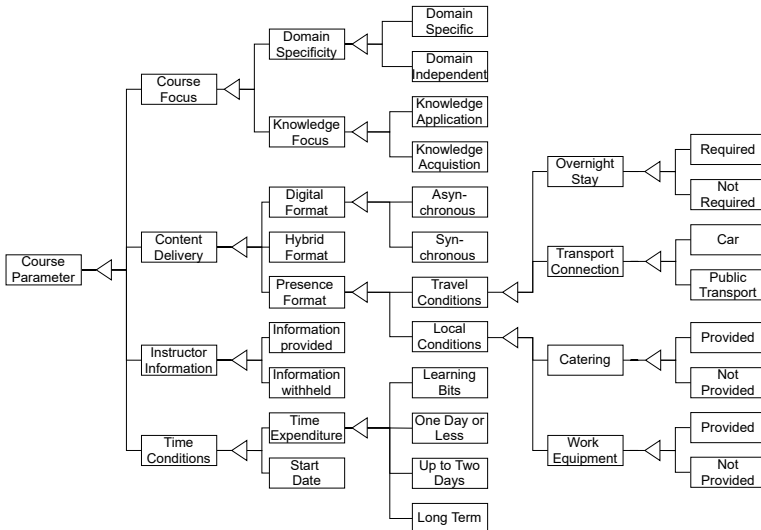


Fig. 2: Specialization for Course Parameters

When considering *presence formats*, several aspects influence participation, especially related to *travel conditions* and the *local conditions* on-site of the education offer. The actual traveling time did not seem to play a big role unless it resulted in a requirement to stay overnight, which would prevent some interviewees from participating. Also, some participants said they would prefer to have the training somewhere well-connected to public transport, as arriving by car would not be an option.

The specialization *instructor information* points to the influence of the instructor of training on participation. The interviewees reported that the quality and likability of an instructor

make a significant difference when choosing a continuous education offer. Thus, the knowledge about the person carrying out the training could heavily influence the intention to participate.

“And I mean, either way, of course, it’s always very dependent on the instructor.”

The last specialization of the *course parameters* are the *time conditions*. On the one hand, this refers to the start date of a continuous education offer, as there should be a direct temporal relationship to the professional necessities of the participants. On the other hand, the total time expenditure to spend on the training was an important influence factor. The interviewees reported strong temporal restrictions in their lives and therefore disliked continuous education offers that contained much more content than they actually needed. Rivetingly, some participants highlighted the need for more so-called *Learning Bits* (or Micro Content), characterized by short inputs consumed during the day, e.g., between meetings [ZW20].

Professional Parameters: The *professional parameters* are the job-related considerations public officials undertake when evaluating continuous education offers. Referring back to the initial problem statement of public officials having varying life- and job circumstances, this parameter category is essential because administrations impose many restrictions upon their employees, e.g., concerning the budget for continuous education, the working time flexibility, or the possibility to work remotely. The most apparent specialization is the current *job situation*, i.e., the area in which one works, the current workload, whether a short-term substitution (for multi-day training) is available, and whether the training is a priority at the moment. Regarding the task area, the interviewees reported the recurring difficulty in deciding whether a certain continuous education offer is appropriate for their own role. For example, some interviewees reported that many would not know which continuous education they need for their tasks and, therefore, would find it useful to receive suggestions based on their task area.

“So in the advance of continuous education, it is always difficult to say, what do I really need personally?”

The specialization *education needs* refers to the formal requirements the public official applies to the continuous education offer of choice. First, the *desired results* from continuous education are manifold. Whereas *competence acquisition* seems obvious, some participants primarily attend continuous education for formal certificates, networking purposes, or materials they can distribute among their colleagues. One interviewee even said the main benefit of continuous education is expanding the horizon and taking creative inspiration regarding alternative ways of working; tangible artifacts were of secondary importance.

“For me, continuous education is also about opening up one’s horizons from routine life in order to be stimulated by new impulses.” Second, training should be appropriate to the own *skill level*. The interviewees spoke about the heavily varying skill level among colleagues. Contents too easy for one could be perceived as too complex for the other.

“It is probably difficult to find the right level if one wanted to make such a training for all because the starting conditions are completely different. I think you have to pick people up at different individual knowledge levels.”

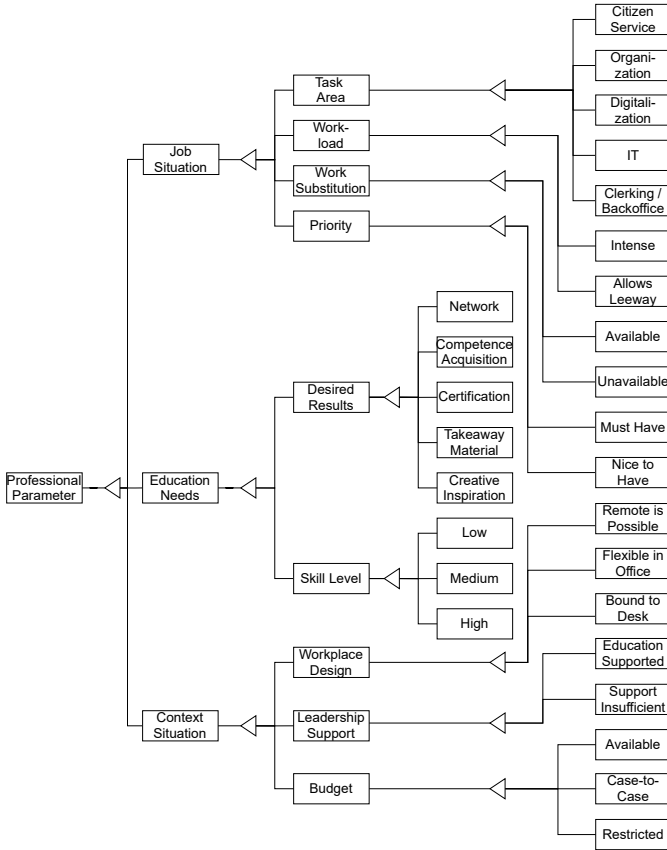


Fig. 3: Specialization for Professional Parameters

The last specialization for professional parameters is the *context situation*, i.e., parameters related to the professional environment. The available *budget* and *leadership support* are well-known but still strongly underlined parameters that often decide whether participation in an offer is possible or not. Another consideration is *workplace design*. The interviewees, e.g., made a difference in the possibility of partaking in digitally mediated offers depending on the possibility of pursuing them from home. Some interviewees highlighted the organization’s scarcity of appropriate technical communication equipment (cameras, headsets, laptops, etc.). This would require them to participate in digital offers in open-plan offices, which was found to be an unpleasant scenario.

“For example, if I attend a webinar, I have to borrow a headset beforehand and I have to borrow a camera, we don’t have those as standard at our workstations.”

Private Parameters: The last parameter category refers to the private lives of public officials and covers those parameters that influence the choice for a continuous education offer. For

once, one interviewee pointed out the desire to attend continuous education dedicated to female development. She discussed the continued unequal treatment of men and women in a professional environment, especially because leadership is often dominated by men. Also, as a young woman with the prospect of being absent for a while due to pregnancy, she wants to ensure she has enough competences to be attractive to the job market even after maternity leave.

“In a classic career, I think especially with regard to women, [...] you recognize the need for education somewhat more and would like to have more support.”

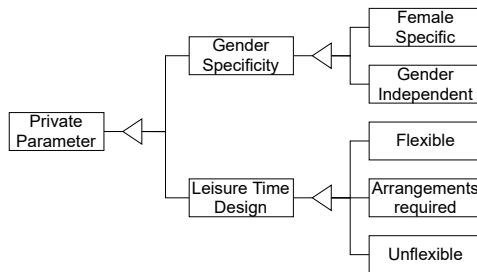


Fig. 4: Specialization for Private Parameters

Furthermore, this regards flexibility in *leisure time design* as a major influence factor. Many interviewees addressed factors in their private lives which would restrict their flexibility, such as children to attend to, social obligations, or commitments in associations or clubs. Some elaborated on social support structures, such as grandparents being able to watch the kids or understanding partners, but highlighted these as not being long-term reliefs but occasional temporary solutions.

5 Discussion

The presented model (and its specializations) depict the parameters public officials apply when selecting an individually appropriate continuous education offer. With our results, we contribute to research and practice in several ways.

Implications for Research: Our model extends the literature on understanding the public official and public sector education. State-of-the-art literature often speaks of *the public official* as if it was a rather homogeneous group of individuals to be treated with similar measures [Ca22, AK10]. While the importance of the public official has been long realized [GGP05, LL01], and we know that there are new, digitalization-induced competence requirements for the public official [MEH19], there is only little work developing an understanding of the different requirements public officials have and realities they operate in. Some scholars have attempted to conceptualize digitalization roles in governments [ORB19] or have generally called for more innovative continuous education formats to foster e-competence development [Ha21]. Other researchers investigate very specific roles, such as the Government Chief Information Officer and their competence requirements [EJ13].

Our results amend these works by providing a hint about the inner life of each public official, independent of certain jobs or roles. The presented model defines important parameters for all public officials, especially when it comes to the evaluation of continuous education offers. It also allows for content-related configuration depending on the circumstances of the public official who uses it, as not all parameters apply to all public officials equally.

Moreover, we add to the stream of government education literature. While there are works discussing the challenges and benefits of continuous education in the public sector [OH18] and developing theoretical constructs for effective e-government education [JEO12], we draw a picture of the requirements of the public official to find the individually suitable offer. For example, Janowski et al. [JEO12] or Ogonek et al. [ORB19] have provided role concepts of who should receive e-government education; amended by our parameters, these roles could be complemented with a role-specific configuration of requirements for continuous education, making the roles more applicable for real public officials.

Implications for Practice: Practitioners can use our model in several ways. It allows public administrations to check their continuous education portfolio for parameter combinations they already address and where they can improve to engage more staff. Moreover, it can help individuals responsible for personnel development in public administrations to curate continuous education offers according to specific parameters. Next, continuous education providers can use the model to align their offers to either span a maximum number of possible parameter combinations or focus on a particular niche (e.g., personnel with a low skill level and restricted time, budget, and personal flexibility). Additionally, the model is a first step towards a configurable system allowing public officials to screen the market for continuous education offers based on their individual needs.

Limitations: First, even though we reached saturation during our interview study, we cannot claim that the results are fully exhaustive. This is further exacerbated by our sample being limited to mid-sized towns; it might be possible that there are selection parameters we could not capture in our model because they originate from other public sector organizations. Next, some specializations (e.g., Budget, Skill Level, Time Expenditure) depict continuous constructs, such as money and time. To make them applicable to the model, we decided to frame them as non-disjoint entities; we are aware of potential issues associated with this, e.g., the unrealistic implementation of these entities. Moreover, the model has not yet been evaluated for its applicability and, therefore, for its purpose. Hence, it is yet to be investigated whether the application of the model actually leads to an increase in e-competence among the workforce.

6 Conclusion

The article at hand presents an Entity-Relationship-Model of the parameters that public officials apply when selecting continuous education offers for e-competences. The results were gained by an in-depth interview study with eleven public officials and subsequent

qualitative content analysis. We contribute to research and practice by extending the role of the *public official* and aim at facilitating e-competence development in the public sector. We call on future research to further extend the model. First, it is necessary to examine the content-related parameters more closely because there is no uniform understanding among public officials of what *e-competence* is. Second, we need to understand the interdependencies between the parameters better. This entails the reciprocal effects, i.e., the influences between parameter preferences, and includes potential recurring combinations of parameter choices, e.g., for specific roles. Third, the model must be applied and evaluated in real-life, potentially through a first working prototype of an operable system.

Acknowledgements: This research article has received funding from the project DFG FOR3539 (GZ: BE 1422/28-1).

Bibliography

- [ABD15] Albino, Vito; Berardi, Umberto; Dangelico, Rosa Maria: Smart cities: Definitions, dimensions, performance, and initiatives. *J. urban Technol.*, 22(1):3–21, 2015.
- [AK10] Aikins, Stephen Kwamena; Krane, Dale: Are Public Officials Obstacles to Citizen-Centered E-Government? An Examination of Municipal Administrators' Motivations and Actions. *State Local Gov. Rev.*, 42(2):87–103, 2010.
- [Br13] Brinkmann, Svend: *Qualitative Interviewing*. Oxford University Press, 05 2013. Doi: 10.1093/acprof:osobl/9780199861392.001.0001.
- [Ca22] Cagigas, Diego; Clifton, Judith; Díaz-Fuentes, Daniel; Fernández-Gutiérrez, Marcos; Echevarría-Cuenca, Juan; Gilsanz-Gómez, Celia: Explaining public officials' opinions on blockchain adoption: a vignette experiment. *Policy Soc.*, 41(3):343 – 357, 2022.
- [DOB19] Distel, Bettina; Ogonek, Nadine; Becker, Jörg: eGovernment Competences Revisited – A Literature Review on Necessary Competences in a Digitalized Public Sector. In: *Proc. 14th Int. Conf. Wirtschaftsinformatik*. Siegen, pp. 286–300, 2019.
- [Eg18] Egloffstein, Marc: Massive open online courses in digital workplace learning: Current state and future perspectives. In: *Digit. Work. Learn. Bridg. Form. Informal Learn. with Digit. Technol.*, pp. 149–166. 2018.
- [EJ13] Estevez, Elsa; Janowski, Tomasz: Landscaping government chief information officer education. In: *2013 46th Hawaii Int. Conf. Syst. Sci.* IEEE, pp. 1684–1693, 2013.
- [Eu22] European Commission: , eGovernment Benchmark 2022. <https://digital-strategy.ec.europa.eu/en/library/egovernment-benchmark-2022>, July 2022. (accessed May 10, 2023).
- [GGP05] Gil-García, J Ramón; Pardo, Theresa A: E-government success factors: Mapping practical tools to theoretical foundations. *Gov. Inf. Q.*, 22(2):187–216, 2005.
- [Ha21] Halsbenning, Sebastian; Koddebusch, Michael; Niemann, Marco; Becker, Jörg: How to Foster e-Competence in the Public Sector? A Mixed-Method Study Using the Case of BPM. In: *EGOV-CeDEM-ePart*. pp. 141–151, 2021.

- [HPS15] Hunnius, Sirko; Paulowitsch, Benedikt; Schuppan, Tino: Does E-Government education meet competency requirements? An analysis of the German university system from international perspective. In: 2015 48th Hawaii Int. Conf. Syst. Sci. IEEE, pp. 2116–2123, 2015.
- [HS13] Hunnius, S; Schuppan, T: Competency Requirements for Transformational E-Government. In: 2013 46th Hawaii Int. Conf. Syst. Sci. pp. 1664–1673, jan 2013.
- [JEO12] Janowski, Tomasz; Estevez, Elsa; Ojo, Adegboyega: Conceptualizing electronic governance education. In: 2012 45th Hawaii Int. Conf. Syst. Sci. IEEE, pp. 2269–2278, 2012.
- [Ko22] Koddebusch, Michael; Halsbenning, Sebastian; Kruse, Paul; Räckers, Michael; Becker, Jörg: The Increasing e-Competence Gap: Developments over the Past Five Years in the German Public Sector. In: Int. Conf. Human-Computer Interact. Springer, pp. 73–86, 2022.
- [Le02] Lenk, Klaus: Electronic Service Delivery—A driver of public sector modernisation. Information polity, 7(2-3):87–96, 2002.
- [Li19] Lindgren, Ida; Madsen, Christian Østergaard; Hofmann, Sara; Melin, Ulf: Close encounters of the digital kind: A research agenda for the digitalization of public services. Gov. Inf. Q., 36(3):427–436, 2019.
- [LL01] Layne, Karen; Lee, Jungwoo: Developing fully functional E-government: A four stage model. Gov. Inf. Q., 18(2):122–136, 2001.
- [LMS21] Lindgren, Ida; Melin, Ulf; Sæbø, Øystein: What is e-government? Introducing a work system framework for understanding e-government. Communications of the Association for Information Systems, 48(1):43, 2021.
- [Ma00] Mayring, Philipp: Qualitative Content Analysis. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 1, 06 2000.
- [MEH19] Mergel, Ines; Edelmann, Noella; Haug, Nathalie: Defining digital transformation: Results from expert interviews. Gov. Inf. Q., 36(4), may 2019.
- [OH18] Ogonek, Nadine; Hofmann, Sara: Governments’ Need for Digitization Skills: Understanding and Shaping Vocational Training in the Public Sector. Int. J. Public Adm. Digit. Age, 5(4):61–75, 2018.
- [ORB19] Ogonek, Nadine; Räckers, Michael; Becker, Jörg: How to Master the “E”: Tools for Competence Identification, Provision and Preservation in a Digitalized Public Sector. In: Proc. 12th Int. Conf. Theory Pract. Electron. Gov. pp. 56–64, 2019.
- [Ro12] Rowley, Jennifer: Conducting research interviews. Management Research Review, 35(3/4):260–271, 2023/01/05 2012. Doi: 10.1108/01409171211210154.
- [Sa22] Sarantis, Demetrios; Dhaou, Soumaya Ben; Alexopoulos, Charalampos; Ronzhyn, Alexander; Mureddu, Francesco: Digital Governance Education: Survey of the Programs and Curricula. Public Adm. Inf. Technol., 38:101 – 119, 2022.
- [St19] Stern, Sebastian; Klier, Julia; Kirchherr, Julian; Pethö-Schramm, Anneke; Susnjara, Isabella; Ohloff, Timur; Sönmez, Neslihan Ana: Die Besten, bitte: Wie der öffentliche Sektor als Arbeitgeber punkten kann. Technical report, McKinsey & Company, 2019.
- [ZW20] Zhang, Jiahui; West, Richard E: Designing Microlearning Instruction for Professional Development Through a Competency Based Approach. TechTrends, 64(2):310–318, 2020.

Vernetzung der Behörden-IT mit Lehre, Wissenschaft und Forschung

IT-Recruiting im Bundesamt für Migration und Flüchtlinge

Uta Katja Schlichte¹, Lisa Schröder-Puls², Carol Marmor-Drews³, Maryna Butara⁴, Dr. Sabine Schlick⁵ und Dr. Martin Kohls⁶

Abstract: Das Bundesamt für Migration und Flüchtlinge (BAMF) hat in den letzten Jahren eine enorme Transformation als Behörde erlebt - für die Bewältigung dieses Wandels wurden vielseitige strategische Maßnahmen und Projekte ins Leben gerufen. Mit Gründung des IT-Research-Teams 2019 wurde das Augenmerk auf moderne Recruiting- und Personalqualifizierungsmethoden von IT-Fachkräften gerichtet. Im Fokus des IT-Research-Teams steht die Vernetzung des Bundesamtes mit Wissenschaft und Forschung. Neben dem gemeinsamen Voranbringen der digitalen Transformation und dem Schaffen von Innovationen hat das Team die Aus- und Weiterbildung von internen Mitarbeitenden sowie die Gewinnung von neuen IT-Fachkräften im Blick. Neben Forschungsk Kooperationen mit Professorinnen und Professoren werden studentische Arbeiten in Zusammenarbeit mit der IT-Abteilung erstellt. Auch interne Mitarbeitende nutzen die Chance, sich neben ihrer Arbeit im Bundesamt durch die Betreuung ihrer Abschlussarbeiten weiter zu qualifizieren. Wir wollen unsere bisherigen Erfahrungen teilen, darstellen und diskutieren, um den Austausch von Behörden mit Hochschulen und Wissenschaft mit Blick auf Ausbildung und Weiterbildung weiter voranzubringen.

Keywords: IT-Recruiting, IT-Ausbildung, IT-Fachkräfte, Weiterbildung, Praxis, Vernetzung, Wissenschaft, Forschung, Lehre, Erfahrungsaustausch

1 Ausgangslage: Eine Behörde im schnellen Wandel

Das Bundesamt für Migration und Flüchtlinge hat vielfältige Aufgaben in den Bereichen Asyl, Migration und Integration. Die starke Fluchtmigration vor allem aus Syrien in den

¹ Bundesamt für Migration und Flüchtlinge, Referat 21 A Digitalstrategie, IT-Research, Chief Digital Office (CDO), Frankenstraße 210, 90461 Nürnberg, UtaKatja.Schlichte@bamf.bund.de

² Bundesamt für Migration und Flüchtlinge, Referat 21 A Digitalstrategie, IT-Research, Chief Digital Office (CDO), Frankenstraße 210, 90461 Nürnberg, Lisa.Schroeder-Puls@bamf.bund.de

³ Bundesamt für Migration und Flüchtlinge, Referat 21 A Digitalstrategie, IT-Research, Chief Digital Office (CDO), Frankenstraße 210, 90461 Nürnberg, Carol.Marmor-Drews@bamf.bund.de

⁴ Bundesamt für Migration und Flüchtlinge, Referat 21 A Digitalstrategie, IT-Research, Chief Digital Office (CDO), Frankenstraße 210, 90461 Nürnberg, Maryna.Butara@bamf.bund.de

⁵ Bundesamt für Migration und Flüchtlinge, Referat 21 A Digitalstrategie, IT-Research, Chief Digital Office (CDO), Frankenstraße 210, 90461 Nürnberg, Sabine.Schlick@bamf.bund.de

⁶ Bundesamt für Migration und Flüchtlinge, Referat 21 A Digitalstrategie, IT-Research, Chief Digital Office (CDO), Frankenstraße 210, 90461 Nürnberg, Dr.Martin.Kohls@bamf.bund.de

Jahren 2016 bis 2018⁷ führte zu einer tiefgreifenden Umstrukturierung und einem starken personellen Ausbau als Reaktion auf die wachsenden Herausforderungen in allen Aufgabenbereichen. Aktuell sind rund 8.300 Mitarbeitende beim dezentral strukturierten BAMF beschäftigt, davon gut ein Drittel in der Zentrale in Nürnberg und zwei Drittel in den ca. 60 Standorten des BAMF, welche im gesamten Bundesgebiet vertreten sind [Bu23].

Die Umstrukturierung und der personelle Ausbau steigerten auch die Anforderungen an die IT, boten aber gleichzeitig auch neue Chancen. Ging es anfänglich vor allem um eine schnelle Stabilisierung und Modernisierung der IT-Infrastruktur, steht nun zunehmend die Neugestaltung der informationstechnologischen Prozesse im Fokus. 2016 erstellte das BAMF eine erste Digitalisierungsagenda, die regelmäßig evaluiert und weiterentwickelt wird [Bu20]. Mit der ersten Digitalisierungsagenda wurde das IT-Labor in der Zentrale in Nürnberg eingerichtet, um Räume für die Umsetzung agiler Softwareentwicklungsmethoden mit Scrum-Teams zu schaffen. Mittlerweile hat sich das IT-Labor zum I-Lab weiterentwickelt, welches innovative Arbeitsmethoden über die IT hinaus fördert und unterstützt.

2 Dringend gesucht: Mehr Personal und Wissenstransfer

Das BAMF hat innerhalb kürzester Zeit sein Personal nahezu verdoppelt. Mit der Gründung des Qualifikationszentrums 2015 wurde bereits die Grundlage geschaffen, Fachkräfte für den Einsatz in den operativen Bereichen Asyl, Migration und Integration sowohl fachlich als auch in der Nutzung der eigenen IT-Anwendungen aus-, weiter- und fortzubilden. Auch die IT-Abteilung wurde in kurzer Zeit um- und ausgebaut, um die Asylprozesse bestmöglich zu unterstützen und zu beschleunigen. Möglich wurde dies durch die Einbeziehung externer Dienstleister sowie den Ausbau des eigenen Personals für zentrale Bereiche der Entwicklung sowie der Betreuung laufender IT-Produkte. Dies erfordert nicht nur die Gewinnung von neuen IT-Fachkräften, sondern auch das Wissen um neue Entwicklungen und Anwendungserfahrungen für interne Mitarbeitende in der IT. Zudem ist die Schaffung von Möglichkeiten für den Austausch und die Weiterbildung in verschiedenen Formen von besonderem Interesse für eine langfristige Bindung der Mitarbeitenden.

Durch die fortschreitende Digitalisierung hat sich der Wettbewerb um geeignete Fachkräfte verschärft. Laut dem Branchenverband bitkom e.V. fehlten 2022 über 137.000 IT-Fachkräfte [Bi22]. McKinsey gab allein für den Öffentlichen Sektor eine Zahl von 39.000 fehlenden IT-Fachkräften an [Mc23]. Der Öffentliche Dienst konkurriert dabei mit der Wirtschaft unter ungleichen Bedingungen. Beamtenrecht und Tarifvertrag setzen einen

⁷ Von 2014 nach 2015 verdoppelte sich die Anzahl der Asylanträge insgesamt auf 476.649 und erreichte mit 745.545 Anträgen 2016 den höchsten Wert seit Bestehen des Bundesamtes. Nach einem Rückgang in den folgenden Jahren lag die Zahl der Anträge 2022 bei 244.132. Bundesamt für Migration und Flüchtlinge: Das Bundesamt in Zahlen 2022, Nürnberg, 2023, S. 11

relativ strikten Rahmen bezüglich Einstellungsverfahren, Gehaltstrukturen und Karrierewegen im Öffentlichen Dienst [Sc18]. In einer Umfrage 2019 gaben nur 32 Prozent der Informatik-Studierenden an, eine Beschäftigung im Öffentlichen Dienst anzustreben. [Ne19]⁸ Dass dies auch für das BAMF eine Herausforderung ist, zeigen zahlreiche offene Stellen im Bereich der Software-Entwicklung und Hardware-Spezialisierung. Die Suche nach IT-Professionals stellt auch die Personalabteilung vor Probleme. Standard-Recruiting-Maßnahmen bei der Suche nach professionellen IT-Fachkräften greifen nicht mehr ausreichend und Alternativen müssen gefunden werden. Neben der Personalsuche über Online-Jobportale und Social-Media-Kanäle geht die IT-Abteilung des BAMF daher nun andere Wege, um das BAMF als möglichen Arbeitgeber für IT-Fachkräfte bekannter zu machen, Personal für das BAMF zu begeistern und gleichzeitig die bereits anwesenden IT-Fachkräfte zu unterstützen und zu halten.

Zeitgleich kam es zu zahlreichen Anfragen zur IT-Entwicklung und Digitalisierung des BAMF von Hochschulen und Forschungseinrichtungen, weshalb bereits 2017 eine IT-Forschungskontaktstelle eingerichtet wurde. Die Anfragen aus Wissenschaft und Forschung zeigten, dass ein grundsätzliches Interesse an den BAMF-IT-Themen besteht und ein Austausch hier für beide Seiten lohnenswert ist. Das Bundesamt entschied sich nicht mehr nur reaktiv zu handeln, sondern aktiv ein Netzwerk aufzubauen und zu pflegen. Daher wurde 2019, direkt in der IT-Abteilung ein eigener Bereich aufgebaut, der den Aufbau und die Gestaltung von Kontakten zu Wissenschaft und Forschung mit der Notwendigkeit neue Wege zur Gewinnung von Nachwuchskräften verbindet.

Das ins Leben gerufene IT-Research-Team ist im Referat für Digitalstrategie zusammen mit dem Chief Digital Office (CDO) angesiedelt. Das cross-funktionale Team besteht derzeit aus sechs Personen mit unterschiedlichen beruflichen Hintergründen und Erfahrungen aus Wissenschaft, IT, Wirtschaft, öffentlicher Verwaltung, Kommunikation und Relationsmanagement, die sich das breite Spektrum an Aufgaben aufteilen:

- Vernetzung, Kooperationen und Austausch mit Hochschulen und Forschungseinrichtungen
- Gewinnung von Fachkräften und Nachwuchs für die IT durch Employer-Branding und aktives Marketing
- Wissenstransfer, Innovationsimpulse und Erfahrungsaustausch mit Hochschulen und Forschung für IT-Projekte sowie die Möglichkeit von Forschungsprojekten zur Digitalisierung
- fachliche Unterstützung von IT-Bereichen und Mitarbeitenden des BAMF bei wissenschaftlichen und nichtwissenschaftlichen Publikationen
- Seminarbetreuung und Organisation in Kooperation mit Hochschulen

⁸ In einer nicht repräsentativen Umfrage von Masterstudierenden im IT-Bereich, die die Zollhof Betreiber GmbH für das IT-Research-Team des BAMF im Jahr 2022 durchführte, gab es vergleichbare Ergebnisse. Nur 31 Prozent gaben an, sie könnten sich vorstellen, nach dem Studium beim BAMF anzufangen.

- Betreuung studentischer Abschlussarbeiten wie auch Betreuung von internen Nachwuchskräften und Werkstudierenden
- Entwicklung und Betreuung eines Promotionsprogramms

3 Erste Schritte des IT-Research-Teams

Kaum nahm das Team Anfang 2020 seine Arbeit auf, brachte die Corona-Pandemie ungeahnte, neue Herausforderungen mit sich. Die verschiedenen Maßnahmen und Kontakteinschränkungen an Hochschulen und beim BAMF verlangten eine Umstellung auf alternative Formen der Vernetzung. Dies brachte Nach- und Vorteile mit sich: Zum einen mangelte es an spontanen Kontaktmöglichkeiten und einem Austausch vor Ort, zum anderen entstanden neue Formen der Veranstaltungsorganisation und Vernetzung mittels der Wahrnehmung von Terminen über digitale Tools. Insbesondere das Programm zur Betreuung von studentischen Abschlussarbeiten hat hiervon profitiert.

Trotz der zusätzlichen Herausforderung ist es gelungen, verschiedene Kooperationen aufzubauen:

- Forschungsprojekte zum einen des Max-Planck-Instituts für ethnologische Forschung Halle und der Martin-Luther-Universität Halle-Wittenberg, zum anderen mit der OHM-Hochschule Nürnberg inklusive einer Promotionsstelle
- U. a. Seminare mit der FAU Erlangen-Nürnberg (Digitalisierung, Kommunikation und Integration), der HTW Berlin (Blockchain-Infrastruktur Asyl), der Ostbayerischen Technischen Hochschule (OTH) Amberg-Weiden (Geschäftsprozesse in der Softwareentwicklung)
- Kooperation mit der Universität Hohenheim im Rahmen des AIDAHO, der AI & Data Certificate Hohenheim-Projektgruppe
- Rahmenorganisation sowie häufige inhaltliche Organisation verschiedener Veranstaltungen im Rahmen der CIC-Thementage in virtueller Form

Veranstaltungen und Forschungsprojekte werden vom IT-Research-Team in Zusammenarbeit mit verschiedenen internen IT-Fachbereichen betreut und ausgestaltet.

Ein Schwerpunkt der Arbeit des IT-Research-Teams ist die Betreuung von studentischen Projekt- und Abschlussarbeiten. Nach Entwicklung eines Konzepts wurden die IT-Fachreferate zur Themenfindung eingebunden, die interessierten Studierenden angeboten werden. Abschlussarbeiten schreiben Studierende, die über die Kooperationen vermittelt werden oder von selbst auf das BAMF zukommen. Außerdem nutzen auch interne Mitarbeitende die Möglichkeit, ihre Qualifikationen im IT-Bereich zu erweitern. Stand Sommer 2023 wurden bisher 6 Master- oder Bachelorarbeiten erfolgreich abgeschlossen, 6 sind aktuell in der Betreuung und weitere in Planung.

Auch wurde ein Konzept für ein Promotionsprogramm entwickelt und mit der Personalabteilung abgestimmt. Eine erste Stellenausschreibung blieb leider erfolglos, weshalb sich das IT-Research-Team aktuell in engem Austausch mit zahlreichen Professuren befindet, um Möglichkeiten und Anforderungen des Konzepts neu auszuloten.

4 Ein neuer Weg für Recruiting und Retention Management?

Angesiedelt in der IT-Abteilung in der Gruppe für die Digitale Transformation, bleibt der Bereich IT-Research vorrangig mit dem Fachbereich verknüpft. Ein modernes IT-Recruiting dagegen setzt voraus, dass die Kompetenzen für Personalvermittlung über den Fachbereich hinaus geteilt sind. Wenn eine Vermischung von Fach- und Human Resources-Kenntnissen vorausgesetzt wird, kann eine klassische Trennung von Zuständigkeiten in der Gemengelage von vielfältigen Interessen und rechtlich-organisatorischen Spielräumen hinderlich werden. Eine Bindung von potentiellen Mitarbeitenden und Fachkräften an das Bundesamt mittels Angeboten des IT-Research-Teams könnte sich hier als erfolgsversprechend erweisen.

Die Kooperationen und Veranstaltungen des IT-Research-Teams bewirken zudem, dass externes Wissen in die IT-Abteilung und zum Teil sogar darüber hinaus fließt und BAMF-IT-Projekte neue Impulse erhalten.⁹ Wissenstransfers werden erfolgsversprechend im Bereich der Personalgewinnung und der Gewinnung von Werkstudierenden für die BAMF-IT eingesetzt. Auch neue Grundlagen für das BAMF-Promotionsprogramm konnten so erarbeitet werden. Die alltägliche Suche nach Lösungsansätzen bei der Erweiterung von Spielräumen, um bedarfsgerecht agieren zu können, wird so maßgeblich unterstützt.

Jenseits der nachhaltigen positiven Effekte von Wissenstransfers aus den bestehenden Kooperationen entsteht die Frage, inwiefern öffentliche Verwaltung überhaupt mit modernen Methoden der IT-Recruiting mithalten könnte. Der Mangel an IT-Fachkräften ist allgegenwärtig und wird so in kostspieligen Karrierenetzwerken, Firmen-eigenen Programmen und anderen Förderprogrammen manifestiert. Dem öffentlichen Dienst fehlen hierfür Titelausschreibungen und Spielräume für einen geeigneten Einstieg. Daher verbleibt als ein erfolgsversprechendes Mittel der Wahl eine thematische Bindung von potentiellen IT-Fachkräften und interessierten Forschungseinrichtungen an BAMF-IT-Themen sowie die Öffnung und Nahbarkeit. Mehrwert für die Verwaltung, Wissenschaft und Forschung entsteht erst, wenn eine Behörde sich öffnet, für Zusammenarbeit bereit ist und das interne Wissen teilen lernt. Den Weg der Transparenz, der kurzen Wege, des direkten Drahtes in die Behörde hinein und aus ihr hinaus, wird das IT-Research-Team weiterhin verfolgen, um den Zugang von IT-Fachkräften und aktuellstem Fachwissen für das BAMF unkompliziert und zeitgemäß weiterzuführen.

⁹ Die von der IT-Research organisatorisch und z. T. inhaltlich gestalteten CIC-Thementage bieten dabei nicht nur dem IT-Bereich die Möglichkeit, sich mit externen Gästen und Referierenden zu Themen der Digitalisierung auszutauschen, sondern sind offen für Mitarbeitenden des BAMF über den IT-Bereich hinaus.

Literaturverzeichnis

- [Bi22] bitkom e.V: Trotz Krieg und Krisen: In Deutschland fehlen 137.000 IT-Fachkräfte, 2022, URL: <https://www.bitkom.org/Presse/Presseinformation/Deutschland-fehlen-137000-IT-Fachkraefte>, abgerufen 06.07.2023
- [Bu20] Bundesamt für Migration und Flüchtlinge: Digitalisierungsagenda 2022. Digitale Initiativen im Bundesamt für Migration und Flüchtlinge, Nürnberg, 2020
- [Bu23] Bundesamt für Migration und Flüchtlinge: Von der Dienststelle zur Bundesbehörde, 2023, URL: <https://www.bamf.de/DE/Behoerde/Chronik/Bundesamt/bundesamt-node.html>, abgerufen 06.07.2023
- [Ne19] Next: Public Beratungsagentur: Nachwuchsbarometer Öffentlicher Dienst. Gradmesser der Attraktivität des Öffentlichen Dienstes als Arbeitgeber bei Studierenden aller Fachrichtungen bundesweit, Berlin, 2019, URL: https://nextpublic.de/wp-content/uploads/Nachwuchsbarometer_Oeffentlicher_Dienst_2019.pdf, abgerufen 06.07.2019
- [Mc23] McKinsey & Company: Action bitte! Wie der öffentliche Sektor den Mangel an digitalen Fachkräften meistern kann, URL: <https://www.mckinsey.de/~media/mckinsey/locations/europe%20and%20middle%20east/deutschland/publikationen/2023-01-25%20it%20talent%20im%20public%20sector/action%20bittemckinsey.pdf>, abgerufen 06.07.2023
- [Sc18] Schütze, Julia: Warum dem Staat IT-Sicherheitsexpert:innen fehlen. Eine Analyse des IT-Sicherheitsfachkräftemangels im Öffentlichen Dienst, Berlin, 2018, URL: <https://www.stiftung-nv.de/sites/default/files/it-sicherheitsfachkraeftemangel.pdf>, abgerufen 06.07.2023

On which legal regulations is a public service based? Fostering transparency in public administration by using knowledge graphs

Leila Feddou¹, Maximilian Raupach¹, Felicitas Löffler² †, Samira Babalou¹, Jonas Hoyer¹,
Marianne Mauch¹, Birgitta König-Ries^{1,3,4}

Abstract: Knowledge about digitization in the public administration is complex and scattered. Information about legal regulations, methods, processes, APIs, metadata and data standards, registers, and terminologies are spread across different platforms. Hence, it is difficult for employees, developers, and decision makers to figure out what data standards, formats, and APIs are relevant for the digitization of a specific public service. Training administrative employees and IT companies requires to gather and link the required knowledge first in a well-structured and accessible manner. We address this need for more shared and transparent knowledge in the digitization of public services in Germany. We propose a first version of an ontology (GerPS-onto) for public administration instantiated by one example german public service. We utilize semantic modelling and Linked Data technologies to enable appropriate data and process descriptions that are readable for humans and machines. We also demonstrate how an existing process description of a public service can be linked to existing terminologies and evaluate the resulting ontology using domain-specific competency questions that are translated to SPARQL queries.

Keywords: Knowledge Graph; Ontology; Public Service; IT competence; Public Administration

1 Introduction and Related Work

A solid knowledge base is essential for a variety of tasks in public administration. The current digital transformation requires dedicated training and innovative concepts to enhance IT competences in public administration and its stakeholders. Besides trainings in using new software applications, public administration must be empowered in new arising tasks being relevant for artificial intelligence, e.g., annotation tasks to support the training of machine learning models or logic tasks to define and classify domain specific terms into hierarchical structures. Setting up such a knowledge base is challenging in the e-Government domain in Germany as data is scattered and spread across various platforms. For instance, digitizing public services requires to first understand and analyze the processes and data, including required activities, actors, and legal foundations. However, in Germany, there are different

† Corresponding author: felicitas.loeffler@tfm.thueringen.de

¹ Heinz Nixdorf Chair for Distributed Information Systems, Friedrich-Schiller-University, Jena, Germany

² Thuringian Ministry of Finance, Ludwig-Erhard-Ring 7, 99099 Erfurt, Germany

³ Michael-Stifel-Center for Data-Driven and Simulation Science, Jena, Germany

⁴ German Center for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Germany

approaches for processual descriptions. One concept is based on the legal foundations, the Federal Information Management - FIM⁵. It is the result of the analysis of legal resources (norm analysis) and consists of three basic building blocks: public service descriptions, data fields, and processes. The need to provide a broad spectrum of actors access to an integrated, harmonized view on complex information and the need to provide easy-to-use ways for information modeling, integration and harmonization is not unique to public administration. In many other fields, in the last decade, Knowledge Graphs (KGs) [Ho21] have been proposed as a promising approach to meet both needs. They encode the semantics of the data by representing it as real world concepts and their relations and have, among others, the following benefits: (1) Independent development by different actors with different expertise and world interpretation, (2) enabling data integration by interlinking concepts and relations, and (3) allowing machine readability and interpretation. In this paper, we use the term KG to describe information on the schema and instance level, while ontology is used to describe schema level information (e.g., the term “process” is a concept in our ontology, the concrete process of a specific service is part of the KG). Even if KGs are primarily instruments for developers, KG applications offer a wide range of opportunities for further stakeholders. For instance, semantic search and question answering approaches integrate KGs to allow searches beyond keywords (e.g., on what legal basis is a public service based?), which are relevant for decision makers, administrative staff, and interested citizens.

Related work. The importance of the semantic description of e-Government knowledge has been intensively discussed in the last decade [CM09, HTW10]. In the context of public administration, a very early ontology [Of12] was developed with the aim of semantically annotating law texts to ease transforming them into process models. The ontology combines concepts from different perspectives: law application (e.g., notification), administrative actions (e.g., application), and the law text (e.g., legal consequence). [SEW13] propose an ontology in UML format for describing public administration processes based on a specific use case (vehicle registration). It relates relevant concepts such as Public service, Activities, or Law. However, both approaches do not follow any agreed upon standard (e.g., FIM), [SEW13] is not available in a semantic common format, and [Of12] does not capture the process-oriented nature of public services (e.g., process steps). There are already initiatives at the European Union (EU) level to provide common vocabularies. The Publications Office of the EU provides a catalog of vocabularies (e.g., EuroVoc⁶) and models (e.g., the European Legislation Identifier (ELI) ontology⁷). However, semantic descriptions of the public administration domain are, to the best of our knowledge, still missing. This aspect is included in the the e-Government Core Vocabularies⁸ that was released by the the ISA² Programme⁹ and that includes a collection of upper level ontologies for basic

⁵ FIM, <https://fimportal.de>

⁶ EuroVoc, <https://op.europa.eu/en/web/eu-vocabularies/thesauri>

⁷ ELI, <https://op.europa.eu/en/web/eu-vocabularies/eli>

⁸ e-Government Core Vocabularies, <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/e-government-core-vocabularies>

⁹ ISA, https://ec.europa.eu/isa2/isa2_en/

concepts in public administration (e.g., public service or legal entity). Nevertheless, the proposed vocabularies offer only very high level concepts. There exist also country-specific initiative [HTP05, SB09, BC20, Be22] e.g., the Finnish Ontology for Public Administration Services (JUPO)¹⁰. The latter are mostly too specific and follow the properties and standards of the corresponding country. Overall, existing efforts are in most of the cases not process-oriented, do not follow a known standard, are either too granular or too general, and do not link to existing terminologies.

We aim to close this semantic gap and model the missing context knowledge for the digitization of public services in Germany. Following the FAIR data [Wi16] and 5 stars principles¹¹, we reuse existing terminologies and align our model with upper level ontologies. We also demonstrate the usage of these terminologies for an existing process to populate a first KG. In summary, our contribution is as follows: (1) A first version of an ontology for describing processes of German public services. (2) Alignment of our concepts to the e-Government Core Vocabularies, and a first identification of semantic gaps (e.g., missing concepts). (3) A KG for a specific public service: KG population by automatically transforming an XML-based format for processes into a semantic format. (4) Evaluation of the ontology using SPARQL queries corresponding to specific competency questions. In the following section, we introduce our proposed ontology, preliminary evaluation results, and a first demonstration of the populated KG of one process description of a digitized service. All results are publicly available in our GitHub repository¹².

2 GerPS-onto: An ontology for German public service processes

The primary goal of our proposed GerPS-onto¹³ is the semantic modeling of processes in the German public administration. For its development, we combined two methods: *Ontology Development 101* [No01] and *Conducting Literature Search for Artifact Reuse (CLeAR)* [Ca20]. At first, we defined the scope of the ontology by formulating possible competency questions. Next, we assessed existing ontologies for reuse. For this purpose, we used the following aspects from *CLeAR*, because the *Ontology Development 101* methodology contains only very generic specifications regarding reuse. We looked for suitable ontologies with specific keywords such as “BPMN”¹⁴, “Knowledge Graph”, “Public Administration”, “Public Service”, “Process”, “Ontology” in *Google Scholar*. This resulted in ten ontologies to be further analyzed. Afterwards, we applied a set of inclusion (e.g., based on BPMN) and exclusion (e.g., online unavailability) criteria to select candidates that are suitable for reuse. Based on these criteria the single result is the BPMN 2.0 based Ontology for Business Process Representation (BBO) [AAGK19]. To build our

¹⁰ JUPO, <https://finto.fi/jupo/en/>

¹¹ 5 stars open data, <https://5stardata.info>

¹² GerPS-onto GitHub repository, <https://github.com/fusion-jena/GerPS-onto>

¹³ GerPS-onto URL, <https://w3id.org/GerPS-onto/ontology/>

¹⁴ BPMN, <https://www.omg.org/spec/BPMN/>

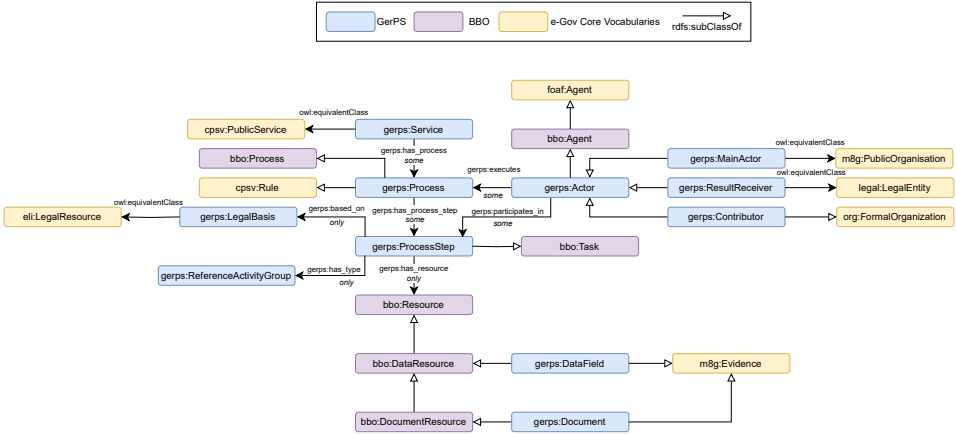


Fig. 1: Conceptual model of GerPS-onto

conceptual model, we identified important terms from our competency questions, precisely defined them, decided if they should be modeled as *property values*, *properties*, or *classes*, and mapped some classes and properties to BBO. For the latter step, we compared our competency questions to the ones provided by the selected ontology and analyzed the definitions of the BBO concepts and their conformity with our scope. The criteria, term definitions, and comparison of competency questions are available in GitHub. Figure 1 presents GerPS-onto. We reused the following BBO main classes: *Process*, *Task*, *Data Resource*, *Document Resource*, and *Agent* by defining them as superclasses of our added domain specific concepts (e.g., *gerps:Actor* *rdfs:subClassOf* *bbo:Agent*). We also reused two BBO properties (*gerps:has_resource* *rdfs:subPropertyOf* *bbo:has_resource* and *gerps:has_process* *rdfs:subPropertyOf* *bbo:has_process*). We added new properties to link the different domain specific concepts (e.g., *gerps:has_process_step*). The ontology was developed using Protégé¹⁵.

Ontology Alignment to e-Government Core Vocabularies. Apart from the BBO ontology, we also aligned our concepts to the e-Government Core Vocabularies of the EU. Therefore, we manually inspected a collection of ten available vocabularies to find classes corresponding to the GerPS-onto concepts. We then map them using two different properties (*owl:equivalentClass* and *rdfs:subClassOf*) to express either equivalence or generalization/specialization. In total, we mapped nine GerPS-onto classes (e.g., *gerps:Service* *owl:equivalentClass* *cpvs:PublicService*). However, for some other classes, we did not find any suitable candidates. This includes the core classes involved in modeling the actual business process reused from the BBO ontology (e.g., *gerps:ProcessStep*). For one of our new added GerPS-onto classes, (*gerps:ReferenceActivityGroup*), no candidate was found. It defines a concrete type of a specific process step, e.g., provide information or formally

¹⁵ Protégé, <https://protege.stanford.edu/>

examine the facts, and is a specific term of the FIM method. Overall, most of the existing vocabularies provide only high level concepts and do not model the underlying process.

Ontology Population and Evaluation. We populated the ontology with instances from one German public service¹⁶ to create a first KG. It is populated using sources in XML-based formats like *XProzess*¹⁷ and *XDatenfelder*¹⁸, for describing administrative processes and form fields involved in executing the specific service. The population is automatically done by parsing the XML-based files after manually mapping specific XML-nodes to our ontology classes. It is an example for the transformation of available domain knowledge (in non-semantic formats) into semantic formats. We evaluated the ontology by verifying its capability to correctly answer the competency questions expressed as SPARQL queries for one public service. The queries were correctly answered and are publicly available together with their corresponding results on GitHub. One example query with some results is shown in Figure 2 together with the concrete relations to other entities for one specific result.

```

1 PREFIX gerps: <https://w3id.org/GerPS-onto/ontology#>
2
3 SELECT DISTINCT ?Handlungsgrundlage
4 WHERE {
5   ?Prozess gerps:hat_prozessschritt ?Prozessschritt.
6   ?Prozess gerps:hat_leikaID "99006028261000".
7   ?Prozessschritt gerps:basiert_auf ?Handlungsgrundlage
8 }
9 # legal foundations for a specific public service
10 # with LeikaID "99006028261000"

```

Repository: (from context)

Query Execution Time: 130 ms / 0.13 s

Handlungsgrundlage

- http://www.gesetze-im-internet.de/muschg_2018/_29.html
- http://www.gesetze-im-internet.de/muschg_2018/_27.html
- https://www.gesetze-im-internet.de/arbschj/_5.html

Anzeige zur Beschäftigung ein...
Process

URI: <https://w3id.org/GerPS-onto/ontology...>
Type: Process

Quick search

name: Anzeige zur Beschäftigung einer schwangeren oder stillenden Frau bearbeiten

has_leikaID: 99006028261000

has_process_step: **Anzeige fachlich prüfen**
ProcessStep, Subprocess

based_on: **§ 29 MuSchG**
LegalBasis

has_process_step: Zusätzliche Informationen einholen, Erhebung der Geldbuße veranlassen, Zuständigkeit prüfen, Notwendigkeit von Überwachungsmaßnahmen prüfen, Anzeige empfangen, Korrektur anfordern, Anzeige fachlich prüfen, Überwachungsmaßnahmen einleiten, Anzeige zurückweisen, Anzeige formell prüfen, Über Erhebung einer Geldbuße entscheiden, Auf zuständige Behörde hinweisen

Fig. 2: SPARQL query to obtain the legal foundations for a given public service and some corresponding results

3 Conclusion

We introduced GerPS-onto, an ontology for modeling processes of German public services. For this purpose, we reused concepts of an existing ontology modeling business processes and linked some of our concepts to the upper level e-Government Core Vocabularies. We populated the ontology with one German public service. In future work, we aim to model

¹⁶ FIM portal, <https://fimportal.de/detail/L/99006028261000>

¹⁷ XProzess, <https://www.xrepository.de/details/urn:xoev-de:mv:em:standard:xprozess>

¹⁸ XDatenfelder, <https://www.xrepository.de/details/urn:xoev-de:fim:standard:xdatenfelder>

missing knowledge and to extend the ontology with more specific concepts (e.g., technical information about standards, registers, and architectures). We hope that GerPS-onto can serve as a starting point for a comprehensive KG of digitization in public administration. We envision others to contribute to this KG by modeling other public services and linking their model to GerPS-onto.

Bibliography

- [AAGK19] Annane, A.; Aussenac-Gilles, N.; Kamel, M.: BBO: BPMN 2.0 based ontology for business process representation. In: ECKM 2019. volume 1, pp. 49–59, 2019.
- [BC20] Bastola, R.; Campus, P.: Developing domain ontology for issuing certificate of citizenship of Nepal. *Journal of Information Technology*, 2(02):73–90, 2020.
- [Be22] Benaddi, H.; Laaz, N.; Kettani, E. E.; Hannad, Y.: Ontology Model for Public Services in Morocco Based on 5W1H Approach: PSOM-eGovMa. *Procedia Computer Science*, 198:429–434, 2022.
- [Ca20] Campos, P. M.C. et al.: Finding reusable structured resources for the integration of environmental research data. *Environmental Modelling & Software*, 133:104813, 2020.
- [CM09] Charalabidis, Y.; Metaxiotis, K.: Ontology-based management of e-Government knowledge. In: *Social and Political Implications of Data Mining: Knowledge Management in E-Government*, pp. 221–234. IGI Global, 2009.
- [Ho21] Hogan, A. et al.: *Knowledge Graphs. Synthesis Lectures on Data, Semantics, and Knowledge* 22. Springer, 2021.
- [HTP05] Hinkelmann, K.; Thönssen, B.; Probst, F.: Referenzmodellierung für E-Government-Services. *Wirtschaftsinf.*, 47(5):356–366, 2005.
- [HTW10] Hinkelmann, K.; Thönssen, B.; Wolff, D.: Ontologies for E-government. In: *Theory and Applications of Ontology: Computer Applications*. Springer Netherlands, Dordrecht, pp. 429–462, 2010.
- [No01] Noy, N.; McGuinness, D. L. et al.: *Ontology development 101*. Knowledge Systems Laboratory, Stanford University, 2001, 2001.
- [Of12] Off, T.: *Durchgängige Verfolgbarkeit im Vorfeld der Softwareentwicklung von E-Government-Anwendungen: Ein ontologiebasierter und modellgetriebener Ansatz am Beispiel von Bürgerdiensten*. Doctoral thesis, Universität Potsdam, 2012.
- [SB09] Savvas, I.; Bassiliades, N.: A process-oriented ontology-based knowledge management system for facilitating operational procedures in public administration. *Expert Systems with Applications*, 36(3, Part 1):4467–4478, 2009.
- [SEW13] Schminck, A.; Eid-Sabbagh, R.-H.; Weske, M.: eGovernment Process Knowledge Ontology - Business Process Knowledge Interdependencies in the Public Administration. In: *INFORMATIK 2013*. volume P-220 of LNI. GI, pp. 722–735, 2013.
- [Wi16] Wilkinson, M. D. et al.: The FAIR Guiding Principles for scientific data management and stewardship. *3(1):160018*, 2016.

Modularisierung von digitalen Weiterbildungsangeboten für die öffentliche Verwaltung

Konzeption und Perspektiven am Beispiel des eGov-Campus

Marc Egloffstein¹, Philipp Kuscher², Dirk Ifenthaler³

Abstract: Modularisierung beschreibt die Transformation von bestehenden Online-Kursen in eigenständige, kleinformate, kompetenzorientierte und flexible Lern-Module. Modularisierte Angebotsformate können als eine der Voraussetzungen für die erfolgreiche Implementierung von digitalen Lernangeboten und Micro-Credentials in der beruflichen Aus- und Weiterbildung angesehen werden. Dieser Beitrag beschreibt die Entwicklung und Anwendung eines Modularisierung-Frameworks für offene Online-Kurse am Beispiel des eGov-Campus. Dabei werden die strukturellen Rahmenbedingungen für modularisierte Online-Angebote, ein Vorgehensmodell zur Kurstransformation sowie exemplarische Umsetzungsaspekte skizziert. Der Beitrag schließt mit einer Diskussion von Herausforderungen aus den Bereichen Technik, Organisation und Didaktik und einem Ausblick auf die Perspektiven modularisierter Lernangebote für die Aus- und Weiterbildung für die digitale Verwaltung.

Keywords: Modularisierung, MOOC, digitale Weiterbildung, eGov-Campus, Micro-Credential

1 Problemstellung

Die Digitalisierung der öffentlichen Verwaltung wird in Deutschland bislang nur unzureichend umgesetzt. Neben technischen und organisatorischen Rahmenbedingungen wurden insbesondere Kompetenzdefizite auf Seiten der Mitarbeitenden als Ursachen identifiziert [Me20]. Angesichts der immer größer werdenden Lücke in Bezug auf „E-Kompetenzen“ in der öffentlichen Verwaltung [Ko22] wächst der Bedarf an adäquaten Bildungsangeboten. Skalierbare digitale Bildungsplattformen wie der eGov-Campus oder der KommunalCampus, die Inhalte zu Kernthemen der Verwaltungsdigitalisierung u.a. in Massive Open Online Courses (MOOCs) anbieten, gewinnen zunehmend an Bedeutung [HNB21].

MOOCs bezeichnen (zugangs-)offene, videobasierte Lernangebote, die entweder als „getaktete“, meist betreute Kurse von mehreren Wochen Dauer oder als

¹ Universität Mannheim, Lehrstuhl für Wirtschaftspädagogik - Technologiebasiertes Instruktionsdesign, L4, 1, 68161 Mannheim, egloffstein@uni-mannheim.de,

² Universität Potsdam, Lehrstuhl Public und Nonprofit Management, August-Bebel-Str. 89, 14482 Potsdam, kuscher@uni-potsdam.de

³ Universität Mannheim, Lehrstuhl für Wirtschaftspädagogik - Technologiebasiertes Instruktionsdesign, L4, 1, 68161 Mannheim, ifenthaler@uni-mannheim.de

Selbstlernumgebungen umgesetzt werden [Eg18]. Über den Hochschulkontext hinaus haben sich MOOCs inzwischen auch in Weiterbildung und Personalentwicklung etabliert [Ha23]. Niedrige Abschlussraten und offene Fragen hinsichtlich der Instruktionsqualität [Bo21] deuten allerdings auf Optimierungspotenziale hin. Auch passt das akademische Kursformat oft nicht sehr gut in den organisatorischen Rahmen von Weiterbildung und den beruflichen Alltag der Mitarbeitenden. Anknüpfend an aktuelle Bildungstrends wie Microlearning [ZW20] und Micro-Credentials [IBM16] erscheint eine Transformation von MOOCs hin zu kleinformatischen und flexiblen Lernangeboten daher zielführend und schlüssig. Während mögliche Effekte einer solchen "Modularisierung" bereits untersucht werden [Se22], existiert bis dato noch keine Methodik zur Umsetzung.

Vor diesem Hintergrund berichtet dieser Beitrag aus dem Modularisierungsprojekt im eGov-Campus. Ausgehend von der übergreifenden Fragestellung „Wie können bestehende MOOCs sinnvoll modularisiert werden, um didaktisch und organisatorisch stimmige Aus- und Weiterbildungsangebote für die digitale Verwaltung zu entwickeln?“ wird ein Modularisierungsframework aus der Instructional Design-Perspektive vorgestellt. Neben den strukturellen Vorgaben und einem (generischen) Vorgehensmodell werden erste Ansätze aus der konkreten Umsetzung berichtet. Der Beitrag schließt mit einer Diskussion von konzeptuellen und praktischen Herausforderungen der Modularisierung und einem Ausblick auf die Perspektiven modularisierter Lernangebote für die Aus- und Weiterbildung für die digitale Verwaltung

2 Digitale Aus- und Weiterbildung im eGov-Campus

Der eGov-Campus ist die Lernplattform für eGovernment in Deutschland. Als bundesländerübergreifendes Verbundprojekt wird er vom IT-Planungsrat gefördert, bis dato über die Föderale IT-Kooperation (FITKO) verwaltet und perspektivisch an die Deutsche Universität für Verwaltungswissenschaften angebunden. Aus technischer Sicht wird der eGov-Campus mit der MOOC-Plattform des Hasso-Plattner-Instituts und einem vorgeschalteten Portal realisiert. Die Kurse des eGov-Campus werden dezentral von unterschiedlichen Hochschulen erstellt, durchgeführt und gepflegt. Das Nutzungskonzept des eGov-Campus zielt auf eine übergreifende Einbindung der Inhalte in unterschiedlichen Aus- und Weiterbildungskontexten ab. Dabei werden die Kurse direkt von den beteiligten Hochschulen sowie von Bildungsträgern und spezialisierten Weiterbildungsanbietern bereitgestellt. Teilnehmende sind Bachelor- und Masterstudierende sowie Mitarbeitende aus dem Öffentlichen Sektor, die, auch unter Einbeziehung der jeweils zuständigen Personalentwicklung, auf die Angebote zugreifen können. Unter den deutschsprachigen MOOC-Providern nimmt der eGov-Campus damit eine Vorreiterrolle ein, was die Verankerung von MOOCs in Hochschul- und Weiterbildungscurricula angeht. So wurde bspw. der Kurs "Verwaltungsportale" der Universität Potsdam und des Stein-Hardenberg Instituts bereits mehrfach über ein Blended-Learning-Format als Wahlpflichtfach in das Bachelor-Studium "Public Management" der Hochschule Nordhausen integriert.

Um das skizzierte Nutzungskonzept weiter auszubauen, wurde im Herbst 2022 ein Projekt zur Modularisierung der eGov-Campus-Angebote initiiert. Modularisierung meint dabei die Transformation von bestehenden Kursen in eigenständige, kleinformative, kompetenzorientierte und flexible Lernangebote. Wesentliche Ziele der Modularisierung im eGov-Campus sind a) die erleichterte Integration von eGov-Campus-Inhalten in unterschiedlichen Szenarien v.a. in der Weiterbildung, b) die inhaltliche Differenzierung des Angebots zur Ermöglichung von Schwerpunktsetzungen und individuellen Lernpfaden, c) die curriculare Restrukturierung des Angebots, insbes. durch inhaltliche und organisatorische Einführungen zu den behandelten Themengebieten sowie d) die Vergabe von Micro-Credentials für Lernangebote unterhalb des Kurslevels.

3 Modularisierung im eGov-Campus

3.1 Struktur und Terminologie

Die Einführung einer bisher fehlenden einheitlichen Begriffsbasis markiert den ersten Schritt im Modularisierungsprojekt des eGov-Campus. Die Festlegung folgt einem pragmatischen Ansatz, der die Konsistenz mit etablierten Bezeichnungen, dem Vorgehen in Referenzprojekten sowie der wissenschaftlichen Terminologie sicherstellt. Abbildung 1 zeigt die verwendeten Konzepte in vereinfachter Modellierung [TB22]:

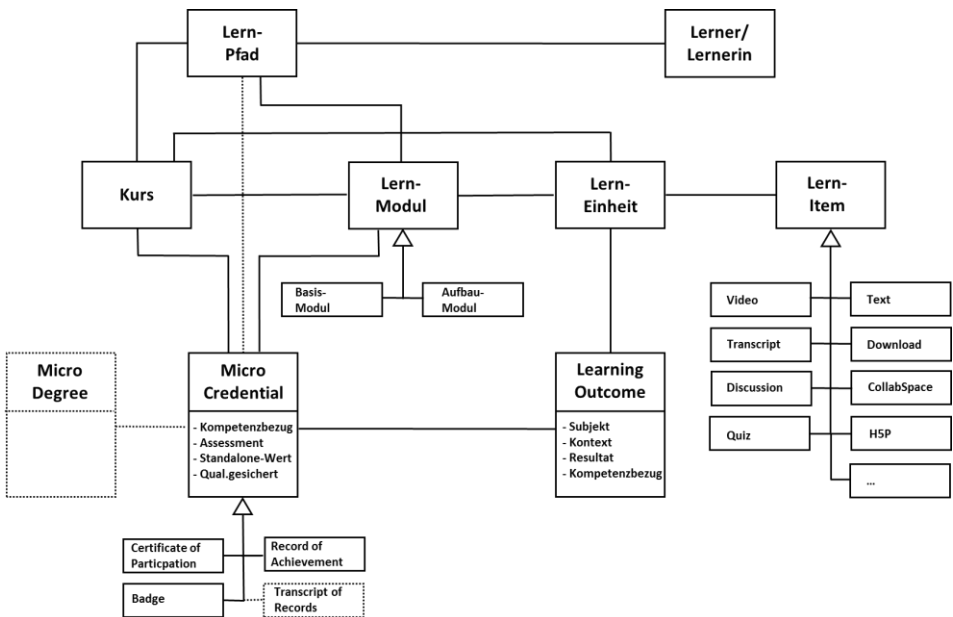


Abb. 1: Struktur der modularisierten eGov-Campus-Plattform

Zentrales Element des Modularisierungsansatzes ist das Konzept eines "Lern-Moduls", d. h. eines neu eingeführten Lernangebots unterhalb des Kurslevels (kürzere Dauer, praxisnahe Inhalte, weniger Aufwand). Bestehende Kurse sollen perspektivisch in flexible Modul-Serien aus verpflichtendem Basis-Modul (Einführung und Grundlagen) und mehreren Aufbau-Modulen (Spezialisierung und Vertiefung) überführt werden. Basis-Module bieten einen grundlegenden Überblick über das Fachgebiet und geben Orientierung für den weiteren Lernprozess, um initialen Schwierigkeiten entgegen zu wirken, Interesse bei den Lernenden zu wecken und Drop-Outs vorzubeugen. Analog zu den bestehenden Kursen können Lern-Module als eigene Einheiten auf der Plattform ausgewählt werden. Kurse wie Lern-Module setzen sich aus Lern-Einheiten (i.d.R. „Sections“ auf der Plattform) zusammen, die an Learning Outcomes gekoppelt sind, um die Kompetenzorientierung sicherzustellen. Lern-Einheiten wiederum bestehen aus verschiedenen Lern-Items (Features der Plattform), die beliebig je nach didaktischer Ausrichtung und Learning Outcomes variiert werden können (Videos, Texte, Übungen, gemeinsame Lernräume, etc.). Die Umsetzung der Learning Outcomes wird u.a. mit Assessments überprüft und analog zu den bestehenden Kursen über Micro Credentials zertifiziert, welche perspektivisch in neu zu entwickelnde Micro Degrees münden sollen.

3.2 Referenzmodell zur Kurstransformation

In Ergänzung des Referenzmodells zur Kurserstellung auf dem eGov-Campus [SHN21] werden die Aufgaben im Prozess der Modularisierung in einem heuristischen Vorgehensmodell zusammengefasst, das als erste „Blaupause“ in unterschiedlichen Transformationsprojekten adaptiert werden kann. Die Schritte beschreiben dabei keine starre Abfolge, sondern sollen Rücksprünge und Iterationen ermöglichen.

Lern-Module initial abgrenzen. Die initiale Abgrenzung der zukünftigen Lern-Module erfolgt auf Basis von inhaltlichen (konsistente Einheiten) und organisatorischen (angemessener Aufwand, einheitliche Vorgaben) Überlegungen heraus. Die Aufbau-Module haben einen Aufwand von 4 Stunden (0,5 Study Days) oder 8 Stunden (1 Study Day).

Basis-Modul („Grundlagen und Einführung“) spezifizieren. Die Learning Outcomes, Inhalte, Aktivitäten und (ggf.) Assessments für das Basis-Modul werden spezifiziert. Das Basis-Modul hat einen Aufwand von 8 Stunden (1 Study Day). Anpassungen für die Zielgruppe „Weiterbildung“ werden vorgenommen (Praxisbeispiele o.Ä.).

Learning Outcomes spezifizieren. Um die Kompetenzorientierung sicherzustellen, werden für die zukünftigen Aufbau-Module Learning Outcomes übernommen oder formuliert. Lern-Einheiten und Lern-Items spezifizieren

Lern-Einheiten und Lern-Items spezifizieren. Die für die Erreichung der Learning Outcomes notwendigen Lern-Items (Inhalte, Aktivitäten, Assessments) werden spezifiziert. Gleichzeitig werden die unterschiedlichen Lern-Einheiten (als Teile der Aufbau-Module) abgegrenzt.

Abhängigkeiten zwischen und innerhalb von Aufbau-Modulen prüfen und modellieren. Aufbau-Module stellen abgeschlossene Lernangebote dar, die einzeln absolviert werden können. Im Zuge der Modularisierung müssen daher (potenzielle) Abhängigkeiten zwischen Lern-Modulen und zwischen den Lern-Einheiten innerhalb eines Aufbau-Moduls geprüft werden. Dabei müssen Eingangsvoraussetzungen, Anschlussmöglichkeiten, Schwierigkeitsgrad, inhaltliche Tiefe vs. Breite sowie Fragen der Sequenzierung (lineare vs. parallele Bearbeitung) in den Blick genommen werden. Weiterhin müssen didaktische Abhängigkeiten innerhalb und zwischen den Aufbau-Modulen in den Blick genommen und ggf. angepasst werden (z.B. bei durchgängigen Fallstudien).

Lern-Module mit Metadaten beschreiben (Modulbeschreibung). Auf Basis der vorherigen Schritte werden die Lern-Module (Basis- und Aufbau-Module) mit Metadaten beschrieben, die u.a. für eine Zuordnung zu Lernpfaden sowie deren (perspektivische) Automatisierung notwendig sind. Eine entsprechende Vorlage wird bereitgestellt.

Modularisierung auf der Lernplattform umsetzen. Die Lern-Module werden auf der HPI-Plattform umgesetzt. Auf Basis der vorangegangenen Analysen werden Lern-Items kuratiert (ggf. hinzugefügt, geändert, oder weggelassen). Anpassungen der Inhalte werden vorgenommen, falls nötig.

Lern-Pfade spezifizieren. Ausgehend von vordefinierten Profilen innerhalb der Zielgruppe werden Lernpfade (als „Wege“ über verschiedene Lern-Module) spezifiziert und den Profilen (oder Schwerpunkten) zugeordnet.

Modularisierung validieren (Modularisierungsworkshops). Die Abgrenzung der Lern-Module, Zuordnung zu Profilen sowie die Spezifikation von Lern-Pfaden werden durch Teilnehmende und/oder externe Expertinnen und Experten validiert. Eine Einbindung von Nutzerinnen und Nutzern erfolgt so weit wie möglich.

Anpassungen vornehmen und Qualität sichern. Auf Basis der Rückmeldungen werden ggf. Anpassungen der Modularisierung vorgenommen. Eine fortlaufende Qualitätssicherung erfolgt auf Basis der Rückmeldungen aus der Evaluation. Entsprechende Items werden bereitgestellt.

3.3 Umsetzung

In Anlehnung an den Educational Design Research-Ansatz [MR21] wird die Modularisierung derzeit für drei bestehende Kurse prototypisch umgesetzt. Als unmittelbares Ergebnis dieser Pilotphase werden „Instructional Design Recommendations“ für Anbieter bestehender und Entwickler zukünftiger eGov-Campus-Kurse abgeleitet. Die neu eingeführten Lern-Module dienen zudem als Proof of Concept für den Modularisierungsansatz und das skizzierte Framework.

Ein besonderes Augenmerk liegt dabei auf der Gestaltung der neuen Basis-Module, die, aus Sicht des Kursdesigns, einen Perspektivenwechsel von der Hochschule hin zur

Weiterbildung implizieren, vor allem auch in der praxisnahen Sprache des Lernangebots. Über „Praxis Inputs“ (z.B. Interviews mit Praktikern und Praktikerinnen, praktischen Übungen sowie Videos, die Praxis-Dilemmata aufzeigen) soll die Relevanz der Inhalte unterstrichen werden. Zusätzliche „Guide-Videos“ am Anfang einer Lern-Einheit werden zur Begleitung durch das Basis-Modul erstellt und dienen dort als niedrigschwellige Lernnavigatoren. In einem kurzen Skript samt Executive Summary werden die wesentlichen Lerninhalte komprimiert und adressatengerecht aufbereitet. Abbildung 2 zeigt die Struktur des Basis-Moduls am Beispiel eines Pilotkurses der Universität Potsdam und des Stein-Hardenberg Instituts.

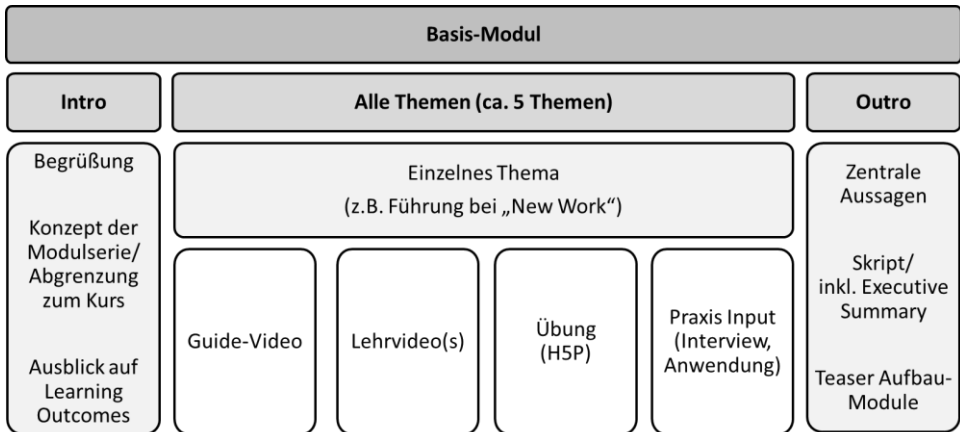


Abb. 2: Exemplarischer Aufbau eines Basis-Moduls

4 Diskussion und Ausblick

In diesem Beitrag wurde ein Modularisierungsframework für den eGov-Campus vorgestellt. Das Konzept erscheint geeignet, um die Qualität, Passung und Reichweite der Lernangebote zu verbessern. In der konkreten Umsetzung bestehen allerdings noch eine Reihe von Herausforderungen. Hinsichtlich der Technik ist bspw. noch nicht abschließend geklärt, wie die neuen Lern-Module auf Portal und Lernplattform implementiert und sinnvoll verknüpft werden können. In punkto Organisation stellen sich Fragen nach der Sequenzierung von Lern-Modulen und der Zuordnung zu inhaltlichen Profilen oder „Tracks“. Aus didaktischer Sicht sind vor allem die Gestaltung der (kleinformatigen) Lern-Module samt notwendiger Anpassungen im Inhalts- und Aufgabendesign sowie die zukünftige Abbildung von Kommunikation und Kooperation von Belang. Dabei besteht die besondere Herausforderung, einen Ausgleich zwischen dem wissenschaftlichen Anspruch der Lernangebote und den praktischen Anforderungen auf einer tendenziell niedrigschwelligem Ebene herzustellen, der den Bedürfnissen der Lernenden gerecht wird. Um die Umsetzung der angestrebten Ziele insbesondere im Hinblick auf Weiterbildung

und Transfer zu gewährleisten, müssen die modularisierten Lernangebote wissenschaftlich evaluiert werden, was zu Anpassungen und Feinjustierungen am Vorgehensmodell führen kann. Nicht zu unterschätzen ist bei alledem der Aufwand: Modularisierung ist keine simple Übernahme bestehender Kursinhalte und somit stets auch eine Ressourcenfrage.

Diesen Herausforderungen steht allerdings eine Reihe von Potenzialen gegenüber, die perspektivisch weit über die unmittelbaren Projektziele hinausweisen. So kann über modularisierte Angebotsformate eine Individualisierung von Lernangeboten erreicht werden, die über adaptive, KI-basierte Weiterbildungsportale für die öffentliche Verwaltung abgebildet werden kann. Individuelle Lernpfade über modularisierte Lernangebote können spezifisch zertifiziert werden, was, in Konsequenz, in neue Micro-Degrees oder Online-Studiengänge zum Themenbereich eGovernment münden kann. Modularisierte Weiterbildungsangebote mit entsprechenden Meta-Daten können darüber hinaus in übergreifende Vernetzungsstrukturen (Stichwort: Nationale Bildungsplattform) eingebunden und dadurch noch breiter verfügbar gemacht werden, was einen erleichterten Zugang für (zukünftige) Mitarbeitende in der öffentlichen Verwaltung bedeutet. In der Summe bietet die Modularisierung vielfältige Anknüpfungspunkte für die praxisgerechte Weiterentwicklung des eGov-Campus und die Aus- und Weiterbildung für die digitale Verwaltung insgesamt.

Literaturverzeichnis

- [Bo21] Bozkurt, A.: Surfing on three waves of MOOCs: An examination and snapshot of research in Massive Open Online Courses. *Open Praxis* 13, pp. 296–311, 2021.
- [Eg18] Egloffstein, M.: Massive open online courses in digital workplace learning: current state and future perspectives. In (Ifenthaler, D., Ed.): *Digital workplace learning: bridging formal and informal learning with digital technologies*. Springer, Cham, pp. 149–166, 2018.
- [Ha23] Hamori, M.: Self-directed learning in massive open online courses and its application at the workplace: Does employer support matter? *Journal of Business Research* 157, 2023.
- [HNB21] Halsbenning, S.; Niemann, M.; Becker, J.: Rettung der Verwaltungsdigitalisierung mittels Bildungsplattform? *Wirtschaftsinformatik & Management* 13/3, S. 174–183, 2021.
- [IBM16] Ifenthaler, D.; Bellin-Mularski, N.; Mah, D.-K. (Eds.): *Foundation of Digital Badges and Micro-Credentials. Demonstrating and recognizing knowledge and competencies*. Springer, Cham, 2016.
- [Ko22] Koddebusch, M., et al.: The Increasing e-Competence Gap: Developments over the Past Five Years in the German Public Sector. In (Fui-Hoon Nah, F; Siau, K., Eds.): *HCI in Business, Government and Organizations. HCII 2022. LNCS, vol. 13327*. Springer, Cham, pp. 73–86, 2022.

- [Me20] Mergel, I.: Kompetenzen für die digitale Transformation der Verwaltung. *Innovative Verwaltung* 4, S. 34–36, 2020.
- [MR21] McKenney, S.; Reeves, T.: Educational design research: Portraying, conducting, and enhancing productive scholarship. *Medical Education* 55/1, pp. 82–92, 2021.
- [Se22] Serth, S. et al.: Measuring the effects of course modularizations in online courses for life-long learners. *Frontiers in Education* 7, 2022.
- [SHN21] Spitzer, V.; Halsbenning, S.; Niemann, M.: Referenzmodell für MOOCs auf dem eGov-Campus. <https://www.zu.de/institute/togi/assets/pdf/symposium-2021/EG07-Spitzer-Halsbenning-Niemann-210614-Praesentation-eGov-Campus-V1.pdf>, Stand:19.05.2023.
- [TB22] Thurner, V.; Böttcher, A.: An Architectural Concept for Didactics that Integrates Technology into Teaching, Learning and Assessment. In (Auer, M. E.; Pester, A.; May, D., Eds.): *Learning with Technologies and Technologies in Learning*. LNNS, vol. 456. Springer, Cham, pp. 391–418, 2022.
- [ZW20] Zhang, J.; West, R. E.: Designing microlearning instruction for professional development through a competency based approach. *TechTrends* 64, pp. 310–318, 2020.

Wegbereiter: Konzeptuelle Arbeitsinstrumente zur Unterstützung interdisziplinären Denkens in der digitalen Verwaltungspraxis

Nassrin Hajinejad ¹, Alinka Rother², Jana Plomin³, Dorian Grosch⁴


Abstract: Der Kompetenzaufbau ist essenziell für die Verwaltungsdigitalisierung, da die Einführung neuer Technologien und Arbeitsweisen eine Anpassung der Fähigkeiten erfordert. Dies umfasst insbesondere das Verständnis der Zusammenhänge zwischen technischen Möglichkeiten, gesellschaftlichen Bedürfnissen, organisatorischen Strukturen und rechtlichen Rahmenbedingungen. In diesem Beitrag stellen wir konzeptuelle Arbeitsinstrumente für die Gestaltung öffentlicher IT vor und zeigen, wie diese Verwaltungsmitarbeitende dabei unterstützen können, ein interdisziplinäres Verständnis in der Projektarbeit zu entwickeln. Am Beispiel eines Nachhaltigkeits-Canvas, beleuchten wir das Potential dieser Arbeitsinstrumente für den Kompetenzaufbau und diskutieren die aufgabenorientierte Aneignung von Kompetenzen als komplementäre Ergänzung zu input- bzw. wissensorientierten Lernansätzen.

Keywords: Kompetenzaufbau, Öffentliche IT, Arbeitsinstrumente, Digitale Verwaltung, Interdisziplinäres Verständnis.

1 Einleitung

Die mit der digitalen Transformation des öffentlichen Sektors einhergehende Anforderung an eine Veränderung der Denkweisen, der Prozesse und der Fähigkeiten innerhalb der Verwaltung, hat eine Fokussierung auf neue Kompetenzen und ihren fortlaufenden Aufbau mit sich gebracht. Im Hinblick auf die Entwicklung vernetzter öffentlicher Leistungen im E-Government Kontext, die behördenübergreifende Zusammenarbeit erfordert, benötigen Verwaltungsmitarbeiter:innen „interdisziplinäres Wissen“ [Schu08], um IT-Lösungen im Zusammenhang mit organisatorischen, betriebswirtschaftlichen und gesellschaftlichen Entwicklungen zu denken und zu gestalten. Denn für öffentliche IT-Lösungen gilt: „Jede Lösung hat eine rechtliche, organisatorische, kulturelle, wirtschaftliche und technische Dimension, die in einem kontingenten, d. h. in

¹ Kompetenzzentrum Öffentliche IT am Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS,

Kaiserin-Augusta-Allee 31, 10589 Berlin, nassrin.hajinejad@fokus.fraunhofer.de,  <https://orcid.org/0000-0002-6587-3502>

² Kompetenzzentrum Öffentliche IT am Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, alinka.rother@fokus.fraunhofer.de

³ Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, jana.plomin@fokus.fraunhofer.de

⁴ Kompetenzzentrum Öffentliche IT am Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, dorian.grosch@fokus.fraunhofer.de

verschiedener Weise realisierbaren Zusammenspiel betrachtet werden müssen.“ [KrTH22, S.9]. In einem Literaturüberblick zu den Kompetenzbedarfen von Fach- und Führungskräften der öffentlichen Verwaltung, leiten die Autoren „Interdisziplinäres Verständnis“ [ScBr20, S.23] als Kernkompetenz für die berufliche Handlungsfähigkeit ab. Auch Studien zu den Rollenprofilen innerhalb der zukünftigen Verwaltung und den dazugehörigen Kompetenzen, stellen neben den grundlegenden technischen Kompetenzen, die es den Beschäftigten ermöglichen, digitale Werkzeuge im Arbeitsalltag zu nutzen, die zunehmende Bedeutung von soziotechnischen Kompetenzen (als fachliche Kompetenzen) heraus [OGZR16].

Werden diese Kompetenzen in den Inhalten der formalen Ausbildung stärker fokussiert, können künftige Verwaltungsmitarbeitenden davon profitieren. Eine besondere Herausforderung stellt jedoch der Kompetenzaufbau bei bereits in der öffentlichen Verwaltung Beschäftigten dar. Schließlich sind Verwaltungsmitarbeitende bereits heute mit neuen Problemstellungen konfrontiert, die ein vernetztes Denken erfordern. Zugleich aber lassen begrenzte (zeitliche) Ressourcen, Arbeitsbelastung, Fristen und tägliche Aufgaben oft wenig Raum für zusätzliches Lernen und Kompetenzentwicklung. Schon der Kompetenzbegriff selbst, der die *Anwendung* des Wissens akzentuiert, verweist auf die Ergebnisorientierung beim Lernen und damit auf die Anforderung an die Beschäftigten, berufliche Aufgaben bewältigen zu können [Schu08][Gill13]. So adressiert das Programm UPO⁵ (Unterstützung bei der Projektorganisation) die akuten Kompetenzbedarfe im Bereich IT- und Projektmanagement zur Umsetzung von OZG-Projekten in den Bundesbehörden. Neben der individuellen Beratung von Bundesbehörden stellt das UPO auch standardisierte Werkzeuge⁶ (z.B. Vorlage Projektplan oder Stakeholderanalyse) bereit.

Während also der Bedarf an zusätzlichen, interdisziplinär ausgerichteten Kompetenzen steigt, ist das gesonderte Erlernen genau dieser Kompetenzen für viele Beschäftigte der öffentlichen Verwaltung nur schwer und selten möglich. In diesem Beitrag adressieren wir die Herausforderung des Kompetenzaufbaus in der Verwaltungspraxis und fokussieren dabei auf interdisziplinäres Denken. Wir illustrieren am Beispiel von konzeptuellen Arbeitsinstrumenten für die öffentliche IT (ÖFIT-Wegbereiter), wie interdisziplinäres Denken und Gestalten in der Projektarbeit unterstützt werden kann. Abschließend besprechen wir den mit den Wegbereitern verfolgten Ansatz eines aufgabenorientierten Kompetenzaufbaus als komplementäre Ergänzung zu vorhandenen Bildungsangeboten im Bereich der digitalen Transformation des öffentlichen Sektors.

⁵ www.onlinezugangsgesetz.de/Webs/OZG/DE/themen/digitalisierungsprogramm-bund/upo/upo-node.html

⁶ www.onlinezugangsgesetz.de/Webs/OZG/DE/themen/digitalisierungsprogramm-bund/upo/upo-services/upo-services-node.html

2 ÖFIT-Wegbereiter – konzeptuelle Arbeitsinstrumente

Öffentliche IT-Lösungen und Dienstleistungen sind komplex, sie umfassen rechtliche, organisatorische, kulturelle, wirtschaftliche und technische Dimensionen. Sie sind mit Gestaltungszielen verbunden, die über gängige Zielsetzungen in privatwirtschaftlichen IT-Projekten hinausgehen. Neben der Gewährleistung von Nutzer:innenzufriedenheit, müssen IT-Lösungen für den öffentlichen Raum gesellschaftliche Bedarfe wie Gemeinwohl, Nachhaltigkeit und Teilhabe befriedigen. Damit erfordert die Gestaltung und Entwicklung Öffentlicher IT erweiterte Kompetenzen sowohl der Mitarbeitenden im öffentlichen Sektor als auch von IT-Dienstleistern aus der Privatwirtschaft wie z. B. GovTech-Unternehmen.

Das Kompetenzzentrum Öffentliche IT⁷ (ÖFIT), forscht in einem interdisziplinären wissenschaftlichen Team zu fortschreitenden Digitalisierung der Gesellschaft und erarbeitet Handlungsempfehlungen für Politik und Entscheidungsträger aus Verwaltung und Zivilgesellschaft (s. [GrWJ23, KrTH22, MHAB00]). Um die Verbreitung wissenschaftlicher Erkenntnisse zur Unterstützung von Digitalisierungs- und Innovationsprojekten in der öffentlichen Verwaltung weiter voranzutreiben und Entscheidungsträger auf verschiedenen Verwaltungsebenen zu erreichen, entwickelt das ÖFIT eigenständig nutzbare Arbeitsinstrumente, die als "ÖFIT-Wegbereiter" bezeichnet werden. Jeder Wegbereiter basiert auf wissenschaftlichen Publikationen des ÖFIT und ist im Unterschied zu von unternehmensorientierten Beratungshäusern angebotenen Materialien frei zugänglich. Die Wegbereiter adressieren u.a. Themen wie Technikgestaltung im kommunalen Kontext und nehmen Organisationsstrukturen und Arbeitsprozesse innerhalb der Verwaltung in den Blick. Im Zentrum der Wegbereiter steht stets die Perspektive der öffentlichen IT – also die Übersetzung gesellschaftlicher Werte in konkrete Anforderungen an Digitalisierungsprozesse und Lösungen.

Im Unterschied zu Methodensammlungen, die ursprünglich für den Unternehmenskontext entwickelt und im Nachgang auf den Verwaltungskontext zugeschnitten wurden (z.B. [Wiss00]), entstehen ÖFIT-Wegbereiter aus der Begleitung von Projekten innerhalb der öffentlichen Verwaltung. Sichergestellt wird der Praxisbezug durch eine Einbeziehung von Beschäftigten der öffentlichen Verwaltung bereits im Entstehungsprozess der Wegbereiter. Dies geschieht sowohl durch den bilateralen Austausch und qualitative Interviews (s. [GrWJ23]) als auch durch eine Überprüfung und Feinjustierung der Wegbereiter in Workshops mit Verwaltungsmitarbeiter:innen. Dieses Vorgehen gewährleistet einen engen Bezug zu den Arbeitsstrukturen und -prozessen in der öffentlichen Verwaltung und hat das Ziel, *tatsächliche* Bedarfe zu adressieren. Vor diesem Hintergrund gehen ÖFIT-Wegbereiter auf drei Bedarfe der digitalen Verwaltung ein: a) Die Adressierung von praxisbezogenen Herausforderungen, b) den eigenständigen und bedarfsorientierten Kompetenzerwerb im Arbeitsprozess sowie c) die Unterstützung eines interdisziplinären Verständnisses vom Beginn bis zum Abschluss des Projektes. Im

⁷ Kompetenzzentrum Öffentliche IT, gefördert vom Bundesministerium des Innern, für Bau und Heimat, www.oeffentliche-it.de

Folgenden illustrieren wir die Umsetzung der drei Anforderungen am Beispiel des ÖFIT-Wegbereiters *Nachhaltigkeits-Canvas*: *Ein visuelles Bewertungsinstrument* [GrWJ23].

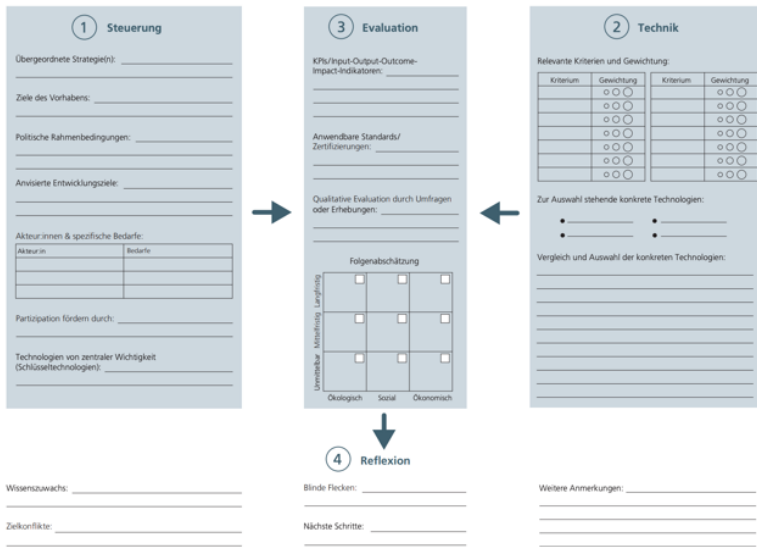


Abbildung 1: Ausschnitt des ÖFIT-Wegbereiter Nachhaltigkeits-Canvas. Die zur Strukturierung genutzten Konzepte werden im dazugehörigen Whitepaper erläutert [GrWJ23].

2.1 Der ÖFIT-Wegbereiter Nachhaltigkeits-Canvas

Wie alle Bereiche unserer Gesellschaft sieht sich auch die öffentliche Verwaltung mit grundlegenden und strukturellen Herausforderungen durch die großen Transformationen unserer Zeit konfrontiert: Nachhaltigkeit und Digitalisierung [JMSJ22]. Die öffentliche Hand hat hier die Chance, eine Vorreiterrolle einzunehmen: Eine gemeinwohlorientierte Gestaltung der öffentlichen Informations- und Kommunikationstechnologien gehört zu ihren zentralen Aufgaben. Eine Verknüpfung von Digitalisierung und Nachhaltigkeit im Rahmen eines konkreten Digitalisierungsvorhabens stellt Beschäftigte der öffentlichen Verwaltung vor die große Herausforderung, ihr Digitalisierungsprojekt (auch) anhand von Nachhaltigkeitskriterien zu gestalten⁸ - und erfordert damit zusätzliche interdisziplinäre Kompetenzen, möglicherweise das Erschließen eines neuen Themengebiets.

Das Nachhaltigkeits-Canvas (NC) ist ein Arbeitsinstrument, mit dem eine ganzheitliche Einschätzung eines Digitalisierungsprojektes mit Blick auf Nachhaltigkeit unterstützt werden soll. In seiner Form ist das NC (Abbildung 1) am Geschäftsmodell-Canvas [OsPi10], einem visuellen Instrument des Projektmanagements, angelehnt. Das NC nimmt

⁸ Green-IT-Strategie des IT-Planungsrates, www.it-planungsrat.de/fileadmin/beschluesse/2022/Beschluss2022-18_Green_IT_Strategie.pdf

die Verknüpfung von Digitalisierung und Nachhaltigkeit aus verschiedenen Perspektiven in den Blick und macht eine Zusammenführung dieser möglich: Auf der Steuerungsebene eines Projekts adressiert das NC die übergeordnete Strategie und konkrete Ziele ebenso wie politische Rahmenbedingungen und Aspekte der politischen Partizipation. Diese Perspektive führt er zusammen mit Kriterien an nachhaltige Technologie selbst - anhand ökologischer, ökonomischer und sozialer Dimensionen der Nachhaltigkeit. Zugleich richtet das NC den Blick auf die Technikfolgenabschätzung und macht gesellschaftliche, wirtschaftliche und umweltpolitische Auswirkungen des Vorhabens anhand von Indikatoren sichtbar. Durch die Bearbeitung des NC lässt sich erheben, an welchen Stellen sich ein Vorhaben am Idealtypus der wertebasierten Digitalisierung für nachhaltige Entwicklung orientiert und an welchen Stellen Verbesserungspotenzial besteht. Idealerweise nehmen Akteur:innen aus verschiedenen Stakeholdergruppen an diesem Prozess teil, um eine möglichst breite Perspektive abzudecken. Das Ausfüllen des NC kann blinde Flecken, wie bspw. ein vernachlässigtes Nachhaltigkeitskriterium, einen Zielkonflikt im Vorhaben oder bisher nicht erkannte Technikfolgen aufdecken (s. [GrWJ23]).

2.2 Projektarbeit durch interdisziplinäre Expertise unterstützen

Das NC unterstützt die Bearbeitung einer spezifischen Projektaufgabe durch eine vorgegebene Struktur. Damit ebnet dieser Wegbereiter den Einstieg in das Projekt, bietet in der Durchführung Orientierung anhand verschiedener Dimensionen und beleuchtet anschließend verschiedene Gestaltungsoptionen sowie deren mögliche Konsequenzen. Die grafische Struktur des NC kondensiert interdisziplinäre Expertise zum Thema Nachhaltigkeit und Digitalisierung und zeigt die unterschiedlichen und wechselseitig wirkenden Dimensionen niedrigschwellig, pragmatisch und ergebnisorientiert auf. Die Anwender:innen können sich die abstrakten Dimensionen erschließen, indem sie diese anhand des eigenen Projektes konkretisieren und damit die theoretischen Konzepte in Austausch mit dem eigenen Praxiswissen bringen. Damit weisen die Wegbereiter als Arbeitsinstrumente Anknüpfungspunkte zum sogenannten problembasierten Lernansatz auf: Problembasiertes Lernen nimmt „Probleme der Praxis als Ausgangs- und die Problemlösung zum Zielpunkt.“ [ScPh21, S.248]. Die Wegbereiter können insbesondere bei der Strukturierung von multidimensionalen Anforderungen unterstützen und damit die Problemlösefähigkeiten stärken. Um Theorie und Praxis an einem authentischen Problem zu integrieren, müssen Wegbereiter in unterschiedlichen Kontexten anwendbar sein und einen projektspezifischen Zugang ermöglichen. Im Falle des NC liegt bspw. das Ausfüllen ebenso wie das Priorisieren einzelner Dimensionen bei den Anwender:innen. Zudem kann das Canvas in unterschiedlichen Projektphasen eingesetzt werden: In der Entscheidungsfindung zu Beginn eines Vorhabens ebenso wie bei der späteren Reflektion und Kommunikation bereits getroffener Entscheidungen. Um der zunehmenden Bedeutung behörden- und sektorenübergreifender Kooperation im Kontext der digitalen Transformation gerecht zu werden, sind die Wegbereiter maßgeblich als kollaboratives Instrument angelegt und weisen damit Schnittstellen zum “kooperativen Lernen”

Inhalte entsprechend eines Baukastensystems nach individuellem Bedarf zur Verfügung stellen - wie bspw. vom Qualifizierungsprogramms Qualifica Digitalis empfohlen [ScBr20]. Mit Blick auf eine stärkere Vernetzung unterschiedlicher Lehrangebote gilt es zukünftig zu diskutieren, wie sich individuelle und stärker ineinandergreifende Lernpfade gemeinsam gestalten lassen.


Literaturverzeichnis

- [Digi00] Digitalakademie - Die Lernreisen.
www.digitalakademie.bund.de/DE/Unser_Angebot/Lernreisen/Lernreisen_node.html.
Stand: 2023-05-17
- [Egov00] eGov-Campus. URL <https://egov-campus.org/>. Stand: 2023-05-17
- [Gill13] Gillen, J.: Kompetenzorientierung als didaktische Leitkategorie in der beruflichen Bildung – Ansatzpunkte für eine Systematik zur Verknüpfung curricularer und methodischer Aspekte. In: bwp@ (2013), Nr. 24
- [GrWJ23] Grosch, D.; Wachsmann, D.; Jörs, V.: Wertebasierte Digitalisierung für nachhaltige Entwicklung im öffentlichen Sektor. Berlin : Kompetenzzentrum Öffentliche IT, 2023
- [JMSJ22] Joint Research Centre (European Commission) ; Muench, Stefan ; Stoermer, Eckhard ; Jensen, Kathrine ; Asikainen, Tommi ; Salvi, Maurizio ; Scapolo, Fabiana: Towards a green & digital future: key requirements for successful twin transitions in the European Union. LU : Publications Office of the European Union, 2022
- [Ki-c00] KI-Campus. <https://ki-campus.org/front>. Stand: 2023-05-17.
- [Komm00] KommunalCampus. www.kommunalcampus.net/. Stand: 2023-05-17.
- [KrTH22] Krenn, Karoline ; Tiemann, Jens ; Hajinejad, Nassrin: Ein Ko2mpass für IT im öffentlichen Raum. Berlin : Kompetenzzentrum Öffentliche IT, 2022
- [MHAB00] Mergel, I.; Haug, N.; Albrecht, V.; Brahimi, A. ; Edelman, N. ; Hajinejad, N.; Hölscher, I.; Plomin, J.: Erfolgreiche Innovationsfellowships in der Verwaltung umsetzen, S. 23
- [OGZR16] Ogonek, N. ; Greger, V. ; Zepic, R.; Räckers, M.; Becker, J. ; Krcmar, H.: Auf dem Weg zu einer innovativen Verwaltung: Rollen und Kompetenzen der Verwaltung im E-Government-Kontext : Gesellschaft für Informatik e.V., 2016.
- [OsPi10] Osterwalder, A.; Pigneur, Y.: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. 1. Aufl. Hoboken, NJ : Wiley, 2010
- [ScBr20] Schmeling, J.; Bruns, L. Kompetenzen, Perspektiven und Lernmethoden im digitalisierten Öffentlichen Sektor. Berlin : Fraunhofer FOKUS, 2020
- [Schu08] Schuppan, T.: Kompetenzen für vernetztes E-Government. In: eGov Präsenz Bd. 9 (2008), Nr. 1, S. 62–65
- [ScPh21] Schmohl, T. ; Philipp, T. (Hrsg.): Handbuch Transdisziplinäre Didaktik : transcript Verlag, 2021

- [ScZK00] Schmeling, J. ; Zouagui, J. ; Klessmann, J.: Fort- und Weiterbildung. In: Schröder, C. ; Schnurbus, L. ; Kumar, A. ; Lühr, H. (Hrsg.): Qualifica Digitalis: Strategien und Handlungsempfehlungen zur digitalen Qualifizierung in der öffentlichen Verwaltung. Bd. 5. Bremen : Kellner Verlag, S. 78–100
- [Wess05] Wessner, Martin: Kontextuelle Kooperation - Unterstützung kooperativen Lernens auf Basis des Kontextes : Gesellschaft für Informatik e.V., 2005.
- [Wiss00] Wissensspeicher: Handbuch Öffentliches gestalten. <https://citylab-berlin.org/de/wissensspeicher-handbuch/>. Stand: 2023-05-17.

(K)ein funktionierender Markt?

Nachfrage und Angebot für das Studium der Verwaltungsinformatik

Tim Pidun ¹, Silko Grellmann, Ludwig Schönthier, Pascal Thielemann,
Stefan Handke²

Abstract: Die Verwaltungsdigitalisierung in Deutschland braucht gut ausgebildetes Fachpersonal insbesondere im technischen Bereich, wobei es derzeit insbesondere an einem Anforderungskatalog oder einem Kompetenzprofil für die Ausbildung mangelt. Der Beitrag erhebt mittels strukturierter Reviews die in der Literatur zur Verwaltungsinformatik geforderten und in der Praxis angebotenen Kompetenzen zur Digitalisierung in der öffentlichen Verwaltung in verschiedenen Studiengängen in Deutschland. Neben übergeordneten Kompetenzen aus der Domäne der Informatik konnte eine Rangliste mit fachlichen Inhalten herausgearbeitet werden, die beide in der Kombination einen ersten Ansatz für ein Kompetenzprofil darstellen können. Des Weiteren werden für verschiedene Rollen in Organisationen Empfehlungen zur Schwerpunktsetzung der einzelnen Studiengänge gegeben.

Keywords: Kompetenzen, Verwaltung, Verwaltungsinformatik, Informatik, Studium

1 Einleitung

Im Bereich der Verwaltungsdigitalisierung in Deutschland mangelt es an fachlich qualifiziertem Personal insbesondere in informationstechnischen Aufgabenbereichen der Verwaltung (z. B. [Be16]; [Og20]; [SB20]). Heuberger führt dazu eher organisational-technische und persönlich-motivationale Hindernisse als behindernde Faktoren des digitalen Organisationswandels insbesondere im öffentliche Dienst an [He20], nicht aber vordergründig persönlich-fachliche Kompetenzen der Mitarbeitenden. Diese werden allerdings in einer Fraunhofer-Metastudie [SB20] in der Form von „persönlicher beruflicher Handlungsfähigkeit“ gefordert, ebenso neben generellen Kompetenzen organisational-technischer Art wie z.B. Suchen, Kommunizieren, Produzieren und Analysieren in digitalen Umgebungen. Zur Definition eben dieser durch eine Aus- oder Weiterbildung bereitgestellten beruflichen Kompetenzen (oft synonym gebraucht: e-Kompetenzen) existiert zwar ebenso Literatur, meist aber auch eher in der Form eher übergeordneter Ansätze. Becker et al. definieren z.B. insgesamt 19 verschiedene Referenzrollen der öffentlichen Verwaltung mit IT-Bezug [Be16], verweisen aber im

¹ Hochschule für Technik und Wirtschaft Dresden, Fakultät Informatik/Mathematik, Friedrich-List-Platz 1,
01069 Dresden, tim.pidun@htw-dresden.de,  <https://orcid.org/0000-0003-1331-1732>

² Hochschule für Technik und Wirtschaft Dresden, Fakultät Wirtschaftswissenschaften, Friedrich-List-Platz 1,
01069 Dresden, stefan.handke@htw-dresden.de.

unbeachtet bleiben. Dies bestätigen auch 2023 Räckers und Halsbenning, die feststellen, dass die Bedeutung technischer IT-Kompetenz seitdem auf etwa 77 % gestiegen ist, aber immer noch etwa 67 % der befragten Beschäftigten angaben, dass Personen mit dieser Kompetenz kaum auf dem Markt verfügbar sind. [RH23, S. 10]. Bezüglich der e-Kompetenzen fordern sie das Vorhandensein der Fähigkeit, operativ und konkret IT bei der Aufgabenerfüllung zu nutzen und zielgerichtet einzusetzen. Außerdem wird die Fähigkeit, die IT-Landschaft einer öffentlichen Verwaltung strategisch zu gestalten, zielgerichtet die richtige IT auszuwählen und in die Prozesse und Abläufe einzubinden [RH23, S. 3] betont, um damit eher gestaltende Rollen einzunehmen. Im Ergebnis liegt derzeit in der Literatur aber noch kein konkretes fachlich-soziotechnisch ausgerichtetes Kompetenzprofil für Mitarbeitende in der Verwaltungsdigitalisierung vor. Es existieren zwar einige einschlägige grundständige Studiengänge der Verwaltungsinformatik in Deutschland, die auf dieses Profil hinweisen, diese sind allerdings von ihrer Zielsetzung und ihrem Anwendungsbereich in der Regel eher organisatorisch-verwaltungswissenschaftlich mit starkem Bezug auf die Länderebene ausgelegt [HP22]. Mithin kann unterstellt werden, dass insbesondere für die Rollen, die unabhängig von Körperschaftsebenen die fachlich-soziotechnische Organisation mit IT gestalten sollen, keine sichtbaren speziellen Bildungsangebote an Hochschulen bestanden [If14, S. 3] und auch derzeit offenbar immer noch bestehen [RH23, S. 12]. Dies zeigt sich auf der Abnehmerseite durch einen Mangel an Mitarbeitenden in der Verwaltungsdigitalisierung.

Im Lichte dieser Versorgungslücke identifiziert dieser Beitrag, welche Kompetenzen und konkreten Inhalte zur Digitalisierung in der öffentlichen Verwaltung einerseits in der Literatur gefordert werden und andererseits derzeit in Deutschland an Hochschulen gelehrt werden. Dies geschieht mit besonderem Augenmerk auf geforderte resp. vermittelte fachlich-soziotechnische Aspekte. Er gleicht im Rahmen eines strukturierten Reviews Realität und Projektion der Ausbildung für die digitale Verwaltung miteinander ab und bildet eine Empfehlungsgrundlage für ein technisches Kompetenzprofil für Mitarbeitende in der Verwaltungsdigitalisierung.

2 Untersuchung

2.1 Methodik und Durchführung

Obwohl der Anwendungsbereich der Disziplin Verwaltungsinformatik in Deutschland mit dem Öffentlichen Dienst offenbar zu sein scheint, existieren hierzu in Literatur und Praxis verschiedene Definitionen, Synonyme und Auslegungen, insbesondere in der Kombination der Domänen Verwaltung und Informatik. Insofern wurde im Rahmen dieser Untersuchung die Methode des strukturierten Reviews nach Fettke und Loos [Fe06] angewendet, um Klarheit über den Untersuchungsgegenstand und die darauf anwendbaren Hochschulausbildungen zu gewinnen. Das Review besteht hierbei aus fünf Schritten: der Problemformulierung, Explikation aller Suchterme und -schritte und anschließender Literatursuche, statistische Aufbereitung und Interpretation sowie Präsentation (S. 260).

Im Zuge des Reviews wurden von Oktober 2022 bis März 2023 frei zugängliche Webseiten (z.B. Google Scholar), Fachdatenbanken (z.B. SpringerLink) und -suchmaschinen wie z.B. das Datenbank-Informationssystem DBIS der Sächsischen Landes- und Universitätsbibliothek (SLUB) und der Hochschule für Technik und Wirtschaft Dresden (HTWD) verwendet. Darüber hinaus wurden gezielt dedizierte Webseiten untersucht, in denen einschlägige Informationen vermutet wurden, wie z.B. Seiten von Bundesministerien, Ländern und Kommunen. Die Problemstellungen, die dabei bearbeitet wurden, waren:

- Welche Kompetenzen werden aus der Literatur gefordert (auf der Nachfrageseite)?
- Welche Hochschulausbildungen der Verwaltungsinformatik existieren (auf der Angebotsseite) und
- zu welchem Grad werden dabei fachlich-soziotechnische Kompetenzen gefordert/gelehrt?

2.2 Review der Nachfrageseite

Die prinzipielle Herangehensweise der Recherche nach in der Literatur geforderten Kompetenzen war, die Disziplin der Informatik in einem verwaltungsrechtlichen Kontext zu betrachten, da im Rahmen der Fragestellung vor allem die technischen Kompetenzen gesucht werden, die in der speziellen Domäne der Verwaltung benötigt werden. Die einzelnen Suchen nebst verwendeter Terme sind in Tabelle 1 zusammengefasst.

Suchterme		Ergebnis
Auf Webseiten des Bundes, der Länder und der Kommunen		
Verwaltungsinformatik eGovernment Public Service		keine direkten Bezüge zu Kompetenzen oder Hochschulausbildung
Alle Webseiten/Datenbanken		
Verwaltungsinformatik		Beschreibung der Disziplin; keine direkten Bezüge zu Kompetenzen oder Hochschulausbildung
Verwaltungsinformatik	Anforderungsprofil Profil	Ohne nutzbare Erkenntnisse
Informatik	Anforderungsprofil Profil	Ein nutzbares Ergebnis
Verwaltungsinformatik	Kompetenzen, Befähigung Fertigkeit, Fähigkeit	154 Literaturquellen; keine direkten Bezüge zu Kompetenzen oder Hochschulausbildung
Rechtswissenschaft	Informatik	Ohne nutzbare Erkenntnisse

Tab. 1: Suchterme Review Nachfrageseite

an diesen Stellen und in diesen Ebenen der originäre zukünftige Arbeitsbereich eines Verwaltungsinformatikers erwartet wird, doch konnten hier weder ein klares Anforderungsprofil noch notwendige Kompetenzen herausgearbeitet werden. Im Anschluss wurde die Suche auf Synonyme wie „eGovernment“ und „Public Service“ ausgedehnt, doch auch hier wurden keine auswertbaren Ergebnisse erzielt.

Danach wurde die Suche auf alle in 2.1. erwähnten Datenbanken und Webseiten ausgedehnt. Zunächst wurde dort direkt nach dem Term „Verwaltungsinformatik“ gesucht, dabei aber hauptsächlich Informationen über Studiengänge gefunden, die zu unspezifisch für die detaillierte Suche nach Anforderungen waren. Die Kombination mit dem Suchterm „Anforderungsprofil“ oder „Profil“ erzielte keine brauchbaren Ergebnisse. Danach wurde der Begriff „Kompetenz“ mit mehreren Synonymen als Basis mehrerer Suchvarianten zusammen mit „Verwaltungsinformatik“ verwendet, was in 154 Literaturquellen resultierte. Da die Verwaltungsinformatik auch als ein Unterthema der Rechtswissenschaften gesehen werden kann [Be18], [Di19], wurde in verschiedenen Datenbanken der Term „Rechtswissenschaften“ nebst „Informatik“ gesucht, was allerdings ohne weitere fachliche Erkenntnisse blieb. Die dezidierte Suche nach Publikationen in die Richtung eines Anforderungsprofils für die Informatik im Generellen resultierte in einer Quelle, dem Hochschul-Bildungs-Report 2019 [St19].

Die insgesamt 154 sinnhaften Literaturquellen wurden anhand von Titel und Abstract stichprobenartig untersucht, um die Aussagefähigkeit hinsichtlich der gesuchten Kompetenzen zu überprüfen. Insgesamt waren diese wenig aussagekräftig hinsichtlich geforderter Kompetenzen oder Bezügen zur Hochschulbildung. Beispielhaft sei hier der Beitrag „Neue Kompetenzen für Allianzen und Verwaltung“ [HRH21] genannt, der sich zwar mit den Voraussetzungen zur Arbeit in der digitalen Welt beschäftigt, dies geschieht aber nur im Hinblick auf Anforderungen in privatrechtlichen Unternehmen. Dabei wird vor allem die prinzipielle Forderung einer „digitalen Affinität“ beschrieben. Auch die weiteren relevanten Stellen der Publikation beschäftigen sich zwar mit den Grundlagen digitaler Verwaltung, Anwendungsfällen von Technologien wie z.B. einer Blockchain im Verwaltungskontext oder eGovernment, aber dies durchgehend ohne Bezug zu konkreten Kompetenzen. Desgleichen zeigte sich in einem weiteren Beitrag „Der große Klick“ [Me22], der Probleme der Verwaltungsstruktur im Allgemeinen behandelt und ebenso nicht auf Kompetenzen des zukünftigen Personals eingeht. Im Hochschulbildungsreport [St19]. sind zwar wichtige technische Kompetenzen für zukünftige Arbeitnehmer in der Domäne der Digitalisierung und der Informatik beschrieben, sie beziehen sich allerdings nur auf Arbeitgeber im unternehmerischen Umfeld und nicht explizit auf die Verwaltung. Dabei bleiben die dort geforderten generellen Kompetenzen (Tabelle 2) zunächst ebenso technisch unkonkret auf einem höheren Anwendungs-Abstraktionsniveau.

Technological Skills	Digital Citizenship Skills	Classic Skills
Komplexe Datenanalyse	Digital Literacy	Problemlösungsfähigkeit
Smart-Hardware-/ Roboterentwicklung	Digitale Interaktion	Kreativität
Webentwicklung	Kollaboration	Unternehmerisches Handeln und Eigeninitiative
Nutzerzentriertes Designen (UX)	Agiles Arbeiten	Adaptionsfähigkeit
Konzeption und Administration vernetzter IT-Systeme	Digital Learning	Durchhaltevermögen
Blockchain-Technologie- Entwicklung	Digital Ethics	
Tech Translation		

Tab. 2: Kompetenzen („Skills“) des Hochschulbildungsreport 2019

2.3 Review der Angebotsseite

Um derzeit vermittelte Kompetenzen eines/einer Verwaltungsinformatiker/-in aus der Angebotssicht genauer zu erfassen, wurde das zweite Review der in Deutschland angebotenen grundständigen Studiengänge mit öffentlichem Zugriff auf die darin angebotenen Module (die im Bologna-Prozess die eigentlichen Kompetenzen vermitteln, [Ke06]) per reiner Online-Recherche durchgeführt. Da es nach einer ersten Durchsicht einige Studiengänge gibt, die das Thema der Verwaltung und Informatik behandeln, nicht aber direkt „Verwaltungsinformatik“ benannt sind, wurden verschiedene Synonyme und Alternativen verwendet. Die Übersicht der einzelnen Suchanfragen nebst verwendeten Terme sind dabei in Tabelle 3 zusammengefasst.

Suchterme	Ergebnis
Verwaltungsinformatik eGovernment	8 originäre Studiengänge
Digitale Verwaltung IT-Management Public Governance Öffentliche Verwaltung Public Informatics	5 ähnliche Studiengänge

Tab. 3: Suchterme Review Angebotsseite

insgesamt acht angebotene Studiengänge an der Hochschule Hof, der Hochschule für öffentliche Verwaltung Rheinland-Pfalz, der Hochschule des Bundes für öffentliche Verwaltung mit zwei Standorten Brühl und Münster, der Hochschule Hannover, der Hochschule für Technik und Wirtschaft Dresden, der Hochschule für Wirtschaft und Recht Berlin, der Hochschule Schmalkalden und der Technischen Hochschule Wildau gefunden. Die Hochschule für Polizei und öffentliche Verwaltung Nordrhein-Westfalen sowie die Hochschule Rhein-Waal bieten zwar zum Redaktionsschluss der Recherche auch entsprechende Studiengänge an, hier waren allerdings die einzelnen Module nicht detailliert abruf- und auswertbar. Diese Studiengänge konnten daher in der weiteren Untersuchung nicht berücksichtigt werden. Des Weiteren existieren zwar entsprechende weiterbildende Masterstudiengänge „eGovernment“ z.B. an der Universität Koblenz und Münster, die aber ebenso nicht berücksichtigt wurden, da hier nach einer ersten Durchsicht unterstellt wird, dass sie auf bereits bestehende Kompetenzen aufbauen und daher eher kein grundlegendes Wissen in der Disziplin der Verwaltungsinformatik vermitteln.

Neben diesen „originären“ Studiengängen wurden stark ähnlich lautende Studiengänge nach einer ersten Durchsicht anhand der Schnittmenge zu den Lehrinhalten von Verwaltung und Informatik ausgewählt. Beispiele hierfür sind unter anderem die Studiengänge „Digitale Verwaltung“, „IT-Management“ oder „Public Governance“. Diese Unterscheidung ist wichtig, da diese „ähnlichen“ Studiengänge auch andere oder weitere Module aufweisen, die in einem Studiengang „Verwaltungsinformatik“ u.U. nicht berücksichtigt werden. In dieser Suchreihe wurden ebenso Studiengänge ohne detaillierte Modulbeschreibungen und Masterstudiengänge, z.B. „Public Governance“ der Hochschule Meißen und Fortbildungszentrum aus der weiteren Untersuchung ausgenommen. Im Ergebnis wurden fünf Studiengänge identifiziert und zur detaillierter betrachtet: Die Studiengänge „Digitale Verwaltung“ der Hochschule Meißen und Fortbildungszentrum und der Hessischen Hochschule für öffentliches Management und Sicherheit, die Studiengänge „IT-Management“ sowie „Verwaltungsdigitalisierung- und Informatik“ der Hochschule Harz und der Studiengang „Öffentliche Verwaltung“ der Hochschule für Wirtschaft und Recht Berlin.

Im weiteren Verlauf der Auswertung wurden die Module der einzelnen Studiengänge einer genaueren Auswertung unterzogen, um den fachlich-soziotechnischen Anteil der vermittelten Kompetenzen herauszuarbeiten. Dabei wurden die Module im Ergebnis in die Modularten „Verwaltung“, „Informatik“, „Schnittstelle“ und „Grundlagen“ kategorisiert. Ein Modul wird dabei der Verwaltung zugeordnet, wenn der Themenschwerpunkt nicht technisch ist und essentielle Bestandteile der Verwaltung abbildet. Dafür kann die Definition zu Verwaltungswissenschaften herangezogen werden, in der die „(öffentliche) Verwaltung und ihre Umwelt, politikwissenschaftliche, soziologische, rechtswissenschaftliche und wirtschaftswissenschaftliche Aspekte“ betrachtet werden [Ma05]. Bei der Modulart der Informatik handelt es sich um Kompetenzen, die „die Wissenschaft von der systematischen Darstellung, Speicherung, Verarbeitung und Übertragung von Daten“ betreffen, wobei „besonders die automatische Verarbeitung mit Digitalrechnern betrachtet wird“ [Du06]. Als ein Schnittstellen-Modul wird kategorisiert, wenn es nicht klar abgrenzbar zwischen den Schwerpunkten Verwaltung und Informatik

ist und eine Verbindung zwischen beiden Arten bildet. Dies kann beispielsweise der Fall sein, wenn sich vordergründig mit Verwaltungswissenschaften, aber im digitalen Kontext auseinandergesetzt wird. Lässt sich ein Modul zu keiner der drei oben genannten Arten zuordnen, so wird es als Grundlagen-Modul kategorisiert. Die Module sind essentiell für das Studium, bringen für die Recherche in die Richtung spezieller fachlich-soziotechnischer Kompetenzen in der der Domäne der Verwaltungsinformatik aber keinen Mehrwert.

Eine Auswahl an untersuchten Modulen, deren Zuordnung in eine Art und die lehrenden Hochschulen und Studiengängen ist in den Tabellen 4 und 6 abgebildet; die komplette Auswertung aller 191 bzw. 146 Module wurde aus Platzgründen diesem Beitrag nicht hinzugefügt. Tabellen 5 und 7 werten dann aber aus dieser Gesamtschau die Anzahl der einzelnen Modularten in den originären und ähnlichen Studiengängen aus.

Modul	Hochschule
Modulart Verwaltung	
IT-Vergabe	Hochschule für Wirtschaft und Recht Berlin
IT-Recht	Hochschule des Bundes für öffentliche Verwaltung
Verwaltungsökonomie	Hochschule für Technik und Wirtschaft Dresden
Öffentliches Finanzmanagement	Hochschule Hannover
Modulart Informatik	
Datenbanken I	Hochschule für Wirtschaft und Recht Berlin
Software Engineering	Hochschule Schmalkalden
Programmierung I	Hochschule für Technik und Wirtschaft Dresden
Algorithmen und Datenstrukturen	Hochschule Hof
Modulart Schnittstelle	
E-Government I	Hochschule für Wirtschaft und Recht Berlin
Informationssicherheit und Awareness	Technische Hochschule Wildau
Informationsmanagement	Hochschule für Technik und Wirtschaft Dresden
IT-Sicherheit und Datenschutz	Hochschule Schmalkalden
Modulart Grundlagen	
Englisch	Hochschule Hof
Statistik	Hochschule für Technik und Wirtschaft Dresden

Projektmanagement	Hochschule Hannover
Wissenschaftliches Arbeiten	Technische Hochschule Wildau

Tab. 4: Beispiele für Module und Modularten originärer Studiengänge

Modulart	Anzahl	Anteil
Verwaltung	53	27,75%
Informatik	78	40,84%
Schnittstelle	22	11,52%
Grundlagen	38	19,90%
Gesamt	191	100,00%

Tab. 5: Anzahl und Anteil der Modularten originärer Studiengänge

Modul	Studiengang	Hochschule
Modulart Verwaltung		
Governance	IT-Management	Hochschule Harz
Digitalisierungsrecht	Digitale Verwaltung	Hessische Hochschule für öffentliches Management und Sicherheit
Allgemeines Verwaltungsrecht	Öffentliche Verwaltung	Hochschule für Wirtschaft und Recht Berlin
Öffentliche Finanzwirtschaft	Verwaltungsdigitalisierung- und Informatik	Hochschule Harz
Modulart Informatik		
Grundlagen digitaler Systeme	Digitale Verwaltung	Hochschule Meißen und Fortbildungszentrum
IT-Architektur	IT-Management	Hochschule Harz
Datenbanksysteme	Digitale Verwaltung	Hessische Hochschule für öffentliches Management und Sicherheit
Softwaretechnik	Verwaltungsdigitalisierung- und Informatik	Hochschule Harz
Modulart Schnittstelle		
E-Government	Digitale Verwaltung	Hochschule Meißen und Fortbildungszentrum
Operatives IT-Management	IT-Management	Hochschule Harz
Anforderungs- und Schnittstellenmanagement	Digitale Verwaltung	Hessische Hochschule für öffentliches Management und Sicherheit

Modul	Studiengang	Hochschule
IT-Sicherheit, vernetzte Verwaltungen und E-Government-Standards	Verwaltungsdigitalisierung- und Informatik	Hochschule Harz
Modulart Grundlagen		
Sozialkompetenz	IT-Management	Hochschule Harz
Gesellschaft und Technik	Digitale Verwaltung	Hessische Hochschule für öffentliches Management und Sicherheit
Sozialwissenschaften	Verwaltungsdigitalisierung- und Informatik	Hochschule Harz
Juristische Falllösungstechniken	Öffentliche Verwaltung	Hochschule für Wirtschaft und Recht Berlin

Tab. 6: Beispiele für Module und Modularten ähnlicher Studiengänge

Modulart	Anzahl	Anteil
Verwaltung	63	43,15%
Informatik	25	17,12%
Schnittstelle	21	14,38%
Grundlagen	37	25,34%
Gesamt	146	100,00%

Tab. 7: Anzahl und Anteil der Modularten ähnlicher Studiengänge

In den originären Studiengängen der Verwaltungsinformatik liegt wie zu erwarten war, der Fokus auf den Informatikmodulen, der Durchschnitt liegt hier bei ca. 41%. Der Anteil der Verwaltungsmodule hingegen beträgt im Mittel 28%. Die Schnittstellenmodule tragen mit 12% zum Modulangebot bei. Die restlichen Module der Verwaltungsinformatik-Studiengänge werden als Grundlagen kategorisiert. Im Gegensatz dazu ist das prinzipielle Verhältnis der Informatik-Module zu den Verwaltungsmodulen in den ähnlichen Studiengängen genau umgekehrt (17% zu 43%). Hier liegt der Fokus stark auf Verwaltungsinhalten, zudem variiert der Anteil der Informatik stark innerhalb der Studiengänge. Der Anteil der Schnittstellenmodule ist ungefähr gleich (14%) wie bei den originären Studiengängen.

Legt man nun aus der Sicht der Identifikation der fachlich-soziotechnischen Kompetenzen die Annahme zugrunde, dass diese eher in den Modulen der Art „Informatik“ zu finden sind, müssen diese Gegenstand erneuter näherer Betrachtung werden. Daher wurden in einer dritten Auswertung alle Informatik-Module aller Studiengänge auf ihren Namen und weitestgehend auf fachliche Gemeinsamkeiten überprüft sowie hinsichtlich der Anzahl an Vorkommnissen in eine Rangfolge einsortiert. Dies geschieht unter der Annahme, dass

Modul	Anzahl	Rang
Datenbanken	11	1
Software Engineering	10	2
Projekte (und -management)	8	3
Betriebssysteme	7	
Informationssicherheit	7	4
Programmierung	7	
E-Government	5	
Grundlagen der (Verwaltungs-) Informatik	5	5
Finanzen	4	
Webentwicklung	4	6
Geschäftsprozesse	3	
Rechnernetze	3	7

Tab. 8: Anzahl der Module der Modulart Informatik und Rang

Die meist genannten Module sind mit Abstand solche mit den Lehrinhalten zu Datenbanken und Software Engineering. Auch Projekte/Projektmanagement und im Folgenden Betriebssysteme, Programmierung, Webentwicklung, Geschäftsprozesse und Rechnernetze sind Lehrinhalte, die in den meisten Studiengängen angeboten werden.

3 Ergebnis und Diskussion

Die Untersuchung insgesamt ergab kein direkt sichtbares Anforderungsprofil oder eine dedizierte Übersicht an Kompetenzen, die eine Hochschulausbildung als Verwaltungsinformatiker/-in beinhalten sollte. Die einzelnen Reviews ergaben, dass auf der Nachfrageseite im Kontext der öffentlichen Verwaltung recht generell „persönliche berufliche Handlungsfähigkeit“ und eine „digitale Affinität“ gefordert wird, ohne aber direkt Bezug zu konkreten fachlich-soziotechnischen Kompetenzen zu nehmen. Aus der Domäne der Informatik werden zwar konkretere Kompetenzen („Future Skills“) beschrieben, die aber leider genauso untechnisch bleiben. Auf der Angebotsseite ergab das Review allerdings, dass eine Rangfolge der wichtigsten fachlich-soziotechnischen Kompetenzen in Form von gelehrten Modulen der Studiengänge der Verwaltungsinformatik abbildbar ist.

Allen diesen gefundenen Kompetenzen ist aber zu eigen, dass sie konkrete Bildungsinhalte sind, die fachlich-logisch den einzelnen Anwendungskompetenzen des

Hochschulbildungsreports zugeordnet werden können, die allerdings aus dem Bereich der privatrechtlichen Unternehmen stammen. Daher müssten diese Skills zunächst für die Domäne der öffentlichen Verwaltung adaptiert werden und könnten dabei in einem nächsten Schritt um die darauf anwendbaren fachlich-soziotechnischen Kompetenzen erweitert werden, um insgesamt eine Grundlage für einen Kompetenzrahmen zu bilden.

Bei den Studiengängen mit ähnlicher Ausrichtung ist wie zu erwarten war das Verwaltungswissen im Vergleich zu Informatikmodulen wichtiger. Daraus ergibt sich aber leider auch eine fachliche Unschärfe: Einerseits werden im Hochschulmarketing neben den originären auch die ähnlichen Studiengänge äquivalent als geeignete Ausbildung für die Digitalisierung der öffentlichen Verwaltung beworben. Damit wären unabhängig vom konkreten Studium die Absolventen nach ihrem Abschluss direkte Konkurrenten für die gleichen Stellen. Je nach Studiengangsart werden aber sehr unterschiedliche fachlich-soziotechnische Kompetenzen gelehrt, was zu einer deutlichen Schwerpunktsetzung je nach Studiengang führt. Mithin ist zu empfehlen, Studiengänge einerseits als eher verwaltungsorientiert mit Bezug zur Digitalisierung zu positionieren (ähnliche Studiengänge) und dazu ausdrücklich komplementäre Studiengänge anzubieten, die eher informatikorientiert mit Bezug auf die Domäne der öffentlichen Verwaltung ausgerichtet sind (originäre Studiengänge). Die Absolventen der einzelnen Studiengänge können so entsprechend besser auf die Rollen direkt in der öffentlichen Verwaltung oder als Mitarbeiter bei z.B. Softwaredienstleistern, die die öffentliche Verwaltung betreuen, vorbereitet werden.



Literaturverzeichnis

- [Be16] Becker, J. et al.: eKompetenz. NEGZ, Berlin, 2016.
- [Be18] Begerow Beratungsgesellschaft: Rechtswissenschaft. Was ist Rechtswissenschaft?, 2018 <https://www.rechtswissenschaft-verstehen.de/rechtswissenschaft-grundlagen/rechtswissenschaft>, Stand: 22.2.23.
- [Di19] Disterer, G.: Was ist Verwaltungsinformatik? In (Schmid, A. Hrsg.): Verwaltung, eGovernment und Digitalisierung. Springer Fachmedien Wiesbaden, Wiesbaden, S. 41–51, 2019.
- [Du06] Duden Verlag: Duden Informatik A - Z. Fachlexikon für Studium, Ausbildung und Beruf. Dudenverl., Mannheim, Leipzig, Wien, Zürich, 2006.
- [Fe06] Fettke, P.: State-of-the-Art des State-of-the-Art. WIRTSCHAFTSINFORMATIK 4/48, 2006.
- [He20] Heuberger, M.: Digitaler Organisationswandel. In (Klenk, T.; Nullmeier, F.; Wewer, G. Hrsg.): Handbuch Digitalisierung in Staat und Verwaltung. Springer VS, Wiesbaden, Heidelberg, S. 587–598, 2020.

- Digitalisierung der Verwaltung. FIF-Kommunikation 3, S. 22–28, 2022.
- [HRH21] Halsbenning, S.; Räckers, M.; Hünemohr, H.: Neue Kompetenzen und neue Allianzen für Staat und Verwaltung – Aus- und Weiterbildung in Zeiten fortschreitender Digitalisierung. In (Seckelmann, M.; Brunzel, M. Hrsg.): Handbuch Onlinezugangsgesetz. Potenziale - Synergien - Herausforderungen. Springer, Berlin, Heidelberg, S. 423–440, 2021.
- [If14] IfG.CC: Aktuelle Ausprägung sowie Gestaltungsmöglichkeiten der E-Government-Aus- und Fortbildung von Fach- und Führungskräften der Verwaltung. Management Summary, Potsdam, 2014.
- [Ke06] Keller, H.-J.: Die Modularisierung und der Bologna-Prozess. Beiträge zur Lehrerbildung 24 (2006) 3, S. 303-314. Beiträge zur Lehrerbildung 24, 2006.
- [Ma05] Machura, S.: Politik und Verwaltung. VS Verlag für Sozialwissenschaften, Wiesbaden, 2005.
- [Me22] Mengers, C. et al.: KOMKIS Analyse. Leipzig University, Leipzig, 2022.
- [Og20] Ogonek, N. et al.: E-Kompetenzen. In (Klenk, T.; Nullmeier, F.; Wewer, G. Hrsg.): Handbuch Digitalisierung in Staat und Verwaltung. Springer VS, Wiesbaden, Heidelberg, S. 611–622, 2020.
- [RH23] Räckers, M.; Halsbenning, S.: E-Kompetenzen. In (Klenk, T.; Nullmeier, F.; Wewer, G. Hrsg.): Handbuch Digitalisierung in Staat und Verwaltung. Springer VS, Wiesbaden, S. 1–13, 2023.
- [SB20] Schmeling, J.; Bruns, L.: Kompetenzen, Perspektiven und Lernmethoden im digitalisierten öffentlichen Sektor. Qualifika Digitalis, Berlin, 2020.
- [St19] Stifterverband für die Deutsche Wissenschaft e. V.: Hochschul-Bildungs-Report 2020 - Jahresbericht 2019.
<https://www.hochschulbildungsreport2020.de/download/file/fid/163>,
Stand: 10.05.2023.

Öffentliche Infrastruktur -
Interoperabilität und
Standardisierung der digitalen
öffentlichen Beschaffung

Nutzung von elektronischen Katalogformaten in öffentlichen Beschaffungsverfahren

Cedric Pauken¹, Andreas Schmitz ², Maria A. Wimmer ³


Abstract: Die öffentliche elektronische Beschaffung sieht sich derzeit mit einer Vielzahl an Herausforderungen wie einer unstrukturierten pre-award Phase, einem nicht regulierten Übergang zwischen pre- und post-award und einer Vielzahl von Barrieren für kleine und mittlere Unternehmen (KMU) konfrontiert. Ziel dieses Beitrags ist es, zu analysieren, inwieweit die Nutzung elektronischer Kataloge zu einer Bewältigung aktueller Herausforderungen in öffentlichen Beschaffungsverfahren beitragen kann. Dazu wird Design Science Research als übergeordnetes Forschungsparadigma verwendet, um im Rahmen des Rigor Cycles mit einer Literaturanalyse das wissenschaftliche Fundament des Beitrags abzusichern und im Rahmen des Relevance Cycles den Bezug zum relevanten Anwendungskontext deutscher Beschaffungsverfahren zu gewährleisten. Praxisbeispiele aus Europa zeigen, dass die identifizierten Herausforderungen durch elektronische Kataloge erfolgreich bewältigt werden können. Ihre Nutzung ist daher empfehlenswert. Als Lösungsbaustein für die Nutzung elektronischer Kataloge in Deutschland wird im Design Cycle ein UBL-Syntaxmapping zwischen dem Peppol pre-award Catalogue und dem deutschen post-award BMEcat mit positivem Ergebnis durchgeführt.


Keywords: öffentliche Beschaffungsverfahren, elektronische Beschaffung, elektronische Kataloge, Katalogstandards, UBL-Syntaxmapping, Peppol, BMEcat

1 Einleitung

Die öffentliche Beschaffung umfasst den Vorgang der Planung, Beschaffung, Lieferung und Bewertung von Waren und Dienstleistungen, die von öffentlichen Stellen erworben werden [UF09] und lässt sich in die Phasen vor (pre-award) und nach (post-award) der Zuschlagserteilung (award) einteilen [Bo17], [Pr19]. Der Beschaffungsvorgang erfolgt gemäß eines öffentlichen Vergabeverfahrens, welches bei Schwellenwerten ab 428.000€ für Waren und Dienstleistungen oder 5.350.000€ für Bauleistungen durch die EU-Richtlinien 2014/24/EU und 2014/25/EU reguliert wird [Eu19]; werden diese Schwellenwerte nicht überschritten, so gelten die deutsche Verordnung über die Vergabe öffentlicher Aufträge (VgV) und die Vergabe- und Vertragsordnung für Bauleistungen

¹ Universität Koblenz, Forschungsgruppe E-Government, Universitätsstraße 1, Koblenz, 56070, cpauken@uni-koblenz.de

² Universität Koblenz, Forschungsgruppe E-Government, Universitätsstraße 1, 56070 Koblenz, andreaschmitz@uni-koblenz.de, <https://www.uni-koblenz.de/agvinf/>,  <https://orcid.org/0000-0001-9535-4001>

³ Universität Koblenz, Forschungsgruppe E-Government, Universitätsstraße 1, 56070 Koblenz, <https://www.uni-koblenz.de/agvinf/>, wimmer@uni-koblenz.de,  <https://orcid.org/0000-0002-8460-1027>

[Bu16], [De19]. Die öffentliche elektronische Beschaffung sieht sich derzeit mit einer Vielzahl an Herausforderungen wie dem Bedarf innovativer Rechtsvorschriften [Am13], einer unstrukturierten pre-award Phase, einem nicht regulierten Übergang zwischen pre- und post-award mit zahlreichen Medienbrüchen und einer Vielzahl an Barrieren für kleine und mittlere Unternehmen (KMU) konfrontiert [MBK20], [IT21], [Eu21], [Lo15]. Die Nutzung elektronischer Kataloge (siehe § 27 VgV) kann dazu beitragen, den administrativen Aufwand innerhalb des Beschaffungsprozesses für alle beteiligten Akteure zu reduzieren, Medienbrüche zu vermeiden und die Lücke zwischen pre- und post-award zu füllen, indem nachnutzbare Katalogdaten bereits im pre-award definiert und dann über weitere Prozessschritte hinweg genutzt werden können (siehe Kapitel 5). Elektronische Kataloge werden bereits erfolgreich im Rahmen des EHF-Beschaffungsstandards in Norwegen und des MePA-Standards in Italien verwendet und werden derzeit auch als europaweiter Peppol-Standard ausgearbeitet.

Ziel dieses Beitrags ist es, die aktuellen Herausforderungen des öffentlichen elektronischen Beschaffungsvorgangs zu analysieren und zu untersuchen, inwieweit die Nutzung elektronischer Kataloge dazu beiträgt, diese Herausforderungen zu beheben. Dazu wird die sogenannte „Three-Cycle-View“ der Design Science Research [He04], [He07] verwendet, um im Rigor Cycle eine Literaturanalyse als wissenschaftliches Fundament zu erarbeiten (Kapitel 3, 4 und 5) und im Relevance Cycle den Anwendungskontext im Projekt KleBe.digital zu erörtern und abzusichern (Kapitel 4 und 5). Auf Basis der Ergebnisse dieser iterativ durchgeführten Kreisläufe wird im Design Cycle ebenfalls inkrementell und iterativ ein UBL-Syntaxmapping zwischen dem Peppol pre-award Catalogue und dem BMEcat konzipiert, das den administrativen Aufwand reduziert sowie einen strukturierten Übergang zwischen pre- und post-award ermöglicht (Kapitel 6). Abschließend werden in Kapitel 7 Handlungsempfehlungen für Wirtschaftsteilnehmer, insbesondere KMUs, und öffentliche Stellen formuliert.

2 Forschungsdesign

Als übergeordnetes Forschungsdesign wird Design Science Research (DSR) angewendet, um die Entwicklung des Datenmappings zwischen dem Peppol pre-award Catalogue und dem deutschen post-award BMEcat mit einem wissenschaftlichen Fundament sowie dem relevanten Anwendungskontext absichern zu können. Dazu werden die drei Design Science Research Cycles iterativ angewendet [He04], [He07]. Aus dem in der Einleitung vorgestellten Kontext ergibt sich folgende Forschungsfrage: *„Welche Herausforderungen können im öffentlichen elektronischen Beschaffungsvorgang durch die Nutzung von elektronischen Katalogstandards bewältigt werden?“*

Das wissenschaftliche Fundament wird im Rahmen des Rigor Cycles durch eine strukturierte Literaturanalyse gemäß Webster und Watson erarbeitet [WW02] und durch graue Literatur wie Gesetzestexte und EU-Richtlinien ergänzt. Hinzu kommen interne Dokumente der Projekte „Kooperationsprojekt zur Digitalisierung der Beschaffung“ und

„Kleinteilige Beschaffung standardisiert.digital.vernetzt“. Gemäß der Methodik von Rowley und Slack wird wissenschaftliche Literatur über Suchmaschinen wie Google Scholar und Literaturdatenbanken wie IEEE gesucht und durch sukzessives Scannen des Titels, des Abstracts, des Fazits und dann des gesamten Beitrags als relevant identifiziert [RS04]. Als Ergebnis dieser iterativen Analyse im Rigor Cycle [He07] werden zuerst allgemeine Herausforderungen im elektronischen Beschaffungsprozess (Lücken zwischen pre- und post-award) identifiziert, welche dann näher analysiert werden, um zu untersuchen, inwieweit elektronische Kataloge einen Teil der formulierten Herausforderungen bewältigen können. In diesem Kontext wird auch die Relevanz des Beitrags durch den Relevancy Cycle der DSR [He07] abgesichert, in welchem die Projektarbeit in KleBe.digital, Peppol Spezifikationen und auch die Analyse relevanter Katalogstandards iterativ und in Bezug auf ihren Anwendungskontext diskutiert worden sind. Durch die iterative Durchführung beider Zyklen kann die rigorose und relevante Datenbasis für den Design Cycle [He07] geschaffen werden, in welchem ein konzeptionelles UBL-Syntaxmapping zwischen dem Peppol pre-award catalogue und dem post-award BMEcat durchgeführt werden konnte, um die im Rahmen der Rigor und Relevancy Cycles identifizierte Lücke zwischen pre- und post-award im öffentlichen Beschaffungsprozess zu schließen. Das Mapping wurde inkrementell im Projekt „Kleinteilige Beschaffung standardisiert.digital.vernetzt“ durchgeführt und iterativ diskutiert und verifiziert.

3 Theoretische Grundlagen

Für das theoretische Fundament werden verschiedene Konzepte aus der aktuellen wissenschaftlichen Literatur wie E-Procurement, Interoperabilität und Standardisierung sowie E-Catalogue entlang der Literaturanalyse nach [He07] erläutert.

Öffentliche Beschaffungsverfahren umfassen gemäß Uyarra und Flanagan den Vorgang der Planung, Beschaffung, Lieferung und Bewertung von Waren und Dienstleistungen, die von öffentlichen Stellen erworben werden [UF09]. Der Beschaffungsvorgang erfolgt gemäß eines öffentlichen Vergabeverfahrens, welches bei Schwellenwerten ab 428.000€ für Waren und Dienstleistungen oder 5.350.000€ für Bauleistungen durch die EU-Richtlinien 2014/24/EU und 2014/25/EU reguliert wird [Eu19]; werden diese Schwellenwerte nicht überschritten, so gelten die deutsche Verordnung über die Vergabe öffentlicher Aufträge (VgV) [Bu16] und die Vergabe- und Vertragsordnung für Bauleistungen [De16]. Zu Beginn des Vergabeverfahrens kann auf Wunsch der zuständigen Behörde die Verwendung elektronischer Kataloge erfolgen, um gemäß § 27 VgV standardisierte Beschreibungen der zu beschaffenden Waren oder Dienstleistungen zu erhalten. Elektronische Kataloge bieten öffentlichen Verwaltungen gemäß der World Bank die Möglichkeit, eine Liste von Waren und Dienstleistungen zu erhalten, die bereits im Voraus anhand von vorab definierten Kriterien als relevant eingestuft worden sind [Wo19]. Auch müssen zu Beginn Zuschlagskriterien festgelegt werden, anhand derer ein Angebot durch die öffentliche Stelle ausgewählt und ein Zuschlag erteilt wird. Diese

müssen gemäß EU-Richtlinien dem Gleichbehandlungsgrundsatz folgen und Bietern eine relative Gewichtung aller Zuschlagskriterien bereitstellen [Eu14a], [Eu14b].

Elektronische Beschaffungsverfahren, im Englischen auch als E-Procurement abgekürzt, lassen sich gemäß Abu Bakar et al. als die Integration, Verwaltung, Automatisierung oder Optimierung des Beschaffungsprozesses unter Verwendung elektronischer Technologien oder webbasierter Anwendungen definieren [Ab16]. Sie lassen sich innerhalb der EU in die Phase vor der Zuschlagserteilung (pre-award) und nach der Zuschlagserteilung (post-award) aufteilen [Bo17], [Pr19] und machen ungefähr 12% des gesamten BIPs der EU aus [OE19]. Deutschland nimmt eine feinere Unterteilung in insgesamt drei verschiedene und voneinander bisher eher unabhängige Phasen durch: Bedarfserhebung, Vergabeverfahren und Vertragsdurchführung, und Einkauf.⁴ Die Prozesse der Bedarfserhebung und des Vergabeverfahrens lassen sich dem pre-award zuordnen, wohingegen alle Prozessschritte der Vertragsdurchführung und des Einkaufs bereits im post-award sind. Elektronische Kataloge können im elektronischen Beschaffungsverfahren in Form von pre-award Katalogen im Vergabeverfahren und post-award Katalogen im Katalogmanagement eingesetzt werden [IT21]. Sie ermöglichen das Festhalten beschaffungsbezogener Informationen in einem standardisierten Katalog [Hu19].

Im Rahmen der elektronischen Beschaffung gibt es zudem diverse nationale und internationale Datenstandards, welche zum Zwecke der Interoperabilität verschiedene Aspekte des Beschaffungsprozesses wie Nachrichtenflüsse und Beschaffungsdokumente zu standardisieren bestreben. Interoperabilität wird gemäß des European Interoperability Frameworks (EIF) als *“the ability of organizations to interact towards mutually beneficial goals, involving the sharing of information and knowledge between these organizations (...) by means of the exchange of data between their ICT systems”* definiert [Eu17] und ermöglicht einen reibungslosen Informationsaustausch sowie grenzüberschreitende Zusammenarbeit zwischen verschiedenen Akteuren [PNB12]. De Vries versteht Standardisierung als *“the activity of establishing and recording a limited set of solutions, to actual or potential matching problems, directed at benefits for the party or parties involved, (...) and expecting that these solutions will be repeatedly or continuously used, (...) by a substantial number of the parties for whom they are meant”* [dV97]. Die Standardisierung von Daten ist von wesentlicher Bedeutung für den öffentlichen Sektor, da die Verwendung offener Standards zu Herstellerunabhängigkeit sowie einer Vielzahl technischer Vorteile und Anpassungsfreiheiten führt [Jo15].

Ein wesentlicher Beschaffungsstandard für die Mitgliedsländer der EU ist „Pan European Public Procurement OnLine“ (Peppol)⁵, welcher auf Basis der Universal Business Language (UBL)⁶ diverse XML-Datenartefakte des Beschaffungsprozesses im pre-⁷ und

⁴ Siehe Planungs- und Architekturkonzept zur Digitalisierung der öffentlichen Beschaffung:

https://www.finanzen.bremen.de/sixcms/media.php/13/Architekturkonzept_OZG-Koop.59013.pdf

⁵ <https://peppol.eu/> letzter Zugriff: 04.05.2023

⁶ https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ubl letzter Zugriff: 08.05.2023

⁷ <https://docs.peppol.eu/pracc/> letzter Zugriff: 08.05.2023

post-award⁸ standardisiert. Peppol ermöglicht Akteuren den Austausch dieser standardisierten Datenartefakte über das Peppol eDelivery Network, das auf dem 4-Corner-Modell⁹ basiert. Der Peppol Standard wurde in Teilen auf den Spezifikationen des Comité Européen de Normalisation (CEN) aufgebaut, welches von 2009 bis 2015 initiale Datenartefakte und Standards für den europäischen Beschaffungsprozess in Profilen spezifiziert hat¹⁰, die Peppol für die Entwicklung eigener Profile als Grundlage verwenden konnte. Ein weiterer relevanter Datenstandard ist der deutsche BMEcat¹¹, welcher bereits seit 2001 als post-award Produktkatalog produktiv in Deutschland genutzt wird¹². Peppol und BMEcat sind für diesen Beitrag relevant, da ein standardisiertes Datenmapping zwischen dem noch in der Entwicklung stehendem Peppol pre-Award Catalogue und dem BMEcat stattfindet (Kapitel 6), um die in Kapitel 4 identifizierte Herausforderung des kritischen Übergangs zwischen pre- und post-award bewältigen zu können. Dazu wird ein UBL-Syntaxmapping [OA06] auf Wertebereichsebene zwischen den Katalogstandards durchgeführt, um die Abbildbarkeit aller Datenelemente untersuchen zu können.

4 Herausforderungen in der elektronischen Beschaffung

Öffentliche elektronische Beschaffungsverfahren sind ein wesentlicher Bestandteil öffentlicher Stellen, die 12% des Bruttoinlandsprodukts der EU sowie fast ein Drittel aller Verwaltungsausgaben der Mitgliedsländer ausmachen [OE19]. Boykin erläutert, dass das öffentliche Beschaffungswesens trotz dieses erheblichen Einflusses auf das Wachstum des europäischen Wirtschaftsraums in der wissenschaftlichen Literatur vernachlässigt wird [Bo22]. Mehrere Veröffentlichungen skizzieren Herausforderungen elektronischer Beschaffung. Amemba et al. etwa diskutieren, dass ein fundamentales Problem darin liegt, wie innovative Technologien bei den stark von bestehenden Rechtsvorschriften getriebenen Umsetzungen in den öffentlichen Beschaffungsprozess Einzug finden [Am13]. Mohungoo et al. identifizieren implementationsbezogene Herausforderungen wie disruptive Innovationsaspekte, Veränderungsresistenz und die Einbindung von kleinen und mittleren Unternehmen (KMU) [MBK20]. Letzteres ist eine wesentliche Herausforderung für die elektronische Beschaffung, da die erfolgreiche Implementation neuer Technologien und Innovationen erfolgen muss, ohne dabei zusätzliche Hindernisse für die Teilnahme von KMUs zu schaffen (SMW22). Diese Erkenntnis wird durch weitere Studien in Europa gestützt [Eu21], [Lo15]. Bei der näheren Analyse dieses Sachverhalts wird die Herausforderung insbesondere für das kleinteilige Beschaffungswesen erkenntlich. Beschaffungsvorgänge werden innerhalb dieses Beitrags als kleinteilig definiert, wenn der Rechnungswert eines Vorgangs unter 10.000€ beträgt. Nach Angaben

⁸ <https://docs.peppol.eu/poacc/upgrade-3/> letzter Zugriff: 08.05.2023

⁹ <https://www.e-rechnung-bund.de/faq/peppol/> letzter Zugriff: 08.05.2023

¹⁰ <https://www.uni-koblenz-landau.de/de/koblenz/fb4/iwvi/agvinf/projekte/abgeschlossene/cen-bii> letzter Zugriff: 09.05.2023

¹¹ <https://www.bme.de/services/bmecat/> letzter Zugriff: 08.05.2023

¹² <https://www.bme.de/services/bmecat/info> letzter Zugriff: 08.05.2023

des Landesbetriebs für Mobilität in Rheinland-Pfalz (LBM)¹³ aus dem Jahre 2019 im Projekt „Kleinteilige Beschaffung standardisiert.digital.vernetzt (KleBe.digital)“¹⁴ bestehen 95% aller Beschaffungsvorgänge aus kleinteiligen Vorgängen, wovon knapp die Hälfte einen Rechnungswert von über 250€ besitzen [Pr21]. Kleinteilige Beschaffungsvorgänge verursachen demzufolge aufgrund der fehlenden Standardisierung im pre-award einen erheblichen administrativen Aufwand mit zahlreichen Medienbrüchen für öffentliche Verwaltungen [Hu19], [IT21], [Pr22]. Hinzu kommt gemäß eines EU-Reports, dass KMUs aufgrund ihrer Spezialisierung und des geringen logistischen Aufwands zwar gut geeignet wären, um diese Aufträge zu erfüllen, aber die aktuelle Organisation der Beschaffungsverwaltung dazu führt, dass in der Praxis nur wenige KMUs an öffentlichen Vergabeverfahren teilnehmen. KMUs sehen sich demnach mit einer Vielzahl von Barrieren wie geringem Vertrauen, unausgewogenen Vergabekriterien und einem hohen Verwaltungsaufwand konfrontiert, die ihr Interesse an Vergabeverfahren einschränken [Eu21]. Eine weitere Studie in Großbritannien betont ebenfalls, dass KMUs frustriert sind über die Teilnahme an elektronischen Beschaffungsverfahren. Zu strenge Qualifikationskriterien, schlecht verfasste Ausschreibungsunterlagen und ein zu hoher Ressourcenbedarf sind gemäß der Studie die häufigsten Barrieren für KMUs [Lo15]. Folglich sind sowohl öffentliche Stellen als auch die am Vergabeverfahren teilnehmenden Unternehmen, insbesondere KMUs, daran interessiert, den administrativen Verwaltungsaufwand innerhalb des elektronischen Beschaffungsprozesses zu reduzieren. Während in Deutschland bereits post-award Standards¹⁵ der Koordinierungsstelle für IT-Standards (KoSiT)¹⁶ wie XRechnung¹⁷ und XBestellung¹⁸ den Weg in die Produktivität finden, um dort Vorteile wie prozessübergreifende Interoperabilität, Harmonisierung bestehender Lösungen und prozessliche (Teil-)Automatisierungen zu ermöglichen, fehlt in Deutschland im pre-award bisher die aktive Einbindung standardbasierter Lösungen [IT21]. Hierzu gibt es aber bereits Bestrebungen in Form eines Prototyps zur Verwaltung von elektronischen pre-award Katalogen im Rahmen von KleBe.digital¹⁹. Des Weiteren wird im Projekt Digitalisierung der öffentlichen Beschaffung sowie in Peppol an der Entwicklung von elektronischen Katalogstandards gearbeitet²⁰. Damit soll ein erheblicher Anteil zur standardbasierten Digitalisierung der pre-award Phase beigetragen werden.

5 Elektronische Kataloge als Erfolgsfaktoren in der Praxis

Basierend auf den Ergebnissen des vierten Kapitels wird nun entlang des Rigor Cycles und des Relevance Cycles der DSR [He07] tiefer analysiert, inwieweit elektronische

¹³ <https://lbm.rlp.de/de/startseite/> letzter Zugriff: 08.05.2023

¹⁴ <https://www.digitale-beschaffung.de/> letzter Zugriff: 08.05.2023

¹⁵ <https://xeinkauf.de/> letzter Zugriff: 08.05.2023

¹⁶ <https://www.xoev.de/> letzter Zugriff: 08.05.2023

¹⁷ <https://xeinkauf.de/xrechnung/> letzter Zugriff: 08.05.2023

¹⁸ <https://xeinkauf.de/xbestellung/> letzter Zugriff: 08.05.2023

¹⁹ <http://141.26.157.226/menu> letzter Zugriff: 08.05.2023

²⁰ https://test-docs.peppol.eu/pracc/Pre-Award_Catalogue/profiles/p035/index.html letzter Zugriff: 08.05.2023

Kataloge bereits erfolgreich in der Praxis eingesetzt werden, um die diskutierten Herausforderungen öffentlicher elektronischer Beschaffungsvorgänge bewältigen zu können. Ein Bericht der „Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH“ über die Nutzung von elektronischen Katalogen in der öffentlichen Beschaffung stellt fest, dass in der Praxis bisher nur wenige Länder zentrale elektronischen Kataloge verwenden. Als mögliche Gründe nennt die GIZ Herausforderungen wie den hohen Initialaufwand für die Erstellung und Pflege von elektronischen Katalogen sowie die Notwendigkeit spezifischer Institutionen und rechtlicher Rahmenbedingungen für diese [De22]. Im Folgenden wird die Nutzung von elektronischen Katalogen in diversen europäischen Ländern näher analysiert und diskutiert.

In Italien gibt es seit 2004 den „Elektronischen Marktplatz für öffentliche Behörden“ (MePA) sowie den MePA-Katalog²¹, der als elektronischer Katalog von öffentlichen Auftraggebern genutzt werden kann, um strukturiert Produkte zu vergleichen und zu beschaffen. Mithilfe des MePA-Katalogs können direkte Aufträge erteilt oder Angebote von Lieferanten angefordert werden, um weitere Verhandlungen über Preis und Lieferbedingungen zu führen [De22]. Gemäß eines OECD-Reports aus dem Jahre 2019 enthielt der MePA-Katalog im Jahr 2017 über 10 Millionen Artikel und verarbeitete ungefähr 600.000 Transaktionen. Die OECD stellt fest, dass MePA zu wesentlichen Zeitersparnissen für öffentliche Auftraggeber und Lieferanten sowie einem deutlich erhöhten Wettbewerb geführt hat, da sich Unternehmen auf mehrere ähnliche Ausschreibungen bewerben können und öffentliche Auftraggeber somit eine größere Auswahl an Produkten und Dienstleistungen erhalten [OE19].

In Norwegen werden im "Elektronisk handelsformat" (EHF)²², einem auf Peppol basierenden Standard, der von der norwegischen Agentur für öffentliche und finanzielle Verwaltung (DFØ / Anskaffelser)²³ gepflegt wird, Profile für die Nutzung elektronischer Kataloge definiert, welche die bisher noch in der Anfangsphase stehenden Katalogbestrebungen in Peppol innovativ erweitern²⁴. Dazu werden im pre-award zwei elektronische Katalogstandards festgelegt²⁵: EHF pre-award Catalogue request²⁶ und EHF pre-award Catalogue response²⁷. Der EHF pre-award Catalogue request kann von einer öffentlichen Stelle verwendet werden, um ihren Bedarf der zu beschaffenden Leistung oder Produkte katalogisiert im Detail darzustellen. Die öffentliche Stelle kann dazu beispielsweise genaue Anforderungen über Produkteigenschaften, Zertifikate und Preismini und -maxima in der Kataloganfrage definieren, die anschließend im Rahmen der Ausschreibung zusammen mit weiteren initialen Beschaffungsdokumenten allen interessierten Wirtschaftsteilnehmern zur Verfügung gestellt wird. Ein interessierter

²¹ <https://www.mepa.it/> letzter Zugriff: 09.05.2023

²² <https://www.digdir.no/standarder/ehf-elektronisk-handelsformat/1678> letzter Zugriff: 08.05.2023

²³ <https://anskaffelser.no/> letzter Zugriff: 08.05.2023

²⁴ <https://anskaffelser.dev/preaward/g2/spec/> letzter Zugriff: 08.05.2023

²⁵ <https://anskaffelser.dev/preaward/g2/spec/draft/pacr-1.0/spec/> letzter Zugriff: 08.05.2023

²⁶ <https://anskaffelser.dev/preaward/g2/spec/draft/pacr-1.0/syntax/PreAwardCatalogueRequest/tree/> letzter Zugriff: 08.05.2023

²⁷ <https://anskaffelser.dev/preaward/g2/spec/draft/pacr-1.0/syntax/catalogue/tree/> letzter Zugriff: 08.05.2023

Bieter kann diese Anfrage dann mithilfe des EHF pre-award Catalogue response Standards ausfüllen und um die geforderten Informationen und Nachweise ergänzen, bevor er den nun vollständigen elektronischen Katalog seinem Angebot anhängt und der öffentlichen Stelle zurücksendet²⁸. In einem Interview vom 15.10.2021 mit einem Mitglied des EHF-Projekts wurde zusätzlich untersucht, inwieweit elektronische Kataloge (wie die EHF pre-award Kataloge) als Erfolgsfaktor die Herausforderungen elektronischer Beschaffungsvorgänge bewältigen können. In Norwegen werden elektronische Kataloge bereits produktiv in Beschaffungssystem eingesetzt und lässt sich gemäß EHF als sehr erfolgreich bewerten. Durch den Einsatz von Katalogen können öffentliche Stellen ihre Ausschreibungen effizienter verwalten und interessierte Wirtschaftsteilnehmer mit geringem Mehraufwand an mehreren ähnlichen Ausschreibungen teilnehmen, da sich die zu erstellenden Angebote meist kaum unterscheiden. Dies hat dazu geführt, dass die Zahl der an Vergabeverfahren teilnehmenden Unternehmen und somit auch der Wettbewerb insgesamt gestiegen ist. Die Vorteile seien lt. EHF Vertreter vor allem für KMUs von großer Bedeutung, da diese nur begrenzte Ressourcen für die Verwaltung von Angeboten für verschiedenen Ausschreibungen haben. Zusammenfassend wird die Annahmequote für Kataloge bisher als sehr hoch eingeschätzt, obwohl einige KMUs aufgrund des Mangels an technischen Möglichkeiten noch keine elektronischen Kataloge nutzen können. Daher rät EHF den öffentlichen Stellen, einfache Zugangspunkte einzurichten, um auch sehr kleine Unternehmen zur Nutzung elektronischer Kataloge zu bewegen. Die Ergebnisse des Interviews legen weiter nahe, dass elektronische Kataloge vor allem für kleinteilige Beschaffungsvorgänge von Produkten und nicht für komplexe Ausschreibungen wie umfangreiche Bauleistungen als Erfolgsfaktor anzusehen sind [Sc21].

Auf Basis der erfolgreichen Nutzung von elektronischen Katalogen in Norwegen wurde auch im deutschen Projekt KleBe.digital ein prototypisches Lieferantentool entwickelt, das sogenannte pre-Award Catalogue Tool (pACT)²⁹. Es ermöglicht auf Basis der EHF-Kataloge die Erstellung eines pre-award Catalogue request durch eine öffentliche Stelle, der dann von interessierten Wirtschaftsteilnehmern ausgefüllt und als pre-award Catalogue response exportiert werden kann. Im Kontext des Projekts „Digitalisierung der öffentlichen Beschaffung“ wird nun in Abstimmung mit der „Peppol Pre-Award Community“ ein Peppol-Standard für elektronische Katalogformate im pre-award ausgearbeitet (Peppol pre-award Catalogue request und Peppol pre-award Catalogue)³⁰.

In Deutschland existiert auch der BMEcat. Anders als im norwegischen Beispiel ist der BMEcat jedoch ein post-award Katalog, der somit erst nach Vertragsschluss vom Lieferanten erstellt und anschließend an einkaufende Unternehmen oder öffentliche Behörden übermittelt wird. Auch der BMEcat kann eingesetzt werden, um den öffentlichen Beschaffungsvorgang standardbasiert zu vereinfachen und administrativen Aufwand zu reduzieren, indem Lieferanten und öffentliche Stellen kontinuierlich einen

²⁸ https://anskaffelser.dev/preaward/g2/spec/draft/pacr-1.0/spec/#_process_and_typical_use_cases letzter Zugriff: 08.05.2023

²⁹ <http://141.26.157.226/menu> letzter Zugriff: 08.05.2023

³⁰ https://test-docs.peppol.eu/pracc/Pre-Award_Catalogue/profiles/p035/index.html letzter Zugriff: 08.05.2023

gemeinsamen post-award Katalog pflegen, aus dem die öffentliche Stelle bei Bedarf immer wieder unkompliziert vom Lieferanten bestellen kann³¹.

Elektronische Kataloge können also in der Praxis genutzt werden, um den administrativen Aufwand für öffentliche Behörden und Unternehmen im Beschaffungsverfahren zu verringern. Die strukturierte Erfassung von Katalogdaten ermöglicht die Nachnutzbarkeit dieser Daten in weiteren Prozessphasen und kann dazu dienen, Medienbrüche zu überwinden und den kritischen Übergang zwischen pre- und post-award zu schließen. In KleBe.digital wurde festgestellt, dass ein konkretes Datenmapping zwischen dem Peppol pre-award Catalogue und dem BMEcat im post-award dazu beitragen kann, diese Vorteile zu realisieren. Dieses Datenmapping wird nachfolgend erarbeitet.

6 Abbildung des Peppol Pre-Award Catalogue auf den deutschen BMEcat 2005.2

Das Mapping zwischen dem Peppol pre-award Catalogue und dem BMEcat wird als Lösungsbaustein für die Bewältigung diverser Herausforderungen im elektronischen Beschaffungsprozess identifiziert. Etwa konnte im Projekt KleBe.digital festgestellt werden, dass die Nutzung beider Kataloge im elektronischen Beschaffungsverfahren dazu führt, dass Daten bereits im pre-award strukturiert erfasst und dann ebenfalls strukturiert und medienbruchsfrei in den post-award überführt werden könnten. Um dies gewährleisten zu können, bedarf es eines UBL-Syntaxmappings zwischen dem Peppol pre-award Catalogue³² und dem BMEcat 2005.2³³, welches auf Wertebereichsebene die direkte Übertragbarkeit von Datenelementen aus dem pre-award Catalogue in den BMEcat prüft. Auf Basis dieses Mappings kann zukünftig eine XSL-Transformation durchgeführt werden, welche automatisiert alle relevanten Daten aus dem Peppol pre-award Catalogue in den BMEcat transformiert (s. Kapitel 7). Das vollständige Datenmapping³⁴ wird im Design Cycles der DSR [He07] inkrementell und iterativ durchgeführt. Im Folgenden werden ausgewählte Ergebnisse des Datenmappings diskutiert.

Die Abbildbarkeit zwischen den Katalogen kann zufriedenstellend bewertet werden. Allgemeine Informationen zu Lieferanten und beschaffenden Behörden sind vollständig abbildbar. Auch Rahmenverträge, bei denen z.B. zwischen der den Katalog erhaltenden (Catalogue Receiver) und der vom Katalog bestellenden (Catalogue Customer) Behörde unterschieden werden kann, sind durch das direkte Mapping der „Catalogue Customer“ und „Catalogue Supplier“ auf die „BUYER“ und „SUPPLIER“ Klassen des BMEcat möglich. Auch spezifische Informationen zur späteren Bestellung sowie diverse Elemente zur Beschreibung, Bepreisung und Besteuerung von Produkten sind fast vollständig abbildbar. Dennoch sind große Unterschiede in der Philosophie und dem

³¹ https://a.storyblok.com/f/104752/x/9344382bd6/bmecat_2005-2_deu.pdf letzter Zugriff: 08.05.2023

³² https://test-docs.peppol.eu/pracc/Pre-Award_Catalogue/syntax/catalogue/tree/ letzter Zugriff: 09.05.2023

³³ https://a.storyblok.com/f/104752/x/9344382bd6/bmecat_2005-2_deu.pdf letzter Zugriff: 09.05.2023

³⁴ Internes Dokument, 2023

Verwendungszweck der Kataloge festzustellen, die einen wesentlichen Einfluss auf einige Elementgruppen sowie die Katalogstruktur haben. Der Peppol Catalogue erfasst als elektronischer Katalog im pre-award möglichst viele produkt- und vertragspezifische Informationen wie diverse Nachweise, Zertifikate, und Eigenschaften. Hinzu kommt, dass die Strukturierung von Produkten minimalistisch auf Positionsebene erfolgt. Im Gegensatz dazu ermöglicht der BMEcat eine sehr detaillierte, hierarchische Definition von ganzen Produktbäumen und arbeitet mit nur minimalen Elementen zur Erfassung vertragspezifischer Informationen, da diese zum Verwendungszeitpunkt des BMEcat im post-award schon feststehen müssen. Der BMEcat ist um ein Vielfaches umfangreicher als der Peppol Catalogue, weil durch den BMEcat gemäß Spezifikation auch umfassende, langfristige Rahmenverträge über eine Vielzahl von Produkten mit einem sehr tiefgreifendem Detailgrad ermöglicht werden sollen³⁵, wohingegen der Peppol Catalogue eher als strukturierter Katalog-Anhang eines Angebots zu verstehen ist, der minimal alle erforderlichen Informationen für die Bewertung der Zuschlagserteilung enthält³⁶. Dieses Ergebnis ist als hoch signifikant zu bewerten, da zwar gemäß des Datenmappings eine Überführung vieler grundsätzlicher Elemente in den BMEcat möglich ist, aber eben die Darstellung weiterführender Produktstrukturen und anderer post-award spezifischer Informationen die Grenze für den Peppol pre-award Catalogue bilden.

Das Mapping leistet einen wesentlichen Beitrag zur Schließung der Lücke zwischen pre- und post-award. Die Abbildbarkeit vieler grundlegender Elemente vom Peppol pre-award Catalogue in den BMEcat zeigt, dass auch eine spätere XSL-Transformation zum BMEcat mit einigen Ergänzungen möglich ist. Im Fazit ist die Verwendung beider elektronischer Katalogstandards im deutschen Beschaffungsprozess als Erfolgsfaktor für die Bewältigung diverser Herausforderungen in der elektronischen Beschaffung anzusehen.

7 Handlungsempfehlungen

Die Ergebnisse dieses Beitrags zeigen, dass diverse Herausforderungen der öffentlichen elektronischen Beschaffung durch die Nutzung elektronischer Kataloge bewältigt werden können. Das UBL-Syntaxmapping zwischen dem Peppol pre-award Catalogue und dem BMEcat zeigt, dass die kritische Lücke zwischen pre- und post-award durch die Verwendung beider Katalogstandards geschlossen werden kann. Um dies weiterführend zu überprüfen, muss jedoch in der Praxis eruiert werden, wie viele relevante Informationen nach der Transformation im BMEcat vorhanden sind und wie viele Elemente nach der Transformation ergänzt werden müssen, um einen minimalen BMEcat zu erhalten. Dazu muss auf Basis des Datenmappings eine XSL-Transformation programmiert und anhand mehrerer Use-Cases für unterschiedliche Beschaffungsverfahren getestet werden. Im positiven Testfall kann die Transformation funktional in das pACT aufgenommen und anschließend in der Praxis prototypisch pilotiert werden. Die Nutzung elektronischer

³⁵ https://a.storyblok.com/f/104752/x/9344382bd6/bmecat_2005-2_deu.pdf letzter Zugriff: 09.05.2023

³⁶ https://test-docs.peppol.eu/pracc/Pre-Award_Catalogue/transactions/Catalogue/index.html#benefits letzter Zugriff: 09.05.2023

Kataloge wird für öffentliche Behörden und Wirtschaftsteilnehmer, insbesondere KMUs, auch im pre-award empfohlen, denn dies trägt dazu bei, den administrativen Aufwand innerhalb des Beschaffungsprozesses zu reduzieren, Medienbrüche zu vermeiden und die Lücke zwischen pre- und post-award zu füllen, indem nachnutzbare Katalogdaten bereits im pre-award definiert und dann über weitere Prozessschritte hinweg genutzt werden. Dadurch können auch KMUs durch die Strukturierung eines Produktkatalogs an mehreren ähnlichen Ausschreibungen mit nur geringem Mehraufwand teilnehmen.

8 Fazit und Ausblick

In diesem Beitrag wurde im Rahmen der verschiedenen, iterativ durchgeführten Cycles des Design Science Research [He04], [He07] untersucht, inwieweit die Nutzung elektronischer Kataloge dazu beitragen kann, diverse Herausforderungen (Bedarf innovativer Rechtsvorschriften [Am13], unstrukturierte pre-award Phase, unregulierter Übergang zwischen pre- und post-award mit zahlreichen Medienbrüchen, Vielzahl von Barrieren für KMUs [MBK20], [IT21], [Eu21], [Lo15]) in der öffentlichen elektronischen Beschaffung zu bewältigen. Dazu wurde im Rigor Cycle eine Literaturanalyse als wissenschaftliches Fundament erarbeitet und durch graue Literatur und Gesetzestexte ergänzt. Hieraus wurden obige Herausforderungen in der elektronischen Beschaffung extrahiert und die Rolle von elektronischen Katalogen bei der Bewältigung dieser analysiert. Die Nutzung elektronischer Kataloge kann dazu beitragen, den administrativen Aufwand innerhalb des Beschaffungsprozesses für alle beteiligten Akteure zu reduzieren, Medienbrüche zu vermeiden und die Lücke zwischen pre- und post-award zu füllen. Des Weiteren können KMUs durch die Strukturierung eines Produktkatalogs an mehreren ähnlichen Ausschreibungen mit nur geringem Mehraufwand teilnehmen (Kapitel 5). Als Ergebnis des Relevance Cycles wurde das UBL-Syntaxmapping zwischen Peppol pre-award Catalogue und BMEcat als Lösungsbaustein für die Bewältigung des Übergangs zwischen pre- und post-award identifiziert (Kapitel 6). Die Abbildung des Peppol Catalogue auf den BMEcat wurde dabei inkrementell und iterativ in KleBe.digital entwickelt, diskutiert und bewertet. Die Standards unterscheiden sich darin, dass der BMEcat dazu dient, umfassende, langfristige Rahmenverträge über eine Vielzahl von Produkten mit hierarchischer Baumstruktur zu ermöglichen, wohingegen der Peppol Catalogue als strukturierter Katalog-Anhang eines Angebots dient und die minimal erforderlichen Informationen für die Bewertung der Zuschlagserteilung enthält. Insgesamt sind die Kataloge aber zufriedenstellend aufeinander abbildbar, da sich fast alle Unterschiede aus den unterschiedlichen Prozessphasen des pre- und post-awards ergeben und die meisten allgemeinen Elemente für Parteien, Produkte, Preise und Steuern erzeugbar sind. Zukünftig soll das theoretische Datenmapping durch eine XSL-Transformation mit praxisnahen Use-Cases getestet werden, die bei Erfolg im Rahmen des pACT prototypisch in der Praxis pilotiert wird. Weiter wird untersucht, inwieweit Transformationen zwischen den elektronischen Katalogstandards und post-award Artefakten wie Bestellung und Rechnung möglich sind. Dabei kann die Pflege eines zentralen, strukturierten Datenartefakts, das direkt in die verschiedenen Datenstandards

im pre- und post-award transformiert werden kann, zielführend sein.

Literaturverzeichnis

- [Ab16] Abu Bakar, N. et al.: Abridgment of Traditional Procurement and E-Procurement: Definitions, Tools and Benefits, *J. Emerg. Econ. Islam. Res.*, Bd. 4, Nr. 1, S. 74, 2016.
- [Am13] Amemba, C.S. et al.: Challenges affecting public procurement performance process in Kenya, *International Journal of Research in Management*, 3(4), pp.41-55. 2013.
- [Bo17] Bobowski, S.: E-Procurement in the European Union, *Asia-Pacific J. EU Stud.*, Bd. 17, Nr. 2, S. 23–35, 2017.
- [Bo22] Boykin, E.A.: Public procurement and European Union integration: A systematic review, *Journal of European Integration*, pp.1-20., 2022.
- [Bu16] Bundesamt für Justiz: Verordnung über die Vergabe öffentlicher Aufträge (VgV), 2016, https://www.gesetze-im-internet.de/vgv_2016/VgV.pdf, letzter Zugriff: 10.05.2023.
- [De19] Deutscher Vergabe- und Vertragsausschuss für Bauleistungen: Vergabe- und Vertragsordnung für Bauleistungen Teil A, 2019, <https://dejure.org/gesetze/VOB-A>, letzter Zugriff: 10.05.2023.
- [De22] Deutsche Gesellschaft für internationale Zusammenarbeit (GIZ): Use of E-catalogues in Sustainable Public Procurement (SPP): Overview of current practices, <https://www.oneplanetnetwork.org/sites/default/files/from-crm/Use%2520of%2520E-catalogues%2520in%2520Sustainable%2520Public%2520Procurement.pdf>, letzter Zugriff: 10.05.2023
- [dV97] de Vries, H.: Standardization - What's in a name? *Terminology* 4(1), S. 55–83, 1997.
- [Eu14a] Europäische Union: Richtlinie über die öffentliche Auftragsvergabe (Nr. 2014/24/EU), *Amtsblatt der Eur. Union*, Bd. 2014, Nr. 4, S. 65–242, 2014.
- [Eu14b] Europäische Union: Richtlinie über die Vergabe von Aufträgen durch Auftraggeber im Bereich der Wasser-, Energie- und Verkehrsversorgung sowie der Postdienste (Nr. 2014/25/EU), *Off. J. Eur. Union*, S. 243–374, 2014.
- [Eu17] Europäische Kommission: New European Interoperability Framework: Promoting seamless services and data flows for European public administrations, *Publications Office of the European Union*. S. 1–48, 2017. doi: 10.2799/78681.
- [Eu19] Europäische Union: Delegierte Verordnung (EU) 2019/1830 der Kommission, *OJ*, 2019.
- [Eu21] Europäische Kommission: SME Needs Analysis in Public Procurement, 2021, <https://op.europa.eu/sk/publication-detail/-/publication/ee874832-decc-11eb-895a-01aa75ed71a1/language-en>, letzter Zugriff: 10.05.2023.
- [He04] Hevner, A.R. et al.: Design Science in Information Systems Research, *MIS Q.*, Bd. 28, Nr. 1, S. 75–105, 2004.
- [He07] Hevner, A.R.: A Three Cycle View of Design Science Research“, *Scand. J. Inf. Syst.*, Bd. 19, Nr. 2, S. 87–92, 2007, doi: <http://aisel.aisnet.org/sjjs/vol19/iss2/4>.

- [Hu19] Hudrasyah, H. et al.: E-Catalogue Attractiveness study to increase suppliers participation. *Int. Journal of Accounting, Finance and Business*, 4(20), 14-31, 2019
- [IT21] IT-Planungsrat: Planungs- und Architekturkonzept: Digitalisierung der Beschaffung - Kooperationsprojekt zur standardbasierten Digitalisierung des öffentlichen Einkaufs- und Beschaffungsprozesses, 2021.
- [Jo15] Jokonya, O.: Investigating open source software benefits in public sector. *Proceedings of 48th Hawaii Int. Conference on System Sciences, IEEE*, S. 2242-2251, 2015.
- [Lo15] Loader, K.: SME suppliers and the challenge of public procurement: Evidence revealed by a UK government online feedback facility. *Journal of Purchasing and Supply Management*, 21(2), 103-112, 2015.
- [MBK20] Mohungoo, I.; Brown, I.; Kabanda, S.: A systematic review of implementation challenges in public E-Procurement. In *Proceedings of 19th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E, Part II*, S. 46-58, Springer Int. P., 2020.
- [OA06] OASIS: Universal Business Language (UBL) Naming and Design Rules, 2006, <https://www.oasis-open.org/committees/download.php/22992/UBL-NDR-2-20070207.pdf> letzter Zugriff: 10.05.2023.
- [OE19] OECD: Report on the Implementation of the Recommendation of the Council on Public Procurement. OECD Publishing, Paris, 2019.
- [PNB12] Pardo, T. A.; Nam, T.; Burke, G. B.: E-Government Interoperability: Interaction of Policy, Management, and Technology Dimensions, *Soc. Sci. Comput. Rev.*, Bd. 30, Nr. 1, S. 7–23, 2012, doi: 10.1177/0894439310392184.
- [Pr19] Prałat, E.: Public e-procurement tools in European Union, *Nierówności społeczne a wzrost Gospod.*, Bd. 58, Nr. 2, S. 188–197, 2019, doi: 10.15584/nsawg.2019.2.14.
- [Pr21] Projektteam KleBe.digital: Projektsteckbrief Kleinteilige Beschaffung standardisiert.digital.vernetzt, internes Dokument, 2021.
- [Pr22] Projektteam KleBe.digital: KleBe.digital Architektur, internes Dokument, 2022.
- [RS04] Rowley, J.; Slack, F.: Conducting a literature review, *Manag. Res. News*, Bd. 27, Nr. 6, S. 31–39, 2004, doi: 10.5124/jkma.2017.60.5.384.
- [Sc21] Schmitz, A.: Development of an interoperable standard data specification for small-scale procurement, Masterarbeit, Universität Koblenz-Landau, 2021.
- [SMW22] Schmitz, A.; Mondorf, A.; Wimmer, M. A.: Framework for designing interoperable public service architectures with exemplification along small-scale procurement and PEPPOL. In *Proceedings of DG. O 2022, ACM*, S. 22-34, 2022.
- [UF09] Uyarra E.; Flanagan, K.: Understanding the Innovation Impacts of Public Procurement, *Manchester Business School Working Paper No. 574*, 2009.
- [Wo19] World Bank: Green Public Procurement: An Overview of Green Reforms in Country Procurement Systems, *Climate Governance Papers Series*, Washington, DC, 2019.
- [WW02] Webster, J.; Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a literature review, *MIS Q.*, Bd. 26, Nr. 2, S. iii–xiii, 2002.

Datenmanagementkonzepte und -frameworks im öffentlichen Beschaffungswesen

Aktuelle Trends und innovative Lösungen aus Wissenschaft und Praxis

Matthias Hartmann¹, Andreas Schmitz², Maria A. Wimmer³


Abstract: Die Digitalisierung wird aktuell stark durch Daten-getriebene Ansätze vorangetrieben, da mehr und mehr Daten von vielen verschiedenen Akteuren erhoben werden. Dieser Trend ist auch im Bereich öffentliche Verwaltung spürbar. Bisher gab es in diesem Kontext jedoch wenig Standardisierung. Auf EU-Ebene werden erste Konzepte und Verordnungen beschlossen, die diese Lücke adressieren. Dieser Beitrag untersucht verschiedene innovative Datenmanagementkonzepte aus der Literatur und Praxis, die für das öffentliche Beschaffungswesen relevant sind. In dem Vergleich werden auch die Nutzungsmöglichkeiten, Stärken und Schwächen der Konzepte dargestellt und weiterer Forschungsbedarf aufgezeigt. Die wesentlichen Ergebnisse sind eine komprimierte Darstellung verschiedener Konzepte, sowie Handlungsempfehlungen für weitere Entwicklungsrichtungen. So wird etwa der Bedarf einer unabhängigen, zentralen Datensammelstelle identifiziert, um die Daten nach Maßgabe von vorher definierten Vorschriften entgegenzunehmen und weiter zu verteilen.


Keywords: Öffentliches Beschaffungswesen, datengestützte Entscheidungsfindung, Datenmanagement, Literaturanalyse

1 Einleitung

Das öffentliche Beschaffungswesen ist ein wichtiger Zweig des staatlichen Handelns [Bü02]. Es umfasst den Einkauf von Waren-, Bau- und Dienstleistungen durch die öffentliche Hand⁴. Da die Digitalisierung, bspw. im Rahmen des Onlinezugangsgesetzes (OZG)⁵, und die datengestützte Entscheidungsfindung im öffentlichen Sektor immer wichtiger werden, gewinnen die Verwaltung und die Steuerung von Daten im öffentlichen

¹ Universität Koblenz, Forschungsgruppe E-Government, Universitätsstraße 1, 56070 Koblenz, mhartmann@uni-koblenz.de, <https://www.uni-koblenz.de/agvinf/>

² Universität Koblenz, Forschungsgruppe E-Government, Universitätsstraße 1, 56070 Koblenz, andreas-schmitz@uni-koblenz.de, <https://www.uni-koblenz.de/agvinf/>,  <https://orcid.org/0000-0001-9535-4001>

³ Universität Koblenz, Forschungsgruppe E-Government, Universitätsstraße 1, 56070 Koblenz, <https://www.uni-koblenz.de/agvinf/>, wimmer@uni-koblenz.de,  <https://orcid.org/0000-0002-8460-1027>

⁴ https://mbei.nrw/sites/default/files/asset/document/anlage_1_begriffe_und_definitionen.pdf, Zugriff 15.07.23

⁵ <https://www.onlinezugangsgesetz.de/>, letzter Zugriff 15.07.23

Beschaffungswesen zunehmend an Bedeutung [Ot19][So22][ASD16]. Um auch die Qualität, Integrität und Sicherheit der Daten während des gesamten Beschaffungsprozesses zu gewährleisten, ist ein wirksames Datenmanagementkonzept erforderlich. Dies gilt für alle Daten entlang des Beschaffungsprozesses, wie bspw. die Lieferantenauswahl, die Vertragsverwaltung und die Leistungsüberwachung. Die Datenverwaltung bzw. das Datenmanagement im öffentlichen Auftragswesen stellen jedoch besondere Herausforderungen dar. Darunter fallen die Notwendigkeit, Transparenz und Vertraulichkeit in Einklang zu bringen, die Komplexität des Datenaustausches zwischen mehreren Behörden und Abteilungen zu managen, sowie Bedrohungen für Datenschutzverletzungen und Cyberangriffe gering zu halten [Sa22][So22]. Darüber hinaus müssen im öffentlichen Beschaffungswesen verschiedene rechtliche und regulatorische Rahmenbedingungen eingehalten werden, z.B. die Datenschutzgrundverordnung (DSGVO)⁶ und das öffentliche Vergaberecht⁷. Ungeachtet dieser Herausforderungen bietet ein effektives Datenmanagement im öffentlichen Beschaffungswesen erhebliche Vorteile. Durch die Festlegung klarer Richtlinien und Verfahren für die Datenverwaltung können Behörden die Datenqualität verbessern, Risiken verringern und die Effizienz ihrer Beschaffungsprozesse steigern. In diesem Beitrag analysieren wir wissenschaftliche und graue Literatur über Datenmanagement im öffentlichen Beschaffungswesen hinsichtlich verschiedener Konzepte und Frameworks aus der Wissenschaft, Wirtschaft und dem öffentlichen Sektor. Die Arbeit zielt darauf ab, innovative Lösungen und aktuelle Trends im Datenmanagement zu identifizieren und mögliche Vorteile einer Nutzung dieser Konzepte im öffentlichen Beschaffungswesen aufzuzeigen. In Kapitel 2 stellen wir die Forschungsfragen und das methodische Vorgehen für die Untersuchung vor. Kapitel 3 erarbeitet durch eine semi-strukturierte Literaturanalyse nach Snyder [Sn19] das Verständnis von Datenmanagement und Daten-Governance. Mit der gleichen Methodik werden in Kapitel 4 Datenmanagementkonzepte und -frameworks untersucht. In Kapitel 5 werden aus Kap. 3 und Kap. 4 die Erkenntnisse aufgearbeitet und in Handlungsempfehlungen für innovatives Datenmanagement und Daten-Governance im öffentlichen Beschaffungswesen überführt. Der Beitrag schließt mit einer Zusammenfassung und Limitationen in Kapitel 6, einschließlich eines Ausblicks auf weitere Forschung.

2 Methodisches Vorgehen

Die Zielsetzung des Beitrags ist es, aktuelle und innovative Datenmanagementkonzepte und -frameworks zu identifizieren und daraus Handlungsempfehlungen für die Digitalisierung des öffentlichen Beschaffungswesens herzuleiten. Folgende Forschungsfragen motivieren die Untersuchungen: (1) *Wie wird der Begriff Datenmanagement/ Data Governance in der Literatur definiert?* (2) *Welche Datenmanagementkonzepte/-frameworks dominieren in der Forschung und der Wirtschaft?*

⁶ <https://dsgvo-gesetz.de>, letzter Zugriff 15.07.23

⁷ https://www.gesetze-im-internet.de/vgw_2016/, letzter Zugriff 15.07.23

Die beiden Forschungsfragen werden entlang einer semi-strukturierten Literaturanalyse nach Snyder [Sn19] aufgearbeitet, indem zunächst die Literatur nach Begriffen, Definitionen, Konzepten, Methoden und Ansätzen für Datenmanagement untersucht wird. Folgende Suchbegriffe sind aus den Forschungsfragen extrahiert und werden für die Suche verwendet: „Datenmanagement Definitionen“, „Datenmanagement im öffentlichen Beschaffungswesen“, „Innovative Datenmanagement Konzepte“, „Innovative Datenmanagementkonzepte im öffentlichen Beschaffungswesen“, „Data governance in public procurement“, „Data governance concepts in public procurement“. Für die Suche der Literatur werden Suchmaschinen, wie Google Scholar, Web of science o.ä. verwendet. Zur weiteren Untermauerung werden Whitepaper und Webseiten von Behörden und Unternehmen herangezogen, welche über die Google-Suche gefunden werden. Die gefundenen Konzepte/Frameworks sind anhand folgender Faktoren ausgewählt: a) Erscheinungsjahr nicht älter als 2018, da innovative Konzepte gesucht werden, die entweder noch nicht getestet wurden oder bereits getestet werden und eventuell dann einen großen Nutzungsbereich abdecken können; b) die Beiträge sollen einen Bezug zum öffentlichen Beschaffungswesen aufweisen; c) die Werke sollen sich auf Datenmanagement beziehen. Im nächsten Schritt werden die Konzepte und Rahmenwerke für Datenmanagement gegenübergestellt, indem sie näher analysiert, evaluiert und anhand ihres Untersuchungsobjektes, verwendeter Standards und Datensets sowie des Ziels gegenübergestellt werden. Die Vergleichskategorien und -kriterien sind aus der Literaturrecherche und aus dem Kontext des Projekts Digitale Beschaffung⁸ des IT-Planungsrats⁹ hergeleitet. Die gefundenen Ansätze aus der Literatur werden in diese Vergleichskriterien eingeordnet und geben einen Überblick über die wichtigsten Frameworks und Konzepte. Aus den Untersuchungen werden im letzten Schritt Handlungsempfehlungen

3 Theoretische Grundlagen

Gemäß des semi-systematischen Ansatzes nach Snyder [Sn19] wird die Literatur nach Definitionen für Datenmanagement und innovativen Datenmanagementkonzepten bzw. -frameworks untersucht. Ein zentraler Aspekt des Beitrages ist das Datenmanagement, engl. Data Governance, im öffentlichen Beschaffungswesen. Wende definiert Datenmanagement *„als das Rahmenwerk, das Entscheidungsrechte und Verantwortlichkeiten festlegt, um ein wünschenswertes Verhalten bei der Nutzung von Daten zu fördern“* [We07, Al16]. Nach Abraham et.al. zielt Datenmanagement darauf ab, eine organisationsweite Datenagenda umzusetzen, den Wert von Datenbeständen in einer Organisation zu maximieren und datenbezogene Risiken zu managen [ASV19]. Damit ist das Datenmanagement keine technische Applikation, sondern umfasst Richtlinien, Organisationen und Standards [KC18]. Im Unternehmenskontext wird diese Definition von Meier und Kauf-

⁸ <https://www.digitale-beschaffung.de/>, Beschluss des IT-Planungsrats zum Kooperationsprojekt Digitalisierung der Beschaffung: <https://www.it-planungsrat.de/beschluss/beschluss-2019-52>, letzter Zugriff 15.07.23

⁹ <https://www.it-planungsrat.de/>, letzter Zugriff 15.07.23

mann weiter spezifiziert, indem die Autoren das Datenmanagement aufteilen in Datenarchitektur, Datenadministration, Datentechnik und Datennutzung. Diese Aspekte unterstützen dabei die Datenhaltung und Datenpflege sowie Business Analytics [Me16]. Die *Datenarchitektur* beschäftigt sich mit der Festlegung von Datenschutzkonzepten sowie dem Formulieren und Pflegen der organisationsweiten Datenarchitektur [Me16]. Die Datenadministration verwaltet die Daten anhand von nationalen und internationalen Standards, Richtlinien und Normen [Me16]. Die Strukturierung der Daten soll anhand eines Klassifizierungsschemas erfolgen, sodass nach einer Ablage wieder auf die Daten zugegriffen werden kann [Bo06]. *Datentechnik* wird als Entwicklung von Datenverarbeitungssystemen zur Automatisierung von Bürotätigkeiten definiert [Ko89]. Meier und Kaufmann definieren Datentechnik weiter, als diejenige Technik, die sich mit dem Installieren, Reorganisieren und Sicherstellen von Datenbeständen beschäftigt. Dabei spielen die Entwicklung und Festlegung eines Datenverteilkonzepts, neben der Katastrophenvorsorge, eine weitere Rolle. Bei der Datennutzung steht die Datenanalyse und -interpretation sowie die Wissensgenerierung bzw. Erstellung von Prognosen und Erkennung von Mustern eine zentrale Rolle [Me16]. Doch was sind eigentlich Daten? Voß untersucht verschiedene Definitionen [Vo13]. Voß greift dabei die Argumentation aus [Ba11] auf, dass Daten Fakten sind, also „objektive, reproduzierbare Ergebnisse von Messungen, die wahre Aussagen über die Realität liefern“ [Vo13]. Machado und Parente de Oliveira untergliedern Daten weiter in strukturierte, semi-strukturierte und unstrukturierte Daten. Strukturierte Daten sind nach starren, vordefinierten Kriterien organisiert, wohingegen bei semi-strukturierte Daten einige Attribute bekannt sind und später weitere hinzugefügt werden können. Unstrukturierte Daten sind Daten, für die kein Schema vordefiniert ist und sie spiegeln nur den Inhalt wider [MP11]. Zudem greift Voß den Diskurs zu Linked Data, Big Data und dem Semantic Web auf, in dem Daten als „kombinierbare Fakten“ dargelegt werden [Vo13] und im Falle der Big Data als aufgezeichnete Beobachtungen aufgefasst werden [Vo13] (vgl. auch [TV22, AK19]). Obige Aspekte des Datenmanagements tragen mittels eines Datenmanagementkonzepts zu einer Steigerung der Effektivität und Effizienz bei [GR22]. Basierend auf dieser Grundlage werden in Kapitel 4 nun konkrete Konzepte und Frameworks analysiert und evaluiert, um diese anhand ihres Untersuchungsobjektes, Standards, Datensets und Ziels gegenüberzustellen.

4 Datenmanagementkonzepte & -frameworks

In der EU gibt es bis dato kein einheitliches Vorgehen für das Datenmanagement. Im Rahmen der Europäischen Datenstrategie¹⁰ wurde das Europäische Daten-Governance-Gesetz¹¹ verabschiedet und ein Vorschlag für harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung¹² entworfen. Damit liegen Bestrebungen vor, Daten in der EU zu standardisieren und zu zentralisieren. Für den Beschaffungsmarkt gibt

¹⁰ <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020DC0066>, letzter Zugriff 24.04.23

¹¹ <https://digital-strategy.ec.europa.eu/de/policies/data-governance-act>, letzter Zugriff 24.04.23

¹² <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52022PC0068>, letzter Zugriff 24.04.23

es für Tender Electronic Daily¹³ zwar einheitliche Vorschriften/Formulare für das Vergabeportal, aber Zugänglichkeit und Transparenz müssen weiter ausgebaut werden [Fa17]. Mit dem Daten-Governance Gesetz sollen verschiedene Sektor-Daten, wie Gesundheits-, Mobilitäts-, Umwelt-, Landwirtschaftsdaten und Daten der öffentlichen Verwaltung, leichter zugänglich und nachnutzbar gemacht werden.¹⁴ Neben diesen Bestrebungen wird im Vorschlag für harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung das Ziel, einen echten Binnenmarkt für Daten zu schaffen, der Digitalstrategie angesteuert und stellt eine Ergänzung des Europäischen Daten-Governance Gesetzes dar, wodurch grundlegende Vorschriften für alle Sektoren erlassen werden sollen.¹⁵ Die Mitgliedsstaaten bleiben von diesen Regelungen nicht unberührt, da sie ebenso zu den Zielen der EU beitragen. Auf Seiten Deutschlands hat daher der IT-Planungsrat bspw. das Kooperationsprojekt zur standardbasierten Digitalisierung des öffentlichen Einkaufs- und Beschaffungsprozesses¹⁶ eingerichtet. In diesem Projekt wird zur Entwicklung der von der EU geforderten Standards beigetragen. Die Beispiele zeigen auch, dass auf EU-Ebene und in Deutschland noch keine Datenmanagementkonzepte/-frameworks existieren, dass diese aber in der Entwicklung sind. In dem EU-Projekt E-Procurement Ontology¹⁷ untersucht das „Publication Office“ der EU in einer Arbeitsgruppe ein mögliches Datenmanagement im Beschaffungswesen der EU. Guasch et.al. haben dieses in einem Pilotprojekt anhand des Tender Electronic Daily (TED)¹⁸ mittels Semantic Knowledge Graphs [AHG11] für Distributed Data Spaces getestet [OHW22][GLV22]. Dabei werden Datensets aus verschiedenen Quellen in ein einheitliches Format transformiert [GLV22]. Die Datentransformation erfolgt mithilfe der RDF Mapping Language (RML) [Di14][GLV22]. RML ist eine generische Abbildungssprache, die individuelle Abbildungsregeln von heterogenen Datenstrukturen und Serialisierungen auf das RDF-Datenmodell ausdrücken [Di14]. Ihr Vorgehen basiert dabei auf DCAT-AP¹⁹, dem deutschen Standard für den Metadatenaustausch für Open Data Portale, und der PROV-Ontologie (PROV-O)²⁰ [GLV22]. PROV-O definiert die OWL2 Web Ontology Language Kodierung [Hi09] des PROV-Datenmodells, einschließlich einiger erweiterter Klassen und Eigenschaften. DCAT-AP wurde ebenso erweitert mittels eines Datenverteilungskonzepts [GLV22]. Damit möglichst keine manuellen Schritte nötig sind, ist eine eigene Web Ontology Language (OWL)²¹ definiert worden, damit der Prozess im Apache Airflow Workflow²² ablaufen kann [GLV22]. Guasch et.al. argumentieren, dass die Daten im öffentlichen Sektor damit leicht verstanden werden können. DCAT-AP wird immer populärer aufgrund der Anforderungen seitens der

¹³ <https://ted.europa.eu/TED/main/HomePage.do>, letzter Zugriff 15.07.23

¹⁴ <https://digital-strategy.ec.europa.eu/de/policies/data-governance-act>, letzter Zugriff 24.04.23

¹⁵ Ibid.

¹⁶ https://www.it-planungsrat.de/fileadmin/beschluesse/2022/Beschluss2022-11_XBeschaffung.pdf, letzter Zugriff 28.04.23

¹⁷ <https://joinup.ec.europa.eu/collection/eprocurement/solution/eprocurement-ontology/about>, letzter Zugriff 28.04.23

¹⁸ <https://ted.europa.eu/TED/main/HomePage.do>, letzter Zugriff 29.04.23

¹⁹ <https://www.dcat-ap.de/def/>, letzter Zugriff 24.05.23

²⁰ <https://www.w3.org/TR/prov-o/>, letzter Zugriff 24.05.23

²¹ <http://www.w3.org/TR/owl-features/>, letzter Zugriff 15.07.23

²² <https://airflow.apache.org>, letzter Zugriff 16.07.23

EU und der Mitgliedsstaaten an Datenkataloge. Weiterhin sei die Bereitstellung eines interoperablen Datenraums der öffentlichen Beschaffung in der EU mittels ihres Vorgehens realisierbar und effektiv [GLV22]. Dadurch wird das Problem der heterogenen IT-Landschaft der verschiedenen Mitgliedsstaaten angegangen und die Zugänglichkeit und die Wiederbenutzbarkeit der Daten erhöht. Einen ähnlichen Ansatz verfolgen Soyly et.al., die sich ebenfalls auf einen Knowledge Graph konzentrieren, indem sie Beschaffungsdaten mit Unternehmensdaten verlinken [So22]. Anders als Guasch et.al. wenden sie den Prozess Extract-Transform-Load (ETL) [CPC20] an, um die Daten über Eigenschaften zusammen zu bringen. Die Beschaffungsdaten werden über Web Scraping, also 400 Scripts, welche Webseiten und Portale nach Daten durchsuchen, erfasst, gesäubert und anschließend auf das Open Contracting Data Specification (OCDS)-Datenmodell²³ gemappt. Ebenfalls werden die Daten auf ein Standardvokabular gemappt. Dabei werden zwei Ontologien verwendet, einerseits zur Darstellung der Beschaffungsdaten anhand des OCDS-Modells und andererseits wird der euBusinessGraph zur Darstellung der Unternehmensdaten wiederverwendet. Die Unternehmensdaten werden hier nicht über vorher bereitgestellte Datensets verfügbar gemacht, sondern über eine Programmierschnittstelle, die täglich Daten von OpenCorporates abrufen. Unternehmensdaten und Beschaffungsdaten werden über LinkedData zusammengebracht und in einem strukturierten Format wieder veröffentlicht. Damit sollen Lücken in bestehenden Datenformaten aufgezeigt werden, wie bspw. ein Mangel an Informationen, eindeutigen Identifizierungen, verschiedene Datenformate und fehlende Standards. Zudem haben die Autoren anhand von Sloweniens öffentlichen Ausgabedaten Anomalien aufgedeckt, die Hinweise auf Korruption bzw. ungleichen Wettbewerb geben [So22]. Tarkovski et.al. setzen auch auf EU-Level an und beziehen sich dabei auf den „Vorschlag für harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung“²⁴. Hierfür stellen sie drei Kriterien auf, erstens muss der Datenzugang und die Fähigkeit diese zu nutzen für mehr Stakeholder sichergestellt werden. Zweitens muss die Kontrolle von Unternehmen über Daten verringert werden. Und drittens sollte die Informationsbereitstellung des Staates gestärkt werden. Hierfür wird eine Verpflichtung zur Bereitstellung von Daten auf der Grundlage klar definierter Kriterien des öffentlichen Interesses vorgeschlagen. Sie schlagen weiter eine Einrichtung einer öffentlichen Datengemeinschaft vor, die diese Daten nicht als ‚Ware‘, sondern als ‚Gemeingut‘ behandelt. Für die Veröffentlichung von Privatdaten schlagen sie einen Test des öffentlichen Interesses der Daten auf der nationalen Ebene und der öffentlichen Datengemeinschaft als letzte Instanz vor. Der Begriff „öffentliches Interesse“ kann allerdings verschieden ausgelegt werden. Die Definition sollte daher laut der Autoren Gegenstand demokratischer, partizipativer Beratungen sein. Allerdings gehen sie davon aus, dass es weitere Spezifizierungen während der Umsetzung des Datengesetzes geben wird. Besonders betonen sie, dass die öffentliche Datengesellschaft ein starkes partizipatives Steuerungsmodell haben soll, um die verschiedenen zivilgesellschaftlichen und wissenschaftlichen Stakeholder miteinzubeziehen [TV22]. Micheli et.al. schlagen ein Modell des Datentreuhändlers vor und stellen vier Datenmanagementkonzepte gegenüber. Sie teilen

²³ <https://standard.open-contracting.org/latest/en/>, letzter Zugriff 16.07.23

²⁴ <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52022PC0068>, letzter Zugriff 28.04.23

die Systeme in Datenaustauschpools, Datenkooperationen, öffentliche Datentreuhändler und die Souveränität über persönliche Daten ein. In diesem Beitrag werden davon zwei im Folgenden aufgegriffen. Datenaustauschpools dienen dem Austausch von Daten zwischen verschiedenen Entitäten (Unternehmen, Behörden, NGO's) auf der Basis von vertraglichen Regelungen über die Datennutzung. Die Kombination der Daten soll laut den Autoren einen Mehrwert schaffen und die Datenzugänglichkeit für Entitäten, die nicht über das nötige Know-How verfügen, erleichtern. Herausforderungen hierbei sind Transaktionskosten und Interoperabilitäts Herausforderungen, die besonders für kleine Unternehmen bzw. Entitäten mit wenig finanziellen Mitteln eine Benachteiligung mit sich bringen [Mi20]. Nach Micheli et.al. verwalten öffentliche Datentreuhändler sämtliche Daten, die von Behörden, Bürgern oder Unternehmen erhoben werden, sodass die Daten zugänglich gemacht werden können. Damit sollen datengetriebene Services und eine Beratung zum Datenaustausch bereitgestellt werden. Ein Hauptziel der öffentlichen Datentreuhändler ist die Integration von Daten aus verschiedenen Quellen, um die Politik zu informieren, Innovationen zu fördern und gesellschaftliche Herausforderungen zu bewältigen, während gleichzeitig ein verantwortungsvoller Ansatz für die Nutzung personenbezogener Daten verfolgt wird. Damit dies gelingt, werden unabhängige Zwischenhändler vorgeschlagen, die dem Bürger Vertrauen vermitteln sollen. Eine weitere Annahme ist, dass Daten, die einen Mehrwert oder eine Service-Generierung für die Öffentlichkeit bieten, auch über die nationale Infrastruktur bereitgestellt werden, selbst wenn sie kommerziell erhoben wurden. Dies funktioniert in der Praxis bisher allerdings nur auf freiwilliger Basis [Mi20]. Projekte werden hier z.B. von der Open Data Initiative (ODI²⁵) durchgeführt.

5 Diskussion & Handlungsempfehlungen

In diesem Kapitel werden die Datenmanagementkonzepte diskutiert und Handlungsempfehlungen benannt. Der Beitrag dieser Arbeit ist die Zusammenfassung von fünf aktuellen Datenmanagementkonzepten und -bestrebungen auf EU-Ebene. Die in Tabelle 1 vollzogene Synthese der verschiedenen Konzepte kann für weitere Ausarbeitungen genutzt werden. Wie in Kapitel 4 bereits beschrieben, folgen einige Konzepte einer ähnlichen Vorgehensweise bzw. enthalten ähnliche Artefakte. Auch wenn es noch kein einheitliches Datenmanagement in der EU gibt, lassen die in Kapitel 4 betrachteten Vorgehen darauf schließen, dass es erste Konzepte und weitere Planungen hinsichtlich eines gemeinsamen Datenmanagements gibt. Ein ausgereiftes Datenmanagementframework liegt jedoch noch nicht vor. Die Ziele der Konzepte umfassen dabei Transparenz schaffen, Datenaustausch fördern, Datenqualität sicherstellen, Daten zugänglich machen und Datenanomalien aufdecken.

²⁵ <https://www.theodi.org/projects-services/projects/>, Letzter Zugriff: 15.07.2023

	Guasch et.al. [GLV22]	Soylu et.al. [So22]	Tarkovski et.al. [TV22]	Micheli et.al. [Mi20]
Anz. Konzepte	1	1	1	4
Untersuchungsobjekt	E-Procurement Ontology	Beschaffungsdaten	Erweiterung harmonisierter Vorschriften	Neu auftretende Datenmanagementkonzepte
Standards	DCAT-AP, RML, Knowledge Graph, PROV-O, OWL	euBusinessGraph, ETL, OCDS, Kriterien-Abgleich	keine	keine
Datensets	Bereitgestellte Datensets	Sammlung von Daten über Web Scraping (400 Scripts), Validierung und Mapping anhand OCDS	Datenstrategie, Datengesetz und Vorschlag für harmonisierte Vorschriften für einen fairen Datenzugang	Literatur und graue Literatur (Zeitschriften, Whitepaper etc.)
Ziel	Semantische Graphen, für einen verteilten Datenraum für das öffentliche Auftragswesen	Anomaly Detection und Unternehmensdaten mit Beschaffungsdaten zu verlinken & zu veröffentlichen	Datenproduktion, Datenaustausch und Datenqualität im B2G fördern	Vier aufstrebende Datenmanagementkonzepte
Ergebnis	1. Deklarativer Ansatz zur Datenumwandlung mit RDF2. Transformations-Ontologie entw. 3. Nutzung EU-Infrastruktur (BDTI) 4. Gemeinsame Sprache im Datenraum & optimierte ETL-Prozessen 5. Probleme: verfügbare Open Data	Klassifizierungsalgorithmus für Beschaffungsdokumente/-daten; Aufdecken von Anomalien am Beispiel von Sloweniens öffentlichen Ausgabedaten	Datentreuhändler als unabhängige Institution errichten, die Daten verteilt und auf ihre Qualität und Relevanz überprüft	Vier Modelle für den öffentlichen Sektor, Wissenschaft, Wirtschaft und die Zivilgesellschaft

Tabelle 1: Vergleich der vier Konzepte

Die vorgestellten Konzepte von Micheli et.al. & Tarkovski et.al. stellen kein Implementationsvorgehen dar, sondern sind als konzeptuelle Modelle zu verstehen. Sie machen Vorschläge für verschiedene Formen einer Institution, einerseits Datenkooperationen, bei denen Daten freiwillig bereitgestellt werden, öffentliche Datentreuhändler, die Datenqualität und Sicherheit gewährleisten und die Pflicht Daten, die für das öffentliche Interesse sind, bereitzustellen. Bei Datenkooperation besteht das Problem, dass Unternehmen bewusst

Daten zurückhalten können, um im Wettbewerb einen Vorteil zu haben, wodurch eine Wettbewerbsverzerrung immer noch stattfinden kann. Anders sieht es bei den Ansätzen für Datentreuhändler aus, die auf Basis von Gesetzen, Standards etc. die Daten erheben, verarbeiten und weiterverbreiten. Damit wäre eine Wettbewerbsverzerrung bzw. ein Vorteil aufgrund von Datenmonopolen ausgeschlossen, sofern der Zugang nicht ausgehebelt wird. Dazu sollte es ein strenges Regelwerk und Vorschriften geben, sodass ein fairer Wettbewerb stattfinden kann, ohne dass große Konzerne Sonderkondition bzw. einzigartige Zugänge zu Daten bekommen. Weiterhin muss ein hohes Maß an Sicherheit gewährleistet werden, gerade in Bezug auf Bürgerdaten. Beim letzten Punkt zur Pflicht der Bereitstellung der Daten, wenn es im öffentlichen Interesse ist, ist ähnlich zu dem öffentlichen Datentreuhändler, wobei hier noch geklärt werden muss, was das öffentliche Interesse ist, und andere Fragestellungen wie Sicherheit und Qualität sind dort nicht mitinbegriffen. Zurzeit fehlen konkrete Vorgaben für die Entscheidung welche Daten von öffentlicher Relevanz sind und wie konkrete Verarbeitungs- und Verbreitungsvorschriften, also Datenaustauschvorschriften, aussehen. Darüber hinaus sollte eine unabhängige Instanz, als zentrale Datensammelstelle, errichtet werden, die für die Datennachfrage und das Daten zur Verfügung stellen, nach Maßgabe der Vorschriften, verantwortlich ist. Die zentrale Datensammelstelle sollte sich dabei an bestehenden Standards orientieren und es sollte ein Regelwerk ausgearbeitet werden, anhand dessen die Datensammelstelle organisiert ist. Dieses Regelwerk sollte die Organisation der Datensammelstelle festlegen, also welche Leitungs- & Steuerungsfunktionen etabliert werden sollten, um den optimalen und rechtlich sicheren Betrieb zu gewährleisten, sodass eine Verletzung oder Missachtung der Datenaustauschvorschriften, Datenschutzvorschriften und der bestehenden Standards ausgeschlossen werden kann. Weiterhin sollte dieses Regelwerk die Interaktion mit den verschiedenen Stakeholdern, Bürgern, Unternehmen, NGOs, Öffentlichen Verwaltungen/Einrichtungen, regeln, sodass ein gleicher und fairer Zugang zu Daten gewährleistet wird. Ein Missbrauch von Daten im Sinne von Auswertungen, die Rückschlüsse auf die Herkunft der Daten geben können, muss vermieden werden. Im Kontext der großen Datenmengen gilt es vor allem die Zugänglichkeit zu den Daten zu erhöhen, bei Gewährleistung der Transparenz, Integrität, Vertraulichkeit und Qualität der Daten. Neben den theoretischen Grundlagen zu Datenmanagement (Kapitel 3) wurden verschiedene Datenmanagementkonzepte im öffentlichen Beschaffungswesen betrachtet (Kapitel 4). Dabei wurde der Fokus besonders auf die EU, Deutschland und die damit verbundenen Strategien, Projekte und Vorschriften gelegt. Dieser Beitrag ist limitiert durch die Suchstrategie, da dort kategorisch alle Werke älter als 2018 ausgeschlossen werden und der Fokus auf dem Beschaffungsmarkt liegt, wodurch andere Datenmanagementkonzepte, die zwar nicht direkt mit dem Beschaffungsmarkt zu tun haben, aber durchaus wichtige Faktoren, die auch eine Rolle für den Beschaffungsmarkt spielen könnten, außen vorgelassen werden. Weiterhin liegt der Fokus sehr stark auf der Europäischen Union, wodurch andere Konzepte ausgeschlossen worden sein könnten. Zudem erheben die vorgestellten Konzepte nicht den Anspruch auf Vollständigkeit der derzeit entwickelten Modelle.

6 Zusammenfassung und Limitation

In dieser Arbeit wurde zunächst ein grundlegendes Verständnis für Datenmanagement bzw. Data Governance geschaffen. Die Untersuchung vorhandener Konzepte hat ergeben, dass in der EU noch viel Unklarheit über standardisierte Formate und ein generelles Konzept besteht, das alle Anforderungen an Open Data erfüllt. Die betrachteten Konzepte widmen sich unter anderem der Qualitätsüberprüfung von Beschaffungsdaten, der Zugänglichkeit, Einheitlichkeit und Transparenz von Daten und äußerten Vorschläge zur Lösung dieser. Der Vergleich der Konzepte zeigt, dass eine zentrale Datensammelstelle eingerichtet werden sollte, um die Daten zugänglich zu machen. Wie das in der Praxis konkret aussieht, ist noch unklar, da dazu eine solche Stelle errichtet und getestet werden muss, um auch konkrete Vorgaben, wie Datenqualität, öffentliches Interesse etc., umsetzen zu können. Weitere Forschung könnte auf die Konzeption einer solchen Institution abzielen und konkrete Vorschläge machen, wie (1) die Institution technisch ausgestattet werden sollte, (2) welche Akteure miteinbezogen werden sollten, (3) welche rechtlichen Rahmenbedingungen geschaffen werden müssen, (4) wer diese Institution leitet und finanziert, und (5) welche Standards entwickelt werden müssen. Weiterhin sind weiterführende Analysen, wie die von Soylu et.al. erforderlich, um Nutzen nicht nur für Slowenien, sondern für alle Mitgliedsstaaten zu identifizieren. Die Liste der in [So22] identifizierten Probleme soll weitergeführt bzw. Lösungen für bereits identifizierte Probleme entwickelt werden.

Literaturverzeichnis

- [ASV19] Abraham, Rene; Schneider, Johannes; Vom Brocke, Jan.: Data governance: A conceptual framework, structured review, and research agenda. *Int. J. of Inform. Mgmt* 49, S. 424–438, 2019.
- [ASD16] Alhassan, Ibrahim; Sammon, David; Daly, Mary: Data governance Activities. An analysis of the literature, *Journal of Decision Systems*, 25:sup1, 64-75, DOI: 10.1080/12460125.2016.1187397, 2016.
- [AHG11] Allemang, Dean; Hender, James; Gandon, Fabien: *Semantic Web for the Working Ontologist. Effective Modeling in RDFS and OWL*, Elsevier, 2011.
- [AK19] Al-Mekhlal, Monerah; Khwaja, Amir Ali: A Synthesis of Big Data Definition and Characteristics, *IEEE*, 2019.
- [Ba11] Ballsun-Stanton, Brian: Asking about data. *Experimental philosophy of Information Technology*, *IEEE Xplore*, 2011.
- [Bü02] Büsing, Heiko: *Das WTO-Übereinkommen über das öffentliche Beschaffungswesen. Eine Untersuchung seines Regelungsgehalts unter besonderer Berücksichtigung der local remedies rule*, Cuvillier Verlag, Göttingen, 2006.
- [Bo06] Bodendorf, Freimut, Hrsg.: *Daten- und Wissensmanagement*. Springer, 2006.

- [CPC20] Corcho, Oscar; Priyatna, Freddy; Chaves-Fraga, David: Towards a new generation of ontology based data access. *Semantic Web*, Vol.11, no.1, pp. 153-160, 2020.
- [Di14] Dimou, A. et.al.: Mapping Hierarchical Sources into RDF using the RML Mapping Language. *IEEE International Conference on Semantic Computing*. DOI 10.1109/ICSC.2014.25, 2014.
- [Fa17] Fazekas, Mihaly: Assessing the quality of government at the regional level using public procurement data, Working Paper 12/2017, Europäische Kommission, 2017.
- [GR22] Gehring, Hermann; Roland, Gabriel: Ausblick – Weiterentwicklungen und Herausforderungen der Wirtschaftsinformatik. In: *Wirtschaftsinformatik*. Springer Gabler, Wiesbaden, 2022.
- [GLV22] Guasch, Cecile; Lodi, Giorgia; Van Dooren, Sander: Semantic Knowledge Graphs for Distributed Data Spaces. *The Public Procurement Pilot Experience. Proc. ISWC 2022*, Springer, 2022.
- [Hi09] Hitzler, Pascal et.al.: OWL 2 web ontology language primer. W3C recommendation, 27(1), 123, 2009.
- [KC18] Kim, Hee Yeong; Cho, June-Suh: Data governance framework for big data implementation with NPS Case Analysis in Korea. *J. of Bus. & Retail Mgmt. Research (JBRMR)* 12, 2018.
- [Ko89] Koller, Rudolf: *Grundlagen der Datentechnik*, Springer, Berlin, Heidelberg, 1989.
- [MP11] Machado, Alexandre Lopes; Parente de Oliveira, Jose Maria: DIGO: An Open Data Architecture for e-Government: 2011 IEEE 15th International Enterprise Distributed Object Computing Conference Workshops. *IEEE*, S. 448–456, 2011.
- [Me16] Meier, Andreas; Kaufmann, Michael: *SQL- & NoSQL-Datenbanken*. Springer, 2016.
- [Mi20] Micheli, Marina et.al.: Emerging models of data governance in the age of datafication. *Big Data & Society*, 7(2). <https://doi.org/10.1177/2053951720948087>, 2020.
- [Ot19] Otto, Boris et.al.: Data Economy – Status Quo der Deutschen Wirtschaft & Handlungsfelder in der Data Economy, Fraunhofer-Institut für Software- und Systemtechnik ISST, 2019.
- [OHW22] Otto, Boris; then Hompel, Michael, Wrobel, Stefan: *Designing Data Spaces. The Ecosystem Approach to Competitive Advantage*. Springer, Cham. 2022.
- [Sa22] Sanchez Graells, Albert: *Identifying Emerging Risks in Digital Procurement Governance*. SSRN, 2022.
- [Sn19] Snyder, Hannah: Literature review as a research methodology: An over-view and guidelines. *Journal of Business Research* 104, S. 333–339, 2019.
- [So22] Soylyu, Ahmetet.al.: Data Quality Barriers for Transparency in Public Procurement. *Information* 2/13, 2022.
- [TV22] Tarkowski, Alek; Vogelzang, Francesco: *Public data commons - A public interest framework for B2G data sharing in the Data Act*, 2022.

- [Vo13] Voß, Jakob: Was sind eigentlich Daten? In: Library Ideas, Nr. 23, S. 4–11. DOI: 10.25969/mediarep/4093, 2013.
- [We07] Wende, Kristin: A Model for Data Governance – Organising Accountabilities for Data Quality Management, 2007.

Overcoming Inefficiency in Public Procurement: An OpenData Approach

Gerhard Klassen¹ , Raphael Palombo¹ , Luca T. Bauer¹ , Björn Niehaves¹ 

Keywords: OpenData; Procurement; Machine Learning; Artificial Intelligence

Abstract: This paper discusses the need for an OpenData platform with data-based services to address the challenges facing the public procurement market. In the last 15 years, public procurement has doubled and now accounts for 15% of GDP. However, there is a shortage of skilled workers, and many tenders are still created manually. By leveraging advanced technologies such as machine learning and predictive analytics, we aim to improve the efficiency and effectiveness of public procurement. Our paper highlights the urgent need for a data-driven approach to public procurement and presents our plans for an OpenData platform that can deliver significant benefits to both the public sector and private enterprises.

1 Introduction

The German government's commitment to public data is in line with its open data strategy [BdI21], which aims to promote transparency, participation, and innovation. The strategy encourages the provision of data by public institutions in a machine-readable format that is easily accessible to the public. Additionally, the OECD Government Data Report [OE18] highlights the importance of data sharing by which governments can improve decision-making processes, enhance service delivery, and stimulate economic growth.

The public procurement holds a particular significance in this context as it represents an important resource for public institutions and is utilized by businesses to bid for public contracts. However, the public procurement process is often characterized by opacity and inefficient processes. In fact, the procurement process in Germany alone cost taxpayers approximately €14.3 billion in 2022, with research and document creation accounting for 22.7% (€3.25 billion) and 19.3% (€2.76 billion) of the costs, respectively [FW19]. The bidding companies also incurred expenses totalling €16.5 billion. Overall, the public procurement in 2022 had a market volume of €500 billion, equivalent to 35% of government spending or 15% of the GDP [FW19]. Moreover, political developments, such as the provision of aid to Ukraine [Bu23], arms procurement [Bu22], and energy-related procurements, are expected to drive up the procurement volume and costs even further.

¹ University of Bremen, AG Digital Public, Bibliothekstraße 1, 28359 Bremen, Germany
{gerhard.klassen,palombo,lbauer,niehaves}@uni-bremen.de

In the last 15 years, there has been a marked increase in the volume of tenders, indicating a growing demand for procurement services. However, this increase in volume may have imposed a significant burden on employees responsible for tender processing. Recent research conducted by PricewaterhouseCoopers International [Fa] suggests that, by 2030, the public sector in Germany may face a shortage of approximately one million skilled workers. Given this prediction, it is imperative to ensure that employees receive adequate support to continue producing high-quality work. Failure to provide such support could not only lead to an increase in processing errors and a longer processing time, but also increase the costs.

The potential of OpenData in the public procurement has been recognized by the German government, as evidenced by the establishment of the *Bekanntmachungs-Service*² and *service.bund.de*³. Although these platforms provide structured data on tenders, they only do so at a meta level. To achieve sustainable efficiency improvements, it is essential to provide standardized, machine-readable data that includes the contents of the tenders, enabling procurement professionals to better analyze the data, build higher level services and make informed decisions. This working paper proposes an OpenData platform for structured access to tender contents, with the aim of enhancing transparency and efficiency in public procurement. Our paper proposes an OpenData platform for structured access to tender contents, with the aim of enhancing transparency and efficiency in public procurement. Through a literature review and analysis of existing technologies, we have developed a framework that has the potential to transform the tendering process.

2 Status Quo

The comprehension of the present procurement process is of pivotal significance in facilitating any necessary alterations to the process. With the aim of improving understanding of the current procurement process, we present an abbreviated example of a German school seeking to purchase new chairs:

The procurement process of a German school that needs to purchase chairs typically follows a specific procedure. First, a needs assessment is conducted to determine the exact requirement for new chairs. This includes identifying the number of chairs needed as well as the desired size, material, and functionality. Next, a requirement profile is created, which summarizes the needs and criteria for the chairs.

The next step involves the creation of the procurement documents, which contain all the necessary information that potential vendors need to submit a bid. This includes the requirements for the chairs, as well as the delivery and payment terms. The tender notice is electronically published on one of the procurement platforms. Potential vendors can then submit their bids, which are subsequently evaluated and assessed by the school.

² <https://www.bescha.bund.de/>

³ <https://www.service.bund.de/>

In the final step, the contract is awarded to the vendor who offers the most economically advantageous bid and demonstrates the best qualifications. The school then enters into a contract with the chosen vendor, agreeing on delivery and payment terms. Once the chairs have been successfully delivered and accepted, the procurement process is complete.

The most elaborate part of the whole process is the creation of the necessary documents to comply with the legal requirements. Currently, the creation of tender documents is being carried out manually for 65% [FW19]. This is not only prone to errors but also cost-intensive and therefore has a high potential for optimization.

Artificial intelligence can be used to assist the tender document creation, resulting in lower expenses, workload and less time. These benefits not only extend to businesses, which can use tools to automatically apply, but also to the public sector, which can perform tasks more efficiently.

3 System Design

In our work, we regard the underlying problem as a data mining process and have employed relevant methodologies to address it. Currently the DCAT-AP.de metadata format is established for OpenData publishing in Germany. [We21]. In contrast to the DCAT-AP.de metadata format, textual content data from previous procurement processes is provided and proceeded in the system. No sufficient data standard for content data operations in OpenData platforms is known to us, therefore we define an own data format that is suitable to our system. Our approach draws upon an adaption of CRISP-DM (Cross-Industry Standard Process for Data Mining)[Sh00], a de facto standard [Sc20] and widely recognized for its effectiveness in the data mining and analytics domain [Lo22, Sc20]. The original approach provides guidance through six distinct phases, namely Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment, providing a clear framework for managing the complexities inherent in such projects. Figure 1 presents a comparative analysis of two models.

To enhance the efficiency of our project, we decided to merge the *Business Understanding* and *Data Understanding* phases into the *Procurement Understanding* phase. By doing so, we are better equipped to identify any minor data quality issues earlier in the development process. Moreover, we believe that a comprehensive understanding of procurement is closely linked with the understanding of tender documents, thereby making the process more agile. In addition, we included the *Data Extraction* phase, which involves scraping tasks, as the tender documents are not available through an API on the platforms. We kept the *Data Preparation* phase unchanged since it is heavily dependent on model selection and should be revisited multiple times. The next two phases of the CRISP-DM model, *Modelling* and *Evaluation*, were merged again since the evaluation is crucial during the model creation stage. Separating the two could potentially create an impression that evaluation is

not as important as modelling. Finally, we renamed the *Deployment* phase to better reflect our goal of providing specific *Data based services*.

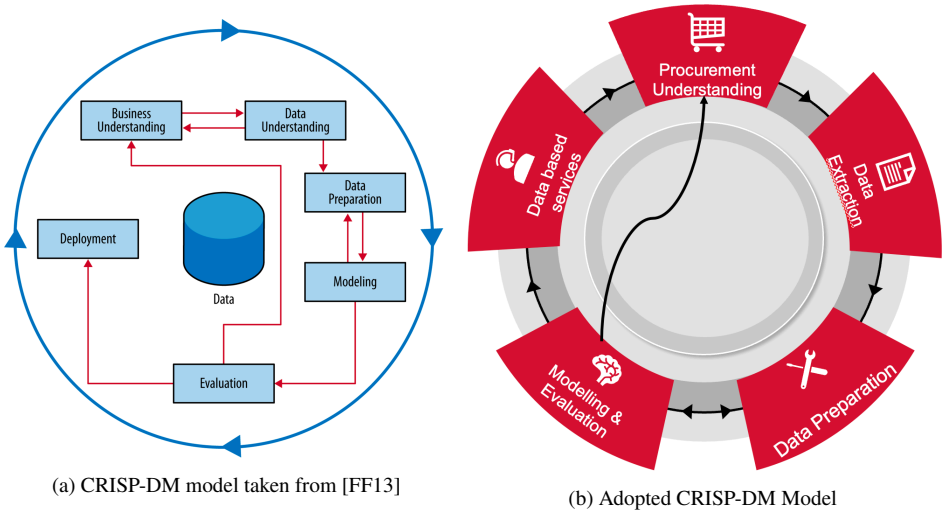


Fig. 1: CRISP-DM and our adopted version of CRISP-DM

4 Outlook

Our forthcoming objective is to proceed with the implementation of the system. Therefore we identified a total of 491 online-procurement platforms that exist within Germany. Our proposed strategy entails utilizing either web scrapers or, if available, API based extraction methods. Subsequently, we intend to employ a language model to standardize the extracted data. However, the main challenge here is to define a data standard and choose an appropriate language model.

In advance to the definition of the data standard it is precisely examined, how historic procurement data can be provided without violating confidential aspects. Defining a data standard involves determining the structure, format, and quality of the data to ensure its compatibility with the system. Choosing a suitable language model is crucial as it plays a significant role in the accuracy and efficiency of the system. Our team is working diligently to address these challenges and ensure a smooth implementation process.

Our primary focus is on implementing various services to provide maximum value. We have already identified several potential services and intend to refine and develop them further during the implementation phase.

In the following we briefly introduce two of the planned services:

1. Searching for similar procurements

The process of searching for similar tenders has been shown to be a highly effective means of reducing the time and resources required to create new tender documents. By leveraging the knowledge and insights gained from previous tenders, organizations can streamline the creation process and minimize the need for extensive rework. In many cases, the only adjustments that need to be made to the tender documents are related to specific requirements or unique aspects of the current tender. By reducing the need for extensive rework, this approach can help organizations to save valuable time and resources, while still ensuring that the tender meets all necessary criteria and requirements. The focus lays on the evaluation of service descriptions, which are highly time consuming in the procurement process. To find similar tenders from different contexts the algorithm is additionally trained to compare tender formats from different contexts (e.g. construction or delivery) and considers these aspects in the prediction.

2. Generating tenders

On basis of the identified previous tender documents a generative adversarial network (GAN) can be trained to generate new tender-text which contains central elements of the identified documents. Additionally the network is trained to condense the contents from previous tenders with new requirements from the actual procurement to create direct suggestions on basis of the actual needs. This process involves additional machine learning techniques, including Natural Language Processing (NLP) to identify key-data and replace it with data and terms that are needed for the new procurement process.

3. Predicting the best bid

We plan to predict the most optimal bid, even before the tender is published. This advanced insight allows the tenderer to assess whether they have the financial capacity to proceed with the tender, and make any necessary adjustments as required. This approach is highly beneficial for the public sector as it allows to make informed decisions and reduce the risk of financial strain or overcommitment. For this purpose a machine learning data service operates on the successful bids and predicts a success rate for previous actors in similar procurement processes.

4. Support decision making

The system primarily supports the user in the decision making process. The final decision has to be made by a qualified procuring employee in public bodies. In complex procurement situations, the system must assist the user by giving predictable and clear information, why a specific bid is outstanding in comparison to the others. For this purpose the system aligns tenders and bids while operating on the prepared data sets of the CRISP-DM to support the user in the selection process with data-based services. The data is presented in an evaluation matrix that contains hard and soft criteria to choose the optimal bid. The hard criteria are nonnegotiable aspects of a tender that are crucial for the success of a bid. Soft criteria are negotiable, but can be decisive in a close race between bids. In favour of the user the result can be altered. In this case the user can set parameters to alter the decision making process.

Bibliography

- [BdI21] Bundesministerium des Innern, für Bau und Heimat: , Open data strategy by the German Federal Government. <https://rdm.mpd1.mpg.de/2021/09/06/open-data-strategy-by-the-german-federal-government/>, 2021. Accessed on March 28, 2023.
- [Bu22] Bundesregierung: , 100 Milliarden Euro für eine leistungsstarke Bundeswehr. <https://www.bundesregierung.de/breg-de/themen/sicherheit-und-verteidigung/sondervermoegen-bundeswehr-2047518>, 2022. Accessed on March 28, 2023.
- [Bu23] Bundesregierung: , Military support for Ukraine. <https://www.bundesregierung.de/breg-en/news/military-support-ukraine-2054992>, 2023. Accessed on March 28, 2023.
- [Fa] Fachkräftemangel im öffentlichen Sektor - PwC. <https://www.pwc.de/de/branchen-und-markte/oeffentlicher-sektor/fachkraeftemangel-im-oeffentlichen-sektor.html>. (Accessed on 05/09/2023).
- [FF13] Foster; Fawcett, Tom: Data science for business. O'Reilly Media, Sebastopol, CA, August 2013.
- [FW19] Ferber, Thomas; Wagner, Christian-David: Erfolgreich öffentliche Aufträge gewinnen (E-Book). Praxisratgeber Vergaberecht. Bundesanzeiger, Köln, Germany, 2 edition, February 2019.
- [Lo22] Lofandri, Wiki: Analisis Predictive Maintenance Peralatan Lab Berbasis Machine Learning. Jurnal Sistim Informasi dan Teknologi, 2022.
- [OE18] OECD: Open Government Data Report. 2018.
- [Sc20] Schröder, Christoph: A Systematic Literature Review on Applying CRISP-DM Process Model. 2020.
- [Sh00] Shearer, Colin: The CRISP-DM model: the new blueprint for data mining. Journal of data warehousing, 5(4):13–22, 2000.
- [We21] Wenige, Lisa; Stadler, Claus; Martin, Michael; Figura, Richard; Sauter, Robert; Frank, Christopher: Open Data and the Status Quo – A Fine-Grained Evaluation Framework for Open Data Quality and an Analysis of Open Data portals in Germany. 06 2021.

Öffentliche Infrastruktur -
CMS-Verteilung im öffentlichen
Bereich in Deutschland -
Erkenntnisse und Vorschläge für
Konsolidierung

CMS-Verteilung im öffentlichen Bereich in Deutschland - Erkenntnisse und Vorschläge für Konsolidierung

Julian Fastnacht¹, Sebastian Kreideweiß², Phoenix Lußky³

Abstract: Demografischer Wandel, nicht nur Fach- sondern Arbeitskräftemangel werden auch im Bereich der Digitalisierung in öffentlichen Einrichtungen den Ruf nach Konsolidierung, Automatisierung und Vereinfachung weiter erhöhen. Der Workshop soll einen Lagebericht über den Digitalisierungsstand des zentralen Kommunikationsmittels Website, z.B. von Ländern, Landkreise und Gemeinden und den Möglichkeiten (Online-Zugangsgesetz/OZG, Open Data), aber auch von anderen Einrichtungen mit öffentlichem Interesse, wie Sportvereinen, Kultur- und Bildungseinrichtungen in Deutschland geben. Im Workshop wird das Potential von Open Data mit Methoden der Data Science (Wikidata, SPARQL) interaktiv erfahrbar.

Keywords: Statistik, WWW, Website, Content Management Systeme, CMS, Verteilung, Nutzungshäufigkeit, www.cmscensus.eu, HTML, Detektion, FOSS

1 Motivation

Die Motivation für diese Arbeit stützt sich auf 4 Säulen: Transparenz des Marktes, Entscheidungshilfe für Vergabeverfahren, Unterstützung bei der Konsolidierung und Nachvollziehbarkeit von strategischen Entscheidungen.

Die Transparenz des Marktes soll eine Übersicht geben, welche Technologien verwendet werden. Dabei wird keine konkrete Zielgruppe angesprochen, sondern es stellt vielmehr ein Angebot dar, die erhobenen Daten für verschiedene Zwecke nutzen zu können.

Einer dieser Zwecke ist eine Entscheidungshilfe in Vergabeverfahren, öffentlich wie privatwirtschaftlich. Dabei richtet sich das Angebot an Entscheider, die im Vergabeverfahren eine Übersicht über die verwendeten Technologien bekommen wollen. Dies ist insbesondere dann hilfreich, wenn für den jeweiligen Einsatz auf bewährte Technologien gesetzt werden soll.

Weiterhin soll es in einem ähnlichen Kontext als Unterstützung bei der Konsolidierung von heterogenen Systemlandschaften dienen. Mit einer entsprechend guten Datengrundlage, lässt sich schneller die Verteilung der verschiedenen Systeme in einer

¹ coding. powerful. systems. CPS GmbH, Entwicklung, Gustav-Meyer-Allee 25, 13355 Berlin, j.fastnacht@familie-redlich.de

² coding. powerful. systems. CPS GmbH, Entwicklung, Gustav-Meyer-Allee 25, 13355 Berlin, s.kreideweiss@familie-redlich.de, <https://cmscensus.eu/about#c289>

³ coding. powerful. systems. CPS GmbH, Entwicklung, Gustav-Meyer-Allee 25, 13355 Berlin, p.lussky@familie-redlich.de

Organisation ermitteln. Und selbst bei der Nutzung eines Systems können verschiedene Versionen im Einsatz sein, bei denen sich eine Konsolidierung lohnen könnte.

Zuletzt ist im Nachgang einer Entscheidung die Nachvollziehbarkeit wichtig. Insbesondere in Hinblick auf öffentliche Ausschreibungen und Vergabeverfahren bei denen Open Source Software sowohl auf EU-Ebene, wie auf nationaler Ebene (Vgl. [BUND2021, S. 15]) im Fokus steht. Dabei sollte neben der Nachvollziehbarkeit der Datengrundlage die Nachvollziehbarkeit des Prozesses gewährleistet sein. Das lässt sich bei der Automatisierung des Vorgangs vor allem durch den Einsatz von Open Source Technologien ermöglichen, die jedem Außenstehenden einen Einblick in die Vorgänge ermöglichen. Die Ausführung erfolgt dabei nach wissenschaftlichen Methodiken.

2 Fallstudie Deutschland

Durch Judith Kölbel wurde im Jahr 2022 im Rahmen ihrer Bachelorarbeit [KÖLBEL2022] eine Reihe von Websites des öffentlichen Raums untersucht. Dabei entstand eine Kategorisierung des öffentlichen Raums, welche als Grundlage für die Datenrecherche diente. Die Websites wurden auf ihr Erstellungssystem analysiert und erlauben die Betrachtung einzelner Kategorien oder Einrichtungen, sowie eine zusammengenommene Betrachtung aller öffentlichen Einrichtungen. Die jeweiligen Linklisten stammen dabei jeweils aus unterschiedlichen Quellen.

Gruppierung für die Auswertung nach [KÖLBEL2022, S. 12]:

- Bildung, Erziehung und Betreuung (Kinderbetreuung, Schulen, Hochschulen, Akademien, Fachschulen, Erwachsenenbildung)
- Öffentliche Verwaltung, Verteidigung und Sozialversicherung (Bundes-, Landes-, Kommunalverwaltung)
- Gesundheit, Pflege und Fürsorge (Kliniken und Krankenhäuser, Heime und Pflegeeinrichtungen, Medizinische Versorgungszentren, soziale Einrichtungen)
- Kunst und Kultur (Museen, Theater, Galerien, Musikschulen)
- Sport und Freizeit (Sportvereine, Schwimmbäder, Zoos)
- Wissenschaft und Forschung (Bibliotheken, Archive, Institute, Forschungsanstalten/-einrichtungen/ -gesellschaften)
- Verkehr, Mobilität und Kommunikation (Postwesen, Telekommunikation, Nachrichtenübermittlung, Personenbeförderung, Häfen/Wasserstraßen, Flughäfen)
- Ver- und Entsorgung (Müll, Wasser, Energie)

Als Ergebnis der Analyse werden in der folgenden Tabelle die jeweils fünf bestplatzierten Content-Management-Systeme (CMS) aufgelistet. Nicht gelistet, jedoch zum Teil

signifikant, ist der Anteil der nicht-detektierbaren Erstellungssysteme. Weitere Details zur Anzahl untersuchter Websites, zur Anzahl der Gesamtmenge und zur prozentualen Verteilung lassen sich in [KÖLBEL2022] nachlesen. Weitere nachträglich analysierte Kategorien wie Gemeinden in Deutschland sind auf [CMSCENSUS2023] bereitgestellt.

KAT.	PLATZ 1	PLATZ 2	PLATZ 3	PLATZ 4	PLATZ 5
Schulen	WordPress	Joomla!	TYPO3	Jimdo	Drupal
Hochschulen	TYPO3	WordPress	Drupal	Contao	Joomla!
Einrichtungen Bund	TYPO3	Government Site Builder	CoreMedia CMS	Adobe Experience Manager	WordPress
Städte	TYPO3	WordPress	iKISS	Joomla!	Contao
Krankenkassen	TYPO3	WordPress	Contao	Adobe Experience Manager	Drupal
Sportvereine	WordPress	Joomla!	TYPO3	Jimdo	1&1 IONOS Website Builder
Krankenhäuser	TYPO3	WordPress	Contao	Joomla!	Weblication
Rehabilitation	TYPO3	WordPress	Government Site Builder	Joomla!	Contao
Forschung	TYPO3	WordPress!	Drupal	Adobe Experience Manager	Joomla!

Tab. 1: Top 5 CMS der untersuchten öffentlichen Einrichtungen [KÖLBEL2022, S. 44]

3 Methodik

Das Vorgehen gliedert sich in 3 Schritte, die Erhebung von Analysezielen, die Ermittlung der verwendeten Technologien dieser Ziele und der Erkenntnisgewinn.

3.1 Erhebung von Untersuchungsgegenständen

Um eine Auswertung durchzuführen, müssen zunächst Analyseziele bestimmt werden. Diese werden nach den in Kapitel 2 benannten Kategorien eingeordnet. Die Erhebung von Analysezielen ist dabei vorrangig eine manuelle Aufgabe.

Eine gute Quelle für die Erhebung von Untersuchungsgegenständen stellt Wikidata dar. Wikidata ist ein Projekt der Wikimedia, welche für die freie Enzyklopädie Wikipedia verantwortlich ist. Es stellt dabei vor allem der Wikipedia strukturierte Daten zur Verfügung, die in Wikipedia-Artikeln referenziert werden können [WIKIDATA2023]. Durch die Nutzung des SPARQL Query Service von Wikidata

(<https://query.wikidata.org/>) können zielgerichtet Informationen abgefragt werden. So ist es bspw. möglich eine bestimmte Kategorie von Daten abzufragen, die eine Website enthalten.

Daneben gibt es bspw. Branchenlisten, die manuell zusammengetragen wurden. Dabei sind nicht alle Branchenlisten frei verfügbar und es müssen die jeweiligen Nutzungsrechte beachtet werden. Nicht immer sind diese auf den Seiten erkennbar hinterlegt.

Unabhängig von der jeweiligen Datenquelle sollten die Analyseziele plausibilisiert werden. Das kann bspw. durch den Abgleich mit den Statistikdaten des Bundesamtes für Statistik erfolgen, um eine Referenzgröße zu bestimmen. Weicht diese Referenzgröße nicht zu sehr von der Menge der erfassten Analyseziele ab, so können die Analyseziele zumindest als plausibel angesehen werden.

Sind die Analyseziele erfasst und plausibilisiert, dann erfolgt eine weitere Bereinigung der Daten. Dabei sollen Dopplungen und unpassende Datensätze entfernt und unvollständige Datensätze ggf. angereichert werden. Das trifft z. B. auf fehlende oder veraltete URLs zu, kann aber ebenso Schreibfehler in den URLs betreffen, die korrigiert werden müssen. Ist dies nicht direkt über die jeweilige Datenquelle möglich (z. B. durch Anpassung des Datensatzes in Wikidata), muss die Anwendung eine Anreicherung der Daten zulassen. Außerdem wird strichprobenartig geprüft, ob die Kategoriezuteilung stimmt.

Dabei ist weiterhin zu beachten, dass der gesamte Vorgang wiederholbar sein muss. Das ist wichtig, um spätere Änderungen nachvollziehen und abbilden zu können. Umso weniger Arbeit dabei in manuelle Prozesse fließt, umso besser für die spätere Wiederholung des Prozesses. Sollte als Datenquelle bspw. Wikidata genutzt worden sein, so wird die für die Erhebung genutzt SPARQL-Query hinterlegt, um ihn zu einem späteren Zeitpunkt wieder auszuführen. Des Weiteren ergibt sich daraus eine Nachvollziehbarkeit bei der Datenerhebung.

3.2 Ermittlung der verwendeten Technologien

Sobald die Daten erhoben und so weit wie möglich bereinigt wurden, erfolgt die Ermittlung der verwendeten Technologien. Dafür wird der Dienst WhatCMS genutzt.

WhatCMS analysiert Websites anhand ihrer URL auf das jeweils verwendete Content Management System. Dazu verwendet es verschiedene Strategien der Erkennung von CMS, wie etwa HTML-Elemente, HTTP Header oder Verzeichnisstrukturen, die sich über die URLs ermitteln lassen. [WHATCMS2023]

Die zu analysierenden Zielsysteme werden dafür zunächst in die Plattform www.cmscensus.eu [CMSCENSUS2023] eingetragen und anschließend durch einen automatisierten Prozess an WhatCMS gesendet. CMSensus ist eine auf TYPO3 basierende Web-Anwendung und wurde von [ULLRICH2022] konzipiert und entwickelt.

Die Rückgabewerte von WhatCMS werden in der Datenbank von CMSscensus gespeichert.

Sobald der Prozess abgeschlossen ist, lassen sich die neuen Kategorien über CMSscensus aufrufen. Auf der Seite wird anschließend jeweils eine tabellarische Aufstellung und ein Tortendiagramm für die Verteilung der verwendeten Content Management Systeme dargestellt. Einzelergebnisse einer URL sind einsehbar und somit überprüfbar.

3.3 Auswertung der Ergebnisse

Je nach Kategorie und Anwendungszweck können nun Erkenntnisse aus den jeweiligen Ergebnissen gezogen werden. Diese können in Entscheidungsprozesse, ähnlich [OSOR2023a], zusammen mit anderen Faktoren einfließen. Dazu kann es eventuell notwendig sein, die Daten aus CMSscensus in andere Werkzeuge für die weiterführende Auswertung zu überführen.

4 Erkenntnisse

Nach Abschluss des Vorgangs ließen sich bisher verschiedene Erkenntnisse aus den Daten ziehen.

Trotz hoher Anzahl detektierbarer Systeme (ca. 1.500) besteht eine beträchtliche Anzahl von unerkannten Erstellungssystemen (nicht immer Content Management Systemen), was auch [KÖLBEL2022, S. 47] bereits abschließend festgestellt hat. Da WhatCMS weiterhin an der Erkennung neuer und bestehender Content Management Systeme arbeitet, kann sich daher bei einer wiederholten Ausführung ohne signifikante Änderungen an den Analysezielen ein leicht unterschiedliches Ergebnis zeigen.

Des Weiteren kam [KÖLBEL2022, S. 44] zu dem Schluss, dass in den betrachteten Teilbereichen des öffentlichen Raums in Deutschland vorrangig TYPO3 CMS und WordPress eingesetzt werden.

5 Fazit und Ausblick

Das Verfahren hat das Potenzial, strategische Entscheidungen für Technologien in EU-Ländern und darüber hinaus anhand von Daten zu untermauern [OSOR2023a].

Nach Deutschland sollen mit Estland (begonnen durch [LYUO2023]) noch weitere Länder der EU in [CMSSENSUS2023] hinzukommen und analysiert werden, um länderbezogene Entscheidungsprozesse analog Deutschland (siehe [OSOR2023b]) zu unterstützen. Ähnliche Analysen wie in den Niederlanden sollen eingebunden und angereichert werden [CMS_NL2023a], [CMS_NL2023b], [CMS_NL2023c]. Anschließend soll eine

Empfehlung in Richtung der Europäischen Kommission zum Einsatz von Open Source im Bereich Websites für öffentliche Einrichtungen in der EU gegeben und vollzogen werden.

Ausblick auf beabsichtigte Weiterentwicklungen von CMSscensus:

- Weitere Analyse nicht detektierbarer CMS-Systeme
- Darstellung von Versionsverteilung detektierter CMS
- Darstellung zu Grunde liegender und eingesetzter Technologien, z.B. Programmiersprache, Datenbank, Web-Server, TLS/SSL, CDN, etc
- Weiterführende Informationen zum CMS, insb. zur Open Source/Closed-Source Einordnung
- Semantische Auszeichnung der Dienstleistungsangaben von CMS-Anwendungen anhand [HABICH2023b] in Anlehnung an Auszeichnung von Hochschuldaten [HABICH2023a]

Ziel: Als Entscheider*in sollen der TechStack und alle Referenzen einer dienstleistenden Firma betrachtet werden können. Auch relevant aus Sicht von Firmenfusionen. Zudem: Als Betreiber mehrerer Web-Anwendungen soll die Heterogenität der IT-Landschaft dargestellt werden können, um daraus Konsolidierungsstrategien abzuleiten. Für Transparenz in Vergabeverfahren soll die Dauer bestehender Dienstleisterbeziehungen aus vorangegangenen Verfahren dargestellt werden, um eine Chancenbewertung von Marktteilnehmenden zu ermöglichen.

Literaturverzeichnis

- [BUND2021] Bundesregierung, Koalitionsvertrag 2021 - Mehr Fortschritt Wagen, <https://www.bundesregierung.de/resource/blob/974430/1990812/1f422c60505b6a88f8f3b3b5b8720bd4/2021-12-10-koav2021-data.pdf>, 30.06.2023
- [CMSCENSUS2023] Statistische Verteilung von CMS nach Kategorien, <https://cmscensus.eu/>, 30.06.2023
- [CMS_NL2023a] CMS usage in Netherlands, <https://basisbeveiliging.nl/>, 30.06.2023
- [CMS_NL2023b] CMS usage in Netherlands, <https://www.200ok.nl/cms-gemeenten/>, 30.06.2023
- [CMS_NL2023c] CMS usage in Netherlands, <https://www.digimonitor.nl/cms-en/gemeenten/>, 30.06.2023
- [HABICH2023a] Bereitstellung offener Daten aus öffentlichen Einrichtungen in Deutschland durch Content-Management-Systemen am Beispiel von Hochschuldaten und TYPO3, Marco Habich, BSc-Thesis TH

Brandenburg,

https://cmscensus.eu/fileadmin/introduction/content/BSc_HabichMarco_BoDaoE_Webversion.pdf, 06.07.2023

[HABICH2023b]

Beispielhafter Aufbau JSON-LD zur semantischen Auszeichnung von Dienstleistern für Websites, Marco Habich, TH Brandenburg, <https://cmscensus.eu/fileadmin/introduction/content/cms-provider-JSON-LD.zip>, 14.07.2023

[KÖLBEL2022]

Statistische Verteilung von Content-Management-Systemen im öffentlichen Bereich in Deutschland, Judith Kölbl, Deutschland, BSc-Thesis BHT Berlin, https://cmscensus.eu/fileadmin/introduction/content/BA_JudithKoelbel_2022.pdf, 30.06.2023

[LYUO2023]

Statistical distribution of content management systems in the public sector in Estonia, Yunao Lyu, China/UK, BSc-Thesis TH Lübeck, <https://cmscensus.eu/fileadmin/introduction/content/BachelorThesisYunaoLyu.pdf>, 30.06.2023

[OSOR2023a]

CMS Census helps public administrators choose the right Open Source solutions, Open Source Observatory (OSOR), <https://joinup.ec.europa.eu/collection/open-source-observatory-osor/news/cms-census-helps-ctos-choose-right-oss-solutions>, 30.06.2023

[OSOR2023b]

German government CMS switches to TYPO3, Open Source Observatory (OSOR), <https://joinup.ec.europa.eu/collection/open-source-observatory-osor/news/german-government-cms-switches-typo3>, 30.06.2023

[ULLRICH2022]

Konzeption und Implementierung einer TYPO3-Extension zur Analyse und Darstellung von statistischen Daten gegebener Content Management Systeme, Alexander Ullrich, Deutschland, BSc-Thesis BHT Berlin, 01/2022, <https://cmscensus.eu/fileadmin/introduction/content/Bachelorarbeit-Ullrich.pdf>, 30.06.2023

[WHATCMS2023]

WhatCMS.org, <https://whatcms.org/>, 30.06.2023

[WIKIDATA2023]

Wikidata, https://www.wikidata.org/wiki/Wikidata:Main_Page, 30.06.2023

GI Quantum Computing Workshop

quark: QUantum Application Reformulation Kernel

Elisabeth Lobe ¹

Abstract: Quantum annealers solve Ising problems heuristically. Several standard methods have been established to transform more complex problems into the Ising problem format, which are commonly still applied by hand. In this work, we present our software package quark, automating the full transformation process from an arbitrary discrete optimization problem to the corresponding Ising problem. Based on a parameterized formulation of the original problem, a series of easily reproducible experiments can thus be set up. This allows users to evaluate the suitability of the annealing machines in solving their specific problem without a deeper knowledge about the Ising problem specifics.


Keywords: Ising Problem, QUBO, Discrete Optimization, Quantum Annealing, Quantum Computing

1 Introduction

Quantum annealing is a method aiming at optimizing certain objective functions by taking advantage of a quantum mechanical process, the adiabatic evolution towards a quantum system representing the optimal solution in its ground state [Fa00]. Hereby, several physical effects might disturb the annealing process leading to sub-optimal solutions, which is why the corresponding devices, the quantum annealers, are considered as heuristic samplers [Mc20].

The implemented problems are known as Ising problems, which search for the minimum of a quadratic function over spin (+1/−1) variables without any further constraints. In the mathematical community, the notation of quadratic unconstrained binary optimization (QUBO) problems is more common, which refers to the same type of problem but acting on binary (0/1) variables. These problems are known to be NP-hard in general [Ba82] and therefore also relate to a multitude of other problems for which the best known classical algorithm takes exponential runtime, and which are thus intractable by classical devices above a certain size of the problem instance.

For a lot of, in general rather academic, combinatorial optimization problems the corresponding Ising problem formulations are known, such as for maximum cut or graph coloring. See [Lu14] for an extensive list of examples. Several more advanced applications have also already been reformulated, e. g., the flight-gate assignment (FGA) problem as part of the airport planning [SLJ19]. A lot of such industrial use cases refer back to some academic example, like the FGA to the quadratic assignment problem. However, the experience from

¹ Institute for Software Technology, German Aerospace Center (DLR),
Lilienthalplatz 7, 38108 Braunschweig, Germany
elisabeth.lobe@dlr.de  <https://orcid.org/0000-0002-3473-8906>

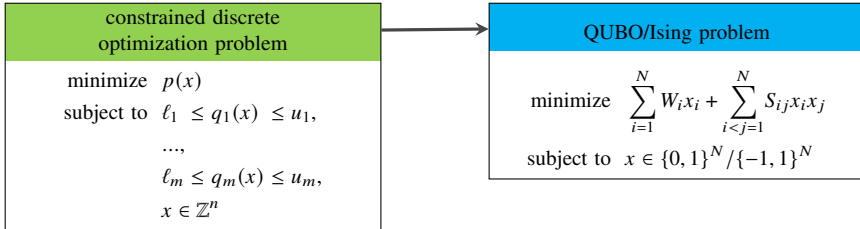


Fig. 1: Main transformation step automatically performed with quark for providing the problem in a format the quantum annealers can process

these applications shows that their actual formulations differ in the details, hence also require individually derived Ising/QUBO formulations.

In Fig. 1, we have depicted the general main transformation step with the corresponding mathematical formulations of the problems. Several standard methods have been established to perform this step and to reformulate given arbitrary discrete optimization problems [G122]. They are usually applied manually, requiring a deeper understanding of the properties of Ising problems rather than only the domain-specific knowledge of the original optimization problem. This demands for an automation of the full process to shift the complexity away from the end-users, such that they are able to easily produce a large test set and evaluate the capabilities of the solver for their specific problem.

2 Package Description

We have developed and constantly extend a software library at <https://gitlab.com/quantum-computing-software/>, whose core package is quark [Ge23]. In this section, we briefly summarize the main goals of quark and explain its general usage along an example.

2.1 Goals and Main Features

Our package quark is designed to provide a user-friendly, low-level entry point for domain experts to deal with quantum annealers. Based on a parameterized formulation of the problem, several problem instances can be generated with the same structure. The corresponding optimization problems can then be reformulated to Ising problems by automated transformation steps, which enables the user to perform easily reproducible experiments over a large test set. This allows an analysis of the machine behaviour and experimental results more suited to the domain demands, providing hints whether the problems are suitable to be solved with the annealers. This is supported by an interface to a classical optimizer.

D-Wave already provides an extensive software library surrounding their quantum annealers, also already implementing several transformation steps [D-23a]. We do not want to compete

with but rather accompany D-Wave's API with our instance-centric approach focusing on the original problem. A particular example where we simplify the provided functionality is the handling of polynomials with degree larger than two, which is available through the D-Wave API only via detours, cf. [D-23a], which means that the user needs to know about the structural differences. In quark no additional but just the base classes are necessary. With the step from a constrained to an unconstrained problem, reduction variables with corresponding penalty terms are introduced and the degree of the polynomials is thus reduced automatically.

From our own experience in experimenting with the annealing machines, we have also seen the necessity of being able to store or load intermediate data, such that scripts can be interrupted and restarted at any stage of the transformation process. Therefore, all of the base classes are accompanied with input-output functionality. As we use the Hierarchical Data Format (HDF), stored data can quickly be retrieved and analysed by the user.

2.2 Usage

In this section, we briefly summarize the main steps and concepts necessary to perform the transformation with quark, where the highlighted terms indicate classes available in quark. In Fig. 2 an example of a Jupyter Notebook for the usage of quark is shown correspondingly, to demonstrate the easy accessibility for users with a bit of experience in programming with Python. It is based on a quadratic formulation of the maximum colorable subgraph problem.

In general, the start is to implement an `Instance`, the container for the problem defining parameters, as shown in cell 2 of Fig. 2. From the instance, the `ConstrainedObjective` is constructed, which is a factory processing the instance data to obtain the objective function and a collection of constraints, cf. cell 3. The latter can then be automatically transformed into the corresponding penalty objective terms, which are contained in the `ObjectiveTerms` together with the actual objective function. The weighted sum of the objective terms forms the `Objective`, the final Ising/QUBO problem. The above steps are all executed in cell 5, starting with the instantiation of a concrete instance with the parameters defined in cell 4.

All of the mentioned objective objects as well as the constraints contain `Polynomials` representing the functions. The special polynomials `PolyBinary` and `PolyIsing` take advantage of the restriction to either binary or spin variables and add specific preprocessing or conversion methods. As apparent in the figure, also multi-index variables can be processed.

The `ScipModel`, used in cell 6, is an interface to the classical MILP solver SCIP [Be21] based on the Python package PySCIPOpt [Ma16], which can solve a `ConstrainedObjective` or a (small enough) `Objective` for comparison. The result of an optimization is stored in the `Solution`. This includes not only the found variable assignment but also further information, like the runtime, which are obtained during the solving process. In cell 8, we then show

```
[1]: from quark import PolyBinary, ConstraintBinary, ConstrainedObjective, ScipModel

[2]: class MCSInstance():

    def __init__(self, edges, colors):
        self.edges = edges
        self.nodes = set(node for edge in self.edges for node in edge)
        self.colors = colors

[3]: class MCSConstrainedObjective(ConstrainedObjective):

    @staticmethod
    def _get_objective_poly(instance):
        # sum_[c in Colors] sum_[(n, m) in Edges] (1 * x_n_c * x_m_c)
        return PolyBinary({('X', node1, color), ('X', node2, color)}: 1
                           for node1, node2 in instance.edges
                           for color in instance.colors})

    @staticmethod
    def _get_constraints(instance):
        constraints = {}
        # for all n in Nodes: sum_[c in Colors] x_n_c == 1
        for node in instance.nodes:
            poly = PolyBinary({('X', node, color)}: 1 for color in instance.colors)
            constraints[f'one_color_for_{node}'] = ConstraintBinary(poly, 1, 1)
        return constraints

[4]: edges = [('a', 'b'), ('b', 'c'), ('b', 'd'), ('c', 'd')]
colors = ['red', 'blue', 'green']

[5]: instance = MCSInstance(edges, colors)
constrained_objective = MCSConstrainedObjective(instance=instance)
objective_terms = constrained_objective.get_objective_terms()
terms_weights = objective_terms.get_default_terms_weights()
objective = objective_terms.get_objective(terms_weights)
print(objective.polynomial)

+4 -1 X_a_blue -1 X_a_green -1 X_a_red -1 X_b_blue -1 X_b_green -1 X_b_red
-1 X_c_blue -1 X_c_green -1 X_c_red -1 X_d_blue -1 X_d_green -1 X_d_red
+2 X_a_blue X_a_green +2 X_a_blue X_a_red +1 X_a_blue X_b_blue +...

[6]: model = ScipModel.get_from_objective(objective)
solution = model.solve()
print(solution)

X_a_blue = 1   X_a_green = 0   X_a_red = 0
X_b_blue = 0   X_b_green = 1   X_b_red = 0
X_c_blue = 1   X_c_green = 0   X_c_red = 0
X_d_blue = 0   X_d_green = 0   X_d_red = 1

[7]: from dwave.samplers import SimulatedAnnealingSampler

[8]: sampler = SimulatedAnnealingSampler()
objective_ising = objective.to_ising()
sample = sampler.sample_ising(objective_ising.polynomial.linear_plain,
                             objective_ising.polynomial.quadratic,
                             num_reads=10)

sample.first.sample

[8]: {('X', 'a', 'blue'): -1, ('X', 'a', 'green'): -1, ('X', 'a', 'red'): 1,
      ('X', 'b', 'blue'): 1, ('X', 'b', 'green'): -1, ('X', 'b', 'red'): -1,
      ('X', 'c', 'blue'): -1, ('X', 'c', 'green'): 1, ('X', 'c', 'red'): -1,
      ('X', 'd', 'blue'): -1, ('X', 'd', 'green'): -1, ('X', 'd', 'red'): 1}
```

Fig. 2: Jupyter Notebook exemplarily showing the usage of quark to obtain a corresponding Ising problem from a given maximum colorable subgraph problem instance (where only the output was slightly adjusted for better readability)

how to solve the constructed objective with a D-Wave sampler, which works analogously to the access of an actual machine but is freely available over the D-Wave API [D-23b].

The package `quark` contains more classes which take further specific hardware restrictions into account. They form an anchor to the other packages `complete_graph_embedding` and `weight_distribution` of our library, supporting embedded Ising problem formulations [LK23]. In the package `quapps`, we have assembled a library of exemplary applications, such as prime factorization or the FGA problem, with individual parameterized implementations of the `ConstrainedObjective` or the `ObjectiveTerms` based on the `Instance` definition.

References

- [Ba82] Barahona, F.: On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General* 15/10, pp. 3241–3253, 1982, URL: <https://doi.org/10.1088/0305-4470/15/10/028>.
- [Be21] Bestuzheva, K.; Besançon, M.; Chen, W.-K.; Chmiela, A.; Donkiewicz, T.; van Doornmalen, J.; Eifler, L.; Gaul, O.; Gamrath, G.; Gleixner, A.; Gottwald, L.; Graczyk, C.; Halbig, K.; Hoen, A.; Hojny, C.; van der Hulst, R.; Koch, T.; Lübbecke, M.; Maher, S. J.; Matter, F.; Mühmer, E.; Müller, B.; Pfetsch, M. E.; Rehfeldt, D.; Schlein, S.; Schlösser, F.; Serrano, F.; Shinano, Y.; Sofranac, B.; Turner, M.; Vigerske, S.; Wegscheider, F.; Wellner, P.; Weninger, D.; Witzig, J.: The SCIP Optimization Suite 8.0, Technical Report, Optimization Online, Dec. 2021, URL: http://www.optimization-online.org/DB_HTML/2021/12/8728.html.
- [D-23a] D-Wave Systems Inc.: D-Wave System Documentation, accessed 29/08/2023, 2023, URL: <https://docs.dwavesys.com/docs/latest/index.html>.
- [D-23b] D-Wave Systems Inc.: `dwave-ocean-sdk`, version 6.4.0, 2023, URL: <https://github.com/dwavesystems/dwave-ocean-sdk>.
- [Fa00] Farhi, E.; Goldstone, J.; Gutmann, S.; Sipser, M.: Quantum computation by adiabatic evolution, 2000, arXiv: quant-ph/0001106.
- [Ge23] German Aerospace Center (DLR) – Institute for Software Technology: `quark`, version 1.0, 2023, URL: <https://gitlab.com/quantum-computing-software/quark>.
- [Gl22] Glover, F.; Kochenberger, G.; Hennig, R.; Du, Y.: Quantum bridge analytics I: a tutorial on formulating and using QUBO models. *Annals of Operations Research* 314/1, pp. 141–183, 2022, URL: <https://doi.org/10.1007/s10479-022-04634-2>.
- [LK23] Lobe, E.; Kaibel, V.: Optimal sufficient requirements on the embedded Ising problem in polynomial time. *Quantum Information Processing* 22/305, pp. 1–44, 2023, URL: <https://doi.org/10.1007/s11128-023-04058-2>.

- [Lu14] Lucas, A.: Ising formulations of many NP problems. *Frontiers in Physics* 2/5, 2014, URL: <https://doi.org/10.3389/fphy.2014.00005>.
- [Ma16] Maher, S.; Miltenberger, M.; Pedroso, J. P.; Rehfeldt, D.; Schwarz, R.; Serrano, F.: PySCIPOpt: Mathematical Programming in Python with the SCIP Optimization Suite. In: *Mathematical Software – ICMS 2016*. Springer International Publishing, pp. 301–307, 2016, URL: https://doi.org/10.1007/978-3-319-42432-3_37.
- [Mc20] McGeoch, C. C.: Theory versus practice in annealing-based quantum computing. *Theoretical Computer Science* 816/, pp. 169–183, 2020, URL: <https://doi.org/10.1016/j.tcs.2020.01.024>.
- [SLJ19] Stollenwerk, T.; Lobe, E.; Jung, M.: Flight Gate Assignment with a Quantum Annealer. *Lecture Notes in Computer Science* 11413/, pp. 99–110, 2019, URL: https://doi.org/10.1007/978-3-030-14082-3_9.

Constrained Grover Adaptive Search for Optimization of the Bidirectional EV Charging Problem¹

Marika Federer², Steve Lenk³, Daniel Müssig⁴, Markus Wappler⁵, Jörg Lässig⁶

Abstract: The optimization problem of bidirectional electrical vehicle charging (Vehicle-to-Home) becomes more and more important with rising energy prices and the required reduction of CO_2 emissions. We maximize the usage of local solar power generation, while minimizing the power grid usage. This is constrained by the energy demand of the household and the required state of charge at departure as well as the idle times of the car at home. The problem is formulated as a Constrained Polynomial Binary Optimization (CPBO) problem, which is convenient for Grover Adaptive Search by representing the objective function and the constraints as a Quantum Dictionary.

Keywords: Quantum Computing; Bidirectional EV Charging; Grover Adaptive Search; Constrained Optimization; Quantum Dictionary

1 Introduction

Quantum Computing [NC10] promises an advantage over conventional computing in solving many NP-hard problems faster [Pi23], without changing their complexity class. The current generation of Quantum Computers is small in number of qubits and not fault-tolerant, thus called *Noisy Intermediate Scale Quantum computers* (NISQ). To use current generation Quantum Computers, programs are usually written on low abstraction level by using circuits composed from quantum gates.

A very promising application of Quantum Computing are optimization problems. With current hardware limitations only small problem instances can be solved practically, e.g. combinatorial problems [LKP21, Ka17]. One such problem is the optimization problem of bidirectional electrical vehicle charging. In a previous work, we have optimized the charging schedule of battery electric service vehicles of the Erfurt airport on a quantum computer [Fe22a]. We have modeled the problem as a *Quadratic Unconstrained Binary*

¹ Authors contributed equally

² Fraunhofer IOSB-AST, Dep. Cognitive Energy Systems, Görlitz, Germany marika.federer@iosb-ast.fraunhofer.de

³ Fraunhofer IOSB-AST, Dep. Cognitive Energy Systems, Görlitz, Germany steve.lenk@iosb-ast.fraunhofer.de

⁴ Fraunhofer IOSB, branch Advanced System Technologies IOSB-AST, Cognitive Energy Systems, Wilhelmsplatz 11, 02826 Görlitz, Germany daniel.muessig@iosb-ast.fraunhofer.de

⁵ Fraunhofer IOSB, branch Advanced System Technologies IOSB-AST, Cognitive Energy Systems, Wilhelmsplatz 11, 02826 Görlitz, Germany markus.wappler@iosb-ast.fraunhofer.de

⁶ University of Applied Sciences Zittau/Görlitz, Brückenstraße 1, 02826 Görlitz, Germany jlaessig@hszg.de

Optimization (QUBO) problem and solved it using *Quantum Approximate Optimization Algorithm* (QAOA) [Ha19]. However, in this industrial use case we have not considered a bidirectional setting. In a follow-up paper we have analyzed the results and investigated the complexity as well as the landscape of the objective function [Fe22b]. We have found that the introduction of constraints increases the complexity of the landscape. As an extension to the first paper, we have implemented the optimization problem of bidirectional EV charging with QAOA [Mü23]. We were able to find the optimal solution for all of our problems. However, problems requiring four qubits had only a probability of 13% (out of 1024 shots) finding the correct solution and problems with seven qubits had only a probability of 2% (out of 1024 shots).

In this paper we aim to solve a very specific problem for bidirectional EV charging. First, we consider a vehicle to home use case, where the car discharges only into the home and not into the grid. Further, we want to maximize the local usage of the solar power generation. Therefore, we use the battery of the electrical vehicle also as a battery for the house. We always want to fulfill the minimum charging energy e_{min} (here and in the following all variables are considered to be integral) and not exceed the maximum charging energy e_{max} at the end of the charging duration, i.e. plug-off time. For each timestep t we have a maximum charging power $j_{max,t}$ and a maximum discharging power $j_{min,t}$. Both, e_{min} and $j_{min,t}$ are defined as positive values. Further, the difference of the solar power generation p_{v_t} and the power demand of the house d_t is continued as $p_t = p_{v_t} - d_t$, which should be always positive in this setting. We minimize the cost function over the charging power j_t^c and discharging power j_t^d with their difference $j_t = j_t^c - j_t^d$, which is either positive or negative. To solve this problem, we use Eq. 1, which is a *Quadratic Constrained Integer Optimization* (QCIO) problem.

$$\begin{aligned}
 C(j_t) = \min \sum_t (j_t - p_t)^2 \\
 \text{s.t. } -j_{min,t} \leq j_t \leq j_{max,t} \\
 e_{min} \leq \sum_t j_t \leq e_{max}
 \end{aligned} \tag{1}$$

We have the following constraints from top to bottom: we are not allowed to discharge more than our maximum discharging power and we are not allowed to charge more than our maximum charging power, additionally, we must fulfill the minimum charging energy and shall not exceed the maximum charging energy.

We decided to formulate the problem as a QCIO problem, because it can be translated into a *Constrained Polynomial Binary Optimization* (CPBO) problem, which can be solved with *Grover Adaptive Search* (GAS).

In this paper, we will first model the Bidirectional EV charging problem as a CPBO problem.

Afterwards, the model will be fed into the described GAS machinery. The paper is organized as follows: Sec. 2 gives an overview of Grover Adaptive Search. In Sec. 3 we reformulate the problem and give implementation details for the Quantum Dictionaries and Grover Adaptive Search. In Sec. 4 we discuss current limitations of the approach and present future research directions.

2 Tools

While extensively studying QAOA-based attempts to tackle problem (1), the authors also considered an alternative approach proposed in [GWG21]. This work aims to encode CPBO problems with *Quantum Dictionaries* (introduced in [Gi21]) to be solved by *Grover Adaptive Search*. We will introduce the main ideas of both techniques, before presenting implementation details in the subsequent section.

Article [Gi21] introduces the concept of Quantum Dictionaries, a specific method to define key-value-pairs by entangling basis states of a key register with basis states of a value register. As shown in the same paper, Quantum Dictionaries can be used to represent a polynomial by encoding its coefficients. Figure 1 shows the implementation of $f(x) = 4x_0 + x_0x_1$.

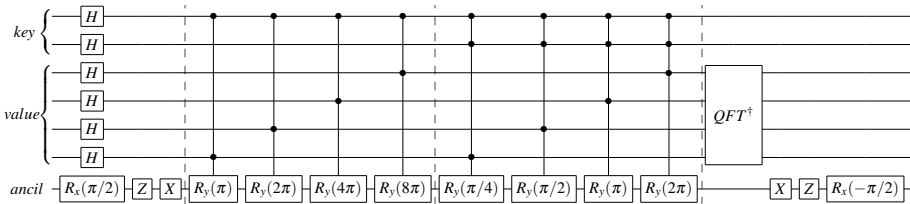


Fig. 1: Quantum Dictionary implementing circuit for function table of $f(x) = 4x_0 + x_0x_1$.

After execution, the key register contains all integral arguments in equal superposition (domain given by width of key register). These are entangled with respective images in value register. Negative function values are represented by Two’s complement, setting the leading qubit to one. Note, that Quantum Dictionary circuits make use of the inverse Quantum Fourier Transform (QFT^\dagger), thus having a certain depth.

For optimizing polynomials, the *Quantum Amplitude Amplification* algorithm can be used [Br02], generalizing Grover’s algorithm [Gr96]. Core of both algorithms is a phase oracle U_c , encoding a search condition $c : \{0, 1\}^n \rightarrow \{0, 1\}$ by

$$U_c : |x\rangle \mapsto \begin{cases} -|x\rangle & \text{for } c(x) = 1, \\ |x\rangle & \text{for } c(x) = 0. \end{cases} \tag{2}$$

Together with the Diffusion Operator $D = 2|0\rangle\langle 0| - I$ and an arbitrary unitary operator A , an iteration of Quantum Amplitude Amplification is given by

$$Q_c = ADA^\dagger U_c. \tag{3}$$

The algorithm applies an appropriate number of iterations k to the initial state $A|0\rangle$, hence Quantum Amplitude Amplification reads

$$|\psi\rangle = Q_c^k A|0\rangle. \tag{4}$$

As the name suggests, $|\psi\rangle$ will contain exactly the solution states $|x\rangle$, marked by $c(x) = 1$, with high amplitudes. Thus, measurement will yield solutions with higher probabilities compared to non-solutions.

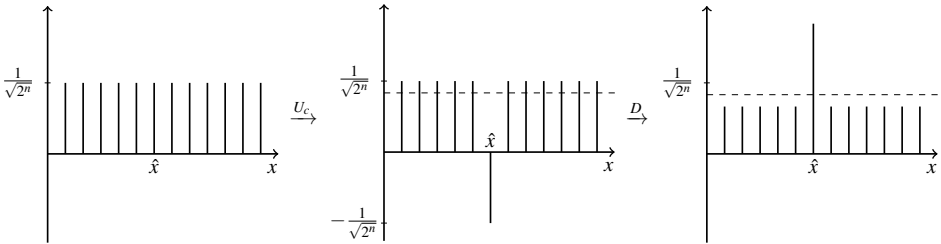


Fig. 2: Illustration of one iteration of Quantum Amplitude Amplification with $A = H^{\otimes n}$, thus yielding the special case of Grover Search. Computational basis state \hat{x} is the unique solution with amplitude first being negated by oracle U_c before being amplified by diffuser D .

Setting operator A of (3), (4) to a Quantum Dictionary representing polynomial $f(x)$, and setting oracle U_c to differentiate negative and non-negative function values, Quantum Amplitude Amplification can be utilized to find all arguments x with $f(x) < 0$. Moreover, with appropriate constant term y , only minimal values of f fulfill $f(x) + y < 0$. Hence, variation of y yields a solution process for a optimization problem $\min f(x)$. GAS is executing this variational principle by using Quantum Amplitude Amplification to find solutions of $f(x) + y < 0$ while performing stochastic search over parameter y .

In a similar way (polynomial) constraints can be treated, by adding a Quantum Dictionary for each constraint respectively. This enables GAS to solve CPBO problems

$$\begin{aligned} &\min f(x), \\ &\text{s.t. } g_i(x) < 0, \quad 1 \leq i \leq m, \\ &\quad h_j(x) = 0, \quad 1 \leq j \leq p, \\ &\quad x \in \{0, 1\}^n, \end{aligned} \tag{5}$$

with f, g_i, h_j being polynomials.

3 Formulation of the CPBO and Implementation of GAS

Grover Adaptive Search is able to find the optimal solution of a QUBO problem or even a CPBO problem utilizing Quantum Dictionaries to represent the equation as shown in the

previous section. In this paper we want to utilize CPBO, since we have seen in previous works, that adding constraints as penalty factors to a QUBO reformulation makes the problem harder to solve.

For the proposed algorithm we have to reformulate the QCIO into a CPBO. The result is shown in Equation 6.

$$\begin{aligned}
 C(j_t) = \min \sum_t [j_t^2 - 2j_t p_t + p_t^2] \\
 \text{s.t. } 0 > -j_{min,t} - j_t - 1 \\
 0 > j_t - j_{max,t} - 1 \\
 0 > e_{min} - \sum_t j_t - 1 \\
 0 > \sum_t j_t - e_{max} - 1
 \end{aligned} \tag{6}$$

For an easier translation into a Quantum Dictionary we have used the binomial formula in the objective function C . In the next step we have reformulated all inequality constraints as non-negativity conditions, e. g. $0 > x$. This will help us later to save $CNOT$ -gates and results in easier to implement oracles utilizing the Two's complement. With this formulation we can start implementing the problem with Grover Adaptive Search (GAS). However, we also need to convert the equations into binary variables.

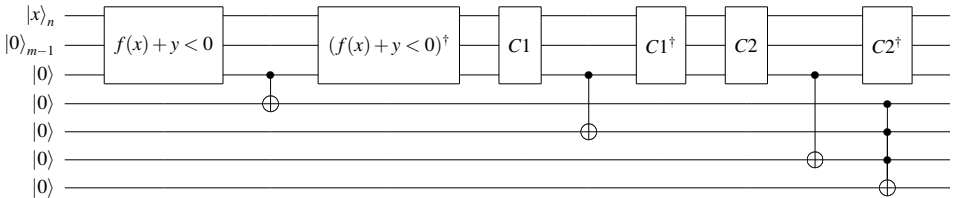


Fig. 3: Exemplary circuit for Grover Adaptive Search. In a first step we calculate the objective function. If the value is less than 0 the first $|0\rangle$ is flipped (in this diagram we use the least significant bit first notation). This result is 'copied' via a $CNOT$ -gate to the next qubit. Afterwards, we need to uncompute the objective function. Then we can implement each constraint on their own. Here, we show it only for two constraints. The sequence is the same as for the objective function. However, we always need a new qubit to 'copy' the bitflip. The last step is a multiple controlled NOT -gate. Thereby, if the objective function with its threshold and all constraints are met the last qubit is in state $|1\rangle$ and we need to search only for this $|1\rangle$.

The core part of the proposed algorithm combines Grover Adaptive Search with Quantum Dictionaries. They are used to implement the equations of QUBOs to produce an equal superposition over all possible inputs with their corresponding objective function values as shown in Fig. 1. In this way we implement the objective function and each constraint with

individual Quantum Dictionaries. Using a threshold value in the objective function we can make this value negative for the optimal values. This results in a circuit similar to the one shown in Fig. 3.

4 Discussion and Future Work

In the last chapter we have shown, how one could implement the problem of finding an optimal charging and discharging schedule for bidirectional EV charging with Grover Adaptive Search. However, using Quantum Dictionaries is not convenient for NISQ-devices. Since the inverse QFT is required for their implementation, it is very costly in overall gate depth. Further, with increasing number of key-value pairs the probability for each individual pair is decreasing. Therefore, we require a high accuracy to measure each individual pair. In Sec. 3 we proposed to use a Quantum Dictionary for all constraints, however for a specific problem it could be more suitable to implement it in a different way, e.g. using Quantum Arithmetic.

Moreover, the application of the Quantum Dictionaries in GAS (operator A) becomes even more costly, since A^\dagger and A are part of the Diffusion operator in each iteration. Thus, our resulting circuits get very deep. Because of the noisiness of current generation Quantum Computers it is very important to do further research on an efficient implementation for the specific use case.

Our intended next steps are: Implementation and tests on a real Quantum Computer, comparison with other algorithms such as QAOA and other hardware such as Quantum Annealers. Overall, a further research direction is the evaluation of the overhead, which is created by implementing constraints in quantum optimization algorithms.

Acknowledgements

Funded by the European Union under Horizon Europe Programme - Grant Agreement 101080086 — NeQST.

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

This work was financially supported by the Bundesministerium für Wirtschaft und Energie through the project “EnerQuant” (Project- ID 03EI1025C) and by Thüringer Aufbaubank through the project “Quantum Hub Thüringen”.

Bibliography

- [Br02] Brassard, Gilles; Hoyer, Peter; Mosca, Michele; Tapp, Alain: Quantum Amplitude Amplification and Estimation. arXiv:quant-ph/0005055, 305:53–74, 2002.
- [Fe22a] Federer, Marika; Müssig, Daniel; Klaiber, Stefan; Lässig, Jörg; Bretschneider, Peter; Lenk, Steve: Application Benchmark for Quantum Optimization on Electromobility Use Case. In: 2022 IEEE Vehicle Power and Propulsion Conference (VPPC). pp. 1–6, 2022.
- [Fe22b] Federer, Marika; Müssig, Daniel; Klaiber, Stefan; Lenk, Steve: Application-Oriented Quantum Computing Benchmark for an Electromobility Use Case. In: 2022 IEEE International Conference on Quantum Computing and Engineering (QCE). pp. 749–752, 2022.
- [Gi21] Gilliam, Austin; Venci, Charlene; Muralidharan, Sreraman; Dorum, Vitaliy; May, Eric; Narasimhan, Rajesh; Gonciulea, Constantin: Foundational Patterns for Efficient Quantum Computing. arXiv:1907.11513, 2021.
- [Gr96] Grover, Lov K.: A Fast Quantum Mechanical Algorithm for Database Search. arXiv:quant-ph/9605043, 1996.
- [GWG21] Gilliam, Austin; Woerner, Stefan; Gonciulea, Constantin: Grover Adaptive Search for Constrained Polynomial Binary Optimization. *Quantum*, 5:428, 2021.
- [Ha19] Hadfield, Stuart; Wang, Zhihui; O’Gorman, Bryan; Rieffel, Eleanor G.; Venturelli, Davide; Biswas, Rupak: From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz. *Algorithms*, 12(2):34, 2019.
- [Ka17] Kandala, Abhinav; Mezzacapo, Antonio; Temme, Kristan; Takita, Maika; Brink, Markus; Chow, Jerry M.; Gambetta, Jay M.: Hardware-Efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets. *Nature*, 549(7671):242–246, 2017.
- [LKP21] Luckow, Andre; Klepsch, Johannes; Pichlmeier, Josef: Quantum Computing: Towards Industry Reference Problems. *Digitale Welt*, 5(2):38–45, 2021.
- [Mü23] Müssig, Daniel; Federer, Marika; Lenk, Steve; Lässig, Jörg: Solving the Optimization Problem of Bidirectional EV-Charging on a Quantum Computer. *Harzer Hochschultexte, NWK 2023 – Tagungsband der 23. Nachwuchswissenschaftler*innenkonferenz*:94–101, 2023.
- [NC10] Nielsen, Michael A.; Chuang, Isaac L.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge ; New York, 10th anniversary ed edition, 2010.
- [Pi23] Pirnay, Niklas; Ulitzsch, Vincent; Wilde, Frederik; Eisert, Jens; Seifert, Jean-Pierre: A Super-Polynomial Quantum Advantage for Combinatorial Optimization Problems. arXiv:2212.08678, 2023.

Providing Quantum Readiness: The Vision of the ProvideQ Toolbox

Domenik Eichhorn¹, Maximilian Schweikart¹, Nick Poser¹, Tobias Osborne²,
and Ina Schaefer¹

Abstract: Quantum computing has the potential to exponentially accelerate the solution of specific problems compared to classical computing. However, the accessibility to quantum computing is currently limited due to the technical challenges posed by quantum devices. Additionally, implementing quantum algorithms is challenging because of the complex nature of quantum systems. To address these challenges, we introduce the ProvideQ Toolbox, a framework designed to enhance the accessibility of quantum computing, especially for optimization problems. Our toolbox includes a range of classical and quantum state-of-the-art optimization algorithms and employs meta-solver strategies to determine the best quantum optimization algorithm or quantum subroutine in a classical algorithm for a given optimization problem. The ProvideQ Toolbox can be used via a web-based frontend and an API and can be seamlessly integrated with multiple quantum computing backends and classical optimization frameworks.

Keywords: quantum software engineering, quantum algorithms, meta-solver strategy

1 Introduction

Quantum computing is a major topic of research since the first ideas for quantum computers were proposed in the 1980s [Be80; Fe85]. Even though no quantum hardware was available during that time, researchers were able to show that quantum computing has the potential to solve certain algorithmic problems exponentially faster than classical computing [DJ92; Sh99]. Recent advantages in quantum hardware started the so-called noisy intermediate-scale quantum (NISQ) era, which led to prototypes of noisy quantum computers and raised hope that practical quantum advantages can be achieved within the following years. However, quantum computing is still in its early stages. Quantum devices available today provide up to a few hundred logical qubits [CN23], but due to technical challenges, high maintenance costs, and high demand, their accessibility is very limited.

Quantum software development is supported by specialized programming languages, libraries, and frameworks [He20]. However, quantum software engineering generally lacks tool support and established processes [De22]. Classical software developers are used to

¹ Karlsruhe Institute of Technology, Institute of Information Security and Dependability (KASTEL), Am Fasanengarten 5, 76131 Karlsruhe, Germany, provideq@lists.kit.edu

² Leibniz University Hannover, Institute of Theoretical Physics, Appelstraße 2, 30167 Hannover, Germany, tobias.osborne@itp.uni-hannover.de

software engineering that provides many high-level development tools and abstraction layers, well-established processes, and other state-of-the-art practices that greatly simplify modern software development. Features like this do not yet exist in the quantum field. Implementing quantum software demands significant effort. Code optimization for underlying quantum devices is essential, and incorporating complex error correction techniques becomes necessary to mitigate the effects of decoherence and noise. Furthermore, quantum computing and quantum algorithms operate fundamentally differently from their classical counterparts. Quantum computing has the potential to highly impact aspects of the world by providing large-scale computational speedups and allowing the calculation of solutions that would not be possible with classical computers. However, its high entry barrier makes it inaccessible for most people and companies, which highly reduces the potential impact of this technology.

To alleviate this problem, we aim to improve the accessibility of quantum computing with the ProvideQ toolbox: a framework that implements a range of classical and quantum state-of-the-art optimization algorithms and acts intermediate layer that can be used to speed up combinatorial optimization problems with quantum subroutines. Furthermore, we employ so-called meta-solver strategies, which determine the best solution approach to an optimization problem by analyzing problem instances based on pre-defined heuristics.

2 Making Quantum Computing Accessible

There are many recent initiatives to improve quantum software engineering, for instance: the development of quantum SDKs such as Qiskit [Qi23], Cirq [Ci22], Forest [SCZ16], or Strawberry Fields [Ki19], the development of specialized libraries and frameworks such as PennyLane [Be18] for quantum machine learning and chemistry, QuTiP [JNN12] for quantum simulation, or Qiskit-optimization [Qi23] for solving optimization problems. There are initiatives like PlanQK³ that focus on the deployment of quantum applications and providing knowledge exchange platforms, and some companies already work on the commercialization of hybrid solutions for optimization problems.

In the ProvideQ project we want to build on these existing initiatives and design a framework that acts as an abstraction layer when using hybrid approaches to solve optimization problems. We focus on the implementation of so-called meta-solver strategies, which combine classical and quantum algorithms to provide highly efficient solutions to optimization problems. A meta-solver strategy predicts the best solution path for a given problem instance by making assumptions about its characteristics and then selecting optimal solvers based on empirical data and expert knowledge. The proposed solution paths consist of several classical and quantum subroutines, which the ProvideQ toolbox will orchestrate and execute on suitable backends (classical computation clusters or quantum hardware). The main challenge for the ProvideQ project is identifying potentials for quantum speedups, which is why another

³ <https://planqk.de/> Accessed: 29.08.2023

main effort of the ProvideQ toolbox is the analysis of classical state-of-the-art algorithms and gathering knowledge about the practical performance of quantum algorithms.

As practical quantum advantages will most likely only be achieved once post-NISQ quantum hardware is available, the benefits of the ProvideQ toolbox are currently mainly from a theoretical perspective. Thus, when developing the ProvideQ toolbox, we have to focus on designing a future-proof framework that can be maintained over a long period. We ensure this by addressing the following requirements:

1. **Easy to Use:** The toolbox should be easy to understand. It has to be well-documented and easy to integrate into other tools. Inputs in the form of standardized file formats have to be supported.
2. **Extendable:** It should be easy for third-party developers to extend the ProvideQ toolbox with new input formats and quantum and classical subroutines.
3. **Exchangeable Quantum Backends:** Quantum backends constantly evolve while new quantum hardware is built. The backends on which the ProvideQ toolbox can deploy quantum code should be exchangeable.
4. **Relevance:** Algorithms and problems implemented by the ProvideQ toolbox should focus on relevant optimization problems and quantum algorithms.
5. **Open Source:** We want to establish a development community around the toolbox and provide all code (including frontend, backend, classical and quantum algorithms, and meta-solver strategies) open-source on GitHub using the MIT license.

3 The ProvideQ Toolbox

The ProvideQ toolbox is currently under development, and a first prototype is already released. The code of the project is available on GitHub⁴, and the prototype can be accessed via a web interface⁵ and an API⁶. The toolbox consists of two main components: the server and the web client. Figure 1 shows an overview of the architecture.

The *toolbox server* component is the heart of our framework. It implements problem definitions, meta-solver strategies, and concrete solvers. Problem definitions are representations of different optimization problems (such as MaxCut or Knapsack). They define a standardized input and output format for every algorithmic problem supported by the ProvideQ toolbox. For every problem definition, we implement one meta-solver strategy and a set of concrete solvers. Concrete solvers are implementations of classical or quantum algorithms. They can be either monolithic, standalone algorithms, or polyolithic algorithms

⁴ <https://github.com/ProvideQ> Accessed: 29.08.2023

⁵ <https://provideq.kit.edu/> Accessed: 29.08.2023

⁶ <https://api.provideq.kit.edu/> Accessed: 29.08.2023

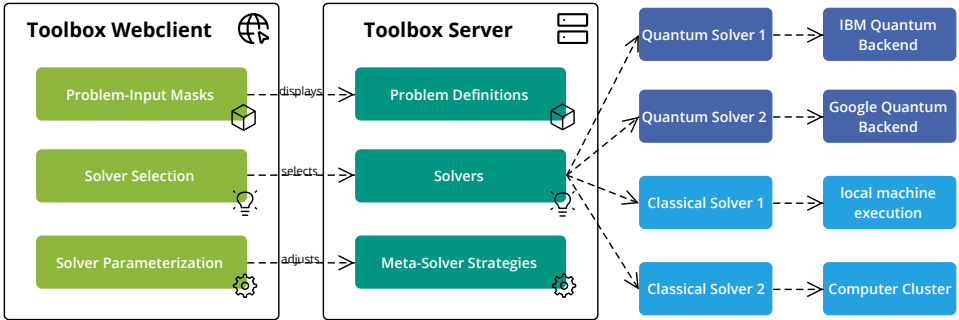


Fig. 1: Overview of the ProvideQ toolbox software architecture.

with multiple exchangeable subroutines. Meta-solver strategies are implemented as decision trees that perform multiple analyses on a given problem instance, make assumptions about its characteristics, and then iterate through the decision tree to decide which solvers should be used to solve a problem optimally. The toolbox-server component also implements multiple interfaces that (1) allow external programs to access the problem definitions, solvers, and meta-solver strategies over an API and (2) implement communication channels with classical computing networks and quantum hardware to execute the concrete solvers.

The *web client* component is one example of an external program using the API of the toolbox server. It is a web-based graphical user interface that visualizes all components of the server. A user can click on a problem definition to open a site where he can insert an instance of the associated algorithmic problem. He can then select a specific solver to calculate a solution or ask the ProvideQ toolbox to select the solver for him based on a meta-solver strategy. He can also help the ProvideQ toolbox to select an optimal solver by providing additional assumptions about the given problem instance through extra text fields, sliders, and checkboxes.

4 Conclusion and Future Work

Quantum computing has a high potential to speed up the solving of optimization problems drastically, but its true potential can only be realized when it is accessible to a broad audience. We presented the vision of the ProvideQ toolbox, a framework designed to enhance the accessibility of quantum computing by providing an abstraction layer between optimization problems and quantum computing hardware. Implementing effective meta-solver strategies and quantum subroutines within ProvideQ necessitates extensive research efforts. ProvideQ is an ongoing project, and our future work will revolve around the constant improvement of the ProvideQ toolbox. We aim to build a community around our framework and expand its capabilities by incorporating additional meta-solver strategies, hybrid algorithms, and interfaces to more quantum backends.

Acknowledgements

This work has been supported by the German Federal Ministry for Economics and Climate Action in project ProvideQ (reference number: 01MQ22006F).

References

- [Be18] Bergholm, V. et al.: PennyLane: Automatic differentiation of hybrid quantum-classical computations. arXiv preprint arXiv:1811.04968/, 2018.
- [Be80] Benioff, P.: The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of statistical physics* 22/, pp. 563–591, 1980.
- [Ci22] Cirq Developers: Cirq, version v1.1.0, 2022, URL: <https://doi.org/10.5281/zenodo.7465577>, visited on: 08/29/2023.
- [CN23] Collins, H.; Nay, C.: IBM Unveils 400 Qubit-Plus Quantum Processor and Next-Generation IBM Quantum System Two, 2023, URL: <https://newsroom.ibm.com/2022-11-09-IBM-Unveils-400-Qubit-Plus-Quantum-Processor-and-Next-Generation-IBM-Quantum-System-Two>, visited on: 09/06/2023.
- [De22] De Stefano, M. et al.: Software engineering for quantum programming: How far are we? *Journal of Systems and Software* 190/, p. 111326, 2022.
- [DJ92] Deutsch, D.; Jozsa, R.: Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 439/1907, pp. 553–558, 1992.
- [Fe85] Feynman, R. P.: Quantum mechanical computers. *Optics news* 11/2, pp. 11–20, 1985.
- [He20] Heim, B. et al.: Quantum programming languages. *Nature Reviews Physics* 2/12, pp. 709–722, 2020.
- [JNN12] Johansson, J. R.; Nation, P. D.; Nori, F.: QuTiP: An open-source Python framework for the dynamics of open quantum systems. *Computer Physics Communications* 183/8, pp. 1760–1772, 2012.
- [Ki19] Killoran, N. et al.: Strawberry fields: A software platform for photonic quantum computing. *Quantum* 3/, p. 129, 2019.
- [Qi23] Qiskit contributors: Qiskit: An Open-source Framework for Quantum Computing, 2023.
- [SCZ16] Smith, R. S.; Curtis, M. J.; Zeng, W. J.: A practical quantum instruction set architecture. arXiv preprint arXiv:1608.03355/, 2016.
- [Sh99] Shor, P. W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review* 41/2, pp. 303–332, 1999.

Towards Performance Benchmarking for Quantum Reinforcement Learning

Volker Reers¹

Abstract: This research paper explores the potential benefits and challenges of using Quantum Computing to enhance Reinforcement Learning (RL), a subset of Machine Learning that involves learning through trial and error. RL algorithms aim to learn a policy that maximizes a reward signal by interacting with an environment. Quantum Computing, which uses quantum-mechanical phenomena to process information, has been applied to several models in the realm of Machine Learning (ML). It therefore has the potential to accelerate computations of ML models which may be used to improve the efficiency of RL. However, current hardware limitations of quantum computers and the difficulty of designing quantum algorithms that outperform classical algorithms are major challenges to overcome. This paper discusses different approaches to using Quantum Computing for RL, such as using quantum algorithms to solve specific subproblems. In particular, we aim towards a performance benchmark of Quantum Reinforcement Learning using an implementation of Quantum Deep Q Learning as an example. While there is still much research to be done, the potential benefits of using Quantum Computing for RL are significant and warrant further exploration as quantum computing technology continues to evolve.

Keywords: Quantum Reinforcement Learning, Quantum Deep Q Learning, Quantum Performance Benchmarking

1 Introduction

Machine Learning (ML) is a subfield of Artificial Intelligence (AI) that focuses on developing algorithms that enable machines to learn from data and make decisions or predictions based on the patterns they discover and without being explicitly programmed. Reinforcement Learning (RL), among Supervised and Unsupervised Learning one of the three main domains of ML, is concerned with how agents can learn to make optimal decisions in a given environment by trial and error. It has gained significant attention in recent years due to its applications in various domains such as robotics, game playing and autonomous systems. However, the current limitations of traditional computers have constrained the capacity of RL. This might be revolutionized by the emergence of Quantum Computing (QC). QC is a relatively new field of study that is concerned with the development of computer technology based on the principles of quantum mechanics. QC has been integrated with ML in several promising proof-of-concepts, leading to the research area of Quantum Machine Learning (QML). In this area, one question of interest is whether the combination of RL and QC has the potential to significantly improve the performance of RL algorithms and enable the development of new RL applications. As of

¹ Qseidon GmbH, Langer Weg 7, Gütersloh, 33332, volker.reers@qseidon.de

today, however, Quantum Reinforcement Learning (QRL) still lags behind Quantum Supervised and Unsupervised Learning [Je21]. One more reason to intensify research and to integrate QC and RL more successful.

In this paper, we will discuss the fundamentals of reinforcement learning, quantum computing, and the potential impact of utilizing quantum computing for RL. We will also discuss the current state of research in this field and identify possible metrics for benchmarking QRL models to evaluate a model's performance. We will demonstrate this with an example of a Quantum version of a Deep Q Network (QDQN).

2 Quantum Reinforcement Learning

2.1 Reinforcement Learning

Reinforcement Learning is a subfield of Machine Learning, where an agent learns to interact with an environment in a way that maximizes its reward. The agent learns through trial and error, exploring the environment, and learning from the consequences of its actions. In RL, the agent is not provided with labeled examples like it would be in Supervised Learning. Instead, it is given a reward signal for every action it takes, and the agent has to learn a sequence of actions that maximizes the cumulative reward over time.

RL has been successfully applied in various domains, from robotics to autonomous systems to game playing. In robotics, RL algorithms have been applied to develop controllers for robots that enable them to perform tasks such as grasping objects, walking, and running. In comparison, RL algorithms have been employed in the domain of autonomous systems to develop agents that can control autonomous vehicles, drones, and spacecraft. Furthermore, RL algorithms have been used in game playing to develop agents that can play games such as chess, Go, and Atari games at a superhuman level [Sa21, Sk22, Ne23].

2.2 Quantum Computing

Quantum computing (QC) is a new paradigm of computing which uses quantum bits or qubits instead of classical bits to represent information. Qubits can be set to superposition, allowing them to exist in multiple states simultaneously. This allows quantum computers to perform certain computations polynomially to exponentially faster than classical computers [NC10]. This makes quantum computing an attractive technology for solving complex problems that are currently intractable with classical computers, including many problems from ML domains.

Quantum computing relies on quantum mechanics, which is the study of the behavior of particles at the quantum level. Quantum mechanics introduces several strange phenomena.

Besides superposition, the most important ones are entanglement and quantum tunneling. These enable quantum computers to perform certain calculations that are not possible with classical computers.

2.3 Quantum Reinforcement Learning

The use of quantum computing in machine learning, including RL, has attracted much attention in recent years, leading to the combination of both to the research field of QRL. The potential impact of quantum computing on RL can be twofold. First, quantum computing can be used to speed up classical RL algorithms, improving their performance and scalability. In this regard, one potential application of QC in RL is in the optimization of the RL process itself. RL algorithms typically involve the use of complex optimization techniques to find the optimal policy. QC can be used to perform these optimization tasks much faster than classical computers, which can lead to significant improvements in the performance of RL algorithms. Second, quantum RL algorithms can be developed that take advantage of the unique properties of quantum mechanics to solve RL problems that are not tractable with classical computers [Me22, Ne23].

Another potential application of QC in RL is in the simulation of complex environments. Simulation plays a crucial role in RL, as it allows agents to learn and make decisions in virtual environments before deploying them in the real world. Simulations provide a controlled and cost-effective way to train RL agents, enabling them to explore a wide range of scenarios and learn from a vast number of virtual experiences [Bi17].

However, simulating complex environments can be computationally intensive and time-consuming, especially when dealing with large state and action spaces or intricate dynamics. This computational burden limits the scalability and efficiency of RL algorithms, hindering their ability to tackle real-world problems with high-dimensional and complex environments.

The research field of QRL can potentially accelerate the simulation of complex environments, allowing RL agents to learn and optimize their policies more efficiently. By harnessing quantum algorithms and techniques, RL agents can explore a much larger state and action space in the simulated environment, enabling them to learn optimal policies in a shorter amount of time. Quantum simulation can potentially offer substantial speedups compared to classical simulation methods, providing a significant advantage in training RL agents to tackle complex real-world problems [DB18].

3 Performance Metrics for Reinforcement Learning

The potential of QML, especially QRL, often leads researchers to concentrate primarily on algorithmic advancements. This includes the translation of classical algorithms into the quantum domain, the development of novel quantum algorithms, and the comparison of

different approaches. However, it is equally important to address the crucial aspect of performance evaluation, both in terms of overall performance and the ability to generalize across diverse problem domains. In this section, we present considerations on how performance metrics for reinforcement learning can be applied to QRL.

In the context of RL, it is highly desirable to have algorithms that can achieve high performance levels in real-world applications without relying heavily on expert knowledge or constant human intervention. Researchers would greatly benefit from such algorithms as they would spend less time fine-tuning them for standard tasks and instead focus on tackling more challenging problems. However, the current evaluation practices overlook the uncertainty in results and fail to acknowledge the difficulties associated with applying RL algorithms to specific problems. Consequently, existing RL algorithms are neither general purpose nor generally valid applicable to real-world scenarios [Jo20].

In order to establish an evaluation procedure with a specific objective, Jordan et al. present the following question as the guiding principle for their study: Which algorithms demonstrate strong performance across a diverse set of environments while requiring minimal or no environment-specific tuning? It is important to distinguish this question from the one typically addressed in articles introducing new algorithms, where the focus is on whether an algorithm under study can outperform other algorithms on specific tasks. Unlike the conventional question, the desired outcome of the general evaluation question is not to identify methods that maximize performance through optimal hyperparameter tuning, but rather to identify algorithms that exhibit robust performance without extensive hyperparameter adjustments, thereby facilitating their applicability to novel problem domains [Jo20].

Le Lan et al. approach this in a similar way. According to the authors, a fundamental principle for achieving generalization in RL involves ensuring that similar states receive similar predictions. One approach to address this principle is through state aggregation, which groups states together based on a notion of similarity. Another approach is to employ adaptive similarity measures, such as a nearest neighbor scheme using representative states. This adaptive approach is commonly used in algorithms designed for continuous state spaces, assuming the existence of a metric that quantifies the distance between states [LBC21].

Reinforcement learning algorithms, particularly Deep RL algorithms, are known to exhibit significant performance variability and sensitivity to various factors, such as implementation details, hyperparameters, choice of environments, and random seeds. This variability poses challenges for reproducibility in research and can have implications for real-world applications, potentially incurring costs or risks. Moreover, it hampers scientific progress as practitioners face difficulties in reliably evaluating or predicting algorithm performance, comparing different algorithms, or comparing different implementations of the same algorithm.

Chan et al. propose a set of metrics for evaluating the general performance and reliability of RL algorithms. Since these are not only qualitative evaluations, but can be computed

quantitatively, these metrics are well suited for application to the QRL example in our work. The metrics presented below are derived from Chan et al. and are suitable for our further considerations [Ch20].

Dispersion across Time (DT): IQR across Time: In order to assess the dispersion across time (DT), the aim is to focus on capturing higher-frequency variability rather than longer-term trends. It is important to be independent of positive training trends, which are considered desirable sources of variation in performance. To achieve this, a detrending technique prior to computing dispersion metrics can be applied. Specifically, differencing can be utilized, where from each data point (y_t) its following data point (y_{t-1}) will be subtracted to obtain the detrended series ($y_t' = y_t - y_{t-1}$). This detrended series can then be used to calculate the inter-quartile range (IQR) within a sliding window along the training curve, providing the final measure of dispersion.

Short-term Risk across Time (SRT): CVaR on Differences: To capture the most significant short-term risk over time, a measure that utilizes Conditional Value at Risk (CVaR) on the performance changes between consecutive evaluation points is introduced. First, the differences between time-points for each training run are calculated. These differences are then normalized by the distance between time-points, ensuring that the measure remains unaffected by variations in evaluation frequency. Subsequently, the distribution of these differences is built, and the α -quantile is identified. Finally, the expected value of the distribution below the α -quantile is calculated. This provides an estimation of the worst-case expected performance decline during training, specifically from one evaluation point to the next.

Long-term Risk across Time (LRT): CVaR on Drawdown: In order to capture the potential for significant performance losses relative to the peak, even over an extended period, a measure that utilizes CVaR on the Drawdown is introduced. The Drawdown at time T represents the decline in performance relative to the highest peak achieved thus far and is a concept borrowed from economics. Mathematically, the Drawdown at time T (Drawdown_T) is calculated as the difference between the current performance (R_T) and the maximum performance achieved until time T ($\max_{t \leq T} R_t$). Similar to the SRT metric, this measure is capable of capturing unusually large short-term performance drops. However, it also has the ability to capture unexpected large drops that occur over longer timeframes, providing a comprehensive assessment of performance volatility.

Dispersion across Runs (DR): IQR across Runs: Typically, the dispersion across runs involves computing the variance or standard deviation across training runs at specific evaluation points. In this approach, the practice is extended by recommending the application of a low-pass filter to the training data. This filtering step effectively removes high-frequency variability within individual runs, which is instead captured by the DT metric. Additionally, the use of robust statistics such as the IQR as an alternative to variance or standard deviation is advocated.

Risk across Runs (RR): CVaR across Runs: To quantify the Risk across Runs (RR), the concept of CVaR can be employed on the final performance of all the training runs. By

calculating the CVaR, an estimate of the expected performance for the worst-performing runs can be obtained. This measure allows to assess the potential risk associated with different training runs and provides insights to the overall performance distribution.

4 Experimental Implementation

With the rapid advancements in quantum computing and quantum machine learning research, numerous development libraries have emerged to aid researchers in their endeavors. TensorFlow Quantum, Qiskit, and PennyLane are among these libraries. PennyLane, specifically designed to facilitate quantum machine learning research, offers a user-friendly platform for evaluating the performance of quantum circuits using quantum simulators. Notably, the quantum simulator supported by PennyLane enables CPU emulation of QPU operations and provides extensive support for parameter usage in the form of PyTorch tensors and gradient operations [Pa19]. These capabilities make PennyLane particularly accessible to researchers familiar with machine learning using PyTorch, thereby easing the transition into quantum machine learning investigations. Building upon this background, our paper presents the implementation of a QRL model using PennyLane in conjunction with PyTorch.

4.1 Cartpole

In this research, we utilize OpenAI's Cartpole environment for our RL models. Cartpole is a game where an agent maneuvers a cart to prevent a pole balanced on top of it from falling. The agent receives rewards based on the duration it successfully maintains the pole's balance. The agent's decision-making relies on observing the position and velocity of the cart, as well as the angle and angular velocity of the pole. We use a 4-qubit Variational Quantum Circuit (VQC) shown in Figure 1, which is well-suited for policy determination in this specific environment. The four pieces of information obtained from the environment are normalized and put into the circuit as values ranging between $-\pi$ and π . By making use of the softmax function, these values are transformed into probability values representing the likelihood of taking various actions. This process iterates until the agent can make optimal decisions based on the provided state information [Kw21].

4.2 Experimental Setup

Variational quantum circuits offer a promising avenue for emulating the principles and functionalities of artificial neural networks, enabling the integration of quantum calculations into conventional machine learning algorithms. This advancement has emerged as a significant trend in quantum machine learning research, leading to numerous active studies. One notable application of VQC is its utilization as a Quantum Neural Network (QNN) in various domains, including reinforcement learning among others. In

our experimental configuration, we follow the proposal of Kwak et al. and leverage the VQC illustrated in Figure 1 as a replacement for the artificial neural network in our deep Q network [Kw21]. This results in our QDQN.

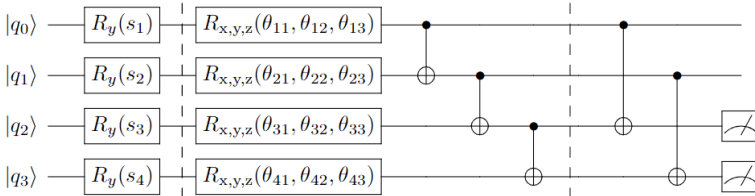


Figure 1: Policy VQC for Deep Q-Learning

The experimental setup employs PyTorch and PennyLane software packages to facilitate efficient tensor operations and quantum circuit simulation. Although PennyLane's quantum simulator offers convenient usage, its computational speed limitations restrict the number of qubits that can be effectively utilized. Hence, we opted for the CartPole environment, which provides a simple setting manageable with a small quantum circuit. For the quantum policy optimization, we employed the classical Adam optimizer with a learning rate of 0.001, while the classical critic utilized a learning rate of 0.00001. Additional hyperparameter settings included $\gamma = 0.98$, $\lambda = 0.95$, and $\varepsilon = 0.01$. As for the baseline model, it was designed as a random variation of the proposed model, where random parameters were assigned at each time step without any optimization process.

5 Results

Figure 2 presents the performance results of the experimental QDQN, compared with a classical version of a DQN (CDQN) with the same parameters, where applicable.

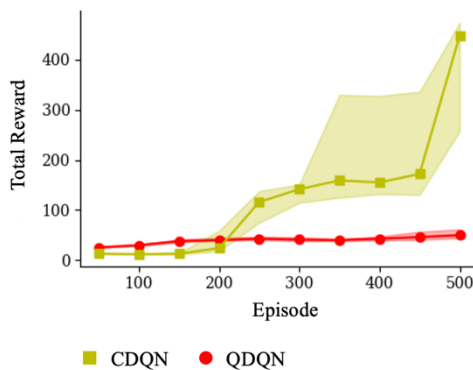


Figure 2: Comparison of total reward on average in the CartPole environment

In this proof-of-concept implementation, the CDQN achieved much better results. It is important to note that the focus was not on beating the classical variant in regards of performance. Therefore, no further tuning has been conducted on the QDQN. However, the significant difference in performance is not to be mistaken as a drawback since both models achieved results which can be used for further evaluation and fulfill the purpose of this paper.

The metrics introduced in chapter 3 are depicted as follows in Figure 3 and Figure 4.

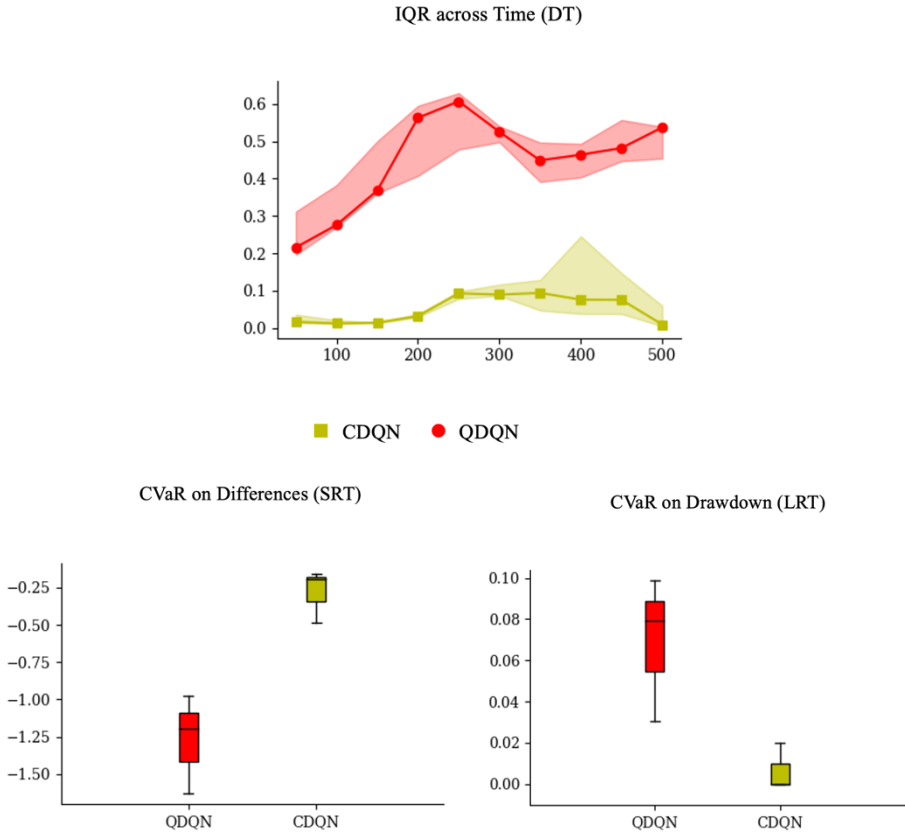


Figure 3: ‘Across Time’ metrics. The IQR across time in the line chart indicates a broader range of variance for the achieved reward for the QDQN. The boxplots show the short and long term CVaR. The left plot indicates a worse estimate of performance for the QDQN on the short term. The right plot shows a higher volatility in performance for the QDQN on the long run.

Depicted in Figure 3 are the ‘Across Time’ metrics. The line chart illustrates the IQR across time, revealing a wider variance in the achieved figure of merit for the QDQN. The boxplots provide insights into the short-term and long-term CVaR. Specifically, the left

plot suggests a poorer performance estimation for the QDQN in the short term. Conversely, the right plot demonstrates higher volatility in the QDQN's performance over the long run. The IQR and CVaR across time suggest that the QDQN in the underlying model will be less reliable over time than the CDQN and that it may have a greater risk of low performance.

Figure 4 shows the ‘Across Runs’ metrics. The line chart on the left illustrates the variability across training runs. As the number of episodes increases, the CDQN outperformed the QDQN. This metric can be seen as a measure of performance, showing that the CDQN achieved better results, starting around the 250th episode. The line chart on the right provides an estimation of the potential worst-case performance across training runs, highlighting a higher risk of poorer performance for the QDQN compared to the CDQN.

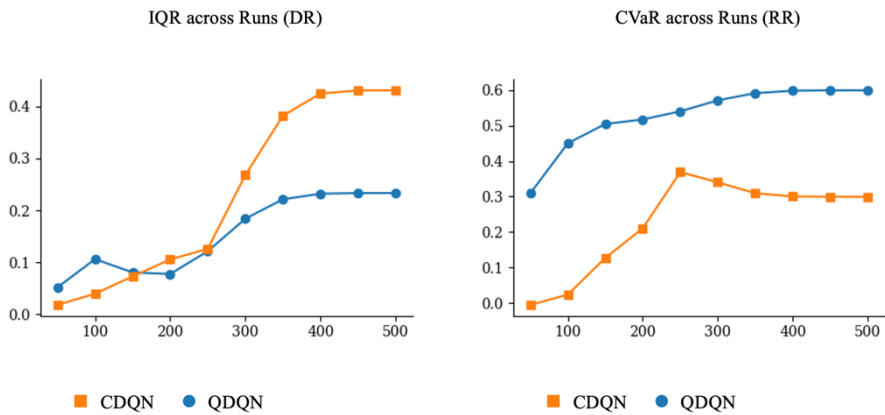


Figure 4: ‘Across Runs’ metrics. The left line chart shows the dispersion across training runs. The CDQN achieved higher results than the QDQN. The right line chart shows an estimate of the potentially worst performance across training runs, indicating that the QDQN has a higher risk of worse performance than the CDQN.

In summary, QRL in our experimental setup does not yet achieve the reliability and performance of the classical model. However, the results are not that far apart, so further experiments and new experimental setups seem promising. Moreover, the results are based on simulations on a classical computer. Measurements on a real quantum computer are therefore still pending. However, the measurement framework shows that with the appropriate metrics, comparability between classical and quantum-based RL models can be achieved.

6 Conclusion

This study presents the implementation and evaluation of a quantum reinforcement learning model utilizing the CartPole environment and the Deep Q Learning algorithm, a state-of-the-art deep reinforcement learning technique. The implementation guidelines are provided using the PennyLane library. The paper explores the underlying principles and potential of reinforcement learning with a focus on variational quantum circuits.

In particular, we introduced a comprehensive set of metrics that aim to assess the performance and reliability of RL algorithms from various perspectives. Moreover, practical guidelines for effectively measuring the reliability of QRL algorithms have been presented. To demonstrate the utility of these metrics, we have applied them to our experimental QDQN algorithm and the CartPole environment. The comparison with the CDQN illustrates the metrics capability to uncover both the strengths and weaknesses of the algorithm under study. By going beyond mean or median performance evaluation, these metrics offer valuable insights into the nuanced aspects of algorithmic reliability, enhancing our understanding of QRL algorithms in a more comprehensive manner.

Our objective is to enable researchers who are new to quantum reinforcement learning to embark on their own investigations in this intriguing field. While the performance of quantum reinforcement learning may not currently surpass that of existing methods, we anticipate that further research endeavors will overcome the limitations of traditional reinforcement learning, yielding promising outcomes.

References

- [Bi17] Biamonte, J.; Wittek, P.; Pancotti, N. et al.: Quantum machine learning. *Nature*, Vol. 549, No. 7671, pp. 195-202, 2017.
- [Ch20] Chan, S.; Fishman, S.; Canny, J.; Korattikara, A. et al.: Measuring the Reliability of Reinforcement Learning Algorithms. 8th International Conference on Learning Representations ICLR, 2020.
- [DB18] Dunjko, V.; Briegel, H. J.: Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Reports on Progress in Physics*, Vol. 81, No. 7, 2018.
- [Je21] Jerbi, S.; Trenkwalder, L. M.; Nautrup, H. P. et al.: Quantum enhancements for deep reinforcement learning in large spaces. *PRX Quantum*, Vol. 2, No. 1, 2021.
- [Jo20] Jordan, S.; Chandak, Y.; Cohen, D. et al.: Evaluating the Performance of Reinforcement Learning Algorithms. *Proceedings of the 37th International Conference on Machine Learning*, PMLR, Vol. 119, pp. 4962-4973, 2020.

- [Kw21] Kwak, Y.; Yun, W. J.; Jung, S. et al.: Introduction to quantum reinforcement learning: Theory and pennylane-based implementation. International Conference on Information and Communication Technology Convergence, ICTC, pp. 416-420, IEEE, 2021.
- [LBC21] Le Lan, C.; Bellemare, M. G.; Castro, P.: Metrics and continuity in reinforcement learning. Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, No. 9, pp. 8261-8269, 2021.
- [Me22] Meyer, N.; Ufrecht, C.; Periyasamy, M. et al.: A Survey on Quantum Reinforcement Learning, 2022.
- [NC10] Nielsen, M.; Chuang, I.: Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge: Cambridge University Press, 2010.
- [NHP23] Neumann, N. M. P.; de Heer, P. B. U. L.; Phillipson, F.: Quantum reinforcement learning. Quantum Information Processing, Vol. 22, No. 125, 2023.
- [Pa19] Paszke, A.; Gross, S.; Massa, F. et al.: Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, Vol. 32, 2019.
- [Sa21] Saggio, V.; Asenbeck, B.E.; Hamann, A. et al.: Experimental quantum speed-up in reinforcement learning agents. Nature, Vol. 591, pp. 229–233, 2021.
- [SJD22] Skolik, A.; Jerbi, S.; & Dunjko, V.: Quantum agents in the gym: a variational quantum algorithm for deep q-learning. Quantum, Vol. 6, 2022.

**Ökologische Nachhaltigkeit -
Umweltinformatik zur Gestaltung
einer nachhaltigen Zukunft**

SeaSaver

Serious Game für die Bewusstseinsbildung bei Kindern zur Verschmutzung von Ozeanen mit Plastikmüll

Stefanie Nina Steinbichl, Janosch Grellner, Nico Hannoun¹, Frank Fuchs-Kittowski²

Abstract: Die Verschmutzung der Meere und Ozeane durch Kunststoffe ist ein dringendes weltweites Umweltproblem. Serious Games sind eine neue Strategie, um Verhaltensänderungen zur Verringerung des Plastikmülls zu fördern. In diesem Beitrag wird das neuartige Serious Game *SeaSaver* vorgestellt, das insbesondere Kinder und Jugendliche für dieses Umweltproblem sensibilisieren und das oftmals noch fehlende Bewusstsein dafür schaffen soll, um eine Verhaltensänderung zu bewirken. Hierfür werden der fachliche Hintergrund, die Konzeption und die Implementierung des Serious Game *SeaSaver* beschrieben.


Keywords: Serious Game, Game Design, Bewusstseinsbildung bei Kindern, Umweltschutz, Plastikmüll in Ozeanen.

1 Einleitung

Die massive Einbringung von Kunststoffen aus unterschiedlichen anthropogenen Quellen in die Gewässer stellt ein ernsthaftes Umweltproblem dar [COP18], [HT22]. Diese oft nicht abbaubaren Abfälle gefährden insbesondere die Ökosysteme der Meere zunehmend, da sie sich in immer größeren Mengen ansammeln. Gleichzeitig ist das Bewusstsein in der Bevölkerung für dieses Problem kaum oder gar nicht ausgeprägt, insbesondere bei Kindern. Aus diesem Grund ist es erforderlich, gerade die junge Generation für dieses Problem zu sensibilisieren, Bewusstsein zu bilden und möglichst eine Verhaltensänderung zur Vermeidung oder zumindest Verringerung von Plastikmüll zu erreichen.

Um dieses Umweltproblem anzugehen, wurden bereits verschiedene Strategien zur Förderung von Verhaltensänderungen zur Verringerung des Plastikmülls entwickelt und erprobt, z.B. Workshops in Schulen [LKN20], Dokumentarfilme [MA12] sowie Online-Kurse und E-Learning-Plattformen [TKL19]. Spielbasierte Lösungen (Serious Games, SGs) sind eine neue Strategie zur Bewältigung des Plastikmüllproblems. Inzwischen wurden bereits eine Reihe von spielbasierten Lösungen entwickelt, um

¹ Hochschule für Technik und Wirtschaft, Umweltinformatik, Wilhelminenhofstraße 75A, 12459 Berlin, {Stefanie.Steinbichl, Janosch.Grellner, Nico.Hannoun}@Students.htw-berlin.de

² Hochschule für Technik und Wirtschaft, Umweltinformatik, Wilhelminenhofstraße 75A, 12459 Berlin, frank.fuchs-kittowski@HTW-Berlin.de,  <https://orcid.org/0000-0002-5445-3764>

Verhaltensänderungen zur Verringerung des Plastikmülls zu erreichen.

In diesem Beitrag soll mit dem SG *SeaSaver - Rette die Weltmeere* ein neuartiger Ansatz für eine spielebasierte Lösung vorgestellt werden. Im Unterschied zu bestehenden SGs zum Thema Verschmutzung von Ozeanen mit Plastikmüll betrachtet *SeaSaver* den globalen Zusammenhang aller Weltmeere und die Auswirkungen auf das jeweilige Ökosystem. Das Spiel stellt sukzessive die verschiedenen Meere und die darin bedrohten Tierarten vor. Darüber hinaus wird auf die reale Problematik der Müllstrudel eingegangen. Diese Elemente wurden zuvor noch nicht in thematisch verwandten SGs behandelt. *SeaSaver* soll zu einer nachhaltigen Bewusstseinsbildung zum Thema Plastikmüllverschmutzung bei Kindern beitragen, indem die komplexe Problematik altersgerecht in den globalen Zusammenhang eingeordnet wird. Dabei ist der Spaßfaktor ein wesentliches Element, da andere Studien gezeigt haben, dass zu viele Lerninhalte auf einmal schnell zu Motivationsverlust und Spielabbruch führen können.

Dieser Beitrag ist wie folgt strukturiert: im folgenden Kapitel 2 wird der fachliche Hintergrund zu Plastikmüll in Ozeanen und sowie der Bewusstseinsbildung bei Kindern zu diesem Thema beleuchtet. Anschließend werden in Kapitel 3 verwandte Arbeiten aus dem Bereich SGs zu dem Thema diskutiert. Danach wird in Kapitel 4 das Konzept des SG *SeaSaver* (u.a. Spielprinzip, Game Design, Spielmechanismen und Levelaufbau) vorgestellt. Technische Aspekte der Implementierung werden in Kapitel 5 beschrieben. Die Arbeit endet mit einer Zusammenfassung und einem Ausblick in Kapitel 6.

2 Fachlicher Hintergrund

In diesem Kapitel wird der fachliche Hintergrund zu Plastikmüll in Ozeanen und sowie der Bewusstseinsbildung bei Kindern zu diesem Thema beleuchtet.

2.1 Plastikmüll in Ozeanen

Seit den 1950er Jahren steigt die weltweite Kunststoffproduktion stetig an und erreichte im Jahr 2021 ein erneutes Hoch mit 391 Mio. Tonnen produziertem Kunststoff [St21]. Jährlich werden rund 4,8-12,7 Millionen Tonnen Kunststoffe in die Weltmeere eingetragen, wodurch nicht nur die Umwelt verschmutzt wird, sondern auch erhebliche Konsequenzen für das Ökosystem und die Organismen entstehen [Ww23].

Kunststoffe gelangen nach ihrer Nutzung auf verschiedenen Wegen in die Umwelt: zum einen durch terrestrische Quellen, wie beispielsweise Verkehr, Bevölkerung, Industrie, Baustellen und zum anderen durch marine Quellen wie Schifffahrt, Fischerei-Aktivitäten oder Ölplattformen. Dabei ist der Eintrag durch terrestrische Quellen im Vergleich deutlich höher als die Menge an Kunststoffen, die direkt auf dem Meer emittiert wird [HT22, S. 19ff.].

In die Umwelt ausgesetzte Kunststoffe werden durch Luftmassen- und Meeresströmungen transportiert und sammeln sich in sogenannten Akkumulationszonen. Die größte Ansammlung an Kunst- und Wertstoffen befindet sich im Nordpazifischen Ozean, der sogenannte *Great Pacific Garbage Patch*. Neben dem horizontalen Transport von Kunststoffen können diese auch vertikal transportiert werden. Dabei lagern sich Mikroorganismen an die zu Mikroplastik fragmentierten Kunststoffe und das Mikroplastik sinkt tiefer ins Meer. Beim zunehmenden Absinken fällt die Wassertemperatur und die Lichteinstrahlung wird geringer, wodurch sich die Mikroorganismen wieder lösen, was zu einem Aufstieg des Mikroplastiks führt [HT22, S. 34ff.].

Plastik altert durch verschiedene chemische Prozesse und als Resultat wird es zunächst spröde und porös. Folglich wird es in immer kleinere Teile fragmentiert. Die Zersetzungszeit von Kunststoffen im Wasser kann mehrere hundert Jahre betragen. Die immense Plastikmüllverschmutzung in den Ozeanen hat verheerende Folgen für die dort lebenden Organismen. Zum einen verfangen sich Meerestiere in Netzen, zum anderen halten sie kleinere Gegenstände und Partikel aus Kunststoff für Nahrung und nehmen diese auf. In ihren Organismen akkumulieren sich toxische Stoffe und teilweise können sie keine Nahrung mehr verdauen. Über die Nahrungskette nimmt der Mensch diese Stoffe schließlich auch auf [Eu21].

2.2 Umweltbewusstsein in der Bevölkerung

Aktuelle Studien zum Umweltbewusstsein in der deutschen Bevölkerung zeigen, dass im Zeitraum von 1996 bis 2016 ein leichter Anstieg erfolgte [Sc20, S. 45]. Vor allem aber ist in den letzten Jahren ein größerer Anstieg des Bewusstseins für ökologische Probleme erfolgt. Dabei liegt der Themenbereich Umwelt- und Klimaschutz auf Platz vier der relevantesten gesellschaftlichen Themenfelder im Jahr 2020 direkt hinter den Themen Zustand des Bildungswesens, Zustand des Gesundheitswesens und soziale Gerechtigkeit.

Das Bewusstsein für umweltrelevante Themen ist insbesondere bei jungen Menschen hoch. So erachten 72 Prozent der 14- bis 29-Jährigen Umwelt- und Klimaschutz als sehr wichtig, bei den 40- bis 49-Jährigen sind es nur noch 54 Prozent. Für Frauen spielt das Thema unabhängig vom Alter generell eine größere Rolle (71 Prozent) als bei Männern (59 Prozent) [WG22, S. 25f.].

Gleichzeitig gaben 39 Prozent der Befragten an, zum Thema Klimawandel nur *etwas informiert* zu sein, auch in der Gruppe der Jugendlichen und jungen Erwachsenen, welche die aktuelle Klimabewegung vor allem vorantreibt. Dies zeigt, dass es nicht die eine junge klimabewusste Generation gibt, sondern sich der Wissensstand und die Einstellung bzgl. Klima- und Umweltschutz in der jungen Bevölkerung vielschichtiger und heterogener darstellt als es zunächst vermuten lässt [WG22, S. 47f.].

Um das Bewusstsein und die Veränderungsbereitschaft in der Bevölkerung für den Umweltschutz weiter zu stärken, ist es notwendig, die Wissens- und Erfahrungsvermittlung passend zur Lebenswelt der Zielgruppen zu gestalten. Die

Bewusstseinsbildung und das Schaffen positiver Zielvorstellungen, wie beispielsweise der Zugewinn an Lebensqualität durch umweltbewusstes Verhalten, sind schließlich die Voraussetzung für das Fördern der Handlungsbereitschaft [WG22, S. 79]. Unabdingbar für eine Transformation ist also in erster Linie das Abholen der jeweiligen Bevölkerungs- und Altersgruppen durch passgenaue Maßnahmen, die sie in ihrem individuellen Lebensumfeld erreichen und mit denen sie sich identifizieren können.

2.3 Umweltbewusstsein bei Kindern

Zwar hat das Thema Umweltschutz bereits bei Vorschulkindern Einzug gehalten. So gab es im Jahr 2021 ca. 2.000 Natur- und Waldkindergärten in ganz Deutschland mit steigender Tendenz, welche die naturpädagogische Arbeit mit Kindern in den Mittelpunkt stellen [Bu23], [Ra21]. Bei einer Gesamtzahl von knapp 60.000 herkömmlichen Kindertageseinrichtungen in Deutschland [Bo22] erscheint die Anzahl dieser Natur- und Waldkindergärten jedoch noch verschwindend gering - gerade einmal 3,33%. Von einer flächendeckenden Sensibilisierung geschweige denn Bewusstseinsbildung für Umweltthemen in diesem Alter kann daher keine Rede sein.

Bei Jugendlichen und jungen Erwachsenen zeigt sich ein ambivalentes Bild. Laut einer Studie von Greenpeace und der Leuphana Universität Lüneburg aus dem Jahr 2011 haben zwar 68 Prozent der 1.070 bundesweit befragten 15-24-Jährigen ein starkes Bewusstsein für Umweltprobleme und Nachhaltigkeit. Allerdings gab nur etwa die Hälfte der Befragten an, das Thema nachhaltige Entwicklung in der Schule behandelt zu haben [GKR12, S. 2ff.]. Ob Umweltschutz in jungen Jahren ein Thema ist, hängt also stark vom Elternhaus und dessen sozialen Status ab [GKR12, S. 5]. Inwiefern dieser Privatverantwortlichkeit nachgekommen wird, beleuchtet eine repräsentative Umfrage der Konrad Adenauer Stiftung zum Thema Umweltbewusstsein bei Familien mit Kindern und Kinderlosen aus dem Jahr 2019 sowie eine Studie des Umweltbundesamtes zu Umweltbewusstsein und Umweltverhalten im Jahr 2016. Es zeigte sich, dass bei Erziehungsberechtigten mit minderjährigen Kindern in der *Rush Hour des Lebens*, also unter Doppelbelastung von Erwerbstätigkeit und Kindererziehung, ein sozial-ökologisches Engagement kaum vorhanden ist und andere Themen, insbesondere die Bewältigung des Familienalltags, im Vordergrund stehen. Umweltschutzthemen werden hier folglich weniger diskutiert als bei Nicht-Erziehungsberechtigten [Po20, S. 6], [SHS18, S. 12]. Die Kinder kommen mit dem Thema innerhalb der Familie demzufolge weniger in Berührung.

Das würde wiederum den Schulunterricht als Wissensvermittler in Sachen Umweltschutz und insbesondere Plastikmüll in den Meeren auf den Plan rufen. Allerdings ist das Thema in den Bildungs- und Lehrplänen der unteren Klassenstufen gar nicht oder nur sehr spärlich integriert. Es wird zwar in den Bundesländern Saarland, Bayern, Bremen, Hamburg und Nordrhein-Westfalen in Vor- und Grundschule auf den bewussten Umgang mit Wasser eingegangen, allerdings bleibt das Thema Plastikmüllverschmutzung der Weltmeere außen vor [ACB16].

3 Verwandte Arbeiten – Serious Games zur Verschmutzung von Ozeanen mit Plastikmüll

In diesem Kapitel werden verwandte Arbeiten aus dem Bereich SGs zur Verschmutzung von Ozeanen mit Plastikmüll diskutiert. Hierzu wird zunächst der Begriff SG geklärt und die Eignung von SGs zur Bewusstseinsbildung bei Kindern diskutiert. Anschließend werden existierende SGs vorgestellt und mit dem in diesem Beitrag beschriebenen SG SeaSaver verglichen, um Gemeinsamkeiten und Unterschiede – insbesondere die Neuartigkeit des Ansatzes von SeaSaver – deutlich zu machen.

3.1 Serious Games

SGs verbinden Spielen mit Lernen [RE13] und unterscheiden sich von herkömmlichen Spielen in ihrem Anspruch, nicht nur zu unterhalten, sondern auch Wissen zu vermitteln, Bewusstsein zu schaffen, Handlungsabläufe anhand von Simulationen zu verinnerlichen oder einen therapeutischen Nutzen bereitzustellen, Spiel- und Lernkomponente also geschickt miteinander zu verweben [Gr20].

SGs und Gamification verfolgen das gemeinsame Ziel, zu einer bestimmten Denkweise, inneren Einstellung oder Verhalten zu motivieren [Jo17, S. 4]. Beide Ansätze arbeiten mit dem Prinzip der Belohnung durch das Sammeln von Punkten und Abzeichen sowie durch das Führen von Ranglisten, durch die man sich mit anderen Spielenden messen kann. Der Unterschied zwischen SGs und Gamification liegt dabei im Umfang ihrer Anwendung bzw. ihrer autarken Ausführbarkeit. Während Gamification spielerische Elemente als Zusatz für bereits existierende Anwendungen beschreibt, handelt es sich bei SGs um in sich geschlossene, unabhängig existierende Spiele mit eigenem Game Design, Regelwerk und einer auf Handlung und Leveln basierenden Spielmechanik.

3.2 Eignung von SG zur Bewusstseinsbildung bei Kindern

SGs finden Verwendung in unterschiedlichsten Bereichen und Marktsegmenten mit jeweils spezifischer Ausrichtung und Zielsetzung. Insbesondere können sie Bewusstsein für ein gesellschaftlich bedeutsames Themengebiet schaffen [Gö20, S. 105], [De11, S. 2]. und zur Verhaltensänderung in den Bereichen Bildung und Gesundheit eingesetzt werden, wie zahlreiche Studien zeigen [KH19].

Darüber hinaus ermöglichen SGs, komplexe Sachverhalte verständlich zu kommunizieren und Spieler:innen zu motivieren, die bereit sind, Verantwortung zu übernehmen und sich aktiv an Lösungen zu beteiligen. Abstrakte Themen, die sonst kaum im Fokus stehen, können durch solche Methoden vermittelt und Zusammenhänge anschaulich dargestellt werden [Fo20].

Zu diesen gesellschaftlich relevanten und komplexen Themen zählen vor allem die

drängenden Umweltprobleme unserer Zeit. Es existieren zahlreiche Möglichkeiten, wie frühzeitig das Umweltbewusstsein mit SG gefördert und nachhaltiges Denken verinnerlicht werden kann [WLW17]. So wurde Gamification bereits im Bereich der Umwelt und Nachhaltigkeit eingesetzt [MSZ22], z.B. Energieeinsparung [TOC21] und Erhaltung der biologischen Vielfalt [SAM15].

SGs können im Umweltbereich einen wichtigen Beitrag zur Sensibilisierung und einem höheren Bewusstsein für umweltrelevante Fragestellungen leisten. Insbesondere bieten sie Kindern einen spielerischen Zugang zu diesen ernsten und zugleich zukunftsrelevanten Themengebieten sowie die Möglichkeit, sich altersgerecht damit auseinanderzusetzen.

3.3 Serious Games zur Verschmutzung von Ozeanen mit Plastikmüll

Eine aktuelle Übersicht über SGs für die Bewältigung des Plastikmüllproblems [VG23] zeigt, dass es bereits mehr als 20 SGs zum Thema gibt. Dabei ist in der Regel die Zielgruppe die allgemeine Öffentlichkeit und das Ziel die Sensibilisierung für die Auswirkungen der Plastikverschmutzung sowie die Bewusstseinsbildung. Zudem ergab die Auswertung der Konzeption und Implementierung der SGs, dass das Fortschreiten im Spiel durch bestimmte Herausforderungen oder Levels, die zum Gewinnen durchlaufen werden müssen, das am häufigsten eingesetzte Element ist. Weitere wiederkehrende Bestandteile sind ein Ranglisten- oder Punktesystems sowie Abzeichen. Dadurch können User:innen gegeneinander antreten und ihre Erfolge teilen. Bei der Auswertung der Wirksamkeit der spielbasierten Lösungen, um den Verbrauch von Einwegplastikartikeln bei Einzelpersonen und Organisationen zu verringern, wurde festgestellt, dass bei einem Großteil der Studien eine Zunahme des Bewusstseins oder des Wissens über das Plastikproblem nach der Teilnahme an den Spielen gemessen werden konnte.

3.4 Vergleich und Abgrenzung von existierenden Serious Games zum Thema

In Tabelle 1 werden existierende SGs zum Thema Verschmutzung von Ozeanen mit Plastikmüll aufgelistet und anhand verschiedener Kriterien zu Spielmechanismen und zur Wissensvermittlung verglichen. Damit sollen zum einen gängige und erforderliche Funktionen für SGs dieser Art ermittelt und zum anderen das Konzept von SeaSaver zu existierenden Lösungen abgegrenzt werden, das im folgenden Kapitel beschrieben wird.

Die verwendeten Kriterien zu Spielmechanismen sind:

- **Leveling:** Erreichen höherer Levels (durch Absolvieren bestimmter Aufgaben oder Sammeln von Erfahrungspunkten)
- **Unlocking:** Freischalten von Werkzeugen oder Levels
- **Advancement:** Spielfortschritt durch Abschließen von Missionen oder Erreichen bestimmter Meilensteine
- **Progress-System:** Mechanismus, der den Spielfortschritt verfolgt und Belohnungen oder neue Möglichkeiten bietet

- **Game Flow:** flüssiger Ablauf des Spiels, ohne größere Hindernisse oder Unterbrechungen
- **Story-Driven Progression:** Fortschreiten durch die Spielhandlung mit neuen Ereignissen und Schauplätzen
- **Skill Development:** Verbessern der spielerischen Fähigkeiten
- **Challenge Progression:** Steigern der Schwierigkeit oder Herausforderungen
- **Achievement Unlocking:** Freischalten von Trophäen oder Badges durch das Erreichen bestimmter Ziele oder das Abschließen besonderer Aufgaben
- **Metagame Progression:** Fortschreiten auf einer übergeordneten Ebene außerhalb des eigentlichen Spiels, wie zum Beispiel durch Ranglisten
- **Time Limitation:** zeitliche Begrenzung einzelner Levels
- **Social Sharing:** Teilen von Spielergebnissen mit der Peergroup

Kriterium	Rettet unseren Ozean [Tr23]	Klima Game [Lü23]	Save the Ocean [So19]	Contact from the future [Pa21]	SeAdventure [RC20]	Sea-Saver
Leveling	-	-	+	o	+	+
Unlocking	-	-	-	o	o	+
Advancement	-	-	+	+	+	+
Progress-System	-	-	+	+	+	+
Game Flow	+	-	+	-	-	+
Story-Driven Progression	-	-	-	+	+	+
Skill Development	-	-	-	-	o	+
Challenge Progression	-	-	-	o	o	+
Achievement Unlocking	-	+	+	+	o	+
Metagame Progression	+	-	+	o	+	+
Time Limitation	+	+	-	-	o	-
Social Sharing	+	-	+	-	o	+
Texttafeln	-	+	-	-	+	+
Infografiken	-	+	-	+	-	+
Videosequenzen	-	-		+	+	+
Quizgames	+	-	-	+	-	+
Story	-	-	-	+	-	+
Authenticity	-	+	+	+	+	+
Global Approach	-	-	-	-	-	+

Tab. 1: Vergleich von SGs hinsichtlich Spielmechanik und SG-spezifischer Wissensvermittlung (o: keine Angabe; + vorhanden; - nicht vorhanden)

Als SG-spezifische Kriterien hinsichtlich der Wissensvermittlung wurden verwendet:

- **Texttafeln**
- **Infografiken**
- **Videsequenzen**
- **Quizgames**
- **Story**
- **Authenticity**: realistische inhaltliche Darstellung der Umweltprobleme und -risiken
- **Global Approach**: weltweite anstelle von lokaler Betrachtung

4 Konzeption des SG SeaSaver

In diesem Kapitel werden das Ziel und die Zielgruppe sowie das Prinzip und der Aufbau des Spiels *SeaSaver – Rettet die Weltmeere* erläutert. Dabei werden die Spielmechanismen und -elemente anhand von drei Levels exemplarisch beschrieben.

4.1 Ziel und Zielgruppe

Das in diesem Beitrag präsentierte SG *SeaSaver* soll Kinder im Alter von ca. 10-12 Jahren als Zielgruppe adressieren, da diese sich der Tragweite der Plastikmüllverschmutzung der Meere und wie sie diese durch ihr eigenes Handeln eindämmen können, oftmals noch nicht bewusst sind. Gleichzeitig verfügt diese Altersgruppe bereits über die kognitiven Fähigkeiten, sich ein eigenes Bild von der Welt zu machen, eigene moralische Überzeugungen zu entwickeln, eine andere Perspektive einzunehmen und auch schon anspruchsvollere Hintergrundinformationen und komplexere Sachverhalte zu verstehen [Fr22].

Außerdem hat diese Altersgruppe eine hohe Affinität zu digitalen Spielen. So verwenden 60 Prozent der 6- bis 13-Jährigen regelmäßig Zeit auf das Spielen digitaler Games an Computer, Konsole, Tablet und Smartphone in ihrem Alltag. 15 Prozent spielen zumindest selten und nur etwa ein Viertel der Kinder spielt nie. Mit steigendem Alter nimmt auch die Nutzung deutlich zu: 36 Prozent der 6- bis 7-Jährigen spielen mindestens einmal in der Woche, bei den 12- bis 13-Jährigen sind es bereits 75 Prozent [Fe21, S. 61f].

4.2 Spielprinzip

Im Bereich SGs gibt es zwei prinzipielle Herangehensweisen, die bildhaft mit den Termini *chocolate covered brokkoli* und *chocolate with brokkoli chunks* gegenübergestellt werden. Der Vergleich von Brokkoli und Schokolade ist mittlerweile ein gängiger Begriff in der Wissenschaft, um die Dynamik von Lernen und Spielspaß zu bezeichnen. Dabei symbolisiert der Brokkoli den Lerninhalt, der von Kindern nur ungern gegessen wird, wohingegen die Schokolade den Genuss am Spiel darstellt. Somit ist der Ansatz des mit

Schokolade überzogenen Brokkolis nur bedingt geeignet, da den Kindern beim ersten Biss in den Brokkoli bewusst wird, dass es sich nicht um eine Süßigkeit handelt und sie eigentlich nur hinters Licht geführt wurden.

Deshalb soll SG *SeaSaver* der entgegengesetzte Ansatz, also die Schokolade mit Brokkoli-Stückchen, angewandt werden. Das bedeutet konkret, dass der Spielspaß im Vordergrund steht und Lerninhalte Stück für Stück nähergebracht werden. Mit jedem erfolgreich beendeten Level wird ein kleines Stück Wissen vermittelt. Dabei steigen sowohl der spielerische Anspruch, als auch die Lerninhalte in ihrer Komplexität. Auf diese Weise soll eine höhere Bewusstseinsbildung der Kinder für das Thema Verschmutzung der Meere durch Kunststoff erreicht werden, ohne sie bereits zu Beginn des Games aufgrund des fehlenden Spielspaßes zu verlieren. Das Sammeln von Punkten und Abzeichen zur Spielmotivierung soll ebenfalls integriert werden.

4.3 Game Design

Bereits in den 1980er und 90er Jahren wurde festgestellt, dass eine Anlehnung an die aus Fernsehproduktionen vertraute Erzählweise sowie visuelle und auditive Gestaltungselemente wie Schnitt, Kameraeinstellungen oder Soundeffekte Kindern das Lernen durch gezielte Aufmerksamkeitssteuerung erleichtern kann. Diese von Kindern präferierte *Grammatik von TV-Inhalten* hat ihr Pendant in der Sprache von digitalen Spielen. Auch in den bei Kindern sehr beliebten Action Games werden beispielsweise häufig eine Vorgeschichte sowie kurze Filmsequenzen im Spielverlauf integriert, um die Narrative gemäß dem Spielfortschritt voranzutreiben und die User:innen emotional zu involvieren. Daneben spielen vor allem bei jüngeren Kindern auch die cartoonartige Grafik der Level und Spielfiguren sowie ein eingängiger Soundtrack und einfach zu verstehende Spielregeln eine wichtige Rolle [BI09, S. 133f.].

Daher wurde entschieden, ein 2D-Side-Scroller-Game im Stil der Raumschiff-Spielreihe *R-Type* mit dem Titel *SeaSaver - Rette die Weltmeere* zu entwickeln, das in einer Unterwasserwelt spielen soll. In einer kurzen Videosequenz zu Beginn wird erklärt, dass die Aufgabe darin besteht, die Weltmeere vom Plastikmüll zu befreien und die Tiere vor einem grausamen Tod durch Plastiknetze zu bewahren. Ziel des Spiels ist das Beseitigen von Plastikmüll, wie z.B. Autoreifen, Gummistiefeln und anderen Kunststoffserzeugnissen sowie das Befreien der Meeresbewohner mit Hilfe entsprechender Werkzeuge und Fähigkeiten, die nach und nach erspielt werden. Die Reise beginnt in der Ostsee und führt um den gesamten Globus. Level für Level werden die bedrohten Tiere der jeweiligen Meere gerettet und es werden dabei Punkte und Abzeichen gesammelt. Die Kollision mit herumschwimmenden Müll führt zum Abzug der Lebenspunkte. Sind alle Leben aufgebraucht, muss das Spiel wieder von vorne begonnen werden. Im letzten Level wird der Endgegner im Nordpazifik entdeckt, es handelt sich um den gigantischen Müllstrudel *Great Pacific Garbage Patch* und dieser muss durch geschicktes Einsetzen der zuvor erworbenen Kompetenzen besiegt werden. Durch den weltumspannenden Umfang des Spiels soll deutlich gemacht werden, dass es sich bei der Plastikmüllverschmutzung um

ein globales Problem handelt, von dem alle Meere und Ozeane betroffen sind.

Mit dem Verlauf des Spiels soll als roter Faden eine Rahmenhandlung über das Schicksal einer vom Plastikmüll bedrohten Schildkrötenfamilie verwoben werden, die mittels kurzer Filmsequenzen zwischen den Levels sowie kleiner Hinweise innerhalb der Levels episodenhaft erzählt wird. Abhängig von der Leistung der Spielenden können die Mitglieder der Schildkrötenfamilie am Ende gerettet werden und wieder unbeschwert durch die Meere schwimmen oder sie verlieren ihr Leben durch die Folgen der Plastikverschmutzung. Durch die kontinuierliche Konfrontation mit diesem *Einzelschicksal*, das stellvertretend für das allen Meeresbewohnern zugefügte Leid durch den Plastikmüll steht, sollen die Spielenden emotional noch stärker ins Spiel (und die Thematik) involviert und ihre Motivation, die Mission erfolgreich zu Ende zu bringen und das Spiel abzuschließen, gesteigert werden.

4.4 Spielmechanismen und -elemente

Als charakteristischer Bestandteil von digitalen Spielen und damit auch von SGs haben sich bestimmte Spielmechanismen und -elemente herausgebildet, die sich als essentiell zur Motivationsförderung der Spielenden bewährt haben. Dazu zählen u.a. die Abbildung des erreichten Spielstatus, z.B. in Form von Titeln oder Badges, was einen Vergleich mit anderen Spielenden ermöglicht und so den gegenseitigen Wettbewerb bestärkt, ebenso wie einsehbare Rang- oder High-Score-Listen, die im Spiel oft mit einem Belohnungssystem verbunden sind. Eine wichtige Rolle nimmt auch Social Sharing ein, also dem Teilen von Spielerfahrungen und -ergebnissen in sozialen Netzwerken. Des Weiteren sind sogenannte Quests zu nennen, also klar definierte Fleiß- oder Extraaufgaben, die meist unter Zeitbegrenzung zu lösen sind, außerdem Resultattransparenz und Feedback, das heißt eine konsequente Rückmeldung an die Spielenden über die (möglichen) Konsequenzen des Handelns, und Epic Meaning, also das Verbinden des Spiels mit einer erstrebenswerten, sinnstiftenden Mission, mit der man zu einem größeren Ziel beiträgt. Darüber hinaus lässt eine dynamische Fortschrittsanzeige erkennen, wie weit man im Level bzw. im Spiel vorangekommen ist, zum Beispiel in Form eines wachsenden Balkens oder als Prozentangabe, was ebenfalls zum Weiterspielen motiviert. Um User:innen vor allem bei Lernspielen nicht zu überfordern, ist es ferner hilfreich, jeweils nur die Informationen zu offenbaren, die für die aktuelle Aufgabe auch wirklich notwendig sind, was als *Cascading Information* bezeichnet wird [Ko12, S. 4f.], [Jo17, S. 14].

Für das SG *SeaSaver - Rette die Weltmeere* sind einige der genannten Spielmechanismen und -elemente entsprechend angepasst und implementiert worden, um die Spielmotivation zu erhöhen. So bekommen die Spielenden nach erfolgreichem Absolvieren eines Levels ein passendes Abzeichen verliehen (siehe Abb. 1).

Außerdem werden die in den verschiedenen Level gesammelten Punkte durch das Beseitigen von Plastikmüll in einer High-Score-Tabelle für alle Spielenden gesammelt und nach jedem Level präsentiert, damit diese ihren Punktestand vergleichen können und

angespornt werden, weitere Punkte zu sammeln. Über das graue Share-Icon am rechten unteren Bildschirmrand besteht die Möglichkeit, den eigenen Punktestand in den sozialen Netzwerken mit anderen zu teilen (siehe Abb. 2).



Abb. 1: Seal-Saver-Abzeichen (Level 2)



Abb. 2: High-Score-Tabelle

Ein weiterer Bestandteil des Spiels sind auf die jeweiligen Level-Welt zugeschnittene Challenges und Quests, die das Level vertiefen und bei denen die Spielenden Extrapunkte sammeln können. Ein Epic Meaning, also das Beitragen zu einem größeren Ziel, ist schon durch das Spielkonzept an sich gegeben, da das Retten der Meeresbewohner und das Befreien der Meere vom Plastikmüll als zentrale Mission im Mittelpunkt des Spiels steht. Eine Fortschrittsanzeige wird mittels einer Weltkarte zwischen den Levels visualisiert, in der die bereits vom Plastik befreiten Regionen in einer anderen Farbe markiert sind.

Als zusätzliches Element kommen (interaktive) Infografiken zwischen den Levels sowie kleine Info-Pop-Ups innerhalb der Levels zum Einsatz. Diese können detaillierte Hintergrundinformationen zur Thematik, wie zum Beispiel zu bedrohten Tierarten während des Spiels vermitteln, aber auch inwiefern das eigene Handeln eine Rolle spielt und damit die eigene Selbstwirksamkeit bestärkt. Es besteht zum Beispiel die Möglichkeit, auf bestehende Umweltschutz-Projekte wie die Citizen Science-Aktion *Plastik Piraten*³ aufmerksam zu machen sowie positiv formulierte Erreichungs- bzw. Annäherungsziele aufzuzeigen (zum Beispiel weniger Müllproduktion durch achtsamen Konsum, Schutz der Artenvielfalt etc.) anstatt den Fokus auf notwendige Entbehrenungen zu legen. Abb. 4 zeigt, wie eine Infografik für das Level *Arktischer Ozean* aussieht. Die Quellen für die Informationen in der Infografik werden im Spielmenü in einem eigenen Bereich aufgeführt, so dass sich interessierte Nutzer:innen bei Bedarf weiter informieren können, der Spielfluss aber nicht unterbrochen wird. Um die Informationen aus den Infografiken zu verinnerlichen, werden die vermittelten Fakten in kleinen Quizgames zwischen den Levels abgefragt, bei denen wiederum Extrapunkte eingesammelt werden können.

³ <https://www.plastic-pirates.eu/de>

4.5 Levelaufbau und -details

Das Durchlaufen der neun Level des Spiels entspricht der Reiseroute der Spielenden durch die verschiedenen Meere und Ozeane: von der Ost- und Nordsee (Level 1) geht es in den arktischen Ozean (Level 2), von dort in den Nordatlantik (Level 3), dann über das Mittelmeer (Level 4) in den Südatlantik (Level 5), anschließend in den indischen Ozean (Level 6), gefolgt vom Antarktischen Ozean (Level 7) und Südpazifik (Level 8) und schließlich in den Nordpazifik (Level 9) zum *Great Pacific Garbage Patch* (Abb. 3).



Abb. 3: Reiseroute durch die Meere (=Level)

Jedes Level stellt eine eigene Spielwelt mit charakteristischen Eigenheiten dar, entsprechend dem jeweiligen Ökosystem und der tatsächlichen Bedrohungslage durch Plastikmüll. Für die Ausgestaltung jedes Levels wurden folgende Inhalte konzipiert:

- **Ausgestaltung, Besonderheiten:** Wie sieht das Level aus? Welche Besonderheiten gibt es in diesem Meer bzw. Ökosystem? Wie sieht die Verschmutzung in diesem Meer genau aus?
- **Fähigkeiten, Werkzeuge:** Welche Fähigkeiten, Werkzeuge und Tools sind vorhanden?
- **Bedrohte Tierarten:** Welche Tierarten im jeweiligen Ökosystem sind bedroht und können gerettet werden?
- **Endgegner:** Welche Gestalt hat der Level-Endgegner bzw. das zu rettende Tier?
- **Abzeichen:** Welches Abzeichen oder welche Rangstufe erhält man bei erfolgreichem Absolvieren des Levels?
- **Challenges:** Gibt es spezielle Challenges bzw. Quests?
- **Rahmenhandlung:** Was passiert mit der Schildkrötenfamilie in der Rahmenhandlung?

Für die Level *Arktischer Ozean* (Level 2), *Europäisches Mittelmeer* (Level 4) und *Nordpazifik* (Level 9) werden die Konzepte exemplarisch in Tabelle 2 dargestellt.

Level	2: Arktischer Ozean	4: Europäisches Mittelmeer	9: Nordpazifik
Ausgestaltung, Besonderheiten	Eisschollen treiben an Wasseroberfläche, darauf befinden sich gefährdete Tiere, (z.B. strangulierte Bartrobbe), deren Rettung Extrapunkte bringt; im Wasser treiben Plastikteile; mit Mikroplastik belastete Sedimente sind farbig hervorgehoben	Müll im Wasser besteht v.a. aus Einwegprodukten (z.B. Getränkeflaschen, Verschlüsse, Deckel, Zigarettenstummel, Wattestäbchen, Tüten und Verpackungen für Chips und Süßigkeiten, Hygieneartikel wie Feuchttücher, Tampons etc.).	Gigantische Müllteppiche bewegen sich langsam auf dem Wasser. Ein am Plastikmüll verendeter Wal treibt vorbei.
Fähigkeiten, Werkzeuge	Schneidewerkzeug zur Befreiung von in Netzen und Seilen gefesselten bzw. strangulierten Tieren, feines Spezialnetz zum Einsammeln von Mikroplastik	Schleppnetze zur effizienteren Beseitigung des Plastikmülls	Garbage Patch Destroyer, ein äußerst leistungsstarker Unterwasser-Staubsauger, zur Beseitigung von Müll jeglicher Größe
Bedrohte Tierarten	Eisbär, Seehund, Bartrobbe, Seeschwalbe, Großer Sturmtaucher, Eissturmvogel, Blauwal u.a. [Be17], [Se17], [Iu22]	Dick-Stör, Gemeiner Engelhai, Echte Karettschildkröte, u.a. [Iu22]	Sonnenblumen-Seestern, Felsenbarsch, Hundshai, u.a. [Iu22]
Endgegner	Befreiung einer Robbe aus großem Fischernetz	Befreiung eines Tümmers aus einem Netz	<i>Great Pacific Garbage Patch</i> als Endgegner des Spiels
Abzeichen	Seal-Saver-Abzeichen (Abb.1)	Dolphin Saver Abzeichen.	Finales Sea-Saver-Abzeichen

Challenges	Vermeide den Zusammenprall des Boots mit Eisschollen oder Eisbergen durch geschicktes Navigieren sowie das Einfrieren des Boots! Filtere Mikroplastik aus den markierten Sedimenten!	Sammlle 100 Plastiktüten!	Zerstöre den Garbage Patch!
Rahmenhandlung	Die Schildkrötenfamilie kommt an einem Fischernetz vorbei. Wenn das Level erfolgreich abgeschlossen wurde, schwimmt sie sicher daran vorbei. Falls das Level nicht erfolgreich abgeschlossen wurde, verendet eine Schildkröte im Netz.	Die Schildkröten kommen an der Adria vorbei und machen erstmal eine Verschnaufpause.	Die Schildkrötenfamilie ist endlich in Sicherheit und kann ohne Bedrohung durch den Plastikmüll durch die Weltmeere ziehen.

Tab. 2: Inhaltliche Levelgestaltung

Ein wichtiges Element zur Wissensvermittlung stellen Infografiken und Info-Pop-Ups dar. Bspw. wurden für Level 2 (*Arktischer Ozean*) u.a. folgende Zahlen, Daten und Fakten recherchiert und anschließend als Infografik umgesetzt (Abb. 4).

- Mehr als 80 Prozent des vorgefundenen Mülls besteht aus Plastik, das meiste davon stammt aus Fischereien. Damit ist der Plastikanteil im arktischen Ozean sogar etwas höher als der weltweit an Stränden vorgefundene Plastikanteil von 72 % [Be17].
- Der aufgefundene Müll ließ sich in folgende Kategorien einordnen: Plastik allgemein, von Fischereien stammende Plastikteile (Fischernetze, Fender, Seile), Kleidung und Schuhe, Metall, Glas sowie organische Substanzen [Be17].
- Die menschlichen Hinterlassenschaften finden sich nicht nur an den Küsten, sondern auch am Grund der Tiefsee und als Mikroplastik in den Eisschollen auf den Ozeanen. Die Mikroplastik-Konzentration in den Schollen liegt sogar über der von belasteten Meeresgebieten im Bereich der großen Müllstrudel. Wenn die Eisschollen schmelzen, rieseln die Plastikteilchen vermutlich auf den Meeresboden. Dies würde erklären, dass die Sedimente an den nördlichsten Messpunkten am stärksten belastet sind [Se17].



Abb. 4: Beispiel für eine Infografik (erstellt mit VISME)

5 Implementierung und technische Aspekte

In diesem Kapitel werden die für die Umsetzung des SG *SeaSaver* eingesetzten Technologie sowie die visuelle und auditive Umsetzung beschrieben.

Mithilfe von *GameSalad* als Drag-and-Drop-Editor wurde ein 2D-Side-Scroller-Game entwickelt. Mittels sogenannter *Scenes* wurden die verschiedenen Bildschirme, wie z.B. Intro, Levelauswahl, Highscore und die jeweiligen Level selbst gestaltet. Jeder dargestellte Gegenstand wurde im Backend durch einen sogenannten *Actor* erzeugt. Mithilfe von konditionaler Logik konnten Eigenschaften, Verhalten und Zustand angepasst werden. Anhand des Master Controllers wurden u.a. die Spawn-Times, also das Erscheinen bestimmter Objekte, kontrolliert. Zusätzlich wurden *Photoshop* und ein GIF-Splitter verwendet, um die Bilder des Plastikmülls von ihren Hintergründen freizustellen sowie zur Darstellung von Animationen durch Einzelbilder.

Wie in Abb. 5 exemplarisch für Level 2 dargestellt, bewegt sich der Avatar durch eine zweidimensionale Unterwasserwelt und muss diese vom umher treibenden Plastikmüll befreien. Am linken oberen Bildschirmrand befindet sich die Punkteanzeige. Für das

Abschießen des Mülls werden Punkte vergeben. Am Ende des Levels muss mit einer hydraulischen Schere eine Robbe aus einem Fischernetz befreit werden.

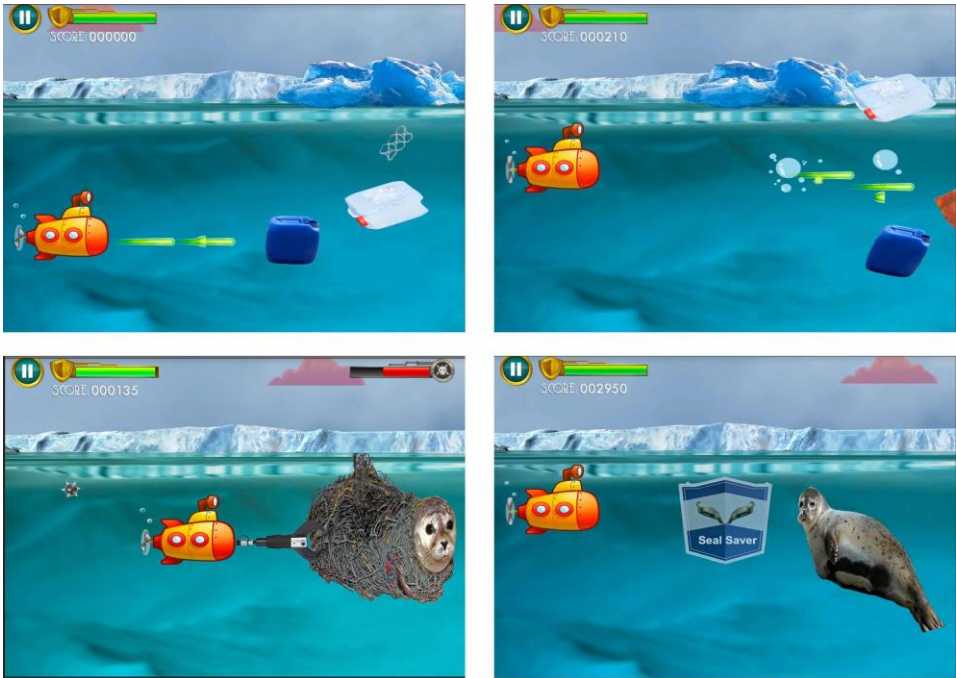


Abb. 5: Screenshots vom Spiel (Level 2)

Allerdings ist hervorzuheben, dass das Abschießen des Mülls möglicherweise suggeriert, dass die Beseitigung einfach wäre. Dahingegen wäre ein größerer Lerneffekt zu erwarten, wenn die Beseitigung mit mehr Aufwand verbunden wäre. Zum Beispiel könnte der Müll mit Keschern verschiedener Größe aufgefangen werden, die der jeweiligen Situation angepasst werden müssen.

Als musikalische Untermalung des Levels wurde der lizenzfreie Track *Draft monk - ambience* von McFunkypants ausgewählt⁴. Dieser schafft die gewünschte atmosphärische Dichte durch die verflochtenen Synthesizer-Pattern sowie die nötige Spannung durch treibende Percussionakzente und Bassläufe. Die explosionsartigen Soundeffekte der auftretenden Laserschüsse aus der GameSalad Vorlage wurden durch neutralere Plopp-Geräusche ersetzt. Der gerettete Seehund wurde mit einem passenden Laut beim Wegschwimmen vertont, um ihn noch lebendiger wirken zu lassen.

⁴ <https://www.chosic.com/download-audio/28363/>

6 Zusammenfassung und Ausblick

Plastikmüllverschmutzung in den Ozeanen stellt ein ernstzunehmendes und drängendes Umweltproblem dar, aber eine Sensibilisierung und Bewusstseinsbildung ist bei Kindern für dieses Thema unbedingt erforderlich. SGs haben ein großes Potential, Kindern die Problematik altersgerecht bewusst zu machen und ihnen spielerisch Sachkenntnis darüber zu vermitteln. In einem Alter ab ca. zehn Jahren verfügen Kinder über die notwendigen kognitiven Fähigkeiten für eine differenzierte Auseinandersetzung damit und zeigen gleichzeitig eine hohe Affinität zu digitalen Spielen, weswegen diese Altersgruppe als Zielgruppe für das SG *SeaSaver - Rette die Weltmeere* ausgewählt wurde.

SeaSaver - Rette die Weltmeere ist als 2D-Side-Scroller-Game mit Fokus auf Spielspaß angelegt, bei dem die Beseitigung des Plastikmülls in den Meeren und die Rettung gefährdeter Tiere die zentrale Mission darstellt. Dabei bewegt sich die spielende Person in einem multifunktionalen U-Boot chronologisch durch neun Level, die jeweils einem bestimmten Meer oder Ozean entsprechen, mit dem Finale im Nordpazifik und dem Great Pacific Garbage Patch als Endgegner. Auf dem Weg dorthin gilt es, möglichst viele Punkte und Abzeichen zu sammeln, um neue Fähigkeiten freizuschalten. Wissenswerte Fakten und Informationen zur Situation in den jeweiligen Meeren und den bedrohten Tierarten werden über Pop-Ups und Infografiken im Spielverlauf vermittelt. Begleitet wird die Mission von einer Rahmenerzählung mittels kurzer Videosequenzen über eine betroffene Schildkrötenfamilie, deren Schicksal durch die spielende Person bestimmt wird und diese emotional involvieren soll.

Bei der Umsetzung der Spielidee mittels der Game Engine GameSalad wurde deutlich, wie umfangreich und vielschichtig die praktische Ausgestaltung eines digitalen SG ist; nicht nur, was die visuelle Animationen und die auditive Gestaltung angeht, sondern vor allem auch hinsichtlich der Entwicklung einer funktionierenden Spielmechanik und eines ausgefeilten Konzepts, das Wissensvermittlung und Spielspaß in einer Weise kombiniert, die die Zielgruppe am Ball bleiben lässt, sie zu weiterem Spielen und Lernen motiviert und darüber hinaus auch emotional so berührt, dass sie dadurch zum Nach- bzw. Umdenken bewegt wird. Da eine Evaluation des SG *SeaSaver* mit Kindern im Alter der Zielgruppe noch nicht durchgeführt wurde, bleibt es spekulativ, ob *SeaSaver - Rette die Weltmeere* seinem Anspruch als wirkungsvolles Instrument zur nachhaltigen Bewusstseinsbildung in Sachen Plastikmüllverschmutzung gerecht werden kann.

Im nächsten Schritt soll der vorhandene Prototyp zu einem Produkt weiterentwickelt werden. Nach mehreren Test- und Optimierungszyklen könnte das Spiel anschließend als Smartphone-App passend zur präferierten Zielgruppe veröffentlicht werden. Damit das Spiel Nutzer:innen findet, wäre es denkbar, an Schulen Informationstage anzubieten, an denen über Plastikmüllverschmutzung aufgeklärt wird. In diesem Kontext wäre es möglich, das Spiel als Teil des Unterrichts zu bewerben, um möglichst viele Kinder zu erreichen. Eine andere Möglichkeit, den Bekanntheitsgrad des Spiels zu erhöhen, wären Kooperationen mit Umwelt- und Meeresschutzverbänden, die auf das Spiel bei ihrer Aufklärungsarbeit verweisen.

Literatur

- [ACB16] Arnold, M.; Carnap, A.; Bormann, I.: Bestandsaufnahme zur Verankerung von Bildung für nachhaltige Entwicklung in Bildungs- und Lehrplänen, https://www.haus-der-kleinen-forscher.de/fileadmin/Redaktion/4_Ueber_Uns/Evaluati on/Abgeschlossene_Studien/170301_BNE_Expertise.pdf, Stand: 01.07.2023, 2016.
- [Be17] Bergmann, M.; Lutz, B.; Tekman, M. B.; Gutow, L.: Citizen scientists reveal: Marine litter pollutes Arctic beaches and affects wild life, *Marine Pollution Bulletin*, 125/1-2, S. 535-540, 2017.
- [BI09] Blumberg, F. C.; Ismaier, S. S.: What Do Children Learn from Playing Digital Games. In (Ritterfeld, U., Cody, M., Vorderer, P. Hrsg.): *Serious Games. Mechanisms and Effects*, Routledge, New York, S. 131-142, 2009.
- [Bo22] Borchert, Y.: Anzahl und Statistik der Kindergärten in Deutschland, <https://www.pflegemarkt.com/2022/10/06/anzahl-und-statistik-der-kindergarten-in-deutschland/>, Stand: 02.07.2023, 2022.
- [Bu23] Bundesverband der Natur- und Waldkindergärten in Deutschland e.V.: Über uns., <https://www.bvnw.de/ueber-uns>, Stand: 02.07.2023, 2023.
- [COP18] Carbery, M.; O'Connor, W.; Palanisami, T.: Trophic transfer of microplastics and mixed contaminants in the marine food web and implications for human health. *Environment International*, 115, S. 400-409, 2018.
- [De11] Deterding, S.; Khaled, R.; Nacke, L.; Dixon, D.: Gamification: Toward a definition. In: CHI 2011 Gamification Workshop Proceedings, Vancouver, S. 12-15, 2011.
- [Eu21] EU-Parlament: Plastik im Meer: Fakten, Auswirkungen und neue EU-Regelungen, <https://www.europarl.europa.eu/news/de/headlines/society/20181005STO15110/plastik-im-meer-fakten-auswirkungen-und-neue-eu-regelungen>, Stand: 01.07.2023, 2021.
- [Fe21] Feierabend, S.; Rathgeb, T.; Kheredmad, H.; Glöckler, S.: KIM-Studie 2020 - Kindheit, Internet, Medien: Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger. Medienpädagogischer Forschungsverbund Südwest (mpfs), Stuttgart, 2021.
- [Fo20] Fox J.; McKnight, J.; Sun, Y.; Maung, D.; Crawfis, R.: Using a SG to communicate risk and minimize psychological distance regarding environmental pollution. *Telematics and Informatics*, 46, Art. 101320, 2020. <http://dx.doi.org/10.1016/j.tele.2019.101320>
- [Fr22] Frädrich, S.: Die kognitive Entwicklung. In welchen Stufen verläuft sie?, <https://greater.com/kognitive-entwicklung/>, Stand: 02.07.2023, 2022.
- [Gö20] Göbel, S.: Serious Games. In (Zimmermann, O., Falk, F., Hrsg.): *Handbuch Gameskultur*. Deutscher Kulturrat e.V., Berlin, S. 104-109, 2020.
- [Gr20] Grüling, B.; Saeidy-Nory, S.; Montag, A.; Hamdorf, T.; Puppe, M.: Serious Games – Überblick. In (game-Verband, Hrsg.): *Fokus Serious Games*, Verband der deutschen Games-Branche e.V., Berlin, S. 6-11, 202, <https://www.game.de/wp-content/uploads/2020/07/game-Fokus-Serious-Games.pdf>
- [GKR12] Grunenberg, H.; Küster, K.; Rode, H.: Greenpeace Nachhaltigkeitsbarometer – Was bewegt die Jugend? Zusammenfassung, <https://epub.sub.uni->

hamburg.de/epub/volltexte/2014/34263/pdf/Zusammenfassung_Nachhaltigkeitsbarometer_0.pdf, Stand: 02.07.2023, 2012.

- [HT22] Hengstmann, E.; Tamminga, M.: Plastik in der Umwelt – Wo kommt es her, wo geht es hin und wie wirkt es sich aus?, Springer, Berlin, 2022.
- [Iu22] IUCN: The IUCN Red List of Threatened Species, International Union for Conservation of Nature and Natural Resources, <https://www.iucnredlist.org/>, Stand: 02.07.23., 2022.
- [Jo17] Johnson, D.; Horton, E.; Rory, M.; Foth, M.: Gamification and serious games within the domain of domestic energy consumption: A systematic review. *Renewable and Sustainable Energy Reviews* 73, S. 249-264, 2017.
- [Ko12] Koch, M.: Gamification – Steigerung der Nutzungsmotivation durch Spielkonzepte, <https://www.soziotech.org/gamification-steigerung-der-nutzungsmotivation-durch-spielkonzepte/>, Stand: 01.07.2023, 2012.
- [KH19] Koivisto, J.; Hamari, J.: The rise of motivational information systems: A review of gamification research. *International Journal of Information Management* 45, S. 191–210, 2019.
- [LKN20] Löffström, E.; Klöckner, C.A.; Nesvold, I.H.: Nature In Your Face - Disruptive Climate Change Communication and Eco-visualisation as part of a Garden-Based Learning Approach involving primary school children and teachers in co-creating the future. *Frontiers in Psychology*, 11, 2876, 2020. <https://doi.org/10.3389/fpsyg.2020.568068>
- [Lü23] Lüdeke, C.: Die Klima App – Klima Game, <https://www1.wdr.de/schule/digital/unterrichtsmaterial/klima-game-100.html>, Stand: 30.05.2023, 2023.
- [MA12] Males, J.; Van Aelst, P.: Did the blue planet set the agenda for plastic pollution? An explorative study on the influence of a documentary on the public, media and political agendas. *Environmental Communication* 15/1, S. 40–54, 2012.
- [MSZ22] Miao, H.; Saleh, M.S.M.; Zolkepli, I.A.: Gamification as a Learning Tool for Pro-Environmental Behavior: A Systematic Review. *Malaysian Journal of Social Sciences and Humanities (MJSSH)* 7/12, e001881, 2022.
- [Pa21] Panagiotopoulou, L.; Cía, G.N.; Scurati, G. W.; Etzi, R.; Massetti, G.; Gallace, A.; Ferrise, F.: Design of a SG for Children to Raise Awareness on Plastic Pollution and Promoting Pro-Environmental Behaviors. *Journal of Computing and Information Science in Engineering* 21/6, 064502, 2021. <https://doi.org/10.1115/1.4050291>
- [Po20] Pokorny, S.: Umwelt- und Klimaschutz in den Familien, *Analysen & Argumente* Nr. 395, Konrad Adenauer Stiftung, Berlin.
- [Ra21] Rau, L.: Waldkindergarten - So funktioniert der Naturkindergarten, <https://utopia.de/ratgeber/waldkindergarten-so-funktioniert-der-naturkindergarten/>, Stand: 02.07.2023.
- [RC20] Rossano, V.; Calvano, G.: Promoting Sustainable Behavior Using SGs - SeAdventure for Ocean Literacy. *IEEE Access* 8, S. 196931-196939, 2020.
- [RE13] Reckien, D.; Eisenack, K.: Climate Change Gaming on Board and Screen. *Simulation*

& Gaming 44/2–3, S. 253–271, 2017. <https://doi.org/10.1177/1046878113480867>

- [SAM15] Sandbrook, C.; Adams, W.M.; Monteferri, B.: Digital games and biodiversity conservation. *Conservation Letters* 8/2, S. 118–124, 2015.
- [Sc20] Schipperges, M.: Umwelteinstellungen in Deutschland von 1971 bis 2019 - Zeitreihenanalyse anhand externer Datenquellen. Umweltbundesamt, Dessau-Roßlau, <https://www.umweltbundesamt.de/publikationen/umwelteinstellungen-in-deutschland-von-1971-bis>, Stand: 02.07.2023, 2020.
- [Se17] Seidler, C.: Die Müllkippe im Norden. Plastikmüll in der Arktis Der Spiegel, <https://www.spiegel.de/wissenschaft/natur/plastikmuell-in-der-arktis-die-muellkippe-im-norden-a-1184545.html>, Stand: 02.07.2023.
- [SHS18] Schipperges, M.; Holzhauer, B.; Scholl, G.: Umweltbewusstsein und Umweltverhalten in Deutschland 2016. Vertiefungsstudie: Sozial-ökologischer Wandel – Anschlussfähigkeit und Engagement-Potenziale, Umweltbundesamt, Dessau-Roßlau, <https://www.umweltbundesamt.de/publikationen/umweltbewusstsein-umweltverhalten-in-deutschland-1>, Stand: 02.07.2023, 2018.
- [So19] Souza, G., Mesquita, R., Coders, D., Hub, G. H. C., Lima, P., Guerra, E., Capistrano, I.: Save the Ocean: A Game for Environmental Awareness. In *Proceedings of SBGames Conference 2019*, S. 28-31, 2019.
- [St21] Statista: Zustand der Weltmeere, <https://de.statista.com/statistik/studie/id/111017/dokument/meere-und-ozeane/>, Stand: 01.07.2023, 2021.
- [TKL19] Tabuenca, B.; Kalz, M.; Löhr, A.: Massive open online education for environmental activism: The worldwide problem of marine litter. *Sustainability* 11/10, 2860, 2019.
- [TOC21] De la Torre, R.; Onggo, B.S.; Corlu, C.G.; Nogal, M.; Juan, A.A.: The Role of Simulation and Serious Games in Teaching Concepts on Circular Economy and Sustainable Energy. *Energies* 14/4, 1138, 2021.
- [Tr23] Triboot: Umweltschutz für unsere Ozeane - 100% aufgeklärt, <https://triboot.de/projekt/umweltschutz-rettet-unseren-ozean-lernspiel-app/>, Stand: 02.07.2023.
- [VG23] Vecchio, L.P.; Del Greco, A: Game-Based Solutions on the Plastic Problem: A Systematic Review. *Sustainability*. 15/6, 5558, S. 1-33, 2023.
- [WG22] Williams, H.; Gellrich, A.: Umweltbewusstsein in Deutschland 2020. Ergebnisse einer repräsentativen Bevölkerungsumfrage. Umweltbundesamt, Dessau-Roßlau, <https://www.umweltbundesamt.de/publikationen/umweltbewusstsein-in-deutschland-2020>, Stand: 02.07.2023, 2022.
- [WLW17] Willenbacher, M.; Lepiorz, R.; Wohlgemuth, V.: Serious Games, Umweltbewusstsein und Nachhaltigkeit. In (Eibl, M., Gaedke, M., Hrsg.), *INFORMATIK 2017. Gesellschaft für Informatik*, Bonn, S. 2017-2026, 2017. DOI: 10.18420/in2017_202
- [Ww23] WWF: Das kann kein Meer mehr schlucken: Unsere Ozeane verschlucken im Plastikmüll, <https://www.wwf.de/themen-projekte/plastik/unsere-ozeane-versinken-im-plastikmuell>, Stand: 02.07.2023, 2023.

Konzeption einer webbasierten Anwendung zur Unterstützung und Akzeptanzsteigerung des Photovoltaikausbaus in Betrieben

Lucas Tiedemann¹, Jochen Wittmann² und Michaela Zoll³

Abstract: Der vorliegende Beitrag stellt die Entwicklung eines webbasierten Rechners dar, der zur Entscheidungsunterstützung und Akzeptanzsteigerung für Photovoltaik (PV) und Speichersystemen im gewerblichen Umfeld beitragen soll. Grund hierfür ist das hohe Dachflächenpotenzial bei gewerblichen genutzten Gebäuden sowie der Bedarf an einem schnellen Ausbau erneuerbarer Energien, um die Klimaschutzziele einzuhalten. Die wirtschaftlichen und ökologischen Kriterien sind die wichtigsten Gründe für die Investition in eine PV-Anlage. Für eine detaillierte Prognose der Rentabilität ist das verwendete Betreibermodell des Unternehmens relevant. Die Auswahl eines passenden Modells ist wiederum abhängig von solarer Eigenversorgung. Aufgrund von Unterschieden zwischen den Lastprofilen verschiedener Gewerbetypen bietet der entwickelte PV-Rechner die Option, ein spezielles Lastprofil für die Berechnungen hochzuladen. Anhand dieser Prognose wirtschaftlicher und ökologischer Faktoren soll die Anwendung zur betrieblichen Entscheidungsunterstützung beitragen. Ein schnell erhältlicher Überblick der Ergebnisse kann zusätzlich die Akzeptanzsteigerung für PV ermöglichen. Ergänzende Simulationsrechnungen und weitere Ergebniskennzahlen können zukünftig zur optimierten Prognose relevanter Faktoren einer PV-Anlage beisteuern.

Keywords: Photovoltaik, Gewerbebetriebe, Entscheidungsunterstützungssysteme, Akzeptanzsteigerung, Energiespeichersysteme

1 Einleitung

Für die Energiewende ist ein schneller Ausbau erneuerbare Energien erforderlich. Im Erneuerbare-Energien-Gesetz (EEG) 2023 ist deshalb ein starker Anstieg des jährlichen PV-Zubaus vorgesehen [Wi22]. Um diese Ziele zu erreichen, ist es notwendig möglichst viel des Potenzials für PV-Anlagen zu nutzen. Ein hohes Potenzial für den Zubau besteht unter anderem in gewerblich genutzten Gebäuden oder Flächen.

¹ Hochschule für Technik und Wirtschaft Berlin, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin, lucas.tiedemann@posteo.de

² Hochschule für Technik und Wirtschaft Berlin, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin, wittmann@htw-berlin.de

³ Hochschule für Technik und Wirtschaft Berlin, Fachbereich 1, Wilhelminenhofstraße 75A, 12459 Berlin, michaela.zoll@htw-berlin.de

Um dieses Potenzial erschließen zu können, gilt es Unternehmen die Vorteile aufzuzeigen und mögliche Hemmnisse auszuräumen. Dabei können in der Wirtschaftlichkeit einer PV-Anlage je nach Gewerbetyp unterschiedliche Ergebnisse erzielt werden.

Im Rahmen einer Simulationsstudie zu PV wurde durch die Analyse von gewerblichen Lastprofilen ermittelt, wie sich diese nach Gewerbetypen unterscheiden. Besonders die Variation der solaren Eigenversorgung, welche durch den unterschiedlichen Energieverbrauch im Tagesverlauf begründet ist, beeinflusst die Rentabilität einer PV-Anlage [We21].

Im Rahmen dieses Beitrags wird die Entwicklung einer Anwendung beschrieben, welche durch das Einbeziehen spezifischer Lastprofile eine optimierte Prognose für eine PV-Anlage geben soll.

2 Photovoltaik im Gewerbe

Wie in der Einleitung bereits beschrieben wurde, bestehen optimistische Ziele für den Ausbau der PV in Deutschland. Insgesamt bieten Dächer und Fassaden in Deutschland ein Potenzial von 1000 Gigawatt Peak (GWp). Die Nutzung dieses hohen Potenzials ist allerdings weiterhin begrenzt. Es werden weniger als 10% des Dachflächenpotenzials genutzt [Wi22].

Einen großen Teil des Dachflächenpotenzials machen Gewerbe-Gebäude aus. Dabei besteht die Besonderheit, dass sich die Anforderungen an eine PV-Anlage je nach Betrieb deutlich unterscheiden können. Die Höhe des Eigenverbrauchs beeinflusst z.B. wie rentabel die unterschiedlichen Betreibermodelle für einen bestimmten Betrieb sind [We21].

Mit Hilfe von Lastprofilen können Aussagen über die solare Eigenversorgung der Betriebe getroffen werden. Je nach Gewerbetyp können sich die Verbrauchskurven der Lastprofile deutlich unterscheiden, woraufhin sich Unterschiede in der solaren Eigenversorgung ergeben. Es bestehen innerhalb eines Gewerbetyps teilweise relevante Unterschiede in der Eigenversorgung. Eine Untersuchung von 35 Lebensmittelmärkten hat für diese Branche keine relevanten Unterschiede ermitteln können. In Betrieben der Metall- und Holzverarbeitung sowie in der Abfallentsorgung bestehen auch innerhalb der Branche größere Unterschiede [We21].

2.1 Simulation des Photovoltaik- und Batteriesystems

Die Simulation des PV-Batteriesystems bildet die Datengrundlage der Anwendung. Das Systemverhalten wird dabei mit einer zeitlichen Auflösung von einer Minute über den Zeitraum von einem Jahr simuliert. Das Hauptziel ist dabei die Bestimmung des Autarkiegrads der jeweiligen Systemkonfiguration. Er entspricht dem Anteil des kumulierten Stromverbrauchs, der mithilfe von eigens erzeugtem PV-Strom gedeckt werden kann. Je höher der Autarkiegrad, desto mehr Netzbezugskosten können eingespart werden und desto mehr CO₂-Emissionen können vermieden werden. Die Autarkie stellt deshalb eine entscheidende Kenngröße zur Bewertung des ökonomischen und ökologischen Nutzens des PV-Speichersystems dar.

Das simulierte PV-Speichersystem ist AC-gekoppelt. Die PV-Anlage und die Batterie sind demnach nicht direkt miteinander verbunden, sondern jeweils an einen eigenen Wechselrichter angeschlossen. Für den PV-Wechselrichter werden die Kennwerte des SMA Sunny Tripower 10.0 herangezogen und für den Batteriewechselrichter die des SMA Sunny Island 8.0 [We21].

Die Zeitreihe der PV-Leistung wird in Abhängigkeit der Umwelteinflüsse Bestrahlungsstärke und Temperatur sowie der Gerätedaten des PV-Generators bestimmt. Die Wetterdaten zur globalen Direkt- und Diffusstrahlung und der Umgebungstemperatur stammen vom meteorologischen Observatorium Lindenberg (Brandenburg) aus dem Jahr 2017 und liegen in einminütiger Auflösung vor. Die Bestrahlungsstärke auf den PV-Generator, der als Abstraktion wie eine geneigte Ebene behandelt wird, wird anhand der jeweiligen Ausrichtung (Azimutwinkel der Ebene α_E) und Neigung (Höhenwinkel der Ebene γ_E) berechnet. Dabei wird auch die Sonnenhöhe abhängig von Datum und Uhrzeit berücksichtigt. Diese wurde ebenfalls an den Standort Lindenberg (52° 20' 83" N, 14° 11' 78" O) angepasst. Der diffuse Strahlungsanteil wird nach dem Modell von Klucher berechnet. Für den PV-Generator werden die Kennwerte des LG NeON R genutzt [We21]. Sein Wirkungsgrad wird zu jedem Zeitschritt in Abhängigkeit von der Umgebungs- und Betriebstemperatur bestimmt [Be04; Qu19].

Die Simulation des Speichersystems wird nach dem Modell „simbat“ der Forschungsgruppe Solarspeichersysteme durchgeführt [WO23]. So werden unter Berücksichtigung der Batterieverluste für jeden Zeitschritt die Lade- und Entladeleistung sowie der Ladezustand (State of Charge) bestimmt. Die verwendeten Umwandlungswirkungsgrade für Speicherung, Laden und Entladen entsprechen dem System Tesvolt TS 48 V, die nominale AC-Leistungsaufnahme sowie -abgabe wird auf 0,35 kW/kWh Speicherkapazität festgelegt [We21].

Für einige ausgewählte Gewerbe sind real gemessene, normierte Lastprofile in 15-minütiger Auflösung hinterlegt. Diese werden entsprechend des angegebenen Jahresstromverbrauchs für die Simulation skaliert. Des Weiteren besteht die Möglichkeit eigene Lastprofile hochzuladen, um eine individuelle Systembewertung zu erzielen.

2.2 Photovoltaik Einspeisevergütung

Die Einspeisevergütung einer PV-Anlage ist von mehreren Faktoren abhängig. Grundsätzlich ist zunächst das gewählte Betreibermodell ausschlaggebend. Für gewerbliche Betriebe in Deutschland gibt es sieben unterschiedliche Modelle, welche sich zunächst nach den Eigentumsverhältnissen unterscheiden. Es gibt somit Modelle für Gebäudebesitzer und für Gebäudemietler [Uh21]. Dieser Beitrag legt den Fokus auf den Betrieb als Gebäudebesitzer, weshalb Mieter-Betreiber-Modelle nicht genauer betrachtet werden. Zudem werden die Modelle Anlagenmiete, Contracting, Direktstromlieferung und Verpachtung der Dachfläche aus der Betrachtung ausgeschlossen. Dieser Ausschluss erfolgt, da bei diesen Modellen die finanzielle Vergütung von Vertragspartnern abhängig ist und somit nicht ohne weiteres zur Verfügung steht. Ein finanzieller Vergleich mit den anderen Modellen ist deshalb nicht möglich.

Zunächst erfolgt die Betrachtung des Eigenverbrauchs-Betreibermodell. Bei diesem Modell wird der erzeugte Strom direkt vor Ort verbraucht. Im EEG wird Eigenversorgung als der Verbrauch von Strom unmittelbar vor Ort durch den Betreiber der Anlage definiert. Dabei darf der Strom nicht durch ein Netz durchgeleitet werden (§3 Nr. 19 EEG 2021). Dieses Modell wird häufig mit der Einspeisung nach EEG verbunden. In dem Fall handelt es sich um eine Überschusseinspeisung. In der folgenden Abbildung ist das Zusammenspiel zwischen dem Eigenverbrauch, Netzbezug und der Überschusseinspeisung dargestellt.

Würde auf den Eigenverbrauchsanteil verzichtet werden, kann der gesamte produzierte Strom ins Netz eingespeist und nach EEG vergütet werden. Die Volleinspeisung wird höher vergütet als die Überschusseinspeisung. Bei einer Anlage wie aus dem Beispiel würde die Vergütung 10,90 Cent anstatt 5,80 Cent betragen.

Im Gegensatz zur Einspeisung nach EEG erfolgt beim Betreibermodell der Direktvermarktung der Verkauf zum Marktpreis an der Börse. Liegt der Erlös unter der EEG-Vergütung wird die Differenz mittels einer Marktprämie ausgeglichen.

Die Höhe der Einspeisevergütungen ist im EEG verankert. Die Vergütungshöhe wird durch verschiedene Faktoren bestimmt, darunter die installierte Leistung der Anlage, die Art der Einspeisung (Voll- oder Überschusseinspeisung) sowie die Wahl zwischen dem Marktprämienmodell und der Festvergütung.

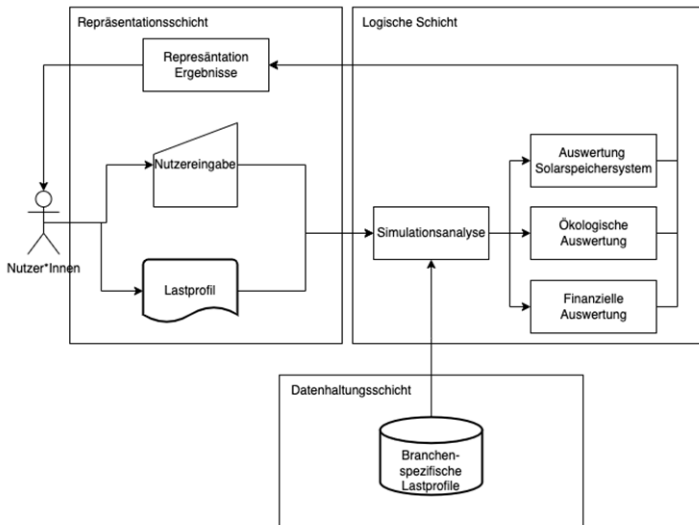


Abb. 1: Beispielrechnung Betreibermodell Eigenverbrauch

3 Methodik

Im Rahmen dieses Beitrags wird die Problemstellung verfolgt, wie die Akzeptanz für PV in Gewerbebetrieben durch eine Web-Applikation zur Entscheidungsunterstützung gesteigert werden kann. Als zweiten Punkt der Problemstellung wird untersucht, welchen Beitrag Lastprofile in den Simulationsanalysen zur solaren Eigenversorgung spielen und inwiefern die Web-Applikation durch die Option zum Hochladen von spezifischen Lastprofilen profitieren kann. Die Anwendungsumgebung begrenzt sich dabei auf gewerbliche Betriebe innerhalb Deutschlands, welche Eigentümer des betrachteten Gebäudes sind.

Um die Relevanz einer Applikation zur Entscheidungsunterstützung von Betrieben in Form eines PV-Rechners besser abschätzen zu können, wurde eine Literaturrecherche zu PV im Gewerbe durchgeführt. Der Fokus lag dabei auf den Entscheidungskriterien, Vorteilen und Hemmnissen.

Zusätzlich zu den Akzeptanzkriterien wurde sich mit der Fragestellung beschäftigt, welche Besonderheiten bei der Betrachtung von PV im Gewerbe bestehen. Ein besonderer Fokus lag hierbei auf den Einfluss des Eigenverbrauchs auf die Rentabilität und wie Lastprofile in der Analyse die Prognosen zu PV-Anlagen beeinflussen.

Neben der Literaturrecherche erfolgte eine Analyse bestehender PV-Rechner. Diese soll einen Überblick über den Funktionsumfang bestehender Lösungen geben. Dafür wurden

Input- und Output Parameter erhoben, welche ebenfalls zur Aussage der Komplexität der Anwendungen herangezogen wurden. Für die Analyse wurden die Rechner pv@now easy, Mysolvation (PV-Check), Meyerburger, Fronius und Varta betrachtet [Fr23; MB23; PN23; Va23]. Neben den reinen Vergleich der Parameter erfolgte eine subjektive Bewertung in den Kategorien Bedienbarkeit und übersichtliche Ergebnisdarstellung. Die Analyse der Rechner hat ergeben, dass die Option des Hochladens eines spezifischen Lastprofils bei den untersuchten Optionen nicht vertreten ist.

4 Web-Applikation

Die Anwendung ist in eine dreischichtige Architektur aufgeteilt, welche in der folgenden Abbildung schematisch dargestellt ist. Da keine Eingaben der Nutzer*innen persistiert werden müssen, werden nur die branchengemittelten Lastprofile in der Datenhaltungsschicht vorgehalten. Innerhalb der logischen Schicht erfolgt zum einen die Simulationsanalyse, welche auf Basis der Eingangsdaten mit Hilfe unterschiedlicher Python-Skripte Ergebniswerte berechnet, welche die Basis für die drei Auswertungsklassen bilden. Innerhalb dieser werden die Werte zur Ergebnisdarstellung an die Nutzer*innen berechnet und daraufhin in der Repräsentationsschicht visualisiert.

Das User Interface (UI) der Web-Applikation ist in einen Teil für die Benutzereingabe und die Ergebnisdarstellung aufgeteilt. Im Frontend wurde React eingesetzt und die meisten der UI-Elemente basieren auf Komponenten der Material UI (MUI) Bibliothek.

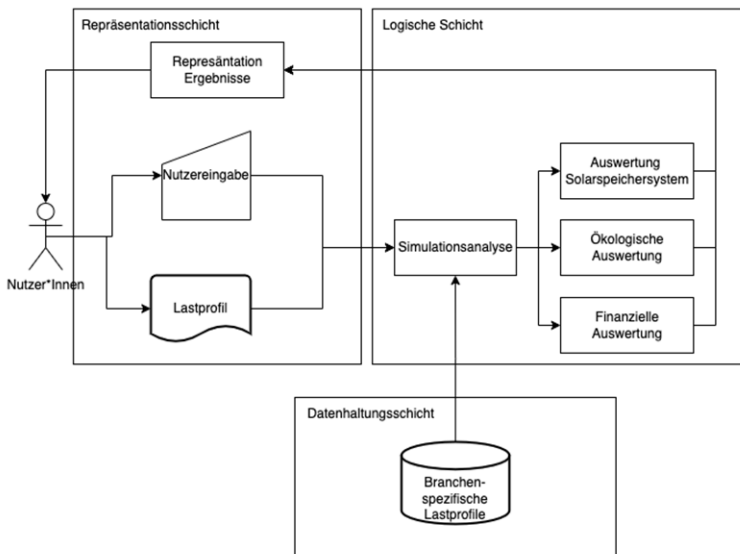


Abb. 2: Schema der Anwendung

The screenshot shows a user interface with three main sections:

- Lastprofil auswählen (Select Load Profile):** A grid of buttons with icons for 'Eigenes Lastprofil' (Own Load Profile), 'Supermarkt' (Supermarket), 'Büro' (Office), 'Landwirtschaft' (Agriculture), 'Metall' (Metal), 'Schule' (School), and 'Sonstige Gewerbe' (Other Commercial).
- Dachfläche bestimmen (Determine Roof Area):** A compass-like slider for 'Himmelsrichtung' (Orientation) with markers for Nord, West, Süd, Ost, and Nord. A 'Süd' button is also present. Below are two numeric input fields: 'Dachfläche in m²' (400) and 'Dachneigung in °' (45), each with up/down arrow controls.
- Wirtschaftliche Eingaben (Economic Inputs):** Two numeric input fields: 'Strompreis-Szenario' (Hochpreis) and 'Laufzeit in Jahren' (20), each with up/down arrow controls.

Abb. 3: Screenshot Ausschnitt Benutzereingabe

Für die Ergebnisdarstellung werden, ebenfalls die bereits erwähnten Technologien eingesetzt. Um die Ergebnisse zu visualisieren und für Nutzer*innen durch eine interaktive Bedienung näher zu bringen wird auf die Nivo-Bibliothek zurückgegriffen, welche sich gut in React-Komponenten integrieren lässt. Die Möglichkeit mit der visuellen Darstellung der Daten zu interagieren, soll Nutzer*innen eine gute User Experience (UX) gewährleisten und zudem für ein besseres Verständnis der Informationen sorgen (Sacket et al. 2016). Die Ergebnisdarstellung besteht aus zwei Arten von wiederverwendbaren Komponenten. In einer Liste werden die Ergebniswerte mit einer Miniatur-Ansicht des jeweiligen Diagramms angezeigt, um die Ergebnisse auf einen Blick erfassbar zu machen. Die detaillierte Ansicht ermöglicht einen genaueren Blick auf die jeweiligen Ergebniskategorien. Mit Hilfe von Tooltips werden hier den Nutzer*innen hilfreiche Informationen zur Verfügung gestellt. Die Abbildung 4 zeigt, wie diese Funktionen im UI der Anwendung umgesetzt wurden.

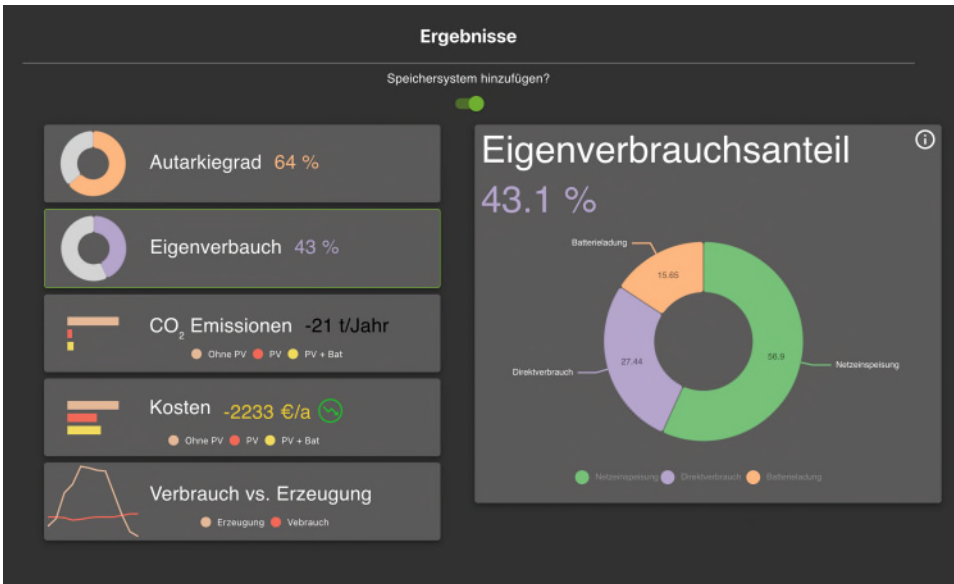


Abb. 4: Screenshot Ergebnisdarstellung der Anwendung

5 Ergebnisse

Die Recherche zu PV im Gewerbe hat gezeigt, dass zur Akzeptanzsteigerung für PV im Gewerbe entsprechende Vorteile aufgezeigt werden müssen. Dem finanziellen Vorteil durch PV wird dabei die größte Bedeutung für Unternehmen zugeschrieben. Ebenfalls sind die Punkte *Unternehmens- bzw. Umweltziele umsetzen*, *Marketing* und *Wettbewerbsfähigkeit sichern* mit hoher Wichtigkeit bewertet worden [Uh21].

Mit Hilfe von PV-Rechnern lassen sich schnell und effizient die Vorteile von PV im Kontext eines konkreten Beispiels berechnen. Allerdings wird bei bestehenden Rechnern auf gemittelte Lastprofile zurückgegriffen. Daraus ergibt sich für den zu entwickelnden Prototypen die Lücke innerhalb der betrachteten Anwendungsumgebung. Durch spezifische oder branchenspezifische Lastprofile sollen genauere Ergebnisse erzielt werden. Denn im Vergleich zu Wohngebäuden ergeben sich in gewerblichen Betrieben je nach Typ unterschiedliche Verbrauchsprofile. Durch diese im Tagesverlauf unterschiedlichen Lastspitzen entstehen Unterschiede in der solaren Eigenversorgung [We21]. Da die solare Eigenversorgung wiederum ein relevantes Entscheidungskriterium für die Auswahl des Betreibermodells darstellt und somit direkten Einfluss auf die Wirtschaftlichkeit einer PV-Anlage hat, soll durch diese Web-Applikation ein möglichst genau zum betrachteten Betrieb passendes Lastprofil in die Simulationsanalyse gegeben werden [Uh21].

Der Vergleich der Simulationsrechnungen von einem branchenspezifischen und einem gemittelten gewerblichen Lastprofil zeigt, dass bei ansonsten gleichen Faktoren ein Unterschied von 10% im Eigenverbrauch besteht. Somit ist in der Beispielrechnung bei der gleichen PV-Anlage abhängig vom Lastprofil jeweils ein anderes Betreibermodell lukrativer. Durch die detaillierte Analyse können somit höhere finanzielle Vorteile innerhalb des PV-Rechners gezeigt werden. Da die finanziellen Vorteile für Unternehmen das wichtigste Entscheidungskriterium darstellen, kann ein höherer Ertrag beim Entscheidungsprozess relevant sein und bei der Akzeptanzsteigerung helfen.

6 Diskussion

Dieser Beitrag hatte zum Ziel, eine Anwendung zur Entscheidungsunterstützung für die Anschaffung von PV-Anlagen im gewerblichen Umfeld zu entwickeln. Dies sollte den Nutzer*innen ohne vorherige Recherchen und Fachkenntnisse einen Überblick zu den wichtigsten Fragestellungen der Investitionsentscheidung geben.

6.1 Theorie- und Praxisbeitrag

Auf Grundlage bestehender Literatur zeigt sich die Relevanz der fokussierten Betrachtung von PV-Ausbau im Gewerbe [Uh21; We21].

In diesem Zusammenhang wurde der ermittelte Einfluss unterschiedlicher Lastprofile auf die solare Eigenversorgung im Rahmen der speziellen Vorteile und Anforderungen von PV im Gewerbe betrachtet [Uh21; We21]. Auf dieser Basis an Informationen haben weitere Recherchen zu Fragen des Einflusses der aktuellen EEG-Novelle auf die Betreibermodelle für PV-Anlagen die Untersuchungsergebnisse weiter vertieft. Dieser Untersuchungsrahmen hat in der Arbeit somit einen Beitrag zur Theorie beitragen können.

Die Untersuchungsergebnisse bilden die Grundlage für die entwickelte Anwendung. Da dieses in einer finalen Version der Praxis kostenfrei zur Verfügung gestellt werden soll, stellt es eine Brücke zwischen dem Theorie- und dem Praxisbeitrag dar.

6.2 Limitation

Für eine einfache Bedienbarkeit der Anwendung wurden im Voraus einige Annahmen getroffen. Bei diesen Annahmen handelt es sich hauptsächlich um die Konfiguration des PV-Speichersystems, diese wurden bereits im Kapitel 2.1 beschrieben. Dadurch ergeben sich Verluste in der Genauigkeit der berechneten Vorhersagen der Simulationsanalyse. Diese Limitation ist aber bewusst getroffen worden, da anderenfalls ein Expert*innen-Tool entstanden wäre.

Da die ökonomische Auswertung lediglich auf einer Prognose für die Entwicklung des Strompreises beruht, ergeben sich weitere Unsicherheiten in der Genauigkeit der berechneten Ergebnisse. In der Anwendung wird für die Berechnung der wirtschaftlichen Ergebnisse aktuell nur das Betreibermodell Eigenverbrauch mit Überschusseinspeisung nach EEG berücksichtigt. Falls das Betrachtungsobjekt des PV-Rechners einen sehr geringen Eigenverbrauch hat, kann dies zu einer wirtschaftlich schlechteren Einschätzung führen. In diesem Fall würde bspw. die Volleinspeisung höhere Prämien durch das EEG ergeben.

6.3 Zukünftige Forschung

Die Weiterentwicklung der Web-Applikation in Verbindung mit zusätzlicher Forschung zum Thema PV im Gewerbe ist sinnvoll. Schwerpunkte könnten dabei das Betreibermodell und die Lastspitzenkappung sein. Um die Relevanz der Wahl des Betreibermodells in der Anwendung zu berücksichtigen, sollte ein Modul implementiert werden, welches den Nutzer*innen auf Basis der Daten des Untersuchungsobjektes eine Empfehlung für die Wahl des Betreibermodells gibt.

Eine zusätzliche Erweiterung könnte sich auf die Lastspitzenkappung konzentrieren, um die optimale Kapazität eines Speichersystems zu bestimmen. Durch die Optimierung der gewählten Kapazität ist ggf. eine Verbesserung der Rendite möglich [Ch16].

Zur Validierung der Applikation sollte diese durch Personen aus dem betrachteten Anwendungsgebiet getestet werden. Dadurch könnte eine bessere Einschätzung des Mehrwerts sowie Probleme und fehlende Funktionen ermittelt werden.

Neben der Weiterentwicklung der Applikation hat dieser Beitrag aufgezeigt, dass der Themenbereich zu PV im Gewerbe in der Forschung intensiver untersucht werden sollte. Zukünftige Studien könnten so einen umfassenderen Blick auf PV im Gewerbe werfen und die Themen Wirtschaftlichkeit, Akzeptanzkriterien und die technische Betrachtung von PV verbinden.

7 Fazit

In dieser Arbeit wurde untersucht, wie die Akzeptanz von PV-Anlagen im gewerblichen Bereich gesteigert werden kann. Dabei wurden ökonomische und ökologische Kriterien erarbeitet. Die Entwicklung einer Web-Applikation, welche die wichtigsten Kriterien mit Hilfe einer Simulationsanalyse prognostiziert, unterstützt Unternehmen bei Investitionsentscheidungen für PV-Anlagen.

Eine Weiterentwicklung des Prototypens könnte durch die Ergänzung weiterer Betreibermodelle eine ausführlichere Analyse der Wirtschaftlichkeit von PV-Anlagen

ermöglichen. Dadurch würden auch die präziseren Prognoseergebnisse durch die Integration spezifischer Lastprofile nutzbar gemacht werden.

Zusammenfassend bietet die Arbeit einen Ansatz zur Förderung von PV im Gewerbe und bietet mit den entwickelten Prototypen die Grundlage für eine Applikation, welche eine Lücke innerhalb der untersuchten Anwendungsumgebung schließt.

Literaturverzeichnis

- [Be04] Beyer H. G. et al.: Identification of a General Model for the MPP Performance of PV-Modules for the Application in a Procedure for the Performance Check of Grid Connected Systems. In 19th European Photovoltaic Solar Energy Conference, Paris, S. 3073–3076, 2004.
- [Ch16] Chua, K. H. et al.: Energy storage system for peak shaving. International Journal of Energy Sector Management, S. 3–18, 2016
- [Fr23] Fronius Solar Creator, creator.fronius.com, 20.03.2023
- [MB23] Meyer Burger, meyerburger.com/de/solar modul/solarrechner, 10.04.2023
- [PN23] PV-Now-Easy, pv-now-easy.de, 12.03.2023
- [Qu19] Quaschnig, V. Regenerative Energiesysteme. Technologie – Berechnung – Klimaschutz. 10., aktualisierte und erweiterte Auflage, München: Hanser, 2019.
- [Sa16] Saket, B et al.: Beyond usability and performance: A review of user experience-focused evaluations in Visualization. In ACM International Conference Proceeding Series, S. 133–142, 2016.
- [Uh21] Uhland, T. et al.: Photovoltaik in Gewerbe und Industrie – Solarenergieerfolgreich einsetzen, Solar Cluster Baden-Württemberg e.V., 2021
- [Va23] Varta Energiespeicher Berechnungstool, varta-ag.com/de/konsument/produktkategorien/energiespeicher/berechnungstool, 20.03.2023
- [We21] Weniger, J. et al.: Photovoltaik und Stromspeicher im Gewerbe. In PV-Symposium, 2021
- [Wi22] Wirth, H.: Aktuelle Fakten zur Photovoltaik in Deutschland, 2022
- [WO23] Weniger J. und Orth, N.: 'simbat: simple PV-battery system model', Hochschule für Technik und Wirtschaft HTW Berlin, geplant 2023. solar.htw-berlin.de

Analyse und Entwicklung eines Prototyps zur Optimierung des Verbrauchs von Ökostrom in Privathaushalten

Malina Bergmann¹, Sebastian Rauch¹, Josephine Brömme, Jochen Wittmann²

Abstract: Die gesamte Stromnachfrage der BRD kann bisher noch nicht mit regenerativem Strom abgedeckt werden. Um privaten Nutzern die Möglichkeit zu geben, ihren Strombedarf zeitlich an das Angebot von Ökostrom anpassen zu können, werden zunächst frei verfügbare Daten für die gesamte Netzlast und das Angebot an erneuerbarem Strom analysiert. Darauf aufbauend wird eine einfache Prognose erstellt, in welchen Zeitfenstern der Anteil von Ökostrom im Gesamtnetz besonders hoch ist. Private Nutzer können mit dieser Info den Anteil von Ökostrom an ihrem persönlichen Stromverbrauch optimieren.

Keywords: Stromnetz, Ökostrom, intelligente Stromverbraucher.

1 Einleitung

Der Bruttostromverbrauch in Deutschland hat im Jahr 2007 mit 624 TWh seinen bisherigen Höhepunkt erreicht. Obgleich in den vergangenen fünf Jahren: 2018 (592 TWh), 2019 (575 TWh), 2020 (555 TWh), 2021 (570 TWh) [Um22] und 2022 (547 TWh) [St22] eine abnehmende Tendenz im Stromverbrauch zu verzeichnen ist, widersprechen mittelfristige Prognosen diesem aktuellen Trend. In einer vom Bundesministerium für Wirtschaft und Energie beauftragten Studie, aus dem Jahr 2021, wird von einer Steigerung um elf % auf 658 TWh im Jahr 2030 ausgegangen [Ke21]. In einer langfristigen Studie, mit dem Titel „Energieziel 2050: 100 % Strom aus erneuerbaren Quellen“, die in Zusammenarbeit von Umweltbundesamt und Fraunhofer-Institut für Windenergie und Energiesystemtechnik entstand, wird von einem Bruttostromverbrauch von 506 TWh im Jahr 2050 ausgegangen [Um10].

Unabhängig davon, welche dieser Szenarien tatsächlich eintreten, hat der Stromverbrauch in Deutschland einen wesentlichen Anteil an den, den Klimawandel weiter vorantreibenden, Emissionen.

¹ HTW Berlin, Fachbereich 2, Wilhelminenhofstr. 75A, 12459 Berlin, Deutschland,
{ Malina.Bergmann | Sebastian.Rauch | Josephine.Broemme }@Student.HTW-Berlin.de

² HTW Berlin, Fachbereich 2, Wilhelminenhofstr. 75A, 12459 Berlin, Deutschland,
Jochen.Wittmann@HTW-Berlin.de

Bei dem in CO₂-Äquivalenten angegebenen Emissionsfaktor des deutschen Strommix ist ein negativer Trend zu beobachten. Durch eine steigende Nutzung von dem aus Kohle erzeugten Strom ist jener pro Kilowattstunde von 438 g CO₂-Äquivalente 2020 auf 485 g 2021 gestiegen [Um22a]. Eine klimafreundliche und nachhaltige Kehrtwende und somit massive Reduktion der Emissionen im Energiesektor, ist nur durch das Senken des Stromverbrauchs in allen Sektoren zu bewältigen. Dies erfordert den vollständigen Ausbau erneuerbarer Energien.

Das Angebot an regenerativen Energien wächst kontinuierlich. Nicht zuletzt durch das Erneuerbare-Energien-Gesetz und ein steigendes Bewusstsein in großen Teilen der Bevölkerung. Betrug der Anteil erneuerbarer Energien am gesamten Bruttostromverbrauch Deutschlands im Jahr 2000 lediglich 6,3 %, so waren es im Jahr 2020 bereits 45,2 % [Um22a]. Für das vergangene Jahr 2022 wurde der Stromverbrauch bereits zu 48,4 % durch erneuerbare Energien gedeckt [In22]. Dies bedeutet jedoch auch, dass noch über die Hälfte der Netzlast durch die Stromerzeugung mit konventionellen Energieträgern gedeckt werden muss. Welcher Strommix dann bei Nutzenden ankommt und aus dem Netz entnommen wird, hängt, ganz unabhängig davon, ob ein Ökostromtarif vorliegt oder nicht, von dem Angebot der erneuerbaren Energien im Netz ab und kann von Endverbrauchenden derzeit nicht transparent eingesehen werden. Dieses Paper und der zugehörige Prototyp der Applikation zeigen einen Ansatz auf, zurückliegende Verbrauchsdaten und Prognosen zur Stromnachfrage und den Anteil an Ökostrom im aktuellen Strommix so aufzuarbeiten, dass daraus individuelle Empfehlungen zum bestmöglichen Ökostromverbrauch ausgegeben werden können. Die Ableitung dieser Prognosen basiert sowohl auf einem möglichst hohen Anteil erneuerbarer Energien am Strommix als auch auf einer möglichst geringen Auslastung des Stromnetzes. In den folgenden Kapiteln werden die Herangehensweise an diese Problematik, als auch die angewandten Modellierungen näher beschrieben, gefolgt von der Beschreibung der prototypischen Umsetzung einer Webapplikation und einem Ausblick hinsichtlich Verbesserungspotentialen der Datenmodellierung und Erweiterungsmöglichkeiten der Applikation.

2 Zielsetzung: Vorhersage umweltfreundlicher Stromverbrauchszeiten

Aufgrund der großen Schwankungen von Stromangebot und -nachfrage und den diversen Einflussfaktoren kommt es oft dazu, dass sich das Angebot an Strom aus erneuerbaren Energien und die Nachfrage nicht decken. So reicht der Strom aus erneuerbaren Energien oft nicht aus, um die Stromnachfrage zu Zeiten der Nachfragespitzen zu befriedigen. Gleichermäßen ist oft zu viel Strom aus erneuerbaren Energien vorhanden, wenn er gar nicht verwendet werden kann.

Ziel dieser Arbeit ist daher, diese Leistungsspitzen mit einhergehendem geringem Angebot zu ermitteln. Die rechnerische Vorhersage von Angebot und Nachfrage wird zur

Entwicklung einer Empfehlung genutzt, zu welchen Zeitpunkten in den nächsten 72 Stunden es im Hinblick auf den Umwelt- und Klimaschutz am sinnvollsten ist, Strom zu verbrauchen. Um die Prognosen möglichst plausibel zu modellieren, werden zunächst vorliegende Angebots- und Nachfragedaten untersucht, bevor das Konzept für den Aufbau der Webapplikation vorgestellt wird.

3 Datenanalyse: Struktur des Angebots und der Nachfrage von Strom in Deutschland

Um Konzepte für die Modellierung von Nachfrage und Angebot entwickeln zu können, ist es sinnvoll, zunächst die aktuelle Strommarktsituation in Deutschland zu untersuchen. Die gesamte Stromnachfrage kann bisher noch nicht mit regenerativem Strom abgedeckt werden. Die Struktur und Analyse der verwendeten Daten für die gesamte Netzlast und das Angebot an Strom aus erneuerbaren Energien wird in den folgenden beiden Kapiteln näher beschrieben.

3.1 Struktur der Gesamtnachfrage von Strom in Deutschland

Um eine Prognose für die künftige Stromnachfrage modellieren zu können, war zunächst eine Datengrundlage gemessener historischer Nachfragewerte nötig. Die Datenquelle für die Netzlast, also den gesamten Stromverbrauch in Deutschland, stellt die Plattform SMARD, Strommarktdaten, der Bundesnetzagentur dar [SM22]. Zunächst wurden die Daten, gefiltert nach Stromverbrauch, realisierter Stromverbrauch und Region Deutschland, als csv Daten für die Analyse heruntergeladen. Die Daten in Originalauflösung bestehen aus sechs Datenspalten: Datum, Anfang, Ende, Gesamt (Netzlast)[MWh], Residuallast [MWh] und Pumpspeicher [MWh]. Die Daten können dabei noch nach Energieträgern gefiltert werden, für die Verbrauchsanalyse wurde jedoch der kumulierte Verbrauch benötigt und verwendet. Anfang und Ende der Daten beschreiben dabei jeweils einen Zeitraum von 15 Minuten. Für die Datenanalyse wurde die Gesamt Netzlast in MWh untersucht, dabei wurde die Residuallast (Netzlast ohne Wind- und Solarstrom) und Pumpspeicher (Energiespeicher unter Nutzung der potenziellen Energie von Wasser) außer Betracht gelassen, da für die Verbrauchsanalyse der tatsächlich verbrauchte Strom untersucht wurde.

Im nächsten Schritt der Datenanalyse wurden die verfügbaren Datensätze zunächst über Jupyter Notebooks formatiert und geplottet. Diese Analysen haben mit Daten aus den Jahren 2016 bis 2022 stattgefunden und werden im Folgenden anhand von Diagrammen für das Jahr 2022 kurz erläutert.

In Abbildung 1 wird der Stromverbrauch im Jahresverlauf dargestellt und, wie auch in anderen untersuchten Jahren, wird deutlich, dass es im Verlauf eines Jahres Schwankungen gibt. Diese jahreszeitlichen Schwankungen sind besonders deutlich zu

erkennen, wenn man die Sommermonate Juli bis September mit durchschnittlich 65.000 MWh und die Wintermonate Januar bis März, als auch Dezember, mit durchschnittlich 75.000 MWh, schwankend zwischen 70.000 MWh und knapp 80.000 MWh, miteinander vergleicht.

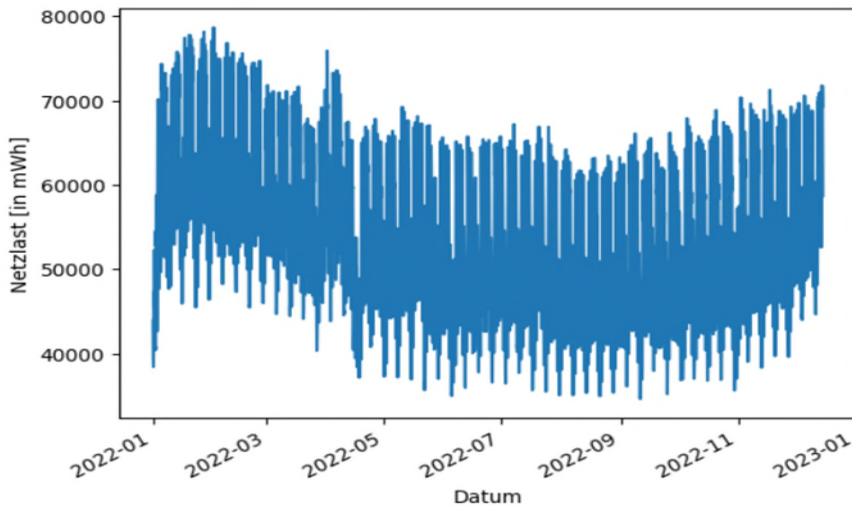


Abbildung 1 - Diagramm Netzlaster Jahresverlauf (2022) in MWh

Die Stromverbräuche schwanken außerdem auch im Verlauf eines Tages, dargestellt in Abbildung 2. Diese Woche zeigt beispielhaft an der ersten August Woche 2022 die Schwankungen im Tagesverlauf, beispielsweise zwischen 10 und 14 Uhr dienstags und mittwochs im vergangenen Jahr, mit Spitzen von bis zu 65.000 MWh. Zwischen 0 und 5 Uhr ist der Verbrauch hingegen an allen Wochentagen vergleichsweise niedrig und fängt erst gegen 5 Uhr am Morgen an, 45000 MWh zu überschreiten. Ausnahmen hierzu stellen Samstag und Sonntag dar, hier steigen die Verbräuche erst ab 7 Uhr morgens. An allen Wochentagen nimmt der Stromverbrauch ab 20 Uhr kontinuierlich ab. Während an Sonntagen am wenigsten Strom verbraucht wird, wird an Samstagen etwas mehr Strom verbraucht als an Sonnabenden, jedoch erheblich weniger als an Wochentagen.

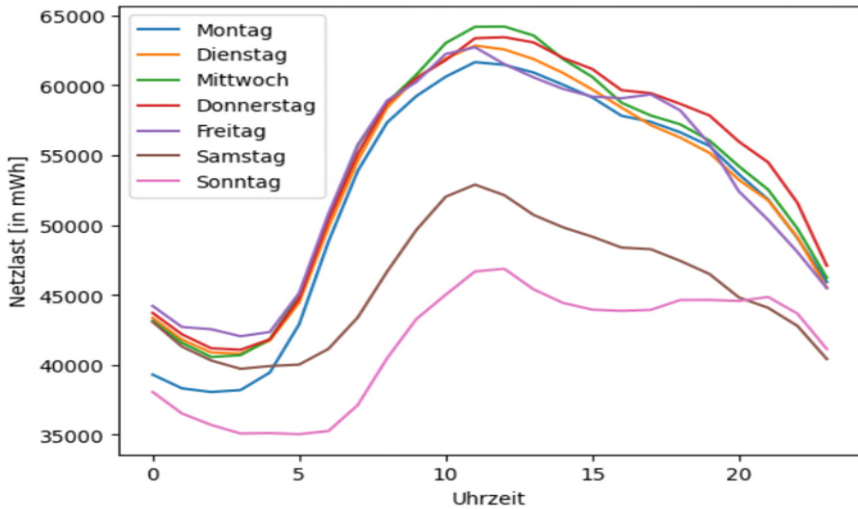


Abbildung 2 - Stromverbrauch in MWh - Schwankungen im Tagesverlauf (1.8.-7.8.22)

Neben den Schwankungen im Jahres- und Tagesverlauf, zeigte die Analyse auch starke Schwankungen in der Stromnachfrage an Feiertagen auf, dargestellt in Abbildung 3. So wird an bundesweit einheitlichen Feiertagen, beispielsweise Neujahr, Ostern oder Pfingsten, weniger Strom verbraucht. Im Gegensatz dazu stehen andere, lokale Feiertage, an denen der Stromverbrauch im Bereich eines normalen Wochentages liegt und die daher auch als solche betrachtet werden.

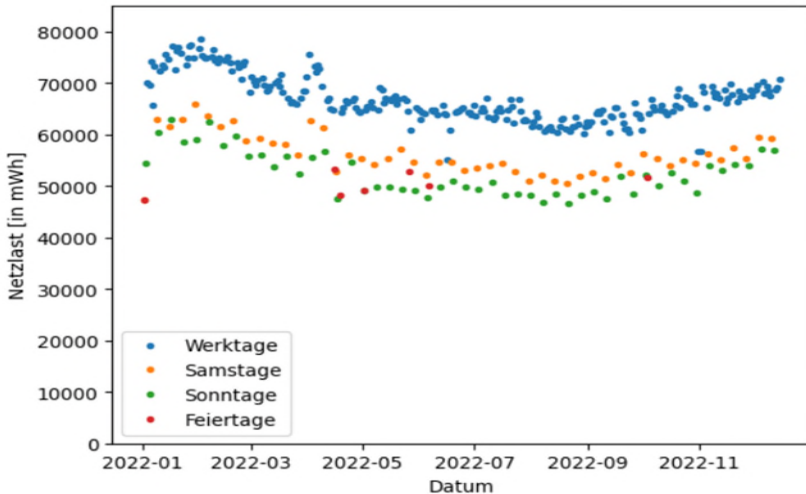


Abbildung 3 - Stromverbrauch in MWh nach Tagesart

Die Unterschiede in den Stromverbräuchen an verschiedenen Wochentagen und Feiertagen führten zu der Unterscheidung nach Tagesart, dabei wurden Werktage, Samstage, Sonntage und Feiertage als Kategorien festgelegt. Zu Feiertagen werden dabei nur die bundesweiten gesetzlichen Feiertage Neujahr, Karfreitag, Ostermontag, Tag der Arbeit, Christi Himmelfahrt, Pfingstmontag, Tag der deutschen Einheit, Erster- und zweiter Weihnachtsfeiertag gezählt. Die auf dieser Basis entstandene Modellierung der Prognosen der Gesamtnachfrage an Strom wird im Kapitel 3.3 erläutert. Zunächst wird jedoch das Angebot an regenerativem Strom, in Kapitel 3.2, näher beschrieben.

3.2 Struktur des Gesamtangebots von Strom aus erneuerbaren Energien in Deutschland

Neben den ständigen Änderungen der Gesamtnachfrage an Strom in Deutschland treten solche auch im Gesamtangebot an Strom aus erneuerbaren Energien auf. Es ist wichtig, das Stromangebot aus erneuerbaren Energien zu betrachten, da es nur möglich ist, den Fußabdruck der eigenen Stromverbräuche zu reduzieren, wenn genug Strom aus erneuerbaren Energien im gewünschten Zeitraum verfügbar ist.

Die Abbildung des Stromangebots aus erneuerbaren Energien ist komplex, da es verschiedene kurzfristige und langfristige Einflüsse gibt.

Zu den kurzfristigen Einflüssen zählen zum Beispiel das Wetter und die Tageszeit. Das Wetter und die Tageszeit haben einen großen Einfluss auf das Angebot aus erneuerbaren Energien, da erneuerbare Energiequellen wie Solar- und Windenergie von der Verfügbarkeit an Sonnenlicht und Wind abhängig sind. Zum Beispiel kann ein bewölkter Tag die Menge an Strom reduzieren, der von Solaranlagen produziert wird, da die Wolken die Sonnenstrahlen abschwächen. Gleiches gilt zur Nachtzeit, da hier nicht ausreichend Sonnenlicht vorhanden ist. [De23]

Langfristige Einflüsse stellen zum Beispiel die Verfügbarkeit von Speicherkapazitäten oder Kapazitäten zur Stromerzeugung dar. Das Vorhandensein an Speicher ermöglicht, dass erzeugter Strom auch zu einem späteren Zeitpunkt verwendet werden kann. Die Möglichkeiten zur Stromerzeugung wiederum, dass die Quellen regenerativer Energien überhaupt effektiv und effizient genutzt werden können. Je größer die Speicher- und Erzeugungskapazitäten also sind, desto mehr Strom aus erneuerbaren Energien kann erzeugt und über einen längeren Zeitraum verwendet werden. [Se09]

4 Konzept: Modellierung der Nachfrage- und Angebotsprognosen

Nach der Analyse der Verfügbarkeit und Struktur der Daten werden im Folgenden Konzepte für die Modellierung der Nachfrage und des Angebots vorgestellt. Anschließend folgt das Zusammenbringen der beiden Prognosen über einen Ranking-Algorithmus.

4.1 Ermittlung der prognostizierten Gesamtnachfrage mit Hilfe mathematischer Näherungsverfahren

Die Analyse der Daten der gesamten Stromnachfrage in Deutschland, beschrieben in Kapitel 3.1, zeigt Schwankungen auf, die für die Prognose des zukünftigen Stromverbrauchs berücksichtigt wurden. Die erheblichen Unterschiede in den Verbräuchen liegen sowohl in den tageszeitabhängigen und wochentagabhängigen Schwankungen als auch den Schwankungen im Verlauf eines Jahres. Vor diesem Hintergrund wurde die stündliche Taktung der Daten, wie sie bereits von der Bundesnetzagentur über die SMARD angeboten wird, für eine möglichst präzise Modellierung beibehalten. Die für jeden einzelnen Wochentag verfügbaren Daten wurden für die Prognose gemäß den Schwankungen zwischen Werktagen, Samstagen, Sonntagen und Feiertagen unterteilt und als Tages-Art in die Modellierung mit einbezogen und gewichtet. Um jährliche und temporäre Schwankungen, die Auswirkungen auf die Prognosen haben können, zu berücksichtigen, werden vor jeder Prognose die aktuellen Strommarktdaten abgefragt. Der genaue Aufbau der Modellierung, die Befüllung der Datenbank und die Berechnung der Prognosen werden im Folgenden anhand der Darstellung des Konzepts in Abbildung 4 näher erläutert.

Die Datenbank besteht aus zwei wesentlichen Tabellen, zum einen aus der Tabelle mit den historischen Daten (*HistoricalDemand*) und zum anderen aus der Tabelle mit den prognostizierten Daten (*PrognoseDemand*). Die Grundlage der Daten, und somit für die Berechnung der Prognosen, stellen die Strommarktdaten der Bundesnetzagentur dar, deren Struktur im Kapitel 3.1 bereits beschrieben wurde. Die Datenbank wird durch das Skript *initializeDbHistoricalDemand* initial befüllt. Für das regelmäßige Aktualisieren der historischen Datenbanktabelle (*updateDatabase*) wurde die SMARD API verwendet [Wi23]. Aus dieser API werden sowohl die Zeitstempel als auch die Zeitreihen URL genutzt. Die Zeitreihen enthalten dabei die Angaben zu Datum, Uhrzeit und Netz-Last. Um diese Daten abfragen zu können, muss jedoch erst der richtige Zeitstempel ermittelt werden. Dazu wird der Unix Code aus der Zeitstempel URL zunächst in das in der Datenbank genutzte Datumsformat konvertiert. Damit wird aus der historischen Datenbanktabelle der zuletzt gespeicherte Eintrag abgerufen und dessen Zeitstempel, bestehend aus Datum und Uhrzeit, wieder zurück in Unix Code konvertiert und damit dann die neuesten Daten aus API abgefragt, dies erfolgt über die Zeitreihen URL. Alle verfügbaren Einträge, die neuer sind als die bereits in der historischen Datenbanktabelle verfügbaren, werden mit Datum, bestehend aus Datum und Uhrzeit, und Netzlast übernommen, leere Einträge werden dabei herausgefiltert.

Anschließend werden diese neuen Datensätze formatiert, zudem wird deren Tages-Art ermittelt (*defineHolidays*) und die daraus entstehenden Dataframes, bestehend aus Datum, Uhrzeit, Tages-Art und Netzlast, in der historischen Datenbanktabelle ergänzt.

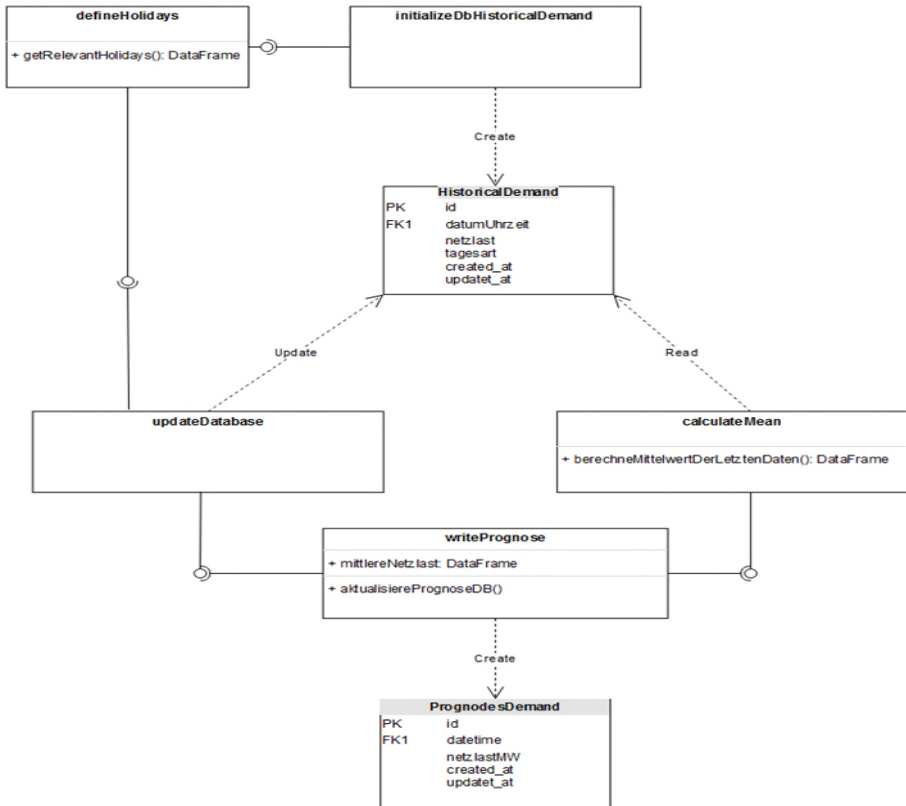


Abbildung 4: Grafische Darstellung des Modellierungskonzepts der Stromnachfrage

Vor jeder Aktualisierung der Prognosen werden die aktuellen Verbrauchsdaten auf diese Art und Weise ergänzt, um möglichst aktuelle, auf kurzfristige Veränderungen in der Marktnachfrage reagierende Prognosen zu erhalten. Die Prognose-Tabelle wird dann durch das Skript *writePrognose*, basierend auf den zuvor aktualisierten Daten, mit den neuesten Prognosen, jeweils für den heutigen und die drei darauffolgenden Tage, ergänzt. Sind für einen dieser vier Tage bereits Prognosen vorhanden, so werden diese mit neuen Daten überschrieben. Zusätzlich zu den aktuellen Datensätzen enthält diese Datenbanktabelle auch alle vergangenen Prognosen. Die Berechnung der Prognosen erfolgt über die Mittelwertberechnung der historischen Daten (*calculateMean*), dazu werden jeweils die letzten Werte der Uhrzeit und Tages-Art herangezogen und deren Mittelwert ermittelt.

Die prognostizierten Daten werden dann für das Matching von Nachfrage und Angebot verwendet, die genaue Zusammenführung dieser Datensätze wird im folgenden Kapitel beschrieben.

4.2 Ermittlung des prognostizierten Gesamtangebots mit Hilfe des Grünstromindex

Für die Prognose des Gesamtangebots an Strom aus regenerativen Energiequellen wurde innerhalb des Projekts die Grünstromindex API eingebunden [Gr22]. Mit der API ist es möglich, unter Angabe der gewünschten Postleitzahl, GSI-Werte für die nächsten 72 Stunden zu erhalten. Die GSI-Werte repräsentieren die dynamischen Tarife für Elektrizität (EFA10) des Stufenmodells zur Weiterentwicklung der Standards für die Digitalisierung der Energiewende. Die Tarifstufen liegen zwischen den Werten 0 und 100 und sagen aus, wie viel regenerativer Strom sich im Netz befindet und vom Verbraucher in Anspruch genommen werden kann. Als Datengrundlage werden für die Prognose unter anderem die Netzstruktur, historische Schaltzustände, Mittelwerte von verschiedenen Wettervorhersagen und Marktstammdaten verwendet. [Zo22]

4.3 Zusammenführung von Angebot und Nachfrage und Empfehlung

Da es sich bei dem GSI des Angebots pro Postleitzahl und der ermittelten prognostizierten gesamten Netzlast auf Seiten der Nachfrage um Werte mit verschiedenen Einheiten handelt, können diese nicht direkt miteinander verrechnet werden, um eine prognostizierende Aussage zu treffen. Aus diesem Grund wurde ein Matching Algorithmus entwickelt, der näherungs- und schrittweise vorgeht, um zu einer Empfehlung zu kommen. Die Schritte werden in Abbildung 5 dargestellt und im Anschluss näher erläutert.

- Schritt 1: Abruf der API
Die GSI-Werte und prognostizierten Netzlasten werden mit einem GET-Request abgerufen. Anschließend werden die Werte im JSON-Format für die nächsten 72 Stunden zurückgegeben.
- Schritt 2: Speichern der Werte
Die abgerufenen Werte der Nachfrage und des Angebots werden jeweils als Datensätze in zwei assoziativen Arrays gespeichert. Dabei wird eine ID als Zahl von 1-72 zugewiesen. Diese Zahl beschreibt den zeitlichen Abstand in Stunden, von der aktuellen Uhrzeit ausgehend.
- Schritt 3: Ranking der Nachfrage und des Angebots
Die vorher erstellten Arrays werden jeweils anhand des GSIs und der prognostizierten Netzlast sortiert. Die Sortierung findet bei dem Angebot absteigend und bei der Nachfrage aufsteigend statt, weil ein hohes Angebot und eine niedrige Nachfrage in Hinblick auf die Gesamtverfügbarkeit von Strom aus erneuerbaren Energien als positiv zu bewerten sind. Die Sortierung stellt den Rang eines jeden Datensatzes dar und wird dem Datensatz hinzugefügt.
- Schritt 4: Summierung der Rankings

Die beiden Arrays werden anschließend aufsteigend anhand ihrer ID sortiert, sodass sie wieder zeitlich chronologisch nach der jeweiligen Stunde angeordnet sind. Anschließend werden die Ränge der Datensätze beider Arrays summiert und pro ID und Stunde in ein neues Array eingetragen. Je niedriger die Nachfrage und je höher das Angebot, desto niedriger ist dabei die Summe. Die Datensätze, beziehungsweise Stunden mit niedrigerem summiertem Rang sind deshalb empfehlenswerter für den Verbrauch von Strom.

• Schritt 5: Darstellung

Nachdem Angebot und Nachfrage für jede der nächsten 72 Stunden bewertet und addiert wurden, können Empfehlungen farblich formatiert und abhängig von der Eingabe der Nutzenden ähnlich einer ABC-Analyse auf dem Bildschirm ausgegeben werden. Näheres zu der Darstellung der Ergebnisse und der Bedienoberfläche wird im nächsten Kapitel erklärt.

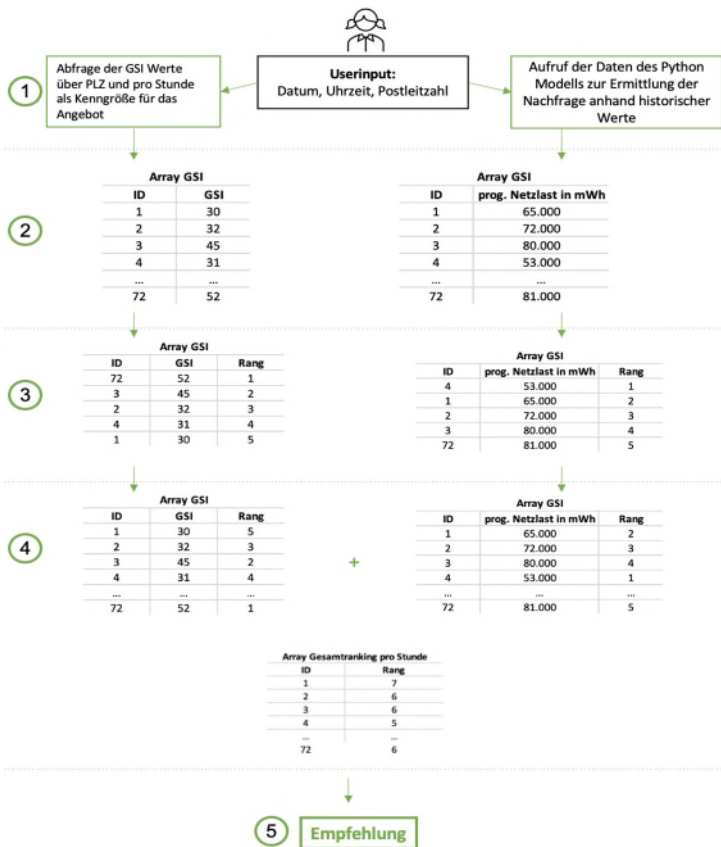


Abbildung 5 - Matching Algorithmus

5 Umsetzung: App mit Dashboard zur Prognose der Gesamtnachfrage und des Gesamtangebots

Die modellierten Prognosen könnten auf unterschiedliche Weisen genutzt werden. Eine vergleichsweise einfache Möglichkeit ist das Nutzen der Daten für das möglichst sinnvolle Betreiben elektrischer Geräte in Privathaushalten. Anwendende sind in diesem Use-Case Privatpersonen, die eine Information darüber erhalten möchten, wann es innerhalb der nächsten Stunden ökologisch sinnvoll bzw. weniger sinnvoll ist, Strom zu verbrauchen. Zur Erfüllung dieses Use-Cases wurde der Prototyp einer Webanwendung entwickelt. Der Aufbau der Infrastruktur und der GUI dieser Webapplikation werden im Folgenden näher erläutert.

5.1 Aufbau der Infrastruktur des Prototyps

Die Infrastruktur der Anwendung basiert auf der Containerisierung bzw. Virtualisierung von Systemressourcen. Jeder Service der Anwendung wird in einem eigenen Docker-Container betrieben. Hierbei wird bei der eingesetzten Virtualisierungstechnologie Docker ein eigenes Subnetting auf dem Host-System erstellt, wobei die verschiedenen Services ein in sich geschlossenes Netz bilden. Um die Kommunikation innerhalb der Service-Struktur zu vereinfachen und um bei ggf. vielen Anfragen von Nutzenden bestimmte Services hochskalieren zu können, (z.B. die Berechnungsumgebung für Nachfrage oder die Datenbankinstanz) wurde ein Proxy-Server in die Infrastruktur integriert, welcher als Stellvertreter für den eingehenden Datenverkehr der Applikation fungiert.

5.2 Darstellung der Bedienoberfläche und der Ergebnisse

Innerhalb der Applikation haben Nutzende die Möglichkeit, verschiedene Eingaben zu tätigen. Wie in Abbildung 6 dargestellt, ist eine dieser die Postleitzahl. Diese wird für die Abfrage des GSI-Wertes benötigt. Zwei weitere Werte sind Prozentwerte für den Bereich A und den Bereich B. Der Prozentwert für den Bereich C wird aus $1-A-B$ berechnet. Die Bereiche bestimmen im Nachgang, ähnlich wie in einer ABC-Analyse, welcher Anteil der Werte in der Empfehlung mit grün, gelb oder rot dargestellt wird. Wählt der Nutzende beispielsweise für A 10 %, B 20 %, und C 70 %, werden die Top 10 % der Datensätze mit den niedrigsten summierten Rängen in Grün, die nächsten 20 % in Gelb und die verbleibenden 70 % in Rot angezeigt.

Zudem haben Nutzende die Möglichkeit, mit einem Klick auf den Button "Prognose Starten" die Abfrage zu starten und die Ergebnisse in der Übersicht darzustellen.

Die übersichtliche Darstellung der Empfehlung zu jeder Stunde erfolgt in tabellarischer Form und chronologischer Reihenfolge (siehe Abbildung 6). Hier wird je Stunde angezeigt, ob das Stromangebot und die Stromnachfrage relativ hoch, mittel oder niedrig

sind. Auf Basis der summierten Ränge wird jeder Datensatz zudem farblich markiert und so der jeweilige Bereich gemäß der Einteilung der Nutzenden angezeigt.

Nachdem die Nutzenden eine Abfrage durchgeführt haben, kann über ein Drop-Down Menü ein Zeitpunkt, zu dem nähere Details über das Angebot und die Nachfrage angezeigt werden sollen, gewählt werden. Wie in Abbildung 6 deutlich wird, können der GSI, die Angebots- und Nachfrage-Ränge sowie die Verbraucherempfehlung (Bereich) eingesehen werden. Auch wird gezeigt, dass aufgrund einer niedrigen Netzlast und dem vergleichsweise hohen Grünstromindex die Ampel auf Grün schaltet.



Abbildung 6 - GUI und Ergebnisübersicht des Prototyps

6 Ausblick: Potential und Möglichkeiten zur Weiterentwicklung

In diesem Kapitel werden die Erweiterungsmöglichkeiten und andere zukünftige Potentiale der Applikation zur Optimierung des Ökostromverbrauchs beschrieben. Diese umfassen zunächst die möglichen Verbesserungen hinsichtlich der Nachfragemodellierung, zeigen Möglichkeiten in der Optimierung der Infrastruktur auf und bieten einen Ausblick hinsichtlich des Einbeziehens weiterer Einflussfaktoren, als auch Erweiterungsmöglichkeiten über die ursprünglich angedachte, manuelle Nutzung der Daten und Applikation hinaus.

Die Stromnachfrage und das Angebot von Ökostrom richten sich in Deutschland nach vielen Faktoren, welche, wie in den vorangegangenen Kapiteln beschrieben, aktuell vor allem basierend auf den historischen Daten zur Netzlast, und unter Berücksichtigung der jeweiligen Tagesart, in die Modellierung des prognostizierten Stromverbrauchs und den aktuell im Netz vorhanden Anteil regenerativen Stroms in die Applikation eingeflossen sind. Eine Erweiterung der Applikation um verbesserte Datenmodellierung, zusätzliche Einflussfaktoren wie Wetter, individuelles Nutzungsverhalten und langfristige Prognosen des Stromverbrauchs werden im folgenden Abschnitt näher beschrieben. Die momentan genutzte Nachfragemodellierung kann in zukünftigen Versionen verbessert werden, indem ausführliche Prüfungen und Bewertungen der Qualität der aktuellen Prognosedaten durchgeführt werden. Diese sollte sowohl eine Fehleranalyse als auch einen Abgleich der Prognosen mit den tatsächlichen Stromverbrauchsdaten, also eine detaillierte Korrelationsanalyse umfassen. Basierend auf diesen Analysen könnten weitere Modellierungsalgorithmen angewandt und ebenfalls bewertet werden. Der Prototyp wurde zunächst mit Mittelwerten modelliert, eine Erweiterung um lineare oder quadratische Regressionen könnten die Prognosen weiter verbessern.

Des Weiteren könnten örtliche Schwankungen im Hinblick auf die Nachfragedaten mit einbezogen werden. Da die Postleitzahl für die Ermittlung des Stromangebots bereits von Nutzenden abgefragt wird, kann diese zukünftig auch für lokale Schwankungen in der Nachfrage herangezogen werden. Diese könnten allgemeine Schwankungen und darüber hinaus auch die Unterschiede an regionalen Feiertagen mit einbeziehen und somit die Prognosen für den Stromverbrauch weiter verbessern.

Eine weitere Möglichkeit, die Prognosen zu verbessern, ergibt sich aus dem Einbeziehen von Wetterdaten. Auch wenn der Einfluss des Wetters (vor allem Sonne und Wind) eine größere Auswirkung auf die Erzeugung und somit das Angebot regenerativen Stroms hat, so sind die Auswirkungen auf das Nutzungsverhalten und somit den Stromverbrauch nicht zu vernachlässigen.

Zusätzlich zu diesen allgemeinen Verbesserungen der Prognosen, könnte den Nutzenden in der Applikation eine weitere Eingabemaske zur Verfügung gestellt werden, über welche das individuelle Nutzungsverhalten erfasst werden kann. In Kombination aus einer verbesserten Modellierung und unter Einbeziehen weiterer Einflussfaktoren in die Prognosen, kann der Zeithorizont der Prognosen von aktuell 72h auf weitaus langfristige Prognosen erweitert werden.

Über die zuvor beschriebenen Ausbaumöglichkeiten hinsichtlich der Modellierung hinaus, gibt es außerdem Potentiale bei der Verbesserung der service-basierten Infrastruktur. Hier ist davon auszugehen, dass eine rein relationale Datenbank nur begrenzte Möglichkeiten zur Analyse von großen Datenmengen bietet und zudem wenig flexibel im Hinblick auf Datenformate oder Datentypen ist. Dabei könnte eine NoSQL- oder Dokumentenbasierte Datenbank eine höhere Flexibilität bei der Analyse von Dokumenten (bspw. CSV-Dateien) bieten, ohne die Inhalte von Input-Daten in eine relationale Datenbank migrieren zu müssen. Der Einsatz einer NoSQL Datenbank würde

sowohl die Performance der Anwendung verbessern als auch die Möglichkeit schaffen, verschiedene Daten- bzw. Dateiformate als Berechnungsgrundlage zu nutzen.

In dem folgenden und letzten Abschnitt dieses Papers werden Erweiterungsmöglichkeiten über die ursprünglich angedachten Anwendungsfälle hinaus, vor allem in Hinblick auf die Einbeziehung aktueller Strompreise und Erweiterungen durch Anbindung an Smart Home Geräte und Applikationen, dargelegt.

In Anbetracht der Tatsache, dass die Preisschwankungen am deutschen Energiemarkt nicht mit aktuellen Tagespreisen an die Endverbrauchenden weitergegeben werden, ist eine Erweiterung der Applikation, um den Einflussfaktor Preis, nur begrenzt sinnvoll. Entwickelt sich der deutsche Strommarkt und vor allem die Abrechnung der Energiekonzerne dahingehend weiter, dass der Verbrauch auf mit Tagesaktuellen oder zumindest pro Monat gemittelten Strompreisen abgelesen und entsprechend abgerechnet wird, würde das Einbeziehen der Marktdaten die Attraktivität und den Nutzen der Applikation enorm steigern.

Dazu wäre jedoch eine Reform des deutschen Strom- und Energiemarktes, nach Vorbild des norwegischen Modells erforderlich. Bei diesem können die Endverbrauchenden entscheiden, ob sie einen Stromtarif mit Festpreis oder einen mit schwankenden Preisen wählen. Die Preisstruktur richtet sich dabei nach den skandinavischen Börsenpreisen, welche wiederum stark durch das besonders im Sommer und Winter schwankende Angebot des Ökostroms, welcher in Norwegen rund 99,1 % ausmacht und vorrangig durch Wasserkraft erzeugt wird [St23], beeinflusst wird. Um die Preisstruktur in Deutschland anpassen zu können, bedürfte es einer anderen Erfassung der individuellen Stromverbräuche, dazu müssten in Deutschland jedoch flächendeckend intelligente Stromzähler installiert und genutzt werden [Uk12]. Gemäß dem Beschluss des Bundeskabinetts vom 11. Januar 2023, werden, sollte dem Gesetzentwurf in allen Instanzen zugestimmt werden, die sogenannten Smart Meter ab 2032 Pflicht [Du23]. Sobald die Smart Meter dann entsprechend verbreitet sind und die Stromanbieter Verträge mit schwankenden Strompreisen anbieten, könnte die App auch dahingehend erweitert werden, dass sowohl die Daten vom Stromzähler im 15 Minuten Takt erfasst werden als auch in ähnlich regelmäßigen Abständen die Marktdaten vom Stromanbieter direkt oder der Europäischen Energiebörse EEX [EE23] abgerufen und in die Modellierung von Angebot und Nachfrage mit einbezogen werden.

Außerdem stellt sich die Frage, ob mittel- und langfristige Prognosen zu steigenden Bruttostromverbräuchen, beispielsweise 11 % Steigerung des Bruttostromverbrauchs bis 2030 [3], anteilig mit in die Modellierung einbezogen werden sollten, oder ob im Rahmen der aktuellen Energiekrise (2022/2023) davon abgesehen werden kann, da mit niedrigeren Verbräuchen durch erhebliche Einsparungen in Industrie und Privathaushalten zu rechnen ist.

Überdies könnte eine Erweiterung der App für die Nutzenden darin bestehen, dass zwischen privater und gewerblicher Nutzung von Ökostrom unterschieden werden kann.

Eine weitere Möglichkeit die Applikation so zu verbessern, dass regenerativer Strom dann genutzt wird, wenn die Nachfrage am geringsten und der im Netz befindliche Ökostrom am höchsten ist, wäre eine zukünftige Anbindung an andere Applikationen, beispielsweise Smart Home Apps, mit denen Nutzende ihre elektrischen Geräte genau für die prognostizierten Zeiträume fernsteuern und starten, oder stoppen könnten. Darüber hinaus könnte die Anbindung an Applikationen und smarte Geräte noch weiter optimiert werden, indem diese nicht mehr durch Nutzende gesteuert, sondern automatisiert die Daten aus der in diesem Projekt geprototypen App beziehen, auswerten und dann den Ökostrom entsprechend nutzen, wenn dieser das Stromnetz potenziell sogar überlasten könnte.

Literaturverzeichnis

- [De23] Deutscher Wetterdienst (2023): Wettervorhersage für erneuerbare Energien. https://www.dwd.de/DE/forschung/wettervorhersage/num_modellierung/07_wettervorhersage_erneuerbare_energien/vorhersage_erneuerbare_energien_node.html (21.05.2023)
- [Du23] Duy, S. (2023): Das Smart Meter kommt: Das sind Deine Vorteile. <https://www.finanztip.de/stromzaehler/> (09.03.2023)
- [EE23] EEX (2023): Marktdaten Strom EEX. <https://www.eex.com/de/marktdaten/strom> (04.03.2023)
- [Gr22] Grünstromindex - Corrently API (2022): <https://corrently.io/books/grunstromindex> (27.03.2023)
- [In22] Informationsportal Erneuerbare Energien (2022): Grafiken und Diagramme unter Verwendung aktueller Daten der Arbeitsgruppe Erneuerbare Energien-Statistik. https://www.erneuerbare-energien.de/EE/Navigation/DE/Service/Erneuerbare_Energien_in_Zahlen/Entwicklung/entwicklung-der-erneuerbaren-energien-in-deutschland.html (13.03.2023)
- [Ke21] Kemmler, A.; Wünsch, A.; Burret, H. (2021): Entwicklung des Bruttostromverbrauches bis 2030. [https://www.prognos.com/de/projekt/entwicklung-des-bruttostromverbrauches-bis-2030#:~:text=Im%20Zielszenario%20steigt%20der%20Bruttostromverbrauch,Jahr%20030%20\(%2B11%20Prozent\).](https://www.prognos.com/de/projekt/entwicklung-des-bruttostromverbrauches-bis-2030#:~:text=Im%20Zielszenario%20steigt%20der%20Bruttostromverbrauch,Jahr%20030%20(%2B11%20Prozent).) (06.11.2022)
- [Se09] Seifert, J. (2009): Preismodellierung und Derivatbewertung im Strommarkt – Theorie und Empirie. https://www.google.de/books/edition/Preismodellierung_und_Derivatebewertung/0fAq6QmbFc8C?hl=de&gbpv=1
- [SM22] SMARD (2022, 2023): Strommarktdaten. <https://www.smard.de/home/downloadcenter/download-marktdaten/#!?downloadAttributes=%7B%22selectedCategory%22:false,%22selectedSubCategory%22:false,%22selectedRegion%22:false,%22from%22:1666821600000,%22to%22:1667775599999,%22selectedFileType%22:false%7D> (13.03.2023)

- [St22] Statista (2022): Bruttostromverbrauch in Deutschland in den Jahren 1990 bis 2022. <https://de.statista.com/statistik/daten/studie/256942/umfrage/bruttostromverbrauch-in-deutschland/> (13.03.2023)
- [St23] Statista (2023): Anteil der Energieträger an der Nettostromerzeugung in Norwegen in den Jahren von 2021 bis 2022. <https://de.statista.com/statistik/daten/studie/1292636/umfrage/struktur-der-stromerzeugung-in-norwegen/> (04.03.2023)
- [Um10] Umweltbundesamt (2010): Energieziel 2050. https://www.umweltbundesamt.de/sites/default/files/medien/378/publikationen/energieziel_2050.pdf (13.03.2023)
- [Um22] Umweltbundesamt (2022): Stromverbrauch. <https://www.umweltbundesamt.de/daten/energie/stromverbrauch> (19.12.2022)
- [Um22a] Umweltbundesamt (2022): CO₂-Emissionen pro Kilowattstunde Strom steigen 2021 <https://www.umweltbundesamt.de/themen/co2-emissionen-pro-kilowattstunde-strom-steigen> (13.03.2023)
- [Uk12] Uken, M. (2012): Strompreise Norwegen Deutschland. <https://www.zeit.de/wirtschaft/2012-09/strompreise-norwegen-deutschland> (04.03.2023)
- [Wi23] Wittmann, L.; Fischer, A. (2023): bundesAPI – SMARD API. <https://github.com/bundesAPI/snard-api> (18.03.2023)
- [Zo22] Zoerner, T. (2022): EAF-10 - Dynamische Tarife für Elektrizität <https://corrently.io/books/grunstromindex/page/eaf-10-dynamische-tarife-fur-elektrizitat> (27.03.2023)

Design and implementation of a lecture for teaching current Green Coding approaches and practices at the HTW Berlin

Dennis Junger ¹, Volker Wohlgemuth ¹

Abstract: At the Industrial Environmental Informatics Unit (HTW Berlin), a test integration for a future 'Green Coding' course was developed to implement the concept of environmentally conscious and energy-efficient software development. The course titled "Current Development Trends in Environmental Informatics" now educates students on fundamental concepts and techniques in this field. The course emphasizes the development of software that is both energy efficient and environmentally friendly. Students learn how to optimize software applications' energy efficiency by designing energy-efficient algorithms and enhancing server and data center performance. The course covers essential principles of 'Green IT' and explores the 'Green Coding' sub-field. The course is divided into three modular parts: internal lectures, external lectures, and a practical part to manifest the theoretical knowledge. The content of the internal lectures was significantly influenced by the current state of research regarding 'Green IT' at the Industrial Environmental Informatics Unit of the HTW Berlin, in which current and past master theses, papers, and publications were presented. Current known German 'Green Coding' researchers were invited for external lectures, partly on-site and online. A practical measurement course lasting several days, which provided the students with several test setups according to the "Blue Angel for software products" standard to pursue their research questions, was created to deepen their knowledge. The following article presents and discusses the topics and results of these short papers and evaluations of the course concept and the course itself.

Keywords: Green Coding, Potentials of Green Coding, Curricula, Software Engineering, ISOC Foundation, Trier University of Applied Sciences, HTW Berlin, University of Applied Sciences

1 Introduction: Motivation and Relevant Terminology

Sustainability as relating to information and communication technology (ICT) generally and software development/engineering more specifically is a research field of growing interest. While the focus of sustainability considerations in ICT, for the longest time, was on hardware, software behavior is the driver of hardware energy consumption and additionally affects the hardware life-cycle e. g. as ever more demanding software systems quickly outgrow older hardware [Wi95]. To enable sustainable software production, access to models, methods, and tools geared toward sustainability is elementary. Furthermore, it is essential to make future IT professionals and researchers aware of the sustainability aspects of ICT and equip them with the skills to consider them in their work adequately. The terms 'Green IT' and 'Green By IT' separate the field into two main areas, 'Green By IT' designates ICT that aims to have a positive impact on sustainability in other areas, whereas 'Green IT' deals with making ICT itself more sustainable [Uh12]. 'Green Coding' is a part of 'Green IT'. 'Green

¹ HTW Berlin, University of Applied Sciences, Industrial Environmental Informatics Unit, Wilhelminenhofstr.75A, 12459 Berlin, Germany dennis.junger@htw-berlin.de, volker.wohlgemuth@htw-berlin.de

IT' focuses on all the aspects of reducing emissions of ICT and focuses on the methods and tools that can be leveraged during development to impact software sustainability in all of its life-cycle phases (development, usage, and disposal phase). The authors expected the conception and evaluation of the 'green coding' course to increase the understanding of green coding and thus to make the future generation of computer scientists more aware of green coding. The current state of the art will be presented in the following article. Subsequently, the conception of the course, divided into the selected didactic concept and the grading concept, as well as the structure of the theoretical and practical parts, are presented to be able to classify and evaluate them in the following evaluation part.

2 Background and Related Work

2.1 Current and Suggested Green Coding Practices for Instruction

To determine the course's relevant 'Green Coding' content, we need to explore existing concepts of environmentally friendly software development and its impact on energy consumption. In addition to models that make the software development cycle itself green, for example, using a sustainability report such as the one designed by the UCB in the 'GREENSOFT model' [Na11, Di13], it is not only the engineering process itself that is important but also the methods for technically measuring and improving energy consumption. These methods and models are considered suitable for providing software developers with insights into their code's potential for energy optimization. ([Ke13, Hö13]). A recurring theme in numerous studies is the absence of a universally optimal solution regarding energy and resource consumption[Pe17]. This demonstrates the complexity of software sustainability and highlights why IT professionals need to be able to assess which measures are suited for a specific project. However, the exact steps necessary to improve upon a given piece of source code's induced energy consumption are not as clear. This is because every software and every algorithm has different solutions and third-party code such as libraries, frameworks, and Software Development Kits (SDK) are only sometimes known or disclosed. Some work provides concrete hints [Hö14] or details where the energy hot spots occur in the source code [Wa12, Ve18]. To develop a general approach to sensitize the students, it is essential to understand the basics and to get a feeling for the impact of software on the environment. Topics like the impact of hardware and software on the desktop device or in the cloud, for example, data centers, are reasonable. Unfortunately, the sending of data through networks can be taught poorly, so this needs to be addressed. Further steps are to learn the basics of software energy consumption measurement so that students can apply this in future implementations. It has been observed that integrating green software into the computer science curriculum presents a notable difficulty, as engineers often need more knowledge, expertise, and tools to create environmentally sustainable software. According to Saraiva [Sa21], recent studies have shown that just these current tools should be taught. To address this, current projects were also presented. Recent projects and foundations, such as the

green web foundation², Sustainable Digital Infrastructure Alliance (SDIA)³ and the Green Software Foundation (GSF)⁴, which are each concerned with individual approaches to solutions, should be presented so that students learn explicit solution strategies and message groups in addition to the holistic theoretical approach. Furthermore, the first attempts to implement online software courses like this one of the GSF are presented⁵. Besides the parallel literature analysis for the research project "Potentials of Green Coding"[Ju23], which summarizes state of the art as of 2010, similar conclusions could be drawn. Deciding which content is relevant for providing a reasonable basis is difficult. It is also a challenge to implement curricula and discussions as an initial step [Ma11]. In 2010 [Ca10] invented a hypothetical agenda for integrating computing education for sustainability. The focus of [Ca10] is on waste management, eco-labeling, energy management, carbon management, reuse, environmental reporting, and open-source software. Most of these courses are already lectured in the HTW Masters program of the Industrial Environmental Informatics Unit⁶, where the pilot course will be part.

2.2 Integration of Green Coding Concepts into Existing Study Program Curricula

It is critical to greening ICT to give future IT professionals the skills needed to produce sustainable software. The focus of research question 3 of the "Potentials of Green Coding" project is to find possibilities for implementing the approaches into the education of future IT professionals. As most software developers have an academic background, this review focuses on implementing aspects of 'Green IT' into the curricula of higher education institutions. The basis for this idea is a statistic from 2022 [St22], which describes the levels of formal education for software developers worldwide in 2022, according to which 75,19 % of the software developers have a higher education background.

2.3 Current Implementation of Green Coding Modules and Courses

To address this question, the initial approach involved researching the courses and programs offered by higher education institutions. However, the search turned out to be more difficult than anticipated. Many universities provide specialization details for their study subjects in module handbooks, while others only make course information and disciplines accessible through an application management system after creating an account. These factors influenced our search, which primarily relied on online sources. In addition to sources from cooperating universities, we consulted job platforms such as LinkedIn⁷ and

² <https://www.thegreenwebfoundation.org/> [2023-05-22]

³ <https://sdialliance.org/> [2023-05-22]

⁴ <https://greensoftware.foundation/> [2023-05-22]

⁵ <https://learn.greensoftware.foundation/> [2023-05-22]

⁶ <https://bui.htw-berlin.de/> [2023-05-22]

⁷ <https://www.linkedin.com/> [2023-05-16]

institutions known in the 'Green Coding' community. Currently, courses designed for Green Software Engineering or 'Green Coding' are mainly available in Europe. We found the following universities and universities of applied sciences in Europe that teach related courses.

- England - University of Bristol
- Finland - LUT University
- Germany - Hasso Plattner Institute (HPI)
- Germany - University of Applied Sciences HTW Berlin
- Germany - Umwelt-Campus Birkenfeld (UCB)
- Netherlands - Amsterdam School of Data Science
- Netherlands - DELFT University of Technology
- Sweden - Mälardalen University
- Sweden - Royal Institute of Technology (KTH)
- Sweden - University of Gothenburg
- Ukraine - Kharkiv Aviation Institute (current status unknown)

The following two universities were identified within the United States of America:

- Carnegie Mellon University
- The University of California Berkeley

2.4 Creating a Green Coding Curriculum

For future curricula, Drake and Reid [Dr18] propose to, on the one hand, teach the big picture of a topic and, on the other hand, the main competencies. To convey this knowledge, technical, practical, and participatory interests should be addressed [Fr06]. Mishra and Mishra [Mi21] describe how a sustainable software engineering curriculum based on ACM/IEEE could be created. Since it is challenging to design entire curricula such as the PERCCOM [Ko19] Masters program, the first step should be to develop individual courses, workshops, discussion groups [Ma11] or university initiatives like in the HAW Hamburg [Ei23]. This can be followed by a module like "Software Engineering Sustainability", introduced at the 'Kharkiv Aviation Institute' [Tu19].

3 Conception of a Green Coding Course at HTW Berlin

In the context of the third research question, which was led by HTW Berlin, the conception of a module in the Master's program "Industrial Environmental Informatics" took place as

a pilot project. A test integration at HTW Berlin implemented the concept. The teaching subject 'Current Development Trends in Environmental Informatics'⁸ now introduces students to the basic concepts and techniques of environmentally conscious and energy-efficient software development. The course focuses on teaching students how to develop energy-efficient and environmentally friendly software by optimizing the energy efficiency of software applications by, for example, writing energy-efficient algorithms or optimizing the performance of servers and data centers. For this purpose, the basic concepts of 'Green IT' are taught, and the Green Coding subfield is addressed. The course is divided into three modular parts: internal lectures, external lectures, and a practical part to manifest the theoretical knowledge. Further grading was divided into a theoretical and a practical part. The academic content was examined through oral examinations focusing on 'Green IT', 'Green Computing', and 'Green Coding'.

3.1 Didactic Concept

A well-designed didactic concept aims to optimize the learning experience by promoting active engagement, fostering critical thinking, and facilitating the acquisition and application of knowledge and skills. For the course design, it was important that the learning objectives be achieved through theoretical and practical exercises. The students learn about selected current technologies that are relevant for the development of 'Environmental Management Information Systems' (EMIS) (e.g. 'mobile computing', 'Green IT'). The students learn to assess current technologies regarding their relevance to industrial environmental informatics. The course content should reflect the most relevant principles identified by previous research in the "Potentials of Greencoding" project by providing relevant literature, multimedia materials, technologies, tools and online resources, primarily via the learning platform Moodle. Feedback through sub-tasks and collaboration in the practical part will deepen the learning experience and guide learning. The practical units should activate the students and prepare them for later tasks. To check the effectiveness of the didactic concept, there will be evaluations and revisions, which will assess the course itself as well as the knowledge level of the students and the perceived knowledge level by means of anonymous surveys and evaluations to present the outcome as a key figure.

3.2 Grading System

To ensure a comprehensive and differentiated grading system that accurately reflects the skills acquired by participating lecturers, a multi-level grading system was implemented. As a prerequisite for final exam admission, students were required to evaluate the impact of ICT hardware. In this context, a lecturer of the Öko-Institut e. V. introduced the digital

⁸ https://bui.htw-berlin.de/fileadmin/HTW/Zentral/Rechtsstelle/Amtliche_Mitteilungsblaetter/2013/30_13.pdf [23.06.2023]

carbon footprint online calculator with a life cycle cost module where the students could contribute to a new functionality based on their theoretical research.

3.3 Theoretical Course Content - Presentations and Talks

The content of the internal lectures was significantly influenced by the current state of research regarding 'Green IT' at the Industrial Environmental Informatics Unit of the HTW Berlin, in which current and past master theses, papers, and publications were presented. External researchers gave lectures, partly on-site and online. For example, experts from the Öko-Institut e. V. gave lectures on the life cycle costs of ICT, the digital footprint of ICT, and the "Blue Angel for software products". The lectures of the LMU Munich⁹ presented the basics of 'Green IT' in data centers and Green Coding Berlin GmbH¹⁰ showed the current state of the art in the software industry. To learn more about the topic of implementation, the 'Greensoft Model' [Na11] was presented by the participating researchers of the UCB, with a focus on 'Green Software' and 'Green Software Engineering'.

3.4 Practical Course Content - Measurement Laboratory

To deepen the knowledge and gain experience with the "Blue Angel for software products" and the underlying methodologies and measurement methods, a measurement lab was designed. The following describes the design and implementation of the measurement lab.

3.4.1 Set-up and Preparation of the Measurement Laboratory

In the context of previous research at the HTW Berlin in 2022 [Ju22], a test stand with a system under tests was modeled on the example of the "Blue Angel for software products". Creating a second test stand was essential to enable the students to run the measurements smoothly and in parallel. To be able to make measurements with an increased GPU share, an additional computer and power meter were prepared for a second measuring station with an extended graphical processing unit. To diverse student needs and enhance efficiency in the lab, the standard image was equipped with additional automation software called 'WinAutomate' and 'Actiona', alongside the existing 'Power Automate'. A standardized workflow was developed, based on the measurement type described in the earlier research in 2022 [Ju22], incorporating fixed measurement duration and cooldown periods. In the measurements carried out later, this was 120 seconds of automation followed by 60 seconds of cooldown time for fitting it to the planned measurement practical time. This was repeated for statistical means 30 times according to the Blue Angel [Na21] model.

⁹ <https://www.lmu.de/de> [2023-05-22]

¹⁰ <https://www.green-coding.berlin/> [2023-05-22]

3.4.2 Measurement Laboratory and Its Results

The results of the measurement lab are presented below. The following topics were part of the laboratory and targeting to figure out differences in resource and energy usage of (proprietary and open source) applications in different fields, like image- (GIMP & Photoshop) and sound processing (Spotify & VLC Media Player) and statistical analyses (Matlab & R-Studio [Se23]). Also, programming technology topics, such as the differences between programming languages (C++ & Python) and (single- & multi-)threading technologies in GO¹¹ have been studied. The gathered results were presented to the entire course as lessons learned and written as a 10-page report.

3.5 Evaluation

To conduct a targeted evaluation, the participant's level of knowledge at the beginning and the end of the module needed to be captured and compared. For this reason, an anonymous Moodle survey was created for students that had been filled out by the students at the beginning and the end of the semester. The curve shown in figure 1 shows the survey results before and after course completion. The survey was designed in such a way that participants could choose between "Strongly Agree (1)", "Agree (2)", "Disagree (3)", and "Strongly Disagree (4)". A mean value was calculated before and after the course to determine the degree to which the content was conveyed. This mean value has to be seen about the respective questions. The questions were asked so that for each question the counter "Strongly Agree (1)" corresponds to the desired goal and "Strongly Disagree (4)" corresponds to the undesirable output. A detailed list of the questions with the corresponding results as RAW Data and detailed questions can be found in the supplementary material at <https://doi.org/10.5281/zenodo.7970613>. In the shown Figure 1, we find the evaluation of the surveyed fields of 'Green IT' for the review. There is continuous progress since each question's average value has improved. As a limitation of the survey, it should be mentioned that the students who were registered for the second exam period but already had to hand in their paper, probably have not yet worked on learning the lecture notes since the second period took place about three weeks after the paper hand-in. This could slightly worsen the survey results of the individual progress. The students successfully learned the criteria and indicators to evaluate the environmental impact of software (Question 4), got ideas about how to produce software that is more energy efficient (Question 7), learned how to evaluate (Question 15) and compare (Question 18) software in terms of power and resources and also knew how to conclude energy consumption to the cause of the problem (Question 19). If you would like more details about the evaluation, follow the link to the above RAW data.

¹¹ <https://go.dev/> [2023-05-22]

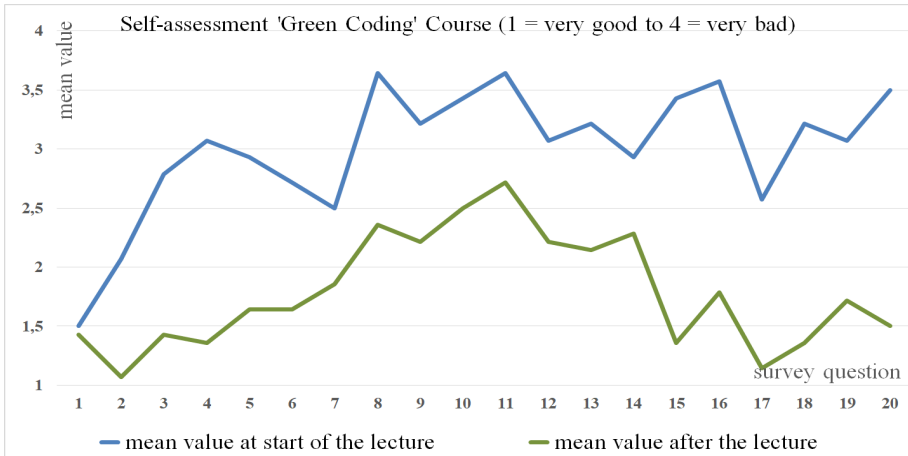


Fig. 1: Self-assessment 'Green Coding' - Pilot Integration

3.6 Challenges During the Lab and Solution Approaches

As described in the 3.2 grading section, a part of the final grade was designed for students to review each other's learning and the difficulties of the topic through lessons learned, which were also made available to the instructors. The most frequently mentioned feedback is listed below to identify and address weaknesses in future course designs. The measurements had to be repeated several times for the statistical mean. This resulted in long waiting and idle times during the scenarios. For the future, it's essential to define measurement blocks parallel to theoretical input so students can use their time more wisely. In addition, an open-door measurement lab would be good. Also to post-process the action, resource, and energy data and evaluate these data with 'OSCAR'¹² was very time-consuming and difficult. To solve this, 'OSCAR' and the measuring station should be reviewed or even extended to compare several measurements with the baseline. Besides data evaluation problems, there is also a problem with comparing scripts, due to no clear measurement instructions for this task. A universal approach could be researched and developed.

4 Conclusion and Outlook

It was possible to find the state of the art regarding the teaching of 'Green Coding' content. This was divided into teaching strategies and teachable topics, nonetheless supported by the previous quantitative literature analysis of the authors as project participants of "Potentials of Green Coding". The focus of 'Green Coding' could be identified and compared to current universities offering it. An example implementation with theoretical and practical

¹² <https://oscar.umwelt-campus.de/> [2023-05-16]

parts could be designed and implemented. The evaluation of the example showed mainly positive study results but also weaknesses in the organizational implementation of the lab part. Weaknesses in the current measurement methodology, mainly in the data post-processing, were identified. As a by-product of the course concept students could contribute to current 'Green Coding' research by means of papers that now can and should be published successively at conferences.

Acknowledgements

This work was funded by the Internet Society Foundation project "Potentials of Green Coding". We would also like to thank our partner, the German Informatics Society, the Institute for Software Systems at Environmental Campus Birkenfeld, especially Stefan Naumann, Achim Guldner, Max Westing, and Franziska Mai for talking and supporting the measurement lab. Also, we want to thank Arne Tarara from Greencoding Berlin GmbH, Jens Gröger from Öko-Institut e. V. and Maximilian Hüb from LMU Munich for their talks about their 'Green Computing' and 'Green Coding' expertise.

Bibliography

- [Ca10] Cai, Yu: Integrating sustainability into undergraduate computing education. In: Proceedings of the 41st ACM Technical Symposium on Computer Science Education. SIGCSE '10, pp. 524–528, 03 2010.
- [Di13] Dick, Markus et al.: Green software engineering with agile methods. In: 2013 2nd International Workshop on Green and Sustainable Software (GREENS). pp. 78–85, 2013.
- [Dr18] Drake, Susan et al.: Integrated Curriculum as an Effective Way to Teach 21st Century Capabilities. *Asia Pacific Journal of Educational Research*, 1:31–50, 01 2018.
- [Ei23] Eickstädt, Elina et al.: Computer Science for Future - Sustainability and Climate Protection in the Computer Science Courses of the HAW Hamburg, 2023.
- [Fr06] Fraser, Sharon P. et al.: The curriculum? That's just a unit outline, isn't it? *Studies in Higher Education*, 31(3):269–284, 2006.
- [Hö13] Hönig, Timo et al.: Proactive Energy-Aware System Software Design with SEEP. *Softwaretechnik-Trends*, 33(2):6–7, 2013.
- [Hö14] Hönig, Timo et al.: Proactive Energy-Aware Programming with PEEK. In: 2014 International Conference on Timely Results in Operating Systems. USENIX, 2014.
- [Ju22] Junger, Dennis et al.: Conception and test of a measuring station for the analysis of the resource and energy consumption of material flow-oriented environmental management information systems (EMIS). In: *EnviroInfo 2022*. Gesellschaft für Informatik e.V., 09 2022.
- [Ju23] Junger, Dennis et al.: Potentials of Green Coding - Findings and Recommendations for Industry, Education and Science, 2023. to be published.

- [Ke13] Kern, Eva et al.: Green software and green software engineering—definitions, measurements, and quality aspects. In: First International Conference on Information and Communication Technologies for Sustainability (ICT4S2013), 2013b ETH Zurich. pp. 87–91, 2013.
- [Ko19] Kor, Ah-Lian et al.: Education in Green ICT and Control of Smart Systems : A First Hand Experience from the International PERCCOM Masters Programme. IFAC-PapersOnLine, 52(9):1–8, 2019. 12th IFAC Symposium on Advances in Control Education ACE 2019.
- [Ma11] Malik, Alam Sher et al.: Twelve tips for developing an integrated curriculum. Medical teacher, 33(2):99–104, 2011.
- [Mi21] Mishra, Alok et al.: Sustainable Software Engineering: Curriculum Development Based on ACM/IEEE Guidelines. In: Software Sustainability. Springer, Cham, 2021.
- [Na11] Naumann, Stefan et al.: The greensoft model: A reference model for green and sustainable software and its engineering. Sustainable Computing: Informatics and Systems, 1(4):294–304, 2011.
- [Na21] Naumann, Stefan et al.: The eco-label blue angel for software - Development and components. In: Advances and New Trends in Environmental Informatics: Digital Twins for Sustainability. Springer, pp. 79–89, 2021.
- [Pe17] Pereira, Rui et al.: Energy Efficiency across Programming Languages: How Do Energy, Time, and Memory Relate? In: SIGPLAN International Conference on Software Language Engineering. ACM, New York, USA, pp. 256–267, 2017.
- [Sa21] Saraiva, João et al.: Bringing Green Software to Computer Science Curriculum: Perspectives from Researchers and Educators. In: ITICSE. ACM, 2021.
- [Se23] Seegert, Tim et al.: Energy and resource comparison of current applications with a focus on statistical analyses and evaluations using the example of Matlab and R/RStudio, 2023. to be published.
- [St22] Statista: Worldwide levels of formal education for software developers, 2022. <https://www.statista.com/statistics/793568> [2/24/2023].
- [Tu19] Turkin, Igor et al.: Software Engineering Sustainability Education in Compliance with Industrial Standards and Green IT Concept. In: Green IT Engineering: Social, Business and Industrial Applications. Springer, pp. 579–604, 2019.
- [Uh12] Uhl, Axel et al.: Beyond Green IT - Die Symbiose von IT und nachhaltiger Energie. In: Smart Energy. Springer, pp. 193–205, 09 2012.
- [Ve18] Verdecchia, Roberto et al.: Code-Level Energy Hotspot Localization via Naive Spectrum Based Testing. In: Advances in Environmental Informatics: Managing Disruption, Big Data and Open Science. Springer, 2018.
- [Wa12] Wang, Shinan et al.: Safari: Function-level power analysis using automatic instrumentation. In: 2012 International Conference on Energy Aware Computing. IEEE, dec 2012.
- [Wi95] Wirth, Niklaus: A Plea for Lean Software. Computer, 28(02):64–68, 1995.

State of the Art Open Access Remote Sensing with ESA Sentinel 1 SAR Data

Jennifer McClelland,¹ Tanja Riedel,² Florian Beyer,³ Heike Gerighausen,⁴ Burkhard Golla⁵

Abstract:

Tackling the consequences of climate change has become a global issue. Climate change will clearly influence our common lifestyle enormously in near future. This involves increasingly frequent sudden weather changes and extreme temperatures as well as drastic changes in water quality and availability. Because of our constant growing global population, nutritional habits and agricultural practices, the share of the agricultural impact on global anthropogenic greenhouse gas emissions take on an estimated 10-12 %. At the same time, fulfilling the agricultural demand is becoming increasingly challenging due to unpredictable farming conditions. Without immediate collaborative efforts including focused research, employment and adaptation of state of the art technologies, this issue will not be tackled soon enough, to avoid massive limitations and enormous losses.

A very promising large-scale technology to monitor agricultural ecosystems and activities is by means of earth observation imagery derived by Synthetic Aperture Radar (SAR). Radar backscatter e.g. allows insights to crop conditions, soil properties and direct mapping of vegetation growth.

Open access technologies offer the best solutions for collaborative efforts, thus minimising financial and legal constraints in comparison to technologies residing in the commercial sector.

Here, we combine and build on state-of-the-art tools and technologies to provide an easy to employ Sentinel-1 SAR pre-processing tool as well as a Germany wide, open access, pre-processed, analysis-ready database of Sentinel-1 SAR data. All tools used and developed are open source and freely available. With the employment of modern software developing methods and tools for a scalable and maintainable architecture, these products can be easily extended and adapted. By deployment of up to date machine learning methods, combining the resulting datasets with other relevant parameters, not to say the least, e.g. early prediction of optimal sowing, harvesting and fertilisation time points can be determined as well as many more valuable insights for successful, resource-efficient and environmentally friendly farming. Furthermore, the pre-processing of SAR datasets is not only substantial for the field of agriculture but for a wide range of other fields concerning environmental observations.

¹ JKI Institute for strategies and technology assessment (JKI-SF), research center for remote sensing in agriculture (JKI-FLF), Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany, jennifer.mcclelland@julius-kuehn.de

² JKI-SF, JKI-FLF, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany, tanja.riedel@julius-kuehn.de

³ JKI-FS, JKI-FLF, Bundesallee 58, 38116 Braunschweig, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany, florian.beyer@julius-kuehn.de

⁴ JKI-FS, JKI-FLF, Bundesallee 58, 38116 Braunschweig, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany, heike.gerighausen@julius-kuehn.de

⁵ JKI-SF, JKI-FLF, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany, burkhard.golla@julius-kuehn.de

Keywords: Sentinel-1 SAR; Analysis ready data; Agriculture; Open access technologies; Geoinformation systems

1 Motivation

Since the first personal computer and satellite in the 1950s technology has generally grown exponentially in developed economies worldwide. Lagging behind has always been the trust of the individual, institutions and organizations in the use and deployment of new technologies. In parts for good reason, considering the human inappropriate usability, the state of development when entering the market and the gradually growing misuse of generated data. Whilst on the one hand, an internationally wide engagement toward technological issues has led to a multitude of creative perspectives, approaches and problem solving, on the other hand, the competitive focus on gaining a large market share in promising discoveries and owning a new technology has led to a general non collaborative attitude.

Whilst competition and market share offer an important drive to investment and dedication, by now this has led to an overwhelming complexity of interfaces, protocols, languages and formats, where the same solutions are tackled over and over again becoming a strong limitation to progress. Unfortunately this phenomenon not only resides in business fields but also in science, often indirectly bound and strongly influenced by economy growth. Despite science targeting the general public for non-commercial reasons, the open source paradigm offering technological and general usage solutions for the community has been driven from the commercial-software industry beginning in 1997 and still only proven in necessity and more extensively supported over the last few years, finding its place in the open market. The open source paradigm as well as diverse online platforms offer technical solutions to combine and extend technical knowledge and expertise.

However, there not only lies a gap in the communication between technical tools and interfaces. There lies a gap between fields of expertise. Not only has technology advanced in parallel but also the impact of technologies, acceptance towards sharing solutions and data as well as deployment varies strongly in originally non tech fields. To solve the global challenge of climate crisis, barriers must be overcome and common languages found not only between technical approaches but also between application fields. It has become more important than ever to combine and expand skills from all scientific areas.

Besides the lack of trust in sharing data and solutions, two strongly limiting factors remain. Lack of technical knowledge and lack of skills, within increasingly critical fields, in reference to environmental sustainability, such as agriculture, geographic information systems and remote sensing. A lot of expertise situated in purely digital fields have not yet found their place in areas of natural science, although the demand for them has truly grown.

Thus, our approach attempts to combine expertise from modern software engineering, as practiced in industrial software companies, classic agricultural knowledge and the typical approach as a geographic information system data scientist.

2 State of the Art Remote Sensing

Satellite remote sensing earth observation is performed either by means of imagery with optical sensors or by interferometric synthetic aperture radar (InSAR). Although spectral analysis of images offer great insight to ground events, InSAR enhances this insight by its immunity to clouds and weather changes. Ever since the deployment of the ERS satellites in 1991 the InSAR technique has strongly matured and with advancing technologies and large financial investments in the deployment of the Sentinel-1 satellite infrastructure within the Copernicus project, access to SAR Data across the planet has been made publicly accessible since 2014. The open-source Sentinel Application Platform (SNAP) offer tools to process the Sentinel data provided by the ESA. While the, in Slovenia headquartered, GIS IT company Sinergise, operated “Sentinel Hub” platform offers a commercial solution to a processing engine and “free” user friendly access to Sentinel databases, Copernicus Data and Exploitation Platform – DE (CODE-DE) is a German solution, where the data center is situated in Frankfurt am Main (Germany) with a stronger orientation towards research and development. CODE-DE likewise offers direct access to all Sentinel satellite data as well as “free” user specific, dedicated virtual machines (VMs) for further processing of the data.

3 State of the Art Algorithm/Software Development in Science

While Python as a programming language has become most common in the fields of data science, it’s easy to learn language design and wide scope of software library’s lead to a broad quality in usage. Non-programmers can apply this language without any application of software programming paradigms such as procedural or object oriented programming. The language offers run-time compilation, transferring disclosure of syntax-, memory-access- and logic errors to run-time, although the initial awareness of insuring memory and logic safeness is not implied, as in e.g. Java or C++. In science the general approach of initially designing a suitable, modular, thus scalable and maintainable software architecture when approaching a technical problem is not taken. “Clean code”, code documentation and reviewing principles are not made use of as well as the equivalent abundance of tools, e.g. Git, to simplify these procedures. Regular structures and infrastructures to document and process technical errors, bug reports as well as desired features are not necessarily established. Project management systems such as agile scrum methodologies are not common or applied. As painfully learned in a long history of software development and legacy code, lack of these practices lead to error-prone, non-transparent, non-reusable and non-transferable algorithms. Furthermore, no thorough test-ability is possible, leaving the results subject to human error and limiting the reliability of the outcome as well as the scope of possible complexity. Lack of incremental feedback to results and usability, as otherwise given within agile systems, allow long-term work efforts to possibly go in the wrong direction and production of large data-sets and algorithms to remain incomplete and faulty.

While from a short term perspective, establishing such globally standard practises and tools from the tech industry might appear like a large overhead in reference to the lifetime of time-limited, externally funded scientific projects, from a long term perspective these practices can be transferred to a multitude of projects, allowing better interaction between works with similar goals, integration of results, higher efficiency within the project lifetime and better, more reliable results.

4 Previous Work

Of course since 2014 until now the open availability to the Sentinel-1 data as well as the tools, processing engines and databases gave motivation to research and development in various fields concerning the deployment and analysis of SAR data. Relevant discoveries have been made in the importance and impact of using SAR data in remote sensing for crop growth analysis, as seen in Mandal et al. [Ma20] and recent publications from Salma et al. [SKD22] and Yadav et al. [Ya22]. Of course SAR has been applied in other fields as e.g. for rapid damage assessment after general disasters (Plank [Pl14]), in detecting small scale illegal mining (Janse van Rensburg; Kemp [JK22]) and detecting the scope of storm damage to orchards (De Petris et al. [De21]). Most relevant and demanded analysis-ready data (ARD) is frequently the normalized radar backscatter (γ^0 , Ticehurst et al. [Ti19]), eigenvector-based dual-polarisation decomposition vegetation index (Mandal et al. [Ma20], Mishra et al. [Mi22]) and interferometric coherence (Villarroya-Carpio et al. [VLE22], Villarroya-Carpio; Lopez-Sanchez [VL23]).

Never the less, most of the technical solutions for the derivation of these parameters remain solely in the scope of specific application scenarios and, besides in the work of Ticehurst et al. [Ti19], there are no open source solutions or insight to the processes.

There likewise, have been multiple approaches to providing pre-processed SAR ARD cubes, as previously noted from CODE-DE, by researchers in Australia (Ticehurst et al. [Ti19]) and in Switzerland (Chatenoux et al. [Ch21]) as well as from further commercial suppliers. However, in dependence from application field important pre-processing steps are missing, incompleteness in time and space, or non-availability for Germany.

All of this previous work has been done with the Sentinel-1 Toolboxes including SNAP. SNAP not only offers the operators for the mathematical calculations to derive the ARD but also a series of further operators including spacial fusion of the data. Although there has been one publication including the latter (Mandal et al. [Ma19]) with an introduction and insight to the workflow of deriving the dual pol covariance matrix with use of spacial operators, to our knowledge non of the other solutions include the spacial data fusion operators.

Despite all previous dedication to the topic, re-use of open source solutions to derive Sentinel-1 ARD still remains a time consuming and technical challenge.

5 Our Contribution

For this reason we propose two solutions for public use. One solution is given with a Sentinel-1 SAR pre-processing tool we developed in the Julius-Kuehn-Institute (JKI). This tool is designed to be user friendly, generalised for diverse use cases and offers a range of convenient functionality as well as optimised processing chains for the calculation of the normalized radar backscatter (γ nought, 0), eigenvector-based dual-polarization decomposition vegetation index and interferometric coherence. Below our workflow for the individual parameters (Figure 1). The Copernicus products we apply and process to derive ARD are the Level-1 Single Look Complex (SLC) in Interferometric Wide Swath (IW) Mode as standard for land observation applications and the Level-1 Ground Range Detected (GRD) in IW Mode.

The other solution is given by providing easy and open access to a data cube of pre-processed and pre-structured ARD in single data tile units of respectively 10x10 km² ground coverage for Germany.

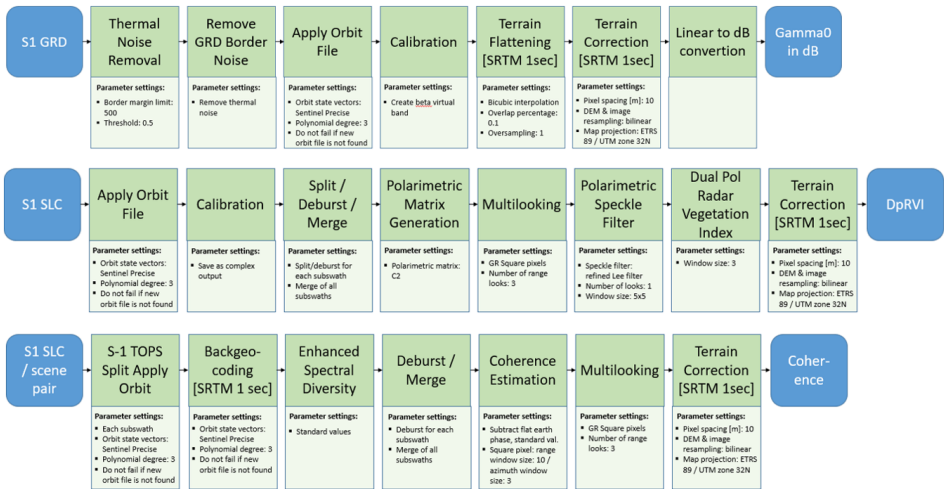


Fig. 1: Processing Workflow

5.1 The Sentinel-1 SAR Processing Tool Software Architecture

Besides with the deployment of Python, SNAP and CODE-DE, the processing tool was developed and tested with extensive use of JupyterLab, Gitea and Git as demonstrated in Figure 3. The tool is roughly designed with a modular Model View Controller (MVC) software architecture (Figure 2). All functional logic is contained in the “Controller” area, where all modules remain and can be easily extended. “Model” files to be accessed by the controller logic for processing and display as well as the SNAP graph operator files are held

separately with a clear interface to the controller. The programming is done in a procedural, object oriented manner with flat hierarchy. The internal communication is based on XML as the file type already introduced by the SNAP tool boxes and Sentinel-1 metadata files as visualised in Figure 4. The internal SNAP graph operator files are likewise held in a modular manner and can be easily adapted and extended. The “View” user interfacing and execution of the tool can be done over a minimal GUI or directly over SSH and any command line interface (Figure 4). The geo specific input and processed analysis ready data (ARD) output are held in the Geojson and Geotiff formats as common and widely supported formats. Since the installation of all necessary Python packages as well as the SNAP Toolboxes can be quite challenging and time consuming, a Dockerfile has been created as well as the necessary structure to create a Docker container wherein the processing can be performed.

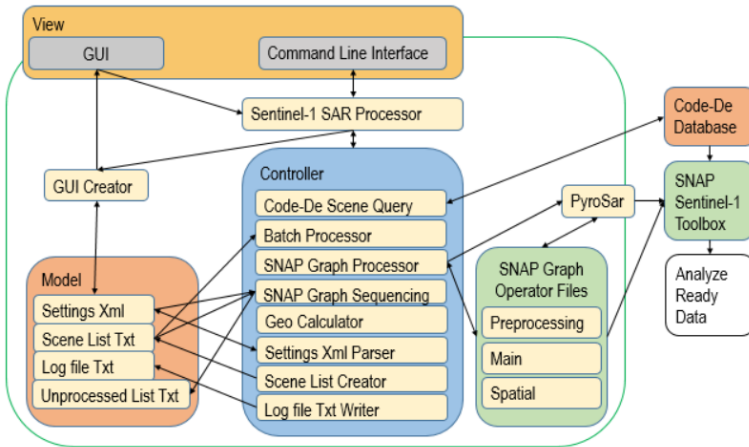


Fig. 2: Software Architecture

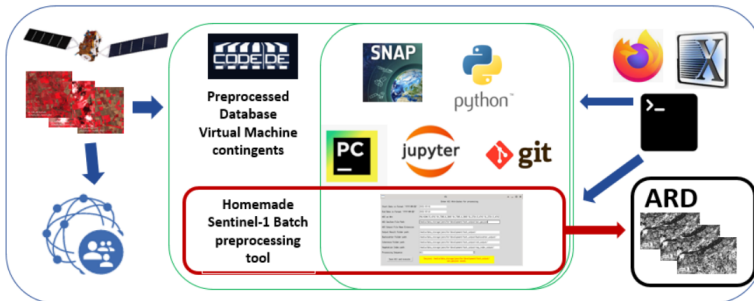


Fig. 3: Tools and Technologies

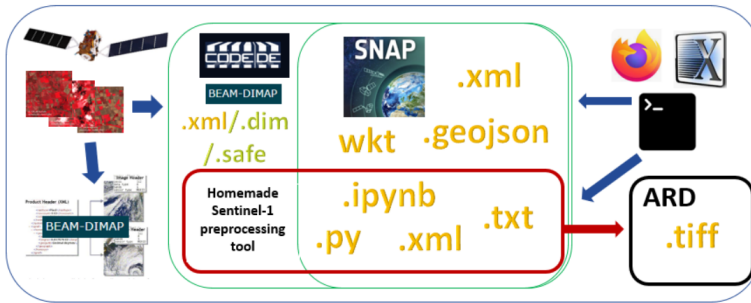


Fig. 4: Internal Formats

5.2 Our Main Focus

Complex processing performed on such voluminous data sets as the SAR data demands high computational power and storage. This is why the availability of VMs as offered by CODE-DE is essential for regular users to be able to generate ARD. Nevertheless, despite the high performance of the CODE-DE VMs at the JKI, processing times can surpass an acceptable maximum. The provided memory and internal cache size of the VMs can also become a bottleneck, additionally extending the processing times enormously. Due to this, one of our main points of focus was on optimising the performance of our workflows.

This we achieved with following solutions: 1. Deployment of full multiprocessing on all available kernels, 2. Temporary storage of intermediate results during the workflow, 3. Limiting the file size to small data units, resulting from the desired area of interest (AOI) for the SAR data, as early as possible. For 1. and 2. we used the already available functionality of the Python package PyroSar. For 3. we integrated the SNAP spacial operators with the general calculation operators in the SNAP graph XML structure. Here, our intention was to stage any spacial operator that reduces the size of the original scenes to the size of the AOI as early as possible. This is demonstrated in Figure 5 where the the induced spacial operators are marked in blue.

However, since we discovered that some of these SNAP operators can be prone to error in specific cases, depending on the geo-location and size of the desired AOI and are not well documented, we included an operation mode as feature of the tool, where some of this functionality can be deactivated. Processing times of specific periods and AOIs are displayed in Figures 6, 7 and 8. AOI “RLP1“ consists of 19.153,612 km² and is the upper half of Rhineland-Palatinate, the other “RLP2” of 21.843,920 km² the lower half of Rhineland-Palatinate, whereas a regular scene consists of approximately 47.783,981 km². Processing times fluctuate mainly because of the position of the processed scene in reference to the AOI.

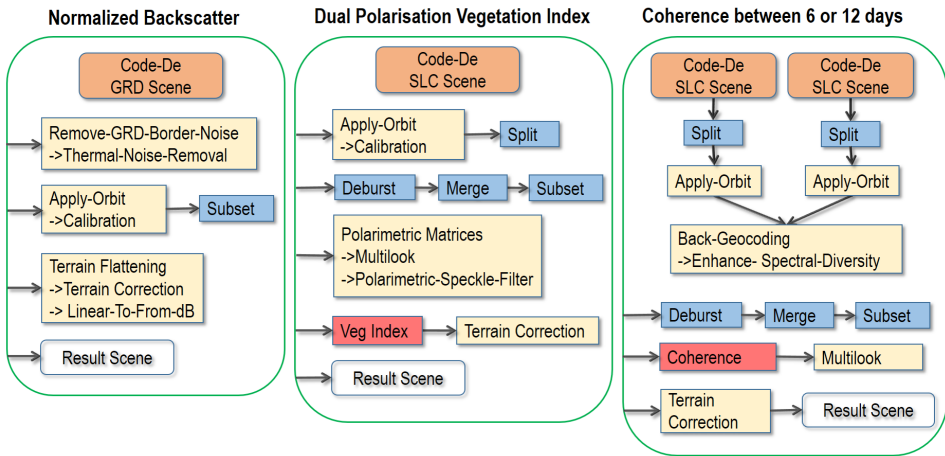


Fig. 5: Workflow with Induced Spatial Operators

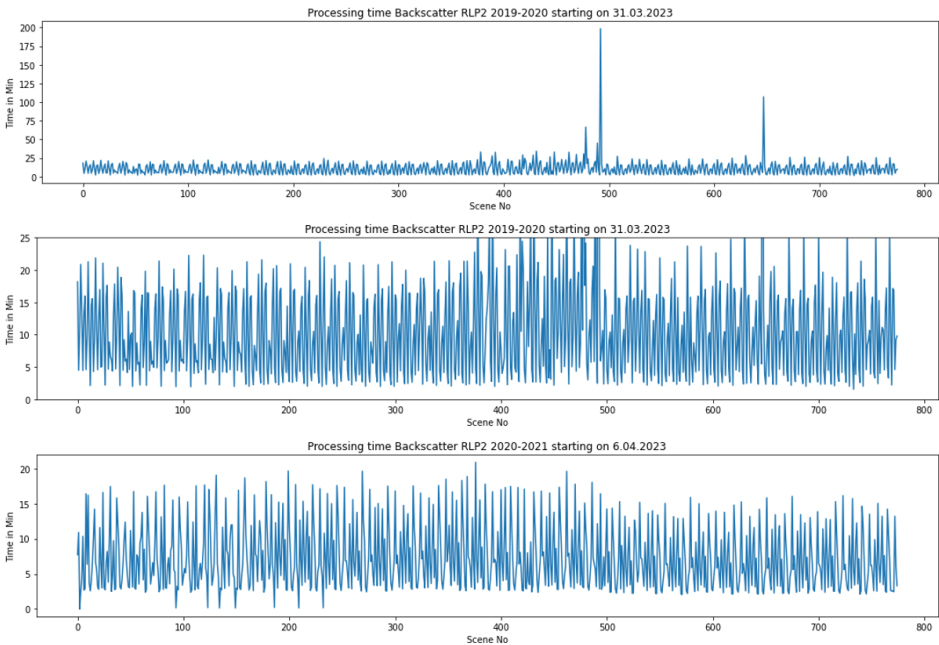


Fig. 6: Processing Times Backscatter

With PyroSar’s internally implemented multiprocessing the performance improved significantly (approx. 10-12 hours -> approx. 2-3 hours) in comparison to using the ESA SNAP/Python interface package “Snappy”. Likewise, with temporary storage of intermediate

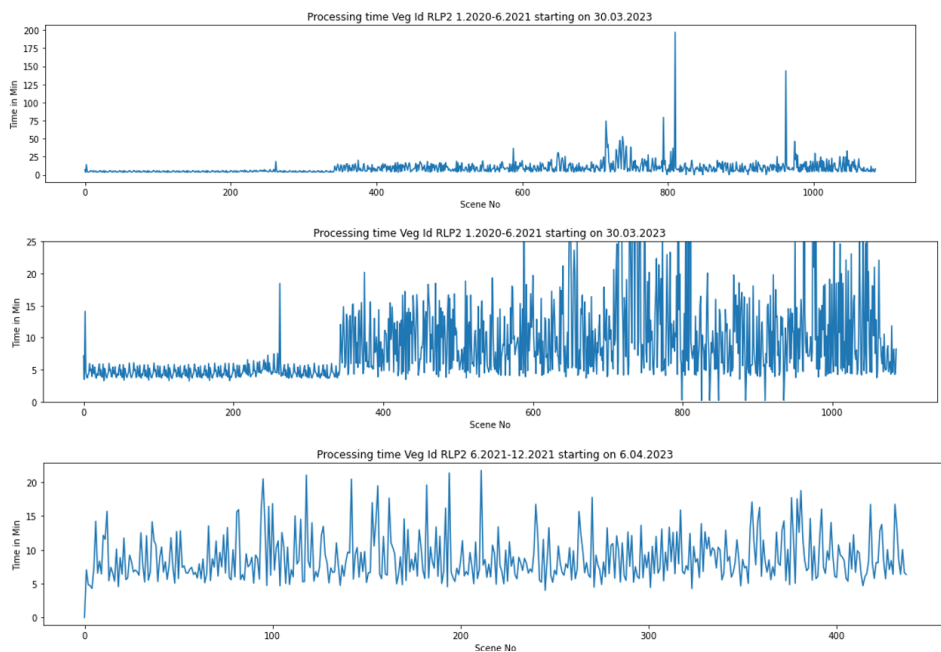


Fig. 7: Processing Times Dual Pol Vegetation Index

data the performance could be enhanced (approx. 2-3 hours -> approx. 16-25 minutes). By then reducing the intermediate file sizes at an early stage, we achieved even better performance (e.g. 16 minutes -> 1,33, scene to AOI in a 50:1 relationship in km² for backscatter processing). Curiously, when comparing processing times on VMs with different CPUs (16 kernels vs. 32 kernels), within the JKI CODE-DE infrastructure, no significant improvement was detected.

Another of our main points of focus was to enhance the precision of the ARD by analysing, adapting and extending the parameters of the workflow in comparison to the available ones we extracted and applied from previous work. For example when generating the backscatter output, for one, we use the ESA pre-processed GRD data, for the other, we integrated the Terrain-Flattening operator in the processing graph. Further adaptations of the parameters are displayed in Figure (1)

5.3 Challenges and Limitations

While developing the S1 SAR pre-processing tool and generating ARD for diverse observation requests and applications, we tried to vaguely apply an agile, user and outcome driven procedure. This however was only possible within the JKI given infrastructure and without

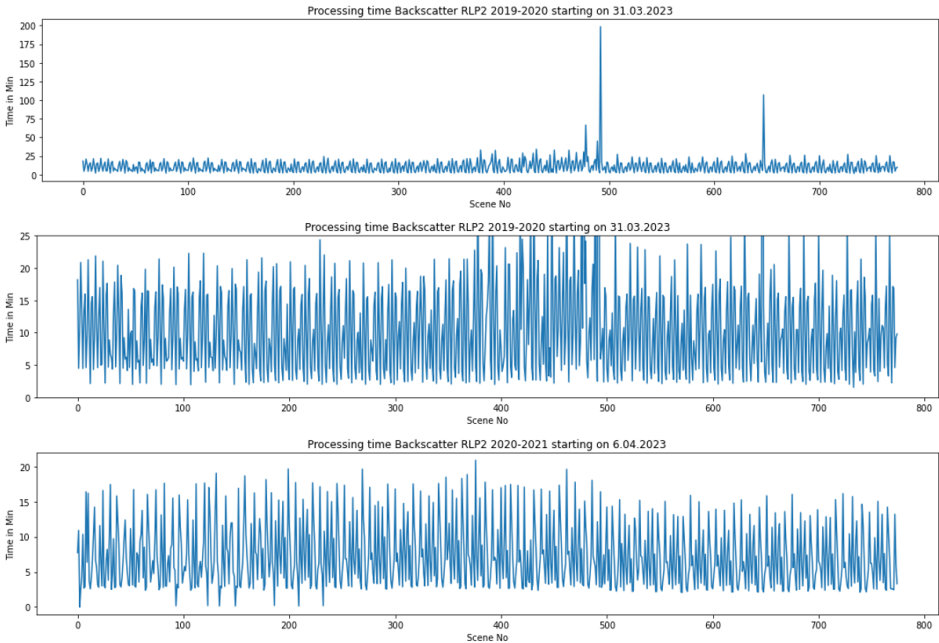


Fig. 8: Processing Times Coherence

an automatised testing framework. In evaluation of the available tools and packages for the S1 SAR data access, these partially appeared, as noted, due to the sometimes limited documentation and transparency of the internal functioning, leading to time-consuming trial-and-error attempts of combining the available features. Not only the ESA STEP forum for exchange on the use of the S1 Tools, but also the CODE-DE forum, with a small active community, still lacks content as a sufficient source of information to all relevant topics. The most thorough source of information is the open SNAP Project Development platform, however very developer specific and not adequate for general use. PyroSar is well documented and allows good insight to the internals. Despite this, the interface is, though simple and clear, narrow and not suitable for the direct execution of predefined SNAP graph XML files. For this we had to make adaptations. Furthermore, the maintenance of PyroSar in near future is not secured.

Whilst Copernicus offers a multitude of handbooks and reports to the satellite infrastructure and products, the exact behaviour of the satellites and structure of the resulting products can likewise only be determined by extensive use and observation. We also detected, that list of available Copernicus scenes over a specific area and time differs between the Sentinel-Hub and CODE-DE databases. When accessing the CODE-DE database to derive pre-processed ARD, a lot of pre-filtering is still necessary beforehand, e.g. the exclusion of duplicates, to receive a uniform data-set suitable for the SNAP graph operator workflow.

Due to the lack of precise documentation on how to integrate the SNAP graph spacial fusion operators, we limited the functionality of the tool to operating on an upper boundary AOI size of 200% of the original S1 scene size.

When executing the tool to process ARD the performance is not only dependant on our algorithms and applied tools but also on the performance of CODE-DE. We noticed that the performance varies strongly depending on time and day in-dependant of our JKI internal workload. At times the processing stagnates completely making the expected calculation times unpredictable which is seen by the outstanding maximum peaks in Figures 6 - 8.

In general, as in software development, insuring the completeness and correctness of the algorithms and, in this case, the resulting ARD without a testing infrastructure is a demanding task. Up until now we only randomly checked individual samples of the data for each individual processing sequence we performed. As a result we can not ensure completeness and correctness for all application scenarios and further testing and adaptations will be necessary.

6 Conclusion

Earth observation with SAR data and farming predictions with machine learning (ML) methods remain a promising solution to upcoming environmental changes. These solutions should remain in the public domain. Despite a lot of efforts in the deployment and provision of SAR data relevant technologies, reliability on the outcome remains vague. Additionally to the challenge of acquiring adequate reference data for ML, which is limited by the lack of trust in sharing data and agreements on standards in diverse fields, the topic of open source SAR ARD remains open. More work must be done and collaborative efforts performed to insure the prerequisite of a continuously up to date and world wide, complete and ready-to-use database of SAR data. This will be necessary to be able to divert focus more strongly toward the ML analysis. Not only in the scope of agriculture, but in all scientific areas of earth observation by means of SAR data.


References

- [Ch21] Chatenoux, B.; Rööslü, C.; Wingate, V.; Poussin, C.; Rodila, D.; Peduzzi, P.; Steinmeier, C.; Ginzler, C.; Psomas, A.; Schaeppman, M.; Giuliani, G.: The Swiss data cube, analysis ready data archive using earth observations of Switzerland. *Scientific Data* 8/, p. 295, Nov. 2021.
- [De21] De Petris, S.; Sarvia, F.; Gullino, M.; Tarantino, E.; Borgogno-Mondino, E.: Sentinel-1 Polarimetry to Map Apple Orchard Damage after a Storm. *Remote Sensing* 13/5, 2021, ISSN: 2072-4292, URL: <https://www.mdpi.com/2072-4292/13/5/1030>.

- [JK22] Janse van Rensburg, G.; Kemp, J.: The Use of C-Band and X-Band SAR with Machine Learning for Detecting Small-Scale Mining. *Remote Sensing* 14/4, 2022, ISSN: 2072-4292, URL: <https://www.mdpi.com/2072-4292/14/4/977>.
- [Ma19] Mandal, D.; Vaka, D. S.; Bhogapurapu, N.; Vanama, V. S. K.; Kumar, V.; Rao, Y.; Bhattacharya, A.: Sentinel-1 SLC Preprocessing Workflow for Polarimetric Applications: A Generic Practice for Generating Dual-pol Covariance Matrix Elements in SNAP S-1 Toolbox./, Nov. 2019.
- [Ma20] Mandal, D.; Kumar, V.; Ratha, D.; Dey, S.; Bhattacharya, A.; Lopez-Sanchez, J. M.; McNairn, H.; Rao, Y. S.: Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sensing of Environment* 247/, p. 111954, 2020, ISSN: 0034-4257, URL: <https://www.sciencedirect.com/science/article/pii/S0034425720303242>.
- [Mi22] Mishra, D.; Pathak, G.; Singh, B. P.; Mohit; Sihag, P.; Rajeev; Singh, K.; Singh, S.: Crop classification by using dual-pol SAR vegetation indices derived from Sentinel-1 SAR-C data. *Environmental Monitoring and Assessment* 195/1, p. 115, Nov. 2022, ISSN: 1573-2959, URL: <https://doi.org/10.1007/s10661-022-10591-x>.
- [PI14] Plank, S.: Rapid Damage Assessment by Means of Multi-Temporal SAR — A Comprehensive Review and Outlook to Sentinel-1. *Remote Sensing* 6/6, pp. 4870–4906, 2014, ISSN: 2072-4292, URL: <https://www.mdpi.com/2072-4292/6/6/4870>.
- [SKD22] Salma, S.; Keerthana, N.; Dodamani, B.: Target decomposition using dual-polarization sentinel-1 SAR data: Study on crop growth analysis. *Remote Sensing Applications: Society and Environment* 28/, p. 100854, 2022, ISSN: 2352-9385, URL: <https://www.sciencedirect.com/science/article/pii/S2352938522001628>.
- [Ti19] Ticehurst, C.; Zhou, Z.-S.; Lehmann, E.; Yuan, F.; Thankappan, M.; Rosenqvist, A.; Lewis, B.; Paget, M.: Building a SAR-Enabled Data Cube Capability in Australia Using SAR Analysis Ready Data. *Data* 4/3, 2019, ISSN: 2306-5729, URL: <https://www.mdpi.com/2306-5729/4/3/100>.
- [VL23] Villarroya-Carpio, A.; Lopez-Sanchez, J. M.: Multi-Annual Evaluation of Time Series of Sentinel-1 Interferometric Coherence as a Tool for Crop Monitoring. *Sensors* 23/4, 2023, ISSN: 1424-8220, URL: <https://www.mdpi.com/1424-8220/23/4/1833>.
- [VLE22] Villarroya-Carpio, A.; Lopez-Sanchez, J. M.; Engdahl, M. E.: Sentinel-1 interferometric coherence as a vegetation index for agriculture. *Remote Sensing of Environment* 280/, p. 113208, 2022, ISSN: 0034-4257, URL: <https://www.sciencedirect.com/science/article/pii/S0034425722003169>.

- [Ya22] Yadav, V. P.; Prasad, R.; Bala, R.; Srivastava, P. K.; Vanama, V. S. K.: Appraisal of dual polarimetric radar vegetation index in first order microwave scattering algorithm using sentinel – 1A (C - band) and ALOS - 2 (L - band) SAR data. *Geocarto International* 37/21, pp. 6232–6250, 2022, eprint: <https://doi.org/10.1080/10106049.2021.1933209>, URL: <https://doi.org/10.1080/10106049.2021.1933209>.

Informationsbedarf von Bürger*innen bezüglich Niedrigwasser, Trockenheit und Dürre - Ergebnisse einer Umfrage

Eva Wein¹, Fabian Bressel², Julian Berndt³, Hellen Siewert⁴, Rosemarie Bähne⁵, René Krüger⁶, Frank Fuchs-Kittowski ⁷

Abstract: In den letzten Jahren gab es einen deutlichen Rückgang der Niederschläge in Deutschland. Niedrigwasser, Trockenheit und Dürre können weitreichende negative Auswirkungen auf die Umwelt, Wirtschaft und Gesellschaft haben. Um den Informationsbedarf zum Thema Niedrigwasser, Trockenheit und Dürre innerhalb der Bevölkerung zu ermitteln, wurde eine Online-Befragung mit 450 Teilnehmenden durchgeführt. Aus den Ergebnissen werden Empfehlungen abgeleitet, um passende Informationsangebote zur Sensibilisierung und Bewusstseinsbildung der Bevölkerung zum Thema Niedrigwasser, Trockenheit und Dürre zu entwickeln.

Keywords: Umfrage, Niedrigwasser, Trockenheit, Informationsbedarf, Bürger*innen

1 Einleitung

Die Sommer 2018, 2019 2020 und 2022 zeigten die Verwundbarkeit vieler Regionen Deutschlands gegenüber Niedrigwasser, Trockenheit und Dürre. Laut dem Deutschen Wetterdienst gab es in diesen Jahren einen starken Rückgang der Niederschläge im Vergleich zum langjährigen Durchschnitt [Bi22, Um23]. Diese Phänomene haben weitreichende Auswirkungen auf die Umwelt, Wirtschaft und Gesellschaft. Sie können die Verfügbarkeit von Trinkwasser verringern, Ernteausfälle, Waldbrände und andere Beeinträchtigungen des Ökosystems verursachen. In einigen Orten Deutschlands gab es bereits Engpässe in der Wasserverfügbarkeit [Ri21, Um23].

Die Ergebnisse einer repräsentativen Umfrage des Instituts Civey im Auftrag der Eon Stiftung haben gezeigt, dass die Besorgnis der Bevölkerung bezüglich zukünftiger Wetterereignisse steigt [Wd22]. Dazu ist das Informationsangebot zum Thema Niedrigwasser, Trockenheit und Dürre bislang eher spärlich. Um den Informationsbedarf der Bürger*innen zu ermitteln, wurde eine Online-Befragung durchgeführt. Ziel der Umfrage war es, Handlungsempfehlungen für passende Informationsangebote für eine bessere Bewusstseinsbildung der Bevölkerung zum Thema Niedrigwasser, Trockenheit

Hochschule für Technik und Wirtschaft Berlin, Umweltinformatik, Wilhelminenhofstraße 75A, 12459 Berlin, {¹eva.wein, ²fabian.bressel, ³julian.berndt, ⁴hellen.siewert, ⁵rosemarie.baehne, ⁶rene.krueger}@student.htw-berlin.de

⁷ Hochschule für Technik und Wirtschaft Berlin, Wilhelminenhofstraße 75A, 12459 Berlin, frank.fuchs-

kittowski@htw-berlin.de,  <https://orcid.org/0000-0002-5445-3764>

und Dürre zu entwickeln. Die Umfrage umfasste 25 Fragen zur Ursache, Auswirkung, Handlungsoptionen und bevorzugten Informationsaufbereitung. Die Ergebnisse der Umfrage, an der sich 450 Personen beteiligten, werden in diesem Beitrag erläutert.

Dieser Beitrag ist in sieben Kapitel strukturiert, die den aktuellen Forschungsstand, die Forschungsfragen und Hypothesen (Kapitel 2), die Methodik und das Vorgehen (Kapitel 3), die Ergebnisse der Umfrage (Kapitel 4), die Diskussion der Ergebnisse (Kapitel 5), die Schlussfolgerungen (Kapitel 6) sowie eine Zusammenfassung und einen Ausblick (Kapitel 7) präsentieren.

2 Stand der Forschung, Forschungsfragen und Hypothesen

Zum aktuellen Zeitpunkt gibt es für Deutschland keine Umfragen oder Studien, welche sich mit dem Informationsbedarf oder Informationsstand der Bevölkerung zum Thema Niedrigwasser, Trockenheit und Dürre befassen.

Das ein Handlungsbedarf besteht, zeichnet sich bereits aus den beschlossenen Maßnahmen der Bundes- und Landesregierungen zur Bildung und Informierung der Bevölkerung ab [Bu23, Mi21]. So wird in der nationalen Wasserstrategie 2021 und 2023 ein Maßnahmenblock der Stärkung des Bewusstseins für die Ressource Wasser gewidmet. Diese Maßnahmen umfassen Schulungen und Bildungsangebote für viele verschiedene Zielgruppen [Bu23]. Land Brandenburg hat im Zuge des Landesniedrigwasserkonzept Brandenburg ebenfalls ein Informationsangebot zu Niedrigwasser in der Region geschaffen. Dieses ist sehr technisch und erfordert ein gewisses Grundwissen der Thematik, um das Dargestellte zu verstehen [Mi21].

Vom Umweltbundesamt wurde 2021 bereits ein umfassendes Dokument zu Niedrigwasser, Trockenheit und Dürre veröffentlicht [Ri21], dieses fasst den aktuellen Stand detailliert zusammen unter Berücksichtigung der verschiedenen Flussgebiete und dem Einfluss auf verschiedene Sektoren (Landwirtschaft, Energieversorgung, etc.). Aufgrund der Komplexität und Länge des Dokumentes stellt es jedoch, wie viele Publikationen zu dieser Thematik, keine geeignete Informationsquelle für Bürger*innen dar.

Eine Studie zum Umweltbewusstsein in Deutschland wurde bereits 2018 und 2020 vom Umweltbundesamt durchgeführt. In der Studie von 2020 gaben 60% der befragten Personen an, sich sehr gut oder gut zu dem Thema „Klimawandel und Klimaschutz“ informiert zu fühlen. In dieser Studie ist gut erkennbar, dass es Interesse an der Umwelt und dem Umweltschutz gibt. Diese Befragung ist allgemein auf die Umwelt und den Umweltschutz ausgerichtet und lässt wenig Rückschlüsse auf das Bewusstsein für konkrete Probleme zu; insbesondere auf den Informationsbedarf zum Thema Niedrigwasser, Trockenheit und Dürre [WBG19, WG22].

Aus diesen Publikationen ist erkennbar, dass das Thema Niedrigwasser, Trockenheit und Dürre bereits in der Wissenschaft aufgearbeitet und von behördlicher Seite als Problem anerkannt wird. Nicht erarbeitet hingegen ist, welches Bewusstsein und welchen Informationsbedarf es in der Bevölkerung zu diesem Thema gibt.

Um das Bewusstsein und den Informationsbedarf der Bevölkerung zu ermitteln, wurden *Forschungsfragen*, orientiert an der Lasswell-Formel aus der Kommunikationswissenschaft, aufgestellt [La48]. Inhaltlich wurden die Fragen in fünf Fragenkomplexe gegliedert: „Wer sagt was, in welchem Kanal, zu wem mit welchem Effekt?“ Außerdem wurde anhand der bereits existierenden Studien zum Thema Niedrigwasser, Trockenheit und Dürre und der noch offenen Fragen bezüglich des Bewusstseins und des Informationsbedarfs der Bevölkerung *Hypothesen* entwickelt, die mithilfe der Umfrage überprüft werden sollen. Dadurch können konkrete *Handlungsempfehlungen* für ein zukünftiges Informationsangebot bereitgestellt werden, welches die Interessen und Wünsche der Bevölkerung berücksichtigt. Die folgende Tabelle 1 bietet einen Überblick über die Forschungsfragen und die damit in Verbindung stehenden Themen und Hypothesen. Die Originalreihenfolge der Lasswells-Formel wurde für eine bessere inhaltliche Fragenführung verändert.

Forschungsfrage	Thema	Hypothesen
An wen sollten die Informationen adressiert werden?	Informationsstand und Betroffenheit	H1a: Die Bevölkerung ist nicht über Niedrigwasser, Trockenheit und Dürre informiert.
		H1b: Je direkter die Bevölkerung betroffen ist, desto höher ist der Informationsbedarf in der Bevölkerung.
Was für Information möchten die Adressaten erhalten?	Inhaltlichen Interessenlage	H2a: Das Interesse der Bevölkerung ist am größten für Auswirkungen von Niedrigwasser, Trockenheit und Dürre auf sich selbst und auf die Umwelt.
		H2b: Rezipienten, die an den Auswirkungen von Niedrigwasser, Trockenheit und Dürre auf sich selbst interessiert sind, haben ein großes Interesse an Informationen zu den Themen „Reglementierung des Wasserverbrauchs“ und „Folgen für Wirtschaft und Gesellschaft“.

		H2c: Die Bevölkerung benötigt Informationen aus verschiedenen Bereichen, um mit Niedrigwasser, Trockenheit und Dürre besser umgehen zu können.
Welchen Effekt wünschen sich die Adressaten von der Information?	Motivation	H3a: Die Bevölkerung erhofft sich von einem Informationsangebot zum Thema Niedrigwasser, Trockenheit und Dürre, Handlungsmöglichkeiten für den Alltag ableiten und selbstständig ergreifen zu können.
Wer sollte die Informationen bereitstellen?	Vertrauenswürdigkeit der Quellen	H4a: Die Bevölkerung vertraut vorrangig Informationen, die von Forschungs- und Bildungseinrichtungen oder von Behörden und Amtsträger*innen bereitgestellt werden.
		H4b: Der Bevölkerung mit einem höheren akademischen Grad ist die Nachvollziehbarkeit der Methodik und die Zugänglichkeit der zugrundeliegenden Daten der bereitgestellten Informationen wichtiger als der Bevölkerung mit einem niedrigeren akademischen Grad.
Über welchen Kanal sollten die Informationen kommuniziert werden?	Bevorzugte Kommunikationsart	H5a: Die Bevölkerung möchte zukünftig vorwiegend über Online-Kanäle wie Webseiten und Apps über das Thema Niedrigwasser, Trockenheit und Dürre informiert werden.
		H5b: Altersgruppen von 0-39 bevorzugen die Kanäle Webseiten und Apps, wohingegen Altersgruppen von 40-70 und älter die Kanäle Printmedien und Fernsehen bevorzugen.
		H5c: Das bevorzugte Format der Informationsaufbereitung sind Texte und Grafiken.

Tab. 1: Forschungsfragen, zugehöriges Thema und Hypothesen

3 Methode und Vorgehen

Bei der Erstellung der Arbeit wurde sich an den von R. Schnell [SHE18] vorgeschlagenen Arbeitsschritten für ein empirisches Forschungsprojekt orientiert.

Forschungsgegenstand der Studie war das Informationsbedürfnis der deutschen Bevölkerung zum Thema Niedrigwasser, Trockenheit und Dürre, was durch die zunehmende Betroffenheit durch Niedrigwasser-, Trockenheit- und Dürre-Ereignisse und die aktuell lückenhafte Forschungslage zum Thema motiviert ist. Einige der in der Phase der Theoriebildung verwendeten Quellen wurden bereits im Kapitel 2 vorgestellt. Bei der *Konzeptspezifikation* wurde sich besonders damit auseinandergesetzt, was unter dem Theorie-Begriff „Umweltbewusstsein“ zu verstehen ist.

Im Rahmen des Schrittes *Operationalisierung* wurde sich für einen Fragebogen entschieden. Die Erstellung beinhaltete auch Pretests (Voruntersuchungen), um herauszufinden, ob mit den erstellten Fragebögen valide Messungen möglich sind. Zusätzlich wurden Expert*innen für inhaltliches Feedback einbezogen. Im Nachgang wurden die Fragen dem Feedback entsprechend angepasst. Der Fragebogen bestand aus insgesamt 26 Fragen. Davon waren 20 inhaltliche Fragen, fünf soziodemografische Fragen und eine offene Frage für Feedback. Inhaltlich wurden die Fragen in fünf Fragenkomplexe gegliedert, orientiert an den Forschungsfragen (siehe Kapitel 2).

Für das *Forschungsdesign* und die Untersuchungsobjekte wurde Folgendes festgelegt: Es handelt sich um eine einmalig durchgeführte Online-Befragung, welche sich an in Deutschland lebende Menschen richtet, ohne diese Zielgruppe weiter einzuschränken. Es wurde mit einer Beteiligung von bis zu 100 Personen gerechnet.

Zur *Durchführung* der Online-Befragung wurde das Tool „SurveyMonkey“ gewählt. Für die Auswahl des Umfragetools wurde eine Auswertung nach GESIS-Ansatz des Leibniz-Instituts für Sozialwissenschaften durchgeführt. Diese kam zu dem Ergebnis, dass das Preis-Leistungs-Verhältnis bzw. Preis-Produktumfang-Verhältnis von dem Werkzeug „SurveyMonkey“ am besten ausgeschöpft wird.

In der Phase der *Datenerhebung* wurde das Layout des Online-Tools „SurveyMonkey“ auf Übersichtlichkeit geprüft (Überschriften, Abstände, Hervorhebungen, Hinweise) und die technische Funktionalität (Links, Auswahlmöglichkeiten, Absenden) sichergestellt. Die Umfrage sollte für circa 7 Wochen online beantwortet werden können. Bei der *Datenerfassung* ließen sich einige Ergebnisse direkt mit Hilfe des Umfrage-Tools Survey-Monkey auslesen. Zusätzlich wurde eine bereinigte CSV-Datei erstellt, welche zur Auswertung mittels Structured Query Language in eine Datenbank geladen wurde.

Die *Datenanalyse* zur Überprüfung der Hypothesen ist in den Kapiteln „Ergebnisse“ und „Diskussion“ nachzulesen.

4 Ergebnisse

Die Online-Umfrage war vom 23.01.23 bis zum 14.03.23 auf der Plattform „SurveyMonkey“ verfügbar und es haben 450 Personen teilgenommen. Da jede Frage übersprungen werden konnte, schwankt die Anzahl der Beantwortungen pro Frage zwischen 384 und 450. Die relativen Angaben in den folgenden Abschnitten beziehen sich auf die tatsächlichen Beantwortungen und nicht auf die Gesamtanzahl der Teilnehmenden. Einen Sonderfall bildet die letzte Frage, bei die Teilnehmenden in einem Freitextfeld Kommentare und Anmerkungen zur Umfrage hinterlassen konnten. Diese Möglichkeit wurde von 37 Personen genutzt.

Aus den soziodemografischen Ergebnissen der Umfrage ging hervor, dass das Geschlecht, sowie das Alter der Befragten repräsentativ für die Gesamtbevölkerung sind. Im Hinblick auf den Bildungsgrad gaben jedoch 76% an, einen Hochschulabschluss zu besitzen und der Wohnort der Befragten bündelte sich stark auf die Bundesländer Berlin und Brandenburg (50% der Teilnehmenden). Aufgrund der insgesamt hohen Anzahl an Teilnehmenden der Umfrage können dennoch wichtige Erkenntnisse, hinsichtlich der Ausgestaltung eines Informationsangebots zu Niedrigwasser, Trockenheit und Dürre, abgeleitet werden.

Nachfolgend werden die Ergebnisse einer Auswahl an Fragen vorgestellt. Die Auswahl basiert auf den wichtigen Erkenntnissen im Hinblick auf die Validierung der Hypothesen. Die Nummerierung der Fragen entspricht nicht dem Fragebogen und dient der vereinfachten Referenzierung im Kapitel „Diskussion“.

4.1 Frage 1: In welchem Umfang fühlen Sie sich von Niedrigwasser, Trockenheit und Dürre betroffen?

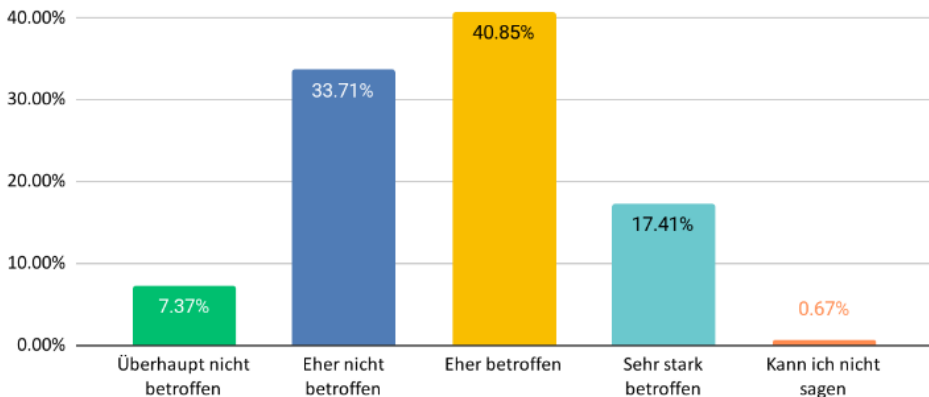


Abb. 1: Einschätzung der Betroffenheit durch Niedrigwasser, Trockenheit und Dürre

Diese Frage erweitert das Feld der Betroffenheit mit dem Umfang, in welchem sich die Befragten mit dem Thema konfrontiert sehen. Ein größerer Teil der Befragten fühlt sich von den Auswirkungen von Niedrigwasser, Trockenheit und Dürre eher betroffen (41 %), 17 % fühlen sich sehr stark betroffen. Ein weiterer großer Teil schätzt sich als eher nicht betroffen ein (34 %) und 7 % als gar nicht betroffen. Weniger als 1 % äußern sich nicht.

4.2 Frage 2: Wie gut fühlen Sie sich über Niedrigwasser, Trockenheit und Dürre informiert?

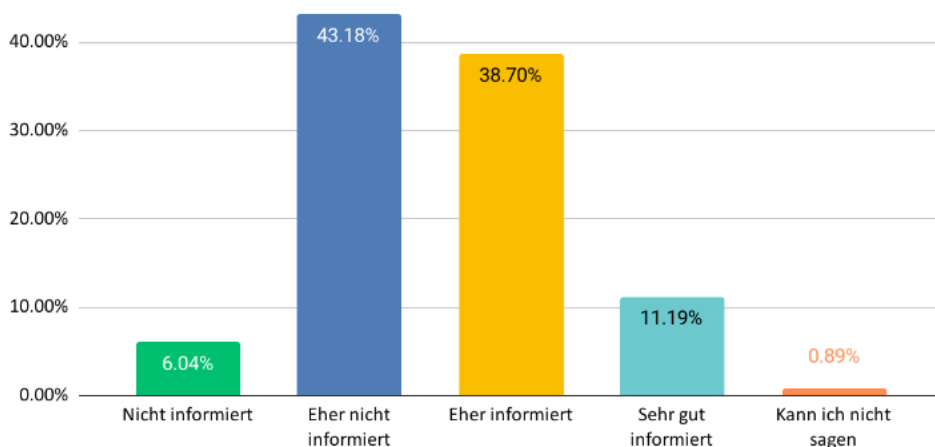


Abb. 2: Einschätzung der Informiertheit über Niedrigwasser, Trockenheit und Dürre

Hierbei wurde der Informationsstand der Befragten thematisiert. Die meisten Personen finden, dass sie über Niedrigwasser, Trockenheit und Dürre eher nicht informiert (43%) oder eher informiert (39 %) sind. Die jeweiligen Extreme wurden von weniger Personen ausgewählt: 6 % fühlen sich nicht informiert und 11 % sehr gut informiert. 0,89 % haben dazu keine Aussage getroffen.

4.3 Frage 3: Über welche Auswirkungen von Niedrigwasser, Trockenheit und Dürre möchten Sie informiert werden?

In dieser Frage ging es darum, über welche Auswirkungen von Niedrigwasser, Trockenheit und Dürre die Befragten informiert werden wollen. Es konnten mehrere Antwortoptionen ausgewählt werden. 85 % der Befragten gaben an, über Auswirkungen auf sich selbst (z.B. Bewässerung des Gartens, lokale Wasserversorgung) informiert werden zu wollen. 89 % wollen über Auswirkungen auf die Umwelt (z. B. Waldbrände, Fischsterben) informiert werden und 51 % über Auswirkungen auf die Wirtschaft. 4 % möchten nicht über Auswirkungen informiert werden.

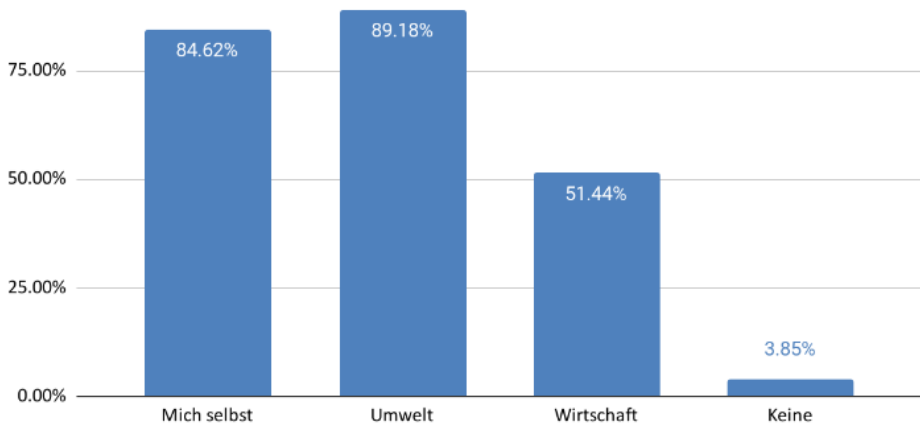


Abb. 3: Präferenzen für Informationen über Auswirkungen von Niedrigwasser, Trockenheit und Dürre

4.4 Frage 4: Informationen zu den Ursachen und Auswirkungen von Niedrigwasser, Trockenheit und Dürre interessieren mich in folgendem Ausmaß

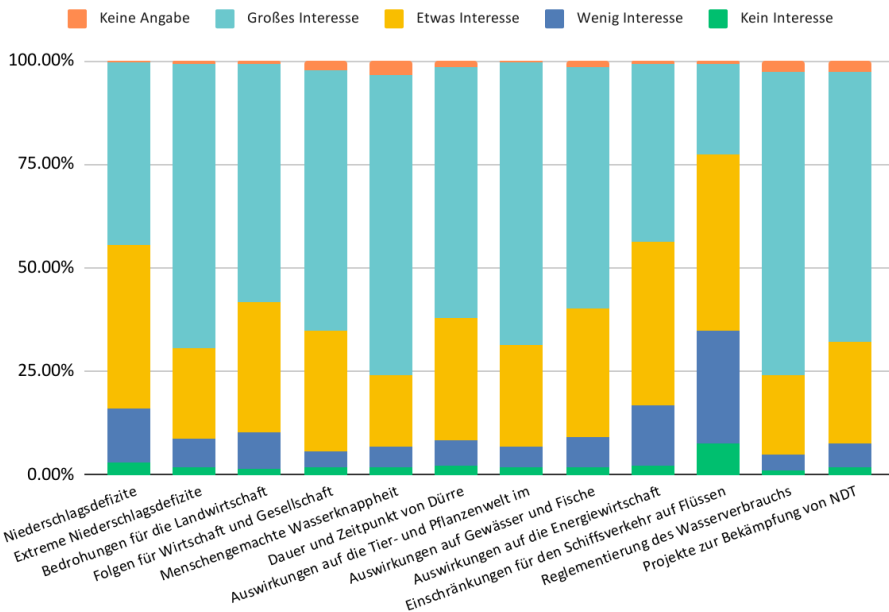


Abb. 4: Interesse an Informationen zu Ursachen und Auswirkungen von Niedrigwasser, Trockenheit und Dürre

Bei dieser Frage sollten die Teilnehmenden beantworten, welche Ursachen und Folgen von Niedrigwasser, Trockenheit und Dürre sie sehr interessieren. Die Frage war in 12 Unterthemen aufgeteilt. Pro Unterthema konnte eine Antwortoption (1= Kein Interesse, 2= Wenig Interesse, 3= Etwas Interesse, 4= Großes Interesse) ausgewählt werden. Bei allen Unterthemen hat ein Großteil der Teilnehmenden großes oder etwas Interesse an Informationen angegeben. Das größte Interesse besteht dabei mit einem gewichteten Mittelwert von 3,7 an den Themen „Reglementierung des Wasserverbrauchs“ und „Menschengemachte Wasserknappheit“. Nach Angaben der Rezipienten ist das Interesse an dem Thema „Einschränkungen für den Schiffsverkehr auf Flüssen“ mit einem gewichteten Mittelwert von 2,8 am geringsten.

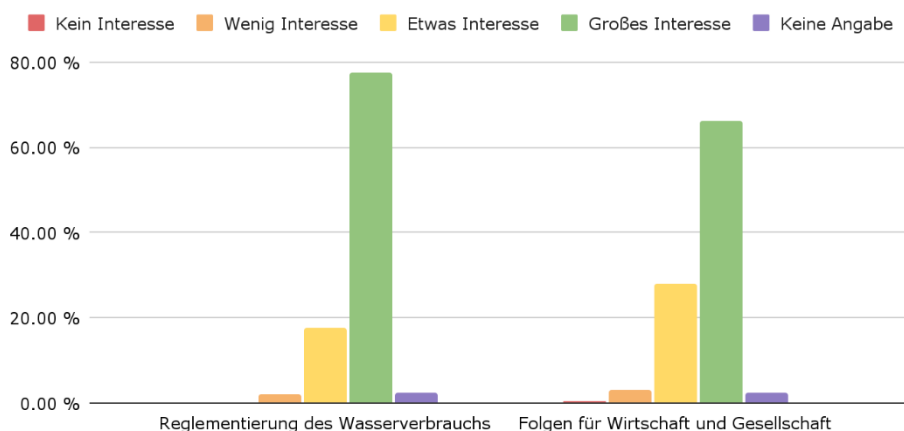


Abb. 5: Antworten aus Frage 4, welche bei Frage 3 „Auswirkungen auf mich selbst“ wählten

Die zwei Antworten „Reglementierung des Wasserverbrauchs“ und „Folgen für die Wirtschaft“ aus Frage 4 wurden noch einmal separat betrachtet, wobei nur Antworten von Teilnehmenden betrachtet wurden, welche in Frage 3 „Auswirkungen auf mich selbst“ wählten. 78 % haben hierbei großes Interesse an der „Reglementierung des Wasserverbrauchs“ und 68 % Interesse an „Folgen für die Wirtschaft“. Weniger als 3 % der Teilnehmenden gaben in beiden Fällen an wenig oder kein Interesse an den Themen zu haben.

4.5 Frage 5: Welche Informationen würden Sie benötigen, um mit Niedrigwasser, Trockenheit und Dürre besser umgehen zu können?

Bei dieser Frage gaben 65 % der Befragten an, dass ihnen „Informationen zur Landschaftsgestaltung“ helfen würden, besser mit Niedrigwasser, Trockenheit und Dürre umgehen zu können. Die anderen Antwortmöglichkeiten lagen sehr nah beieinander mit

ca. 50 % der Befragten. So gaben 55 % an, dass ihnen „Gesundheitsinformationen“ helfen würden, 53 % „Informationen zum Wassersparen“, 52 % „Informationen zur Unterstützung lokaler Behörden und Organisationen im Bereich Wassermanagement“, 51 % „Kommunikationskanäle für up-to-date Informationen“ und jeweils 49 % „Informationen zur Vorsorge“ und „Informationen zu aktuellen Restriktionen und ihrer Konsequenzen bei Nichteinhaltung“. Etwas abgeschlagen sind „Informationen zu Möglichkeiten der finanziellen Unterstützung“, mit 26 % und stark abgeschlagen sind die Befragten, die keine der genannten Informationen benötigen, mit 2 % .

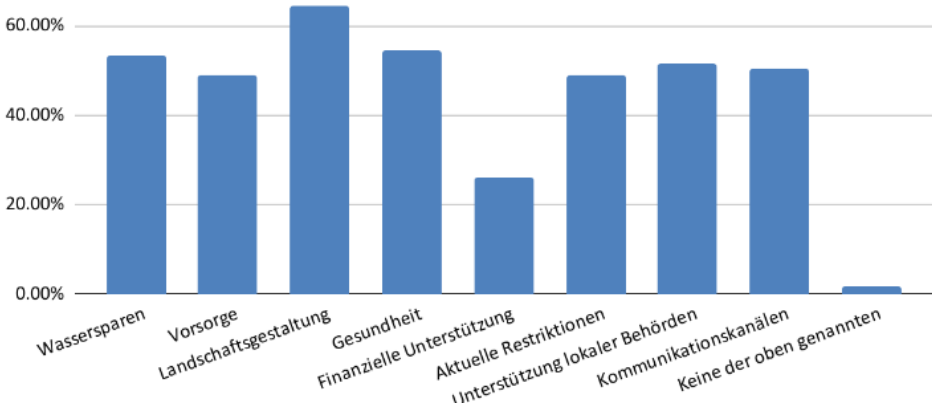


Abb. 6: Benötigte Informationen für einen besseren Umgang mit Niedrigwasser, Trockenheit und Dürre

4.6 Frage 6: Was erhoffen Sie sich von einem Informationsangebot über Niedrigwasser, Trockenheit und Dürre?

Hierbei wurde erfragt, was sich die Befragten persönlich von einem Informationsangebot erhoffen. Die Teilnehmenden konnten aus einer 4-stufigen Likertskala wählen (1= Stimme überhaupt nicht zu, 2= Stimme eher nicht zu, 3=Stimme eher zu, 4= Stimme voll und ganz zu). Dabei stimmte die große Mehrheit den gegebenen Optionen zu. Am relevantesten ist das Ableiten von Handlungsmöglichkeiten (gew. Mittelwert: 3,61) während mehr über das Thema zu erfahren und das Interesse zu decken die geringste Zustimmung erhält (3,33).

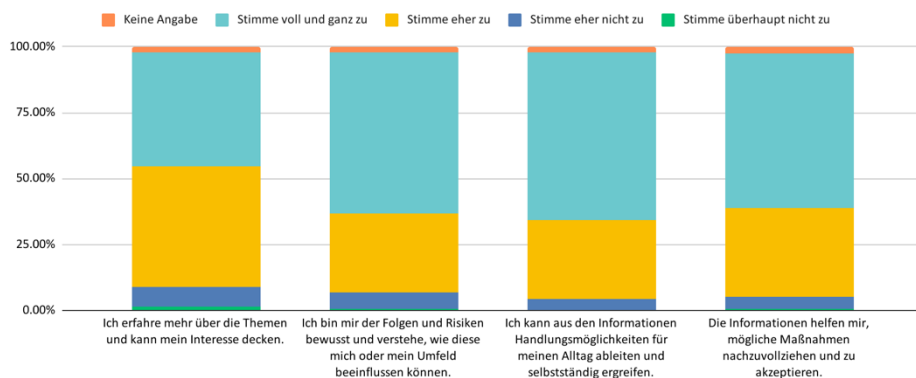


Abb. 7: Erwartungen an ein Informationsangebot zu Niedrigwasser, Trockenheit und Dürre

4.7 Frage 7: Wie sehr vertrauen Sie den folgenden Informationsquellen zu den Themen Niedrigwasser, Trockenheit und Dürre?

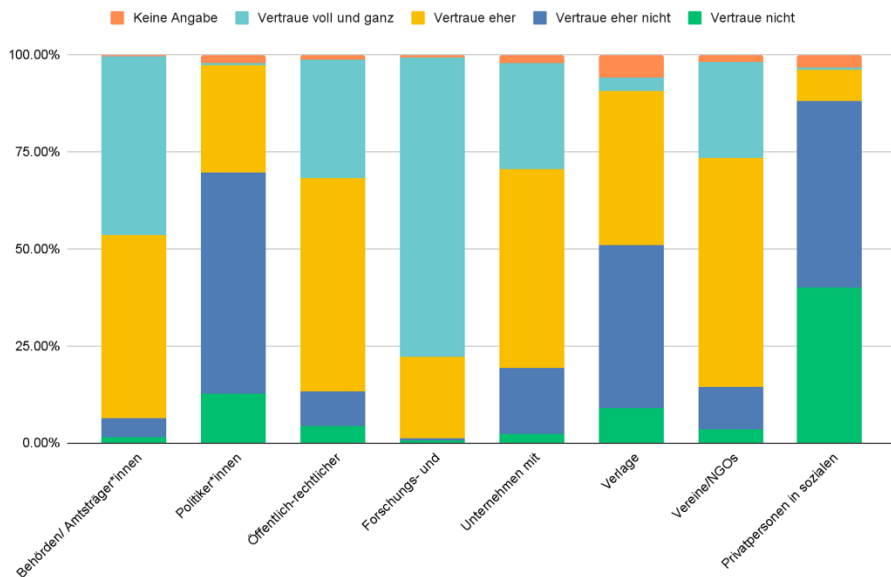


Abb. 8: Vertrauen in Informationsquellen zu Niedrigwasser, Trockenheit und Dürre

Die Frage greift das Vertrauen der Befragten in verschiedene Informationsquellen auf. Das Vertrauen der Befragten in die zur Auswahl stehenden Informationsquellen variiert

sehr stark. Dabei erlangen Forschungs- und Bildungseinrichtungen das größte Vertrauen mit 77 % der Befragten, die voll und ganz vertrauen und nur etwas mehr als 1 % die diesen eher nicht oder nicht vertrauen. Das geringste Vertrauen erlangen Privatpersonen in sozialen Medien. Weniger als 1 % vertrauen ihnen voll und ganz. 88 % vertrauen diesen eher nicht oder vertrauen diesen gar nicht.

4.8 Frage 8: Wie wichtig ist es Ihnen, dass Sie nachvollziehen können, mit welcher Methodik die Informationen gewonnen wurden?

In dieser Frage wird sich mit der Nachvollziehbarkeit der bereitgestellten Informationen beschäftigt. Dem größten Teil der Befragten ist es sehr wichtig, nachvollziehen zu können, mit welcher Methodik die Informationen gewonnen wurden. 71 % finden es sehr wichtig, 24 % eher wichtig, 4 % weniger wichtig und weniger als 1 % nicht wichtig.

Die Abb. 9 zeigt die Korrelation zwischen der Frage 8 und dem Bildungsgrad der Befragten. Dabei ist zu erkennen, dass die Nachvollziehbarkeit der Methodik für Befragte mit einem Hochschulabschluss wichtiger ist (73 % sehr wichtig, 23 % eher wichtig) als für Befragte mit Abitur als höchsten Bildungsgrad (64% sehr wichtig, 31 % eher wichtig) oder der mittleren Reife (67 % sehr wichtig, 24 % eher wichtig). Lediglich eine Person mit Volks-/Hauptschulabschluss nahm an der Umfrage teil und beantwortete diese Frage mit weniger wichtig. Da weit über die Hälfte der Teilnehmenden einen Hochschulabschluss besitzen, wurden die anderen Abschlüsse für den Vergleich und aus Gründen der Übersichtlichkeit in Abb. 9 zusammengefasst. Insgesamt lässt sich erkennen, dass in allen repräsentativen Bildungskategorien die Nachvollziehbarkeit der Methodik als wichtig erachtet wird.

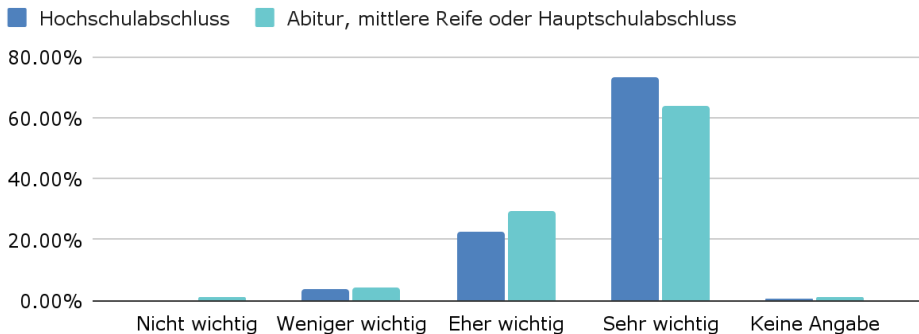


Abb. 9: Auswertung von Frage 8 mit Unterteilung nach dem Bildungsgrad

4.9 Frage 9: Wie wichtig ist es Ihnen, dass die Daten (z. B. Studien und Messwerte), auf denen die bereitgestellten Informationen beruhen, jederzeit leicht zugänglich sind?

Die Frage nimmt Bezug auf die Zugänglichkeit der bereitgestellten Informationen. Den meisten Befragten ist es wichtig, dass die Daten, auf denen die bereitgestellten Informationen beruhen, jederzeit leicht zugänglich sind. 64 % bewerteten das als sehr wichtig, 30 % als eher wichtig, 6 % als weniger wichtig und weniger als 1 % als nicht wichtig.

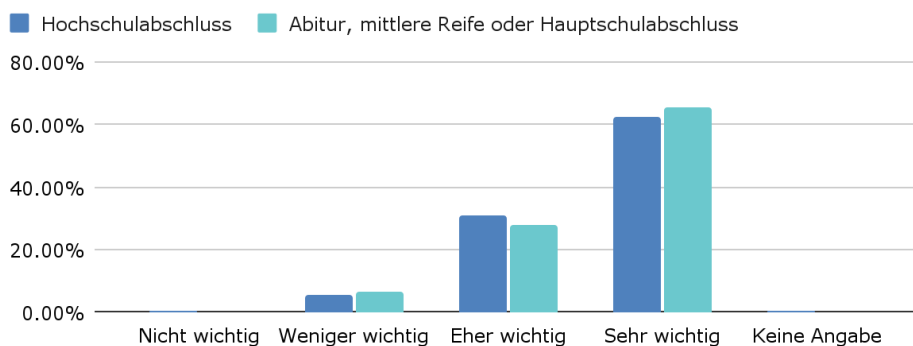


Abb. 10: Auswertung von Frage 9 nach dem Bildungsgrad

Auch diese Frage wurde entsprechend Frage 8 hinsichtlich der Korrelation mit dem Bildungsgrad untersucht. Dabei gaben 71% der Befragten mit mittlerer Reife an, einen leichten Zugang zu Daten als sehr wichtig zu empfinden und 24% als eher wichtig. Bei den Befragten mit Abitur waren 65% für sehr wichtig und 29% für etwas wichtig. Im Unterschied dazu gaben die Personen mit Hochschulabschluss zu 63 % sehr wichtig und 31 % eher wichtig an. Werden die Antworten der drei zusammengefassten Bildungskategorien und der Ergebnisse von Personen mit Hochschulabschluss gewichtet (1= Nicht wichtig, 2= Weniger wichtig, 3= Eher wichtig, 4= Sehr wichtig), so fällt der gewichtete Mittelwert von Teilnehmenden mit Abitur, mittlerer Reife oder Hauptschulabschluss knapp höher aus als der gewichtete Mittelwert der Personen mit Hochschulabschluss.

4.10 Frage 10: Haben Sie Interesse, über die folgenden Kanäle zukünftig über das Thema Niedrigwasser, Trockenheit und Dürre informiert zu werden?

In dieser Frage ging es darum, über welche Kanäle die Befragten in Zukunft über Niedrigwasser, Trockenheit und Dürre informiert werden möchten. Pro Kanal konnte eine Antwortoption (1= Kein Interesse, 2= Wenig Interesse, 3= Etwas Interesse, 4= Großes Interesse) ausgewählt werden. Allgemein hat ein Großteil entweder großes oder etwas Interesse daran. Das mit Abstand größte Interesse an der zukünftigen

Informationsbereitstellung über Kanäle liegt bei den „Websites“ (gewicht. Mittelwert 3,5). Alle weiteren Kanäle haben einen ähnlichen gewichteten Mittelwert von 3,0. Am wenigsten interessieren sich die Teilnehmenden für die Informationsbereitstellung über „Warnsysteme“ (gewicht. Mittelwert 2,8).

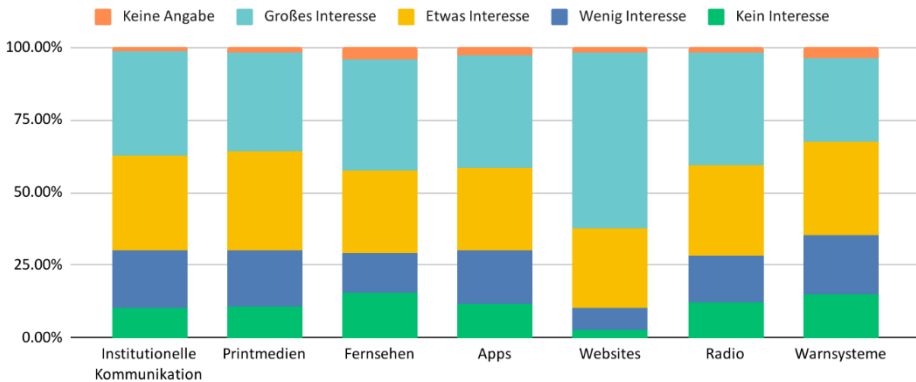


Abb. 11: Präferenzen für Informationskanäle zu Niedrigwasser, Trockenheit und Dürre

Des Weiteren wurde die Frage 10 im Hinblick auf die Korrelation mit den Altersklassen 0-39 und 40+ betrachtet. Die Altersklasse 40+ weist ein größeres Interesse für alle Kanäle, bis auf „Apps“ auf. In den Kanälen „Printmedien“ und „Fernsehen“ gibt es die größten Unterschiede. So haben in der Altersgruppe 0-39 26 % der Teilnehmenden großes Interesse am „Fernsehen“, während in der Altersklasse 40+ 54 % großes Interesse daran zeigen. Bei den „Printmedien“ haben in der Altersklasse 0-39 ebenfalls 26 % der Teilnehmenden großes Interesse, während in der Altersklasse 40+ 44 % großes Interesse aufweisen.

4.11 Frage 11: Welches Format der Information zum Thema Niedrigwasser, Trockenheit und Dürre gefällt Ihnen am besonders?

Bei dieser Frage sollten die Teilnehmenden auf einer Skala bewerten, welches Format der Information Ihnen wie sehr gefällt. Jedes der genannten Formate wurde von einem Großteil mit „Gefällt mir sehr“ oder „Gefällt mir etwas“ bewertet; das ist besonders auffällig bei „Interaktive Karte“ (gewicht. Mittelwert 3,7) gefolgt von den Formaten „Text mit Grafik/Diagramm“ (gewicht. Mittelwert 3,6) und „Video und Animation“ (gewicht. Mittelwert 3,4). „Audio“ wurde als Format hingegen von über 30% der Befragten als „Gefällt mir wenig“ oder „Gefällt mir gar nicht“ bewertet.

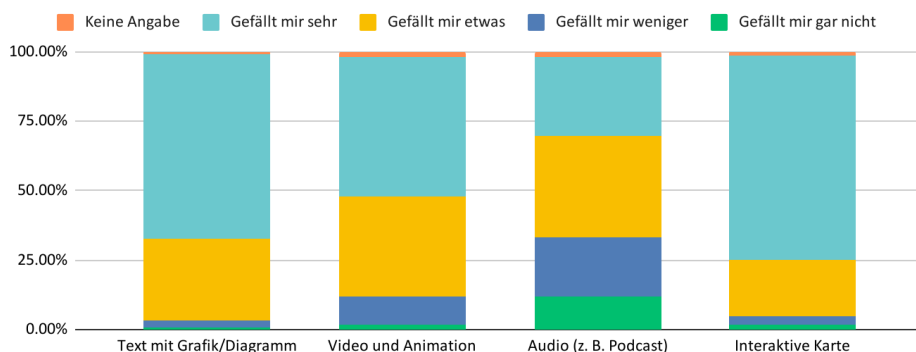


Abb. 12: Beliebtheit von Informationsformaten zu Niedrigwasser, Trockenheit und Dürre

Es wurden diverse weitere Formate im Freitextfeld „Sonstiges“ angegeben, die sich mitunter stark unterscheiden. Manchen Teilnehmenden gefallen eher analoge Formate wie „Exkursionen in der Natur“, „frontale Vorträge“ oder „kompetente Kontaktpersonen bei Behörden“, soweit diese als Format zählen. Anderen Teilnehmenden gefallen etwas spezifischere Online-Formate bzw. Formate, die auf digitalen Medien basieren, wie z.B. „Videospiele“, „VR-Brille“, „interaktive 360°-Rundgänge“ oder „interaktive Datenbanken mit Auswertungs- und Visualisierungsmöglichkeiten“. Auch hier wurden wieder „soziale Medien“ sowie „Blogbeiträge“ genannt und zuletzt auch das Format „Zahlen und Daten in Tabellarischer Form“.

5 Diskussion

5.1 Informationsstand und Betroffenheit

- **H1a: Die Bevölkerung ist nicht über Niedrigwasser, Trockenheit und Dürre informiert.**
 - Anhand der Antworten auf Frage 2 kann die Hypothese weder bestätigt noch widerlegt werden. Es ist festzuhalten, dass nur nach der Selbsteinschätzung des Wissens gefragt wurde. Diese kann vom tatsächlichen Wissensstand abweichen. Außerdem haben 76 % der Befragten einen Hochschulabschluss, wodurch der Bildungsgrad nicht repräsentativ für Deutschlands Gesamtbevölkerung ist und gerade die Ergebnisse dieser Frage, stark vom deutschen Durchschnitt abweichen könnte.
- **H1b: Je direkter die Bevölkerung betroffen ist, desto höher ist der Informationsbedarf in der Bevölkerung.**

- Diese Hypothese lässt sich untersuchen, wenn die Beantwortungen der Fragen 1 und 4 zusammengeführt werden. Dabei wurde betrachtet, ob die Betroffenheit mit dem Informationsbedarf der Bevölkerung korreliert. Circa drei Viertel der Befragten, die sich sehr stark betroffen fühlen, haben großes Interesse für die genannten Themen angegeben. Bei den Befragten, die sich eher betroffen fühlen, haben lediglich etwas mehr als die Hälfte großes Interesse an den Themen und ein Viertel etwas Interesse. Die Befragten, die sich eher nicht betroffen fühlen, haben ein deutlich geringeres Interesse am Thema Niedrigwasser, Trockenheit und Dürre gezeigt. Somit ist eindeutig erkennbar, dass das Interesse höher ist, wenn die gefühlte Betroffenheit höher ist, wodurch die Hypothese bestätigt wird.

5.2 Inhaltliche Interessenlage

- **H2a: Das Interesse der Bevölkerung ist am größten für Auswirkungen von Niedrigwasser, Trockenheit und Dürre auf sich selbst und auf die Umwelt.**
 - Zur Untersuchung dieser Hypothese wurde Frage 3 entwickelt. Die Hypothese hat sich bestätigt, da die große Mehrheit „Informationen über Auswirkungen auf die Umwelt“ und „Informationen über Auswirkungen auf mich selbst“ erhalten möchte.
- **H2b: Rezipienten, die an den Auswirkungen von Niedrigwasser, Trockenheit und Dürre auf sich selbst interessiert sind, haben ein großes Interesse an Informationen zu den Themen Reglementierung des Wasserverbrauchs und Folgen für Wirtschaft und Gesellschaft.**
 - Diese Hypothese kann bestätigt werden. Wie in Abb. 5 erkennbar, hat die Mehrheit in beiden Fällen großes Interesse. Jedoch erkennt man bei dem Thema „Reglementierung des Wasserverbrauchs“ eine deutlich größere Mehrheit mit großem Interesse.
- **H2c: Die Bevölkerung benötigt Informationen aus verschiedenen Bereichen, um mit Niedrigwasser, Trockenheit und Dürre besser umgehen zu können.**
 - Zur Auswertung dieser Hypothese kann Frage 5 betrachtet werden. Fast jede Antwortmöglichkeit wurde von ungefähr der Hälfte der Befragten ausgewählt. Die einzigen Ausreißer sind „Informationen zum Erhalten von finanzieller Unterstützung“ und „Informationen zur Landschaftsgestaltung“. Somit bestätigt sich die Hypothese, da die benötigten Informationen sehr vielfältig und unterschiedlich sind.
 - Eine mögliche Erklärung für die Ausreißer könnte das persönliche Vermögen der Befragten sein, da dieses beide Ausreißer beeinflussen könnte und die Reichweite der Umfrage limitiert ist (76 % der Befragten haben einen Hochschulabschluss). Hat eine Person ein hohes Einkommen, ist für diese

auch finanzielle Unterstützung weniger relevant und die Wahrscheinlichkeit größer, dass sie einen Garten besitzt [Vu21]. Über das Einkommen der Befragten liegt keine Information vor. Es kann jedoch der Bildungsgrad als Indikator für das Einkommen verwendet werden. So ist in Deutschland das durchschnittliche monatliche Einkommen höher, je höher der erlangte Bildungsgrad ist [PKW17]

- Verknüpft man die Antworten mit dem Bildungsgrad, kann gezeigt werden, dass die Befragten mit Hochschulabschluss und ohne Hochschulabschluss zu gleichen Anteilen für „Informationen zur finanziellen Unterstützung“ interessieren. Bei der Landschaftsgestaltung ist ein deutlicher Unterschied hinsichtlich der Korrelation mit dem Bildungsgrad erkennbar. Jedoch sind auch für Befragte ohne Hochschulabschluss „Informationen zur Landschaftsgestaltung“ relevanter als die anderen Informationen. Außerdem interessieren sich der Großteil der Befragten, die sich für finanzielle Unterstützungen interessieren, auch für die Landschaftsgestaltung. Unterschiede im Einkommen sind also keine geeignete Erklärung dafür, warum die beiden Antworten besser bzw. schlechter abgeschlossen haben.

5.3 Motivation

- **H3a: Die Bevölkerung erhofft sich von einem Informationsangebot zum Thema Niedrigwasser, Trockenheit und Dürre, Handlungsmöglichkeiten für den Alltag ableiten und selbstständig ergreifen zu können.**
 - Zur Untersuchung dieser Hypothese kann Frage 6 verwendet werden. Dabei ist erkennbar, dass das Ableiten von Handlungsmöglichkeiten zwar die beliebteste Antwort ist, sich aber auch nicht stark von den anderen Antwortmöglichkeiten abhebt. Um hier einen Unterschied deutlich zu machen und die Hypothese verifizieren zu können, könnten in einer erneuten Umfrage die Antwortmöglichkeiten durch die Befragten in eine Reihenfolge, entsprechend der Wichtigkeit, gebracht werden.

5.4 Vertrauenswürdigkeit der Quellen

- **H4a: Die Bevölkerung vertraut vorrangig Informationen die von Forschungs- und Bildungseinrichtungen oder von Behörden und Amtsträger*innen bereitgestellt werden.**
 - Diese Hypothese hat sich mit Frage 7 bestätigt. „Forschungs- und Bildungseinrichtungen“ und „Behörden und Amtsträger*innen“ sind die vertrauenswürdigsten Informationsquellen und sind damit noch vor dem „öffentlich-rechtlichen Rundfunk“.

- **H4b: Der Bevölkerung mit einem höheren akademischen Grad ist die Nachvollziehbarkeit der Methodik und die Zugänglichkeit der zugrundeliegenden Daten der bereitgestellten Informationen wichtiger, als der Bevölkerung mit einem niedrigeren akademischen Grad.**
 - Zur Überprüfung dieser Hypothesen werden die Fragen 8 und 9 verwendet und hinsichtlich ihrer Korrelation mit dem Bildungsgrad untersucht. Dabei ist zu erkennen, dass die Nachvollziehbarkeit der Methodik erwarteter Weise für Befragte mit einem Hochschulabschluss wichtiger ist als für Befragte mit Abitur oder der mittleren Reife als höchsten Bildungsgrad. In allen repräsentativen Bildungskategorien ist jedoch die Nachvollziehbarkeit der Methodik sehr wichtig und die Unterschiede fallen nur gering aus.
 - Unerwarteterweise ist der leichte Zugang zu Daten, auf denen die bereitgestellten Informationen beruhen, für Befragte mit Abitur, mittlerer Reife oder Hauptschulabschluss wichtiger ist als für Befragte mit Hochschulabschluss.
 - Eine Annahme, warum mit geringerem Bildungsgrad der Zugang zu Daten wichtiger wird, könnte sein, dass das Vertrauen in Informationsgeber bei geringerem Bildungsgrad niedriger ist. Das hat bereits eine Befragung von Media Impact GmbH & Co. KG ergeben [Bp19]. Auch die Ergebnisse dieser Studie zeigen, dass das Vertrauen in die Informationsquellen (entsprechend Frage 7) im Durchschnitt niedriger ist, je geringer der Bildungsgrad ausfällt. Das Vertrauen in die Informationsgeber ist bei den Befragten mit Hochschulabschluss also etwas höher und bestätigt die Annahme zumindest quantitativ. Insgesamt muss die Hypothese entsprechend abgelehnt werden.

5.5 Bevorzugte Kommunikationsart

- **H5a: Die Bevölkerung möchte zukünftig vorwiegend über Online-Kanäle wie Webseiten und Apps über das Thema Niedrigwasser, Trockenheit und Dürre informiert werden.**
 - Diese Hypothese hat sich bestätigt. Bei Frage 10 haben „Websites“ das mit Abstand beste Ergebnis erlangt. Allgemein zeigte sich aber ein großes Interesse bei den Teilnehmenden über diverse Kanäle informiert zu werden.
- **H5b: Altersgruppen von 0-39 bevorzugen die Kanäle Webseiten und Apps, wohingegen Altersgruppen von 40-70 und älter die Kanäle Printmedien und Fernsehen bevorzugen.**
 - In dieser Hypothese wurde untersucht, ob das Alter einen Einfluss auf das Interesse an Informationskanälen zum Thema Niedrigwasser, Trockenheit und Dürre hat. Die Ergebnisse zeigen, dass das Interesse von Befragten bis 39 Jahren an „Apps“ etwas geringer ist als das Interesse aller Befragten. Das

Interesse an „Websites“ ist bei Befragten bis 39 Jahren ebenfalls etwas geringer im Vergleich zu allen Befragten. Bei Befragten über 40 Jahren wurde häufiger die Option „Keine Angabe“ verwendet, was als fehlendes Interesse interpretiert werden kann. Das Interesse an „Apps“ ist bei unter 40-Jährigen jedoch nur geringfügig höher als bei über 40-Jährigen, während das Interesse an „Websites“ bei den über 40-Jährigen sogar größer ist.

- Allgemein ist das Interesse der Befragten unter 40 Jahren an „Websites“ am größten, gefolgt von „Apps“, womit sich der erste Teil der Hypothese bestätigen lässt. Bei Befragten im Alter von 40 Jahren oder älter besteht ein höheres Interesse an „Printmedien“ im Vergleich zu allen Befragten. Das Interesse an Informationen über das „Fernsehen“ ist bei Befragten im Alter von 40 Jahren oder älter signifikant höher als bei allen Befragten. Dennoch ist das Interesse an „Websites“ auch bei Befragten über 40 Jahren am größten, wodurch sich der zweite Teil der Hypothese nur teilweise bestätigt hat.
- **H5c: Das bevorzugte Format der Informationsaufbereitung sind Texte und Grafiken.**
 - Das bevorzugte Format ist die „Interaktive Karte“. „Texte und Grafiken“ erfreuen sich einer nur geringfügig niedrigeren Beliebtheit und sind somit knapp hinter der „Interaktiven Karte“. Da nicht alle Informationen als „Interaktive Karten“ dargestellt werden können und die Beliebtheit verglichen mit „Texten und Grafiken“ fast identisch ist, hat sich die Hypothese eher bestätigt.

6 Schlussfolgerungen

Abschließend können aus den Ergebnissen der Umfrage die ursprünglichen Forschungsfragen beantwortet und verschiedene Handlungsempfehlungen zur Ausgestaltung eines Informationsangebots zu Niedrigwasser-, Trockenheit- und Dürre-Themen für Bürgerinnen und Bürger abgeleitet werden.

- **An wen sollten die Informationen adressiert werden?**
 - Die Informationen sollten an alle Bürger und Bürgerinnen gerichtet sein.
 - Das Publikum wird sich wahrscheinlich besonders aus direkt von Niedrigwasser, Trockenheit und Dürre betroffenen Personen zusammensetzen. Es sollten Überlegungen angestellt werden, wie zum einen auf deren spezielles Informationsbedürfnis (vorwiegend auf dem Land lebend) ausreichend eingegangen werden kann und wie zum anderen weniger direkt betroffene Personen (vorwiegend in der Stadt lebend) trotzdem mit wichtigen Informationen erreicht werden können.
- **Was für Informationen möchten die adressierten Personen erhalten?**

- Inhaltlich sollte der Schwerpunkt auf den Auswirkungen von Niedrigwasser-, Trockenheit- und Dürre-Ereignissen auf die Personen selbst (z.B. durch Reglementierung des Wasserverbrauchs) und auf die Umwelt (z.B. Auswirkungen auf Gewässer und Fische) liegen.
 - Auf das Thema Wasserreglementierung könnte ein extra Schwerpunkt gelegt werden – z.B. mit einer interaktiven Karte, über die die aktuellen Bestimmungen für die verschiedenen Landkreise abrufbar sind.
 - Auf das Thema Gewässer und Fische könnte ebenfalls eine Karte mit einem Ampelsystem für die aktuelle Gefährdung in den verschiedenen Gewässern eingeführt werden.
 - Das Informationsangebot sollte möglichst breit gefächert sein. Möglicherweise wäre es zielführend, auf jeder Seite am Ende die Frage zu stellen, ob diese Information hilfreich war, um das Angebot schrittweise zu verbessern.
 - Die Informationen sollten die lokalen Bedürfnisse und Besonderheiten berücksichtigen, welche im ländlichen und im städtischen Raum voneinander abweichen können.
- **Welchen Effekt wünschen sich die adressierten Personen von den Informationen?**
 - Viele der Befragten wollen sich informieren, um handlungsfähig zu sein. Auf diesen Zweck/ diese Motivation sollte ein Teil des Informationsangebots extra zugeschnitten werden, zum Beispiel mit praktischen Tipps, die helfen, sich auf die Auswirkungen von Niedrigwasser, Trockenheit und Dürre vorzubereiten.
- **Wer soll die Informationen bereitstellen?**
 - Informationen von Forschungs- und Bildungseinrichtungen sollten bevorzugt werden, da ihnen besonders viel Vertrauen entgegengebracht wird. Dennoch sollten für eine effiziente Gestaltung der Informationen verschiedene Organisationen zusammenarbeiten.
 - Zusätzlich wird empfohlen, den interessierten Personen immer die Möglichkeit zu bieten, mehr über die Hintergründe einer Information in Erfahrung zu bringen. So ist es hilfreich, zum Beispiel die Methodik einer Studie verständlich zu erklären und Zugriff auf mögliche weitere Forschungsdaten zu bieten.
 - Es ist wichtig, die Informationen kontinuierlich zu aktualisieren und zu überprüfen, damit sie immer auf dem neuesten Stand sind.
- **Über welchen Kanal sollten die Informationen kommuniziert werden?**

- Es wird empfohlen, Online-Kanäle (Webseiten, Apps) zu nutzen, um eine möglichst große Zielgruppe zu erreichen. Digitale Medienkanäle sprechen bereits viele Leute an. Eine Umsetzung als Hybrid-App würde sich anbieten, da diese eine Darstellung als Webseite im Browser und auf mobilen Plattformen in App-Form ermöglicht.
- Es ist wichtig, dass die Informationen in einer klaren und verständlichen Sprache präsentiert werden, um sicherzustellen, dass die Bevölkerung sie leicht versteht.
- Bei der Darstellung von Informationen sollte dringend auf eine visuell ansprechende und leicht verständliche Aufbereitung geachtet werden. Formate wie interaktive Karten, Grafiken und Texte werden empfohlen.
- Es wird empfohlen, Artikel zusätzlich im Audio-Format anzubieten. Dies erhöht die Barrierefreiheit und ist bereits bei vielen Menschen beliebt.

7 Zusammenfassung und Ausblick

Diese Studie untersucht erstmals das Bewusstsein und den Informationsbedarf zum Thema Niedrigwasser, Trockenheit und Dürre innerhalb der Bevölkerung. Sie bietet ein umfassendes Bild darüber, wie der vorliegende Wissensstand der Teilnehmenden zu bewerten ist und was sich diese für die Zukunft wünschen würden.

Durch die Literaturrecherche konnten essenzielle Fragen für die Ermittlung des Informationsbedarfs aufgestellt und mit Hilfe der Lasswell-Formel strukturiert werden. Nach Erstellung der Hypothesen und Fragen, einem Pretest der Umfrage und einer Validierung der Fragen in Kooperation mit Expert*innen verschiedener Institutionen, wurde der Fragebogen online, über das Portal „SurveyMonkey“, zur Verfügung gestellt. Die Umfrage war sieben Wochen online verfügbar und 450 Personen nahmen daran teil.

Die Auswertung der Umfrage zeigt, dass viele der zuvor erstellten Hypothesen zu weiten Teilen korrekt antizipiert wurden. Die Ergebnisse der Umfrage stammen jedoch nicht zu gleichen Teilen aus den verschiedenen Bevölkerungsteilen Deutschlands. Sie konnten aufgrund der hohen Anzahl an Teilnehmenden aber dennoch für eine Ableitung eines Kataloges mit Handlungsempfehlungen, zur Ausgestaltung eines Informationsangebots über Niedrigwasser, Trockenheit und Dürre für Bürgerinnen und Bürger, genutzt werden.

Diese Handlungsempfehlungen sollten nun durch zuständige Ämter und Institutionen geprüft werden. Es sollte getestet werden, ob und wie die bereits zur Verfügung gestellten Informationen von der Bevölkerung aufgenommen werden. Nach der Umsetzung oder Überarbeitung eines Informationsangebotes könnten die Ergebnisse der Umfrage genutzt werden, um die Effektivität des Angebots zu überprüfen und eine Verbindung zu den Ergebnissen der Ursprungsumfrage zu schaffen.

Abschließend kann festgehalten werden, dass das Thema des Informationsbedarfs der Bevölkerung bezogen auf Niedrigwasser, Trockenheit und Dürre in der Zukunft weitere Forschung benötigt. Weiterhin ist es wichtig, neue Erkenntnisse in die Informationsverbreitung einfließen zu lassen, um die Bevölkerung effektiv über die Themen Niedrigwasser, Trockenheit und Dürre informieren zu können.

Literatur

- [Bi22] Bissoli, P. et. al.: Trockenheit in Europa 2022. Offenbach, 2022.
- [Bp19] b4p trends: Welchen News-Quellen vertrauen die Onliner?, 2019.
- [Bu23] Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz: Nationale Wasserstrategie. Berlin, 2023.
- [La48] Lasswell, H. D.: The structure and function of communication in society. In: The Communication of Ideas : Bryson, Lyman, S. 216–228, 1948.
- [Mi21] Ministerium für Landwirtschaft, Umwelt und Klimaschutz Brandenburg: Landesniedrigwasserkonzept Brandenburg. Potsdam, 2021.
- [PKW17] Piopunik, M.; Kugler, F.; Wößmann, L.: Einkommenserträge von Bildungsabschlüssen im Lebensverlauf: Aktuelle Berechnungen für Deutschland. In: ifo Schnelldienst Bd. 07, Nr. 70, S. 19–30, 2017.
- [Ri21] Riedel, T. u. a.: Niedrigwasser, Dürre und Grundwasserneubildung – Bestandsaufnahme zur gegenwärtigen Situation in Deutschland, den Klimaprojektionen und den existierenden Maßnahmen und Strategien. Essen, 2021.
- [SHE18] Schnell, R.; Hill, P. B.; Esser, E.: Methoden der empirischen Sozialforschung: De Gruyter Oldenbourg, 2018.
- [Um23] Umweltbundesamt: Trends der Niederschlagshöhe, <https://www.umweltbundesamt.de/daten/klima/trends-der-niederschlagshoehe>, Stand: 14.07.2023.
- [Vu21] VuMa: Gartenbesitzer in Deutschland nach Netto-Einkommen im Vergleich mit der Bevölkerung im Jahr 2021, Arbeitsgemeinschaft Verbrauchs- und Medienanalyse, <https://de.statista.com/statistik/daten/studie/316331/umfrage/umfrage-in-deutschland-zum-netto-einkommen-von-gartenbesitzern/>, Stand: 14.07.2023.
- [Wd22] WDR: Umfrage zur Trockenheit: Zwei von drei Deutschen machen sich Sorgen, <https://www1.wdr.de/nachrichten/angst-klima-zukunft-umfrage-100.html>, Stand: 14.07.2023.
- [WBG19] Williams, H.; Benthin, R.; Gellrich, A.: Umweltbewusstsein in Deutschland 2018. Umweltbundesamt, Dessau-Roßlau, 2019.
- [WG22] Williams, H.; Gellrich, A.: Umweltbewusstsein in Deutschland 2020. Umweltbundesamt, Dessau-Roßlau, 2022.

Ökologische Nachhaltigkeit - 11.
Workshop Umweltinformatik
zwischen Nachhaltigkeit und
Wandel (UINW 2023)

Digital Transformation in Forestry - Stakeholders and Data Collection in German Forests

Michael Pleger ¹, Ina Schiering ²

Abstract: This paper examines the process of data collection within forestry and its individuals and organizations in Germany as a first step towards digitization. Germany has one of the largest forest areas in Europe and plays a significant role in mitigating the impacts of climate change and providing ecosystem services to society. To fulfill this role, data collection is required by law for most forests in Germany. This data collection has been historically labor intensive and time consuming. Data collection through *Internet of Things* (IoT) devices has the potential to improve the efficiency and accuracy of forestry operations while also providing valuable insights into forest health and productivity. Modern data collection through drones and satellite imagery already provide significant benefits to the economy of forestry. This could be further enhanced by low-cost IoT devices bundled as sensor networks in forests to gather data over a lifespan of a forest.

Keywords: forestry; climate change; digitization; data collection; sensor networks

1 Introduction

Forestry in Germany plays a significant role in mitigating the impact of climate change and providing ecosystem services to society. Not only provides forestry a crucial industry that plays an important role in the country's financial economy and environment with a net income of more than 1.3 billion € in 2020 [St22]. It also offers huge ecological and social benefits, including carbon sequestration, biodiversity conservation and recreational opportunities. According to the *Food and Agriculture Organization* (FAO) of the United Nations [FO20], Germany is one of the leading countries in Europe in terms of forest area and forest management. As of the third *National Forest Inventory* (NFI) in 2012, Germany had a total forest area of 11.4 million hectares, which represents approximately 32% of the country's land area. Forestry in Germany is managed largely by private forest owners, who own 48% of the country's forested land. The remaining 52% is split between public corporations and entities, including federal, state and municipal governments, according to the results of the NFI in 2012 [WB19]. Forestry itself involves a wide range of activities, including planting, harvesting and managing forests. By nature these activities have long cycles of up to 100 years between planting and harvesting, depending on the tree species.

¹ Ostfalia University of Applied Sciences, Faculty of Computer Science, Salzdahlumer Straße 46/48, Wolfenbüttel, Germany, mic.pleger@ostfalia.de

² Ostfalia University of Applied Sciences, Faculty of Computer Science, Salzdahlumer Straße 46/48, Wolfenbüttel, Germany, i.schiering@ostfalia.de

The German government has established laws and regulations to ensure sustainable forest management in order to protect this valuable asset. These include the *Federal Forest Act* ('Bundeswaldgesetz') and the *Forest Management Guidelines* ('Waldpflegegesetz'), which promote ecological, economic, and social sustainability of forest management, and are supported by a range of institutions and organizations involved in forestry research and education.

In regards to sustainability, climate change is posing significant challenges to the forestry sector in Germany. Rising temperatures, changing precipitation patterns, and more frequent extreme weather events are leading to increased challenges in forestry, including insect outbreaks (e. g. bark beetle infestation), wildfires, and wind-throw. These disturbances not only impact the health and productivity of forests but also threaten the provision of ecosystem services to the society. To address these challenges, the forestry sector in Germany is undergoing a transformation towards more sustainable and resilient forest management practices. These practices aim to promote forest health and productivity while also enhancing the ability of forests to mitigate and adapt to climate change. For example, forest managers are adopting measures to increase the diversity of tree species, promote natural regeneration, and reduce the risk of forest disturbances.

To adapt to changing conditions, monitoring and collection of data are important aspects. Data collection involves the systematic gathering of information about forest ecosystems, including their composition, size and age, structure, local terrain features and dynamics. A data collection is necessary to monitor forest health, identify potential risks and threats, and inform management decisions. To achieve this, Germany mandates to gather such a data collection in a long-term interval, which depends on the ownership and size of the forest. For example, publicly owned forests require a *Forest Management Plan* (FMP) that has to be conducted every ten years, which contains a predefined set of data points like the size, composition, gain, inventory or the degree of damage. Such a FMP provides the basis for taxation and therefore for the economy of forestry. In contrast, private forest owners are required to provide a subset of the FMP data points as an operations report, which gets used for e. g. tax compensation. The NFI, utilized for research and development, is a monitoring program that contains these data collections, as well as forest mappings from federal entities (e. g. *Chamber of Agriculture* (SCA) of each federal state).

In terms of data collection, the *Thünen Institute of Forest Ecosystems* (TI) is one of the main institutions responsible for forest monitoring and research in Germany. The institute conducts a range of studies to monitor forest health, biodiversity, productivity, and provides scientific advice to support policy development and decision-making in rural areas, agriculture, forests and fisheries. Other organizations involved in forestry research and data collection in Germany include the *Federal Agency for Nature Conservation* (BfN) and the *German Centre for Integrative Biodiversity Research* (iDiv).

Data collection in forestry is based on historical data, statistics, and manual field surveys using a range of sampling techniques. For example, the TI conducts surveys to measure tree

height, diameter and volume on forests, as well as on specific ecological research stations (e. g. in Britz, Germany [Jo23]). To enhance this data collection with much more up-to-date data, an approach would be to utilize modern low-cost devices and sensor networks deployed in a large scale. Those battery operated smart devices gathering information like soil moisture, temperature, humidity or precipitation could be joined to a low powered sensor network to provide real-time data in a timely manner from forests remotely. For example during planting season, this could increase the viability of tree saplings due to a more controlled environment.

As a first step in the process of digitization and based on expert interviews, this paper gives an overview of stakeholders involved in data collection in German forests.

2 Related Work

Germany already has a number of policies for forest monitoring in place (e. g. [BM16; We18]). The German forest act provided the legal grounds for the NFI and the *National Forest Soil Inventory* (NFSI). Thorough monitoring is also conducted through the annual *Crown Condition Survey* (CCS) and continuous *Intensive Monitoring* of 68 test fields in Germany. Data collected by the NFSI and CCS (level I), as well as the intensive monitoring (level II) are nested within the *International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests* (ICP Forest) [BM20; Fl16]. Holzwarth et al. provided a review regarding data collection and purpose in Germany [Ho20]. Remote sensing programs utilizing satellite [Wh16] or *unmanned areal vehicles* (UAV) have been established in many countries, including Germany [Gu20; To17]. Furthermore, researchers are evaluating the prospect of data collection in the harvesting process and possible benefits thereof [Ro19]. Social effects of forests on humans are also an area of research [Si18]. However the general focus of data collection in forestry has been towards research in regards to forestry management and planning and is missing interdisciplinary studies of both the purpose, and views of different stakeholders. Though that is needed as a basis for digital transformation.

3 Analysis of Status Quo

The relationship between stakeholders in German forestry is complex and often collaborative. The stakeholders involved include federal, state, and municipal governmental entities and organizations, as well as private businesses like contractors and individuals like forest owners based on interviews with interdisciplinary experts. To better illustrate these an overview will break down the stakeholders from a federal to a state and municipal level and describe individuals and organizations. Figure 1 visualizes a simplified overview of forestry's stakeholders in Germany.

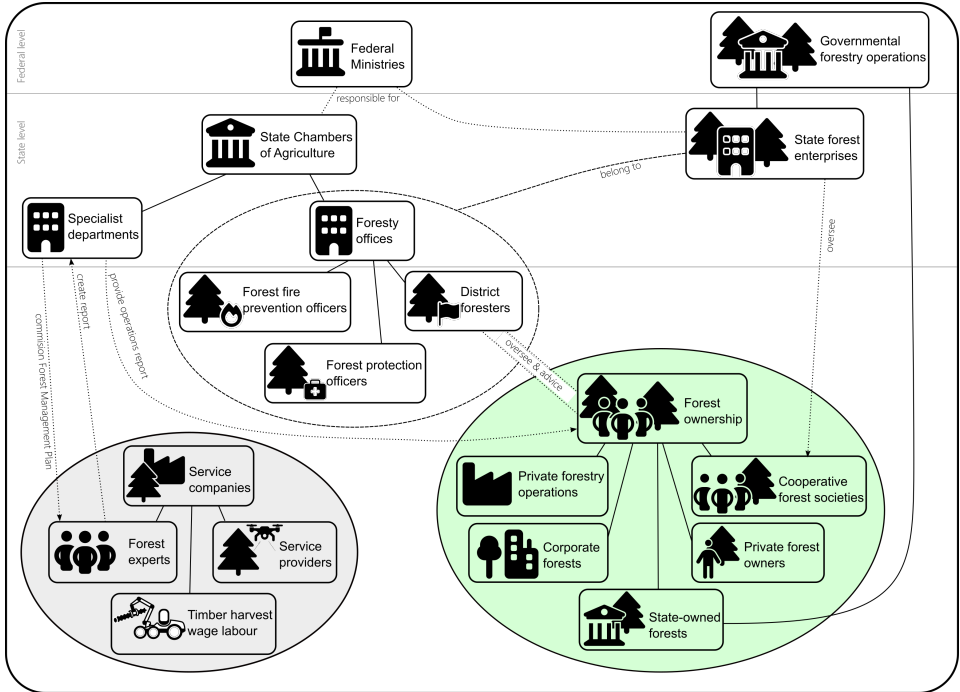


Fig. 1: Overview of stakeholders involved in German forestry - including the process of how *Forest Management Plans* are commissioned to forest owners.

The *Federal Ministry of Food and Agriculture* (BMEL) is responsible for creating policies related to forestry in Germany. The ministry sets standards for sustainable forest management by developing and implementing policies and legislation related to forestry. This includes setting national goals for sustainable forests management, defining regulations and establishing funding programs to support forestry activities. Furthermore it manages the so called '*Bundesforst*', the state-owned forests in Germany by developing management plans, coordinating the harvesting and selling timber. It is the national equivalent to the FAO of the United Nations. In terms of data collection one of the prominent programs developed by the BMEL is the NFI, intended to run from 2013 until 2052. The program aims to systematically survey and monitor the state of forests across Germany. It realized a database that contains data sets about the condition of forests on a national scale. The collection can be used to visualize how forests have grown (or reduced), how the forests are being used (commercially), or how the timber harvesting potential has developed over the years. For example, the non-commercial, protected forest area assigned to the NATURA-2000 amounts to 2,7 million hectares of forest, or roughly 24% of German forest area [BM20]. The NFI collects data on a wide range of forest-related variables, including forest structure, tree species composition, biomass, and forest health, tree ages and value proportions among

others. To gather these information, forest owners have to submit details of their forest operations report. For the NFI 2022 it is intend to add remote sensing data from Sentinel-2 satellite imagery to derive tree species or timber stocks and validate them against locally gather data.

In contrast to the BMEL, which focuses on policies regarding the economy and sustainability, the *Federal Agency for Nature Conservation* (BfN) is responsible for nature and forest conservation. The BfN develops and implements policies and measures to protect and conserve the biodiversity and ecosystems in Germany. It defines constraints regarding forestry and the impact thereof on the environment. The BfN further develops and coordinates research projects related to forest conservation and provides information and training to forest practitioners, policymakers and the general public. The BfN provides easy access to layered cartography based on data collected by the NFI for multiple purposes, for example to showcase how much area of forests are managed under a preservative title (e. g. Forest Stewardship Council).

Isolated from the commercial aspects of forestry, a data collection purpose is based on research. As such, the TI plays a significant role in providing scientific advice and policy recommendations to national and international policymakers, stakeholders, and the public. The institute examines how land use can be managed sustainably to meet the needs of agriculture, forestry, and other land uses while protecting the environment and promoting social and economic development. Therefore, it operates a research site as *a forest-based hydro-logical test field* in Britz (Germany) since 1972. The site covers an area of around 50 hectares and serves as a testing ground for various agricultural and forestry research projects. The location is equipped with various research facilities and resources, including experimental plots for crop cultivation, research forests, a biomass power plant, and research laboratories. The site is used for a wide range of research projects as it provides a controlled environment for conducting experiments and testing new ideas and technologies.

In Germany each national state does have a *Chambers of Agriculture*, which represents the interests of forest owners on a regional and national level and provides advice and support related to forestry management. It's data collection includes both agriculture and forestry, e. g. crop yields, livestock production, land use and forestry management practices. The SCA utilizes the data to identify trends and outlooks in those domains. The Chamber is typically comprised of multiple specialist departments, since it manages all domains of agriculture. In regards to forestry, it also uses data collected by the NFI and provides applications for planning and managing forests. Furthermore, it provides geographic information systems to foresters, forest experts and owners.

Forestry offices are public institutions orientated below the SCA on a state level. Their purpose is to ensure sustainable management of state-owned forests, and to promote multiple functions of forests, including timber production, biodiversity conservation, and recreational activities (e. g. managing and creating hiking trails or tourism). Forestry offices operate

both on a legislative and executive level. The later belongs to state forest enterprises, which is a governmental forestry operation and responsible for state-owned forests.

Subordinated to forestry offices are state or municipal district foresters, who exercise supervision of forestry and forestry management. They oversee and advice forest owners to comply with policies and regulations. They are also responsible for protecting forests from pests, diseases, and natural disasters, such as storms and fires. Regular inspections are required to fulfill those tasks, where data collection is typically handled by hand on paper or via mobile applications. Further entities of the forestry office are forest fire prevention officers, responsible for fire safety measures, as well as forest protection officers which enforce the laws and regulations on a local level.

In Germany, forest ownership is very fragmented, with a large number of small private forest owners. Around half of the forest area in Germany is privately owned, and the rest belongs to public entities such as federal or state governments. These owners are required to collect data of their forests, that includes information on the forest's tree species composition, age, growth rate, and health, as well as data on the topography, soil type, and climate of the forest area. They also monitor wildlife populations and biodiversity, assess potential risks such as forest fires and pest infestations, and keep track of the economic performance of their forests, including income and expenses related to forestry activities. Some of that data has to be provided by a so called forest operations report, which is the financial basis for forest owners regarding taxation. Depending on the size and ownership, this report has to be extended and reflect as a management plan. Owners typically outsource these tasks to professionals, e. g. the creation or a yearly inventory gets conducted by forest experts or service providers.

In a commercial aspect, forestry and specifically timber production is another process which typically involves several stages, including planning, harvesting, transport and processing. The planning phase is developed by forest owners according to their management plans. For harvesting, foresters select individual trees or areas for harvesting based on their age, size, condition and requirement. During harvesting, data such as the number of trees harvested, their species, volume, and quality are collected directly on the harvesting machines. Forest owners might also gather data about the impact of harvesting on the forest ecosystem, such as changes in soil and water quality or wildlife population.

4 Discussion and Conclusion

Traditional data collection in German forestry is often times conducted analog and with manual, re-occurring surveys. Modern approaches for remote sensing, like the Copernicus program of the European Union with Landsat 8 or Sentinel-2 data already provide valuable insights into forest conditions like the temperature or humidity levels, but lack accurate measurements on a local or regional scale. However, such data provides an interesting basis for large scale assessments in regards to forest health. Furthermore the ICP Forest

program already provides valuable insights with over 286 plots of forest monitoring, where measurements like the soil condition and ground water levels can be co-related to dynamic data on e. g. tree growth and the crown condition or ambient air quality. This database provides a significant basis for evaluations and trend analysis, however insights for localized plots and private owners can only be derived of the database in a general sense. Another remote sensing solution currently evaluated within the scientific field are drones or UAVs. These can provide localized insights with optical and thermal imagery on an areal scale, however continuous data can not be collected or requires an operator to be scheduled. On a local scale, the next step towards digitization in forestry could be to conduct a field test together with the stakeholders identified above by holding participatory workshops to address the challenges posed by climate change and changing business models. The results could then be evaluated and applied in a living lab using digital devices for enhanced remote sensing, possibly paired with grid-independent wireless communication to provide decision making information for foresters, researchers and the general public.

Acknowledgements

This work was supported by the Federal Ministry for Digital and Transport as part of 5G Smart Country (45FGU117).

Bibliography

- [BM16] BMEL: Forstliches Umweltmonitoring in Deutschland. Federal Ministry of Food and Agriculture, Berlin, 2016.
- [BM20] BMEL: Am Puls des Waldes - Umweltwandel und seine Folgen – ausgewählte Ergebnisse des intensiven forstlichen Umweltmonitorings. Federal Ministry of Food and Agriculture, Bonn, 2020.
- [FI16] Fleck, S.; Cools, N.; De Vos, B.; Meesenburg, H.; Fischer, R.: The Level II aggregated forest soil condition database links soil physicochemical and hydraulic properties with long-term observations of forest condition in Europe. *Annals of Forest Science* 73/4, Number: 4 Publisher: BioMed Central, pp. 945–957, Dec. 2016, ISSN: 1297-966X, URL: <https://annforsci.biomedcentral.com/articles/10.1007/s13595-016-0571-4>, visited on: 06/27/2023.
- [FO20] FOA: Global Forest Resources Assessment 2020. FOA, 2020, ISBN: 978-92-5-132974-0.
- [Gu20] Guimarães, N.; Pádua, L.; Marques, P.; Silva, N.; Peres, E.; Sousa, J. J.: Forestry Remote Sensing from Unmanned Aerial Vehicles: A Review Focusing on the Data, Processing and Potentialities. *Remote Sensing* 12/6, Number: 6 Publisher:

- Multidisciplinary Digital Publishing Institute, p. 1046, Jan. 2020, ISSN: 2072-4292, URL: <https://www.mdpi.com/2072-4292/12/6/1046>, visited on: 05/13/2023.
- [Ho20] Holzwarth, S.; Thonfeld, F.; Abdullahi, S.; Asam, S.; Da Ponte Canova, E.; Gessner, U.; Huth, J.; Kraus, T.; Leutner, B.; Kuenzer, C.: Earth Observation Based Monitoring of Forests in Germany: A Review. *Remote Sensing* 12/21, Number: 21 Publisher: Multidisciplinary Digital Publishing Institute, p. 3570, Jan. 2020, ISSN: 2072-4292, URL: <https://www.mdpi.com/2072-4292/12/21/3570>, visited on: 05/13/2023.
- [Jo23] Johann Heinrich von Thünen-Institut: Intensivmonitoringfläche Britz, 2023, URL: <https://www.thuenen.de/de/fachinstitute/waldoekosysteme/arbeitsbereiche-neu/oekologie-walddynamik/waldoekologie-versuchsstation-britz>, visited on: 05/13/2023.
- [Ro19] Rossit, D. A.; Olivera, A.; Viana Céspedes, V.; Broz, D.: A Big Data approach to forestry harvesting productivity. *Computers and Electronics in Agriculture, BigData and DSS in Agriculture* 161/, pp. 29–52, June 1, 2019, ISSN: 0168-1699, URL: <https://www.sciencedirect.com/science/article/pii/S0168169917316368>, visited on: 05/13/2023.
- [Si18] Siebert, A.; Bezama, A.; O’Keeffe, S.; Thrän, D.: Social life cycle assessment: in pursuit of a framework for assessing wood-based products from bioeconomy regions in Germany. *The International Journal of Life Cycle Assessment* 23/3, pp. 651–662, Mar. 1, 2018, ISSN: 1614-7502, URL: <https://doi.org/10.1007/s11367-016-1066-0>, visited on: 05/13/2023.
- [St22] Statistisches Bundesamt: Umsatz in der Forstwirtschaft in Deutschland nach Segmenten bis 2020, 2022, URL: <https://de.statista.com/statistik/daten/studie/270514/umfrage/umsatz-in-der-forstwirtschaft-in-deutschland-nach-segmenten/>, visited on: 05/08/2023.
- [To17] Torresan, C.; Berton, A.; Carotenuto, F.; Di Gennaro, S. F.; Gioli, B.; Matese, A.; Miglietta, F.; Vagnoli, C.; Zaldei, A.; Wallace, L.: Forestry applications of UAVs in Europe: a review. *International Journal of Remote Sensing* 38/8, Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/01431161.2016.1252477>, pp. 2427–2447, May 19, 2017, ISSN: 0143-1161, URL: <https://doi.org/10.1080/01431161.2016.1252477>, visited on: 05/13/2023.
- [WB19] Wellbrock, N.; Bolte, A., eds.: Status and Dynamics of Forests in Germany: Results of the National Forest Monitoring. Springer International Publishing, Cham, 2019, ISBN: 978-3-030-15734-0.
- [We18] Wellbrock, N.; Eickenscheidt, N.; Hilbrig, L.; Dühnelt, P.-E.; Holzhausen, M.; Bauer, A.; Dammann, I.; Strich, S.; Engels, F.; Wauer, A.: Leitfaden und Dokumentation zur Waldzustandserhebung in Deutschland./, ISBN: 9781150416781, 2018, URL: https://www.openagrar.de/receive/openagrar_mods_00035667, visited on: 05/13/2023.

- [Wh16] White, J. C.; Coops, N. C.; Wulder, M. A.; Vastaranta, M.; Hilker, T.; Tompalski, P.: Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing* 42/5, Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/07038992.2016.1207484>, pp. 619–641, Sept. 2, 2016, ISSN: 0703-8992, URL: <https://doi.org/10.1080/07038992.2016.1207484>, visited on: 05/13/2023.

Towards Energy-Efficient Large-Scale Artificial Intelligence for Sustainable Data Centers

Dusan Dokic,¹ Hannah Stein,¹ Sabine Janzen,¹ Wolfgang Maaß¹

Abstract: The growing interest in AI services has led to a higher demand for computing power to train and execute complex AI models, causing a surge in power consumption in data centers. Together with rising costs for electricity, gas, petroleum, and coal, and the national target for climate neutrality of data centers by 2027, the ability to operate data centers economically is threatened in Germany. To address these issues, a pressing need to improve the sustainability of data centers and that of artificial intelligence. This paper proposes a roadmap to develop sustainable and resource-efficient data centers and AI systems. The roadmap includes four key building blocks: sustainable data centers, AI algorithms, AI sustainability framework, and economic efficiency analysis. Each building block poses pivotal research questions grounded in contemporary literature to guide the pursuit of environmental sustainability in data centers and AI.

Keywords: Data Centers; Sustainable AI; Energy-Efficient AI Algorithms

1 Introduction

The electricity demand of German data centers has increased by 30 % since 2016 and has more than doubled in the past ten years - from 5.8 billion in 2010 to 16 billion kWh in 2020 [St20, Hi22b], as three quarters of all companies in Germany use data centers for operating and developing business-critical IT applications [Bu21, Hi22b]. The training of artificial intelligence (AI) models can be identified as one of the main drivers of this growth as training large AI models can emit the equivalent of five SUV lifetimes (284 tons of CO₂) [Ne22].

For this reason, the energy consumption of AI models must be addressed as they require more computing power as their complexity increases [SGM19]. Solutions to increase the efficiency of the models are necessary to manage the demand for computing power and minimize the resulting environmental impact. In order to achieve this, tools for measuring and monitoring environmental impact must be used in conjunction with new technologies that surpass the limitations of previous hardware [Th22, Wa20]. Currently, companies developing AI-based services for innovative business models, have no tools to help them to manage the implementation including training and operation of the AI models via energy and sustainability monitoring [Bo20].

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Smart Service Engineering, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Deutschland, {dusan.dokic;hannah.stein;sabine.janzen;wolfgang.maass}@dfki.de

This work addresses these shortcomings by developing a roadmap for improving the sustainability of data center and AI-systems. In the roadmap we outline the necessary research that has yet to be undertaken, such as the development of a concept for the sustainable, resource-efficient design of data centers and AI systems as well as a framework that can be used to measure the sustainability of AI systems so that practitioners can make more informed decisions on the application of AI models in favor of a more sustainable solution, environmentally and ecologically.

The paper is structured as follows: Building on the state of the art on the sustainability of data centers and AI, we propose a research road map consisting of four building blocks. We suggest possible pathways for further research including possible research questions. We conclude the paper by summarizing and giving an outlook on our future work.

2 State of the Art

2.1 Sustainability of Data Centers

Numerous factors influence the sustainability of data centers. In addition to the energy efficiency of the hardware installed in the servers (e.g. GPU, CPU), cooling and energy management systems also play a major role [LDM20]. [FT12] provide an overview of the main components responsible for the majority of energy consumption in data centers. In addition, the share of renewable energy in the total consumption and metrics to measure sustainability represent major factors [Wh14]. [LMG18] propose a model for assessing the environmental impact and operational efficiency of data centers. Research in sustainable data centers has already produced numerous metrics to measure data center energy efficiency in an attempt to measure and manage energy consumption [Wh14] such as the power usage effectiveness value [Th07] or the green energy coefficient [Th14]. However, these metrics are criticized for their limited view as they only include the operating phase of the data centers [Wh14]. For this reason, life-cycle-based approaches are increasingly adopted in an attempt to consider data center sustainability across all phases [Gu22, WAS15a, WAS15b]. Despite current efforts to make data centers more sustainable, existing hardware (CPU, GPU, TPU) is reaching its limits in terms of efficient training of ever larger AI models [Th22]. Because of this, research is investigating new hardware technology such as neuromorphic chips; these are designed to mimic the structure and functioning of the human brain and promise efficiency gains of up to 50 % for training and 80 % in inference for certain AI models [Su20]. Despite the promising positive impacts of neuromorphic technologies, the research is still at a beginning stage and more research into the topic is needed [Sc22].

2.2 Sustainability of AI

According to [vW21], Sustainable AI is ‘a movement to promote change throughout the life-cycle of AI products towards greater ecological integrity and social justice’ and is

divided into ‘AI for sustainability’ and ‘sustainability of AI’. Academia currently focuses on the aspect of ‘AI for sustainability’ and explore how the use of AI can mitigate the effects of climate change [SM22, OE22]. While concentrating on the positive effects AI applications can have on the environment, the resource consumption required to develop and use AI are neglected. ‘Sustainability of AI’ addresses this issue and discusses principles for developing sustainable AI [MI22]. Building on this, [Li22] offered a preliminary approach for implementing an AI life-cycle assessment (LCA) in practice. Their work is foundational in defining the system boundary in the context of AI LCA, contributing significantly to our understanding of AI’s environmental impacts. In a recent paper, [RSC23] review the current body of knowledge and examine 98 paper relating to Sustainable AI or „Green AI“ and conclude that there is little work considering the environmental impact of AI as a whole, apart from energy use. They highlight the lack of tools to implement Sustainable AI and urge researchers to use the current body of knowledge to develop such tools for practitioners.

2.3 A Framework for Measuring the sustainability of AI

The discussed trade-off between ‘AI for Sustainability’ and ‘Sustainability of AI’ poses forces decision makers to carefully consider all positive and negative environmental impact before introducing AI applications [ZGF23]. However, in order to do this, they need information about the consumption and environmental impact of the AI application. In this regard, some scientific works can already be found that aim to measure the power consumption of different algorithms or AI models [He20, Me23, Mi21] and methods to improve energy efficiency during training by applying so-called green coding principles [Gr22] or optimized ‘work flows’ [LCC22]. In addition, work such as that of [Ka22] shows that frameworks for capturing the environmental impact of AI are needed and must capture multiple levels and factors than just the operational phase and power consumption. This finding, combined with the fact that current research in this area primarily lacks tools to implement current research [RSC23], leads us to conclude that a framework is needed that captures the environmental impact of AI systems in their entirety and provides decision support for the sustainable use of AI applications.

2.4 Economic Efficiency Analysis of Sustainable Data Center

Artificial intelligence is already embedded in the business models of many companies and is also already being used to make them more sustainable [To20]. In their comprehensive review on AI and sustainable business models, [Di20] found that it is mostly used to improve sustainability aspects but also aspects of ethics and finance. However, in the context of data centers, it is not only a matter of integrating business models with the help of AI, but also of taking into account the influence of AI as a product, which is directly related to the costs of a data center [Na19], and combining these with sustainability aspects of other business model components such as the building location and technical equipment [LDM20]. To this

end, standards and norms already exist for the operation of conventional data centers from which business models can be derived [DI19]. But there are only a few studies that analyze and evaluate the efficiency of "green data centers"[Pe16, Sh17] and conventional data center models only consider proven hardware such as CPU and GPU. An economic analysis of new hardware such as NPUs [Sc22] for use in the data center, is currently lacking. Even though, the resulting efficiency gains [Su20] could strongly influence the business model, especially of AI-focused data centers, and improve previously unprofitable lines of business [Da09].

3 A Research Roadmap towards Sustainable Data Centers and AI Systems

Building on related work and current shortcomings, we propose a research roadmap for future work in the development of sustainable AI and data centers. We propose four building blocks, namely the conceptualization of sustainable data centers, frameworks for measuring the sustainability of AI, development of sustainable AI algorithms as well as economic efficiency analyses. We summarize the research roadmap in table 1, including short descriptions of the building blocks and potential research questions.

3.1 Conceptualizing Sustainable Data Centers

To conceptualize sustainable data centers, hardware and software-related aspects need to be considered. We argue that sustainable data centers should appraise the utilization of neuromorphic chip technologies in the form of Neuromorphic Processing Units (NPU) as conservative estimates expect NPUs to reduce data center energy consumption up to 47.5 % [Hi22a]. In terms of hardware, hybrid architecture should be exploited, allowing traditional GPU/CPU/TPU-based data centers to be combined with NPUs. Additionally, it needs to be investigated for which use cases classic GPUs can be completely replaced by NPUs, so that a concept for a purely AI focused data center for an energy-efficient training of large, resource-intensive AI models such as Natural Language Processing (NLP) and Visual Computing (VC) can be designed. This leads to the following research questions: RQ1: 'How can NPUs be integrated into existing data center architectures to enable a hybrid concept?' and RQ2: 'Under what circumstances does replacing existing hardware in data centers like GPUs with NPUs have a positive impact on the environment?'

3.2 Sustainable AI Algorithms

The 'sustainability of AI' is becoming increasingly important due to larger and larger AI models with billions of parameters that need to be trained and adapted and consume more and more resources [SGM19]. Neuromorphic chip architectures have potential to reduce

this consumption, but the algorithms for these architectures are not yet as mature as for traditional chip architectures and current algorithms still need to be adapted to the new architecture. Therefore, a re-implementation of traditional AI algorithms on neuromorphic chip architectures is required, taking into account green-coding principles [Gr22], in order to achieve improved energy efficiency and increased sustainability of the developed algorithms. The research questions that need to be answered are therefore RQ3: 'How can traditional AI algorithms be adapted to neuromorphic hardware?' and RQ4: 'Are AI algorithms implemented on neuromorphic hardware more sustainable or energy efficient in training and inference than traditional AI algorithms?'

3.3 The AI Sustainability Framework

Global challenges such as climate change can only be addressed if companies become sustainable and environmentally friendly [vZvT21]. Based on the 17 SDGs of the United Nations, it was found that AI can have a positive impact on a large part of the SDGs [Sc23]. However, the 'lack of understanding of the use of AI in the primary and supporting functions of organizations' [SM22] and the lack of measurement capabilities to verify the environmental sustainability of AI systems [OE22] currently prevent the exploitation of the positive impact AI can have on the environment. Since the data processing of AI systems requires energy, but their application can result in energy savings, this derives an optimization problem via an AI sustainability framework model. The application of the AI-SF requires measurement data collected at different levels: the technical level, the application level and the systemic level [Ka22]. The technical level primarily considers the computing power that must be expended to develop an AI model and apply it. Here, the individual life cycles of model development such as data collection, training, inference and adaptation are considered. In addition to the computing power, other factors need to be taken into account, such as building technology, sustainability of energy used [DI19], total costs of the hardware and software used and personnel deployment [DI17]. The application level considers the positive and negative impacts that result from the use of AI [Ka22]. In addition to direct effects, the use of AI also entails indirect effects, which have an impact on the environment at a systemic level [BM22].

To implement such a framework, it requires the design of models to measure the various aspects of sustainability of AI systems, such as emissions and energy and resource consumption [Ro21]. The measurement models must be able to be used generically for classical and innovative technical architectures and span the range of technology, application, and system. The integrated measurement model is iteratively designed using the framework model (AI-SF). From the AI-SF, an AI-based optimization model will be derived to make predictions about the sustainability of AI applications in AI data centers. This model will be tested and optimized on real data. On the decision-making level, the forecasts will be used to optimize AI applications with respect to their effects on the environment or rebound effects. Research questions will be RQ5: 'How can the overall environmental footprint of AI

applications be measured and optimized?'. These measurement models and the AI-SF will then be translated into software modules and a platform usable by end users in future work.

3.4 Economic Efficiency Analysis

The computational needs of AI systems are growing, leading to a quadrupling of data center power consumption by 2030 [An20]; raising sustainability concerns among policymakers. For example, the German government's coalition agreement stipulates that data centers must become carbon neutral by 2027, while the OECD is pushing to fill current gaps in measuring the environmental impact of AI, including addressing the aspect of computing power [OE22]. Still, efforts to increase energy efficiency in data centers have stagnated over the past decade, as further optimization would only allow marginal returns at high costs [Bi21]. Long-term environmental sustainability in data centers hinges on its integration into their business model, while considering economic factors. Utilizing energy-efficient neuromorphic chips for AI training can boost efficiency and profitability, also opening new business prospects for both newcomers and established entities. The resulting research questions that need to be answered are RQ6 'What economic aspects need to be considered when using neuromorphic hardware in data centers to make it profitable?', RQ7 'What risks to the business model of traditional data centers need to be considered when using neuromorphic hardware in data centers?', and RQ8 'What does an economically and environmentally sustainable business model of a data center based on neuromorphic hardware look like?'

4 Conclusion

The discourse on 'AI for sustainability' highlights the need for a balanced approach that focuses on the environmental benefits AI can bring and addresses the significant resource consumption involved in its development and application. To this end, future research is urged to work on the identified building blocks such as the AI sustainability framework, that optimizes energy savings from AI applications against the energy expenditure in AI data processing. In summary, achieving sustainable AI and data centers has many complex considerations and hurdles. Nevertheless, the potential rewards are equally significant, not only for the environment but also for the data center business model. This work aims to be a starting point for future research and offers a research roadmap, presenting key questions that must be addressed to achieve environmental sustainability of data centers and AI.

Acknowledgement

This work is part-funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) within the funding program "GreenTech innovation competition", managed by the DLR project management agency, under grant number 01MN23004A (www.escade-project.de).

Bibliography

- [An20] Andrae, A. S.: Hypotheses for primary energy use, electricity use and CO2 emissions of global computing and its shares of the total between 2020 and 2030. *WSEAS Transactions on Power Systems*, pp. 50–59, 2020.
- [Bi21] Bizo, D., Ascierio, R., Lawrence, A., Davis, J. (2021). Globale Studie des Uptime Institute zu Rechenzentren 2021. Retrieved May 9, 2023 from <https://uptimeinstitute.com/resources/asset/2021-data-center-industry-survey-de>.
- [BM22] Bundesministerium für Ernährung und Landwirtschaft (BMEL). (2023). Künstliche Intelligenz für eine nachhaltigere Landwirtschaft. Retrieved May 9, 2023 from <https://www.bmel.de/SharedDocs/Downloads/DE/Broschueren/k-i-fuer-nachhaltige-landwirtschaft.html>.
- [Bo20] Boll, S., Schnell, M., others. (2020). Working Group Business Model Innovation: With Artificial Intelligence to Sustainable Business Models [White Paper]. Retrieved May 9, 2023 from https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen_EN/AG4_WP_ES_with_ai_to_sustainable_business_models.pdf. White Paper.
- [Bu21] Bundesregierung. (2021). Klimaschutzgesetz 2021. Retrieved May 9, 2023, from <https://www.bundesregierung.de/breg-de/themen/klimaschutz/klimaschutzgesetz-2021-1913672>.
- [Da09] Data center insider. (2009). Cloud Computing ist kein profitables Geschäftsmodell für Provider – ein Plädoyer gegen den Einsatz im Rechenzentrum. Retrieved May 9, 2023 from <https://www.datacenter-insider.de/cloud-computing-ist-kein-profitables-geschaeftsmodell-fuer-provider-ein-plaedoyer-gegen-den-einsatz-im-rechenzentrum-a-167195/?p=2>.
- [DI17] DIN. (2017). DIN EN 50600-4-1. Retrieved May 9, 2023, from <https://www.vde-verlag.de/normen/0800409/din-en-50600-4-1-vde-0801-600-4-1-2017-06.html>.
- [DI19] DIN. (2019). DIN EN 50600-4-2 VDE. Retrieved May 9, 2023, from <https://www.vde-verlag.de/normen/0800596/din-en-50600-4-2-vde-0801-600-4-2-2019-08.html>.
- [Di20] Di Vaio, A.; Palladino, R.; Hassan, R.; Escobar, O.: Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, pp. 283–314, 2020.
- [FT12] Flucker, S.; Tozer, R.: Data Centre Energy Efficiency Analysis to minimize total cost of ownership. *Journal of Building Services Engineering Research & Technology*, pp. 103–117, 2012.
- [Gr22] GreenPrinciples. (2022). Principles of Green Software Engineering. Retrieved May 9, 2023 from <https://principles.green/>.
- [Gu22] Gupta, U.; Kim, Y. G.; Lee, S.; Tse, J.; Wei, G.; Brooks, D.: Chasing Carbon: The Elusive Environmental Footprint of Computing. *IEEE Micro*, 42, 2022.
- [He20] Henderson, P.; Hu, J.; Romoff, J.; Brunskill, E.; Jurafsky, D.; Pineau, J.: Towards the systematic reporting of the energy and carbon footprints of machine learning. *Journal of Machine Learning Research*, 21, 2020.

- [Hi22a] Hiltsher, J. (2022) Künstliche Neuronen bis zu 16 Mal effizienter als GPUs. Retrieved May 9, 2023 from <https://www.golem.de/news/neuromorphic-computing-kuenstliche-neuronen-bis-zu-16-mal-effizienter-als-gpus-2205-165661.html>.
- [Hi22b] Bitkom (2022). Rechenzentren in Deutschland Aktuelle Marktentwicklungen. Retrieved May 9, 2023, from <https://www.bitkom.org/sites/main/files/2022-02/10.02.22-studie-rechenzentren.pdf>.
- [Ka22] Kaack, L.; Donti, P.; Strubell, E.; Kamiya, G.; Creutzig, F.; Rolnick, D.: Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 2022.
- [LCC22] Lee, D.; Chen, Y.; Chao, S.: Universal Workflow of Artificial Intelligence for Energy Saving. *Energy Reports*, pp. 160–167, 2022.
- [LDM20] Laurent, A.; Dal Maso, M.: Environmental sustainability of data centres: A need for a multi-impact and life cycle approach. Report, 2020.
- [Li22] Ligozat, A.; Lefevre, J.; Bugeau, A.; Combaz, J.: Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions. *Sustainability*, (9), 2022.
- [LMG18] Lykou, Georgia; Mentzelioti, Despina; Gritzalis, Dimitris: A new methodology toward effectively assessing data center sustainability. *Computers Security*, pp. 327–340, 2018.
- [Me23] Menghani, G.: Efficient Deep Learning: A Survey on Making Deep Learning Models Smaller, Faster, and Better. *ACM Computing Surveys*, pp. 1–37, 2023.
- [Mi21] Mikić, V.; Ilic, M.; Zakić, A.; Zlatkovic, D.: Green Cloud Computing in the Purpose of Energy Efficiency. *International Journal of Electrical and Computer Engineering (IJECE)*, pp. 4809–4817, 2021.
- [MI22] Inhabitat. (2022). MIT moves toward greener, more sustainable artificial intelligence. Retrieved May 9, 2023 from <https://inhabitat.com/mit-moves-toward-greener-more-sustainable-artificial-intelligence/>.
- [Na19] Forbes (2019). Exploring The Impact Of AI In The Data Center May 9. , 2023 from <https://www.forbes.com/sites/cognitiveworld/2019/05/31/exploring-the-impact-of-ai-in-the-data-center/>.
- [Ne22] New Scientist. (2022). Creating an AI can be five times worse for the planet than a car. Retrieved May 9, 2023 from <https://www.newscientist.com/article/2205779-creating-an-ai-can-be-five-times-worse-for-the-planet-than-a-car/>.
- [OE22] OECD. (2022). Measuring the environmental impacts of artificial intelligence compute and applications (No. 341). Retrieved May 9, 2023 from <https://www.oecd-ilibrary.org/content/paper/7babf571-en>.
- [Pe16] Pegasus, P.; Varghese, B.; Guo, T.; Irwin, D.; Shenoy, P.; Mahanti, A.; Culbert, J.; Goodhue, J.; Hill, C.: Analyzing the Efficiency of a Green University Data Center. In: *Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering. ICPE '16*, Association for Computing Machinery, New York, NY, USA, p. 63–73, 2016.
- [Ro21] Rohde, F., Wagner J., Reinhard P., Petschow U., Meyer A., Voß M., Mollen A.: , Nachhaltigkeitskriterien für künstliche Intelligenz. Umweltbundesamt, 2021.

- [RSC23] Roberto.; S., June; Cruz, Luis: A Systematic Review of Green AI. WIREs Data Mining and Knowledge Discovery, 2023.
- [Sc22] Schuman, C. D.; Kulkarni, S. R.; Parsa, M. et al.: Opportunities for neuromorphic computing algorithms and applications. Nat Comput Sci, pp. 10–19, 2022.
- [Sc23] Schoormann, T.; Strobel, G.; Müller, F.; Petrik, D.; Zschech, P.: Artificial Intelligence for Sustainability—A Systematic Review of Information Systems Literature. Communications of the Association for Information Systems, 52, 2023.
- [SGM19] Strubell, E.; Ganesh, A.; McCallum, A.: Energy and policy considerations for deep learning in NLP. arXiv preprint arXiv:1906.02243, 2019.
- [Sh17] Sharma, P.; II, P.; Irwin, D.; Shenoy, P.; Goodhue, J.; Culbert, J.: Design and Operational Analysis of a Green Data Center. IEEE Internet Computing, pp. 16–24, 01 2017.
- [SM22] Schober, L.; Mattke, J.: AI for Sustainability in Organisations: A Literature Review. (2022). AMCIS 2022 Proceedings, 15, 2022.
- [St20] Statista. (2020). Stromverbrauch von deutschen Rechenzentren und kleineren IT-Installationen pro Jahr. Retrieved May 9, 2023, from <https://de.statista.com/infografik/27846/stromverbrauch-von-deutschen-rechenzentren-und-kleineren-it-installationen-pro-jahr/>.
- [Su20] Subramoney, A.; Nazeer, K. K.; Schöne, M.; Mayr, C.; Kappel, D.: EGRU: Event-based GRU for activity-sparse inference and learning. 2020.
- [Th07] The Green Grid (2007). Green grid metrics: describing datacenter power efficiency. May 9, 2023 from http://www.thegreengrid.org/~media/WhitePapers/Green_Grid_Metrics_WP.ashx?lang=en.
- [Th14] Harmonizing global metrics for data center energy efficiency. Global taskforce reaches agreement regarding data center productivity. Retrieved May 9, 2023 from <http://www.thegreengrid.org/library-and-tools.aspx>.
- [Th22] Thompson, N. C.; Greenewald, K.; Lee, K.; Manso, G. F.: The computational limits of deep learning. 2022.
- [To20] Toniolo, K.; Masiero, E.; Massaro, M.; Bagnoli, C.: Sustainable Business Models and Artificial Intelligence: Opportunities and Challenges. pp. 103–117, 04 2020.
- [vW21] van Wynsberghe, A.: Sustainable AI: AI for sustainability and the sustainability of AI. AI Ethics, pp. 213–218, 2021.
- [vZvT21] van Zanten, J.; van Tulder, R.: Improving companies' impacts on sustainable development: A nexus approach to the SDGS. Business Strategy and the Environment, pp. 3703–3720, 2021.
- [Wa20] Wang, Y.; Wang, Q.; Shi, S.; He, X.; Tang, Z.; Zhao, K.; Chu, X.: Benchmarking the Performance and Energy Efficiency of AI Accelerators for AI Training. In: 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID). pp. 744–751, 2020.
- [WAS15a] Whitehead, B.; Andrews, D.; Shah, A.: Assessing the environmental impact of data centres part 2: Building environmental assessment methods and life cycle assessment. International Journal of Life Cycle Assessment, pp. 332–349, 2015.

- [WAS15b] Whitehead, B.; Andrews, D.; Shah, A.: The life cycle assessment of a UK data centre. *The International Journal of Life Cycle Assessment*, pp. 332–349, 2015.
- [Wh14] Whitehead, B.; Andrews, D.; Shah, A.; Maidment, G.: Assessing the environmental impact of data centres part 1: Background, energy use and metrics. *Building and Environment*, pp. 151–159, 2014.
- [ZGF23] Zhao, J.; Gómez F., B.: Artificial Intelligence and Sustainable Decisions. *Eur Bus Org Law Rev*, pp. 1–39, 2023.

Appendix

Building Block	Short description	Possible Research Questions
Sustainable Data Centers	Designing a concept for sustainable data centers based on neuromorphic hardware for a 'hybrid' and 'cloud' architecture	How can NPUs be integrated into existing data center architectures to enable a hybrid concept? Under what circumstances does replacing existing hardware in data centers like GPUs with NPUs have a positive impact on the environment?
Sustainable AI Algorithms	Re-implementation of traditional AI algorithms on neuromorphic chip architecture and evaluation of efficiency gains	How can traditional AI algorithms be adapted to neuromorphic hardware? Are AI algorithms implemented on neuromorphic hardware more sustainable or energy efficient in training and inference than traditional AI algorithms?
AI Sustainability Framework	Solving the optimization problem of positive environmental AI impact and the necessary resource consumption by measuring and quantifying the impact and costs on the technical, the application and the systemic level	How can the overall environmental footprint of AI applications be measured and optimized?
Economic Efficiency Analysis	Integrating environmental sustainability into economically sustainable business models to ensure	What economic aspects need to be considered when using neuromorphic hardware in data centers to make it profitable? What risks to the business model of traditional data centers need to be considered when using neuromorphic hardware in data centers? What does an economically and environmentally sustainable business model of a data center based on neuromorphic hardware look like?

Tab. 1: Building blocks of sustainable data centers and AI system and open research questions

Smartmonitoring – eine webbasierte Plattform für das Monitoring von Umweltdaten aus verschiedensten Schnittstellen und Sensoren zur Nutzung für studentische Projekte und Forschungsprototypen

Florian Fehring¹, Grit Behrens², Jannis Thiel³

Abstract: In diesem Beitrag werden das Konzept und die bisherige Implementierung einer integrierenden Softwarelösung für die Erfassung, Visualisierung und Nutzung von Daten aus unterschiedlichen Quellen dargestellt. Ein besonderer Fokus liegt dabei auf Anwendungen im Bereich der Umweltinformatik, die beispielhaft gezeigt werden und auf der Integration der Lösung in Lehre und Forschung.

Keywords: Umwelt Monitoring, Datenerfassung, Datenmanagement, interaktive GIS-Karte, Online Visualisierung,

Einleitung

Dieser Artikel beschreibt eine webbasierte Monitoring-Plattform, die auf einem universellen Datenformat [Fe20] basiert, eine GIS-basierte Speicherung und Darstellung unterstützt und die Integration von Datenquellen verschiedenster Art ermöglicht. Mit der Plattform können u.a. Ertrags-Daten aus dem Monitoring von PV-Anlagen, Einstrahlungs- und Wetterdaten, Wasserqualitätsdaten, Luftqualitätsdaten und Raumklimadaten aus z.B. Sensoren (RaspberryPi, Sunsniffer, Einstrahlungssensoren, Wetterstationen, etc.), REST-Schnittstellen zu Internetdiensten (OpenWeatherMap, SensorWeb, Umweltportal NRW, etc.) erfasst, organisiert und ausgewertet werden. Sie enthält nutzerfreundliche Schnittstellen zum Zugriff auf die verwalteten Daten. Die Plattform wird kontinuierlich im Rahmen des Unterrichts erweitert und modernisiert und ermöglicht so das Erlernen von webbasierter Programmierung und Data-Mining-Technologien mit KI-Fragestellungen bei gleichzeitiger Sensibilisierung der Studierenden für umweltrelevante Fragestellungen.

Ziel der Plattform ist die Bereitstellung einer projektübergreifend nutzbaren Software, die den gesamten Prozess vom Datenerfassen bis zur Auswertung einfacher zugänglich macht. Damit soll der Aufwand beim Start neuer Projekte reduziert werden und insbesondere die wiederholte Implementierung von Standardvorgehensweisen wie Datenabruf und einfacher Visualisierung vermieden werden.

¹ Hochschule Bielefeld, Campus Minden, Artilleriestraße 9, 32427 Minden, florian.fehring@hsbi.de,

² Hochschule Bielefeld, Campus Minden, Artilleriestraße 9, 32427 Minden, grit.behrens@hsbi.de

³ Hochschule Bielefeld, Campus Minden, Artilleriestraße 9, 32427 Minden, jannis.thiel@hsbi.de

1 Vorhandene Lösungen

Die Literatur beschreibt verschiedenste Arten von GIS-basierten Monitoringplattformen in Bezug auf die Applikationsbreite und auf die unterstützten Technologien. Es gibt Systeme, die zum Teil für bestimmte Umwelthanwendungen entwickelt wurden, wie z.B. eine Applikation zur Prognose von Umweltverschmutzungen durch Bergwerke und Chemieindustrie [Be23] oder die Systeme ermöglichen die Verwaltung universeller Umweltdaten wie zum z.B. [Gi22], in dem eine breite GIS-Plattform für das Umweltmonitoring von Verschmutzungsdaten für Boden, Wasser und Luft aus industriellen Datenquellen und Citizen Science Projekten über verschiedene Schnittstellen eingebunden sind. Bei der Einbindung von Sensoren wird in der Literatur häufig die IoT-Technologie genutzt, um Realtime-Monitoring und eine Steuerung der Systeme zu ermöglichen, wie in [Ka22], oder man beschränkt sich auf webbasierte Schnittstellen zum Einlesen der Umwelt-Datenquellen.

Das hier beschriebene System verwaltet Umweltdaten auf Basis offener Webstandards, wobei Daten von Hardware wie RaspberryPi basierten Sensoren, Wechselrichtern, Wetterstationen etc. ebenso eingebunden werden können, wie Daten von FTP-Servern und REST-Schnittstellen.

1.1 Verwaltung von Komponenten

Sensoren, Serverhardware und Datenbanken werden oftmals mit Software zur Überwachung von IT-Infrastruktur überwacht. Beispiele sind hier z.B. Prometheus und Graphana. Diese Tools erstellen Metriken, die den Status der Infrastruktur wiedergeben. Es werden Dashboards eingesetzt, die den Status meist im zeitlichen Verlauf darstellen [Sh23]. Die zahlreichen Metriken sind für die Identifikation von Schwachstellen in der Infrastruktur (Speicherplatz, Prozessorauslastung) nützlich, jedoch für Datenanalysten unnötig.

Das hier vorgestellte SmartMonitoring konzentriert sich auf die Abbildung der Struktur und die Überwachung des Datenflusses. Es wird sichtbar gemacht, welchen Weg Messwerte durch die Infrastruktur nehmen und wo Daten ausblieben, für die Fehleranalyse kann anschließend auf die zuvor genannten, gut entwickelten Tools zurückgegriffen werden.

1.2 Datenerfassung

Die Datenerfassung im System kann mit Sensoren, manuell, aus den in Forschungsprojekten integrierten Citizen-Science Projekten und aus OpenData-Apis erfolgen. Viele Hersteller liefern ihre Sensoren mittlerweile mit eigenen Apps aus. Jedoch mangelt es, insbesondere bei kostengünstigeren Geräten für Endanwender, oft an offenen Schnittstellen. Andere Gerätegruppen wiederum erfordern Hardware und

Programmierkenntnisse, um sie erfolgreich in die Dateninfrastruktur zu integrieren. Das SmartMonitoring möchte hier eine leicht zu verwendende Brücke bauen, indem angeschlossene Sensoren über eine REST-API zugänglich gemacht werden. Die Erfassung von Daten erfolgte früher meist händisch und wird zunehmend digitalisiert. Historische Daten liegen häufig in Form von Excel-Sheets oder älteren Datenbankformaten vor. Der Import dieser Daten stellt einen regelmäßigen, zeitaufwändigen Arbeitsschritt in Forschungsprojekten dar. Das SmartMonitoring bietet mit seiner Oberfläche zum einen eine Möglichkeit einfach zu bedienende Eingabemaschinen zu erstellen und zum anderen über einen erweiterbaren Importer alte Datenbestände aufzunehmen.

Zunehmend werden OpenData- und durch CitizenScience-Projekte (z.B. sensor.community) generierte Daten in Forschungsprojekte einbezogen. Portale wie z.B. das Daten Portal Deutschland: govdata.de, bieten durchsuchbare Metakataloge. Die Verwendung der Daten erfolgt mit im Projekt implementierten Importern oder verschiedenen spezialisierten Programmen wie QGIS. SmartMonitoring unterstützt die Verwendung dieser Datenquellen mit Funktionalitäten, welche die Nutzung vereinfachen und die Integration mit anderen Daten erleichtern.

2 Architektur und Konzepte

2.1 Software-Architektur

Für das SmartMonitoring wurde das Konzept einer Abstraktions-Schichten-Architektur entwickelt. Dabei wird die Implementierung in den unteren Schichten immer abstrakter und unabhängig von einem konkreten Anwendungsfall. Dadurch ergibt sich ein deutlich reduzierter Implementierungsaufwand in der obersten Schicht, welche die anwendungsfallspezifische Geschäftslogik enthält, während die darunter liegenden Schichten in funktionalen Komponenten gegliedert werden können.

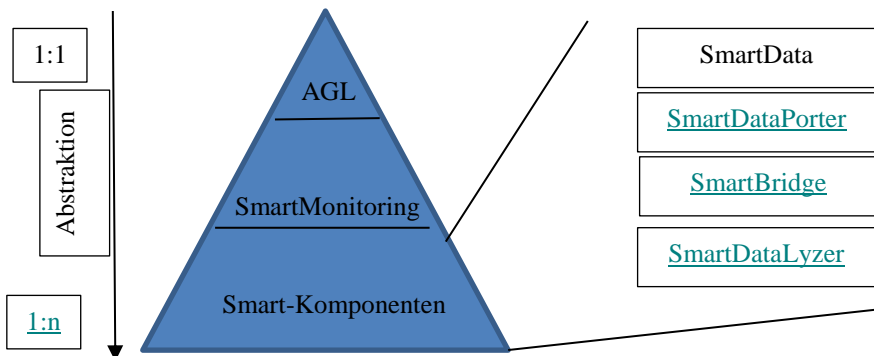


Abbildung 1: Abstraktions-Schichten-Architektur

Die hohe Abstraktion in der untersten Ebene lässt sich durch gewohnte Konzepte der objektorientierten Programmierung nicht erreichen. Auf dieser Ebene wird deshalb informationsorientiert gearbeitet. Informationen und Funktionen werden voneinander getrennt. Funktionen in Komponenten zu funktionalen Einheiten zusammengefasst. Die Information als Basis wird durch ein Objekt mit unbestimmter Anzahl von Attributen abgebildet. Komponenten können alle Informationen verarbeiten, deren (Teil-)Signatur zur Komponente passt. Eine Komponente kennt die Bedeutung der Daten nicht. Beispielsweise könnten die Datensätze die das Navigationsmenü definieren in der selben Auflistung dargestellt werden, wie aufgenommene Messwerte. Es ist die Aufgabe der darüber liegenden Schichten die Komponenten sinnvoll einzusetzen.

Die zweite Schicht, in diesem Fall das SmartMonitoring, implementiert Funktionen in allgemeiner Art, stellt die sinngemäße Verwendung der Daten sicher und gibt Daten aus einer bestimmten Datenquelle eine allgemeine Bedeutung. So liefert das SmartMonitoring die Daten aus der Navigations-Tabelle an die Navigations-Komponente und Datensätze, die Sensoren oder Organisationseinheiten beschreiben, werden als ObservedObjects in Darstellungen verwendet, die beispielsweise in der Objektübersicht oder Objektkarte angezeigt werden. Gesammelte Daten werden auf dieser Ebene noch ganz allgemein als Messwerte betrachtet. Dies erlaubt die Wiederverwendung der gesamten Applikationsfunktionen. Alle Funktionen können unabhängig von der Bedeutung der Daten verwendet werden.

Die dritte Schicht enthält den wenigsten Programmcode und kann zum überwiegenden Teil durch Konfiguration, anstelle von Implementierung genutzt werden. Ihre Aufgabe ist es den Daten ihre Bedeutung zu verleihen, nicht sinnvoll nutzbare Funktionen auszublenden und eine verständliche Oberfläche für den Endnutzer zu schaffen. Dazu können die allgemeinen Bezeichnungen, wie beispielsweise „Beobachtbare Objekte“, zu konkreten Fachbegriffen transformiert werden, beispielhaft „Solarmodule“. Nicht sinnvoll einsetzbare Funktionen können deaktiviert werden, z.B. die Kartenanzeige (datamap) für die einzelnen Ertragsdatensätze, während andere konfiguriert werden können, z.B. Einfärbung nach dem Ertragswert auf der Solarmodulkarte (objectmap).

Neben der Abstraktions-Schichten-Architektur wird eine Model-View-Algorithm Architektur verwendet, die in allen drei Abstraktions-Schichten Verwendung findet. Das Model entspricht der zuvor genannten Information und existiert daher als simples, datentragendes Objekt. Zusammen mit den Algorithmen zum Datenabruf bilden sie den Model-Code. Als View werden alle Menschen- oder Maschinen-lesbaren Schnittstellen bezeichnet. Diese teilen sich wiederum in einen Controller- und einen Design-Aspekt auf. Die bislang nur wenig genutzte Algorithm-Schicht beherbergt auf alle Daten anwendbare Algorithmen und wird zukünftig einen Großteil der Daten-Analyse umfassen.

2.2 Softwarekomponenten und Datenhaltung

Die Verwendung des Microservices Konzepts sorgt für eine funktionale Unterteilung und lose Kopplung der Komponenten. Damit lässt sich das SmartMonitoring für den Anwendungsfall passend zusammenstellen. Einzelne Komponenten können zudem auf unterschiedlichen Servern gehostet werden, wodurch das Gesamtsystem skalierbar wird. Die Kommunikation der Komponenten untereinander erfolgt über REST-Schnittstellen.

Das Frontend zum SmartMonitoring kann Daten aus REST-Schnittstellen (json), JSON-Dateien und JavaScript-Objekten gleichwertig verwenden. Shape-Dateien können direkt auf Kartendarstellungen genutzt werden, während andere Datenformate über den SmartDataPorter in die im Hintergrund arbeitende Postgres-Datenbank importiert werden können. Der Importer kann leicht durch Parser für weitere Datenformate ergänzt werden.

3 Funktionalitäten und Anwendungsgebiete

Das SmartMonitoring ist als allgemeine Software zur Unterstützung von der Datenerfassung bis zu deren Auswertung konzipiert und bietet somit eine Vielzahl von Funktionen. Nachfolgend werden die wichtigsten Funktionen und ihre bisherigen Anwendungsfelder aufgeführt.

- Verwaltung von Geräten, Sensoren, Datenquellen und -zielen
- Datenerfassung von Sensoren
- Erfassung von OpenData aus REST-APIs
- Darstellung von Gerätepositionen
- Schematische Darstellung von Geräten an Messstationen
- Labeln von Datensätzen durch Nutzer
- Erfassung von Multimedia
- Datenvisualisierung als Diagramme und auf Karten 2D/3D

Sensoren können einzeln verwaltet, in Gruppen organisiert oder zu Geräten zusammengefasst werden. Andere Datenquellen können analog zu Sensoren genutzt werden. Berechnungsergebnisse werden in Datenzielen festgehalten, die wiederum als Datenquellen für weitere Algorithmen dienen können.

Das Erfassen von Daten einer Wetterstation, mehrerer Einstrahlungsmessgeräte (Pyrheliometer, Pyranometer, Referenzzellen) sowie mehrere PV-Kennlinien verschiedener Module im Langzeitfeldversuch stellte den ersten Anwendungsbereich dar [We16].

Zusätzlich zum Kennlinienmesstand wurde nachfolgend eine Instanz des SmartMonitoring aufgesetzt, um die 15-minütigen Moduldaten (Temperatur, Spannung) der PV Dachanlage und Stringdaten (Stromstärke) sowie Einstrahlungsmessungen mit Referenzzellen zu erfassen. [He19].

Am Host angeschlossene Sensoren können mit geringem Aufwand über eine REST-API zugreifbar gestaltet werden. Diese REST-API dient ebenso dem Frontend zur Darstellung von Momentanwerten. Im Hintergrund können unter anderem Python- oder Shell-Skripte sowie Java-Programme genutzt werden.

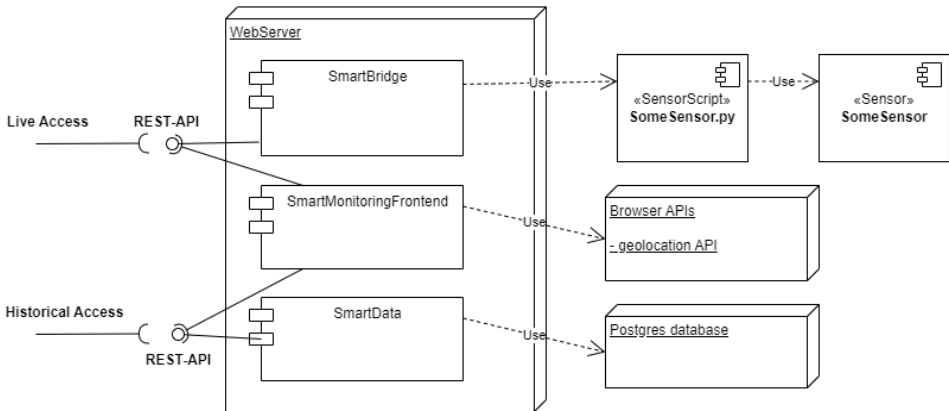


Abbildung 2: Komponenten und Schnittstellen live Datenerfassung

Die Sensoranbindung wurde im Rahmen des Projekts ENVIRON (2018-21) entwickelt, dabei wurden Indoor-Luftqualitätsmessungen (Luft- und Wandtemperatur, Luftfeuchtigkeit, CO₂-Gehalt) mit RaspberryPi-Messgeräten in 30 Bielefelder Wohnungen durchgeführt [Be19],

Über die REST-API ist der Abruf der Daten grundsätzlich immer möglich. Die Suche nach erfassten Daten, statistische Auswertungen, eine Vorschau, sowie Anleitungen zum Abruf und zur Einbindung sind direkt in der GUI verfügbar.

Sofern Geräte über eine Geoposition verfügen, können diese auf der Gerätekarte, auch gemeinsam mit anderen Informationen, dargestellt werden. Hier können die Marker entsprechend eines Messwerts eingefärbt werden. Auch können über diese Karte Messstellen hinzugefügt werden.

Diese Erweiterung wurde im Rahmen des Unterrichts von Studierenden hinzugefügt und soll zukünftig vor allem für Projekte mit Bezug zum Gewässerschutz Verwendung finden.

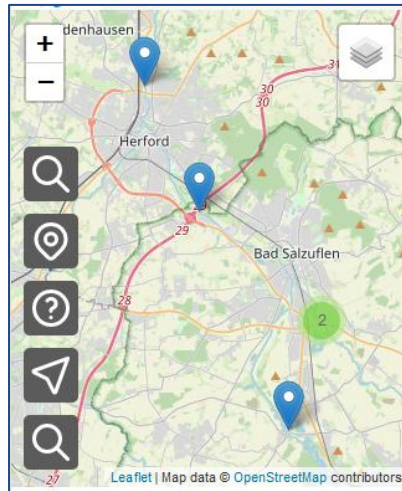


Abbildung 3: Darstellung der Messstationen für Pegelstände und Wasserqualität

Die Zusammenhänge zwischen Geräten können durch automatisch generierte schematische Darstellungen sichtbar gemacht werden. Diese können bei Bedarf auch manuell angepasst werden, um sie verständlicher zu gestalten. Diese Funktion wurde für die Darstellung der Photovoltaikanlage des Campus Minden im Unterricht entwickelt.

Datensätze können gekennzeichnet werden. Dies kann manuell über die GUI erfolgen, oder über eine entsprechende REST-API. Die verfügbaren Label können vom Administrator fest vorgegeben und bei Bedarf ebenfalls über die REST-API erzeugt werden.

Mittels dieser Label soll es zukünftig möglich sein, Datenpipelines zu erstellen, welche Datenquellen Eingangsfunktionen und Datenzielen Ausgangsfunktionen zuordnen und über die Label die zu verwendenden Daten steuern.

LABEL		ID	TEMP	SYNCED
UNVOLLSTÄNDIG +	Auf Karte zeigen Route zeigen	42	3.6	true
+ Ein Label hinzufügen	Auf Karte zeigen Route zeigen	43	2.2	true
+ Unvollständig	Auf Karte zeigen Route zeigen	44	2.9	true
	Auf Karte zeigen Route zeigen	45	2.8	true

Abbildung 4: Labeln von Datensätzen in der Datensatzliste

Bilder und Videos können ebenfalls mit dem SmartMonitoring aufgenommen und verwaltet werden. Ursprünglich sollten damit Aufnahmen durch eine Infrarot-Kamera von Solarmodulen verwaltet und Hotspots automatisch auf den Bildern markiert werden. Zukünftig können damit z.B. auch automatisch gefertigte Aufnahmen einer mobilen Messstation zum Pflanzenbewuchs in der Nähe von Hochwassermessstellen aufgenommen werden.

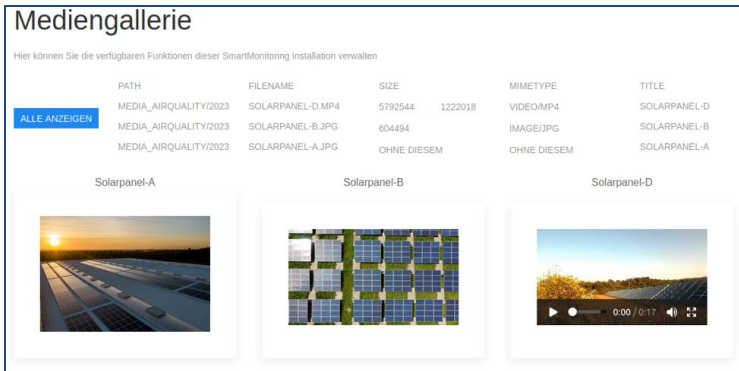


Abbildung 5: Mediengalerie für Bilder einer PV Anlage

Daten können mit Hilfe von Charts (Linien-, Balken-, oder Kreisdiagramm) und Piktogrammen visualisiert werden. SmartMonitoring ermittelt ohne vorherige Konfiguration eine geeignete Einstiegsdarstellung, die nachfolgend durch den Nutzer angepasst werden kann. Abbildung 5 zeigt die Verwendung des Charts im Projekt zur Erstellung eines mobilen Low-Cost-Feinstaubensors, sowie die Darstellung eines Wetterdatensatzes als Piktogramm.

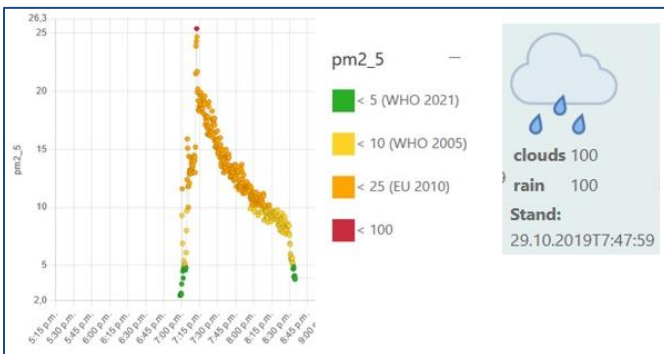


Abbildung 6 Generierte Darstellung von Messwerten als Diagramm und als Piktogramm

Werden in Datensätzen Geodaten ermittelt, können die Datensätze auf einer Karte dargestellt werden und entsprechend einer gegebenen Konfiguration eingefärbt werden. Es werden 2D und 3D Darstellungen unterstützt.

In Studentenprojekten wurde die Möglichkeit entwickelt, Luftqualitätsdaten wie die Feinstaubkonzentration mit einer kompakten, mobilen Messstation zu erfassen. Das SmartMonitoring wurde mit Funktionen für den mobilen Einsatz ergänzt und das in ENVIRON entwickelte Messsystem für diesen Einsatzzweck verwendbar gemacht.



Abbildung 2: Darstellung von Feinstaubmesswerten auf der Karte

4 Fazit

Mit dem SmartMonitoring wird eine Software entwickelt, die besonders für kleine und mittlere Forschungsprojekte interessant ist und den Aufwand, der zur Einrichtung einer dokumentierten Datenerfassung notwendig ist, erheblich reduziert. Durch die informationsorientierte Sicht ist das System flexibel einsetzbar und durch die Konfigurierbarkeit auf den Anwendungsfall anpassbar.

Das System soll auch in den kommenden, eigenen Forschungsprojekten eingesetzt und evaluiert werden. Anschließend ist die Veröffentlichung und Fortführung als OpenSource Projekt geplant.

Literaturverzeichnis

- [Be19] G. Behrens, K. Schlender, M. Brandt, Ph. Kösling “Überwachtes maschinelles Lernen von Personentypen zum Energieverbrauch durch KI-Algorithmen zur Reduzierung der CO2 Emissionen und zur Vermeidung schlechter Raumluftqualität durch Interventionen”, Informatik 2019, UINW 2019 - 7. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel, Kassel 2019, in proceedings: Adjunct Proceedings of the 33rd Enviroinfo conference, Shaker Verlag 2019 ISBN 978-3-8440-6847-4
- [We16] G. Behrens, J. A. Weicht, R. Rasch, G. Behrens, F. U. Hamlemann: „Performance characterization oft thin-film-silicon based solar modules under clouded clear sky conditions in comparision to crystalline silicon modules “, Electronic Materials Letters, July 2016, Volume 12, Issue 4, pp 468-471, 2016 Springer International Publishing, The Korean Institute of Metals and Materials, Print ISSN1738-8090
- [Me21] F. Meier, J. Rüter, G. Behrens, M. Diehl „Intelligente PV-Moduldatenbank-Applikation für Predictive Maintanance von PV-Anlagen“, UINW2021 9. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel, Berlin 2021, Berlin 2021, LNI – Proceedings - Series of GI, ISBN 978-3-88579-708-1
- [He19] S. Hempelmann, M. Grüneis, G. Behrens, Frank U. Hamelmann “Investigating Partial Shadowing of PV Module at Solar Cell Level” Proceedings of 35 st European Photovoltaic Solar Energy Conference and Exhibition, Marseille, 2019
- [Fe20] F. Fehring, G. Behrens, L. Stoll, B. Struzek, A. Rogalski „Adaptive Fusion und Visualisierung inhomogener Umweltdaten aus öffentlichen Quellen“. ISBN 978-3-8440-7579-3 Reihe Umweltinformatik, Shaker-Verlag 2020, Virtueller Workshop 2020 der GI-Fachgruppe 4.6.3 / ASIM-FG SUG „Simulation in den Umwelt- und Geowissenschaften“
- [Gi22] Giglione, G.; Annibaldi, A.; Iaccarino, A.; Capancioni, R.; Borghini, G.; Ciabattini, F.; Illuminati, S.; Pace, G.; Memmola, F.; Giantomassi, G. An Integrated Web-Based GIS Platform for the Environmental Monitoring of Industrial Emissions: Preliminary Results of the Project. *Appl. Sci.* **2022**, *12*,3369. <https://doi.org/10.3390/app12073369>
- [Be23] N. Bernatska, E. Dzhumelia, V. Dyakiv, O. Mitryasova, I. Salamon „Web-Based Information and Analytical Monitoring System Tools – Online Visualization and Analysis of Surface Water Quality of Mining and Chemical Enterprises“. *Ecological Engineering & Environmental Technology* 2023, *24*(3), 99–108 <https://doi.org/10.12912/27197050/159885> ISSN 2719–7050, License CC-BY 4.0
- [Ka23] M. Kandil, M. Aly, M. Akl, M. Iqbal “Monitoring System for a Hybrid Photovoltaic-Diesel Power System: Web-Based SCADA Approach“. *Jordan Journal of Electrical Engineering* ISSN (print): 2409-9600, ISSN (online): 2409-9619 , January 2023
- [Sh23] S. Shethi, 12 Beste Open Source-Überwachungssoftware für die IT-Infrastruktur. <https://geekflare.com/de/best-open-source-monitoring-software/>

Was weiß ChatGPT über Nachhaltige Software-Entwicklung und Green Coding? Erste Tests und Bewertungen

Stefan Naumann¹, Achim Guldner¹, Sebastian Weber¹, Max Westing¹

Abstract: In den letzten Jahren sind Large Language Models (LLM) wie GPT, BERT, PaLM, BLOOM oder LLama durch die Verbesserung der Generierung natürlich wirkender Texte sowie neuer Fähigkeiten (Generierung von Quellcode etc.) stark ins öffentliche Interesse gerückt. Durch den Black Box-Charakter der Modelle und der Unklarheit der Qualität der zugrundeliegenden Trainingsdaten ist die Korrektheit der LLM-generierten Texte jedoch, insbesondere im wissenschaftlichen Umfeld, unklar und bedarf der Bewertung durch Expert:innen. In diesem Beitrag gehen wir daher auf die Frage ein, inwieweit LLM und insbesondere ChatBots wie ChatGPT Themen wie Green Coding und Nachhaltige Software-Entwicklung unterstützen können, und wie zuverlässig die Antworten am Beispiel von ChatGPT (GPT-3.5) sind. Wir stellen fest, dass die Inhalte für einen Überblick über die Themen sowie einen Einstieg in das Thema durchaus nutzbar sind, jedoch insbesondere im Hinblick auf weiterführende Quellen Schwachstellen aufweisen.

Keywords: Green Coding, Nachhaltige Software-Entwicklung, ChatBots, ChatGPT

1 Einführung

Kaum eine Neueinführung eines IT-Systems hat in den vergangenen Jahren für mehr Aufmerksamkeit gesorgt als die öffentliche Bereitstellung des Chatbots *ChatGPT* durch OpenAI². Zu dem Suchbegriff „ChatGPT“ lieferte Google Scholar im Juni 2023 bereits ca. 16.400 Ergebnisse³. Die hohe Qualität der Sprachausgabe, auch in deutscher Sprache, sowie die zumindest auf den ersten Blick schlüssigen Zusammenfassungen und die Fähigkeit zum Dialog sind Eigenschaften, die einen Innovationssprung im Bereich der hybriden künstlichen Intelligenz vermuten lassen. Wie steht es aber mit der Qualität der Ergebnisse aus wissenschaftlicher Perspektive, und welche Chancen, aber auch Grenzen, haben solche Sprachmodelle? Im Rahmen dieses Beitrags wollen wir anhand der Workshop-Themen *Nachhaltige Software-Entwicklung* [Na11] (*Green Software Engineering* als Synonym) und *Green Coding* der Frage nachgehen, wie ChatGPT die Forschung in diesem Bereich unterstützen kann, aber auch welche Fallstricke in der Nutzung lauern. Dabei haben wir uns auf die Version GPT-3.5 konzentriert, da diese frei zugänglich ist⁴.

¹ Hochschule Trier, Umwelt-Campus Birkenfeld, Institut für Softwaresysteme, Postfach 1380, 55761 Birkenfeld, Deutschland {s.naumann,a.guldner,s.weber,m.westing}@umwelt-campus.de

²<https://chat.openai.com> [2023-05-15]

³Suchanfrage „ChatGPT“, keine Zitate, keine weiteren Einschränkungen

⁴Es wird lediglich ein Open AI Account benötigt.

Der Beitrag ist wie folgt aufgebaut: Nach einer knappen Darstellung des Standes der Technik wird die Methodik der Untersuchung dargestellt. Im anschließenden Hauptteil werden eine Auswahl der Fragen, die zugehörige ChatGPT-Antwort und eine Bewertung wiedergegeben. Im Fazit diskutieren wir diese Bewertung und geben einen Ausblick, wie solche und vergleichbare Werkzeuge zukünftig für Wissenschaft und Praxis nutzbar sein können.

2 Stand der Technik

Ein LLM ist ein Sprachmodell, das typischerweise Deep Learning-Algorithmen mit mehreren Millionen Parametern verwendet, um Sprache zu verarbeiten. LLMs sind vielseitige Modelle, die eine breite Palette von Aufgaben erfolgreich bewältigen können. Sie besitzen umfangreiches Wissen und die Fähigkeit, Syntax und Bedeutung menschlicher Sprache zu erfassen. Dies hat zu einem Wandel in der Forschung im Bereich der natürlichen Sprachverarbeitung (Natural Language Processing - NLP) geführt, weg von spezialisierten Modellen für bestimmte Aufgaben hin zu solchen universelleren Modellen [Mi21]. Die Einführung dieser Technologie in der Öffentlichkeit erfolgte mit der Veröffentlichung von ChatGPT, einem zugänglichen Tool, das auf den zugrunde liegenden Sprachmodellen GPT-3.5 bzw. GPT-4.0 basiert. Neben OpenAI gibt es auch weitere Tools wie HuggingChat⁵ und BARD⁶, die ähnliche Funktionen anbieten. Neben den GPT-Modellen gibt es diverse weitere wie bspw. BERT [De19], LLaMA [To23] und PaLM⁷. Es wird bereits untersucht, wie solche Tools in die Forschung und Ausbildung von Wissenschaftler:innen integriert werden können [Ki23]. Dies geschieht z. B. ebenfalls im „Gespräch“ mit ChatGPT, wie Lund und Wang [LW23] gezeigt haben. Neben den allgemeinen Sprachmodellen wurden bereits, unter anderem von Google, mit dem Projekt Tailwind⁸, Anwendungen angekündigt, die es ermöglichen, Modelle auf einen individuellen Anwendungsfall hin zu spezialisieren.

Obwohl OpenAI einen Technical Report zu GPT-4 veröffentlicht hat [Op23], enthält er keine Informationen darüber, wie der Algorithmus funktioniert, welche Daten für das Training verwendet wurden oder auf welcher Hardware das System trainiert und betrieben wird. OpenAI weist jedoch selbst auf bestimmte Safety Challenges hin. Unter anderem, dass GPT-4, ähnlich wie sein Vorgänger GPT-3.5, eine Tendenz zum „Halluzinieren“, also der Generierung falscher Informationen, die jedoch als Fakten dargestellt werden, zeigt, was potenziell zu zahlreichen Problemen führen kann. Der Bericht thematisiert nicht die Umweltauswirkungen, die ein Sprachmodell wie GPT-4 mit sich bringt. Im Bereich der Green AI-Forschung [Sc20] haben beispielsweise Luccioni et al. [LVL22] gezeigt, dass LLMs einen enormen Energieverbrauch, und damit einen hohen CO₂ Ausstoß, haben können. In dieser Studie wurden 50,5 Tonnen CO₂-Äquivalente beim Training eines Sprachmodells (BLOOM) mit 176 Milliarden Parametern ermittelt.

⁵<https://huggingface.co/chat/> [2023-05-15]

⁶<https://bard.google.com/> [2023-05-15]

⁷<https://ai.google/static/documents/palm2techreport.pdf> [2023-05-15]

⁸<https://thoughtful.sandbox.google.com/> [2023-05-15]

Zur Evaluierung von LLMs kommen typischerweise verschiedene Methoden zum Einsatz. Weit verbreitet ist die Bestimmung der *Perplexity*. Dieser Wert kann einen guten Anhaltspunkt darstellen. Nach Russel und Norvig [RN20] sind jedoch Tests unter realen Bedingungen weitaus aussagekräftiger. Im huggingface OpenLLM Leaderboard werden deshalb Evaluationsverfahren wie die AI2 Reasoning Challenge (25-shot), HellaSwag (10-shot), MMLU (5-shot) und TruthfulQA (0-shot) herangezogen. Diese Methoden zielen darauf ab, logisches Denken und Allgemeinwissen zu überprüfen.⁹

Folgende Beschreibungen sollen für die spätere Bewertung der Antworten von ChatGPT dienen. *Green Software Engineering* beschäftigt sich mit allen Aspekten der Softwareentwicklung, die dazu beitragen, den direkten und indirekten Einfluss auf die Umwelt so gering wie möglich zu halten [Na11]. *Green Coding*, als Teil des *Green Software Engineering* wiederum, beschäftigt sich mit der Nachhaltigkeit der Software selbst und bezieht sich auf Methoden und Werkzeuge, die während der Entwicklung eingesetzt werden können, um die Nachhaltigkeit der Software in allen Phasen ihres Lebenszyklus (Entwicklung, Nutzung und Entsorgung) zu optimieren.

3 Vorgehensweise

Um einen Überblick über die Qualität der Antworten von ChatGPT im Bereich *Nachhaltige Software-Entwicklung* und *Green Coding* zu erhalten wurden ChatGPT einschlägige Fragen gestellt und die Antworten von den Autoren hinsichtlich Korrektheit, Vollständigkeit, Verwendbarkeit und Verständlichkeit bewertet. Um die verschiedenen Versionen des LLM zu vergleichen wurden Fragen sowohl auf Deutsch als auch auf Englisch formuliert.

Da der gesamte Dialog mit ChatGPT zu umfangreich für diese Veröffentlichung ist, werden hier zunächst die gestellten Fragen und anschließend ausgewählte Antworten dargestellt, auf die wir auch in Abschnitt 4 eingehen. Eine umfassende Sammlung der Fragen und Antworten stellen wir ergänzend zur Verfügung: <https://gitlab.rlp.net/green-software-engineering/chatgpt-ueber-green-coding>

3.1 Fragen an ChatGPT

Tab. 1 enthält die Fragen/Prompts, die von ChatGPT beantwortet wurden. Die Fragen dienen zum einen zur generellen Einschätzung der Wissensbasis zu Green Software Engineering und Green Coding, und zum anderen dazu, herauszufinden wie die Quellenlage ist, also woher das Wissen extrahiert bzw. zusammengestellt wurde.

⁹https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard [2023-06-15]

Deutsch	Englisch
Was ist Green Coding?	What is Green Coding?
Was ist Green Software Engineering?	What is Green Software Engineering?
Kennst Du Definitionen (von Green Software Engineering) aus wissenschaftlichen Papern?	Do you know definitions (of Green Software Engineering) from scientific papers?
Wie viel Energie benötigt Software?	How much energy does software need?
Wie kann man den Energiebedarf von Software messen?	How to measure the energy consumption of software?
Wie kann man Software ressourceneffizienter machen?	How to make software more resource efficient?
Ich brauche ein Referenzmodell für Nachhaltige Software.	I need a reference model for green software.
Was sind die wichtigsten wissenschaftlichen Quellen zu Green Software Engineering?	What are the most important scientific sources on Green Software Engineering?

Tab. 1: Input-Prompts an ChatGPT

PROMPT: Was ist Green Coding?

ANTWORT: Green Coding ist ein Konzept, das sich auf die Anwendung von nachhaltigen Praktiken und Methoden in der Softwareentwicklung konzentriert, um die Umweltauswirkungen von Software zu minimieren. Das Ziel von Green Coding ist es, eine nachhaltige Entwicklung von Software zu fördern und gleichzeitig eine positive Auswirkung auf die Umwelt zu erzielen.

Konkrete Praktiken des Green Coding können beispielsweise die Verwendung energieeffizienter Algorithmen und Programmiersprachen sein, die Reduzierung des Ressourcenverbrauchs von Anwendungen, die Optimierung von Code für eine effiziente Laufzeit oder die Verwendung von grünen Hosting-Lösungen. Darüber hinaus kann Green Coding auch die Verwendung von umweltfreundlicher Hardware für die Softwareentwicklung oder den Einsatz von Cloud-Lösungen mit erneuerbaren Energiequellen beinhalten.

Insgesamt zielt Green Coding darauf ab, die Umweltbelastung von Softwareentwicklung zu reduzieren, indem es den Fokus auf eine nachhaltige Entwicklung von Software legt. Durch die Anwendung von Green Coding-Praktiken kann der CO₂-Fußabdruck von Software reduziert werden, was wiederum zur Reduzierung des ökologischen Fußabdrucks von Organisationen beitragen kann.

Bewertung: Die Ergebnisse sind insgesamt richtig und als erste Einordnung schlüssig. Die Unterscheidung zwischen Prozess und Produkt ist etwas unscharf, was aber auch in der Literatur zu beobachten ist. Auch die Frage des Einsatzes von Erneuerbaren Energien ist fachlich sinnvoll, aber hängt weniger mit Green Coding zusammen. Aber auch hier ist die Literatur nicht einheitlich.

PROMPT: Kennst Du Definitionen (von Green Software Engineering) aus wissenschaftlichen Papern?

ANTWORT: Ja, hier sind einige Definitionen von Green Software Engineering aus wissenschaftlichen Papern:

Nach Mourad et al. (2017) bezieht sich Green Software Engineering auf den systematischen Ansatz zur Identifizierung, Analyse, Bewertung und Minimierung der Umweltauswirkungen von Softwareprodukten und -systemen.

Lago et al. (2013) definieren Green Software Engineering als den Prozess, der dazu beiträgt, die Nachhaltigkeit von Softwareprodukten und -systemen zu verbessern, indem die Umweltauswirkungen in jeder Phase des Software-Entwicklungsprozesses berücksichtigt werden.

Gemmel et al. (2014) definieren Green Software Engineering als den systematischen Einsatz von Software-Engineering-Prinzipien und -Techniken, um die Umweltauswirkungen von Softwareprodukten und -systemen zu reduzieren.

Stolterman et al. (2013) definieren Green Software Engineering als eine neue Disziplin innerhalb des Software-Engineering, die sich auf die Entwicklung von Softwareprodukten und -systemen konzentriert, die die Umweltauswirkungen minimieren und die Nachhaltigkeit fördern.

Diese Definitionen verdeutlichen, dass Green Software Engineering eine wichtige Rolle bei der Förderung einer nachhaltigen Entwicklung und der Reduzierung der Umweltauswirkungen von Softwareprodukten und -systemen spielt.

Bewertung: Die gegebenen Definitionen treffen den Kern von Green Software Engineering, auch wenn sie recht allgemein formuliert sind. Teile der Definitionen sind sehr nah an Formulierungen, die sich bei einer Google-Suche der Texte finden. So findet man den englischen Satz „*Green Software [Engineering] is an emerging discipline*“ in mehreren Blogbeiträgen¹⁰. Dieser findet sich übersetzt in der von ChatGPT gegebenen Definition von Stolterman et al. als „*definieren Green Software Engineering als eine neue Disziplin*“ wieder. Dies ist ein Beispiel, in dem die Definition von ChatGPT inhaltlich und sprachlich dicht an einem originalen Text liegt. Die angegebenen Quellen passen zunächst zu den gegebenen Definitionen. Die Namen der Hauptautor:innen decken sich mit den angegebenen Namen in den Definitionen, auch das Publikationsjahr ist konsistent. Generell wirken die Quellen authentisch, Titel wirken realistisch (vgl. „*A systematic mapping study of green software engineering. Journal of Cleaner Production*“ mit [Pe14]), und die angegebenen Journale und Konferenzen als solche existieren. Daher haben wir bzgl. der Quellen nachgefragt.

¹⁰z. B. <https://learn.greensoftware.foundation/introduction> [2023-06-15]

PROMPT: Nenne mir die genauen Quellen.

ANTWORT: Gerne! Hier sind die Quellen zu den genannten Definitionen:

Mourad, A., El Louadi, M., & Ounnar, F. (2017). A systematic mapping study of green software engineering. *Journal of Cleaner Production*, 142, 3206-3224.

Lago, P., Espana, S., Prikladnicki, R., & Audy, J. (2013). Green software engineering: From concept to practice. *IEEE Software*, 30(1), 30-38.

Gemmel, J., Rieger, A., & Lutz, R. (2014). Green software engineering – a survey. In *Proceedings of the 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 1302-1307.

Stolterman, E., Wiltse, H., & Valdez, A. (2013). Green software engineering: Defining the discipline and exploring challenges and opportunities. In *Proceedings of the 2013 International Conference on Software Engineering (ICSE)*, 1199-1208.

Bewertung: Die Bewertung der einzelnen Quellen gestaltet sich schwieriger, je detaillierter recherchiert wird. Sucht man lediglich an den angegebenen Stellen in den Journalen und Konferenzbänden, lässt sich schnell feststellen, dass die Quellen dort nicht existieren (Halluzination). Bei der Suche nach den angegebenen Autor:innen wurden in den meisten Fällen Personen gefunden, die auf die Angaben passen. Teilweise allerdings auch mehrere, die im Software Engineering tätig sind (teilweise auch im Bereich nachhaltige Software), teilweise haben sie auch miteinander publiziert oder wurden in Publikationen zitiert. An dieser Stelle wären ausgefeiltere Methoden und Werkzeuge notwendig, um mögliche Ursprünge der Komponenten der Quellenangaben zu suchen. Im Folgenden wird daher nur allgemein auf die Quellen eingegangen. Bei Quelle 1 und 2 wurden mögliche Autor:innen gefunden (teilweise sind mehrere möglich), die zumindest im Bereich Software Engineering, teilweise mit Bezug zu Green IT, arbeiten/publizieren. Zwei der Autor:innen haben gemeinsam publiziert. Bei Quelle 3 wurden die Namen „Gemmel, J.“ und „Rieger A.“ nicht gefunden. Eine Kandidatin für „R. Lutz“ publiziert im Bereich Software Engineering. Bei Quelle 4 wurden Kandidaten für „Stolterman E.“ und „Wiltse H.“, aber nicht für „Valdez A.“ gefunden. Stolterman und Wiltse haben gemeinsam publiziert.

PROMPT: Ich brauche ein Referenzmodell für Nachhaltige Software.

ANTWORT: Ein Referenzmodell für nachhaltige Softwareentwicklung könnte das sogenannte „Sustainable Software Engineering Model“ sein. Dieses Modell wurde entwickelt, um Entwicklern und Organisationen dabei zu helfen, nachhaltige Praktiken in den gesamten Softwareentwicklungslebenszyklus zu integrieren. Es besteht aus mehreren Dimensionen und Aktivitäten, die im Folgenden beschrieben werden:

Ressourcenmanagement: Dies beinhaltet die effiziente Nutzung von Ressourcen wie Energie, Hardware und Infrastruktur. Entwickler sollten sich bewusst sein, wie ihr Code die Ressourcennutzung beeinflusst und nach Möglichkeiten suchen, sie zu optimieren.

Lebenszyklusmanagement: Dies umfasst die Planung und Verwaltung des gesamten Softwarelebenszyklus, einschließlich Anforderungsdefinition, Entwicklung, Bereitstellung, Wartung und Entsorgung. Nachhaltige Softwareentwicklung zielt darauf ab, Verschwendung zu reduzieren und die Lebensdauer von Software zu maximieren.

Umweltverträglichkeit: Hierbei geht es um die Auswirkungen der Software auf die Umwelt. Entwickler sollten umweltfreundliche Praktiken wie die Reduzierung von Energieverbrauch, die Vermeidung von giftigen Substanzen oder die Minimierung von Emissionen in Betracht ziehen.

Soziale Verantwortung: Dieser Aspekt bezieht sich darauf, wie die Software die Gesellschaft und die Benutzer beeinflusst. Entwickler sollten ethische Prinzipien einhalten, Datenschutz sicherstellen und sicherstellen, dass die Software für alle Benutzer zugänglich ist.

Wiederverwendung und Modularität: Nachhaltige Softwareentwicklung fördert die Wiederverwendung von Code und die Schaffung modularer Softwarekomponenten. Dadurch können Entwickler Ressourcen sparen und die Entwicklungseffizienz verbessern.

Qualitätssicherung: Ein nachhaltiges Softwaremodell beinhaltet auch Qualitätsprüfungen, um sicherzustellen, dass die entwickelte Software den Anforderungen entspricht und robust, zuverlässig und sicher ist.

Es ist wichtig zu beachten, dass ein Referenzmodell allein nicht ausreicht, um nachhaltige Softwareentwicklung zu gewährleisten. Es erfordert das Engagement der Entwickler und Organisationen, nachhaltige Prinzipien in ihre Arbeitsweisen zu integrieren und kontinuierlich danach zu streben.

Bewertung: Die Modellbeschreibung ist inhaltlich schlüssig, allerdings wurde das „Sustainable Software Engineering Model“ nicht explizit als Referenzmodell veröffentlicht, sondern eher im Nebentext von wissenschaftlichen Beiträgen als Schlagwort eingeführt. Aufgrund des fehlenden Referenzmodells haben wir nachgefragt ob noch weitere Modelle existieren.

PROMPT: Gibt es noch andere Referenzmodelle?

ANTWORT: Ja, neben dem „Sustainable Software Engineering Model“ gibt es noch

weitere Referenzmodelle und Frameworks, die sich mit nachhaltiger Softwareentwicklung befassen. Hier sind einige davon:

Green Software Foundation's Green Software Capability Model: Die Green Software Foundation hat ein eigenes Referenzmodell entwickelt, das den Reifegrad von Organisationen in Bezug auf grüne Softwareentwicklung bewertet. Es definiert verschiedene Stufen der Reife und umfasst Bereiche wie Energieeffizienz, Lebenszyklusmanagement, Ressourcenoptimierung und soziale Verantwortung.

Sustainable Development Goals (SDGs): Die Vereinten Nationen haben die 17 SDGs definiert, um eine nachhaltige Entwicklung weltweit zu fördern. Softwareentwicklung kann einen Beitrag zu diesen Zielen leisten, indem sie beispielsweise Lösungen für erneuerbare Energien, Bildung oder Gesundheitsversorgung entwickelt. Das SDG-Framework kann als Referenzmodell dienen, um nachhaltige Softwareprojekte zu gestalten.

ISO 14001: Die ISO 14001 ist ein internationaler Standard für Umweltmanagementsysteme. Obwohl er nicht speziell auf Softwareentwicklung ausgerichtet ist, kann er als Referenz für Organisationen dienen, um umweltbezogene Aspekte in ihre Softwareentwicklungsprozesse zu integrieren und umweltfreundliche Praktiken zu fördern.

Diese Modelle und Frameworks bieten jeweils verschiedene Ansätze und Schwerpunkte für nachhaltige Softwareentwicklung. Je nach den spezifischen Zielen und Anforderungen deines Projekts kannst du dasjenige auswählen, das am besten zu deinen Bedürfnissen passt. Es ist ratsam, weitere Informationen zu jedem Modell zu suchen, um ein umfassendes Verständnis für seine Prinzipien und Anwendungen zu erhalten.

Bewertung: Die genannten Modelle und Frameworks sind soweit passend, wobei die Green Software Foundation dies nicht explizit in dieser Form veröffentlicht hat. Weitere konkrete Modelle zu Green Software wurden nicht gefunden. Da das GREENSOFT-Modell [Na11] erneut nicht genannt wurde haben wir explizit danach gefragt.

PROMPT: Was ist mit dem GreenSoft Modell?

ANTWORT: Es tut mir leid, aber ich konnte keine spezifischen Informationen über ein Referenzmodell namens „GreenSoft“ in Bezug auf nachhaltige Softwareentwicklung finden. Es besteht die Möglichkeit, dass es sich um eine spezifische Methode oder ein Modell handelt, das von einer bestimmten Organisation oder Firma entwickelt wurde und möglicherweise nicht weit verbreitet ist.

Es ist wichtig anzumerken, dass der Bereich der nachhaltigen Softwareentwicklung dynamisch ist und sich ständig weiterentwickelt. Es gibt verschiedene

Ansätze, Modelle und Frameworks, die von verschiedenen Organisationen, Forschern und Experten entwickelt werden. Es ist möglich, dass das von dir erwähnte „GreenSoft“ ein Modell ist, das von einer spezifischen Quelle oder Organisation entwickelt wurde.

Wenn du weitere Informationen über das „GreenSoft“-Modell suchst, empfehle ich dir, gezielt nach diesem Modell zu suchen, möglicherweise in Kombination mit spezifischen Schlagwörtern wie „nachhaltige Softwareentwicklung“ oder „grüne Software“. Auf diese Weise könntest du hoffentlich relevante Informationen zu diesem spezifischen Modell finden.

Bewertung: Dieses Ergebnis ist etwas erstaunlich, da das GREENSOFT-Modell ein häufiger zitiertes Referenzmodell für grüne und nachhaltige Software ist.

PROMPT: Was sind die wichtigsten wissenschaftlichen Quellen zu Green Software Engineering?

ANTWORT: Es gibt viele wissenschaftliche Quellen zu Green Software Engineering. Einige der wichtigsten sind:

„Green Software Engineering: Theories, Methods, and Tools“ von Xiaofeng Wang und Patricia Lago (2015): Dieses Buch bietet einen umfassenden Überblick über Green Software Engineering, einschließlich der Theorien, Methoden und Tools zur Messung und Verbesserung der Nachhaltigkeit von Software.

„Energy Efficiency in Large-Scale Distributed Systems: Techniques and Trends“ von Rajiv Ranjan, Samee U. Khan und Albert Y. Zomaya (2015): Dieses Buch beschäftigt sich mit der Energieeffizienz von großen verteilten Systemen und zeigt Techniken und Trends auf, die zur Verbesserung der Energieeffizienz beitragen können.

„Green IT: Technologies and Applications“ von Bhuvan Unhelkar (2011): Dieses Buch behandelt verschiedene Aspekte von Green IT, einschließlich Green Software Engineering. Es bietet einen Überblick über die Konzepte und Technologien, die zur Unterstützung der Nachhaltigkeit von IT-Anwendungen beitragen können.

„Energy-Aware Computing“ von Kaijie Wu und Zili Shao (2016): Dieses Buch beschäftigt sich mit energiebewusstem Computing und bietet einen Überblick über die Konzepte und Technologien, die zur Reduzierung des Energieverbrauchs von Software beitragen können.

„Sustainability in Software Engineering: A Systematic Literature Review“ von Birgit Penzenstadler, Steve Easterbrook, Colin Venters, und Colin C. Venters (2016): Diese systematische Literaturübersicht zeigt verschiedene Ansätze und

Methoden auf, die zur Verbesserung der Nachhaltigkeit von Software beitragen können.

Diese Quellen bieten wichtige Einblicke in Green Software Engineering und bieten eine Fülle von Informationen und Ressourcen zur Unterstützung von Forschung und Entwicklung auf diesem Gebiet.

Bewertung: Ähnlich zu den Quellen zu den Definitionen, fanden wir die Publikationen nicht wie angegeben. Die Angabe vollständiger Namen erleichterte hier die Recherche nach den Autor:innen. Die Suche lieferte aber auch hier nicht immer ein eindeutiges Ergebnis.

1. Die erste Quelle wurde nicht gefunden. Für beide Namen wurden mögliche Autor:innen gefunden. Patricia Lago publiziert im Bereich Nachhaltige Software, Xiaofeng Wang publiziert im Bereich Software Engineering.
2. Zu den Namen wurden mögliche Autor:innen gefunden. Alle sind im Bereich Distributed Computing tätig, teilweise mit Verbindungen zu Green IT. Es wurde eine Publikation mit ähnlichem Titel gefunden [Tr13].
3. Es existiert zwar ein Buch „*Green IT: Technologies and Applications*“, das 2011 publiziert wurde, allerdings nicht vom angegebenen Autor. Bhuvan Unhelkar hat ein Buch zu Green IT geschrieben: „*Green IT Strategies and Applications: Using Environmental Intelligence*“ [Un16].
4. Es wurden Kandidat:innen für beide Autor:innen gefunden. Beide publizieren im Software Engineering. Die angegebene Publikation wurde nicht gefunden.
5. Hier existiert eine Publikation von Birgit Penzenstadler mit dem angegebenen Titel, diese wurde aber 2012 veröffentlicht und nicht mit den angegebenen Co-Autoren [Pe12]. Interessant ist hier auch die doppelte Nennung von Colin C. Venters, der laut seinem IEEE-Profil auch ohne die Initialen seines zweiten Vornamens, als Colin Venters, publiziert hat¹¹. Hier scheint ChatGPT nicht korrekt erkannt zu haben, dass es sich um dieselbe Person handelt.

4 Diskussion und Ausblick

In den vorgestellten Tests sind die von der generativen KI ChatGPT erzeugten Antworten im Grundsatz schlüssig, und geben im textlichen Teil auch richtige Antworten bzw. zumindest Antworten, die in die richtige Richtung weisen, wobei die Antworten häufig nicht über *Wikipedia-Niveau* hinausgehen und teilweise oberflächlich bzw. unkonkret waren. Klare Schwächen zeigt ChatGPT 3.5 in der Quellensicherheit. Dieses *Halluzinieren* ist aus wissenschaftlicher und auch praktischer Sicht durchaus bedenklich, da hier Autorenschaften, Titel und Beitragsessenzen anscheinend ohne nachvollziehbare Regeln durchmischt werden.

¹¹<https://ieeexplore.ieee.org/author/37540614500> [2023-05-15]

Hinzu kommt die Frage des Energieverbrauchs von ChatGPT für Training und auch Inferenz. Hier liegen noch keine belastbaren Zahlen vor, aber Patterson et al. [Pa21] haben bereits gezeigt, dass für das Training von einem GPT3 Modell ca. 1,2 MWh benötigt wurden. Es wird auch aufgezeigt, dass anzunehmen ist, dass der Großteil der Energie bei der Inferenz verbraucht wird. Insgesamt lässt sich festhalten, dass wie häufig bei neuen Anwendungen, die auch Aspekte der Umwelt- und Nachhaltigkeitsinformatik umfassen, die ökobilanzielle Frage nur schwer beantwortbar ist. Dennoch kann ChatGPT dazu beitragen, den Diskurs um Green Coding und Green Software Engineering in Wissenschaft und Praxis zu befördern. Dies wird insbesondere interessant, wenn nachhaltiges Coden auch anhand von konkreten Code-Beispielen und -Überarbeitungen unterstützt wird, und wenn Akteur:innen Werkzeuge wie ChatGPT nutzen, um eigene Anwendungen und auch einfach nur die Nutzung von IKT ressourceneffizienter zu gestalten.

Über die beschriebenen Tests hinaus gibt es zahlreiche weitere Fragestellungen, deren Untersuchung lohnt: Vergleiche mit neueren Versionen von ChatGPT und auch anderen Modellen, tiefer gehende Dialoge sowie mehrfaches Re-Generieren der Antworten. Zudem ließe sich das Prompt Engineering (“In der Frage liegt die Antwort”) deutlich ausbauen. Die Rolle des Fragenden könnte variieren (Nutzer, Software-Entwickler, Studierender, Marketing-Experte etc.). Insbesondere aus Sicht der Software-Entwicklung sind Code-Optimierungen, Debugging hinsichtlich Energy Leaks und Konzepte für effiziente Architekturen Möglichkeiten zum aktiven Green Coding-Beitrag. Gleichzeitig besteht hier die Gefahr eines “Coding-Bias”, wenn bspw. immer eine bestimmte Programmiersprache oder Systemarchitektur empfohlen wird, die aber nicht für den angedachten Anwendungsfall geeignet ist. Zudem kann getestet werden, wie die aktuellste Variante von ChatGPT (GPT-4) diese Aufgaben bewältigt. In ersten Versuchen zeichnet sich dort ab, dass die “Halluzinationen” deutlich nachlassen und die vorgeschlagenen Quellen tatsächlich auffindbar sind.

Im Ausblick wäre zudem ein spannender Ansatz, einen “ChatGPT-Green-Coding-Experten” aufzubauen, der neben seinem Informatik-Basiswissen auch mit einschlägigen Code-, Architektur- und Paper-Beispielen trainiert ist. Hierzu ließen sich bspw. LLaMA oder BLOOM¹² nutzen, um ein eigenes Sprach-Wissens-Modell zu generieren und so das Wissen greifbarer und interaktiver im Sinne hybrider Intelligenz zu machen.

Danksagungen

Diese Arbeit wurde durch das Ministerium für Wissenschaft und Gesundheit des Landes Rheinland-Pfalz (MWG) im Rahmen des Forschungskollegs „AI-CPPS“ sowie durch das Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV) im Rahmen des Forschungsprojekts „KIRA“ unter dem Förderkennzeichen 67KI32013B unterstützt.

¹²<https://huggingface.co/bigscience/bloom> [2023-05-15]

Literatur

- [De19] Devlin, J. et al.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of NAACL-HLT 2019. ACL, S. 4171–4186, 2019.
- [Ki23] Kiliç, S.: Embracing the Future of Distance Science Education: Opportunities and Challenges of ChatGPT Integration. *Asian Journal of Distance Education* 18/1, 2023.
- [LVL22] Luccioni, A. S.; Viguier, S.; Ligozat, A.-L.: Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model. *arXiv:2211.02001/*, 2022.
- [LW23] Lund, B. D.; Wang, T.: Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library Hi Tech News/*, 2023.
- [Mi21] Min, B. et al.: Recent advances in natural language processing via large pre-trained language models: A survey. *arXiv:2111.01243/*, 2021.
- [Na11] Naumann, S. et al.: The greensoft model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 1/4, S. 294–304, 2011.
- [Op23] OpenAI: GPT-4 Technical Report, 2023, *arXiv: 2303.08774 [cs.CL]*.
- [Pa21] Patterson, D. et al.: Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350/*, 2021.
- [Pe12] Penzenstadler, B. et al.: Sustainability in Software Engineering: A Systematic Literature Review. In: 16th International Conference on Evaluation & Assessment in Software Engineering (EASE 2012). Bd. 2012, Jan. 2012.
- [Pe14] Penzenstadler, B. et al.: Systematic Mapping Study on Software Engineering for Sustainability (SE4S). In: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. EASE '14, ACM, London, England, United Kingdom, 2014.
- [RN20] Russell, S.; Norvig, P.: *Artificial Intelligence: A Modern Approach* (4th Edition). Pearson, 2020, ISBN: 9780134610993.
- [Sc20] Schwartz, R. et al.: Green ai. *Communications of the ACM* 63/12, S. 54–63, 2020.
- [To23] Touvron, H. et al.: LLaMA: Open and Efficient Foundation Language Models, 2023, *arXiv: 2302.13971 [cs.CL]*.
- [Tr13] Trobec, R. et al.: Energy efficiency in large-scale distributed computing systems. In: Proceedings of the 36th International convention on information and communication technology, electronics and microelectronics. S. 253–257, Jan. 2013, ISBN: 978-953-233-076-2.
- [Un16] Unhelkar, B.: *Green IT strategies and applications: using environmental intelligence*. CRC Press, 2016.

Potentials of Green Coding - Findings and Recommendations for Industry, Education and Science

Dennis Junger ¹, Max Westing ², Christopher Freitag ¹, Achim Guldner ², Konstantin Mittelbach ¹, Sebastian Weber ², Stefan Naumann ², Volker Wohlgemuth¹

Abstract: Progressing digitalization and increasing demand and use of software cause rises in energy- and resource consumption from information and communication technologies (ICT). This raises the issue of sustainability in ICT, which increasingly includes the sustainability of the software products themselves and the art of creating sustainable software. To this end, we conducted an analysis to gather and present existing literature on three research questions relating to the production of ecologically sustainable software ('Green Coding') and to provide orientation for stakeholders approaching the subject. We compile the approaches to Green Coding and Green Software Engineering (GSE) that have been published since 2010. Furthermore, we considered ways to integrate the findings into existing industrial processes and higher education curricula to influence future development in an environmentally friendly way.

Keywords: Green Coding, Potentials of Green Coding, Curricula, Software Engineering, Literature Analysis, Sustainable, Engineering, Curricula, Education, Higher Education

1 Introduction and Motivation

Sustainability, as relating to ICT, in general, and software development and engineering in particular, is a research field of growing interest [An22]. While sustainability considerations in ICT have focused on hardware for the longest time, software is the driver of hardware energy consumption [Gu17]. Additionally, it affects the hardware life-cycle, e. g. as ever more demanding software systems quickly outgrow older hardware [Wi95]. Access to models, methods, and tools for GSE is elementary to enable sustainable software production. Furthermore, it is essential to make future IT professionals and researchers aware of ICT's sustainability aspects and equip them with the skills to consider them in their work adequately. Thus, in this literature analysis, we'd like to present available Green Coding practices and information on how educators can incorporate them into their teaching.

Sustainability is a complex topic that is commonly subdivided into dimensions or pillars. Within Green Coding, we focus on the environmental aspect of sustainability, dispensing with the other pillars, like ecological, social, and economic sustainability. Of course, these

¹ HTW Berlin, University of Applied Sciences, Industrial Environmental Informatics Unit, Wilhelmshofstr. 75A, 12459 Berlin, Germany {dennis.junger|christopher.freitag|volker.wohlgemuth}@htw-berlin.de, konstantin.mittelbach@student.htw-berlin.de

² Institute for Software Systems, Environmental Campus Birkenfeld, Trier University of Applied Sciences, 55765 Birkenfeld, Germany {m.westing|a.guldner|seb.weber|s.naumann}@umwelt-campus.de

are also important and are introduced for the field of software engineering e. g. in Calero et al. [CMP21].

Research on sustainability in ICT is done under several terms, varied in definition and scope [CP17], e. g. Green IT/ICT, Green by IT/ICT, Sustainable Software, etc. The terms 'Green IT' and 'Green By IT' separate the field into two areas. In this regard, we designate ICT that aims to have a positive impact on sustainability in other areas (e. g. smart grids, building automation, environmental management information systems (EMIS) like life-cycle assessment software (LCA), etc.) as 'Green By IT', whereas 'Green IT' deals with making ICT itself more sustainable (e. g. follow-the-sun load shifting, backward compatible software, designing energy- and resource-efficient hardware, etc.). We consider 'Green Coding' as part of 'Green IT', concerned with the sustainability of the software itself, focusing on methods and tools that can be leveraged during development to impact software sustainability in all of its life-cycle phases (development, usage, and disposal phase).

2 Research Questions and Method

The search and analysis is divided into three research questions (RQ). For RQ1 *"What concepts of programming software in an environmentally friendly way exist in general for software development?"*, Green Coding practices and tools outlined in the academic literature are gathered and presented. This should provide an overview to IT professionals approaching the topic of sustainable software development, aide them to navigate the field, and help them find Green Coding practices and tools applicable to their work. RQ2 *"What concepts of environmentally friendly software engineering processes are already in place in the industry?"* is intended to yield information on Green Coding practices used by software developers. We searched for case studies or white papers on implementing Green Coding practices, and their impact on software performance and sustainability metrics. Methods and tools from industry and community sources are also presented here. RQ3 *"How can Green Coding concepts be implemented into the current software development processes and the curricula of existing study programs?"* explores how Green Coding practices can be communicated to future IT professionals, to establish competence in using available tools and methods. To reduce the complexity of RQ3 and more precisely match the target audiences, RQ3 is divided into the following two sub-questions. RQ3-1 focuses on the software development process, pointing out established software development methods, models, processes, and tools. RQ3-2 focuses on curricula of existing study programs to generate knowledge about teaching approaches for future IT professionals. As most software developers have an academic background, this analysis focuses on implementing aspects of Green IT into curricula of higher education institutions, since over 75 % of software developers attended a college [St22].

The analysis bases on an exploratory literature search using known contributions as well as querying Google Scholar, Digital Bibliography & Library Project (dblp), and the university libraries of the Trier University of Applied Sciences and the Industrial Environmental

Informatics Unit of the HTW Berlin – University of Applied Sciences for the terms 'sustainable software development processes', 'sustainability as non-functional requirement', and 'development of curriculum'. We limited the results to literature published since 2010 and excluded papers referring to sustainability in a different context than defined above (i. e. not focusing on the environmental dimension of sustainability), as well as papers from non-computer science fields (e. g. results about green (building) code). This was decided based on the title, research area, and abstract. We had to exclude some literature that would likely have been applicable, due to access restrictions.

We selected the search strings for the RQs with the intention of acquiring literature summary sources (i. e. previous literature analysis or introductory texts), concrete measures and guidelines, and literature on implementations of green coding practices in software development and education. Keywords to identify case studies, models, best practices, guidelines, processes, and curricula were added to identify the state of the art in Green Coding. The search strings are available in the supplementary material repository³.

3 Processed Sources and Findings

3.1 RQ1 - What concepts of programming software in an environmentally friendly way exist for software development in general?

Energy consumption at the hardware level is well-understood. However, the impact of software on energy consumption has only been regarded in the past decade and has gained more attention since about 2015. There are models for estimating or measuring the energy consumption of software, and tools based on these models that seem suitable to inform software developers about energy optimization potentials in their code such as [As21; Hi14; Hö13; Ka09; Kh22; Le15]. These exist for a range of technologies. Henderson et al. developed a framework to make the climate impacts of machine-learning research more visible [He20]. Furthermore, hardware-based power measurement tools like RAPL⁴ or nvidia-smi⁵ exist, that can be used to quickly assess the energy consumption of individual hardware components.

A common theme in many studies is that there rarely is an optimal solution concerning energy- and resource consumption [Pe17]. Some studies like [Ja16] compare hardware-based measurements with simultaneously obtained estimations and find that there are sometimes large (over 60 % in this case) differences in the measurements vs. the estimations due to unexplained energy overheads. This demonstrates the complexity of software sustainability and highlights why IT professionals need to be able to assess which measures are suited for

³ <https://doi.org/10.5281/zenodo.7920569>

⁴ <https://01.org/blogs/2014/running-average-power-limit-%E2%80%93-rapl> and <https://cdrdv2.intel.com/v1/dl/getContent/671200> [2023-04-28]

⁵ <https://developer.nvidia.com/nvidia-system-management-interface> [2023-04-28]

a project. The steps required to improve upon a given piece of source code's induced energy consumption are also not clear. Some work provides concrete hints [Hö14] or details where energy hot spots occur in the source code [No15; Ve18; Wa12].

Considering the underlying technologies, Kreten [Kr22] describes methods for energy-efficient container technologies. There is research on software sustainability along the product life-cycle: when designing the software architecture (e. g. [Ch19]), during development (e. g. [Hö14; Wa12]), and maintenance (e. g. [Bu17; Ma21; Sa16]). Most of these approaches rely on energy measurements and a broad conceptual basis, such as [Hi14; Ke13; Na11], who define models, process artifacts and evaluation methods for green software engineering processes, not including specific Green Coding methods such as coding guidelines or using specific architectures and technologies. Several studies focus explicitly on applications for mobile devices (e. g. [Ba22; Hi14; Sa16]), due to the impact of energy consumption on battery life, with Anwar et al. providing an overview of available tools for the development of green android apps [An21]. Advances have been made in the development and application of standards for sustainable software products and eco-labels [Ke18; Na21]. However, no sustainability standards applying explicitly to software development have been identified by standards organizations.

3.2 RQ2 - What concepts of environmentally friendly software engineering processes are already in place in the industry?

Few case studies on software sustainability practices in the industry are available. Calero et al. [Ga21] performed an in-depth analysis of software sustainability on a set of personal health record software products. They propose recommendations and measured their impact on software energy consumption. Haldal et al. [He23] identify a need for sustainability competency in IT professionals in the industry. Lammert et al. [La22] found that software engineers are usually not equipped with the training and tools to fulfill this need. In addition to scientific papers, industry and developer communities have also produced methods and tools for sustainable software engineering, some of which are not (yet) published. These often build on the software measurement approaches mentioned in RQ1.

In addition to the blue angel for software products and the criteria developed in [Ke18] and the proposed measurement setups using external power meters like [Ju22], internal loggers (like the aforementioned RAPL and nvidia-smi or eBPF⁶) in combination with resource loggers can be useful to assess the impacts of a software product under development. Here, it is necessary to access the data provided by the hardware drivers. Exemplary tools are the Green Metrics Tool⁷, Code Carbon⁸ Hubblo RAPL exporter⁹ and Scaphandre¹⁰.

⁶ <https://ebpf.io/> [2023-04-28]

⁷ <https://github.com/green-coding-berlin/> [2023-05-09]

⁸ <https://github.com/mlco2/codecarbon> [2023-05-10]

⁹ <https://github.com/hubblo-org/windows-rapl-driver> [2023-05-09]

¹⁰ <https://github.com/hubblo-org/scaphandre> [2023-05-09]

Many available guidelines and best-practice recommendations are in the form of blog posts or lists from developer communities, consulting firms and businesses¹¹. Regardless of the correctness of the information contained in this content, they often do not provide sources for their recommendations.

3.3 RQ3-1 How can Green Coding concepts be implemented into the current software development processes?

Heldal et. al. [He23] identify that several factors motivate organizations to act sustainably, primarily beneficial effects on their business. We identified the following approaches to implementing the findings into the software development process: In addition to the cost savings from more efficient hardware and energy use, the reduction of direct and indirect environmental impacts must be incentivized, e. g. through environmental guidelines and environmental regulations. Ultimately, environmental seals, labels, and awards can provide a competitive advantage. For example, an eco-label is a clear advantage for public procurement in Germany, as it can be part of the specification for a product [Sc22]. However, even without regulations, it can be assumed that a green impact will benefit from transparency.

Additionally, we considered where the knowledge could be implemented in companies, leading to the question of how the knowledge is transferred to the corresponding destinations. In this context, training and education are fundamental anchor points for transfer and implementation [Sa21]. The conceptual and technical basics of Green Coding or Green IT, in general, should be conveyed. Training courses and workshops must bring this knowledge into the companies. These training courses can be given or initiated externally by consultants and governmental actors. However, they can also be quickly and widely disseminated, e. g. through e-learning platforms.

For integrating the fundamentals of Green Coding into the software development process, reference models, like the Greensoft Model [Di13], describe ways to implement them into software development processes like Scrum [Na11]. Investing in (the development of) tools and frameworks for implementing more energy-saving code or adapting the software development process is indispensable to implement the path most effectively. Only by making the consumption transparent for developers, improvements are possible.

3.4 RQ3-2 How can Green Coding concepts be implemented into the curricula of existing study programs?

To answer this question and develop recommendations, our first step was to research existing courses and programs of universities and institutes of higher education. However, the

¹¹ e. g. <https://www.ibm.com/cloud/blog/green-coding>, <https://www.suso.academy/en/2023/03/13/green-coding-the-5-most-important-basics-for-sustainable-software-development-with-code-examples/>, or <https://geekflare.com/green-coding/> [2023-06-27]

search was more challenging than it initially seemed. Most universities have printed the specializations of the study subjects in module handbooks or make the course information with their disciplines available only after an account has been created in an application management system. This influenced our search, which we mainly conducted via online media limited to English publications and announcements. In addition to sources from cooperating universities, we consulted job platforms such as LinkedIn¹² and institutions known in the Green Coding community. Currently, courses designed for GSE or Green Coding are mainly available in Europe. Table 1 lists the universities and universities of applied sciences in Europe and the US, which we found that teach related courses.

School	Country	School	Country
University of Bristol	England	LUT University	Finland
Hasso Plattner Institut (HPI)	Germany	HTW Berlin	Germany
Umwelt-Campus Birkenfeld (UCB)	Germany	Amsterdam School of Data Science	Netherlands
DELFT University of Technology	Netherlands	Kharkiv Aviation Institute*	Ukraine
Royal Institute of Technology (KTH)	Sweden	Mälardalen University	Sweden
University of Gothenburg	Sweden	Carnegie Mellon University	USA
The University of California Berkeley	USA		

*current status unknown

Tab. 1: Universities teaching Green Coding and GSE

To be able to give recommendations for the implementation of Green Coding and GSE, in addition to the search terms listed in the supplementary material, search terms for the creation of curricula are added to the queries.

For future curricula, Drake and Reid [DR18] propose to, on the one hand, teach the big picture of a topic and, on the other hand, the main competencies. To convey this knowledge, technical, practical, and participatory interests should be addressed [FB06]. Mishra and Mishra [MM21] describe how a sustainable software engineering curriculum based on ACM/IEEE could be created. Since it is challenging to design entire curricula such as PERCCOM [Ko19], the first step should be to develop individual courses, workshops, discussion groups [MM11] or university initiatives like in the HAW Hamburg [Ei23]. This can be followed by a module like 'Software Engineering Sustainability', introduced at the Kharkiv Aviation Institute [TV19]. To identify the right target group for these modules, it should be attended by various similar courses because interdisciplinary knowledge is more important than ever in current times [Ra18]. To implement the new knowledge, the Industrial Environmental Informatics Unit at HTW Berlin developed and evaluated a prototypical course that includes the listed points to identify best practices and obstacles, conducted in the winter semester of 2022/23. As a part and also a by-product of the course and the measurement setup [Ju22] used for this purpose, the acquired knowledge was already passed on to students and colleagues. Thus, publications on the latest statistical analysis applications [Se23] and also concerning current AI environments in education [Bü23]. The contents and modalities are currently being published in [Ju23].

¹² <https://www.linkedin.com/> [2023-05-09]

4 Conclusion and Outlook

In this literature analysis, we collected results to classify the state of the art of Green Coding in software engineering and education. The goals were to enable stakeholders to realize the potential of Green Coding, whether well-researched or untapped, to provide them with a solid introduction to the field and possible approaches for their activities. A decade ago, the area of *Sustainable Software Engineering* was still in its infancy, as evidenced by a lack of case studies and evaluations [Pe14]. Today, we conclude that research in the field is steadily increasing and has made significant progress. Several measurement and evaluation systems for assessing software's energy- and resource consumption and -efficiency were developed. However, the low adoption rate of GSE methods still stands, as we could identify only a few case studies. Implementing aspects of sustainability, such as Green Coding techniques, into the training of future IT professionals is a crucial step to bringing sustainability into view for software producers while providing them with a workforce capable of implementing sustainability into their development process. Future work could be done by extracting practices contained in non-academic sources and finding academic papers supporting their impact on sustainability or attempting to prove some of the recommendations experimentally by applying them to a software product and tracking the impact of each recommendation. Surveys and expert interviews should be conducted to get a better overview of green coding in institutes of higher education and industry, as curricula may not be up-to-date and corporations do not always publish the results of their projects.

A logical consequence of this project is that this knowledge is listed and communicated in a way that is as easy to understand as possible. Besides already existing courses, like the course from the green software foundation¹³, it is essential that these concepts are scientifically validated and distributed. Since the expertise in imparting knowledge lies with the institutes of higher education, the design for the course concept to impart this knowledge and practice needs to be expanded. Furthermore, we are currently extending the GSE approach to Artificial Intelligence and also planning to focus on a case study about the conception of a prototype measurement approach to fully virtualized AI laboratories.

Acknowledgements

This work was funded by the Internet Society Foundation¹⁴ project "Potentials of Green Coding". We would also like to thank our partner, the German Informatics Society (Gesellschaft für Informatik e. V.), especially Elisabeth Schauer mann and Carolin Henze¹⁵, the assisting master students, and student assistants.

¹³ <https://learn.greensoftware.foundation/> [2023-05-09]

¹⁴ <https://www.isocfoundation.org/> [2023-05-09]

¹⁵ <https://gi.de/en/aktuelles/projekte/en-green-coding> [2023-05-09]

References

- [An21] Anwar, H. et al.: Tool Support for Green Android Development. In: *Software Sustainability*. Springer, pp. 153–182, 2021.
- [An22] Andrikopoulos, V. et al.: Sustainability in Software Architecture: A Systematic Mapping Study, 2022, URL: <https://arxiv.org/abs/2204.11657>.
- [As21] Aslanpour, M. S. et al.: WattEdge: A Holistic Approach for Empirical Energy Measurements in Edge Computing. In: *Service-Oriented Computing*. Springer, Cham, pp. 531–547, 2021.
- [Ba22] Bangash, A. A. et al.: A Black Box Technique to Reduce Energy Consumption of Android Apps. In: *44th International Conference on Software Engineering: New Ideas and Emerging Results*. ACM, Pittsburgh, Pennsylvania, pp. 1–5, 2022.
- [Bu17] Bunse, C.: On the Impact of Code Obfuscation to Software Energy Consumption. In: *From Science to Society, New Trends in Environmental Informatics - EnviroInfo 2017*, Luxembourg, 13-15September 2017. Springer, pp. 239–249, 2017, URL: https://doi.org/10.1007/978-3-319-65687-8%5C_21.
- [Bü23] Bültemann, M. et al.: Energy Consumption of AI in Education: A Case Study. In: *21. Fachtagung Bildungstechnologien (DELFI)*. Gesellschaft für Informatik e.V., Bonn, 2023.
- [Ch19] Chowdhury, S. A. et al.: GreenBundle: An Empirical Study on the Energy Impact of Bundled Processing. In: *41st International Conference on Software Engineering*. 2019.
- [CMP21] Calero, C.; Moraga, M. Á.; Piattini, M.: Introduction to software sustainability. *Software Sustainability*, pp. 1–15, 2021.
- [CP17] Calero, C.; Piattini, M.: Puzzling out Software Sustainability. *Sustainable Computing: Informatics and Systems* 16/, pp. 117–124, 2017, ISSN: 2210-5379, URL: <https://www.sciencedirect.com/science/article/pii/S2210537916301676>.
- [Di13] Dick, M. et al.: Green software engineering with agile methods. In: *2013 2nd International Workshop on Green and Sustainable Software (GREENS)*. Pp. 78–85, 2013.
- [DR18] Drake, S.; Reid, J.: Integrated Curriculum as an Effective Way to Teach 21st Century Capabilities. *Asia Pacific Journal of Educational Research* 1/, pp. 31–50, Jan. 2018.
- [Ei23] Eickstädt, E. et al.: Computer Science for Future - Sustainability and Climate Protection in the Computer Science Courses of the HAW Hamburg, 2023.
- [FB06] Fraser, S. P.; Bosanquet, A. M.: The curriculum? That’s just a unit outline, isn’t it? *Studies in Higher Education* 31/3, pp. 269–284, 2006.




- [Ga21] García-Berná, J. A. et al.: Energy efficiency in software: A case study on sustainability in personal health records. *Journal of Cleaner Production* 282/, p. 124262, 2021.
- [Gu17] Guldner, A. et al.: Energy Consumption and Hardware Utilization of Standard Software: Methods and Measurements for SoftwareSustainability. In: *Progress in IS*. Springer, Aug. 2017.
- [He20] Henderson, P. et al.: Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning. *Journal of Machine Learning Research*/, Arxiv: 2002.05651, Jan. 2020.
- [He23] Heldal, R. et al.: *Sustainability Competencies and Skills in Software Engineering: An Industry Perspective*, 2023.
- [Hi14] Hindle, A. et al.: GreenMiner: A Hardware Based Mining Software Repositories Software Energy Consumption Framework. In: *Proceedings of the 11th Working Conference on Mining Software Repositories*. ACM, New York, USA, pp. 12–21, 2014.
- [Hö13] Höning, T. et al.: Proactive Energy-Aware System Software Design with SEEP. *Softwaretechnik-Trends* 33/2, pp. 6–7, 2013.
- [Hö14] Höning, T. et al.: Proactive Energy-Aware Programming with PEEK. In: *2014 International Conference on Timely Results in Operating Systems*. USENIX, 2014.
- [Ja16] Jagroep, E. A. et al.: Software Energy Profiling: Comparing Releases of a Software Product. In: *38th International Conference on Software Engineering Companion*. 2016.
- [Ju22] Junger, D. et al.: Conception and test of a measuring station for the analysis of the resource and energy consumption of material flow-oriented environmental management information systems (EMIS). In: *EnviroInfo 2022*. Gesellschaft für Informatik e.V., 2022.
- [Ju23] Junger, D. et al.: Design and implementation of a lecture for teaching current Green Coding approaches and practices at the HTW Berlin. In: *Informatik 2023*. accepted for publication, Gesellschaft für Informatik e.V., Bonn, 2023.
- [Ka09] Kansal, A. et al.: *JouleMeter: Virtual Machine Power Measurement and Management*, 2009, URL: <https://www.microsoft.com/en-us/research/project/joulemeter-computational-energy-measurement-and-optimization/>.
- [Ke13] Kern, E. et al.: Green software and green software engineering—definitions, measurements, and quality aspects. In: *First International Conference on Information and Communication Technologies for Sustainability (ICT4S2013)*, 2013b ETH Zurich. Pp. 87–91, 2013.
- [Ke18] Kern, E. et al.: Sustainable software products—Towards assessment criteria for resource and energy efficiency. *Future Generation Computer Systems*/, pp. 199–210, 2018.

- [Kh22] Khandelwal, S. et al.: Real Time Carbon Emissions Calculator for Personal Computers. In: *Data Management, Analytics and Innovation*. Springer, Singapore, pp. 651–663, 2022.
- [Ko19] Kor, A.-L. et al.: Education in Green ICT and Control of Smart Systems : A First Hand Experience from the International PERCCOM Masters Programme. *IFAC-PapersOnLine* 52/9, 12th IFAC Symposium on Advances in Control Education ACE 2019, pp. 1–8, 2019, ISSN: 2405-8963, URL: <https://www.sciencedirect.com/science/article/pii/S2405896319304495>.
- [Kr22] Kreten, S.: *Modellbildung und Umsetzung von Methoden zur energieeffizienten Nutzung von Containertechnologien*, PhD thesis, Trier University, 2022.
- [La22] Lammert, D. et al.: Software Engineers in Transition: Self-Role Attribution and Awareness for Sustainability. In: *Proceedings of the 55. Hawaii International Conference on System Sciences*. Jan. 2022.
- [Le15] LeBeane, M. et al.: Watt Watcher: Fine-Grained Power Estimation for Emerging Workloads. In: *International Symposium on Computer Architecture and High Performance Computing*. IEEE, 2015.
- [Ma21] Mancebo, J. et al.: Does maintainability relate to the energy consumption of software? A case study. *Software Quality Journal* 29/1, pp. 101–127, 2021.
- [MM11] Malik, A. S.; Malik, R. H.: Twelve tips for developing an integrated curriculum. *Medical teacher* 33/2, pp. 99–104, 2011.
- [MM21] Mishra, A.; Mishra, D.: *Sustainable Software Engineering: Curriculum Development Based on ACM/IEEE Guidelines*. In: *Software Sustainability*. Springer, Cham, 2021.
- [Na11] Naumann, S. et al.: The greensoft model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 1/4, pp. 294–304, 2011.
- [Na21] Naumann, S. et al.: The eco-label blue angel for software - Development and components. In: *Advances and New Trends in Environmental Informatics: Digital Twins for Sustainability*. Springer, pp. 79–89, 2021.
- [No15] Noureddine, A. et al.: Monitoring energy hotspots in software. *Automated Software Engineering* 22/3, pp. 291–332, Oct. 2015.
- [Pe14] Penzenstadler, B. et al.: Systematic mapping Study on Software Engineering for Sustainability (SE4S). *ACM International Conference Proceeding Series*/, May 2014.
- [Pe17] Pereira, R. et al.: Energy Efficiency across Programming Languages: How Do Energy, Time, and Memory Relate? In: *SIGPLAN International Conference on Software Language Engineering*. ACM, Vancouver, Canada, pp. 256–267, 2017.

-
- [Ra18] Ramirez-Mendoza et al.: Engineering Education 4.0: proposal for a new Curricula. In: IEEE Global Engineering Education Conference (EDUCON). Pp. 1273–1282, 2018.
- [Sa16] Sahin, C. et al.: How does code obfuscation impact energy usage? *Journal of Software: Evolution and Process*, Jan. 2016.
- [Sa21] Saraiva, J. et al.: Bringing Green Software to Computer Science Curriculum: Perspectives from Researchers and Educators. In: ITICSE. ACM, 2021.
- [Sc22] Schneider, T.: Rechtsgutachten umweltfreundliche öffentliche Beschaffung - Aktualisierung 2022, Accessed: 2023-04-28, 2022.
- [Se23] Seegert, T. et al.: Energy and resource comparison of current applications with a focus on statistical analyses and evaluations using the example of Matlab and R/RStudio. In: *EnviroInfo 2023*. accepted for publication, Gesellschaft für Informatik e.V., Bonn, 2023.
- [St22] Statista: Worldwide levels of formal education for software developers, <https://www.statista.com/statistics/793568> [2/24/2023], 2022.
- [TV19] Turkin, I.; Vykhodets, Y.: Software Engineering Sustainability Education in Compliance with Industrial Standards and Green IT Concept. In: *Green IT Engineering: Social, Business and Industrial Applications*. Springer, pp. 579–604, 2019.
- [Ve18] Verdecchia, R. et al.: Code-Level Energy Hotspot Localization via Naive Spectrum Based Testing. In: *Advances in Environmental Informatics: Managing Disruption, Big Data and Open Science*. Springer, 2018.
- [Wa12] Wang, S. et al.: Safari: Function-level power analysis using automatic instrumentation. In: *2012 International Conference on Energy Aware Computing*. IEEE, pp. 1–6, 2012.
- [Wi95] Wirth, N.: A Plea for Lean Software. *Computer* 28/02, pp. 64–68, 1995.

**Ökologische Nachhaltigkeit -
Nachhaltige Wertschöpfungssysteme
(NaWerSys) III**

3. Workshop: Nachhaltige Wertschöpfungssysteme


Thorsten Schoormann ¹, Friedemann Kammler ², Paul Christoph Gembariski ³
und Simon Hagen⁴


1 Einleitung


Vor dem Hintergrund aktueller, gesellschaftlicher Herausforderungen [Th20] ist die Entwicklung (digitaler) Produkte und Dienstleistungen gegenwärtig von einem Umdenken hin zu mehr Nachhaltigkeit und Verantwortung geprägt [Di20, Lo21, Sc23]. Dies betrifft den gesamten Lebenszyklus solcher Angebote von der Ideenentwicklung und Gestaltung bis zum Betrieb, zur Nachnutzung und zum Recycling. Informations- und Kommunikationstechnologie nimmt dabei nicht nur die traditionelle Rolle als Intermediär zwischen den anbietenden Unternehmen und dem Kundensegment ein. Sie ist vor allem auch ein Enabler vernetzter, leistungsübergreifender Funktionen [Ha19] bis hin zur Autonomisierung ganzer Prozessabläufe, in denen etwa Umweltveränderungen sensorisch erfasst und Handlungsmöglichkeiten abgeleitet werden.

Für die Implementierung solcher Funktionen werden oftmals datengetriebene Angebote mehrerer Organisationen in Wertschöpfungssystemen (WSS) gebündelt, um Fähigkeiten und Ressourcen übergreifend zu nutzen. WSS können Leistungsangebote ad-hoc auf konkrete und sich verändernde Situationen zuschneiden, um nachhaltige Lösungen zu erbringen [GK21]. Die holistische Perspektive der WSS eröffnet dabei neue Wege, sich von der einzelnen Anwendung zu lösen und unternehmensübergreifende Potenziale zur Verbesserung, auch unter Berücksichtigung von Nachhaltigkeitszielen, umzusetzen. Die neuen Möglichkeiten sowie die damit einhergehenden Herausforderungen für WSS können in die drei Dimensionen für ökonomische, ökologische und soziale Nachhaltigkeit [E197] über den gesamten Lebenszyklus hinweg gegliedert werden.

Vor dem Hintergrund *ökonomischer* Aspekte ergeben sich durch die zentrale Rolle von Daten teils disruptive Veränderungen für den Markt und den Wettbewerb. Die durch den technologischen Fortschritt immer kürzer werdenden Innovationszyklen müssen kontinuierlich überwacht werden, um etwa Kompatibilität von Angeboten langfristig sicherzustellen. Hierbei ergeben sich auch neue eigenständige Leistungen, die sich bspw. mit dem zur Verfügung stellen, dem automatischen Austausch oder der sicheren

¹ Universität Hildesheim, Institut für Betriebswirtschaft und Wirtschaftsinformatik, Universitätsplatz 1, 31141 Hildesheim, schoormann@uni-hildesheim.de; Fraunhofer Institute for Software and Systems Engineering, Industrial Manufacturing, Speicherstraße 6, 44147 Dortmund,  <https://orcid.org/0000-0002-3831-1395>

² UNEOS AG, friedemann.kammler@uneos.io,  <https://orcid.org/0000-0002-2418-4882>

³ Leibniz Universität Hannover, Institut für Produktentwicklung und Gerätebau (IPEG), An der Universität 1, 30823 Garbsen, gembariski@ipeg.uni-hannover.de,  <https://orcid.org/0000-0002-2642-3445>

⁴ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Smart Enterprise Engineering, Hamburger Str. 24, 49084 Osnabrück, simon.hagen@dfki.de

Verwendung von Daten befassen. Beispiele sind Daten-Marktplätze [Va21], datenbasierte Geschäftsmodelle [Sc22] und Plattform-Geschäftsmodelle [He20]. Besonders herausfordernd ist hier die Festlegung von Preisen und Werten [Pe23], die nun dynamisch und kundenspezifisch bestimmt werden können.

Mit dem Fokus auf der Erfüllung individueller Kundenbedürfnisse können Unternehmen und gesamte WSS ihre Aktivitäten so gestalten, dass bspw. der Materialeinsatz in der Produktion und der Ausstoß von Emissionen in der Logistik minimiert werden. Im Zusammenhang mit *ökologischen* Zielen entstehen durch WSS neue Wege Ressourcen einzusparen, wiederzuwenden und/oder anderen zugänglich zu machen. Zu den Ansätzen zählen das Virtualisieren von Produktfunktionen (z.B. durch Digitale Zwillinge [WD21]), das Ersetzen von Dienstleistungen durch datengetriebene Varianten [KB19], das Teilen von Produkten (z.B. Sharing Economy [GK21]) oder das Erkunden alternativer Verwendungen von „ausgemusterten“ Leistungen [GL18, Br20].

Neben den ökonomischen und ökologischen Facetten spielt zunehmend auch die *soziale* Nachhaltigkeitsdimension in WSS eine wichtige Rolle [Wu23, SK20]. Einerseits können neue Infrastrukturen den Zugang zu großen Datenmengen ermöglichen, um auch kleinere Unternehmen und Startups in die Lage zu versetzen digitale Lösungen zu entwickeln. Andererseits besteht die Gefahr von sich verhärtenden Marktstrukturen, in denen wenige Organisationen die Macht über Zugänge und Daten haben und somit starke Abhängigkeiten schaffen [HA17]. Fragen hinsichtlich digitaler Verantwortung, transparentem Umgang mit Daten, Schaffung von Vertrauen und Sicherstellung sozial gerechter Produktions- und Lieferketten gilt es daher weiter zu untersuchen.

Aus wissenschaftlicher Perspektive entsteht im Kontext von WSS und Nachhaltigkeit ein interdisziplinäres Spannungsfeld, das unterschiedliche Fragestellungen und Anwendungsfelder verbindet. Hierzu zählt sowohl der Fortschritt datengetriebener Technologien, insbesondere unter Berücksichtigung innovativer Anwendungen der Künstlichen Intelligenz [Sc23], als auch grundlegende Infrastrukturmaßnahmen wie der Aufbau von sicheren, digitalen Datenökosystemen und Data Spaces. Andererseits ist zu klären, wie technische Innovationen messbar und zuverlässig Beiträge zu ökonomisch, ökologisch und sozial nachhaltigen Angeboten leisten können. Hierfür gilt es, auch das Verhältnis zwischen erzieltm Nutzen und in Kauf genommenen Aufwänden kritisch zu beleuchten und neue, wirksame Geschäfts- und Kooperationsmodelle zu identifizieren.

2 Beiträge für den NaWerSys-Workshop 2023

Der NaWerSys-Workshop möchte in dem oben genannten Spannungsfeld Forscherinnen und Forschern die Möglichkeit geben, aktuelle Ergebnisse und Praxiseinblicke zu präsentieren. Konzeptuelle, empirische und prototypische Beiträge sollen dabei innovative Anwendungen datengetriebener WSS verdeutlichen sowie deren Implementierung in neuen, nachhaltigeren Geschäftsmodellen reflektieren. Ein besonderes Anliegen des Workshops liegt darin, Entwicklungsteams übergreifend zusammenzubringen,

erfolgreiche Anwendungen zu diskutieren und gemeinsame Konzepte, Modelle und Methoden zu prägen, die zur systemischen Gestaltung von WSS beitragen.

Nach einem Peer-Review-Verfahren wurden fünf Beiträge für den dritten NaWerSys-Workshop – im Rahmen der INFORMATIK 2023 – angenommen, die die übergeordnete Motivation zur Gestaltung und zum Management von nachhaltigen WWS aufgreifen:

- In ihrem Beitrag betrachten *Stephanie Winkelmann, Julia Schweihoff, Ilka Jussen* und *Frederik Möller* Ansätze für die Circular Economy in der Automobilindustrie. Basierend auf öffentlichen Daten leiten sie Maßnahmen zur Förderung von Circular Economy ab und klassifizieren diese anhand eines 9R-Frameworks.
- Das Autorenteam um *Christoph Hoppe, Robert Schmelzer, Frederik Möller* und *Thorsten Schoormann* widmet sich der Frage welche Rolle das Teilen von Daten zur Erreichung von Nachhaltigkeitszielen hat. Sie identifizieren Potenziale von Data Spaces und deren Einfluss auf die ökonomische, ökologische und soziale Dimension.
- *Stefan Neubig, Dominik Rebholz, Andreas Hein, Robert Keller* und *Helmut Krmar* analysieren am Beispiel der Tourismusbranche die Potenziale von Knowledge Graphs. Die Betrachtung von Lebenszyklen und die Sicherstellung von Nachhaltigkeit nehmen dabei eine wesentliche Rolle ein.
- Der Kurzbeitrag von *Marco Di Maria, David Walter* und *Ralf Knackstedt* zeigt die Relevanz Interorganisationalem Lernen und stellt Forschungsperspektiven für Ökosysteme vor. Einen besonderen Fokus legen die Autoren auf das sog. Unlearning, das im Spannungsfeld Nachhaltigkeit und Ökosysteme eingesetzt werden kann.
- *Christoph Heinbach, Henning Gössling* und *Oliver Thomas* konzipieren in ihrem Forschungsbeitrag einen Co-bility Hub auf der Grundlage von Gaia-X. Verschiedene Mobilitätskonzepte werden integriert, um Nachhaltigkeit in einem offenen Datenökosystem zu realisieren.

3 Danksagung

Der Workshop wäre nicht möglich gewesen ohne die Unterstützung durch das Organisationsteam der Jahrestagung INFORMATIK 2023. Ein besonderer Dank geht zudem an das NaWerSys-Programmkomitee, das diesen Workshop mit sorgfältigen Gutachten unterstützt hat: *Dennis Behrens* (Data Scientist), *Katja Bley* (Deutsches Zentrum für Luft- und Raumfahrt), *Helge Fischer* (FOM Essen), *Sven Jannaber* (LVM Versicherung), *Ilka Jussen* (TU Dortmund), *Robert Keller* (Hochschule Kempten), *Marvin Auf der Landwehr* (HS Hannover), *Cristina Mihale-Wilson* (Goethe-Universität Frankfurt), *Frederik Möller* (TU Braunschweig, Fraunhofer ISST), *Dimitri Petrik* (Universität Stuttgart), *Gero Strobel* (Universität Duisburg-Essen), *Jannis Vogel* (DFKI Osnabrück) und *Stephanie Winkelmann* (TU Dortmund)


Thorsten Schoormann, Friedemann Kammler, Paul Gembarski und Simon Hagen

Litertaurverzeichnis

- [Br20] Brinker, J.; Gembarski, P.C.; Hagen, S.; Thomas, O.: Anwendungspotenziale von Additive Repair und Refurbishment für Service-orientierte Geschäftsmodelle. In *Konstruktion für die Additive Fertigung 2019* (S. 43-54). Springer Vieweg, Berlin, Heidelberg, 2020.
- [Di20] Di Vaio, A.; Palladino, R.; Hassan, R.; Escobar, O.: Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research* 121, S. 283-314, 2020.
- [El97] Elkington, J.: The triple bottom line. *Environmental management: Readings and cases*, 49-66, 1997.
- [GK21] Gembarski, P.C.; Kammler, K.: Mass Customizing for Circular and Sharing Economies: A Resource-Based View on Outside of the Box Scenarios. In *Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems*, S. 1039-1046. Springer, Cham, 2021.
- [GL18] Gembarski, P.C.; Lachmayer, R.: Product-Service-Systems – What and why Developers can learn from Mass Customization. *Enterprise Modelling and Information Systems Architectures* 13, S. 16-1, 2018.
- [HA17] Hagiú, A.; Altman, E. J.: Finding the platform in your product. *Harvard Business Review* 95/4, S. 94-100, 2017.
- [Ha19] Hagen, S.; Brinker, J.; Gembarski, P.C.; Lachmayer, R.; Thomas, O.: Integration von Smarten Produkten und Dienstleistungen im IoT-Zeitalter: Ein Graph-basierter Entwicklungsansatz. *HMD Praxis der Wirtschaftsinformatik* 56, S. 1220–1232, 2019.
- [He20] Hein, A.; Schreieck, M.; Riasanow, T.; Setzke D.S.; Wiesche, M.; Böhm, M.; Krcmar, H.: Digital platform ecosystems. *Electronic Markets* 30, S. 87-98, 2020.
- [KB19] Kühne, B.; Böhmman, T.: Data-Driven Business Models-Building the Bridge between Data and Value. In: *Proceedings of the European Conference on Information Systems*, Stockholm, Sweden, 2019.
- [Lo21] Lobschat, L.; Mueller, B.; Eggers, F.; Brandimarte, L.; Diefenbach, S.; Kroschke, M.; Wirtz, J.: Corporate digital responsibility. *Journal of Business Research*, S. 122, 2021.
- [Pe23] Petrik, D.; Springer, V.; Strobel, G.; Möller, F.; Schoormann, T.: The Price is Right: Exploring Pricing of Digital Industrial Platforms. *Information Systems Management*, 1-30, 2023.
- [Sc23] Schoormann, T.; Strobel, G.; Möller, F.; Petrik, D.; Zschech, P.: Artificial Intelligence for Sustainability—A Systematic Review of Information Systems Literature. *Communications of the Association for Information Systems*, 52(1), 8, 2023.

- [Sc22] Schweihoff, J. C.; Jussen, I.; Stachon, M.; Möller, F.: Design options for data-driven business models in data-ecosystems. In: Proceedings of the INFORMATIK, Gesellschaft für Informatik, Bonn, pp. 331-345, 2022.
- [SK20] Schoormann, T.; Kutzner, K.: Towards Understanding Social Sustainability: An Information Systems Research-Perspective. In: Proceedings of the International Conference on Information Systems, Hyderabad, India, 2020.
- [Th20] Thomas, O.; Hagen, S.; Frank, U.; Recker, J.; Wessel, L.; Kammler, F.; Zarvic, N.; Timm, I.: Global crises and the role of BISE. *Business & Information Systems Engineering*, 62/4, S. 385-396, 2020.
- [Va21] van de Ven, M.; Abbas, A. E.; Kwee, Z.; de Reuver, M.: Creating a taxonomy of business models for data marketplaces. In: Proceedings of the 34th Bled eConference: Digital Support from Crisis to Progressive Change, pp. 313-325, 2021.
- [WD21] Wache, H.; Dinter, B.: Digital Twins at the Heart of Smart Service Systems-An Action Design Research Study. In: Proceedings of the European Conference on Information Systems, Virtual, 2021.
- [Wu23] Wulfert, T.; Woroch, R.; Strobel, G.; Schoormann, T.; Bahn, L.: Unboxing the role of e-commerce ecosystems to address grand challenges. In: Proceedings of the European Conference on Information Systems, Kristiansand, Norway, 2023.

Turning Old into New – The Lane Change to a Circular Economy in the Automotive Industry

Stephanie Winkelmann ¹, Julia Schweihoff², Ilka Jussen³, Frederik Möller⁴

Abstract: Sustainability and circular economy have gained significant interest in research, politics, and society. The circular economy is a product of the transformation from a traditional "take-make-dispose"-approach to a closed-loop system, intending to reduce waste and reuse existing resources. The circular economy (CE) is particularly important for the automotive industry as it contributes greatly to global waste and pollution. Although many companies see the potential in the circular economy, research on it is still in its infancy. This paper starts precisely at this point, and analyses publicly presented measures for circular behavior in the automotive industry. We constructed a final sample of nine companies and conducted a website analysis. We assigned 150 identified measures to 30 dimensions and the 9R framework. Our findings indicate that recycling is the most applied principle in the automotive industry. Some measures do not fit the 9R framework. Thus, we argue to supplement the CE principles.

Keywords: Circular Economy, 9R framework, Automotive, Measures

1 Introduction

"Circular design means a circular mindset. We have to RE:THINK everything we do. It's about circularity through the whole process and far beyond: new materials, new technologies, new processes and, in general, new ideas." [BM23c]

The circular economy (CE) has become an exciting avenue for research and practice in the automotive industry during the transformation to more sustainable action. Not only political regulations (e.g., [EE00]) increase the pressure on automotive companies, but also customer requirements strengthen the urge to act in a more circular manner. Many associations deal with the integration of CE and argue that it is a relevant aspect of current ecosystems [SA22], [Eu18]. Integrating a CE is equated with the transformative scope of Henry Ford's assembly line system [Wo20]. The CE is an economic system that focuses

¹ Chair for Industrial Information Management, TU Dortmund University, Joseph-von-Fraunhofer-Str. 2-4, 44227 Dortmund; TU Braunschweig, Mühlenpfordtstraße 23, 38106 Braunschweig,

stephanie.winkelmann@tu-dortmund.de,  <https://orcid.org/0000-0002-1391-0180>

² Chair for Industrial Information Management, TU Dortmund University, Joseph-von-Fraunhofer-Str. 2-4, 44227 Dortmund, julia.schweihoff@tu-dortmund.de

³ Chair for Industrial Information Management, TU Dortmund University, Joseph-von-Fraunhofer-Str. 2-4, 44227 Dortmund, ilka.jussen@tu-dortmund.de

⁴ Information Systems and Data-Driven Enterprise, TU Braunschweig, Mühlenpfordtstraße 23, 38106 Braunschweig; Fraunhofer ISST, Speicherstraße 6, 44147 Dortmund, frederik.moeller@tu-braunschweig.de

on using materials as efficiently and sustainably as possible by keeping them in the economic cycle for as long as possible [KRH17]. Recently, CE has been discussed as a key enabler to introduce more sustainable production and consumption practices by decoupling economic performance and emissions [SC18], [GCU16]. Companies can promote their sustainable practices by adopting CE principles. In doing so, they have the potential to minimize their environmental impact, promote social inclusion, and improve economic efficiency. One example is e.GO [Ne23], which focuses on innovative electric vehicles and production solutions. They offset CE with an automotive design and the choice of materials to optimize energy and environmental footprint in their factory. The vehicles can be dismantled at the end of the product life, and most materials are recycled. For replaced components, e.Go Mobile introduces another use, such as using old batteries for home solar storage. They enable customers to share, use, and rent their vehicles to improve economic efficiency. The automotive industry is a major user of natural resources, such as different metals. A study by Capgemini shows that the automotive industry needs to make up a 20% shortfall in its current investment to achieve the international climate targets [Wi20].

Yet, automotive is not the focus of recent research, even though it is an essential industry with a high impact on sustainability. [Ro22] analyzed the impact of implementing CE principles on the financial performance of manufacturing companies in the automotive industry. The results show that the implementation of the principles strengthens performance and that this effect is mediated by the company's commitment to promoting better Sustainable Supply Chain Management (SSCM). Currently, each company has its own system to implement circular economy. The processes run separately without any distinct coordination in the supply chain. The topic of remanufacturing has received little attention to date. For this reason, it seems relevant to define and collect possible measures. They should serve as a basis to implement the circular economy within the entire supply chain. Because of the above, we ask the following research question (RQ):

What measures do companies in the automotive industry present to foster CE?

To answer the RQ, we empirically investigated publicly available information. We examined websites from nine companies and identified actions that they are taking toward CE. We have assigned these measures to the 9R framework to obtain an overview of concrete CE measures in the automotive industry and derive a focus of the companies.

The paper is structured as follows. In the upcoming section, we introduce CE in the theoretical background. Then we explain our research design in detail in Section 3. Section 4 presents our findings from the public data analysis and discusses them in Section 5. Finally, we conclude our paper with an overview of the contributions and limitations and give an outlook for further research in Section 6.

2 Circular Economy

The CE is attracting increasing interest in research and practice [D'17], resulting in the emergence of a range of definitions [Ge17], [GCU16]. [KRH17] synthesized these definitions and proposed the following one, which we apply to our research: "A circular economy describes an economic system that is based on business models which replace the 'end-of-life' concept with reducing, alternatively reusing, recycling and recovering materials in production/distribution and consumption processes, [...] with the aim to accomplish sustainable development [...] to the benefit of current and future generations" [KRH17].

Since CE can have many benefits [St16] but also some barriers [Ki18] for companies, many researchers addressed it in different disciplines, for example, e-commerce, technology and innovation, and management information systems. In Information Systems (IS) research, [Ze19] has developed a taxonomy of information flows as a basis for the successful application of CE practices. They build their taxonomy on the R-framework of [Po17]. The taxonomy illustrates that information plays an essential role in CE, such as the properties and state of material objects, the use of these objects, or instructions for transforming them. Furthermore, they propose which types of information affect which CE principle. Thus, they contribute a basis for further discussion on using data and information for CE. [Ze21] developed a research agenda for IS research for CE. It includes directions to advance knowledge of how information systems and illustrates the iterative progression of actors from understanding to implementing CE principles. [Nu23] analyzed real-world use cases from the building industry and extracted a set of strategies to apply CE. They identified and discussed significant challenges of life cycle assessment. The study shows that future research should identify circular strategies with high potential that are suitable for implementation.

Potting's 9R framework (see Figure 1) is the most prominent and widely used in CE research. The framework consists of the ten principles to adopt CE in business: *Refuse*, *Rethink*, *Reduce*, *Re-use*, *Repair*, *Refurbish*, *Remanufacture*, *Repurpose*, *Recycle*, and *Recover*. They are ranked based on their priority from a high level of circularity (low R-number) to a low level (high R-number).

[Mo20] investigates existing CE targets, proposes new ones, and matches them to the R-Strategies by [Po17]. He argues that targets can facilitate a transition towards a CE in several ways and that most relate to waste management or resource conversation, even if their impact is limited. Through recovery and recycling activities, products do not remain in the economy. He proposed further targets for other CE principles such as remanufacturing, refurbishing, repairing, or re-using. He argues that these principles are more potent in increasing CE.

CE principle	Description
Refuse (R 0)	Make product redundant by abandoning its function or by offering the same function with a radically different product.
Rethink (R 1)	Make product use more intensive (e.g., through sharing products, or by putting multi-functional products on the market).
Reduce (R 2)	Increase efficiency in product manufacture or use by consuming fewer natural resources and materials.
Re-use (R 3)	Re-use by another consume of discarded product which is still in good condition and fulfils its original function.
Repair (R 4)	Repair and maintenance of defective product so it can be used with its original function.
Refurbish (R 5)	Restore an old product and bring it up to date.
Remanufacture (R 6)	Use parts of discarded product in a new product with the same function.
Repurpose (R 7)	Use discarded product or its parts in a new product with a different function.
Recycle (R 8)	Process materials to obtain the same (high grade) or lower (low grade) quality.
Recover (R 9)	Incineration of materials with energy recovery.

Fig. 1: The 9R framework [Po17]

3 Research Design

To answer the research question, we explored public data (e.g., websites, company reports, company news, or newsletter) from companies in the automotive industry. First, we identified relevant automotive companies using the rankings by [Ha20] and [St23]. We defined the following filter criteria for the final sample: (1) the companies must present the topic of CE on their website and (2) suggest concrete measures for CE. Our final sample comprises nine large automotive companies: BMW AG, Mercedes-Benz Group AG, Ford Motor Company, Jaguar Land Rover Automotive PLC, Nissan Motor Co. Ltd., Renault Group, Stellantis N.V., Toyota Motor Corporation, and Volkswagen AG.

Based on their public data, we coded and abstracted each identified measure (see Figure 3) and assigned them to the 9R framework. It concerns the role of innovation in circular economy transitions in product chains and is the most nuanced and integrated conceptualization to date (e.g., [Mo20], [Ze19]). The 9R principles can help companies transform from a throwaway culture to CE by identifying which actions are highly relevant to CE and which have a lower CE-level. By adopting this framework a company can increase their level of CE.

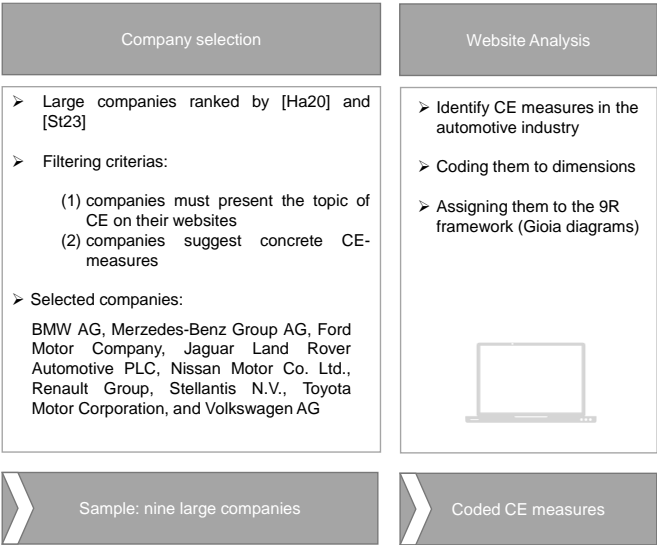


Fig. 2: Research process

To analyze and summarize the findings of the company analysis in a structured way, we created Gioia diagrams. The "grounded theorizing" approach of [GCH13] is used to summarize elements into themes or concepts (see Figure 3). In this paper, we have assigned direct quotes from the website that dealt with concrete activities of CE to superordinate 1st-order concepts. Subsequently, several 1st-order concepts are combined into aggregated terms. In this case, these aggregated terms are the CE principles.

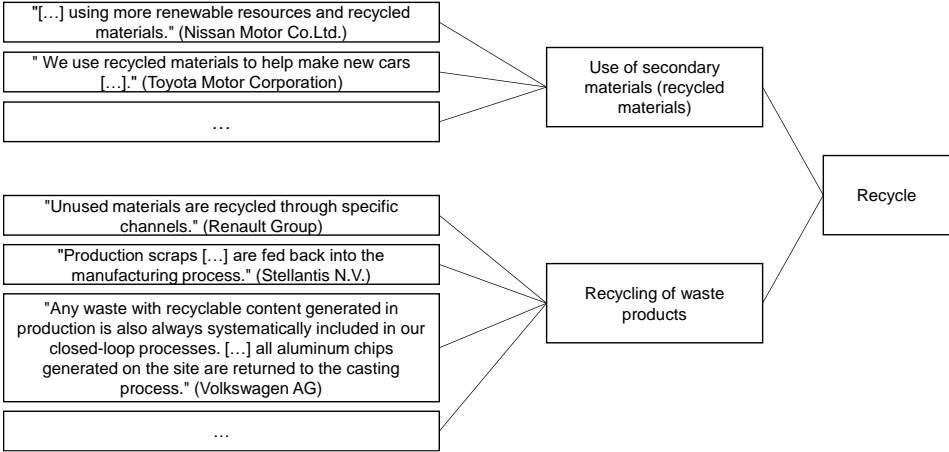


Fig. 3: Illustrative Gioia diagram

[Ni23] and [To23] illustrate using recycled materials to produce new cars. Further, companies report on using secondary materials. We summarized such measures using *secondary materials (recycled materials)* and matched the dimension to the CE principle of *recycling*. Another dimension that fits this CE principle is the *recycling of waste products*. [St22] feeds back production scraps into the manufacturing process, [Re22] recycles unused materials, and [Vo21] includes any waste in their closed-loop processes.

4 The Circular Economy in the Automotive Industry

During the analysis, we identified 150 CE measures listed by the nine companies. For example, [BM23] aims "to reduce consumption of resources by using high-quality secondary material". [Ni23] is "promoting the collection and recycling of used aluminum road wheels" to act in a cycle-oriented manner. In addition to other companies, [Me23] has "circular economy in mind from the very start whenever [they] develop a product". We combined these measures into 30 dimensions, such as building a network, development of new recycling processes, or the preparation of products/components, and assigned them to the 9R framework (see Figure 4).

Our results show that the most significant dimension is the *use of secondary materials (recycled materials)*. 14% of all measures and seven out of nine companies deal with the use of such materials to manufacture new vehicles. [Vo21] state that "the growing use of secondary materials and the establishment of closed material loops is helping to significantly reduce [their] CO₂ emissions". [Fo23] has "a collaboration with McDonald's USA in 2019 to turn coffee chaff, a waste byproduct from McDonald's coffee production, into vehicle parts". This also highlights the relevance of ecosystems to realize CE within the automotive industry, but also cross-sector. Our other results support this finding. More than half of the companies studied are practicing measures to form an ecosystem. The Renault Group built a factory called Refactory in 2020 and issue that "this industrial ecosystem is open to start-ups and partnerships. Its aim is to encourage initiatives, develop innovation in the service of the circular economy [...]" [Re23]. [Ja22] implement collaboration and knowledge sharing within an ecosystem as a cornerstone in their strategy and say that this allows them to reimagine new vehicle features.

Another interesting finding is the relevance of the dimension *circular product design*. 12% of all measures refer to the design of a new automotive. [BM23] see circular product design as a prerequisite "to achieve [CE, and argue that] vehicles must be designed to be as resource-efficient as possible as early as the production phase." While Toyota is designing its vehicles with an easy-to-dismantle concept, Volkswagen takes care of the recyclability of the used materials [To23], [Vo22]. Nissan Motor Corporation developed design guidelines to construct a new vehicle [Ni23d]. The results show that this dimension is important in the automotive industry but not yet focused on CE principles. No principles clearly fit this dimension. We thus allocated it to the dimension *other* and discussed this aspect in a further step (see Section 5).

R0 – R9 high to low level of CE											
None measure	1 – 4 measures										
5 – 9 measures	≥ 10 measures										

	R0	R1	R2	R3	R4	R5	R6	R7	R8	R9	
Sum	Refuse	Rethink	Reduce	Re-use	Repair	Refurbish	Remanufacture	Repurpose	Recycle	Recover	Other
	150	5	6	18	12	2	9	11	3	46	0
offer of sharing services	1	0	1	0	0	0	0	0	0	0	0
building a network	8	0	0	0	0	0	1	0	1	1	0
preparation of products/components	8	0	0	0	0	0	7	1	0	0	0
creation of secondary material (recycled materials)	2	0	0	0	0	0	0	0	2	0	0
digitization	6	0	2	4	0	0	0	0	0	0	0
introduction of new business areas	8	0	0	0	0	0	2	0	1	0	5
saving of critical materials	4	0	0	4	0	0	0	0	0	0	0
development of alternative products	5	0	0	0	0	0	4	1	0	0	0
development of new manufacturing processes	3	0	0	2	0	0	0	0	0	0	1
development of new recycling processes	8	0	0	0	0	0	0	0	7	0	1
development of new repair processes	1	0	0	0	0	1	0	0	0	0	0
development of new technologies	6	0	0	2	0	0	0	0	2	0	2
development of resource-saving products	3	2	0	1	0	0	0	0	0	0	0
investment in new technologies	2	0	0	0	0	0	0	0	0	0	2
promoting rethinking	2	0	0	0	1	0	0	1	0	0	0
circular product design	18	0	0	0	0	0	0	0	0	0	18
material return	8	0	0	0	1	0	0	1	0	6	0
employee training	3	0	0	0	0	0	0	0	0	0	3
material recycling	1	0	0	0	0	0	0	0	1	0	0
repair of products	1	0	0	0	0	1	0	0	0	0	0
collection points	1	0	0	0	0	0	0	0	1	0	0
training of corporate partners	1	0	0	0	0	0	0	0	0	0	1
conversion of old parts	2	0	0	0	1	0	0	1	0	0	0
sale of secondary materials (recycled materials)	6	0	2	0	4	0	0	0	0	0	0
use of durable materials	1	0	1	0	0	0	0	0	0	0	0
use of renewable raw materials	4	2	0	2	0	0	0	0	0	0	0
use of alternatives	2	1	0	1	0	0	0	0	0	0	0
use of secondary materials (recycled materials)	21	0	0	1	1	0	0	1	1	17	0
recycling of waste products	11	0	0	1	1	0	1	0	0	8	0
reuse of transport material	3	0	0	3	0	0	0	0	0	0	0

Fig. 4: Heat map

We also analyzed the expression of CE principles in the automotive industry. Our results show a dissimilarity between the *prioritization of the CE principles* in industry and research. As seen in Figure 1, *recycling* describes a low level of CE. However, most identified measures (31%) refer to this principle. The *refuse* principle is exactly the opposite. According to research, the principle ensures an increasing CE. But it is hardly implemented in the automotive industry. Only three companies develop and use biomaterials or sustainable resources instead of critical ones (e.g., [Ni23]). An interesting aspect is the agreement between research and industry on *recover*. Recover describes the energetic utilization of material (see Figure 1). None of the companies listed such measures on their websites, and even scholars argue that the principle does not support an increasing CE.

In addition to the diversity in the principles, we identified a different understanding of the CE principles in the automotive industry and research. *Rethink* is an excellent example to highlight this. Scholars define this principle as making product use more intensive, e.g., through car sharing. However, from the measures BMW AG assigns to rethink, we deduce that they operationalize the principle as the general rethinking towards CE. They list measures like "invested in key technologies" [BM23a] or "the vision vehicle [...] were

rethought, developed and optimized from the ground up" [BM23b], which are not consistent with the understanding in research. Other measures also address the issue of rethinking. [Ni23], for example, certified their dealerships to "stay up to date on the practices to be carried out through the manual complied with the Automobile Recycling Law". [Re22] is training his employees through a certified program and is "offering training modules on the circular economy open to all audiences". [St22] is scaling up its teams to act more circulation oriented. All measures underline the significance of such rethinking, but a principle in research is still missing.

5 Discussion

Our findings support that the automotive industry already implements CE in its current business. Potential reasons for this development are, in addition to competitive advantages and legal regulations, more efficient use of raw materials and energy, which in turn can lead to cost advantages.

The careful use of resources is becoming increasingly important. The *use of secondary or recycled material* is the most listed measure in the automotive industry. One possible reason is the legal regulation of the European Union which requires that European automotive companies have to fulfill a recovery rate of 85% for reuse and 95% for recycling for each vehicle [EE00]. To ensure that the cycle is as closed as possible, automotive companies are increasingly relying on vehicle return programs. As a result, they get back components that they refurbish and reuse as secondary or recycled material. However, the reuse of materials can also reduce costs. Especially when considering that the costs for raw material extraction are often not integrated into the manufacturing costs of new parts. Many companies promote that they act according to the regulation by [EE00]. But does that also apply to smaller companies such as service providers or suppliers? The investigation of CE implementation in such companies is an interesting aspect for further research. Analyzing existing requirements and barriers is a significant step to enable smaller companies to participate in CE in the automotive industry. Another research topic could be the study of benefits resulting from an ecosystem to support the implementation of CE principles. At this point, it is meaningful to focus on data available throughout the product life cycle and the transparency inside such an ecosystem. This can support the product developers to design a more circular-orientated product.

Advanced technologies, new materials, and highly effective components are obtainable to enhance vehicle design. During vehicle engineering, progress in carefully using non-renewable resources affects the whole ecosystem. We identified that *circular product design* is one of the most named measures and highlighted by six out of nine companies. [WBZ22] developed a roadmap for circular product design and provided practical guidance for businesses and policymakers. Our findings contribute to the understanding that design plays a significant role within CE practices and confirms the findings of [DRB22]. Although our analysis and other research show that it is an essential aspect of the automotive industry,

circular product design does not fit the CE principles. Thus, we argue to supplement the 9R framework and add *redesign* as a further CE principle. This new principle describes all measures that refer to the design phase of a product, e.g., to design and construct mono-materials and detachable connections to allow easy dismantling or to prepare design guidelines to advance the recycling design of new vehicles. Products reach their full potential for CE if developers invest much time in their design. Investments in the design mean fewer costs for the product dismantling. The challenge is holistic evaluation and planning with little product knowledge. Developers have to make many decisions without any funded understanding of the future product life. We suggest investigating the use of technologies during the construction phase, e.g., technologies to gather all data during the product life cycle.

Implementing a successful CE in the automotive industry is based on the products. But these, in turn, are rather only the starting point. Our results show that there is a *different prioritization in industry and research*. This may also be due to the complexity and durability of the vehicles. While products in other industries may comprise only a few components, a vehicle is a construction of many individual components. During product use, cars are exposed to different conditions, thus, many challenges are present, e.g., products are exposed to temperature fluctuations and impacts. This limits the possibilities for the use of alternatives. Components can not simply be made redundant by abandoning their function or by offering the same function with a completely different product (*refuse*). The intensification of product use (*rethink*) is also limited. Car sharing is one option, but it is usually offered by other service providers. At this point, we address our limitations (see Section 6). We did not consider using secondary or recycled material as the intensification of product use. We mostly assign this to the principle *recycle*. These and other reasons may be the cause of different prioritization. Further research can start at this point.

Our results also show that there is a *different understanding of the CE principles* in research and industry. *Rethink* in the automotive industry fits better with changing the company's vision, strategy, and culture than intensifying product use. It has a wider connotation, as it includes new concepts, new processes, or new business models. At this point, the elaboration of the understanding of CE in the automotive industry arises to reflect our findings with industry experts. One aspect is to find out if companies are even aware of the 9R framework and their opinion on the principles *rethink*. A second aspect is the investigation of reasons for the different prioritization to identify reasons for the distinction.

6 Contributions, Limitations, and Outlook

The paper represents a *contribution* to science and practice. Our paper generally provides an overview of CE measures in the automotive industry and shows how CE has been implemented so far. Our research results are based on the 150 measures listed by nine com-

panies we identified through website analysis. For *practical* purposes, we primarily provide an overview of how CE is already being implemented. Based on the implementation possibilities for the CE principles that we have worked out, companies can derive conclusions or initial inspiration for the implementation of CE in their own company. Our contribution to *research* is that our findings are a starting point for working up implementation possibilities of CE in practice for deeper research. Likewise, our work creates an understanding of CE in the automotive industry. Ultimately, we also point out research gaps where research can start in the future to support practice (i.e., companies) in the best possible way in implementing CE.

Due to various aspects, this paper also has some *limitations*. First, our results are limited by our choice of automotive companies and their websites, e.g. selection based on two studies, or our narrowing of the research scope to automotive manufacturers. Second, companies represent themselves on their websites. It could be possible that measures mentioned on these websites are not really implemented. Greenwashing should be noted at this point and limits the results. This limits the validity of our study. Third, the timing of the search is also a limiting aspect. As the automotive industry is dynamically changing, the companies may have edited their websites and supplemented other measures that no longer fell within the period under consideration and were thus not considered. This results in a short applicability of our results. During the analysis, we found that it is sometimes difficult to clearly assign the listed measures to the 9R framework without further context. In the case of principles *remanufacturing* and *repurpose*, information regarding the further use of a component is needed. However, we have always used our best judgment. Other authors could, however, arrive at different results as a result.

Our paper offers starting points for *new research*. In the next step, we would like to expand the database through interviews, expert discussions, or workshops to gain deeper and more detailed insights into the implementation of CE in the automotive industry. This will then allow implementation options for the CE principles to be prepared in more detail. In addition to the more in-depth analysis of the aspects just mentioned, the investigation of the influence of data on the implementation of CE represents an exciting approach. Possible questions at this point could be, for example, whether the use of data can make CE more efficient or whether the sharing of knowledge can create a lasting change in the automotive industry in the direction of CE. Fundamentally, analyzing the whole automotive supply chain regarding CE implementation is an interesting follow-up study.

Acknowledgement

This research and development project is funded by the German Federal Ministry of Education and Research (BMBF) within the Digital ecosystem for a sustainable circular economy in the automotive industry (DIONA) Funding Action (02J21E200). The author is responsible for the content of this publication.

Bibliography

- [BM23a] BMW AG: Circularity. <https://www.bmwgroup.com/en/sustainability/our-focus/circularity.html>, accessed: 08/05/2023.
- [BM23b] BMW AG: Focus on circular economy: sustainable into 2040. <https://www.bmw.com/en/magazine/sustainability/circularity-at-bmw.html>, accessed: 08/05/2023.
- [BM23c] BMW AG: The principle of circular design. <https://www.bmw.com/en/magazine/sustainability/circular-lab/rethink/the-principles-of-circular-design.html>, accessed : 08/05/2023.
- [D'17] D'Amato, D. et al.: Green, circular, bio economy: A comparative analysis of sustainability avenues. *Journal of Cleaner Production* 168, pp. 716–734, 2017.
- [DRB22] Diaz, A.; Reyes, T.; Baumgartner, R. J.: Implementing circular economy strategies during product development. *Resources, Conservation and Recycling* 184, p. 106344, 2022.
- [EE00] Richtlinie 2000/53/EG über Altfahrzeuge. AltfahrzeugRL, 2000.
- [Eu18] European Automobile Manufacturers' Association: Circular economy. <https://www.acea.auto/fact/circular-economy/>, accessed: 08/05/2023.
- [Fo23] Ford Motor Company: Ford commits to carbon neutrality by 2050. <https://corporate.ford.com/articles/sustainability/ford-expands-climate-change-goals.html>, accessed: 08/05/2023.
- [GCH13] Gioia, D. A.; Corley, K. G.; Hamilton, A. L.: Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods* 1/16, pp. 15–31, 2013.
- [GCU16] Ghisellini, P.; Cialani, C.; Ulgiati, S.: A Review on Circular Economy. The Expected Transition to a Balanced Interplay of Environmental and Economic Systems. *Journal of Cleaner Production* 114, pp. 11–32, 2016.
- [Ge17] Geissdoerfer, M. et al.: The Circular Economy – A new sustainability paradigm? *Journal of Cleaner Production* 143, pp. 757–768, 2017.
- [Ha20] Handelsblatt: VW vor Toyota: Das sind die größten Autohersteller der Welt. <https://www.handelsblatt.com/mobilitaet/motor/ranking-vw-vor-toyota-das-sind-die-groessten-autohersteller-der-welt/25560670.html>, accessed: 08/05/2023.
- [Ja22] Jaguar Land Rover Automotive PLC: Annual Report 2021/2022, 2022.
- [Ki18] Kirchherr, J. et al.: Barriers to the Circular Economy: Evidence From the European Union (EU). *Ecological Economics* 150, pp. 264–272, 2018.
- [KRH17] Kirchherr, J.; Reike, D.; Hekkert, M. P.: Conceptualizing the Circular Economy. An Analysis of 114 Definitions. *Resources, Conservation and Recycling* 127, pp. 221–232.
- [Me23] Mercedes-Benz Group AG: Batteries are a key focus of the circular economy. <https://group.mercedes-benz.com/sustainability/resources/battery.html>, accessed: 08/05/2023.

- [Mo20] Morseletto, P.: Targets for a circular economy. *Resources, Conservation & Recycling* 153, pp. 1–12, 2020.
- [Ne23] Next.e.GO Mobile SE: Holistic sustainability. Welcome to the future. <https://e-go-mobile.com/en/sustainability>, accessed: 08/05/2023.
- [Ni23a] Nissan Motor Corporation: Toward Zero New Material Resource Use. <https://www.nissan-global.com/EN/SUSTAINABILITY/ENVIRONMENT/GREENPROGRAM/DEPENDENCY/RESOURCES/>, accessed: 08/05/2023.
- [Ni23b] Nissan Motor Corporation: Activities at Dealers. <https://www.nissan-global.com/EN/SUSTAINABILITY/ENVIRONMENT/GREENPROGRAM/DEPENDENCY/DEALER/>, accessed: 08/05/2023.
- [Ni23c] Nissan Motor Corporation: Initiatives for Zero Emissions. Vehicle initiatives to realize a carbon-neutral society, https://www.nissan-global.com/EN/INNOVATION/TECHNOLOGY/ARCHIVE/ZERO_EMISSION/, accessed: 08/05/2023.
- [Ni23d] Nissan Motor Corporation: Activities in the Development Stage. <https://www.nissan-global.com/EN/SUSTAINABILITY/ENVIRONMENT/GREENPROGRAM/DEPENDENCY/PRODUCT/>, accessed: 08/05/2023.
- [Nu23] Nußholz, J. et al.: From circular strategies to actions: 65 European circular building cases and their decarbonisation potential. *Resources, Conservation & Recycling Advances* 17, p. 200130, 2023.
- [Po17] Potting, J. et al.: *Circular Economy: Measuring Innovation in Product Chain*, The Hague, 2017.
- [Re22] Renault Group: Refactory of Flins. https://www.renaultgroup.com/wp-content/uploads/2022/05/202203_rg_plaquette_refactory_12_en.pdf, accessed: 08/05/2023.
- [Re23] Renault Group: Circular economy. <https://www.renaultgroup.com/en/our-commitments/respect-for-the-environment/circular-economy/>, accessed: 08/05/2023.
- [Ro22] Rodríguez-González, R. M. et al.: Does circular economy affect financial performance? The mediating role of sustainable supply chain management in the automotive industry. *Journal of Cleaner Production* 379, p. 134670, 2022.
- [SA22] SAP SE: Enabling a Circular Economy Through Industry Network Collaboration. <https://news.sap.com/2022/05/circular-economy-automotive-industry-network-collaboration/>, accessed: 08/05/2023.
- [SC18] Stål, H. I.; Corvellec, H.: A decoupling perspective on circular business model implementation: Illustrations from Swedish apparel. *Journal of Cleaner Production* 171, pp. 630–643, 2018.
- [St16] Stahel, W. R.: The circular economy. *Nature* 7595/531, pp. 435–438, 2016.
- [St22] Stellantis N.V.: Stellantis Fosters Circular Economy Ambitions with Dedicated Business Unit to Power New Era of Sustainable Manufacturing and Consumption. <https://www.stellantis.com/en/news/press-releases/2022/october/stellantis-fosters-circular-economy-ambitions-with-dedicated-business-unit-to-power-new-era-of-sustainable-manufacturing-and-consumption>, accessed: 08/05/2023.

- [St23] Statista: Größte Automobilhersteller weltweit nach Fahrzeugabsatz in den Jahren 2021 und 2022. <https://de.statista.com/statistik/daten/studie/173795/umfrage/automobilhersteller-nach-weltweitem-fahrzeugabsatz/>, accessed: 08/05/2023.
- [Ti23] Titan Advanced Energy Solutions: Titan Advanced Energy Solutions. <https://www.titanaes.com/>, accessed: 08/05/2023.
- [To23] Toyota Motor Corporation: Circularity. Building a circular economy. <https://www.toyota-europe.com/sustainability/circularity>, accessed: 08/05/2023.
- [Vo21] Volkswagen AG: Circular Economy. https://www.volkswagenag.com/presence/nachhaltigkeit/documents/sustainability-report/2021/focus-topics/220310_VW_NB21_Circular_Economy_EN.pdf, accessed: 08/05/2023.
- [Vo22] Volkswagen AG: Circular Economy, 2022.
- [WBZ22] Wang, J. X.; Burke, H.; Zhang, A.: Overcoming barriers to circular product design. *International Journal of Production Economics* 243, p. 108346, 2022.
- [Wi20] Winkler, M. et al.: *The Automotive Industry in the Era of Sustainability*, 2020.
- [Wo20] World Economic Forum: The circular economy could forever change how cars are made – here’s how. <https://www.weforum.org/agenda/2020/12/how-the-circular-economy-could-forever-change-how-cars-are-made/>, accessed: 08/05/2023.
- [Ze19] Zeiss, R.: Information Flows in Circular Economy. In: *Proceedings of the 27th European Conference on Information Systems*, Stockholm, Sweden, 2019.
- [Ze21] Zeiss, R. et al.: Mobilising information systems scholarship for a circular economy: Review, synthesis, and directions for future research. *Information Systems Journal* 31 (1), pp. 148–183, 2021.
- [ZE23] ZEVX: Our Mission. <https://zevx.com/about>, accessed: 08/05/2023.

Data Spaces as Enablers for Sustainability

Christoph Hoppe¹, Robert Schmelzer², Frederik Möller³, Thorsten Schoormann⁴

Abstract: One of our society's most fundamental challenges is promoting sustainable development. Data sharing across organizations is one way to spur innovation and address the Sustainable Development Goals (SDGs) through optimizing resource utilization, fostering circular supply chains, and producing accurate information about CO₂ emissions. However, organizations often hesitate to share data given a range of concerns, including the fear of data misappropriation or the lack of control over what others do with their data. Data spaces as a novel artifact seek to tackle these issues by providing an integrated data management approach upholding data sovereignty. These spaces can boost sustainable development by being a secure and trusted digital infrastructure for organizations to share data for sustainable purposes. To disclose how this digital solution fosters sustainability, we present a set of 65 potentials of data spaces along with the dimensions of ecological, economic, and social sustainability.

Keywords: Data Spaces, Sustainability, Sustainable Development, Circular Economy

1 Introduction

Fighting global warming, reducing greenhouse emissions, improving personal health, and enabling sustainable economic growth are just a few of the most pressing issues our society faces today [Mu21], [Ro21], [Sc21]. To make these ecological, economic, and social concerns more tangible, the United Nations (UN) proposed the 17 Sustainable Development Goals (SDGs), which illustrate a shared vision of how to achieve and preserve sustainability for societies and the planet [Un15]. One promising enabler in this context is digitalization [Ri22], [Sc21], [SL17]. Particularly, novel technologies like cyber-physical systems, the Internet of Things (IoT), and data spaces generate new opportunities for inter-organizational data sharing and pave the way toward more sustainable value creation within ecosystems [AUK18], [Wo15], [De15]. This is also indicated by a review of 550 data sharing cases conducted by the Boston Consulting Group, which concludes that data sharing can contribute significantly to sustainable

¹ Fraunhofer Institute for Software and Systems Engineering, Industrial Manufacturing, Speicherstraße 6, 44147 Dortmund, christoph.hoppe@isst.fraunhofer.de

² Technische Universität Chemnitz, Business Informatics, Straße der Nationen 62, 09111 Chemnitz, robert.schmelzer@wirtschaft.tu-chemnitz.de

³ Technische Universität Braunschweig, Institute for Business Informatics, Rebenring 58a, 38106 Braunschweig; Fraunhofer Institute for Software and Systems Engineering, Industrial Manufacturing, Speicherstraße 6, 44147 Dortmund, frederik.moeller@tu-braunschweig.de

⁴ Universität Hildesheim, Department for Information Systems, Samelsonplatz, 31141 Hildesheim; Fraunhofer Institute for Software and Systems Engineering, Industrial Manufacturing, Speicherstraße 6, 44147 Dortmund, thorsten.schoormann@uni-hildesheim.de

development. For example, aggregating shared data from various sources can help address climate change by combining and verifying the same data from multiple sources or complementing different data sources, for example, in optimizing energy efficiency or supporting sustainable agricultural farming through data [Ru21]. In this paper, we take the view of one of these technical enablers, namely *data spaces*, and explore how they contribute to sustainable development by acting as a digital infrastructure for multiple participants to share data in ecosystems.

Data spaces offer the promising potential to address various sustainability goals across ecological, economic, and social concerns in practice. However, given that data spaces as technological artifacts are novel and still emerging, these potentials need to be uncovered rigorously to make them tangible for future research and applications. Consequently, we explore this phenomenon and pursue the following research question (*RQ*): *How do data spaces promote sustainability?* To answer this RQ, we report on a systematic literature review and analyze 60 data spaces working as enablers for sustainability.

This paper is structured as follows: First, the theoretical foundations of sustainability and data spaces are introduced (Section 2). Afterward, we illustrate our research approach, which follows a systematic literature review (e.g., [Br09]) as well as a procedure for clustering and conceptualization of the collected insights [Ku22] (Section 3). Following this, we present a map of data space potentials, structured according to their domain and dimensions of sustainability (Section 4). We then discuss selected use cases in detail and illustrate how they promote sustainable value creation in practice (Section 5). Finally, we conclude with our paper and indicate further research opportunities (Section 6).

2 Research Context

2.1 Sustainability and circular economy

Sustainability describes a quasi-stable state in the long term, which is considered a societal goal or paradigm [BK87]. The path to these goals is usually described as sustainable development. Its three pillars are: ecology, economy and social equity [EI98]. Due to the increasing ecological challenges, the term is strongly charged with regard to the ecological dimension. Apart from the focus on one dimension, this paradigm is increasingly finding its way into all social considerations and disciplines such as information systems and engineering, but also psychology and education [PMR19]. One model that is being discussed in various places for achieving the SDGs is the circular economy (CE). According to Kirchherr et al.'s [KRH17] analysis of 114 definitions, the following characteristics of the concept were proposed: 1. “End of life” is avoided by reducing, reusing, recycling and recovering materials along the value and use chain; 2. It impacts at the micro (products, companies, consumers), meso (eco-industrial parks) and macro (city, region, nation and beyond) levels; 3. It aims to achieve sustainable development while creating environmental quality, economic prosperity and social equity for the benefit of

future generations; 4. Through responsible consumers and new business models [KRH17]. In order to achieve an operating CE in practice, technical support to enable exchange of data is needed. One promising technology to ensure secure data sharing is data spaces.

2.2 Data spaces

Data spaces were introduced in computer science about 15 years ago. At that time, the term referred mostly to data integration concepts, which integrate data without a physical layer [FHM05], [HFM06]. Moreover, a data space refers to the co-existence of structured, semi-structured, and unstructured data distributed over several sources. They manage an extensive collection of data shared over various sources [Si13], [SJ11]. Nowadays, data spaces can be defined as an enabler of data sharing between organizations (e.g., in larger ecosystems), based on a set of relationships among them [Ot22], [SJ11]. According to Otto [Ot22] a data space is “a distributed data integration concept. Thus, there is no central data store or data vault into which data providers deliver their data and from where it can be accessed and retrieved by data consumers. In contrast, the exchange of the data happens directly between the two participants.”

The logic pursued in data spaces is to keep the operational data with the data providers, making common database schemas obsolete [Ot22]. The fundamental traits of data spaces include data integration through a semantic level and vocabularies in accordance with linked data principles, retaining a decentralized data infrastructure and the ability to nest and overlap data [FHM05], [HFM06], [OB21]. The variety of ways data space users can connect and share data leads to the formation of data ecosystems. These data ecosystems are socio-technical systems enabling its participants to (re-)use data, promote collaboration between data providers and consumers, and link data to cutting-edge services [GGO21], [OL18]. The underlying software infrastructure of data spaces and data ecosystems must support interoperability, trust, and data sovereignty. Additionally, data spaces meet the requirements of the Data Governance Act (DGA) of the European Union, which mandates neutrality in data-sharing intermediaries [Eu20], [Ot22]. The implementation of data spaces in practice is supported by concepts such as GAIA-X [Ga19] and the IDS Reference Architecture Model (IDS-RAM) [In23], [Ot22]. The IDS-RAM describes the central software components of data spaces and provides a blueprint of a secure gateway for trusted data exchange [Ot22]. Since GAIA-X includes not only data sharing but also the storage and management of data on cloud platforms, it strives to achieve data sovereignty in a broader sense [Ga19]. In practice, the IDS-RAM and GAIA-X are closely aligned in order to allow seamless integration of architectures and support processes. From a business point of view, the term – data space – led to a common understanding of collaboration on data, driven by the desire to accomplish common business goals [Ot22]. As the centerpiece of multiple European research initiatives and industry projects, data spaces have recently received huge attention [Gi23]. Especially in the scope of sustainability and CE, data spaces are seen as a key enabler for a data-driven, sustainable value creation. A leading example in this regard is the Green Deal Data Space (GDDS) [Ad23], which is the first cross-domain data space for projects and services

related to resilience and sustainability. Moreover, many other industrial data spaces address this context, such as Catena-X, an initiative launched by parts of the German automotive industry and government. Catena-X focuses on building trusted data chains in the automotive supply chain and implementing various use cases towards more sustainable value creation, such as CO₂ monitoring or material traceability [Ca22]. Both examples illustrate the potential of data spaces as enablers for sustainability, making a closer examination of them of major interest.

3 Research Design

In this paper, we report on a systematic literature review uncovering the potential of data spaces regarding sustainability. Although data spaces are a novel technological phenomenon, we observe a variety of data space initiatives entering the market. In consequence, from our perspective, the field has matured and collected a body of objects to study meriting a deeper investigation. Next, orientated at Kundisch et al. [Ku22], we describe our approach to building a potential map of data spaces enabling sustainability. We adapted a taxonomy-based approach because this allows us to make use of both empirical objects (real data spaces) and conceptual knowledge (captured in literature).

Data spaces and their application for sustainability is a field yet largely untapped in IS research. As there is a growing interest in such spaces, our review predominantly draws from the analysis of publically available data spaces and describes how contribute to sustainable development. In total, our review passed three iterations consisting of one conceptual-to-empirical iteration of investigating the literature and two empirical-to-conceptual iterations of examining real-world data spaces to disclose their potential towards sustainably (see Table 1). Below, each iteration is described in greater detail.

Iteration	Source	Sample
Iteration 1	Systematic literature review in Scopus, IEEE Xplore, Science Direct	n=9
Iteration 2	Empirical analysis of International Data Spaces Association (IDSA), EuPro Gigant	n=22 n=16
Iteration 3	Internet Research	n=13

Tab. 1: Overview of data sources

Within the *first iteration (conceptual)*, we conducted a literature review with the scope to find scientific papers regarding data spaces and sustainability, published in common scientific databases, including Scopus, Science Direct, and IEEE Xplore. Therefore, we adapted the established methods of Webster and Watson [WW02] and vom Brocke et al. [Br09]. In order to perform the review, vom Brocke et al. [Br09] proposed a framework that consists of five steps. In the first step, a definition of the review scope must be given. The second step focuses on the conceptualization of the topic to be researched. In step

three, the current literature search took place. Within the fourth step, the literature found is analyzed and synthesized. Subsequently, the research agenda will be defined [Br09]. To conceptualize the topic (step 2), key concepts and terms with attention to real-world data spaces regarding sustainability were identified. For the literature review conducted (step 3), we used the search string 'data space*' AND 'sustainability' OR 'circular economy' among different scientific databases and received an initial amount of 178 papers. During the literature analysis and synthesis phase (step 4), the papers found were subjected to title and abstract analysis and our exclusion criteria were applied. Moreover, all duplicates, papers that could not be accessed and not peer-reviewed papers were excluded. Furthermore, quality aspects proposed in vom Brocke et al. [Br09] and [Co88] like a consistent methodological approach were applied to the papers to be analyzed. This resulted in a total number of 29 remaining papers. Those were taken into account for an introduction and conclusion analysis, which resulted in ten papers that made it to the final corpus for the full-text analysis. After the full-text analysis, we identified nine relevant papers addressing data spaces in the context of sustainability.

Our *second iteration (empirical)* focused on the investigation of real-world data spaces enabling sustainability. In doing so, we aimed to enrich the results of the first iteration with practical initiatives and insights. Hereby, we identified 22 data spaces from the data space radar published by the International Data Spaces Association (IDSA) [In23]. Moreover, we collected 16 real-world data spaces out of the EuPro Gigant repository [Eu23]. The project aims to create a cross-site, digitally networked manufacturing ecosystem through data spaces [Gi23], [Eu23]. Finally, our *third iteration (empirical)* focused on an Internet research, which allowed us to derive additional 13 data spaces. Following our three iteration-based approach, we examined 60 real-world data spaces from the scientific literature and practice during our iterations.

Based on the systematic literature review and an empirical analysis of real-world data spaces, an overview of data space potentials regarding sustainability was derived. In line with the approach informed by [Ku22], we organized and structured our findings and present them in the form of a map. This map provides a description of the identified dimensions and data spaces potential towards sustainability.

4 Data Spaces for Sustainability

To answer the RQ, we formed dimensions and potentials of data spaces regarding sustainability and conceptualize them in a map. The map itself consists of 15 dimensions, derived from the systematic literature review. The lowest level of the map is formed by 65 specific potentials of data spaces enabling sustainability. Following the recommendations of the literature and to structure the potentials more clearly, we grouped our findings into meta-dimensions. Here, the three pillars of sustainability, ecology, economic and social were chosen as meta-dimensions (see Figure 1).

Ecological	Environment	Food	Plants	Energy
	<ul style="list-style-type: none"> • Air pollution tracking • Water quality optimization • Soil quality tracking 	<ul style="list-style-type: none"> • Improving Food supply chain • Improved access to food • Tracking food quality • Optimizing food quality 	<ul style="list-style-type: none"> • Optimizing plant growth 	<ul style="list-style-type: none"> • Analyze wind power systems • Optimizing water power • Reduce waste of energy • Monitor energy consumption • Optimizing wind power to produce hydrogen
Emission	Animals	Waste		
<ul style="list-style-type: none"> • Emission monitoring • CO₂ optimization • Emission comparison between companies • Reducing CO₂ emissions in production • CO₂ transparency 	<ul style="list-style-type: none"> • Monitoring animal health 	<ul style="list-style-type: none"> • Circular waste • Smart waste 		
Economic	Mobility	Reporting	Supply Chain	Manufacturing
	<ul style="list-style-type: none"> • Seamless mobility • Optimizing traffic routes • Live hazard alters • Traffic flow management • Automated speed monitoring • Connected mobility 	<ul style="list-style-type: none"> • Sustainability ratings • SDG-Monitoring • Credibility and Compliance through transparency • Business Partner Data Management 	<ul style="list-style-type: none"> • Due Diligence in supply chains • Supply chain tracking • Resilient supply chain • Predictive risk management in supply chains • Sharing of sensor data for transparency of ships • Digitalization of fresh produce logistics • Identification of delays in the supply chain 	<ul style="list-style-type: none"> • Life cycle management • Optimization of production • Industrial quality platform for Material • Traceability for circular economy • Marketplace for recycled products • Early warning system for quality defects • Optimization of product quality • Automatic product inspection • Optimization of production sequence • Sharing transactional data • Maintenance forecast • Data sharing in production • Marketplace for manufacturing data • Warehouse automation • Shared production • Intelligent maintenance of machines
Tourism	Finance			
<ul style="list-style-type: none"> • Seamless travel • Analyzing tourism data • Building tourism value chains 	<ul style="list-style-type: none"> • Financial fraud detection 			
Social	Health	Education		
	<ul style="list-style-type: none"> • Individual care plans • Personal healthcare services • Secondary use of health data • Self-determined everyday health • Clinical Accompaniment • Share rare disease data 	<ul style="list-style-type: none"> • Free sharing of education data 		

Fig. 1: Conceptualization of data space potentials enabling sustainability

4.1 Meta-dimension: Ecological sustainability

The first meta-dimension summarizes dimensions relevant to *ecological* potentials of data spaces. As a first observation, it can be stated, that the meta-dimension contains dimensions from various domains such as *emissions*, *environment* and *energy*. Moreover, the specific potentials cover a wide range from *air pollution tracking* to *emission comparison between companies* or the *optimization of water power systems*. In doing so, the most addressed SDGs in the meta-dimension of ecology are *zero hunger* (SDG 2) and *climate action* (SDG 13). Overall, the number of potentials in this meta-dimension account for 32 % of the total potentials observed, which indicates a great interest in this pillar of sustainability. One possible reason for the high level of interest could be the requirements of the legislator, who have passed various laws relating to the ecological sustainability of companies in the EU. For example, the European Corporate Sustainability Due Diligence Directive (CSDDD) [Eu22] can be considered as a leading law, that establishes comprehensive rules for companies operating in the EU to respect human rights and the environment along their operations.

4.2 Meta-dimension: Economic sustainability

In the second meta-dimension, we examined potentials of data spaces regarding sustainability purposes in the context of *economy*. As a first observation, it can be stated that most of the potentials in our sample are concerned with this meta-dimension (58 %). It is reasonable to assume that most potentials are involved here, as companies in our sample believe that sustainability from an economic point of view has the most immediate impact on their day-to-day business. Similar to the *ecological* meta-dimension, a wide range of dimensions and potentials can be observed. For instance, the domains of *mobility* as well as *manufacturing* and *supply chain* are represented with a variety of specific potentials like the *optimization of traffic routes*, *material traceability* or *predictive risk management in supply chains*. Another observation shows that most of the potentials inside the meta-dimension of *economic* are located in the dimension of *manufacturing*. Based on these findings, we anticipate that there is a huge interest to create data spaces for manufacturing purposes in practice. As the majority of the data spaces examined in this study are located in Europe, where the sector of industrial manufacturing is very important, we assume that enough funding can be provided in this area to carry out research and industrial projects in the field of data spaces. Furthermore, there are laws like the German Due Diligence Supply Chain Act [Fe21] that demand a sustainable manufacturing and supply chain, considering various sustainability aspects. The SDGs addressed the most in this meta-dimension are *decent work and economic growth* (SDG 8), *industry, innovation and infrastructure* (SDG 9) and *responsible consumption and production* (SDG 12).

4.3 Meta-dimension: Social sustainability

The third meta-dimension deals with *social* potentials of data spaces in terms of sustainability. Compared to the other meta-dimensions, it can be stated that there are significant fewer potentials (10 %) located in this meta-dimension. We assume that the meta-dimension is smaller compared to the others, as the social sector has fewer financial resources to build and operate data spaces in practice. The meta-dimension itself is divided into the dimensions *health* and *education*. While in the dimension of *health* different potentials like *individual care plans* and the *sharing of rare disease data* are located, there is only the potential of *free sharing of education data* in our sample placed in the dimension of *education*. The SDG mentioned the most in the *social* meta-dimension was *good health and well-being* (SDG 3).

5 Illustration of Selected Data Space Potentials

This section gives a deeper insight into selected potentials from our potentials map and demonstrates their applicability and relevance to sustainability. For this purpose, we choose the *traceability* use case from the German data space Catena-X [Ca23] as well as the *predictive risk management in supply chains* case from the PAIRS project [PA22]. Following, both use cases are described in detail based on public data.

The manufacturing of a vehicle is a complex process, involving multiple steps from production over usage to recycling. In order to coordinate all the processes inside the life-cycle of a vehicle, materials used must be traced. The use case *traceability* aims to trace hardware and software applications ranging from production to recycling in the near future [Ca23]. In doing so, information on each work step are recorded and material as well as part lists are developed. This gives rise to continuous data chains that show when and where which materials, components, or software were used in the value creation process of a vehicle. Through data spaces, those continuous data chains can be shared with other participants inside the value chain, in compliance with rules and their respective data sovereignty. All participants use a common ecosystem, through which they document their hardware and software deployments and exchange information. Thereby, Catena-X also explicitly addresses small and medium-sized enterprises (SMEs), providing compatible, interoperable solutions from an Excel upload via standardized interfaces to open-source and manufacturer-specific traceability solutions [Ca23], [Ca22]. As a result, the consistent *traceability* of materials leads to multiple positive impacts on sustainability. For example, the CO₂ footprint calculation of a product across the entire life-cycle, a simpler recall of materials and a successful CE can be realized through continuous data chains, leading to accomplish the SDGs of *decent work and economic growth* (SDG 8), *industry, innovation and infrastructure* (SDG 9), and *responsible consumption and production* (SDG 12).

The collapse of supply chains has now reached global proportions. Behind a final supply chain are numerous other supply chains. If just one of these chains collapses, the entire

process comes to a standstill. The resulting sensitivity extends from raw material extraction to material procurement and includes all logistical imponderables such as the closure of ports and lack of transport options due to staff shortages [PA22]. To solve this, the use case *predictive risk management in supply chains* is developing a service-oriented, open data infrastructure to predict the emergence and impact of crisis situations to create transparency and initiate precautionary measures for risk minimization. The technology will incorporate both the initial crisis event and the reactions of various actors in a cross-domain data space in order to generate targeted recommendations for action [PA22]. In doing so, they aim to support the SDGs of *decent work and economic growth* (SDG 8), *Industry, innovation and infrastructure* (SDG 9), as well as *responsible consumption and production* (SDG 12).

6 Conclusion and Future Work

In this paper, we have developed a map to collect, group and organize real-world sustainability potentials enabled by data spaces. To the best of our knowledge, there is currently no other study that examines the field of data spaces as an enabler for sustainability-related potentials from both a scientific and practical point of view. Therefore, several implications for science and practice can be derived from our findings.

In terms of *scientific contributions*, our paper advances knowledge of data spaces in the context of sustainable development. Our map presents a big picture of the sustainability landscape enabled through data spaces, which is one of the most important and future-oriented research fields for the industrial value creation and society in the 21st century.

Also, our findings have *implications to practice*. As mentioned, various regulations, such as the Due Diligence Supply Chain Act or CSDDD, require companies to act with regard to the three pillars of sustainability, which requires secure cross-company data sharing [Fe21], [Eu22]. As a result, industry needs knowledge of data space potentials related to sustainability in order to support their strategic decisions towards more sustainable value creation. Our paper contributes to solving these problems by providing a conceptualization of the data space potentials in terms of sustainability. In this way, our findings support companies to implement the regulations of the legislators. In particular, the map presented helps practitioners to gain and classify knowledge in this area.

Naturally, the interpretation of results also includes the discussion of *limitations* that affect this study. Data spaces in general, and particularly in the context of sustainable development, are a rapidly evolving field that remains relatively unexplored, even at this point in time. As a result, our study is a time-limited contribution. Therefore, the data sources and potentials map need to be updated regularly to remain relevant. Furthermore, our findings are based on literature derived from common scientific databases and practical data space initiatives. Despite a strict methodical approach, the selection of the underlying data can be subjective and biased. It is unlikely that the scientific literature and

data spaces analyzed in this paper cover all sustainability potentials where data spaces might develop in the future.

However, the aforementioned limitations lead to possibilities for *future research* directions. As a next step, we want to conduct a series of interviews with experts from science and industry to find out how data spaces can support companies in reaching sustainability goals and act as an enabler of a CE. In doing so, we aim to derive transition patterns for companies moving towards a CE and point out how data spaces can support this process. Based on the transition patterns found, the development of archetypical patterns for data spaces enabling a CE could be a common next step.

Acknowledgement

This work was partially funded by the “Digital ecosystem for a sustainable circular economy in the automotive industry” (DIONA) project. The project DIONA is funded by the German Federal Ministry of Education and Research.






Bibliography

- [Ad23] Advaneo GmbH: Green Deal Dataspace - Resilience & Sustainability. <https://green-deal-dataspace.eu/de/>, accessed: 11/04/2023.
- [AUK18] Antikainen, M.; Uusitalo, T.; Kivikytö-Reponen, P.: Digitalisation as an Enabler of Circular Economy. In (Sakao, T. et al. ed.): Proceedings of the 10th CIRP Conference on Industrial Product-Service Systems (IPS2). Elsevier B.V., Amsterdam, pp. 45–49, 2018.
- [BK87] Brundtland, H.; Khalid, M.: Our Common Future: Report of the World Commission on Environment and Development. Oxford University Press, USA, 1987.
- [Br09] vom Brocke, J. et al.: Reconstructing the Giant - On the Importance of Rigour in Documenting the Literature Search Process. In (Newell, S. et al. ed.): Proceedings of the 17th European Conference on Information Systems (ECIS), Verona, Italy, 2009.
- [Ca22] Catena-X Automotive Network e.V.: Catena-X Automotive Network. <https://catena-x.net/de/>, accessed: 11/04/2023.
- [Ca23] Catena-X Automotive Network Consortium: Traceability as the backbone of Catena-X. <https://catena-x.net/en/mehrwerte/traceability>, accessed: 11/04/2023.
- [Co88] Cooper, H. M.: Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society* 1/1, pp. 104–126, 1988.
- [De15] Deloitte: The more things change: Value creation, value capture, and the Internet of Things. <https://www2.deloitte.com/us/en/insights/deloitte-review/issue-17/value-creation-value-capture-internet-of-things.html>, accessed: 30/03/2023.

- [El98] Elkington, J.: Partnerships from cannibals with forks: The triple bottom line of 21st-century business. *Environmental Quality Management* 8/1, pp. 37–51, 1998.
- [Eu20] European Commission: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European strategy for data. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>, accessed: 30/03/2023.
- [Eu22] European Commission: Corporate sustainability due diligence directive. Fostering sustainability in corporate governance and management systems. https://commission.europa.eu/business-economy-euro/doing-business-eu/corporate-sustainability-due-diligence_en#what-are-the-next-steps, accessed: 07/04/2023.
- [Eu23] EuProGigant: Internationale Datenräume. <https://euprogigant.com/wissens-hub/internationale-datenraeume/>, accessed: 05/04/2023.
- [Fe21] Federal Ministry of Labour and Social Affairs: Act on Corporate Due Diligence Obligations in Supply Chains. <https://www.bmas.de/DE/Service/Gesetze-und-Gesetzesvorhaben/Gesetz-Unternehmerische-Sorgfaltspflichten-Lieferketten/gesetz-unternehmerische-sorgfaltspflichten-lieferketten.html>, accessed: 07/04/2023.
- [FHM05] Franklin, M.; Halevy, A.; Maier, D.: From databases to dataspace. *ACM SIGMOD Record* 4/34, pp. 27–33, 2005.
- [Ga19] Gaia-X European Association for Data and Cloud AISBL: Gaia-X: A Federated Secure Data Infrastructure. <https://gaia-x.eu/>, accessed: 11/04/2023.
- [GGO21] Gelhaar, J.; Groß, T.; Otto, B.: A Taxonomy for Data Ecosystems. In (Bui, T. ed.): Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS), Hawaii, USA, 2021.
- [Gi23] Gieß, A.; Möller, F.; Schoormann, T.; Otto, B.: Design Options for Data Spaces. In: Proceedings of the 31st European Conference on Information Systems (ECIS), Kristiansand, Norway, 2023.
- [GV95] Glass, R. L.; Vessey, I.: Contemporary application-domain taxonomies. *IEEE Software* 4/12, pp. 63–76, 1995.
- [HFM06] Halevy, A.; Franklin, M.; Maier, D.: Principles of dataspace systems. In: Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems. ACM, New York, NY, USA, pp. 1–9, 2006.
- [In23] International Data Spaces Association: IDS Reference Architecture Model 4.0. <https://docs.internationaldataspaces.org/knowledge-base/ids-ram-4.0>, accessed: 11/04/2023.
- [KRH17] Kirchherr, J.; Reike, D.; Hekkert, M.: Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling* 127, pp. 221–232, 2017.
- [Ku22] Kundisch, D. et al.: An Update for Taxonomy Designers. *Business & Information Systems Engineering* 4/64, pp. 421–439, 2022.
- [Mu21] Mulhern, O.: 2020: A Climate Change Perspective. https://earth.org/data_visualization/2020-a-climate-change-perspective/, accessed: 30/03/2023.

- [OB21] Otto, B.; Burmann, A.: Europäische Dateninfrastrukturen. *Informatik Spektrum* 4/44, pp. 283–291, 2021.
- [OL18] Oliveira, M. I. S.; Lóscio, B. F.: What is a data ecosystem? In (Janssen, M. et al. ed.): *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*. ACM, New York, NY, USA, pp. 1–9, 2018.
- [Ot22] Otto, B.: *Designing Data Spaces. The Ecosystem Approach to Competitive Advantage*. Springer International Publishing; Springer, Cham, 2022.
- [PA22] PAIRS-Project: Predictive risk management in supply chains. PAIRS, 2022.
- [PMR19] Purvis, B.; Mao, Y.; Robinson, D.: Three pillars of sustainability: in search of conceptual origins. *Sustainability Science* 3/14, pp. 681–695, 2019.
- [Ri22] Rickmann, H.: Digitalization: Enabler for a sustainable future, <https://www.telekom.com/en/company/management-unplugged/details/digitalization-enabler-for-a-sustainable-future-1010186>, accessed: 30/03/2023.
- [Ro21] Robinson, D.: 14 Biggest Environmental Problems of 2021, <https://earth.org/the-biggest-environmental-problems-of-our-lifetime/>, accessed: 30/03/2023.
- [Ru21] Russo, M.; Young, D.; Feng, T.; Gerard, M.: Sharing Data to Address Our Biggest Societal Challenges. <https://www.bcg.com/publications/2021/data-sharing-will-be- vital-to-societal-changes>, accessed: 11/04/2023.
- [Sc21] Schoormann, T. et al.: Achieving Sustainability with Artificial Intelligence. A Survey of Information Systems Research. In: *Proceedings of the 42nd International Conference on Information Systems (ICIS)*, Austin, USA, 2021.
- [Si13] Singh, M.: A Framework for Data Modeling and Querying Dataspace Systems. In: *Proceedings of the 2013 Conference on Data Mining and Warehousing (ICDMW)*, pp. 17–25, 2013.
- [SJ11] Singh, M.; Jain, S. K.: A Survey on Dataspace: *Proceedings of the Fourth International Conference on Network Security & Applications (CSNA)*, pp. 608–621, 2011.
- [SL17] Seele, P.; Lock, I.: The game-changing potential of digitalization for sustainability: possibilities, perils, and pathways. *Sustainability Science* 2/12, pp. 183–185, 2017.
- [SMS22] Schoormann, T.; Möller, F.; Szopinski, D.: Exploring Purposes of Using Taxonomies. In: *Proceedings of the Wirtschaftsinformatik, Nürnberg, Germany, 2022*.
- [Un15] United Nations: Transforming our world: The 2030 agenda for sustainable development, <https://sdgs.un.org/2030agenda>, accessed: 02/02/2023.
- [Wo15] World Economic Forum: *Intelligent Assets Unlocking the Circular Economy Potential. Industry Agenda 2015*, 2015.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 26, pp. 13–23, 2002.

To Graph or Not to Graph: The Missing Pieces for Knowledge Graphs in Sustainable Tourism


Stefan Neubig ¹, Dominik Rebholz ², Andreas Hein ³, Robert Keller ⁴,
Helmut Krcmar ⁵


Abstract: Sustainability is a critical challenge in modern tourism, exacerbated by climate change and globalization. Thanks to digitization, data-driven approaches constitute a key technology for addressing related issues, such as overtourism. However, the overarching complexity of the touristic data landscape, amplified by the interplay of diverse digital platform ecosystems, poses considerable challenges to both data owners and consumers. To mitigate such issues, knowledge graphs (KGs) have received significant attention. KGs focus on data quality by employing unified data models and continuous data refinements, making them well-suited for data-driven applications. Although promising, many challenges must be addressed to make KGs useful in practice. This paper overviews the state of the art of the field and identifies avenues for future research, explicitly focusing on touristic value and sustainability. Following our results, future research should focus on different areas, notably real-time knowledge graph population, distributed and parallelized processes, and ontologies for dynamic data types.


Keywords: Sustainable Tourism, Knowledge Graphs, Semantic Data Integration


1 Introduction


Sustainability is one of the key challenges of modern tourism in the current century. Battered by the adverse effects of climate change and globalization, overcoming challenges such as overtourism is even more critical today [Be20; Ca19]. During the worldwide Covid-19 pandemic, it became even more apparent that not only famous places such as Venice, Dubrovnik, and Barcelona are suffering from these problems [Ca19; GVT20], but they also increasingly apply to smaller destinations [KHL22]. Thanks to digitization, data-driven approaches are a key technology for tackling overtourism and

¹ Technical University of Munich, Boltzmannstr. 3, 85478 Garching & Outdooractive AG, Missener Str. 18, 87509 Immenstadt, stefan.neubig@tum.com,  <https://orcid.org/0000-0002-3794-7260>

² University of Applied Sciences Kempten, WTZ, Bahnhofstraße 61, 87435 Kempten (Allgäu) & Fraunhofer Institute for Applied Information Technology FIT, Augsburg, dominik.rebholz@hs-kempten.de,  <https://orcid.org/0009-0003-8769-0365>

³ Technical University of Munich, Boltzmannstr. 3, 85478 Garching, andreas.hein@tum.de,  <https://orcid.org/0000-0001-9565-5840>

⁴ University of Applied Sciences Kempten, WTZ, Bahnhofstraße 61, 87435 Kempten (Allgäu) & Fraunhofer Institute for Applied Information Technology FIT, Augsburg, robert.keller@hs-kempten.de,  <https://orcid.org/0000-0001-7097-1724>

⁵ Technical University of Munich, Boltzmannstr. 3, 85478 Garching, helmut.krcmar@tum.de,  <https://orcid.org/0000-0002-2754-8493>

related problems [Gr15; Li17; SR21; UN19]. Examples include the use of sensor technology to capture overtourism-relevant data (e.g., POI occupancy and visitor movement patterns [Bo22]) and the development of smart applications that make use of this data (e.g., raising visitor awareness and sustainable recommendations [Sc22]). Unfortunately, the complex nature of touristic data, resulting from both high connectivity as well as a wide range of stakeholders and intersecting digital platform ecosystems [He20], the creation of such applications is complicated. In practice, relevant data is often unavailable, located in isolated data silos, or suffering from uncertain data quality [SR21]. Consequently, this paper proposes using knowledge graphs (KGs) to mitigate these issues. We argue that KGs can effectively capture the complex nature of data in touristic ecosystems (including complex relationships among different entities), making them well-suited for supporting sustainable tourism by facilitating data integration and managing sustainability-related data.

However, most literature primarily focuses on the initial construction processes of KGs [TG23]. It often neglects crucial aspects of maintenance and lifecycle management (e.g., schema management, continuous data updates), a fundamental prerequisite for the widespread adoption of KG technology. Therefore, this paper presents what we consider the most pressing challenges in this regard. Our analysis draws on our previous work regarding the use of data and related service improvements in tourism [Ne22], our practical experiences with the touristic data landscape, and a systematic review of recent literature. More precisely, we aim to answer the following research question:

What are the key components and relationships in a knowledge graph lifecycle and maintenance model to support the management of touristic data?

We review recent literature to answer this research question and synthesize our findings to describe a conceptual perspective of an overall knowledge graph lifecycle model. By doing so, our paper contributes to both theory and practice. First, we provide an overview of the existing knowledge graph lifecycle and maintenance concepts. Second, we synthesize these concepts into a single process and support future research in positioning itself within the overall KG lifecycle. Third, we analyze the potentials and limitations of recent approaches concerning the heterogeneous touristic data landscape and identify the most promising avenues for future research. Besides, our findings support practitioners by evaluating whether knowledge graphs are a good choice for their use cases.

2 Background

Tourism is a complex domain, and so is its related data landscape. The tourism sector has also been an early adopter of semantic data models and has driven a widespread implementation of semantic web technologies. The following sections briefly overview touristic data types, data-driven use cases, and current data management challenges in tourism. Furthermore, we introduce the concept of knowledge graphs as a natural continuation of semantic web technology.

2.1 Touristic data landscape

In tourism, destinations are typically managed and promoted by destination management organizations (DMOs) that work with different stakeholders (e.g., government, businesses, local communities) and are responsible for brand development as well as the implementation of an appropriate tourism strategy [BLR05]. Recently, DMOs have also been increasingly faced with challenges regarding sustainability. In previous work [Ne22], we interviewed DMO representatives to gain insights into the data they collect, the insights they draw from it, and the related use cases they are pursuing to tackle touristic sustainability challenges. According to the identified use cases, there is a shift from static data to the increasing need for dynamic and real-time data, which is necessary for implementing use cases such as cross-DMO data sharing, or visitor guidance. Implementing these use cases at scale can significantly contribute to the sustainability of the overall ecosystem. For instance, the AIR research project [AIR23] focuses on implementing an AI-based recommender system aiming at a more sustainable distribution of visitors among hot and cold spots. Such efforts, however, strictly rely on the availability of real-time data and the broad participation of various DMOs. With many DMOs as data owners, however, data in the touristic ecosystem is highly fragmented and isolated [SR21]. Considering the German-speaking tourism area as an example, a couple of different approaches exist trying to defragment and integrate such data. Commercial service providers, which most DMOs use for managing large subsets of their data, recently initiated the so-called tourism tech alliance (TTA) [ND23], which aims to share data among these providers. Moreover, publicly funded data hubs, data spaces, and related research projects (e.g., [AIR23; Ba19; DAT23; DZT22]) have emerged to enable a standardized and open data exchange. However, since the overall practical adoption is slow and each approach above suffers from its restrictions, the broad availability and integration of touristic data is still in their infancy [SR21].

Another exacerbating problem that complicates data representation and integration lies in the nature of the data itself. Touristic data is heterogeneous [So16], and entities have highly dynamic relations to each other. For example, political and marketing regions co-exist and overlap. Unclear responsibilities lead to duplicates, such as creating multiple instances of the same point of interest (POI) (e.g., there are multiple instances of the Neuschwanstein castle in the Outdooractive platform [OA23]). Due to the increasing need for real-time data, future data management increasingly needs to cope with fast-arriving and frequently changing data types as well as with larger volumes of data. The next section introduces knowledge graphs as an enabling technology for standardized data exchange, even across the borders of a single DMOs. By utilizing knowledge graphs, data can be made broadly available, which allows the widespread implementation of use cases contributing to the sustainable development of the overall tourism ecosystem.

2.2 Knowledge graphs

Since a blog post by Google in 2012 [Si12], the term knowledge graph has received significant attention from companies and the research community. Many of its underlying ideas are borrowed from the semantic web domain [BL01], which aims at employing semantic technologies to create a web of data in which machines (i.e., intelligent agents [Ki21]) can “understand” the content of interlinked web pages. The emerging semantic web research community has since been developing various standards and technologies for this purpose, commonly referred to as the semantic web technology stack [Wc15]. The most notable concepts and technologies include ontologies [Gu09] (i.e., formal conceptualizations of real-world entities and relations), the resource description framework (RDF) [SR14] (i.e., a graph-based representation of facts in triple-form), the web ontology language [Hi12] (i.e., a family logic-based ontology languages), and the shape constraint language (SHACL) [KK17], which enforces specific structures in graph-based data models. Fig. 1 illustrates an exemplary touristic KG. The conceptual layer (i.e., ontological concepts) describes what general data types (i.e., tour, POI, category) exist and how they may be linked to each other (i.e., contains, has_category), the instance layer describes a possible manifestation of these concepts.

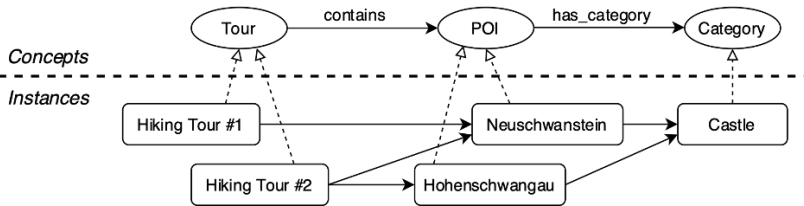


Fig. 1: Exemplary knowledge graph comprising tours, POIs, and POI categories

In a recent review, [Hi21] describes a focus shift within the research community from establishing a “web of data” to knowledge graphs, which represent a natural continuation of the development of the field with some shifts in emphasis, such as higher data quality, reduced openness, and industrial adoption. Considering the large number of related publications, however, the existing literature is surprisingly dividing on what a knowledge graph actually is. [Fä17], for example, define a knowledge graph as “any RDF graph” [Fä17]. In contrast, [VT17] provide a more restrictive definition by adding the requirement that RDF graphs must be linked to terminological knowledge (i.e., types) provided by an ontology. [Hi21] points out that knowledge graphs primarily differ from (possibly ontology-driven) RDF graphs in that they are more self-contained and usually benefit from better internal consistency. Following these definitions, knowledge graphs can be perceived as a new framing and emphasis on old (semantic web) concepts.

However, knowledge graphs have become relevant in various fields and are not limited to a semantic-web-based, technology-specific definition. Thus, [Ho22] provide a more general definition, describing a knowledge graph as “a graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of

interest and whose edges represent potentially different relations between these entities” [Ho22]. The authors explicitly state that their definition is meant to allow different graph models, not limited to RDF. However, [EW16] argue that such a general definition is not really helpful regarding implementing a KG-based system. Instead, while being technology-agnostic, the authors define a knowledge graph as a knowledge-based system that “acquires and integrates information into an ontology and applies a reasoner to derive new knowledge” [EW16]. In line with many of the discussed definitions, this paper considers knowledge graphs as graphs of real-world entities with an attached ontological schema, but without requiring a specific technology (i.e., a trade-off between flexibility and rigor). Similar to [Pa16], we do not expect the presence of a reasoner but require the underlying schema to capture ontological knowledge. Through the lens of sustainable tourism, this constraint is necessary due to the importance of cross-DMO data sharing and the related requirement of well-founded, standardized data models (e.g., [DZT22]).

3 Literature Review

We conducted a systematic literature review based on [WW02] to identify and analyze previous research. Following the interdisciplinary nature of our study, we chose Scopus as our primary literature database, given its comprehensive coverage across technical and domain-specific publications. To ensure capturing a wide range of relevant studies even beyond Scopus, we conducted a rich forward and backward search supported by tools like ResearchRabbit and Scite.ai. Our overall process (Fig. 1) consists of three steps. To retrieve articles that primarily focus on identifying literature on knowledge graph lifecycle and maintenance, we applied the search string “*knowledge graph*” AND (“*lifecycle*” OR “*maintenance*”) to the title, abstract, and keyword fields. We deliberately chose not to use the keyword “tourism” in our search string to avoid limiting our search results to a narrow subset. After screening and pre-filtering the results by their titles and abstracts, we read and analyzed the remaining papers in-depth and complemented this process with a forward and backward search. We included any English article describing at least one potential phase that could be part of a KG lifecycle (e.g., creation, data updates, maintenance tasks). In contrast, we excluded papers that do not match our definition of a knowledge graph.

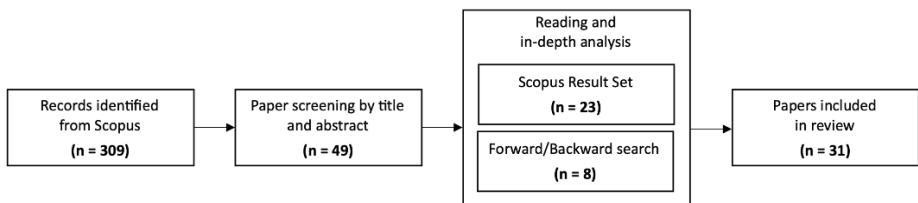


Fig. 2: Literature review methodology, including the number of hits at each step

4 Results

To obtain our overall results, we adhered to the open, axial, and selective coding guidelines presented by [WFW13]. More precisely, we followed an inductive approach to identify a set of key concepts (i.e., characteristics) related to creating or maintaining knowledge graphs. In a subsequent step, we grouped these characteristics to form a set of overarching categories (i.e., dimensions). The first category (data source, D1) involves the diverse data source types knowledge graphs are derived from and comprises structured data (e.g., relational databases), semi-structured data (e.g., documents such as JSON or XML), and unstructured data (e.g., free text or images). The second category (construction, D2) refers to the construction method of a (physical) knowledge graph, which may either follow a top-down or a bottom-up approach. In contrast, the third category (virtualization, D3) refers to knowledge graphs that do not follow any physical construction process but rely on virtualization. The fourth category (refinement, D4) refers to (optional) KG refinement steps (i.e., error detection or graph completion), which were present in most of the identified papers. Furthermore, we introduced a fifth category (tourism-specific, D5) to emphasize approaches developed explicitly for tourism-related use cases. However, only three publications ([Si21; SUF20; Si23]; all authors from the same institute) focused on a touristic use case. Fig. 4 depicts the final dimensions and characteristics.

Dimension	Characteristics		
D1: Data Source	Unstructured (16)	Semi-Structured (19)	Structured (21)
D2: Construction	Top-Down (18)		Bottom-Up (11)
D3: Virtualization	Fully Virtualized (3)		Partly Virtualized (2)
D4: Refinement	Error Detection (13)		Graph Completion (10)
D5: Tourism-Specific	Yes (3)		No (28)

Fig. 3: Identified Dimensions and Characteristics during the coding phase (Brackets indicate the number of representatives of a given characteristic)

4.1 KG lifecycle and maintenance process

Based on this coding, we synthesized different phases and activities to compile a comprehensive knowledge graph lifecycle model (Fig. 4), distinguishing between bottom-up and top-down construction iterations, as well as virtual knowledge graphs (VKGs). The bottom-up construction paradigm does not rely on a previously defined ontology but includes constructing such an ontology as an integral part of its process. In contrast, top-down approaches and VKGs use a previously defined ontology. Top-down construction methods are generally more complex, while VKGs suffer from the limitations of virtualization. Each type comprises various activities representing the fundamental common elements compiled from the analyzed literature. While some of these activities are considered required (i.e., are present in all investigated approaches), activities only present in a subset of the investigated approaches are considered optional (e.g., KG refinement methods). In the following, we describe the individual activities in more detail.

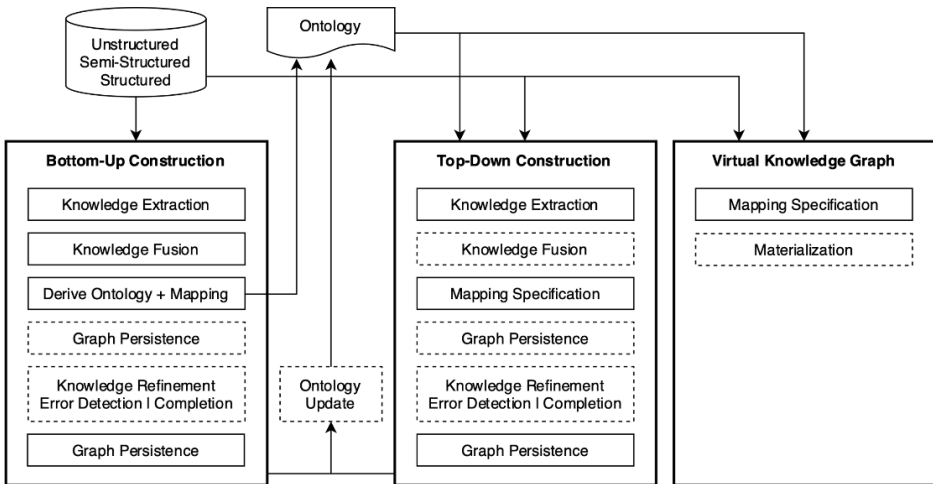


Fig. 4: Synthesized KG lifecycle model considering bottom-up and top-down construction processes, as well as virtual knowledge graphs (Boxes with solid lines refer to required activities, dashed boxes refer to optional activities)

- Knowledge Extraction.* Based on one or multiple data sources, the knowledge extraction activity aims at extracting entities, relations (linking the entities), and attributes [Ch22; La22; TG23; Wa21]. This activity is usually supported by various methods, including named-entity recognition (NER) or different natural language processing (NLP) methods. The reliable extraction of knowledge, especially from unstructured or semi-structured data, is a crucial part of knowledge graph creation and a very active field of research [Si23; TG23]. Note, however, that the extracted knowledge may still contain unconnected entities or isolated subgraphs (e.g., when different data sources are involved).
- Knowledge Fusion.* In its most basic version, knowledge fusion refers to integrating the previously extracted data, thus, dissolving data silos that may have emerged due to using different data sources. Some authors extend this concept by employing error detection and error solving, such as removing duplicates and eliminating ambiguity [La22; TG23; Wa21], while others delay such activities to a dedicated knowledge refinement phase [Ch22; Pa16; SUF20]. Some authors primarily consider the knowledge fusion task for bottom-up approaches (e.g., [Ch22]) and assume that top-down approaches do not necessarily need this task due to the existence of an ontology that may sufficiently define appropriate links, even across data from different sources. The result of the knowledge fusion task is an integrated layer of knowledge without data silos.
- Ontology and Related Mappings.* The bottom-up construction approach does not depend on a previously defined ontology, meaning its creation is integral to the process. Thus, in a bottom-up process, an ontology is created by analyzing the

structure of the fused knowledge of the previous task and by deriving more general concepts [Ch22; TG23]. The ontology creation can be handcrafted or extracted in an automated way. In contrast, top-down construction approaches link the extracted data to the previously defined ontology based on a formal mapping specification [FI22]. In the latter case, the main challenge lies in creating the mapping specification, which usually requires considerable manual effort. The result of this task is a layer of instances connected to an ontological schema.

- *Graph Persistence.* Knowledge graphs can be persisted in a range of different databases, including relational databases (RDBMS), triple stores (i.e., a specific type of an RDBMS), and graph databases [Si23; TG23]. While some authors do not consider the persistence of the graph at this stage (e.g., [Ch22]), other authors base the subsequent refinement task on capabilities specific to an individual (semantic) database system [Si23].
- *Knowledge Refinement.* Once a KG is readily established (persisted or not), it can be further refined and completed using various methods, including KG embeddings and link prediction [Ro21; WQW21; Zh20]. [Pa16] distinguishes approaches for graph completion from those for error detection. Traditionally, knowledge refinement tasks can be performed using logical reasoning or machine learning (ML), but even purely handcrafted approaches exist [Ji20]. Recent approaches [Ch20] leverage representational learning (i.e., learning low-dimensional feature vectors from raw graph data) to overcome scalability issues from a growing graph's increased data sparsity. [Ch20; WQW21] found such approaches to perform significantly better in terms of prediction performance. The result of the KG refinement task is persisted in an appropriate database (see graph persistence).

4.2 Virtual knowledge graphs

Even before the popularity of knowledge graphs, a sub-stream of research evolved focusing on the virtualization of ontologies and RDF graphs. So-called ontology-based data access (OBDA) approaches are often described as a triple $(\mathcal{O}, \mathcal{M}, S)$, where \mathcal{M} is a mapping between a (primarily relational) source database with a schema S and a target ontology \mathcal{O} [Xi18]. Later, with the increasing popularity of knowledge graphs, such systems were frequently referred to as virtual knowledge graphs (VKGs) [Xi19]. One advantage of VKGs is that data resides in its original databases and is always up to date. Although partly materialized approaches exist [SAM14; Vi18], these are usually limited to improving the performance of queries and do not change the overall behavior. Querying VKGs is usually based on query-rewriting techniques, that is, queries directed to the VKG (e.g., SPARQL [HS13]) are translated to queries specific to the source database. Thanks to the introduction of the OWL QL profile, such queries even support basic reasoning (e.g., hierarchical subclass relations) [KEK12]. VKGs, if fully virtualized, do not require any additional storage. On the other hand, VKGs are limited concerning their logical reasoning capabilities and put additional load on the original data sources.

4.3 Updates and maintenance

Updating the knowledge graph is required when the connected data sources change (e.g., new data is emitted, or additional data sources are added) or ontological changes are introduced. For VKGs, updates relate to the underlying mapping specifications and are only necessary if either the ontology or the database schema evolves; integrating new data added to the already mapped databases, however, does not require any action. For physical knowledge graphs, the update task involves performing a new iteration of the overall process [Ch20; Do21; TG23], including the extraction and (optional) fusion of new knowledge, a mapping to the previously defined or updated ontology, and possible employment of additional knowledge graph refinements. In contrast to the initial construction iteration, however, an ontology is guaranteed to exist. Therefore, subsequent iterations usually employ a top-down approach.

5 Discussion and Future Research Directions

As argued in the background section of this paper, a sustainable touristic ecosystem specifically benefits from a wide range of available and reliable data, particularly dynamic and real-time data, such as traffic, weather, or data specific to (local) overcrowding. To reflect such data in knowledge graphs, small and frequent data updates are therefore unavoidable. According to our findings, however, most literature mainly deals with the original creation of a knowledge graph but falls short in providing clear methodological or technical guidelines regarding its subsequent maintenance and data updates. Also, literature specifically addressing the tourism domain is relatively sparse.

In the following paragraphs, we outline what we consider the most pressing research directions focusing knowledge graphs to provide data to build a more sustainable touristic ecosystem.

- *Real-Time Knowledge Graph Population.* Although KGs are dynamic by nature [Ta19], many existing studies are limited to their initial construction phase [TG23]. Moreover, most literature in our study is quite vague about how to perform instance updates to keep a KG current or even requires the repetition of the complete, heavyweight lifecycle for any update, which results in an entire new release of the graph (e.g., [Do21]). While the latter seems reasonable for relatively static data and helps to ensure rigorous quality levels, a more lightweight sub-process (i.e., knowledge graph population) is required to update KG instances in (or close to) real-time based on data streams and frequent updates of data sources (e.g., sensors). Such a process could trade flexibility and consistency guarantees for better performance while being limited to a subset of instance-related operations (e.g., creating or updating instances). Accordingly, rigorous data processing is required upstream to

reliably ensure the quality and consistency of data streams and to avoid quality gaps between the KG population and further graph refinement steps.

- *Distribution and Parallelization.* Depending on the use case, it may be necessary to synchronize KG population processes running in parallel. Current literature often describes the construction of KGs as a rather sequential process, not considering the presence of parallel input streams. However, considering real-world scenarios (e.g., a destination with distributed sensor sets, each driven by different owners), parallel KG population processes could rather be the default than the exception. In even more complex scenarios, preprocessing could not be limited to simple aggregations but could also include more sophisticated ML methods to train different models (e.g., to predict the occupancy of a given area). ML algorithms, however, usually need access to all relevant data to create reliable prediction models, which is likely to raise privacy concerns. As a possible solution, future research could investigate the integration of federated learning [Fe22] into the KG population process to enable the collaborative training of ML models without needing to share raw and confidential data.
- *Touristic Ontology for Dynamic Data.* Due to the rather static nature of past touristic data collections, current data models (e.g., [ABA19; DZT22; FF22]) still need to meet the new challenges of tourism data management. Thus, broadly applicable ontologies are necessary to reflect the nature of dynamic data types, including the online and offline behavior of individual visitors or groups, as well as transport and mobility. Such capabilities should be integrated into existing and accepted models to promote interoperability further and avoid becoming a semantic zoo.

6 Conclusion

This paper reviewed the current state of the art regarding creating and maintaining knowledge graphs. We focused on their use in sustainable touristic ecosystems and analyzed them in the context of the touristic data landscape. We identified the most urgent directions for future research to increase the practical relevance of KGs, including (i) a quality-preserving and scalable knowledge graph population, (ii) distributed data processing, and (iii) the enhancement of existing tourism ontologies. Like other work, our study suffers from certain limitations. First, our investigation is limited to a technical perspective of KG lifecycles, omitting related non-technical challenges. Second, we perceive touristic KGs as graphs with a pre-built schema (i.e., an ontology) and excluded KGs that do not satisfy this definition from our study. Third, the primary sources of literature review are limited to publications available in Scopus, which, while having a broad coverage of multidisciplinary literature, may miss out relevant publications. Fourth, due to its nature, the selection of papers is subjective to a certain degree. Regarding future work, we plan to further strengthen our research by incorporating additional literature and domain experts as well as by widening the scope of our review by further exploring additional needs of touristic data management. Furthermore, we will address the identified

research gaps by focusing on advanced knowledge graph population processes, their synchronization, and the use of federated learning in combination with KG refinement approaches based on graph neural networks (GNNs) [KW17]. Finally, we aim to evaluate these processes based on real-world use cases addressing current touristic challenges such as overtourism and active visitor management.

Acknowledgements

This work is a result of a cooperation between Outdooractive AG, the Technical University of Munich, and the University of Applied Sciences Kempten. This study was supported by the AIR research project (67KI21005B), funded by the federal ministry for the environment, nature conservation, nuclear safety, and consumer protection.

Bibliography

- [AIR23] AIR – AI-basierter Recommender für nachhaltigen Tourismus – Ein Verbundprojekt zur Entwicklung eines Digitalen Besuchermanagements in deutschen Tourismusdestinationen, <https://air-tourism.de/>, accessed 2023-03-26.
- [ABA19] AlpineBits Alliance: AlpineBits. Open Data Exchange in the Alpine tourism, <https://www.alpinebits.org>, accessed 2019-08-30.
- [BL01] Berners-Lee, T.; Hendler, J.; Lassila, O.: The Semantic Web. *Scientific American* 284, Scientific American Inc., No. 5, pp. 34–43, 2001.
- [Ba19] Baecker, J.; Engert, M.; Müller, J.; Neubig, S.; Schaffer, N.; Kunz, I.; Schneider, A.; Krcmar, H.; et al.: *BayernCloud Gesamtreferenzarchitektur*, 2019.
- [Be20] Benner, M.: Overcoming overtourism in Europe: Towards an institutional-behavioral research agenda. *Zeitschrift für Wirtschaftsgeographie* 64, pp. 74–87, 2020.
- [BLR05] Blain, C.; Levy, S. E.; Ritchie, J. R. B.: *Destination Branding: Insights and Practices From Destination Management Organizations*. *Journal of Travel Research*, 2005.
- [Bo22] Bollenbach, J.; Neubig, S.; Hein, A.; Keller, R.; Krcmar, H.: *Using Machine Learning to Predict POI Occupancy to Reduce Overcrowding*, GI, Bonn, 2022.
- [Ch22] Chen, Q.; Li, Q.; Wu, J.; Mao, C.; Peng, G.; Wang, D.: Application of knowledge graph in power system fault diagnosis and disposal: A critical review and perspectives. *Frontiers in Energy Research* 10, 2022.
- [Ca19] Capocchi, A.; Vallone, C.; Pierotti, M.; Amaduzzi, A.: *Overtourism: A Literature Review to Assess Implications and Future Perspectives*. *Sustainability* 11, 2019.
- [Ch20] Chen, Z.; Wang, Y.; Zhao, B.; Cheng, J.; Zhao, X.; Duan, Z.: *Knowledge Graph Completion: A Review*. *IEEE Access* 8, pp. 192435–192456, 2020.

- [DAT23] DATES project Tourism Data Space European Tourism Data Space, <https://www.tourismdataspace-csa.eu/>, accessed 2023-03-26.
- [DZT22] Deutsche Zentrale für Tourismus e.V.: Domain Specifications Browser, <https://open-data-germany.org/domain-specifications-browser/>, accessed 2022-10-03.
- [Do21] Dojchinovski, M.; Forberg, J.; Frey, J.; Hofer, M.; Streitmatter, D.; Yankov, K.: The DBpedia Technology Tutorial 3064, 2021.
- [EW16] Ehrlinger, L.; Wöß, W.: Towards a Definition of Knowledge Graphs, p. 4, 2016.
- [Fäl17] Färber, M.; Bartscherer, F.; Menne, C.; Rettinger, A.: Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. In: Zaveri, A.; Kontokostas, D.; Hellmann, S.; Umbrich, J.; Zaveri, A.; Kontokostas, D.; Hellmann, S.; Umbrich, J. (eds.) *Semantic Web 9*, pp. 77–129, 2017.
- [FF22] FIWARE Foundation: Smart Data Models, <https://www.fiware.org/smart-data-models/>, accessed 2022-10-03.
- [Fl22] Flores, J.; Jamin, E.; Nadal, S.; Romero, O.: The Knowledge Graph Lifecycle in NTT DATA. In: *Proceedings of the ISWC, 3254*, 2022.
- [Fe22] Fernández, J. D.; Menci, S. P.; Lee, C. M.; Rieger, A.; Fridgen, G.: Privacy-preserving federated learning for residential short-term load forecasting. *Applied Energy* 326, p. 119915, 2022.
- [GVT20] Gowreesunkar, V. G.; Vo Thanh, T.: Between Overtourism and Under-Tourism: Impacts, Implications, and Probable Solutions. In: Séraphin, H.; Gladkikh, T.; Vo Thanh, T. (eds.): *Overtourism*. Springer International Publishing, pp. 45–68, 2020.
- [Gu09] Guarino, N.; Oberle, D.; Staab, S.: What Is an Ontology? In: Staab, S.; Studer, R. (eds.): *Handbook on Ontologies*. Berlin, Heidelberg: Springer, pp. 1–17, 2009.
- [Gr15] Gretzel, U.; Werthner, H.; Koo, C.; Lamsfus, C.: Conceptual foundations for understanding smart tourism ecosystems. *Computers in Human Behavior* 50, pp. 558–563, 2015.
- [HS13] Harris, S.; Seaborne, A.: *SPARQL 1.1 Query Language*, 2013.
- [Ho22] Hogan, A.; Blomqvist, E.; Cochez, M.; D’amato, C.; Melo, G. D.; Gutierrez, C.; Kirrane, S.; Gayo, J. E. L.; et al.: *Knowledge Graphs*. *ACM Computing Surveys* 54, pp. 1–37. 2022.
- [Hi21] Hitzler, P.: A review of the semantic web field. *Communications of the ACM* 64, pp. 76–83, 2021.
- [Hi12] Hitzler, P.; Krötzsch, M.; Parsia, B.; Patel-Schneider, P. F.; Rudolph, S.: *OWL 2 Web Ontology Language Primer (Second Edition): W3C-Recommendation*, 2012.
- [He20] Hein, A.; Schreieck, M.; Riasanow, T.; Setzke, D. S.; Wiesche, M.; Böhm, M.; Krcmar, H.: Digital platform ecosystems. *Electronic Markets* 30, pp. 87–98, 2020.
- [Ji20] Jia, Y.; Liu, D.; Sheng, Z.; Feng, L.; Liu, Y.; Guo, S.: EasyKG: An end-to-end knowledge graph construction system. *Communications in Computer and Information Science* 1157, 2020.

- [Ki21] Kirrane, S.: Intelligent software web agents: A gap analysis. *Journal of Web Semantics* 71, p. 100659, 2021.
- [KW17] Kipf, T. N.; Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks, arXiv, 2017.
- [KK17] Knublauch, H.; Kontokostas, D.: Shapes Constraint Language (SHACL), <https://www.w3.org/TR/shacl/>, accessed 2022-10-03.
- [KHL22] Krajčickova, A.; Hampl, F.; Lancosova, E.: Visitors' perception of overtourism impacts in a small destination. *Anatolia* 33, pp. 236–246, 2022.
- [KEK12] Krötzsch, M.: OWL 2 Profiles: An Introduction to Lightweight Ontology Languages. In: Eiter, T.; Krennwallner, T. (eds.): *Reasoning Web. Semantic Technologies for Advanced Query Answering: 8th International Summer School*, Vienna, Austria. Berlin, Heidelberg: Springer, pp. 112–183, 2012.
- [Li17] Li, Y.; Hu, C.; Huang, C.; Duan, L.: The concept of smart tourism in the context of tourism information services. *Tourism Management* 58, pp. 293–300, 2017.
- [La22] Lai, F.; Zhao, Z.; Gao, D.; Luo, W.; Lou, C.; Han, S.; Ma, C.: Using Semantic Dependencies to Realize the Construction of Cloud Data Center Operation and Maintenance Knowledge Graph. In: *Prognostics and Health Management Conference*, pp. 236–240, 2022.
- [Ne22] Neubig, S.; Baecker, J.; Gehring, N.; Hein, A.; Weking, J.; Krcmar, H.: Data-driven Initiatives of Destinations Supporting Sustainable Tourism. In: *Proceedings of the Americas Conference on Information Systems*, 2022.
- [ND23] neusta destination.one: Tourism Tech Alliance, www.tourismtechalliance.org/, accessed 2023-05-19.
- [OA23] Outdooractive, <https://www.outdooractive.com>, accessed 2023-05-20.
- [Pa16] Paulheim, H.: Knowledge graph refinement: A survey of approaches and evaluation methods. In: Cimiano, P. (ed.) *Semantic Web* 8, pp. 489–508, 2016.
- [Ro21] Rossi, A.; Barbosa, D.; Firmani, D.; Martinata, A.; Merialdo, P.: Knowledge Graph Embedding for Link Prediction: A Comparative Analysis. *ACM Transactions on Knowledge Discovery from Data* 15, 2021.
- [Si21] Şimşek, U.; Angele, K.; Kärle, E.; Opdenplatz, J.; Sommer, D.; Umbrich, J.; Fensel, D.: Knowledge graph lifecycle: Building and maintaining knowledge graphs. In: *Second International Workshop on Knowledge Graph Construction*, 2021.
- [SR14] Schreiber, G.; Raimond, Y.: *RDF 1.1 Primer*, 2014.
- [So16] Soualah-Alila, F.; Coustaty, M.; Rempulski, N.; Doucet, A.: DataTourism: Designing an Architecture to Process Tourism Data. In: Inversini, A.; Schegg, R. (eds.): *Information and Communication Technologies in Tourism 2016*. Cham : Springer International Publishing, pp. 751–763, 2016.
- [SR21] Schmücker, D.; Reif, J.: The Big Data Illusion. *Zeitschrift für Tourismuswissenschaft* 13, pp. 157–166, 2021.

- [SAM14] Sequeda, J. F.; Arenas, M.; Miranker, D. P.: OBDA: Query Rewriting or Materialization? In Practice, Both! In: Mika, P.; Tudorache, T.; Bernstein, A.; Welty, C.; Knoblock, C.; Vrandečić, D.; Groth, P.; Noy, N.; et al. (eds.): *The Semantic Web 8796*. Cham: Springer International Publishing, pp. 535–551, 2014.
- [Si12] Singhal, A.: Introducing the Knowledge Graph: things, not strings, <https://blog.google/products/search/introducing-knowledge-graph-things-not/>, accessed 2022-10-02.
- [SUF20] Şimşek, U.; Umbrich, J.; Fensel, D.: Towards a knowledge graph lifecycle: A pipeline for the population of a commercial Knowledge Graph 2535, 2020.
- [Si23] Şimşek, U.; Kärle, E.; Angele, K.; Huaman, E.; Opdenplatz, J.; Sommer, D.; Umbrich, J.; Fensel, D.: A Knowledge Graph Perspective on Knowledge Engineering. *SN Computer Science* 4, 2023.
- [Sc22] Schmücker, D.; Keller, R.; Reif, J.; Schubert, J.; Sommer, G.: *Digitales Besuchermanagement im Tourismus – Konzeptioneller Rahmen und Gestaltungsmöglichkeiten*, 2022.
- [TG23] Tamašauskaitė, G.; Groth, P.: Defining a Knowledge Graph Development Process Through a Systematic Review. *ACM Transactions on Software Engineering and Methodology* 32, pp. 1–40, 2023.
- [Ta19] Tasnim, M.; Collarana, D.; Graux, D.; Orlandi, F.; Vidal, M.-E.: Summarizing Entity Temporal Evolution in Knowledge Graphs. In: *Companion Proceedings of the 2019 World Wide Web Conference*. San Francisco USA: ACM, pp. 961–965, 2019.
- [UN19] UNWTO: *Overtourism? New UNWTO report offers case studies to tackle challenges*. Madrid, Spain, 2019.
- [VT17] Villazon-Terrazas, B.; Garcia-Santa, N.; Ren, Y.; Faraotti, A.; Wu, H.; Zhao, Y.; Vetere, G.; Pan, J. Z.: Knowledge Graph Foundations. In: Pan, J. Z.; Vetere, G.; Gomez-Perez, J. M.; Wu, H. (eds.): *Exploiting Linked Data and Knowledge Graphs in Large Organisations*. Cham: Springer International Publishing, pp. 17–55, 2017.
- [Vi18] Vidal, V.; Magalhães Pequeno, V.; Moura Arruda Júnior, N.; Casanova, M. A.: Publication and maintenance of RDB2RDF views externally materialized in enterprise knowledge graphs. *International Journal of Web Information Systems* 18, pp. 255–285, 2022.
- [WQW21] Wang, M.; Qiu, L.; Wang, X.: A Survey on Knowledge Graph Embeddings for Link Prediction. *Symmetry* 13, p. 485, 2021.
- [Wc15] W3C: *Semantic Web*, <https://www.w3.org/standards/semanticweb/>, accessed 2022-10-02.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 26, pp. xiii–xxiii 2002.
- [WFW13] Wolfswinkel, J. F.; Furtmueller, E.; Wilderom, C. P. M.: Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems* 22, Taylor & Francis, pp. 45–55, 2013.

- [Wa21] Wang, J.; Wang, X.; Ma, C.; Kou, L.: A survey on the development status and application prospects of knowledge graph in smart grids. In: *IET Generation, Transmission and Distribution* 15/3, pp. 383–407, 2021.
- [Xi18] Xiao, G.; Calvanese, D.; Kontchakov, R.; Lembo, D.; Poggi, A.; Rosati, R.; Zakharyashev, M.: *Ontology-Based Data Access: A Survey*, 2018.
- [Xi19] Xiao, G.; Ding, L.; Cogrel, B.; Calvanese, D.: Virtual Knowledge Graphs: An Overview of Systems and Use Cases. *Data Intelligence* 1, pp. 201–223, 2019.
- [Zh20] Zhang, Z.; Cai, J.; Zhang, Y.; Wang, J.: Learning Hierarchy-Aware Knowledge Graph Embeddings for Link Prediction. In: *Proceedings of the AAAI Conference on Artificial Intelligence* vol. 34, pp. 3065–3072, 2020.

Promoting Sustainable Ecosystems through Interorganizational Unlearning – A Call for Research

Marco Di Maria¹, David Walter¹ and Ralf Knackstedt¹

Abstract: In response to changes within and outside their boundaries, ecosystems need to constantly adapt. Thus, interorganizational learning becomes a vital process to facilitate this adaptation. To support interorganizational learning, ecosystems can apply unlearning. As part of overall learning, it refers to the process of deliberately abandoning obsolete beliefs or routines. However, it remains an open issue how unlearning unfolds on the interorganizational level. To address this issue, we provide an initial definition, five research directions, and a research project plan. Thereby, we hope to provide useful stepping stones for researchers of ecosystems, sustainability, and unlearning. For ecosystem practitioners, our goal is to encourage sustainable knowledge management through the design of support tools for interorganizational unlearning.

Keywords: Ecosystems, Interorganizational Unlearning, Sustainability, Knowledge Management.

1 Introduction

As ecosystems become larger and more complex, the efforts for coordinating knowledge-related processes rise [OBF19]. In consequence, continuous management of 'interfirm knowledge' [Ar09] emerges as a key challenge for ecosystems [Zu15], [FH07]. In doing so, ecosystems can create sustained competitive advantages [Ta04] including collaborative know-how useful for the entire ecosystem [Ki16], [Ho03]. To stay up-to-date, they must repeatedly renew their shared knowledge [KTY20]. To achieve that, unlearning can be applied [MN17]. Unlearning refers to deliberately abandoning existing beliefs or behaviors that jeopardize the success of an organization [dHP04], such as overconsumption. Some research already addresses interorganizational unlearning (IU). For instance, [VN22] highlight IU's importance for ecosystems to effectively address ecological burdens. [NJN17] found that IU can support fostering urban biodiversity. Further, [AK16] examined overcoming the use of outdated technology in innovation ecosystems. However, it remains unclear how unlearning operates on the interorganizational level. Concepts and tools to understand and facilitate IU [Be19], [Di23] are still missing. To address this gap, we call for research on IU, starting with ecosystems. Ecosystems offer promising grounds for studying IU, with shared knowledge domains among members and the accumulation of interorganizational knowledge [An21]. This dynamic provides opportunities for learning and unlearning over time, including 'competence traps' [LM88] resulting from experiential learning.

¹ University of Hildesheim, Information Systems, Universitätsplatz 1, 31141 Hildesheim,
marco.dimaria, walter002, knacks}@uni-hildesheim.de

2 Towards Understanding of Inter-Organizational Unlearning

IU applies to diverse interorganizational constructs, including strategic alliances [Ha91], supply chains [Ho85], clusters [Ar09], networks [Be05], and ecosystems [Ad17]. Each construct presents unique features and dynamics relevant to IU. In *strategic alliances*, partners align their activities towards a shared goal, mostly through contractual agreements. Due to the often relatively small number of partners, the scope for exploring phenomena of IU may be limited. Besides alliances, *clusters* are characterized by geographic concentrations of a large number of related organizations within specific industries. However, these organizations are often loosely coupled, with limited cross-integration, e.g., resources. As a result, there is a comparatively reduced potential for exploring IU. In *supply chains*, diverse partners participate in different stages of the value creation process. The primary emphasis on efficiency often leads to limited engagement between organizations across stages, posing a challenge for IU at the overall supply chain level. In contrast to that, *networks* are structures characterized by organizations interconnected across industries, regions, or domains. While they promote learning and knowledge sharing, the focus on IU may be relatively less pronounced compared to constructs with more homogeneous partner sets. Besides that, *ecosystems* are constructs that bring together different organizations in close collaboration, aligning their resources and strategies towards shared goals. The ecosystem's success often depends on shared IT infrastructure. The comprehensive integration within ecosystems creates an ideal environment for exploring the challenges and mechanisms associated with IU [An21] since they are characterized by complex, enduring relationships among multiple heterogeneous partners that concentrate on a shared product, data, or technological platform. As such, they can broadly be described as a “group of interacting firms that depend on each other’s activities.” [JCG18, p. 2256]. In ‘business ecosystems’ [Te07], organizations form a community whose performance affects each member. In ‘innovation ecosystems’ [AK10], organizations collaborate towards a common goal, often creating value in the form of a shared product. In ‘platform ecosystems’ [Ce12], peripheral organizations orbit around a central, leading player that provides a technological platform acting as the main integrator streamlining value creation within the ecosystem. ‘Data ecosystems’ represent another type wherein different partners take multiple roles and form complex relationships to facilitate data-centric processes [Az20]. Thus, partners must manage key resources, particularly knowledge, to ensure the ecosystem’s sustainability [OL18].

To create value for all member organizations, an ecosystem has to learn. As part of the overall learning process [Az08], *unlearning* refers to the process of intentionally questioning and discarding existing knowledge that is no longer relevant [TZ08], as a prerequisite for acquiring new knowledge [ZLW13], [CL90]. For instance, unlearning can support abandoning hindering routines that no longer serve organizational needs. In organizational unlearning, three *types of knowledge* are often differentiated, e.g. [EL11]: cognitive, behavioral, social. *Cognitive* knowledge refers to values, beliefs, assumptions, mental models, and frames of reference [Ak07]. In ecosystems, this can be seen as

interorganizational knowledge [ALP21] shared by the ecosystem's members, including common history and shared values. Besides that, *behavioral* knowledge refers to practices and routines [FO17]. For instance, this can be shared knowledge in-action, such as jointly developing a technology or other forms of implicit knowledge [No94]. Further, *social* knowledge refers to social networks, interactions, and joint sensemaking [EL11]. Thus, an ecosystem writes its own history as the sum of shared experiences. Further, *types of unlearning* [KG19] are used to analyze unlearning. Technical unlearning refers to discarding explicit, technical knowledge, such as operative procedures, in contrast to adaptive unlearning of implicit knowledge [No94], for instance, deeply rooted assumptions. Further, [dHP04] differentiate between directed unlearning of existing knowledge, i.e., 'purging', and avoiding to absorb new knowledge, i.e., 'suspension'. More, knowledge resides on different *levels* [KG19]: individual [Hi14], group/team [Ma19], organizational [Ak14]. [No94, p. 20] adds an "interorganizational level" on top. Similarly, [Ho03, p. 101] views knowledge at this level as "sets of interorganizational experiential rules". Conceptualizing IU in the area of ecosystems and linking it to interorganizational learning, we propose the following *definition* for IU: IU refers to the process through which a collective of organizations intentionally identifies, questions and discards existing knowledge that hampers the collective's ability to adapt to changing conditions and achieve agreed upon goals.

3 Deriving Research Directions for Interorganizational Unlearning

Arguing that ecosystems may benefit from a better understanding of IU, we hereby suggest a non-exhaustive set of five research directions [RD1-5] as an extension to [SL22]'s clusters for future research on unlearning currently lacking aspects of IU.

- [RD1] It is unclear which *purposes* drive ecosystems to apply unlearning. However, the literature on organizational unlearning suggests some purposes, including crisis management [HNS76], organizational change [K189], knowledge transfer support [Ts08], and problem-solving [Br16]. Yet, it remains unclear how these purposes, or the intentional absence of them [Pe19], guide IU in ecosystems. Therefore, additional guidance for ecosystem practitioners would be beneficial to properly apply IU.
- [RD2] The impact of different *knowledge types* on IU is still unknown. Exploring cognitive, behavioral, and social knowledge [EL11], as well as the distinctions between explicit and implicit knowledge [No94], could be insightful. Additionally, investigating technical and adaptive unlearning [KG19] and the effects of purging old knowledge, suspending new knowledge [dHP04], and proactive avoidance of absorption [MN17] on IU is important. Understanding the interplay between IU, absorptive capacity, and varying knowledge similarity among member organizations [CL90] is interesting for further research.

- [RD3] The relationship between different *types of ecosystems* [Ad17] and IU remains unclear. Variations among ecosystem members [Wa17] may impact IU. Exploring phenomena such as 'systems-of-ecosystems' [ELT08], participating in multiple ecosystems with diverse roles [BB15], ecosystem age [Ki16], and the evolutionary patterns of IU in ecosystems (episodic or continuous) [TZ08] would be intriguing. Additionally, research on the influence of co-competition [FBK19] could provide valuable insights. Addressing these issues would assist ecosystem practitioners in decision-making and navigating complex trade-offs in managing knowledge in ecosystems via IU.
- [RD4] It is not clear how concepts of *sustainability* [VW19] intersect with IU. Researchers would benefit from insights regarding *how* ecosystems apply IU 'to sustain' themselves (as a process) and *what* goals they pursue 'for sustainability' (as outcomes). They could explore the influence of IU on, for example, sustainable development processes [RJB12] or the United Nation's 17 Sustainable Development Goals [UN23] as desired outcomes as some researchers have already begun, for example [Pa18], [Sc23], [CN13].
- [RD5] It remains unclear how to *practically support* IU [Di23]. Further research is needed to understand requirements for IU, the types of artifacts involved, and how to measure their effectiveness. Practitioners would benefit from the design of effective digital technologies as vital component [Ce12] to support IU in ecosystems.

4 Planned Study and Expected Overall Contribution

To start with [RD5], we propose a *three-part research project* on IU in ecosystems applying the Design Science Research (DSR) paradigm [He07]. In *Part 1* of our study, we conduct a qualitative review [Sc15] on IU to synthesize existing knowledge. We build upon models, frameworks and concepts grounded in previous works on unlearning and ecosystems. The outcome of this review guides further research on IU in ecosystems. Moving to *Part 2*, we use the insights from Part 1 to conduct expert interviews [FF00] to refine our understanding of IU. Then, we perform a Delphi-Study [OP04] to validate the findings from the interviews. This comprehensive approach allows us to integrate academic and practical viewpoints and establish a well-rounded understanding of IU. With this foundation, we formulate a research agenda on IU in ecosystems with focus on design-related issues. *Part 3* centers around practical support for IU by iteratively designing and testing digital prototypes in a case study [Yi18] with industry partners who seek to promote sustainability in their ecosystem.

Overall, we *contribute* in multiple ways to the nascent research on IU. We provided a definition for IU based on concepts from unlearning, sustainability, and ecosystems. Reflecting on conceptual gaps, we pointed out five research directions in response to calls from academia, e.g., [Be19], [KTY20]. To catalyze research on IU, we presented a

plan for a DSR study to synthesize existing knowledge, formulate a research agenda, and create design knowledge to support IU. So, we promote sustainable ecosystems via IU.

Acknowledgments

We thank the European Social Fund (ESF+) and the State of Lower Saxony (NBank) for partly funding this research in the context of the research project “ProXHybrid” (ZAM 3-87002690).

Bibliography

- [Ad17] Adner, R.: Ecosystem as structure: An actionable construct for strategy. *Journal of Management* 43/1, pp. 39-58, 2017.
- [Ak07] Akgün, A. E. et.al.: Organizational unlearning as changes in beliefs and routines in organizations. *Journal of Organizational Change Management* 20/6, pp. 794-812, 2007.
- [AK10] Adner, R.; Kapoor, R.: Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic Management Journal* 31/3, pp. 306-333, 2010.
- [Ak14] Akhshik, S. S.: Organizational unlearning: an analytical literature review. *International Journal of Economics & Management Science*, 3/3, pp. 1-7, 2014.
- [AK16] Adner, R.; Kapoor, R.: Innovation ecosystems and the pace of substitution: Re-examining technology S-curves. *Strategic Management Journal* 37/4, pp. 625-648, 2016.
- [ALP21] Argote, L. et.al.: Organizational learning processes and outcomes: Major findings and future research directions. *Management Science*, 67/9, pp. 5399-5429, 2021.
- [An21] Anand, A. et.al.: Interorganizational learning: a bibliometric review and research agenda. *The Learning Organization* 28/2, pp. 111-136, 2021.
- [Ar09] Arikán, A. T.: Interfirm knowledge exchanges and the knowledge creation capability of clusters. *Academy of Management Review* 34/4, pp. 658-676, 2009.
- [Az08] Tabassum Azmi, F.: Mapping the learn-unlearn-relearn model: imperatives for strategic management. *European Business Review* 20/3, pp. 240-259, 2008.
- [Az20] Azkan, C. et.al.: Service Dominant Logic Perspective on Data Ecosystems - A Case Study based Morphology. In: *Proceedings of the 28th European Conference on Information Systems*, virtual, 2020.
- [BB15] Bosch-Sijtsema, P. M.; Bosch, J.: Plays nice with others? Multiple ecosystems, various roles and divergent engagement models. *Technology Analysis & Strategic Management* 27/8, pp. 960-974, 2015.


- [Be05] Bell, G. G.: Clusters, networks, and firm innovativeness. *Strategic Management Journal* 26/3, pp. 287-295, 2005.
- [Be19] Becker, K.: Organizational unlearning: the challenges of a developing phenomenon. *The Learning Organization* 26/5, pp. 534-541, 2019.
- [Br16] Brook, C. et.al.: On stopping doing those things that are not getting us to where we want to be: Unlearning, wicked problems and critical action learning. *Human Relations* 69/2, pp. 369-389, 2016.
- [Ce12] Ceccagnoli, M. et.al.: Cocreation of value in a platform ecosystem: The case of enterprise software. *MIS Quarterly* 36/1, pp. 263-290, 2012.
- [CL90] Cohen, W. M.; Levinthal, D. A.: Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35/1, pp. 128-152, 1990.
- [CN13] Cegarra-Navarro, J.G. et.al.: Environmental knowledge, unlearning, and performance in hospitality companies. *Management Decision* 51/2, pp. 341-360, 2013.
- [dHP04] de Holan, P. M.; Phillips, N.: Organizational forgetting as strategy. *Strategic Organization* 2/4, pp. 423-433, 2004.
- [Di23] Di Maria, M. et.al.: Practical Support for Unlearning – A Systematic Review to Organize the Field. In: *Proceedings of the 31th European Conference on Information Systems*, Kristiansand, Norway, 11, 2023.
- [EL11] Easterby-Smith, M.; Lyles, M. A.: In praise of organizational forgetting. *Journal of Management Inquiry* 20/3, pp. 311-316, 2011.
- [ELT08] Easterby-Smith, M.; Lyles, M. A.; Tsang, E. W.: Inter-organizational knowledge transfer: Current themes and future prospects. *Journal of Management Studies* 45/4, pp. 677-690, 2008.
- [FBK19] Fredrich, V.; Bouncken, R. B.; Kraus, S.: The race is on: Configurations of absorptive capacity, interdependence and slack resources for interorganizational learning in coopetition alliances. *Journal of Business Research* 101, pp. 862-868, 2019.
- [FF00] Fontana, A.; Frey, J. H. The interview: From structured questions to negotiated text. In (Denzin, N. K., Lincoln, Y. S., eds.): *Handbook of qualitative research*, 2nd ed., pp. 645-672. Thousand Oaks, CA: Sage, 2000.
- [FH07] Felin, T; Hesterly, W. S.: The knowledge-based view, nested heterogeneity, and new value creation: Philosophical considerations on the locus of knowledge. *Academy of Management Review* 32/1, pp. 195-218, 2007.
- [FO17] Fiol, C. M.; O'Connor, E. J.: Unlearning established organizational routines–Part II. *The Learning Organization* 24/2, pp. 82-92, 2017.
- [Ha91] Hamel, G.: Competition for competence and interpartner learning within international strategic alliances. *Strategic Management Journal* 12/S1, pp. 83-103, 1991.
- [He07] Hevner, A. R.: A three cycle view of design science research. *Scandinavian Journal of Information Systems* 19/2, pp. 87-92, 2007.
- [Hi14] Hislop, D. et.al.: The process of individual unlearning: A neglected topic in an under-researched field. *Management Learning* 45/5, pp. 540-560, 2014.

- [HNS76] Hedberg, B. L.; Nystrom, P. C.; Starbuck, W. H.: Camping on seesaws: Prescriptions for a self-designing organization. *Administrative Science Quarterly* 21, pp. 41-65, 1976.
- [Ho03] Holmqvist, M.: A dynamic model of intra- and interorganizational learning. *Organization Studies* 24/1, pp. 95-123, 2003.
- [Ho85] Houlihan, J. B.: International supply chain management. *International Journal of Physical Distribution & Materials Management* 15/1, pp. 22-38, 1985.
- [JCG18] Jacobides, M. G.; Cennamo, C.; Gawer, A.: Towards a theory of ecosystems. *Strategic Management Journal* 39/8, pp. 2255-2276, 2018.
- [KG19] Klammer, A.; Gueldenberg, S.: Unlearning and forgetting in organizations: a systematic review of literature. *Journal of Knowledge Management* 23/5, pp. 860-888, 2019.
- [Ki16] Kilubi, I.: Strategic technology partnering capabilities: A systematic review of the empirical evidence over two decades. *Journal of Strategy and Management* 9/2, pp. 216-255, 2016.
- [Kl89] Klein, J. I.: Parentetic learning in organizations: Toward the unlearning of the unlearning model. *Journal of Management Studies* 26(3), pp. 291-308, 1989.
- [KTY20] Kano, L.; Tsang, E. W.; Yeung, H. W. C.: Global value chains: A review of the multi-disciplinary literature. *Journal of International Business Studies* 51, pp. 577-622, 2020.
- [LM88] Levitt, B.; March, J. G.: Organizational learning. *Annual Review of Sociology* 14/1, pp. 319-338, 1988.
- [Ma19] Matsuo, M.: Critical reflection, unlearning, and engagement. *Management Learning* 50/4, pp. 465-481, 2019.
- [MN17] Morais-Storz, M.; Nguyen, N.: The role of unlearning in metamorphosis and strategic resilience. *The Learning Organization* 24/2, pp. 93-106, 2017.
- [NJNI17] Nygren, N. V.; Jokinen, A.; Nikula, A.: Unlearning in managing wicked biodiversity problems. *Landscape and Urban Planning* 167, pp. 473-482, 2017.
- [No94] Nonaka, I.: A dynamic theory of organizational knowledge creation. *Organization Science* 5/1, pp. 14-37, 1994.
- [OBF19] Oliveira, M. I. S.; Barros Lima, G. D. F.; Lóscio, B. F.: Investigations into data ecosystems: a systematic mapping study. *Knowledge and Information Systems* 61, pp. 589-630, 2019.
- [OL18] Oliveira, M. I. S.; Lóscio, B. F.: What is a data ecosystem?. In: *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, pp. 1-9, 2018.
- [OP04] Okoli, C.; Pawlowski, S. D.: The Delphi method as a research tool: an example, design considerations and applications. *Information & Management* 42/1, pp. 15-29, 2004.
- [Pa18] Pappas, I. O. et al.: Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. *Information systems and e-business management* 16, pp. 479-491, 2018.

- [Pe19] Peschl, M. F.: Unlearning towards an uncertain future: on the back end of future-driven unlearning. *The Learning Organization* 26/5, pp. 454-469, 2019.
- [RJB12] Rogers, P. P.; Jalal, K. F.; Boyd, J. A.: An introduction to sustainable development. Earthscan, Routledge, London, 2012.
- [Sc15] Schryen, G.: Writing qualitative IS literature reviews—guidelines for synthesis, interpretation, and guidance of research. *Communications of the Association for Information Systems* 37/1, pp. 286-325, 2015.
- [Sc23] Schoormann, T. et.al.: Artificial Intelligence for Sustainability—A Systematic Review of Information Systems Literature. *Communications of the Association for Information Systems*, 52/1 pp. 199-237, 2023.
- [SL22] Sharma, S.; Lenka, U.: On the shoulders of giants: uncovering key themes of organizational unlearning research in mainstream management journals. *Review of Managerial Science* 16/6, pp. 1599-1695, 2022.
- [Ta04] Tallman, S. et.al.: Knowledge, clusters, and competitive advantage. *Academy of Management Review* 29/2, pp. 258-271, 2004.
- [Te07] Teece, D. J.: Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* 28/13, pp. 1319-1350, 2007.
- [Ts08] Tsang, E. W.: Transferring knowledge to acquisition joint ventures: an organizational unlearning perspective. *Management Learning* 39/1, pp. 5-20, 2008.
- [TZ08] Tsang, E. W.; Zahra, S. A.: Organizational unlearning. *Human Relations* 61/10, pp. 1435-1462, 2008.
- [UN23] United Nations, www.sdgs.un.org/goals, accessed: 17/05/2023.
- [VN22] Vu, M. C.; Nguyen, L. A.: Mindful unlearning in unprecedented times: Implications for management and organizations. *Management Learning* 53/5, pp. 797-817, 2022.
- [VW19] Vogt, M.; Weber, C.: Current challenges to the concept of sustainability. *Global Sustainability* 2/E4, pp. 1–6, 2019.
- [Wa17] Wang, X. et.al.: Organizational unlearning and knowledge transfer in cross-border M&A: the roles of routine and knowledge compatibility. *Journal of Knowledge Management* 21/2, pp. 360-378, 2017.
- [Yi18] Yin, R. K.: *Case Study Research and Applications: Design and Methods*, 6th ed., Thousand Oaks, CA: Sage, 2018.
- [ZLW13] Zhao, Y.; Lu, Y.; Wang, X.: Organizational unlearning and organizational relearning: a dynamic process of knowledge management. *Journal of Knowledge Management* 17/6, pp. 902-912, 2013.
- [Zu15] Zuiderwijk, A. et.al.: Open data for competitive advantage: insights from open data use by companies. In: *Proceedings of the 16th annual international conference on digital government research*, pp. 79-88, 2015.

The Hitchhiker's Guide to Urban Spaces

Conceptualizing a Gaia-X-enabled Co-bility Hub Combining Public Transport, Crowd Mobility, and Last Mile Logistics

Christoph Heinbach ¹, Henning Gössling¹ and Oliver Thomas^{1,2}


Abstract: Urban transportation is increasingly challenged by growing populations and the rapid growth of e-commerce, thus, driving data-driven innovations for sustainable mobility services. Shared mobility consequently emerges as a promising city transport concept, while combined service opportunities between public transport, crowd mobility, and last mile logistics are scarcely investigated. In this paper, we explore the co-creation of urban mobility services within federated ecosystems focusing on a transshipment hub, and propose a novel approach called “co-bility.” Following a design science research (DSR) approach, we conceptualize a co-bility hub based on literature and expert interviews with practitioners from the mobility sector. The exchange of data and services in urban spaces is based on the technical framework Gaia-X. Our study findings show that a Gaia-X-enabled co-bility hub can be achieved by (a) a federated ecosystem orchestrating mobility services and resources, (b) municipalities ensuring coherent platform governance, and (c) eclectic incentives to make co-bility successful.

Keywords: Shared Mobility Hub, Co-bility, Public Transportation, Crowd Mobility, Last Mile Logistics, Gaia-X, Design Science Research

1 Introduction

In 2018, about 55% of the world's population lives in cities with estimates of 68% to be reached by 2050 [Un19] resulting in higher volumes of passenger transport. At the same time, worldwide retail e-commerce sales are expected to increase by 23.6% between 2020 and 2025 with a forecasted sales record of approximately \$7.4 trillion [In23], driving a continuously growing “last mile” delivery market [As20]. Although urban transport activities relying on motor vehicles account for substantial amounts of emissions generated, sustainable mobility solutions through technological advancements (e.g., electric vehicles) are yet facing challenges. Furthermore, the pertinent problems from the transportation of passengers and goods in cities are related to severe environmental effects (e.g., traffic congestion) [Eu06]. To address these issues, the

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Smart Enterprise Engineering, Hamburger Str. 24, 49084 Osnabrück, {christoph.heinbach, henning.goesling, oliver.thomas}@dfki.de,

 <https://orcid.org/0000-0003-1602-5192>

² University Osnabrück, Informationsmanagement und Wirtschaftsinformatik, Hamburger Str. 24, 49084 Osnabrück, oliver.thomas@uni-osnabrueck.de

promising concept of shared mobility and transport collaboration in urban areas has leveraged new forms of platform-enabled mobility services [CI19]. For instance, platform providers have recognized shared service capabilities for crowd mobility comprising cars, bikes, and e-scooters facilitating *mobility as a service* (MaaS) business [AG20] or *mobility on demand* concepts [SC20]. Likewise, last mile delivery innovations have demonstrated benefits for achieving automated deliveries by remotely operating driverless drones or robots for the courier, express, and parcel (CEP) providers [As20]. At the same time, commercial road freight transportation is increasingly turned into a data-driven platform business [He22].

In recent years, the integration of passenger, crowds, and freight transport has shaped multifold concepts (e.g., crowdsourcing, co-modality, cargo hitching) [CN22, Ho06] and revealed the idea of a “shared mobility hub” [Ro23] aiming at sustainable urban mobility ecosystems. Nevertheless, public and crowd resources used to support CEP providers in delivering shipments are accompanied by organizational and technological challenges of collaboration [CI19]. However, shared or collaborative mobility remains an isolated phenomenon and lacks a comprehensive view to combine public transport, crowd mobility, and last mile deliveries [SR19], requiring design knowledge to strengthen the development of co-created mobility services for collaborative innovations [SC16]. In this context, the transformation of urban transportation by platform innovations across different research fields [Ce21] motivates the exploration of interconnected mobility systems using emerging federated ecosystems (e.g., Gaia-X) to maintain digital sovereignty and to secure data exchange between the actors involved [Br21, He23]. For this reason, this paper explores the co-creation of mobility services by combining public transportation, crowd mobility, and last mile logistics framed by the novel approach “co-bility” based on a federated architecture. The idea of a co-bility hub arises from a current Gaia-X research project focusing on competitive fleet management in urban areas and combining transport resources from the public (e.g., buses) and last mile delivery operations (e.g., delivery robots). Since value propositions found in higher transport capacities utilization, synchronizing pick-up and drop-off (PUDO) service levels, and socializing opportunities [CN22, Du19] arrive, the research question (RQ) guiding this study to establish a concept for both practice and future research is as follows: *How can a co-bility hub combining public transport, crowd mobility, and last mile logistics in urban areas be conceptualized based on the ecosystem Gaia-X?*

To answer the RQ, the remainder of the paper is divided into seven sections. In Section 2, we provide the fundamentals of our study comprising shared mobility hubs, the co-bility hub approach, and the federated ecosystem Gaia-X. Subsequently, in Section 3, we establish a design science research (DSR) approach comprising a theoretical knowledge base and identified business needs from expert interviews of the current Gaia-X project research community. Afterward, we present the concept of a co-bility hub and describe a Gaia-X-enabled platform for sovereign data exchange (Section 4). Then, in Section 5, we evaluate the derived concept through an architecture analysis. In Section 6, we discuss the finding for academia and practice and draw the limitations. Finally, we conclude with an outlook for future research in Section 7.

2 Foundations

2.1 Emergence of shared mobility hubs in urban areas

Shared mobility hubs represent novel innovations in the field of mobility [Bl22] and emerge growingly by combining different transport modes to foster sustainable mobility in cities [Ro23]. The key value propositions of these hubs are found in the substitution of privately owned assets, vehicle utilization, and social interactions by contributing actors [Co21]. As mobility options, shared cargo bikes, e-scooters, and cars in combination with public transport systems are integrated at a dedicated place [Ro23]. For example, in the Dutch city of Amsterdam, a shared mobility hub was inaugurated in 2021 next to a hotel and within 100m distance of a metro station offering shared e-bikes, e-cars, and cargo bikes [Ro23]. Considering the dramatic change in energy and resource efficiency, the increasing number of related EU-funded projects (e.g., SmartHubs [Sm23] and eHUBS [NW23]) indicate the substantial relevance of future shared mobility hubs for cities [Ro23]. Overall, there are different aspects to be considered for setting up a shared mobility hub at a certain location, such as the expected earnings for the operator, the investment and operating costs, the additional space needed, the generalized travel costs of using the hub by the end-user, the increase in travel time reliability for the end-user, and the impact on quality of life and emissions [Bl22].

2.2 Co-bility hub: Approach and definition

Due to the recent shared mobility trends based on vehicles or passenger rides [SC20], future mobility ecosystems will rely increasingly on advanced connectivity between urban transport modes [SC16]. Shared mobility hubs are, thus, recognized as places, “*where different transport modes are integrated seamlessly, promoting efficient and sustainable urban mobility.*” [Ro23, p. 6]. The hubs are associated with cargo bikes or amenities like stationary PUDO points, making them a place for integrating last mile delivery activities. To integrate mobility resources and provide mobility services to ecosystem participants, digital platforms provide flexible cloud infrastructure. Digital mobility platforms, therefore, emerge as marketplaces [SR19] and are associated with multi-sided markets [HW15] allowing the exploration of value co-creation opportunities [Va08] discussed in the mobility context [Sc23]. In reflecting on the integration of mobility resources and the provision of connected mobility services to be exchanged between ecosystem actors for mutual benefit, we introduce the term “co-bility,” a portmanteau comprising “co-creation” and “mobility.” Following the proposed definition of shared mobility hubs [Ro23], a co-bility hub consequently describes a place in urban areas where passengers, crowds, and freight are integrated into a platform to facilitate seamless collaboration for mutual benefit by promoting data and service exchange between mobility providers and consumers. Capturing value within a co-bility hub, thus, follows different propositions (e.g., co-creation) formed by platform interactions between participants entailing collaborations [Va17].

2.3 Shared mobility in the federated ecosystem Gaia-X

The Gaia-X initiative has been launched primarily by Germany and France in 2019 to provide a (technical) framework offering the European economy a competitive alternative to the dominance of “hyperscalers” in the cloud business [BM19]. Gaia-X consequently aims at providing companies with a trustworthy and decentralized data ecosystem based on a federated data infrastructure. It is expected that Gaia-X will support interoperable data exchange, thus, enabling collaborative value creation and joint business models [Kr23]. However, Gaia-X Federation Services (GXFS) represent the enabling software components currently maturing by ongoing implementations of open-source software [As23]. In this light, shared mobility offers a fruitful research avenue for the Gaia-X development currently addressed, for instance, by current projects focussing on transport automation and data interoperability between freight fleets [He23]. Based on the Gaia-X principles, the Mobility Data Space³ has been launched in 2020 to exchange mobility data [Ot22]. The value propositions of the data space follow the idea of a platform-enabled marketplace where data is offered or purchased by participants (e.g., service providers) accessing the network. Likewise, the data consumed by searching a data catalog allows for creating new services provided by the ecosystem participants [Ot22]. Since the scope of implemented business services based on the architecture has shown benefits to the actors by provisioning information [Mo23], further collaboration opportunities arrive by combining public transport, crowd mobility, and last mile logistics and favor the exploration of a Gaia-X-enabled co-bility hub.

3 Research Method

To answer the guiding RQ in this paper, our study follows the DSR principles to analyze knowledge from both science and practice for the creation of a co-bility hub [He04]. For the entry point of our study, literature from academia was used to ground a scientific knowledge base (rigor). Relevant stakeholders and collaboration approaches related to a platform-enabled co-bility hub were primarily identified in the work of [CI19] and [SR19]. Subsequently, knowledge from practice was gained from expert interviews and document-based analysis to identify business needs (relevance) responding to the practical character of our topic. To evaluate the derived concept, an architecture analysis [He04] was performed to assess the fit with the developing Gaia-X components resulting from the ongoing research project GAIA-X 4 ROMS. The research framework covered in this paper follows the DSR guidelines outlined by [He04]. Overall, the goal is to build a novel design artifact that extends the current knowledge base and respects its applicability in business environments as detailed in Figure 1.

³ <https://mobility-dataspace.eu>, accessed: 14/04/2023.

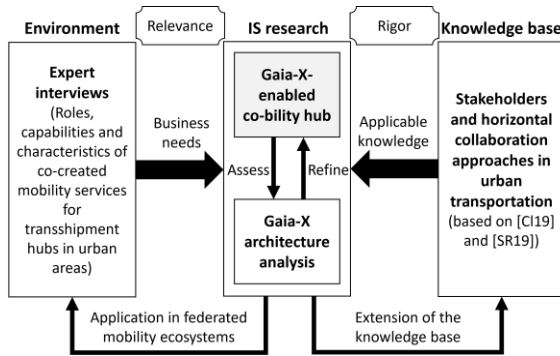


Fig. 1: Research framework for the co-bility hub adapted from [He04]

To gather insights from practice, semi-structured interviews [He04] were conducted with experts involved in the mentioned Gaia-X research project. The research project is part of the project family GAIA-X 4 FUTURE MOBILITY⁴, and the participating organizations forming individual consortiums provide access to potential experts related to the topic of our research study. In Figure 2, an anonymized overview of the 12 interviewed experts including the organization role, organization size (per EU definition), position, work experience, and interview duration is provided.

Org. type	Org. size	Expert position	Experience	Duration
Last mile logistics software provider A	Large	AI solution architect and data spaces	12 yrs	40 min
Last mile logistics software provider B	Large	Last mile logistics innovations manager	4 yrs	44 min
Last mile logistics software provider C	Large	Business developer smart city	5 yrs	46 min
Mobility software provider A	Medium	Sales manager mobility sol.	25 yrs	50 min
Mobility software provider B	Small	Innovation/network manager	10 yrs	50 min
Mobility software provider C	Small	Project manager mobility data spaces	3 yrs	54 min
Mobility software provider D	Small	Senior project manager intelligent mobility	4 yrs	56 min
Municipality of a city	Large	Consulter urban mobility and federated data spaces	15 yrs	49 min
Public transport provider A	Large	Technical consultant mobility	1 yrs	43 min
Public transport provider B	Large	Product owner digital mobil.	2 yrs	44 min
Research institution A	Large	Researcher for intelligent/connected mobility systems	2 yrs	34 min
Research institution B	Medium	Professorship for mobility, retail, and logistics	23 yrs	45 min

Fig. 2: Overview of interviewed experts

⁴ <https://www.gaia-x4futuremobility.dlr.de/>, accessed: 10/05/2023.

An interview guide was sent to the experts beforehand consisting of open questions divided into five parts: (1) introduction, (2) combined mobility, (3) Gaia-X enabled co-creation opportunities, (4) co-bility hub requirements, and (5) challenges and risks. To analyze the collected data, the interviews were coded following deductive and inductive category formation [Ma14]. The interviews were extended with elements for constructing, justifying, and conflicts of interactions for interpretative research [Wa06], and we asked follow-up questions to gain more in-depth information on specific topics (e.g., business models). This strategy allows enriching information for the topic addressed by our research study, which results from discussions with the respondents [Pa15]. To arrive at the desired artifact of a Gaia-X-enabled co-bility hub, the method of conceptual modeling [ET11] was applied to describe the platform application on an abstracted level composed of resources, infrastructure, services, and participants. From the findings of our analysis, we extend the knowledge base in the field of mobility ecosystems and obtain a model that contributes to nascent design theory positioned within the framework of DSR contribution types [GH13]. The intended artifact, thus, represents an exaptation in the DSR knowledge contribution framework [GH13], implying the extension of known solutions to new phenomena and fields of applications (i.e., co-created mobility combining passengers, crowd, and freight transportation).

4 Conceptualization of a Gaia-X-enabled Co-bility Hub

4.1 Design elicitation of a Gaia-X-enabled co-bility hub

Following the work of [SR19], a Gaia-X-enabled co-bility hub grounds on six stakeholder groups directly interacting with each other: (1) MaaS provider (e.g., car sharing provider), (2) CEP delivery provider (e.g., logistics company), (3) craftsman (e.g., construction companies for private persons), (4) crowd (e.g., private person, individual), (5) public transport service provider (e.g., fleet operator of buses), and (6) retailer (e.g., grocery store, entrepreneur). Although additional stakeholders with relevance to the infrastructure of a co-bility hub were mentioned (e.g., provider of electricity stations) during the interviews, no direct linkage was recognized for co-creating platform services. From the expert interviews, we consolidated the requirements and identified transactions between the participants that are manifested by four core services: (a) last mile delivery support (e.g., individuals take parcels from CEP after route matching), (b) retail shopping forwarding (e.g., entrepreneurs offer purchases for delivery to individuals), (c) material supply storing (e.g., craftsman can book dedicated parking zone to ensure supply during constructions), and (d) crowd mobility matchmaking (e.g., individuals share assets or make use of transport rates collaboratively). To realize these services, the interview participants emphasized that transport capacities and vehicle resources require integrating them into the hub system. Furthermore, we noticed a consensus from the experts on the role of the municipality, to be positioned as a platform provider (e.g., server host) to govern secure and reliable

transactions based on GXFS components. In this light, all experts from our interviews have acknowledged the capabilities of Gaia-X to be applied as the technical framework providing federated data exchange components for the participants (e.g., onboarding, logging, identity management, service catalog). However, fundamental requirements to make value co-creation for shared mobility services successful rely on incentive mechanisms and were stressed by all experts. Therefore, incentives are pivotal to building acceptance and making the platform solution attractive [C119], thus, encouraging participants to realize network effects. During our discussions, we have identified versatile incentive opportunities encompassing credit points (e.g., parcels being forwarded are awarded credits for individual benefits), gamification (e.g., hub performance is set as a competition between participants), repayments or discounts (e.g., cost saving of CEP provider are shared with individuals), social interactions (e.g., individuals meet-up for vehicle sharing supporting socializing), and service level improvements (e.g., retailers involve the crowd to deliver purchases to elderly people). Overall, the engaged participants in the co-bility hub appear to follow their personal benefits or business model opportunities likewise to be governed by the municipality.

4.2 Conceptual model of a Gaia-X-enabled co-bility hub

The elicited findings described above form the basis for a co-bility hub focalizing the co-creation of mobility services based on a federated data infrastructure. From the identified stakeholders, co-created services, integrated resources, and incentive opportunities, we constitute a conceptual model and arrive at a generic Gaia-X-enabled co-bility hub combining public transport, crowd mobility, and last mile logistics (Figure 3).

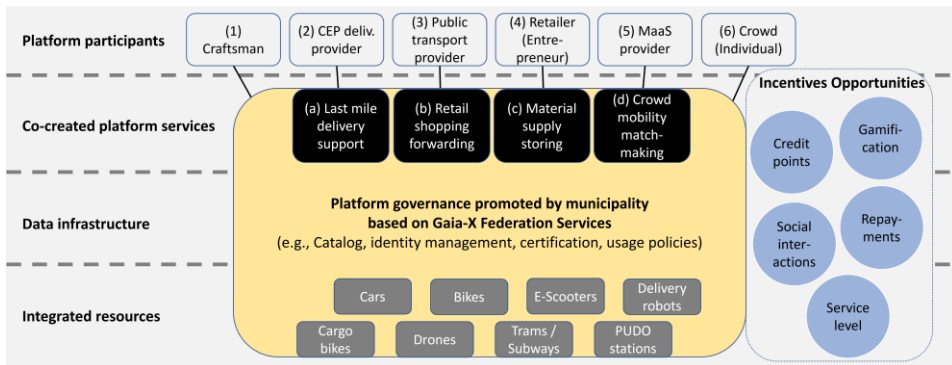


Fig. 3: Generic concept of the Gaia-X-enabled co-bility hub

Within the proposed concept of a co-bility hub, platform interactions remain abstracted according to our findings and miss further information to detail the requirements and functions on a more granular level. Therefore, the provided conceptual model is a first-time approach for grasping co-bility that makes an urban interconnected mobility system associated with emerging federated ecosystems (i.e., Gaia-X) applicable in practice.

5 Evaluation by Gaia-X Architecture Analysis

To evaluate the proposed Gaia-X-enabled co-bility hub, an architecture analysis was performed based on publicly available information from the Gaia-X database and scientific papers addressing federated data infrastructures. The authors focussed on knowledge related to data-driven services offerings in the context of mobility and demonstrations that could serve as a basis to position the derived concept. Subsequently, from the results of our analysis, an abstracted data space architecture was derived for the conceptualized model and further collated with the artifacts from the ongoing research project GAIA-X 4 ROMS. From that basis, the authors present the corresponding federation services and components to obtain a federated data space aligned with the general concept of the nascent Gaia-X ecosystem. As a result, the created co-bility hub data space represents a multi-stakeholder environment and follows the federated data infrastructure implemented for the Mobility Data Space. The participants are connected using a connector (e.g., Eclipse Data Space Connector [Ec23] for providing and consuming data encapsulated in co-created mobility services. A federated catalog allows both data providers and consumers to describe their offerings or search for the desired data or services. Accessing the data or the service is only possible by the data owner's self-determination, enabling data sovereignty. In addition, the catalog is open to other providers (e.g., software developers) and consumers (e.g., vehicle manufacturers) from other domains supporting the interoperability of the data space (Figure 4).

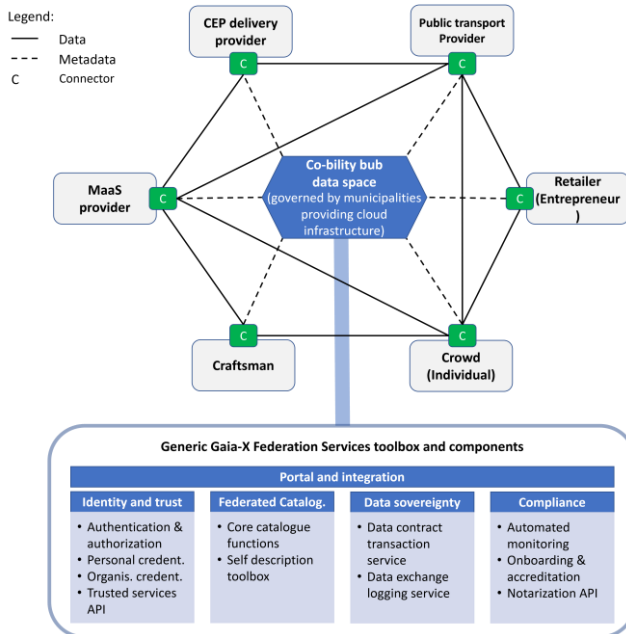


Fig. 4: Overview of the co-bility hub data space derived from analysis and adapted from [Ot22]

Notwithstanding these provisions, data from resources (e.g., PUDO stations, vehicles) use technologies (e.g., Internet of Things) facilitating co-created platform services to be integrated via connectors with other back-end components. Overall, municipalities could govern the service-oriented transactions restricted to the cloud infrastructure and incentive settings. It may comprise the authentication of the participants, the notarization of data and services offered, and the regulations on compensations. A general overview of the generic co-bility hub data space is finally built based on the generic Gaia-X Federation Services toolbox and components provided by the eco – Verband der Internetwirtschaft e.V.⁵ discussed in the context of the research project GAIA-X 4 ROMS. In essence, the authors reflected several benefits during the analysis capturing value through co-created mobility services within a secure and federated data infrastructure. This includes (1) the sovereign service provisioning and consuming, (2) the integration of isolated urban mobility stakeholders, (3) the incentive opportunities for social engagement of the crowd, (4) the business opportunities for commercial transport providers, and (5) the data exchange in compliance with European legislation.

6 Discussion and Limitations

The concept derived from our DSR analysis represents a platform-enabled artifact aiming at sustainable mobility and mitigating traffic impact in urban areas. To this end, co-created mobility services entail value propositions focusing on delivery support of parcels, take away retailer purchases, storing of material supply for urban constructions, and matchmaking opportunities of the crowd mobility offerings. We answer our research question with a designed solution called co-bility hub established by a transshipment point for commodities and co-created services between platform participants. Co-bility is, thus, positioned at the edge of smart cities and smart mobility through digital innovations and collaborative transportation within urban spaces. Our proposed conceptual model responds to the rise of interconnected urban transportation in the sharing economy [SC20] and provides a federated approach that goes beyond collaborative city logistics systems [C119] based on the domain-transferable GXFS components. In addition, the identified provider role of municipalities addressing platform governance corresponds with the findings from [Sc23] and confirms the orchestration of transactions from outside the mobility operations. However, the target of this study was not to understand the service capabilities of existing mobility solutions operated in cities. Rather, our DSR study focuses on the design of a mobility hub for passenger, crowd, and freight incorporating shared service opportunities [Ro23], associated with the Gaia-X framework to achieve an interoperable and sovereign urban ecosystem. The presented co-bility hub provides theoretical design knowledge with valuable insights, likewise relevant for practice. The conceptualized model can be seen as a reference to guide stakeholders to achieve collaborative transport in cities. While the acceptance remains critical to realize a co-bility hub [C119], the concept provides new

⁵ <https://www.eco.de/>, accessed: 10/05/2023.

impetus to make urban spaces a data-driven business field and may attract additional stakeholders such as fleet operators or insurance companies.

Although our DSR study is based on a careful selection of interview participants and rigorous data analysis, our findings are neither complete nor comprehensive. Other experts from the mobility domain and apart from the Gaia-X research community would have identified other stakeholders and service potentials. Due to the novelty of the topic, co-created mobility services derived from the interviews have resulted in the abstraction of service specifications for the concept, thus, entailing an uncertainty to meet all functions required. Not to be forgotten are the challenging aspects related to the concept and discussed with the experts. It comprises the segregation of liability (e.g., loss or damage, of parcel deliveries), planning dimensions for matchmaking resources, business model requirements, tracking capabilities, and usability propositions including the personalization for special groups (e.g., disabled individuals). Our work can only be seen as a starting point and calls for more profound knowledge to make co-bility a sustainable transport service, for instance, by considering emerging “eco-labelling” concepts [Ke22].

7 Conclusion and Outlook

The demand for future urban transport is expected to rapidly grow with severe consequences for the traffic infrastructure and the environment. In this light, our research study contributes to sustainable transportation by exploring platform-enabled co-creation between public transport, crowd mobility, and last mile logistics. Following a DSR approach [He04], we used existing literature and conducted interviews with 12 experts from the digital mobility sector to derive a platform-enabled shared mobility hub. The concept strengthens the co-created mobility in urban ecosystems by combining transport resources from different areas resulting in transaction-based services comprising capacity utilization, advanced service levels, and social interactions termed co-bility as a novel approach. The designed artifact was evaluated by performing an architecture analysis aligned to the Gaia-X ecosystem yielding mutual benefits between platform participants. Whether the presented concept will succeed in the real world by implementing demonstrations addressing the identified services using the Gaia-X framework may result in new questions driving future research activities. We hope that our work will help to span the gap between theory and practice and augment design knowledge with high societal and ecological relevance in the sphere of urban spaces.

Acknowledgements

This research study is part of the project Gaia-X 4 ROMS – Support and Remote Operation of Automated and Networked Mobility Services (FKZ: 19S21005C). The funded project is promoted by the German Federal Ministry of Economics and Climate Protection (BMWK), and the authors are responsible for the content of this paper.

Bibliography

- [AG20] Arias-Molinares, D.; García-Palomares, J.C.: Shared mobility development as key for prompting mobility as a service (MaaS) in urban areas: The case of Madrid. *Case Studies on Transport Policy* 8/03, pp. 846–859, 2020.
- [As20] Asdecker, B.: How Dare You Replace My Deliveryman?! Acceptance of Last-Mile Transportation Innovations—A Qualitative Perspective. In: *Proceedings of the Americas Conference on Information Systems, Salt Lake City*, 2020.
- [As23] Association of the Internet Industry: Implementation. <https://www.gxfs.eu/implementation/>, accessed: 11/04/2023.
- [Bl22] Blad, K. et al.: A methodology to determine suitable locations for regional shared mobility hubs. *Case Studies on Transport Policy* 10/03, pp. 1904–1916, 2022.
- [BM19] BMWi, Bundesministerium für Wirtschaft und Energie: Das Projekt GAIA-X - Eine vernetzte Dateninfrastruktur als Wiege eines vitalen, europäischen Ökosystems. Berlin, 2019.
- [Br21] Braud, A. et al.: The Road to European Digital Sovereignty with Gaia-X and IDSA. *IEEE Network* 35/02, pp. 4–5, 2021.
- [CN22] Cavallaro, F.; Nocera, S.: Integration of passenger and freight transport: A concept-centric literature review. *Research in Transportation Business & Management* 43, 100718, 2022.
- [Ce21] Ceder, A.: Urban mobility and public transport: future perspectives and review. *International Journal of Urban Sciences* 25/04, pp. 455–479, 2021.
- [Cl19] Cleophas, C. et al.: Collaborative urban transportation: Recent advances in theory and practice. *European Journal of Operational Research* 273/03, pp. 801–816, 2019.
- [Co21] Coenegrachts, E. et al.: Business Model Blueprints for the Shared Mobility Hub Network. *Sustainability* 13/12, 6939, 2021.
- [Du19] van Duin, R. et al.: Evaluating new participative city logistics concepts: The case of cargo hitching. *Transportation Research Procedia* 39, pp. 565–575, 2019.
- [Ec23] Eclipse Foundation: Eclipse Dataspace Components, <https://projects.eclipse.org/projects/technology.edc>, accessed: 04/07/2023.
- [ET11] Embley, D.W.; Thalheim, B., eds: *Handbook of Conceptual Modeling – Theory, Practice, and Research Challenges*. Springer Berlin Heidelberg, 2011.
- [Eu06] European Commission: *Integrated Pollution Prevention and Control. Reference Document on Best Available Techniques on Emissions from Storage*. Brussels, Belgium, 2006.
- [GH13] Gregor, S.; Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly* 37/02, pp. 337–355, 2013.
- [HW15] Hagi, A.; Wright, J.: Multi-sided platforms. *International Journal of Industrial Organization* 43, pp. 162–174, 2015.

- [He22] Heinbach, C. et al.: Data-driven forwarding: a typology of digital platforms for road freight transport management. *Electronic Markets* 32, pp. 807–828, 2022.
- [He23] Heinbach, C. et al.: Smart Managed Freight Fleet: Ein automatisiertes und vernetztes Flottenmanagement in einem föderierten Datenökosystem. *HMD Praxis der Wirtschaftsinformatik* 60, pp. 193–213, 2023.
- [He04] Hevner, A. R. et al.: Design Science in Information Systems Research. *MIS Quarterly* 28/01, pp. 75–105, 2004.
- [Ho06] Howe, J.: The Rise of Crowdsourcing. *Wired Magazin* 14, pp. 1–4 (2006).
- [In23] Insider Intelligence: Global Ecommerce Forecast 2022, www.insiderintelligence.com/content/global-ecommerce-forecast-2022, accessed: 11/04/2023.
- [Ke22] Kirschstein, T. et al.: Eco-labeling of freight transport services: Design, evaluation, and research directions. *Journal of Industrial Ecology* 26/03, pp. 801–814, 2022.
- [Kr23] Kraemer, P. et al.: Gaia-X and Business Models: Types and Examples, Gaia-X Hub Germany. White Paper 1, 2023.
- [Ma14] Mayring, P.: *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Klagenfurt, 2014.
- [Mo23] Mobility Data Space: Success through cooperation: MDS Business Cases, <https://mobility-dataspace.eu/business-cases>, accessed: 14/04/2023.
- [NW23] NWE Europe: eHUBS - Smart Shared Green Mobility Hubs. <https://vb.nweurope.eu/projects/project-search/ehubs-smart-shared-green-mobility-hubs/>, accessed: 10/05/2023.
- [Ot22] Otto, B.: A federated infrastructure for European data spaces. *Commun. ACM* 65/04, pp. 44–45, 2022.
- [Pa15] Patton, M.Q.: *Qualitative Research & Evaluation methods*. SAGE Publications, Thousand Oaks, 2015.
- [Ro23] Roukouni, A. et al.: An Analysis of the Emerging “Shared Mobility Hub” Concept in European Cities: Definition and a Proposed Typology. *Sustainability* 15/06, 2023.
- [Sc23] Schulz, T. et al.: Smart Mobility: Contradictions in Value Co-Creation. *Information Systems Frontiers* 25, pp. 1125–1145, 2023.
- [SC16] Shaheen, S.; Chan, N.: *Mobility and the Sharing Economy: Potential to Overcome First- and Last-Mile Public Transit Connections*. UC Berkeley: Transportation Sustainability Research Center, 2016.
- [SC20] Shaheen, S.; Cohen, A.: - Mobility on demand (MOD) and mobility as a service (MaaS): early understanding of shared mobility impacts and public transit partnerships. In (Antoniou, C. et al., eds.): *Demand for Emerging Transportation Systems*, Elsevier, pp. 37–59, 2020.
- [Sm23] SmartHubs: Creating smart shared mobility options for the city of tomorrow, <https://smarthubs.eu/>, accessed: 10/05/2023.

- [SR19] Szmelter-Jarosz, A.; Rześny-Cieplińska, J.: Priorities of Urban Transport System Stakeholders According to Crowd Logistics Solutions in City Areas. *A Sustainability Perspective*. *Sustainability* 12/01, 317, 2019.
- [Un19] United Nations: *World Urbanization Prospects: the 2018 Revision*. United Nations, New York, 2019.
- [Va17] Vargo, S.L. et al.: Conceptualizing Value: A Service-ecosystem View. *Journal of Creating Value* 3/02, pp. 117–124, 2017.
- [Va08] Vargo, S.L. et al.: On value and value co-creation: A service systems and service logic perspective. *European Management Journal* 26/03, pp. 145–152, 2008.
- [Wa06] Walsham, G.: Doing interpretive research. *European Journal of Information Systems* 15/03, pp. 320–330, 2006.

Ökologische Nachhaltigkeit -
Zukunft nachhaltig gestalten durch
digitalisierte
Wertschöpfungsprozesse (DigiWe)

Technology Readiness Levels of Reinforcement Learning methods for simulation-based production scheduling

Arne Seipolt¹, Ralf Buschermöhle¹, Maximilian Höfinghoff¹, Goy-Hinrich Korn²,
Marcel Schumacher¹

Abstract: Digital Twins (DT) are nowadays widely used and provide a benefit for the companies using it. One service of the DT is the simulation of a production process. This enables an optimization of the production process by simulation optimization, for example with Reinforcement Learning (RL). To support researchers and practitioners in deciding which algorithm is suitable for an implementation under real-life conditions, a literature research is performed, and a Machine Learning Technology Readiness Level is assigned to the different RL-Algorithms. It can be shown that recent research focuses mainly on model free value based and evolutionary algorithms, and both are suitable for an implementation in a real-world scenario. Both algorithms can outperform widely applied dispatching rules. Nevertheless, it should be evaluated why other algorithms are not in the focus of recent research and how the algorithms perform in comparison to each other.

Keywords: Technology Readiness, Reinforcement Learning, Simulation, Production Scheduling

1 Introduction

Digital Twins (DT) provide a benefit for the companies using it. A study from 2022 shows, that the companies, which are using Digital Twins, could improve their operational efficiency on average by 15 % [Gy22]. Furthermore, 68 % of the asked organizations see a simulation as a service from the DT, justifying its use [Gy22]. For example, it is possible to optimize the production scheduling by using a simulation [Sc07]. But there is still potential for further development [Mo20]. While there are different methods to optimize the parameters of a simulation [LTD22], these optimizations can often only be done approximately [La17].

Recently, Reinforcement Learning (RL) methods have achieved great successes in playing different games and are able to outperform human players [Mn15]. These algorithms can be used for simulation optimization by suggesting promising simulation parameters [PBG22, KD21]. In a literature research, Panzer et al. reviewed a total of 55 recent scientific articles that deal with production scheduling using reinforcement learning. However, 52 of these were implemented and validated exclusively in a laboratory

¹ Osnabrück University of Applied Sciences, Faculty of Management, Culture and Technology, Kaiserstraße 10c, D-49809 Lingen/Ems, {a.seipolt, r.buschermoehle, m.hoefinghoff, marcel.schumacher}@hs-osnabrueck.de

² Bernard Krone Holding GmbH & Co. KG, CIO & CDO, 48480 Spelle, Heinrich-Krone-Straße 10, goy-hinrich.korn@krone.de

environment. Therefore, no general statements can be made about the reliability of such systems in reality [PBG22].

To support researchers and practitioners in deciding which method is suitable for an implementation under real-life conditions, this paper aims to show different reinforcement learning methods for simulation-based optimization of production processes. In order to show the technology readiness of these methods, they will be assigned a “Machine Learning Technology Readiness Level” (MLTRL), defined in [La22].

For this purpose, a definition of production scheduling and an overview of different reinforcement learning algorithms is given, followed by an introduction to the MLTRL. Then a literature research is conducted, considering only papers published in recent years that deal with reinforcement learning for simulation optimization in the context of production planning. These publications are classified based on the algorithms used, in order to check to which MLTRL the implementations and evaluations performed correspond. This is followed by a discussion to derive further research approaches and implementation strategies.

2 Production Scheduling

“Scheduling problems can be understood in general as the problems of allocating resources over time to perform a set of tasks being parts of some processes, among which computational and manufacturing ones are most important.” [B19] “In manufacturing, the purpose of production scheduling is to minimize production time and costs, by telling a production facility when to make something, with which staff, and using which equipment.” [Ri12] Scheduling problems can be characterized by three sets: $\mathcal{T} = \{T_1, T_2, \dots, T_n\}$ defines the different tasks, $\mathcal{P} = \{P_1, P_2, \dots, P_m\}$ defines the processors or machines and $\mathcal{R} = \{R_1, R_2, \dots, R_s\}$ defines additional resources. Different tasks can be combined to jobs. So, job J_j is divided into n_j tasks: $T_{1j}, T_{2j}, \dots, T_{n_jj}$ while different tasks are performed on different machines. Under the condition, that every machine is specialized for the execution of certain tasks, there are three models for processing: *flow shop*, *open shop* and *job shop*. In an open shop, the number of Tasks is equal for every job, and the task T_{ij} is performed on the machine P_i . Additionally, in flow shop, the processing of T_{i-1j} precedes the processing of T_{ij} . In a job shop, n_j is arbitrary [B19].

To introduce the complexity of Production Scheduling, the number of possible solutions for a simple Production Scheduling problem will be reviewed. It is assumed, that n different tasks should be performed to finish a Job. The number of tasks per job is constant. For every task, there are m machines and s resources which are specialized for one job. The sequence, in which the j Jobs will be performed as well as the assignment of machines and resources must be defined. If the sequence of the tasks is fixed, it is a flow shop Problem, if it must be defined as part of the production scheduling, it is an open shop problem.

The sequence, in which the jobs are performed, is a permutation, therefore there are $j!$ different possibilities. Furthermore, there are m possibilities to choose a machine and s possibilities to choose a resource. Therefore, the number of possibilities for a flow shop scheduling problem $k_{Flow\ Shop}$ is $k_{Flow\ Shop} = m * s * j!$. In an open shop problem, the sequence of the tasks has also be defined during the production scheduling. Therefore, the complexity is enlarged by an additional permutation: $k_{Open\ Shop} = m * s * j! * t!$.

To show the influence of the different Parameters, in Fig. 1 the number of possible solutions for an open shop problem is shown, if only the number of machines or jobs is varied. The numbers, which are not varied, are one.

It is obvious, that the influence of the number of jobs is much higher than the number of machines. The reason for this is that the number of jobs, as well as the number of tasks, enter the complexity with their faculty. A main driver of the complexity of production scheduling problems is the number of jobs to be scheduled, and therefore the time planned in the future. For open shop problems, the number of tasks, which must be performed to finish a Job, also have a great influence since it is also going into the complexity with the faculty.

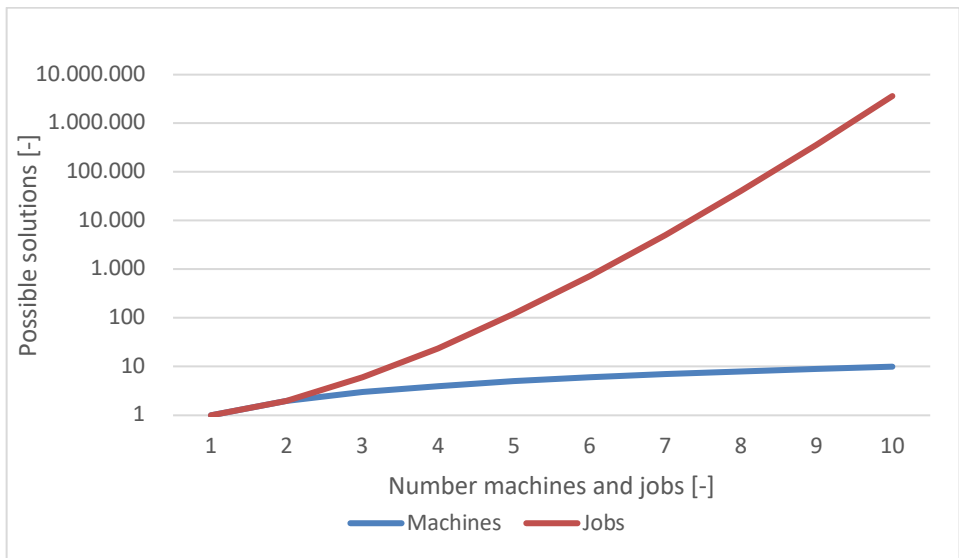


Fig. 1: Varying the number of machines and jobs for an open shop problem.

3 Reinforcement Learning

In Reinforcement Learning (RL), an agent interacts with an environment. Based on the information about the current state of the environment S , it decides which action A it will

perform. The environment changes according to the action and the agent receives a Reward R . This reward is used to train the agent, so it can make better decisions [Di20]. To differentiate the RL-algorithms, a couple of classes will be defined. Zhang and Yu differentiate between model based and model free algorithms. For the model based algorithms, there are some algorithms where the model is given, such as Alpha Go [Si16], where the rules of Go are specified. Other model based algorithms, for example Imagination-Augmented Agents (I2As) [Ra17] learn a model of the environment by themselves [ZY20].

The algorithms, without a model of the environment, are divided in the value-based and the policy based algorithms [ZY20]. Value based algorithms, like DQN [Mn15], try to optimize the action-value function, so the optimal policy is to always choose the action with the highest action value. On the opposite, the policy based algorithms, like Proximal Policy Optimization (PPO) [Sc17], directly optimize the policy, which can be seen as a set of rules that define the action to be performed in a given situation.

Furthermore, there are so-called actor-critic algorithms, which combine both approaches by using a value-based algorithm to learn a value function and a policy-based algorithm to learn the policy function [ZY20, Di20]. One example is the Asynchronous Advantage Actor-Critic (A3C) algorithm [Mn16]. Since it is a combination of value- and policy-based approaches, it will form a new category for this paper.

Zhang and Yu assign evolutionary algorithms to the model free policy based algorithms [ZY20]. Otherwise, Sutton and Barto argue that there are fundamental differences between evolutionary algorithms and other Reinforcement Learning algorithms. For example, they do not notice, which states an individual passes through during its lifetime, or which states it selects [SB18]. Nevertheless, they also have a lot in common with other RL- algorithms, which is why they are included in this paper but as an own category and not as part of the policy-based algorithms.

The selected papers will be categorized in the five categories stated in Fig. 2. Also, different benefits and shortcomings of the algorithm categories are stated, derived from literature.

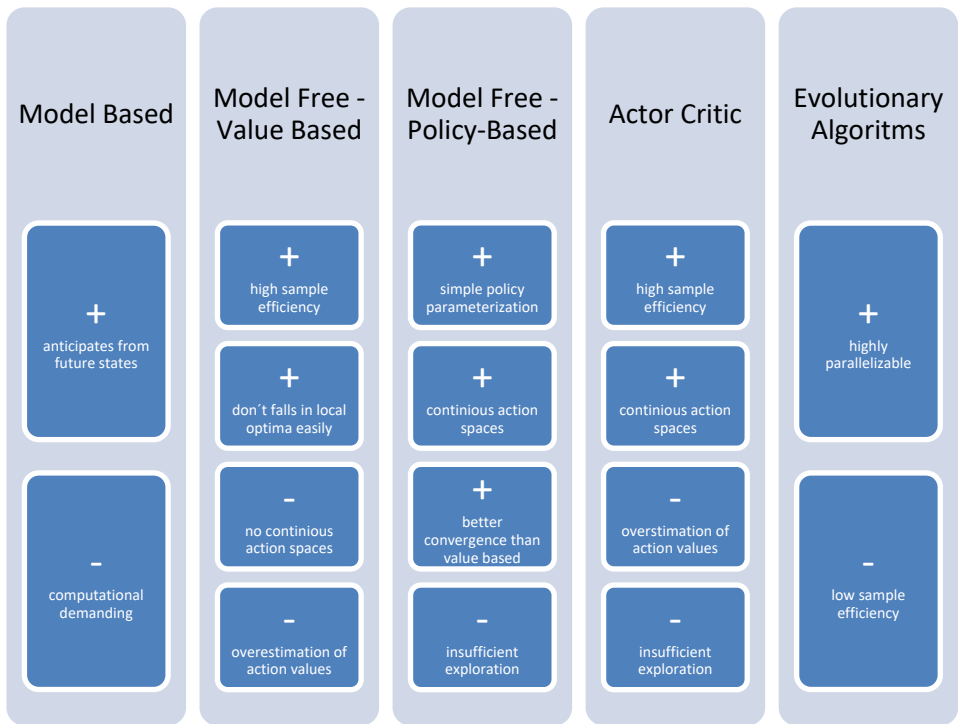


Fig. 2: Categories of Reinforcement Learning algorithms used in this paper, with their benefits and shortcomings.

Every category of Algorithm has different benefits and shortcomings. For example, an model based approach needs a planning algorithm, which is often computational demanding. Therefore, the time constraints have to be taken into accounts for real-time-decision-making [Fr18]. But if it is suitable, it has the advantage of anticipating from future states and rewards in advance [ZY20]. Model free value based methods usually cannot handle continuous action spaces and tend to overestimate the value of an action, but have a high sample efficiency and don't fall in local optima easily [ZY20]. Value based methods have the advantage of simpler policy parameterization, better convergence and are suitable for continuous action spaces [ZY20]. Actor-critic methods, which combine value- and policy-based algorithms, do also combine some of their benefits and shortcomings. For example, they have a high sample efficiency and are suitable for continuous action spaces. But they also have the problem of the overestimation of the value of an action and suffer from insufficient exploration [ZY20]. Evolutionary algorithms are on the one hand less sample efficient than other RL-algorithms, but highly parallelizable on the other hand [Sa17].

4 Technology Readiness Level for Machine Learning

Based on the recognized Technology Readiness Level (TRL), Lavin et al. developed the Machine Learning Technology Readiness Level (MLTRL) which concretizes the requirements of the TRL specifically for Machine Learning technologies [La22]. The aim of this paper is to evaluate, which MLTRL can be assigned to the different Reinforcement Learning algorithms for simulation optimization for production scheduling. Tab. 1 gives a short overview over the different MLTRLs up to MLTRL 4. Higher MLTRLs than 4 indicate, that the integration in a commercial application has started, which is not the case for simulation optimization in production scheduling.

MLTRL	Methodology	Data	Review
0	Literature research, mathematical principles, whiteboarding of concepts and algorithms.	Review of data availability	Team- or laboratory lead
1	Analysis of the model- and algorithm properties	At least representative synthesized data	Research team
2	Application in a testbed	Benchmark-data, partially or completely simulated data	Documented and reproducible achieving of the research claims
3	Ensure interoperability, maintainability, extensibility and scalability	At least according to MLTRL2	Experts for applied AI and engineering should be involved
4	Demonstration in a real-world scenario	Representative, real data	Consideration of security and privacy aspects

Tab. 1: Overview over the MLTRL according to [La22]

For every MLTRL, there is a methodology, a data basis and a review specified. The first MLTRL is Level 0. This stage might be for example greenfield AI research. This is usually done by literature research or whiteboarding of concepts and algorithms. Often, there are no Data available, which is why at this stage, the data availability should be reviewed. Furthermore, at this early stage, the research team- or laboratory lead decides if the MLTRL should be assigned.

To bring the technology to the next level, low-level experiments to analyze the specific model or algorithm properties should be done, rather than end-to end runs for performance benchmark score. At least representative synthesized data should be used, and the research team reviews the first experiments, deciding if further experiments should be done.

To reach MLTRL 2, the Proof of Principle should be done, which means running the algorithm in a test bed, for example in a simulated environment with simulated data. Nevertheless, if available, benchmark data can be used. The research claims made in previous stages needs to be satisfied with analysis well documented and reproducible.

MLTRL 3 is the system development: The Algorithm now does not stand on its own but is prepared to be integrated in a real-world scenario as part of MLTRL4. Therefore, the code must be developed towards interoperability, reliability, maintainability, extensibility and scalability to reach prototype character, data flow and interfaces must be considered. The data are in general consistent with MLTRL 2, but at the review, teammates from applied AI and engineering should be included.

The last MLTRL considered in this literature research is MLTRL 4, the Proof of Concept. This means, that the Technology is demonstrated in a real scenario with real and representative data. During the review, besides evaluating the data quality, validity and availability, security and privacy considerations should be done.

A more detailed description of the different Levels can be found in the original Paper from Lavin et al. [La22].

5 Literature Research

This chapter will describe the performed literature research. To focus only on Reinforcement Learning algorithms for simulation optimization, the keywords “Reinforcement Learning” and “simulation optimization” are included into the search. Furthermore, to focus on the area of production scheduling, one of the following keywords, which are introduced in the definition of production scheduling in Chapter 1, must be included: “job shop”, “open shop”, “flow shop”, “shop floor” or “production scheduling”. Therefore, the following search term was entered at Google Scholar to perform the literature research:

"Reinforcement Learning" AND "simulation optimization" AND (“job shop” OR “open shop” OR “flow shop”, “shop floor” OR “Production Scheduling”)

This leads to 202 results while only considering papers which were published 2022 or later. First, papers which were not accessible, duplicates and papers with a title that does not fit the topic being discussed, were sorted out. After this, 48 Papers were left. Of these Papers, 31 were sorted out after a more detailed review, because they did not use an RL-Algorithm, were not published in a peer viewed paper or conference paper or did not handle with a problem in a production setting. After this, 17 results were left.

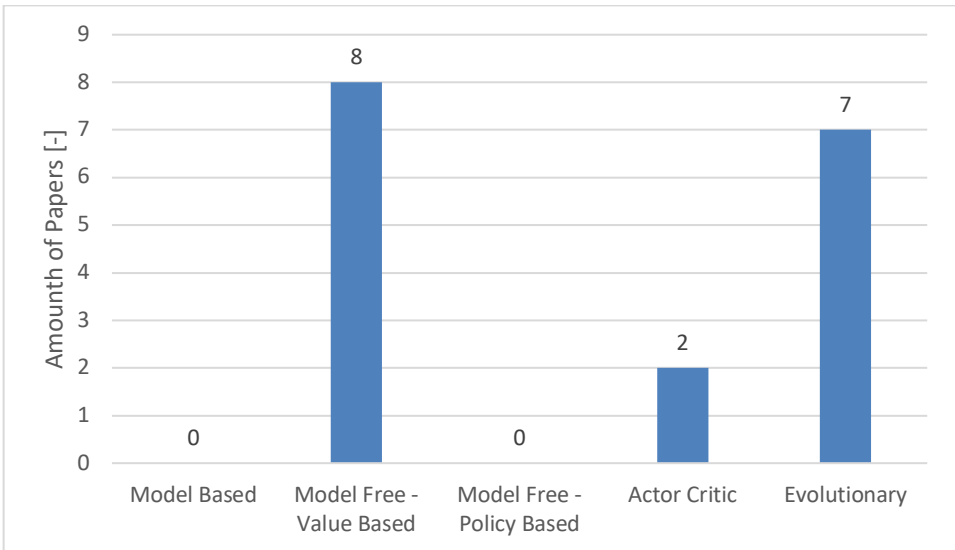


Fig. 3: Amount of papers per category.

As indicated in Fig. 3, no papers in which a model based or a model free policy-based algorithm was used could be found, but there are eight papers that deal with model free value-based algorithms, two with actor critic and seven with evolutionary algorithms. In the following Subchapter, the found papers are briefly summarized. After that, the literature is further analyzed to assign a MLTRL.

5.1 Summary

Devanga et al., Du et al., Schneckenreither et al., Wang et al., Inal et al., Wei et al. and Joo et al. have proven the principal of a model free value-based algorithms in a simulated test bed while using simulated data [DBD22, Du22, Sc22, Wa22, In23, We22, JJS22]. Furthermore, Kuhl et al. and Zhang et al. have developed a framework, how to implement a model free value-based algorithm in a real scenario. Kuhl et al. for a warehouse system [Ku22] and Zhang et al. to assign autonomous guided vehicles for material handling in a production logistic [Zh22]. Since they have not performed an experimental study, there are no data used.

Julati et al., as well as Song et al. have proven the principle of an actor critic algorithm in a simulated test bed while using simulated data. Julati et al. have scheduled in a high mix, low volume manufacturing facility [Ju22] and Song et al. in a biopharmaceutical production process [So23].

Regarding evolutionary algorithms, Aibi and Olfa, Bergmann, Cao et al., Ma et al. and Ghasemi et al. have proven the principle of an algorithm in a simulated test bed [ADE23,

Be22a, Ca22, MZS22, GKH22]. Wurster et al. have compared different algorithms in a simulated test bed [Be22b]. They used benchmark-, simulated and real-world data. Furthermore, Panigrahi et al. have developed a framework to integrate an evolutionary algorithm in a real scenario. The aim is the production scheduling in a semiconductor wafer fabrication [Pa22].

After this short summarization, in the next chapter an MLTRL will be assigned to the different algorithms, based on the information extracted from the literature.

5.2 Analysis of the Literature

To assign the MLTRL to the different algorithms, the methodology of the reviewed papers is illustrated in Fig. 4, ordered by algorithm.

Most of the papers deal with model free value-based algorithms or evolutionary algorithms (see also Fig. 2). While the majority of these papers prove the principle of the algorithm in a simulated test bed, for two model free value based and one evolutionary algorithm a framework for a real scenario was developed. For actor critic algorithms, the principle was only proven in a simulated test bed.

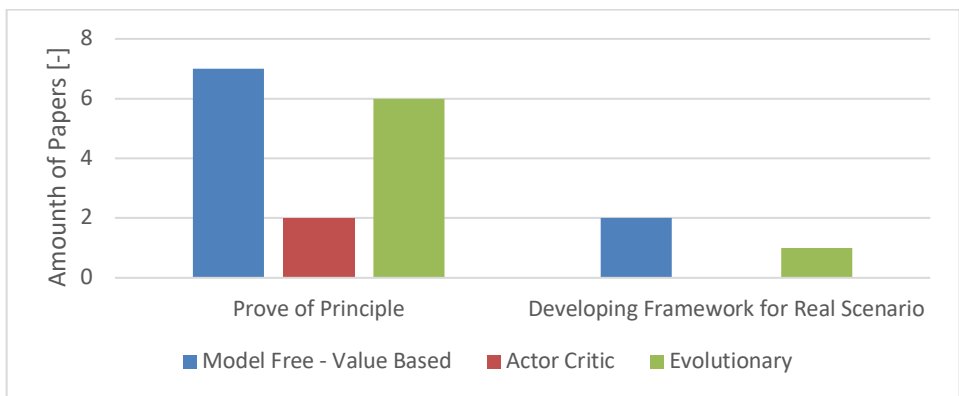


Fig. 4: Methodology of the reviewed Papers.

The Proof of Principle in a simulated test bed corresponds with a MLTRL 2. Furthermore, the peer viewed publication can be seen as a documented and reproducible documentation of the research claims and the use of simulated data also meets the requirements for MLTRL2. Therefore, all three algorithms can be seen as at least MLTRL 2.

The development of a framework for a real scenario is the requirement for MLTRL 3. Since there are no further requirements on the Data and developing the Framework requires the inclusion of applied AI and engineering, the authors of this paper argue to assign the model free value-based algorithms and the evolutionary algorithms an MLTRL 3.

As shown in Chapter 2, production scheduling is a very complex problem, especially if a lot of jobs have to be scheduled. To evaluate the possible performance, the papers in which a problem with at least 10^{10} possible solutions is handled are shown in more detail.

There are four papers with a problem with a complexity of more than 10^{10} possible solutions. The one with the smallest problem was written by Devanga et al. The performance of two value-based RL-methods, namely DDQN and DQN was compared to two dispatching rules, namely First-In-First-Out (FIFO) and Shortest Setup-and-Processing-Time (SSPT). 30 jobs have to be scheduled on one machine. Therefore, there are $30! = 2,65 * 10^{32}$ possible solutions. The agent assigns priority values to the incoming jobs which determine the sequence, in which the jobs are processed. The DDQN outperformed the DQN agent and the FIFO-dispatching rule. With 10 possible priority values, its performance was similar to SPT. Nevertheless, the Devanga et al. argue, that it has the potential to outperform SPT [DBD22].

Julati et al. have scheduled N independent jobs to M machines. The deep deterministic policy gradient with separate handling (SSDDPG) agent assigns a assignment priority to the tasks, and a free machine will process the job with the highest priority. This priority is composed of weights of different heuristics. The target is to minimize the delay past the due date, weighted by a job priority. The performance is compared to different dispatching rules, namely earliest weighted due date (EWDD), maximum mean on-time (MMOT) and shortest processing time (SPT). For the smallest problem, 35 Jobs are assigned to 10 Machines, which leads to a total of $10 * 35! = 1,03 * 10^{41}$ possible solutions, SSDDPG has a 14 % lower weighted delay than the best performing scheduling rule, which is MMOT. The biggest problem is to schedule 100 Jobs to 35 Machines, leading to a total of $35 * 100! = 3,3 * 10^{159}$ possible solutions. In this case, SSDDPG had also the best performance with a 5 % lower weighted delay than the best performing scheduling rule, which is also MMOT [Ju22].

Du et al. have used a DQN agent to schedule the crane, which transport the jobs between the machines. The agent chooses a dispatching rule, every time a transport procedure is necessary. In the most complex problem, 100 jobs and 10 machines have to be handled, leading to a complexity of $10 * 100! = 9,3 * 10^{158}$ possible solutions. The DQN agent outperforms a lot of different dispatching rules by far, for example FIFO, Machine Remaining Process Time (MRT) and Shortest Setup Time (SSU) to name just a few [Du22].

Ma et al. have used two evolutionary algorithms to schedule 2500 jobs to a maximum of 50 machines. The algorithms choose dispatching rules which are used to define the sequence, a machine processes the jobs in the queue. With 2500 jobs to schedule, it is obvious that the number of possible sequences tends to infinity. The best performing evolutionary algorithm leads to an 94 % reduction of tardy jobs compared to the best performing scheduling rule [MZS22].

6 Discussion

To support practitioners in deciding which Reinforcement Learning algorithm for simulation optimization for production scheduling is suitable for an implementation under real-life conditions, this paper firstly gives an overview over the topic of Reinforcement Learning and a classification, to order different algorithms. Namely, the categories model based-, model free value or policy based-, actor critic- and evolutionary algorithms are used. Since the subsequent literature research shows significant differences in the number of papers per category, it is assumed that the categories were chosen appropriately. Secondly, a short introduction into the MLTRL is given, which is an indicator of how ready a technology is for a commercial application. Afterward, the literature research was performed, to give an overview over the Technology Readiness of the different algorithms.

One criterion to assign an MLTRL is the performed review. For MLTRL2, a documented and reproducible achieving of the research claims is needed. Since a peer viewed publication can be seen as documented and reproducible, for this level the review criterion can easily be checked by a literature research. But for MLTRL3 this is not the case, since the expertise of the member of the review team is relevant. Therefore, it is assumed, that the expertise is needed to perform the corresponding methodology and to peer view the publication.

It also has to be mentioned, that a literature research cannot review in-house research and development of companies, since this information is usually confidential. So, a higher MLTRL than assigned in this paper cannot be excluded.

One result of the literature research is that no publications of model based or model free policy based algorithms could be found. One reason could be, that the literature research only takes publications into account, which were published in 2022 or later. This is a relatively short period of time, and it is possible, that there are publications which were not considered since they are older. This should be evaluated in further research. The same is for publications, in which a framework for a real scenario of an actor critic algorithm is developed. Nevertheless, by far the most publications deal with model free value based or evolutionary algorithms, so both seem to be in focus of the development. Possible reasons might be, that research has shown that other algorithms are not applicable to this problem, or just wasn't considered yet. In any case, an MLTRL3 is assigned to both algorithms, which means the next higher Technology Readiness Level can be achieved by demonstrating the technology in a real-life scenario.

When reviewing the paper which handle with a problem with at least 10^{10} possible solutions, these algorithms have comparable results to widely applied dispatching rules or are outperforming them. It is interesting to see, that none of these try to directly schedule a batch of Jobs but to assign priorities or choose a heuristic to be used to schedule the Jobs. For example in the Case of Devanga et al. 30 Jobs have to be scheduled on one machine [DBD22], which leads to $10! = 2,65 * 10^{32}$ possible solutions. If every incoming Job one of ten priority values is assigned, there are 10^{30} possible combinations of priority

values, which is nearly the same complexity as scheduling the Jobs directly. A direct scheduling of the Jobs seems to be possible, and its performance should be reviewed, either by literature research or by implementing and testing the algorithm.

It can also be said, that for all algorithms, for which papers were found, implementations performed better than the dispatching rules used as benchmark. A direct comparison of these algorithms cannot be done based on the literature, since the scheduling problems and benchmark-dispatching rules varying from paper to paper.

7 Conclusion

To implement a Reinforcement Learning algorithm for simulation optimization in production scheduling, the following findings of this paper should be considered:

- Current research focuses mainly on model free value based and evolutionary algorithms. Both algorithms a MLTRL 3 is assigned, therefore the next step in technology development is the demonstration in a real scenario.
- Very few publications could be found, which cover actor critic algorithms and none that cover model based or model free policy-based algorithms. The reason for this should be evaluated, since it is possible, that these algorithms are not applicable or just have not been considered yet.
- For the algorithm categories value based, evolutionary and actor-critic there is evidence, that all of these can perform significantly better than widely used dispatching rules. But there is no direct comparison between these algorithms.

Therefore, the recommended implementation strategy is to check, why there are so few recent papers which deal with actor critic and none which deal with model based and model free policy-based algorithms. Eventually, more research should be done to evaluate the usability of these algorithms for simulation optimization for production scheduling. If this has already been done, model free value based and evolutionary algorithms seem to be ready to be proven in a real scenario, so further research should be done to reach this goal. Also, a direct comparison between different RL-algorithms to evaluate their performance for different scheduling problems could be helpful to choose the best suited algorithm for a specific used case.

Acknowledgement

This work is supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) under grant No. 01MD22001C as part of the “edge data economy initiative”.

Literaturverzeichnis

- [ADE23] Aribi, D.; Driss, O. B.; El Haouzi, H. B.: Multi-Objective Optimization of the Dynamic and Flexible Job Shop Scheduling Problem Under Workers Fatigue Constraints. In (Rocha, A. P.; Steels, L.; van Herik, H. J. den Hrsg.): ICAART 2023. Proceedings of the 15th International Conference on Agents and Artificial Intelligence February 22-24, 2023, Lisbon. SciTePress - Science and Technology Publications, Setúbal, S. 301–308, 2023.
- [Be22a] Bergmann, S.: Optimization of the Design of Modular Production Systems. In (Feng, B. et al. Hrsg.): 2022 Winter Simulation Conference (WSC). 11-14 Dec. 2022. IEEE, Piscataway, NJ, S. 1783–1793, 2022.
- [Be22b] Behrendt, S. et al.: Extended Production Planning of Reconfigurable Manufacturing Systems by Means of Simulation-based Optimization. In (Herberger, D.; Hübner, M. Hrsg.): Proceedings of the Conference on Production Systems and Logistics: CPSL 2022. Hannover publish-Ing, S. 210–220, 2022.
- [Bl19] Blazewicz, J. et al.: Handbook on Scheduling. From Theory to Practice. Springer Nature Switzerland, Cham, 2019.
- [Ca22] Cao, Z. et al.: Two-stage genetic algorithm for scheduling stochastic unrelated parallel machines in a just-in-time manufacturing context. IEEE Transactions on Automation Science and Engineering 2/20, S. 936–949, 2022.
- [DBD22] Devanga, A.; Badilla, E. D.; Dehghanimohammadabadi, M.: Applied Reinforcement Learning for Decision Making in Industrial Simulation Environments. In (Feng, B. et al. Hrsg.): 2022 Winter Simulation Conference (WSC). 11-14 Dec. 2022. IEEE, Piscataway, NJ, S. 2819–2829, 2022.
- [Di20] Ding, Z. et al.: Chapter 2. Introduction to Reinforcement Learning. In (Dong, H.; Ding, Z.; Zhang, S. Hrsg.): Deep Reinforcement Learning. Fundamentals, Research and Applications. Springer Singapore; Imprint Springer, Singapore, 47 - 122, 2020.
- [Du22] Du, Y. et al.: A reinforcement learning approach for flexible job shop scheduling problem with crane transportation and setup times. Transactions on Neural Networks and Learning Systems, 1-15, 2022.
- [Fr18] François-Lavet, V. et al.: An Introduction to Deep Reinforcement Learning. Foundations and Trends® in Machine Learning 3-4/11, S. 219–354, 2018.
- [GKH22] Ghasemi, A.; Kabak, K. E.; Heavey, C.: Demonstration of the Feasibility of Real Time Application of Machine Learning to Production Scheduling. In

- (Feng, B. et al. Hrsg.): 2022 Winter Simulation Conference (WSC). 11-14 Dec. 2022. IEEE, Piscataway, NJ, S. 3406–3417, 2022.
- [Gy22] Gya, R. et al.: Digital Twins. Adding intelligence to the real world. https://www.capgemini.com/gb-en/wp-content/uploads/sites/3/2022/05/Capgemini-Research-Institute_DigitalTwins_Web.pdf, Stand: 24.4.2022.
- [In23] İnal, A. F. et al.: A Multi-Agent Reinforcement Learning Approach to the Dynamic Job Shop Scheduling Problem. *Sustainability* 10/15, S. 1–24, 2023.
- [JJS22] Joo, T.; Jun, H.; Shin, D.: Task Allocation in Human–Machine Manufacturing Systems Using Deep Reinforcement Learning. *Sustainability* 4/14, S. 1–18, 2022.
- [Ju22] Julaiti, J. et al.: Stochastic parallel machine scheduling using reinforcement learning. *Journal of Advanced Manufacturing and Processing* 4/4, 1-17, 2022.
- [KD21] Kumar, A.; Dimitrakopoulos, R.: Production scheduling in industrial mining complexes with incoming new information using tree search and deep reinforcement learning. *Applied Soft Computing* 110, S. 1–15, 2021.
- [Ku22] Kuhl, M. E. et al.: Warehouse Digital Twin: Simulation Modeling and Analysis Techniques. In (Feng, B. et al. Hrsg.): 2022 Winter Simulation Conference (WSC). 11-14 Dec. 2022. IEEE, Piscataway, NJ, S. 2947–2956, 2022.
- [La17] Lamghari, A.: Mine Planning and Oil Field Development: A Survey and Research Potentials. *Mathematical Geosciences* 3/49, S. 395–437, 2017.
- [La22] Lavin, A. et al.: Technology readiness levels for machine learning systems. *Nature communications* 1/13, S. 1–19, 2022.
- [LTD22] Luo, D.; Thevenin, S.; Dolgui, A.: A state-of-the-art on production planning in Industry 4.0. *International Journal of Production Research* 19/61, S. 6602–6632, 2022.
- [Mn15] Mnih, V. et al.: Human-level control through deep reinforcement learning. *Nature* 7540/518, S. 529–533, 2015.
- [Mn16] Mnih, V. et al.: Asynchronous Methods for Deep Reinforcement Learning. In (Balcan, M. F.; Weinberger, K. Q. Hrsg.): *Proceedings of The 33rd International Conference on Machine Learning*. PMLR, New York, New York, USA, S. 1928–1937, 2016.
- [Mo20] Mourtzis, D.: Simulation in the design and operation of manufacturing systems: state of the art and new trends. *International Journal of Production Research* 7/58, S. 1927–1949, 2020.

- [MZS22] Ma, H.; Zhang, C.; Shi, Z.: A Simulation Optimization-Aided Learning Method for Design Automation of Scheduling Rules: 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE). 20-24 Aug. 2022. IEEE, Piscataway, NJ, S. 1992–1997, 2022.
- [Pa22] Panigrahi, S. et al.: Production Scheduling of Semiconductor Wafer Fabrication Facilities Using Real-Time Combinatorial Dispatching Rule. In (Ciobață, D. D. Hrsg.): International Conference on Reliable Systems Engineering (ICoRSE)-2021 // International Conference on Reliable Systems Engineering (ICoRSE) - 2021. Springer International Publishing AG, Cham, S. 78–90, 2022 // 2021.
- [PBG22] Panzer, M.; Bender, B.; Gronau, N.: Neural agent-based production planning and control: An architectural review. *Journal of Manufacturing Systems* 65, S. 743–766, 2022.
- [Ra17] Racanière, S. et al.: Imagination-Augmented Agents for Deep Reinforcement Learning. In (Luxburg, U. von et al. Hrsg.): *Advances in neural information processing systems* 30. 31st Annual Conference on Neural Information Processing Systems (NIPS 2017) Long Beach, California, USA, 4-9 December 2017. Curran Associates Inc, Red Hook, NY, S. 5694–5705, 2017.
- [Ri12] Righi, R. d. R.: Preface. In (Righi, R. d. R. Hrsg.): *Production Scheduling*. InTech, S. X, 2012.
- [Sa17] Salimans, T. et al.: Evolution Strategies as a Scalable Alternative to Reinforcement Learning. ArXiv 1703.03864v2, 2017.
- [SB18] Sutton, R. S.; Barto, A.: *Reinforcement learning. An introduction*. The MIT Press, Cambridge, Massachusetts, London, England, 2018.
- [Sc07] Schulz, A. et al.: Simulation in der operativen Produktionsplanung – Erfolgsfaktoren für KMU. *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 1-2/102, S. 32–36, 2007.
- [Sc17] Schulman, J. et al.: Proximal Policy Optimization Algorithms. ArXiv 1707.06347, 2017.
- [Sc22] Schneckenreither, M. et al.: Average reward adjusted deep reinforcement learning for order release planning in manufacturing. *Knowledge-Based Systems* 247, S. 1–16, 2022.
- [Si16] Silver, D. et al.: Mastering the game of Go with deep neural networks and tree search. *Nature* 7587/529, S. 484–489, 2016.

- [So23] Song, W. et al.: Stochastic Economic Lot Scheduling via Self-Attention Based Deep Reinforcement Learning. *IEEE Transactions on Automation Science and Engineering*, S. 1–12, 2023.
- [Wa22] Wang, X. et al.: Digital Twin-Assisted Efficient Reinforcement Learning for Edge Task Scheduling: 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring). *IEEE*, S. 1–5, 2022.
- [We22] Wei, Q. et al.: A Self-Attention-Based Deep Reinforcement Learning Approach for AGV Dispatching Systems. *IEEE Transactions on Neural Networks and Learning Systems*, S. 1–12, 2022.
- [Zh22] Zhang, L. et al.: Reinforcement learning and digital twin-based real-time scheduling method in intelligent manufacturing systems. *IFAC-PapersOnLine* 10/55, S. 359–364, 2022.
- [ZY20] Zhang, H.; Yu, T.: Chapter 3. Taxonomy of Reinforcement Learning Algorithms. In (Dong, H.; Ding, Z.; Zhang, S. Hrsg.): *Deep Reinforcement Learning. Fundamentals, Research and Applications*. Springer Singapore; Imprint Springer, Singapore, S. 125–133, 2020.

Sensor-based modelling of carbon dioxide emissions for trailer traffic


Justus Leskow ¹, Steffen Greiser ² and Goy-Hinrich Korn³


Abstract: In this research work, the calculation of carbon dioxide emissions for trailer traffic based on sensor data is described. The data consist of trailer type, load, trip and route parameters. While the individual trailer parameters can be derived from the order management, the trip and route parameters are collected by frequently data. With this, the transport carbon footprint is calculated based on DIN EN 16258 by means of the *ProBas* database. The distributions of carbon dioxide emissions in trailer traffic are analysed and discussed. The results of the case study indicate that sensor-based modelling can be a useful tool for an improved transparency and estimation of carbon dioxide emissions.

Keywords: transport carbon footprint, trailer, telemetry data, logistics

1 Introduction

Anthropogenic climate change is currently one of the greatest threats to humanity. The earth is warming up due to anthropogenic carbon dioxide-equivalent (CO_{2e}) emissions and the resulting increase in the greenhouse effect. [RF09], [NO23]. The associated decline in agricultural yields, natural disasters, and water shortages are just a few of the causes leading to international conflicts [MKK21], [UB21]. In order to counteract these negative effects in the best possible way, savings in CO_{2e} emissions are required. Globally, the transport sector contributed to 20 % of CO_{2e} emissions in 2019 [St23]. Due to the restrictions during the corona pandemic, CO_{2e} emissions have temporarily decreased significantly in the year 2020. In the area of road transport, this has even halved [Le20]. Since within the transport sector, road transport is responsible for 73% of global CO_{2e} emissions in 2018, this sector seems particularly suitable for considering improvements [La21]. While the reduction of CO_{2e} emissions due to the corona pandemic has been of short-term importance, long-term savings in the transport sector are mainly possible through more efficient fuels, means of transport or modal shifts [BI20]. For this kind of

¹ University of Applied Sciences Osnabrück, Institute of Management and Technology, Kaiserstr. 10c, 49809 Lingen, Justus.Leskow@hs-osnabrueck.de,  <https://orcid.org/0009-0000-1875-6947>

² University of Applied Sciences Osnabrück, Institute of Management and Technology, Kaiserstr. 10c, 49809 Lingen, s.greiser@hs-osnabrueck.de,  <https://orcid.org/0009-0009-5609-6180>

³ Bernard Krone Holding GmbH & Co. KG, CIO & CDO, 48480 Spelle, Heinrich-Krone-Straße 10, goy-hinrich.korn@krone.de

more efficient transport management, which, for example, protects the environment by avoiding traffic congestions, the collection, analysis, and use of data is the basis [RM22]. Therefore, enhancement in transport efficiency and CO₂e emission reduction is mostly achieved by exploiting telemetry data, that became available due to Global Positioning System (GPS) traces from cell phones, vehicles and anonymized Call Detail Records (CDR) from cell phone providers [Ma19]. Based on these data, it is possible to develop models that represent the spatiotemporal CO₂e emissions [HZJ20].

2 Problem Statement and Scientific Question

Under consortium leadership of the Krone Group, Spelle with partners from industry and science, the three-year *EDNA* project (grant No.: 01MD22001C; funding program: "EDGE Datenwirtschaft") funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) is exploring the potential of digital technologies for a more sustainable production in the transport and logistics sector (www.edna-projekt.de). One aspect of this project is the previously described calculation of the transport carbon footprint (TCF) caused by trailer traffic.

The trailers of Fahrzeugwerk Bernard KRONE GmbH & Co. KG regularly transmit telemetry data. Figure 1 shows a section of telemetry data from a trip. These data are periodically transmitted and contains, among other things, the location and movement data as well as the trailer type. The data discussed in this paper have been anonymised and contain only one customer of Krone fleet, the trailer rental service of the Krone Group. These data have been explicitly approved by [Kemena GmbH](#) and kindly provided for the analysis in this paper.

	‡ BD_GPS_LATITUDE	‡ BD_COUPLED	‡ 67	‡ BD_DOOR_OPEN	‡ 69	‡ BD_GPS_LONGITUDE	‡ BD_GPS_LOCATION	‡ 83	‡ 84	‡ BD_GPS_DIRECTION
634	53.08438	True	nan	False	nan	8.73477	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	0
635	53.08437	True	nan	False	nan	8.73480	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	0
636	53.08439	True	nan	False	nan	8.73480	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	0
637	53.08437	True	nan	False	nan	8.73488	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	0
638	53.08438	True	nan	False	nan	8.73482	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	0
639	53.08438	True	nan	False	nan	8.73483	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	0
640	53.08436	True	nan	False	nan	8.73480	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	136
641	53.08437	True	nan	False	nan	8.73480	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	242
642	53.08437	True	nan	False	nan	8.73488	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	318
643	53.08434	True	nan	False	nan	8.73494	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	5
644	53.08439	True	nan	False	nan	8.73488	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	5
645	53.08438	True	nan	False	nan	8.73491	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	5
646	53.08435	True	nan	False	nan	8.73494	Georg-Henschel-Straße 3, 28197 Bremen, DE	nan	nan	5
647	53.08032	True	nan	False	nan	8.73605	Bremen Georg-Henschel-Straße, 28197 Bremen, DE	nan	nan	69
648	53.07462	True	nan	False	nan	8.76165	A281, 28197 Bremen, DE	nan	nan	125
649	53.08601	True	nan	False	nan	8.79601	Landwehrstraße 4, 28217 Bremen, DE	nan	nan	39
650	53.09567	True	nan	False	nan	8.79471	Parallelweg 49, 28219 Bremen, DE	nan	nan	195
651	53.05992	True	nan	False	nan	8.76180	Emsländstraße 22, 28259 Bremen, DE	nan	nan	243
652	53.04625	True	nan	False	nan	8.70130	B75, 27751 Delmenhorst, DE	nan	nan	259
653	53.03507	True	nan	False	nan	8.62124	Schluttenweg 6, 27755 Delmenhorst, DE	nan	nan	273
654	53.05779	True	nan	False	nan	8.52816	E22, 27777 Ganderkesee, DE	nan	nan	275
655	53.05926	True	nan	False	nan	8.42537	E22, 27798 Hude, DE	nan	nan	264
656	53.07263	True	nan	False	nan	8.32730	E22, 26209 Hatten, DE	nan	nan	315
657	53.12310	True	nan	False	nan	8.27532	A29, 26135 Oldenburg, DE	nan	nan	5
658	53.18635	True	nan	False	nan	8.25922	A29, 26125 Oldenburg, DE	nan	nan	332
659	53.23211	True	nan	False	nan	8.18460	A29, 26180 Rastede, DE	nan	nan	318
660	53.29376	True	nan	False	nan	8.14751	A29, 26180 Rastede, DE	nan	nan	348
661	53.35885	True	nan	False	nan	8.12281	Aeropark 1, 26316 Varel, DE	nan	nan	354
662	53.41644	True	nan	False	nan	8.08264	A29, 26316 Varel, DE	nan	nan	307
663	53.46425	True	nan	False	nan	8.01762	A29, 26453 Sande, DE	nan	nan	354
664	53.52688	True	nan	False	nan	8.00755	A29, 26419 Schortens, DE	nan	nan	36

Figure 1: Snapshot of a telemetry data frame

This telemetry data is not yet used for the consideration of CO₂e emissions. Theoretically, CO₂e emissions can be derived manually from the fuel consumption of logistics service providers. However, as the CO₂e emissions are to be calculated automatically, this method is not suitable for this purpose. The motivation on the part of Fahrzeugwerk Bernard KRONE GmbH & Co. KG is to make logistics more sustainable and to limit and monitor the emissions generated in the process. The first step towards achieving this goal is to determine CO₂e emissions. Therefore, this paper deals with the implementation of a transport carbon footprint calculator using telemetry data.

This paper is structured as follows. First, the general methodology of the transport carbon footprint calculation is explained. This is subsequently illustrated by means of an example. Next, the telemetry data is explained in more detail and the calculation of the transport carbon footprint based on this data is presented. The paper concludes with a summary and an outlook for further studies.

3 Methods and Techniques

DIN EN 16258, developed by the German Institute for Standardisation, includes not only specific calculation and allocation procedures, but also basic definitions as well as regulations regarding system boundaries [De13]. In this context, the standard serves "to calculate and declare energy consumption and greenhouse gas emissions for any transport service" [De13]. The basic requirement is that energy quantities are reported in Joule and Greenhouse Gas (GHG) emissions in gram CO₂e or multiples thereof [De13]. For the evaluation of a system, all consumptions and emissions during the system's usage are to be included [De13]. This involves, among other things, the expenses for the on-board electronics, the propulsion power as well as any additional services required, such as the operation of cooling units [De13]. This applies to both, the loading journeys and any resulting empty journeys [De13]. Additionally, fuels as well as electricity and their corresponding emissions from extraction to provision must be considered [De13]. Excluded from the system boundaries are, for example, emissions originated by spills or short-term support such as navigation by tugboats - an overview of all exceptions can be found in Ref. [De13]. Emissions trading and compensation cannot be a direct part of the transport carbon footprint, so that these are not considered in the calculations. [De13].

The derivation of the transport carbon footprint is divided into three steps. First, the different partial routes and the means of transport used are to be identified based on the activity data. [De13]. The next step is the determination of the following values for each lag of the journey: Well-to-wheel energy consumption, Well-to-wheel GHG emissions, Tank-to-wheel energy consumption, Tank-to-wheel GHG emissions [De13]. In addition to the activity data, emission factors are also required for this. The activity data includes the distance, the truck type and the weight of the load, which finally leads to the fuel consumption [De13]. Ideally, these data together with emission factors are determined by real-time measurements. However, the use of default values from databases is well-aligned with norms and guidelines [De13].

The aforementioned emission values could be calculated as follows:

1. $E_w = F * e_w$ ⁴
2. $G_w = F * g_w$ ⁵
3. $E_t = F * e_t$ ⁶
4. $G_t = F * g_t$ ⁷

The third substep of the calculation includes the addition of the partial results to one total result per value [De13]. However, if statements are not to be made for an entire transport unit, but only for a transported good, allocations are to be carried out [De13]. Results according to DIN EN 16258 may be publicly reported to increase transparency about the CO₂e emissions. These analysis results of the transport carbon footprint may be published in a declaration [De13]. The prescribed formulation of such a declaration can be found with more details in DIN EN 16258 on pages 20 and 21.

DIN EN 16258 specifies the average fuel consumption of a truck by 35 litres per 100 km. Thus, the fuel consumption F can be calculated from the distances travelled and the average fuel consumption. This results in emission values (TCF) for the route, as exemplary shown for three sections of a route in Table 2 (each section is 49, 32 and 23 km long).

4 Application and Results

By adapting this procedure, CO₂e emissions can now be calculated using telemetry data. Since the transmitted telemetry data contain values for the trailers' type and load, it is possible to determine the trucks' load. This load in turn enables databases such as *ProBas* (<https://www.probas.umweltbundesamt.de/php/index.php>) to include more granular emission factors in the calculation of the TCF. The solution presented in this paper uses the *ProBas* database, which yields emission factors for truck utilisation intervals of 10 %. For example, for the truck types used in this paper, *ProBas* gives the values as presented in Table 3 [Um23d], [Um23e], [Um23c], [Um23a], [Um23b]. Since the value of e_t is not available in the *ProBas* database, the value of DIN EN 16258 is still used for this calculation.

For the calculation of the distance travelled by the trucks, the GPS coordinates are converted into a Geo-JSON file. This Geo-JSON file is used to determine the distance

⁴ E_w = Well-to-Wheel-Energy-Consumption (MJ), F = Fuel consumption (l), e_w = Well-to-Wheel-Energy-Factor (fuel) (MJ/l)

⁵ G_w = Well-to-Wheel-Greenhousegas-Emissions (kg CO₂e), g_w = Well-to-Wheel-Emission-Factor (fuel) (kg CO₂e/l)

⁶ E_t = Tank-to-Wheel-Energy-Consumption (MJ), e_t = Tank-to-Wheel-Energy-Factor (fuel) (MJ/l)

⁷ G_t = Tank-to-Wheel-Greenhousegas-Emissions (kg CO₂e), g_t = Tank-to-Wheel-Emission-Factor (fuel) (kg CO₂e/l)

travelled via the API of *Openrouteservice* (<https://maps.openrouteservice.org/>). Figure 2 shows the 137 km route travelled as measured by telemetry GPS coordinates based on the *Openrouteservice* API.

Based on these distances, the CO₂e emissions can be calculated with the load data, truck type and the average fuel consumption. Exemplary results are shown in Figure 3 for the CO₂e emissions of one truck for four trips, whereas the purple curve shows the CO₂e emissions for the distance travelled in Figure 2. Accordingly, it is now possible to analyse the CO₂e emissions along the route.

Mode of transport	e_t (MJ/l)	e_w (MJ/l)	g_t (kg CO ₂ e/l)	g_w (kg CO ₂ e/l)
Truck	35,9	42,7	2,67	3,24

Table 1: Energy and emission factors according to DIN EN 16258

Route	E_w (MJ)	G_w (kg CO ₂ e)	E_t (MJ)	G_t (kg CO ₂ e)
Part 1	732,305	55,566	615,685	45,791
Part 2	478,24	36,288	402,08	29,904
Part 3	343,735	26,082	288,995	21,494
Total	1554,28	117,94	1306,76	97,19

Table 2: Exemplary calculation of the TCF

Utilization rate	e_t (MJ/1000km)	e_w (MJ/1000km)	g_t (kg CO ₂ e/1000km)	g_w (kg CO ₂ e/1000km)
0 %	-	7,49	0,554	0,643
10 %	-	4,22	0,312	0,362
30 %	-	1,59	0,118	0,137
50 %	-	1,07	0,0789	0,0915
70 %	-	0,839	0,0621	0,072

Table 3: Selected emission factors in relation to the utilisation rate

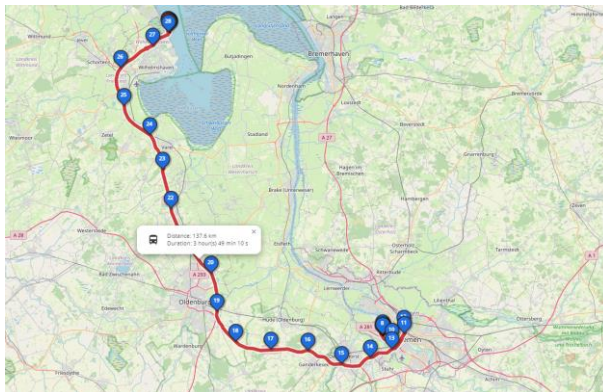


Figure 2: Exemplary travelled Route

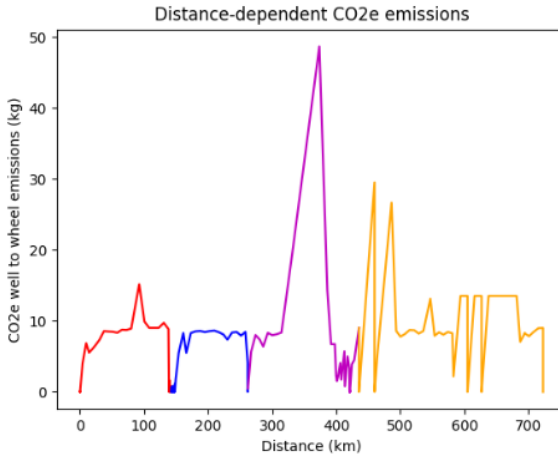


Figure 3: Distance dependent CO₂e emissions

5 Summary and Outlook

In this paper, a model has been proposed which can determine the CO₂e emissions of trailers based on telemetry data. In this proposed model, the routes travelled by the trailers are calculated using an API from Openrouteservice together with the GPS coordinates gathered by telemetry. The CO₂e emissions are then calculated according to DIN EN 16258, considering route length, truck type, load factor and fuel consumption. This will improve the transparency of CO₂e emissions for the logistics of the Krone Group and its customers in the future.

In the further development, an extension of this model will consider the elevation profile of the route and a model-based calculation of the fuel consumption based on the trailer data to further improve the quality of the transport carbon footprint calculation. This enables an even more precise representation of CO₂e emissions and therefore a realistic transport carbon footprint dashboard of trailer logistics in real-time. In addition, it will be possible to offer the service of CO₂e emissions monitoring to customers in the future.

Acknowledgements

The authors would like to thank Kemena GmbH for providing the telemetry data in compliance with data protection. The first author is grateful for the contribution of Leon Lelle, Johannes Rosen and Kieron Stegemann (University of applied Sciences in Osnabrück) in the development of the CO₂e calculator.

Bibliography

- [Bl20] Blok, K. et al.: Assessment of Sectoral Greenhouse Gas Emission Reduction Potentials for 2030, 2020.
- [De13] Deutsches Institut für Normung e. V. (Hrsg.): DIN EN 16258:2013-03, Methode zur Berechnung und Deklaration des Energieverbrauchs und der Treibhausgasemissionen bei Transportdienstleistungen (Güter- und Personenverkehr); Deutsche Fassung EN_16258:2012. Beuth Verlag GmbH, Berlin, 2013.
- [HZJ20] He, Z.; Zhang, W.; Jia, N.: Estimating Carbon Dioxide Emissions of Freeway Traffic: A Spatiotemporal Cell-Based Model. *IEEE Transactions on Intelligent Transportation Systems* 5/21, pp. 1976–1986, 2020.
- [La21] Lamb, W. F. et al.: A review of trends and drivers of greenhouse gas emissions by sector from 1990 to 2018. *Environmental Research Letters* 7/16, p. 73005, 2021.
- [Le20] Le Quéré, C. et al.: Temporary reduction in daily global CO₂ emissions during the COVID-19 forced confinement. *Nature Climate Change* 7/10, pp. 647–653, 2020.
- [Ma19] Markovic, N. et al.: Applications of Trajectory Data From the Perspective of a Road Transportation Agency: Literature Review and Maryland Case Study. *IEEE Transactions on Intelligent Transportation Systems* 5/20, pp. 1858–1869, 2019.
- [MKK21] Malhi, G. S.; Kaur, M.; Kaushik, P.: Impact of Climate Change on Agriculture and Its Mitigation Strategies: A Review, 2021.
- [NO23] NOAA Climate.gov: Global Temperature Anomalies - Graphing Tool. <https://www.climate.gov/maps-data/dataset/global-temperature-anomalies-graphing-tool>, accessed 30 Apr 2023.
- [RF09] Ramanathan, V.; Feng, Y.: Air pollution, greenhouse gases and climate change: Global and regional perspectives. *Atmospheric Environment* 1/43, pp. 37–50, 2009.
- [RM22] Ravi, S.; Mamdakar, M. R.: A Review on ITS (Intelligent Transportation Systems) Technology: 2022 International Conference on Applied Artificial Intelligence and Computing (ICAIC). *IEEE*, pp. 155–159, 2022.
- [St23] Statista: CO₂-Ausstoß weltweit nach Sektoren | Statista. <https://de.statista.com/statistik/daten/studie/167957/umfrage/verteilung-der-co-emissionen-weltweit-nach-bereich/>, accessed 2 May 2023.
- [UB21] Uexkull, N. von; Buhaug, H.: Security implications of climate change: A decade of scientific progress, 2021.
- [Um23a] Umweltbundesamt: ProBas - Prozessdetails: LKW oder Lastzug. <https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={11E18C8C-AAB0-5835-A30C-70023B2A8D21}>, accessed 13 May 2023.
- [Um23b] Umweltbundesamt: ProBas - Prozessdetails: LKW oder Lastzug. <https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={11E18C8C-AAB0-5830-A30C-90023B2A8D20}>, accessed 13 May 2023.

- [Um23c] Umweltbundesamt: ProBas - Prozessdetails: LKW oder Lastzug. <https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={11E18C8C-AAB0-5835-A30C-70023B2A8D27}>, accessed 13 May 2023.
- [Um23d] Umweltbundesamt: ProBas - Prozessdetails: LKW oder Lastzug. <https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={11E18C8C-AAB0-5835-A30C-70023B2A8D23}>, accessed 13 May 2023.
- [Um23e] Umweltbundesamt: ProBas - Prozessdetails: LKW oder Lastzug. <https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={11E18C8C-AAB0-5835-A30C-70023B2A8D25}>, accessed 13 May 2023.

Enhancing Digital Twins for Production through Process Mining Techniques: A Literature Review

Marcel Schumacher¹, Ralf Buschermöhle¹, Liane Haak¹, Max Höfinghoff¹, Arne Seipolt¹, Goy-Hinrich Korn²

Abstract: A digital twin (DT) plays a vital role in the advancement of manufacturers towards Industry 4.0. However, the creation and maintenance of DTs can be time-consuming. One approach to streamline this process is the utilization of process mining (PM) methods and techniques, which can automatically generate valuable information for DTs. Therefore, this paper aims to examine different approaches that augment DTs with PM and explore their effects. The review categorizes these approaches into three groups: theoretical approaches, approaches with laboratory case studies, and approaches with real-world case studies conducted by manufacturers. The review reveals that the use of PM can enhance the flexibility and sustainability of DTs. However, this improvement comes at the cost of requiring high-quality data and more data preparation efforts.

Keywords: Process Mining, Digital Twin, Production

1 Introduction

The creation of a digital twin (DT) is time-consuming [Br21] but at the same time the DT is an essential component for manufacturers advancing to industry 4.0 [LM21]. In order to accurately represent real-world objects, such as the production process, a DT could rely on a process model [Br21]. Traditionally, the manual creation or updating of process models is both time-consuming and susceptible to errors. To overcome these challenges, Process Mining (PM) techniques can be employed to automatically create and update process models using event data. Event data is generated by performing the process and stored in various information systems like enterprise resource planning systems (ERP-Systems), manufacturing execution system (MES), programmable logic controller (PLC), etc. [va16].

The objective of this paper is to provide an overview of different approaches for augmenting DTs using PM and examine their effects on DTs. To achieve this, a comprehensive review of the existing literature will be conducted, categorizing the approaches into three distinct categories. The first category comprises theoretical approaches that have not been tested in any case studies. The second category includes

¹ Hochschule Osnabrück, Fakultät Management, Kultur und Technik, Kaiserstraße 10c, 49809 Lingen (Ems), marcel.schumacher@hs-osnabrueck.de; r.buschermoehle@hs-osnabrueck.de; l.haak@hs-osnabrueck.de; m.hoefinghoff@hs-osnabrueck.de; a.seipolt@hs-osnabrueck.de

² Bernard Krone Holding GmbH & Co. KG, CIO & CDO, 48480 Spelle, Heinrich-Krone-Straße 10, goy-hinrich.korn@krone.de

approaches that have been tested in a laboratory environment, and the third category consists of approaches that have been tested in real-world manufacturing companies.

The paper is structured as follows: Chapter 2 begins by introducing a definition of DTs. Chapter 3 outlines the fundamentals of PM. Chapter 4 explains the research methodology employed in this study. Chapter 5 presents the review itself, providing an analysis of the different approaches. The final chapter initiates the discussion based on the findings and concludes the paper by summarizing the key insights gained from the review.

2 Digital Twin

The concept of a digital twin (DT) encompasses various definitions that vary depending on the specific domain they originate from [Be21]. One definition that encapsulates the essence of a DT in the manufacturing context is as follows: “a digital twin is an integrated multi-physics, multi-scale, probabilistic simulation of a complex product and uses the best available physical models, sensor updates, etc., to mirror the life of its corresponding twin.” [Ta18, p. 3564] This definition highlights the key characteristics of a DT in the production domain. In manufacturing, the primary technologies employed for digital twins include simulation techniques such as discrete event simulation and continuous simulation, communication protocols, and core technologies associated with Industry 4.0 [Kr18]. These technologies enable the creation of a virtual representation that closely mirrors the behavior and lifecycle of the physical counterpart, allowing for advanced analysis and optimization within the manufacturing process. There are multiple benefits from using DTs in the production. For Example, the DT supports the planning and controlling [Ro15], mainting [DUM17] and the layout planning of the production [ULS17]. The next chapter explains the theoretical knowledge of PM and the different ways PM can be used to enhance DTs.

3 Process Mining

Process mining (PM) is focused on capturing and analyzing the workflow of a process based on its digital footprints. When processes are executed in the real world, they often leave behind data that can be considered as event data. This event data can be transformed into an event log, which is a compilation of event data organized by cases, sequences, and activities. For instance, let's consider an example of an order being processed for a customer. All the events associated with that order, such as order placement, order confirmation, production, packaging, and shipping, would be grouped together as a case in the event log. Each event would have a specific order in which it occurred within the overall process. By analyzing this event log, process mining techniques can extract valuable insights about the process, such as the sequence of activities, bottlenecks, and variations. This allows for a comprehensive understanding of the process and facilitates process improvement and optimization efforts [vWM04]. PM can be categorized into three

types, each serving a distinct purpose within the analysis of event logs: The first type is process discovery, which aims to create a process model based on a given event log, utilizing discovery algorithms [va10]. By analyzing the event log, these algorithms uncover the underlying process flow and dependencies, resulting in a process model that represents the observed behavior. The second type is conformance checking, which involves comparing a process model against an event log to evaluate its fitness and precision [va16]. This comparison helps identify any deviations or non-conformances between the model and the actual process execution. Furthermore, conformance checking can also be used to compare two process models, enabling the detection of differences between them [va05]. The third and last type of PM is process enhancement. The process enhancement aims to repair or extend the process model with new perspectives [va12]. The most common ones are the time, organizational and case. All of them show a different view of the process[va12]. The organizational represents the organizational structure, the social network, the role and the behavior of the resources [Sv08]. The timley exposes frequencies, process time, service time, waiting time, etc. of the process and allows to find bottlenecks in the process [va12]. The case perspective uncovers the rules behind a decision in the process and makes them visible as a decision tree [Rv06]. The next chapter explains the used research methodology for the paper and which paper were chosen for the review.

4 Research Methodology

The objective of the papers is to provide an overview of the various approaches of process mining in digital twins specifically within the manufacturing domain. The review follows the structure recommended by Kitchenham et al. in 2013 [KB13] to address the following research question: How could PM enhance DTs in the production and which effects result subsequently?

In order to answer this research question, a systematic approach was taken to gather relevant scientific papers from diverse scientific databases. The details of the search string and library databases used are presented in Table 1. It is important to note that the total number of papers includes the results obtained from multiple databases.

Library	Search in	Search string	Result
IEEE	Metadata	“process mining”	17
Science Direct	Metadata	AND “digital twin”	4
Google Scholar	Title & Keywords	AND “production”	87

Tab. 1: Overview library databases

The total of 108 papers were manual sorted with the help of exclusion (EC) and inclusion criteria (IC) like Kitchenham suggested [KB13].

- EC1: Papers aren't available in English.
- EC2: Papers aren't accessible by the authors.
- EC3: Papers aren't in the context of production or manufacturing
- IC1: Papers discuss a DT with PM as a component

Out of the 108, only 13 satisfy the IC and EC and are checked for their quality. The quality check secures that only paper with a high quality are used for the review, several quality criteria (QC) are introduced to give each paper a quality score. The score reflects if a paper fully (1), partly (0,5), or not (0) meets the criteria.

- QC1: Does the paper explain how PM enhances the DT?
- QC2: Does the paper explain why PM is used for the DT?
- QC3: Does the paper use its own approach for PM and DT?
- QC4: Does the paper explain the effects of PM for the augmentation of the DT?

Paper	QC1	QC2	QC3	QC4	Total
[PLN22]	1	0	0	0	1
[RA20]	1	1	1	1	4
[Tr21]	1	1	1	1	4
[KNB22]	1	1	1	1	4
[Fr22]	1	1	1	1	4
[MLC22]	1	1	1	0	3
[LM21a]	0	0	0	0	0
[NV22]	1	1	1	1	4
[Br21]	1	1	1	1	4
[CAP21]	1	1	1	0	3
[Ya22]	0	0	0	0	0
[Fu23]	1	1	0	0	2

Tab. 2: Quality criteria results

The score of each paper after applying the quality criteria can be viewed in the table 2. If a paper doesn't reach a total score of at least 3, it will be excluded from the review. Only 9 paper could reach the needed quality score of at least three. These papers are reviewed and presented in the next chapter.

5 Literature Review

This chapter explains in three subchapters the different ways PM is used to augment a DT.

5.1 Theoretical approaches

Brockhoff presents in his paper [Br21] an approach for the creation of a digital twin. According to their interpretation, the DT serves the purpose of predicting problems while a process is in operation. They adopt a model-driven approach to generate the DT, and PM techniques are employed to create process models from the stored information within the DT. The discovered process models are stored in the DT and utilized in conjunction with conformance checking methods to determine any kind of deviation between the discovered model and the event-log. If any deviations are found, they are reported to a reasoner, which suggests a reaction. This action is executed by the executor.

A similar approach is used by Chiò et al. to detect changes in the production line. By utilizing process mining techniques, they generate a graphical representation of the event log, which allows for visual analysis by a human operator. Based on this analysis, necessary actions can be taken if any changes are identified. [CAP21].

5.2 Approaches with case study in the lab

The paper by Friedrichs et al. presents a framework for data-driven DTs. This approach uses PM to extract process models from the event logs using process discovery. The discovered model creates the simulation model of the DT. The discovered process model serves as the simulation model for the DT, ensuring that it is always up-to-date and accurately reflects the mirrored production environment [Fr22].

Another approach from Mayr et al. [MLC22] focusses on creating process models from process-state data. This involves generating an event log of the process using clustering techniques. The generated event log is then used to discover the process model, which monitors the control-flow of the production process.

A completely different approach by Novák and Vyskočil [NV22] uses PM to mine the operational data of time for the DT. The time data itself is divided into duration, waiting time and activity times of the process. This allows the DW to be flexible and gives him the ability to react to changes in the production.

5.3 Approaches with case study by manufactures

In the paper from Kumbhar et al., the objective is to create a digital twin (DT) of the production process to identify bottlenecks based on operational data. Process mining (PM) is employed to construct an accurate process map of the as-is process. This involves creating a comprehensive process map, which is then abstracted and tested for conformance in subsequent steps. The outcome is an as-is process map that serves as the foundation for simulating the production process in the DT. The study demonstrates that PM facilitates the creation of the simulation at the expense of data preparation efforts. [KNB22].

Tran et al. [Tr21] adopt a similar approach to create a DT that optimizes the parameters of the real-world system using simulation. PM is utilized to extract a process model from data, which in turn generates and updates the autonomy of the DT. This automated update process ensures that the simulation remains accurate and closely reflects the actual system at all times.

A different approach for PM and DTs is used by Ruppert and Abonyi [RA20]. This conception has the goal of creating a DT for real time tracking of products in the production. The concept involves utilizing several sensors and the manufacturing execution system (MES) to monitor and track products in real time. PM uses this information to extract the process flow and to keep the simulation adaptive.

6 Discussion & Conclusion

The objective of this paper is to present an overview of how process mining is employed to enhance digital twins (DTs) and the resulting effects on DT. Based on a limited number of papers reviewed, it is evident that the majority of approaches focus on achieving information retrieval goals for DTs. Currently, the retrieved information primarily includes the process model, processing time, waiting time, and total process time. In addition to the process model and times, process mining could also generate information about the individuals involved in the process. This is generally not the case. Reasons for this may be the insufficient availability of such information for extraction or the absence of a need for it.

The papers show that the enhancement of DT with PM brings a few positive effects: The DT reflects the actual process, can be adjusted to a newly planned production and the generation of simulation models is supported.

Indeed, the utilization of process mining in enhancing digital twins (DTs) does present certain challenges: The availability and the completeness of data for PM as well as the event data needs to be prepared for PM.

Up to now, there are only a few studies using DTs to enhance with PM. Nevertheless, a clear picture emerges. The approach of a sustainable and flexible DTs can be supported by PM. The ability to extract and prepare information from existing information systems and process it within the digital twin framework is a crucial step in advancing production towards the state of Industry 4.0.

Acknowledgement

This work is supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) under grant No. 01MD22001C as part of the edge data economy initiative.

Bibliography

- [Be21] Bergs, T. et al.: The Concept of Digital Twin and Digital Shadow in Manufacturing. *Procedia CIRP* 101, pp. 81–84, 2021.
- [Br21] Brockhoff, T. et al.: Process Prediction with Digital Twins: 2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C). IEEE, pp. 182–187, 2021.
- [CAP21] Chiò, E.; Alfieri, A.; Pastore, E.: Change-point visualization and variation analysis in a simple production line: a process mining application in manufacturing. *Procedia CIRP* 99, pp. 573–579, 2021.
- [DUM17] D’Addona, D. M.; Ullah, A. M. M. S.; Matarazzo, D.: Tool-wear prediction and pattern-recognition using artificial neural network and DNA-based computing. *Journal of Intelligent Manufacturing* 6/28, pp. 1285–1301, 2017.
- [Fr22] Friederich, J. et al.: A framework for data-driven digital twins of smart manufacturing systems. *Computers in Industry* 136, pp. 103586, 2022.
- [Fu23] Fur, S. et al.: Sustainable Digital Twin Engineering for the Internet of Production. In *Digital Twin Driven Intelligent Systems and Emerging Metaverse*. Springer Nature, Singapore, pp. 101–121, 2023.
- [KB13] Kitchenham, B.; Brereton, P.: A systematic review of systematic review process research in software engineering. *Information and Software Technology* 12/55, pp. 2049–2075, 2013.
- [KNB22] Kumbhar, M.; Ng, A. H.; Bandaru, S.: Bottleneck Detection Through Data Integration, Process Mining and Factory Physics-Based Analytics. In *SPS2022: Proceedings of the 10th Swedish Production Symposium*. IOS Press, 2022.
- [Kr18] Kritzinger, W. et al.: Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* 11/51, pp. 1016–1022, 2018.
- [LM21] Lugaresi, G.; Matta, A.: Automated Digital Twins Generation for Manufacturing Systems: a Case Study. *IFAC-PapersOnLine* 1/54, pp. 749–754, 2021.
- [MLC22] Mayr, M.; Luftensteiner, S.; Chasparis, G. C.: Abstracting Process Mining Event Logs From Process-State Data To Monitor Control-Flow Of Industrial Manufacturing Processes. *Procedia Computer Science* 200, pp. 1442–1450, 2022.
- [NV22] Novák, P.; Vyskočil, J.: Digitalized Automation Engineering of Industry 4.0 Production Systems and Their Tight Cooperation with Digital Twins. *Processes* 2/10, pp. 404, 2022.
- [PLN22] Park, K. T.; Lee, S. H.; Noh, S. D.: Information fusion and systematic logic library-generation methods for self-configuration of autonomous digital twin. *Journal of Intelligent Manufacturing* 8/33, pp. 2409–2439, 2022.
- [RA20] Ruppert, T.; Abonyi, J.: Integration of real-time locating systems into digital twins. *Journal of Industrial Information Integration* 20, pp. 100174, 2020.

- [Ro15] Rosen, R. et al.: About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. *IFAC-PapersOnLine* 3/48, pp. 567–572, 2015.
- [Rv06] Rozinat, A.; van der Aalst, W. M. P.: Decision Mining in ProM. In (Hutchison, D. et al. Hrsg.): *Business Process Management*. Springer, Berlin, Heidelberg, pp. 420–425, 2006.
- [Sv08] Song, M.; van der Aalst, W. M.: Towards comprehensive support for organizational mining. *Decision Support Systems* 1/46, pp. 300–317, 2008.
- [Ta18] Tao, F. et al.: Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology* 9/94, pp. 3563–3576, 2018.
- [Tr21] Tran, T. et al.: Real-time locating system and digital twin in Lean 4.0: 2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI). IEEE, pp. 369–374, 2021.
- [ULS17] Uhlemann, T. H.-J.; Lehmann, C.; Steinhilper, R.: The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. *Procedia CIRP* 61, pp. 335–340, 2017.
- [va05] van der Aalst, W. M. P.: Business alignment: using process mining as a tool for Delta analysis and conformance testing. *Requirements Engineering* 3/10, pp. 198–211, 2005.
- [va10] van der Aalst, W.: Process Discovery: Capturing the Invisible. *IEEE Computational Intelligence Magazine* 1/5, pp. 28–41, 2010.
- [va12] van der Aalst, W. et al.: Process Mining Manifesto. In (Daniel, F.; Barkaoui, K.; Dustdar, S. Hrsg.): *Business Process Management Workshops*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 169–194, 2012.
- [va16] van der Aalst, W.: *Process Mining. Data Science in Action*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.
- [vWM04] van der Aalst, W.; Weijters, T.; Maruster, L.: Workflow mining: discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering* 9/16, pp. 1128–1142, 2004.
- [Ya22] Yang, M. et al.: A Novel Embedding Model Based on a Transition System for Building Industry-Collaborative Digital Twin. *Applied Sciences* 2/12, pp. 553, 2022.

Process Mining und Künstliche Intelligenz in der Beschaffung

Einsatz von Process Mining und KI zur Bewältigung aktueller Herausforderungen in der Beschaffung


Marco Barenkamp ¹, Dany Moualeu-Ngangue ²


Abstract: Die wachsende Bedeutung der betrieblichen Beschaffung für den Unternehmenserfolg ist auf die Internationalisierung der Beschaffungsmärkte und den Einsatz neuer Technologien zurückzuführen. Die steigende Komplexität und mangelnde Transparenz der Lieferketten und Beschaffungsprozesse stellen jedoch Herausforderungen dar, die durch den Einsatz veralteter Technologien und manueller Prozesse oft verschärft werden. Moderne Technologien wie Process Mining und Künstliche Intelligenz können zur Lösung dieser Probleme beitragen, indem sie Prozesse visualisieren, Ineffizienzen aufdecken und große Mengen von Daten schnell analysieren. Durch die Automatisierung von Routineaufgaben wird die Effizienz erhöht und gleichzeitig Raum für strategische Initiativen wie weitere Digitalisierung und Automatisierung geschaffen. Die Analyse und Integration von Unternehmensdaten in Echtzeit ermöglichen fundierte Entscheidungen und eine verbesserte Risikosteuerung. In diesem Paper wird der Einsatz von Process Mining und Künstlicher Intelligenz (KI) in der Beschaffung eines real durchgeführten Projekts vorgestellt, um, unter Berücksichtigung von Konzepten des Maschinellen Lernens (ML) zur Optimierung des Beschaffungsprozesses in Unternehmen, langfristig Kostensenkungen, optimale Lagerbestände, erhöhte Produktivität, bessere Prozesstransparenz sowie Qualitätssteigerungen zu erreichen. Eingesetzt wurde dabei das Konzept der Zeitreihenanalyse in Verbindung mit Maschinellern. Es wurden dabei verschiedene Strategien an Prognosen, einschließlich direkter, rekursiver und mehrstufiger Ansätze, sowie Methoden zur Datenextraktion und Merkmalsextraktion eingesetzt und vorgestellt.

Keywords: Process Mining, Künstliche Intelligenz, Beschaffung, Procurement, Optimierung, Machine Learning, Zeitreihenanalyse

1 Einleitung

Die betriebliche Beschaffung nimmt innerhalb der Geschäftsprozesse eines Unternehmens einen hohen Stellenwert ein und die Signifikanz steigt überdies in der Regel bei Produktionsunternehmen. Bei der Beschaffung handelt sich heutzutage meist um eine

¹ Hochschule Osnabrück, Institut für Duale Studiengänge, Kaiserstr. 10c, 49809 Lingen, m.barenkamp@hs-osnabrueck.de,  <https://orcid.org/0000-0002-6724-5542>

² LMIS AG, AI & Data Science, Neumarkt 1, 49074 Osnabrück, dany.pascal.moualeu-ngangue@lmis.de,  <https://orcid.org/0000-0003-3356-7209>

Routineaufgabe, die in Unternehmen unter Anwendung eines ERP-Systems ausgeführt wird. Um ein möglichst hohes Maß an Effizienz in der Beschaffung zu erreichen, ist nicht nur eine gute Planung der Beschaffungsprozesse erforderlich, sondern ebenso deren Abstimmung und Integration in die Geschäftsprozesse des gesamten Unternehmens [DH20].

Unterschiedliche Einflussfaktoren sorgen in den letzten Jahren dafür, dass die betriebliche Beschaffung einen bedeutenden Wandel durchläuft, der sowohl mit neuen Möglichkeiten als auch mit großen Herausforderungen verbunden ist. Der größte Einfluss geht hierbei von der Internationalisierung der Beschaffungsmärkte aus, die dazu führt, dass der Anteil, den die Beschaffung zum Unternehmenserfolg beiträgt, weiter ansteigt. Zudem ermöglicht das Aufkommen neuer Technologien eine Umstrukturierung der Beschaffungsprozesse, wodurch große Effizienzsteigerungen realisiert werden können. Dazu zählen insbesondere moderne internet- und auf Künstlicher Intelligenz (KI) basierte Softwarelösungen [KR23].

Unternehmen sind heute oft nur einen einzigen Mausklick davon entfernt, einen Kunden an einen Mitbewerber zu verlieren. Daher ist die Kundenzufriedenheit von höchster Bedeutung. Aufgrund steigender Inflation und Störungen in der Supply Chain gestaltet sich dies aber immer schwieriger. Durch das Zusammenwirken von Mitarbeitenden, Prozessen und Technologien können Unternehmen diese Herausforderungen jedoch bewältigen, indem sie Ineffizienzen in ihren Geschäftsprozessen, die die Unternehmensperformance schwächen, aufdecken und beseitigen [CE23]. Insbesondere die Corona Pandemie hat auf eindrucksvolle Weise die Schwachstellen in den Lieferketten offengelegt. Häufig kam es in dieser Zeit zu Beschaffungslücken, die nicht rechtzeitig geschlossen werden konnten. Obwohl die Digitalisierung den Unternehmen umfassende Lösungen bietet, arbeiten zahlreiche Unternehmen im Beschaffungsbereich heute weiterhin mit einem technologischen Rückstand, was durchaus zu Frustration bei den Mitarbeitenden führen kann. Fehlende Transparenz kann zudem dazu führen, dass Risiken häufig nicht oder zumindest zu spät erkannt werden und Optimierungsmöglichkeiten somit ungenutzt bleiben [SA22].

Diese Situation adressiert vor allem das Process Mining. Es handelt sich hierbei um eine moderne Technologie, welche die Unternehmen in die Lage versetzt, die aktuellen Herausforderungen in der Beschaffung zu meistern. Abteilungs- und systemübergreifende Abstimmung sowie die Standardisierung der Beschaffungswege schafft Transparenz in den Lieferketten und ist die Basis für unternehmensweite Ausgabenoptimierung. Die Automatisierung der Beschaffungsprozesse befreit dabei Mitarbeitende von zeitintensiven und fehleranfälligen manuellen Tätigkeiten, wodurch vor allem Kosteneinsparungen realisiert werden können. Weiterhin wird die Zuverlässigkeit der Lieferketten optimiert und dadurch entstehende Produktionsunterbrechungen vermieden [CE23].

Vor diesem Hintergrund zeigt der vorliegende Artikel, wie mit Hilfe von Process Mining und KI der Beschaffungsprozess optimiert und dadurch dessen Beitrag zum Unternehmenserfolg gesteigert werden kann.

2 Grundlagen Process Mining

Process Mining ist ein datengesteuerter Ansatz, der darauf abzielt, aus den Ereignisprotokollen der betrieblichen Informationssysteme Erkenntnisse über die tatsächlichen Prozesse und Aktivitäten in einem Unternehmen zu gewinnen, Ineffizienzen zu erkennen und Initiativen zur Prozessverbesserung voranzubringen. Es geht darum, reale Prozesse zu entdecken, zu überwachen und zu verbessern, indem Prozesswissen aus den Ereignisdaten extrahiert wird [AC12]. Relevante Ereignisdaten werden aus verschiedenen Quellen extrahiert, wie beispielsweise Transaktionsprotokollen, Datenbanken oder Unternehmenssystemen. Diese Ereignisprotokolle enthalten Informationen über Aktivitäten, Zeitstempel, beteiligte Ressourcen und andere relevante Datenpunkte für weitergehende Analysen und Korrelationen [SA13].

Die extrahierten Ereignisdaten werden gereinigt, gefiltert und in ein geeignetes Format, wie in der nachfolgenden Tabelle exemplarisch dargestellt, für die Analyse umgewandelt. Zudem werden möglicherweise fehlende Daten ergänzt, Zeitstempel angepasst und irrelevante oder störende Ereignisse entfernt.

Id	Zeitstempel	Aktivität
LMAS-4234242	12.06.2023 08:12	Rechnungseingang
LMAS-4234242	12.06.2023 08:15	Rechnungskontrolle
LMAS-4234242	16.06.2023 14:36	Rechnungszahlung

Tabelle 1 Informationen für das Process Mining

Das Process Mining verwendet unterschiedliche Algorithmen, die auf mathematischen, statistischen und maschinellen Lernverfahren basieren, um die Ereignisdaten zu analysieren und aussagekräftige Prozesserkennnisse zu erlangen [BA22].

Die Anwendungen des Process Mining sind drei aufeinander aufbauenden Teilbereichen zugeordnet. Dabei handelt es sich, wie in der Abbildung 1 zu erkennen, um die Bereiche der Prozesserkennung (Discovery), der Prozessübereinstimmung (Conformance) und der Prozesserverweiterung bzw. -verbesserung (Enhancement) [BS23].

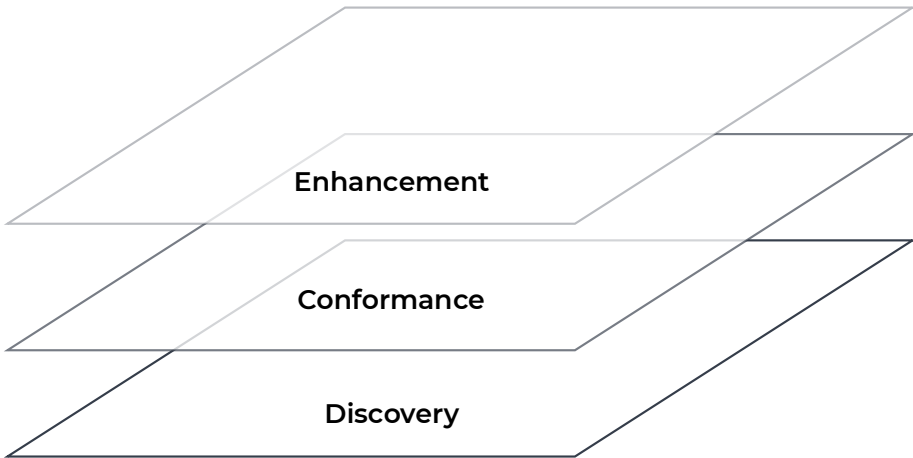


Abbildung 1 Teilbereiche des Process Mining

Bei der Prozesserkennung wird anhand der Ereignisdaten eine visuelle Darstellung des tatsächlichen Prozessablaufs erstellt. Process Mining Algorithmen analysieren dafür die Ereignisprotokolle, um die Abfolge von Aktivitäten, ihre Abhängigkeiten sowie den Kontrollfluss zwischen ihnen zu identifizieren. Die Ergebnisse der Prozesserkennung werden durch Diagramme visualisiert, die einen Überblick über die entdeckten Prozesse geben [SA13]. Im Rahmen der Prozessübereinstimmung wird ein Vergleich des entdeckten Prozessmodells mit dem idealen oder beabsichtigten Prozessmodell durchgeführt, um Abweichungen oder Nichtübereinstimmungen zu identifizieren. Durch die Abstimmung der Ereignisdaten mit dem Prozessmodell ist es möglich, die Übereinstimmung zu messen und Bereiche zu kennzeichnen, in denen der tatsächliche Prozess vom erwarteten Prozess abweicht [SM13]. Weiterhin ermöglicht Process Mining die Analyse von Prozessleistungskennzahlen wie Durchlaufzeit, Durchsatz, Engpässe und Ressourcenauslastung [LF22]. Auf Basis der Analyseergebnisse können Unternehmen somit fundierte Entscheidungen treffen, um ihre Prozesse zu verbessern und zu optimieren. Dies umfasst die Erkennung und Beseitigung von Engpässen, die Optimierung von Aktivitäten, die Umverteilung von Ressourcen oder die Neugestaltung des Prozessflusses, wodurch Effizienz, Qualität und Kundenzufriedenheit verbessert werden können [AC12].

3 Bedeutung der Beschaffung und aktuelle Herausforderungen

Die Beschaffung übernimmt wichtige strategische und operative Aufgaben und spielt für den Erfolg eines Unternehmens eine entscheidende Rolle. Die wichtigste Aufgabe der Beschaffung liegt in der Sicherstellung der Versorgungssicherheit. Eine effektive

Beschaffung gewährleistet die rechtzeitige und kontinuierliche Versorgung des Unternehmens mit allen benötigten Materialien, Produkten oder Dienstleistungen. Eine unterbrochene Lieferkette oder Engpässe in der Beschaffung können zu Produktionsausfällen, verpassten Geschäftschancen und Kundenunzufriedenheit führen. Weiterhin hat die Beschaffung direkten Einfluss auf die Kostenstruktur eines Unternehmens und spielt somit eine wichtige Rolle im Rahmen des Kostenmanagements. Durch die effiziente Verhandlung von Preisen, die Identifikation kostengünstiger Lieferanten und den Einsatz von Einkaufsvorteilen trägt die Beschaffung zur Kostensenkung und zur Steigerung der Rentabilität bei. Die Beschaffung ist zudem für die Bewertung und das Management von Risiken in der Lieferkette, wie z.B. Lieferantenrisiken, Währungsrisiken, politischen Risiken oder Naturkatastrophen verantwortlich und spielt somit eine entscheidende Rolle im Risikomanagement eines Unternehmens. Zuletzt ist die Beschaffung ein strategischer Aspekt bei der Identifikation innovativer Lieferanten und Technologien. Durch die Zusammenarbeit mit Lieferanten in einem frühen Stadium der Produktentwicklung können neue Ideen, Technologien und Markttrends frühzeitig erkannt und genutzt werden. Insgesamt kann festgehalten werden, dass die Beschaffung eines Unternehmens von großer strategischer Bedeutung ist. Sie ist die Grundlage für den reibungslosen Betrieb, die Wettbewerbsfähigkeit, die Rentabilität und ein nachhaltiges Wachstum. Eine effektive und effiziente Beschaffungsfunktion ermöglicht es dem Unternehmen, die richtigen Ressourcen zum richtigen Zeitpunkt zu den richtigen Kosten zu beschaffen und trägt somit maßgeblich zu dessen Erfolg bei [KR23].

Wie bereits eingeleitet, ist die Beschaffung heute mit großen Herausforderungen konfrontiert, welche den Unternehmen jedoch auch die Chance bieten, zu wachsen und neue Wettbewerbsvorteile zu realisieren. Durch den Einsatz moderner Technologien und effektiver Strategien können aktuelle Probleme bewältigt werden, um eine nachhaltige und effiziente Beschaffung sicherzustellen. Zu den bedeutendsten Herausforderungen in der Beschaffung zählen zurzeit fragile Lieferketten in Verbindung mit schwankenden Preisen und hoher Inflation, fehlende Transparenz und Komplexität in den Beschaffungsprozessen sowie fehlende Automatisierung und veraltete Technologien in der Beschaffungsfunktion.

Die Fragilität der Lieferketten wurde durch die Corona Pandemie offenkundig. Fragile Lieferketten sind anfällig für Störungen und Engpässe, die zwangsläufig zu Produktionsunterbrechungen führen können. Das Fehlen von Alternativen oder Pufferbeständen kann zu Lieferverzögerungen, erhöhten Kosten und Produktionsausfällen führen, was die Effizienz und Kundenzufriedenheit beeinträchtigt. Daher sind die Schaffung von Transparenz in den Lieferketten und die Auswahl verlässlicher Lieferanten bedeutende Aufgaben der Beschaffungsfunktion [CE23]. Inflation und schwankende Preise stellen für die Beschaffung zusätzlich eine große Kostenherausforderung dar, denn Unternehmen müssen mit steigenden Preisen für Rohstoffe, Materialien und Dienstleistungen umgehen, was ihre Profitabilität beeinträchtigt und Budgetplanungen erschweren kann. Es erfordert somit eine effektive Preisverhandlung und kontinuierliches Kostenmanagement [SA22].

Mangelnde Transparenz in den Beschaffungsprozessen kann ebenso zu Effizienzverlusten führen. Wenn beispielsweise Informationen über den Bestellstatus, den Lagerbestand oder die Liefertermine nicht in Echtzeit verfügbar sind, kann dies zu Verzögerungen, Überbeständen oder Engpässen führen. Viele Unternehmen haben keinen aktuellen Überblick über ihre Ausgaben, ihre Lagerbestände, ihre Lieferanten sowie deren Lagerbestände. Dies hat in den letzten Jahren mehrfach dazu geführt, dass Unternehmen nicht in der Lage waren, Unterbrechungen in ihrem Lieferantennetz zu kompensieren [SA22]. Beschaffungsprozesse können aufgrund einer Vielzahl von Aktivitäten, Beteiligten und sich ändernden Anforderungen komplex sein. Dies kann ebenfalls zu Ineffizienzen, hohen Kosten und Fehlern führen und insbesondere bei manuellen oder papierbasierten Prozessen besteht ein erhöhtes Risiko für menschliche Fehler [KA21].

Prozessautomatisierung ist daher eine der dringlichsten Aufgaben in der Beschaffung, denn dadurch entsteht Flexibilität und es werden Freiräume geschaffen. Mitarbeitende können wichtige strategische Initiativen umzusetzen, statt langweilige Routineaufgaben durchzuführen [SA22]. In vielen Beschaffungsabteilungen werden Datenanalysen manuell oder mit veralteten Technologien durchgeführt. In diesen Fällen fehlen aktuelle und datengestützte Erkenntnisse, auf deren Grundlage optimale Unternehmensentscheidungen getroffen werden können. Hier können moderne Anwendungen der KI eingesetzt werden, um fundierte Entscheidungen in der Beschaffung zu treffen und dadurch Unternehmensrisiken deutlich zu verringern [FI22] [KA21]. Jedoch besteht in diesem Fall häufig das Problem, dass die Einführung und Integration dieser neuen Technologien die Mitarbeitenden der Beschaffungsabteilung regelmäßig vor große technische Herausforderungen stellt, da die dafür notwendigen Kenntnisse und Fähigkeiten häufig fehlen [FL19]. Daher ist es für Unternehmen in der Regel ratsam, einen qualifizierten externen Dienstleister in Anspruch zu nehmen, um den digitalen Transformationsprozess erfolgreich voranzubringen [LF22].

4 Nutzenpotentiale von Process Mining in der Beschaffung

Mit Hilfe des Process Mining können in der Beschaffung enorme Nutzenpotentiale erschlossen und die aufgeführten Herausforderungen bewältigt werden. Process Mining Lösungen werden dabei von unterschiedlichen Software-Unternehmen am Markt angeboten [LF22].

Die Anwendung und Implementierung von Process Mining in der Beschaffung verläuft in der Regel in drei Schritten, wie in Abbildung 2 illustriert. In einem ersten Schritt werden die Unternehmensdaten aus den unterschiedlichen betrieblichen Informationssystemen mit Hilfe vorgefertigter Konnektoren in Echtzeit in die Process Mining Engine integriert. Dabei werden ebenso Personal-, Vertriebs- und Finanzdaten berücksichtigt, um eine ganzheitliche Perspektive des Beschaffungsprozesses zu erlangen. In einem zweiten Schritt wird Process Mining in Verbindung mit Maschinellem Lernen (ML) angewendet, um die real durchgeführten Prozesse objektiv zu identifizieren. Die Visualisierung der

Prozesse zeigt die digitale Wirklichkeit und offenbart Ineffizienzen innerhalb der Beschaffungsprozesse sowie deren Ursachen. Aufbauend auf diesen Ergebnissen können moderne ML-Algorithmen angewendet werden, um konkrete Handlungsempfehlungen zur Korrektur von Fehlern im Prozessablauf abzuleiten. In einem dritten Schritt werden die Empfehlungen priorisiert und Aktionen zur Behebung der Ineffizienzen ausgelöst [CE23].

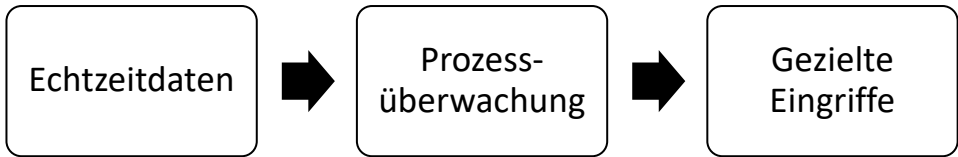


Abbildung 2: Aufdeckung und Korrektur von Ineffizienzen [CE23]

Die Analyseergebnisse durch Process Mining in der Beschaffung liefern damit einen weitreichenden Beitrag zum Gesamtunternehmenserfolg. Dies wird im Folgenden anhand von vier konkreten Anwendungsfällen veranschaulicht. Dabei handelt es sich um die Schaffung von Transparenz, die Automatisierung des Beschaffungsprozesses, die Optimierung des Risikomanagements sowie die Verbesserung der Datengrundlage für fundierte Unternehmensentscheidungen.

Process Mining bietet einen umfassenden Einblick in den Beschaffungsprozess vom Einkauf bis zur Bezahlung und liefert verlässliche Informationen zur Leistung gemäß wichtiger Beschaffungskennzahlen. Dadurch können die Treiber hinter aktuellen Leistungsniveaus sowie die Ursachen für schlechte Leistung erkannt werden. So können beispielsweise nicht nur manuelle Preisänderungen identifiziert, sondern auch ihre Quellen und ihre Auswirkungen auf andere Kennzahlen und vor allem ihren Einfluss auf den Unternehmenserfolg automatisch angezeigt werden. Mit diesem Einblick können schnell diejenigen Personen, Kunden und Prozesse erkannt werden, die manuelle Preisänderungen verursachen, und es können sofort Maßnahmen ergriffen werden, um diese Vorfälle zu reduzieren [KD21].

Process Mining ist weiterhin ein leistungsstarkes Werkzeug zur Identifikation von Automatisierungsmöglichkeiten im Beschaffungsprozess, da es einen zuverlässigen und umfassenden Überblick über die Leistung und Prozesslücken bietet. Es zeigt die Bereiche auf, in denen manuelle Arbeit zu Verzögerungen führt und wo Automatisierung zur Verbesserung der Ergebnisse eingesetzt werden kann. Durch die Darstellung potenzieller Zeiteinsparungen und Effizienzgewinne können Automatisierungsinvestitionen und -bemühungen priorisiert werden. Wichtig ist auch, dass Process Mining durch die Aufdeckung von Reibungspunkten, Prozesslücken und ineffizienten Bereichen dabei unterstützt, Problemprozesse zu identifizieren, die vor der Automatisierung behoben werden müssen, um zukünftige Probleme zu vermeiden [KD21].

Weiterhin kann die Anwendung von Process Mining in der Beschaffung das Risikomanagement sowie die Interne Revision unterstützen. Es ermöglicht eine

umfassende Analyse des Beschaffungsprozesses und identifiziert potenzielle Risiken, Compliance-Verstöße und ineffiziente Aktivitäten. Durch die Überwachung von Ereignisdaten in Echtzeit können Abweichungen von definierten Prozessen erkannt und sofortige Maßnahmen ergriffen werden. Process Mining ermöglicht die Identifizierung von Schwachstellen in der internen Kontrolle, Unregelmäßigkeiten in der Zusammenarbeit mit Lieferanten und potenzielle Betrugsfällen. Durch die datenbasierte Analyse können Risiken proaktiv erkannt, bewertet und minimiert werden, was zu einer verbesserten Governance und einem besseren Risikomanagement in der Beschaffung führt [JO17]

Process Mining in der Beschaffung kann überdies zu besseren Unternehmensentscheidungen beitragen, indem es eine datengetriebene Grundlage für Analysen und Erkenntnisse bietet. Durch die detaillierte Analyse des Beschaffungsprozesses können Engpässe, Flaschenhälse und ineffiziente Aktivitäten identifiziert werden. Dies ermöglicht das Ergreifen gezielter Maßnahmen zur Prozessoptimierung und darüber hinaus liefert das Process Mining Daten und Informationen in Echtzeit, um die Leistungskennzahlen des Beschaffungsprozesses zu überwachen und zu bewerten. Diese Informationen helfen bei der Identifikation von Trends, Mustern und potenziellen Risiken [LF22].

5 Nutzung von KI zur Optimierung der Beschaffung

Das Erstellen von fundierten Prognosen bietet eine Möglichkeit, ML in den Beschaffungsprozess einzubeziehen. Durch die Analyse von historischen Daten können Unternehmen Trends und Muster erkennen und daraus eine Prognose ableiten. Unternehmen benötigen dafür eine gute Vorhersagequalität der Produktnachfrage und des Ressourcenverbrauchs. Die gegenständliche Aufgabe in der Beschaffung bestand daraus, eine möglichst genaue Prognose der zu beschaffenden Materialien eines Fertigungsunternehmens zu erarbeiten. Zunächst mussten die historischen Daten des Verbrauchs verschiedener Materialgruppen durch Process Mining identifiziert werden, damit der zukünftige Bedarf an den gewünschten Materialgruppen prognostizierbar war. Das übergeordnete Ziel war eine Optimierung des Lagers. Diese Aufgabenstellung wurde mit Hilfe von Zeitreihenprognosen in Zusammenarbeit mit ML behandelt, wobei der Umsatz und die Positionsmenge für jede Materialgruppe als Zielvariable verwendet wurde. Die Prognosen sollten somit dazu beitragen, die Effizienz des Beschaffungsprozesses zu verbessern und die Kosten zu senken. Heutzutage sind die Einflussfaktoren auf Beschaffungsprognosen mithilfe von tradierten Zeitreihenanalysen sehr komplex [BO15], daher konzentrierte sich der gewählte Ansatz auf die Anwendung moderner ML-Methoden in Verbindung mit Zeitreihendaten mit dem Ziel, die Prognosegenauigkeit so zu verbessern, dass möglichst viele innerbetriebliche, aber auch makro-ökonomische Aspekte zu berücksichtigen waren.

Für eine erfolgreiche Prognose mit Zeitreihen werden für das so genannte Feature-Engineering diverse Zeitmerkmale (Tage, Wochen, Monate, Jahre), Verzögerungsmerkmale (Werte eines vorangegangenen Zeitraums), technische Merkmale (Trends, Jahreszeiten, Zyklen) sowie weitere Merkmale, die zum Verständnis von Ausreißern (Feiertage, Sonderangebote oder politischen Entscheidungen) beitragen, hinzugezogen [VE21]. Vor der Implementierung des gewählten ML-Algorithmus war es jedoch von entscheidender Bedeutung, die verfügbaren Rohdaten in anwendbare und signifikante Merkmale zu transformieren, welche spezifische und angestrebte Aspekte der beobachteten Phänomene repräsentieren.

Der erste Schritt bei der Merkmalsextraktion bestand darin, die Rohdaten umzuwandeln. Dies geschieht im Falle dessen, dass die Daten streng positiv sind, durch eine Box-Cox-Transformation, also einer statistischen Technik zur Transformation von nicht-normal verteilten abhängigen Variablen in eine Form, die wiederum näher an der Normalverteilung liegt. Für den Fall, dass die Rohdaten negativ sind, wird eine so genannte log+1-Transformation durchgeführt, bei der zu allen Datenwerten der Wert eins hinzuaddiert wird. Dieser Ansatz kann dazu beitragen, die Symmetrie der Verteilung zu verbessern und den Einfluss von Ausreißern zu verringern.

Merkmale wie „Wochen des Jahres“ oder auch „Tag des Jahres“ können aus den in den Rohdaten vorhandenen Datumsspalten abgeleitet werden. Andere Merkmale wie Verzögerungsdaten, gleitender Durchschnitt, Zyklus, Trend oder Fourier-Koeffizienten können auf der Grundlage dieser transformierten Daten geschätzt werden. Eine der Herausforderungen bei der Nutzung von Trendmerkmalen in Zeitreihen lag darin, möglichst verlässliche Informationen über zukünftige Entwicklungen zu erlangen, die in die Analyse einfließen können. Aus diesem Grund wurden die deterministischen Komponenten der transformierten Daten mit Hilfe der so genannten Out-of-Sample-Schätzung geschätzt. Diese bezieht sich auf den Prozess der Verwendung eines Teils der verfügbaren Daten zur Erstellung eines Modells und dann des Testens der Leistung dieses Modells mit Daten, die während der Modellbildung bewusst nicht verwendet wurden. Hierbei konnten bei der Zeitreihenanalyse die Daten der letzten fünf Jahre verwendet werden, um ein Prognosemodell zu erstellen, das dann die Leistung des Modells testet, indem die Prognosen mit den tatsächlichen Daten des darauffolgenden Jahres verglichen wurden. Mit diesem Ansatz konnte in der Tat bereits ein durchschnittlicher deterministischer Wert ermittelt werden, der allerdings noch mithilfe zusätzlicher Merkmale angepasst wurde. An dieser Stelle sei gesagt, dass wenn die Anzahl der Merkmale sehr hoch ist, die Einbeziehung multivariater Merkmale in Betracht zu ziehen ist wie beispielsweise eine Hauptkomponentenanalyse (PCA), eine Singulärwertzerlegung (SVD), eine (nicht-negative) UV-Zerlegung, eine Tensor Zerlegung oder eine lineare Diskriminanzanalyse (LDA), die versucht, eine optimale lineare Transformation der Daten zu finden, so dass jede hochdimensionale Beobachtung in eine niedrigdimensionale Darstellung projiziert wird [VE21].

Sobald Merkmale wie Trends, Zyklen, Saisonalität und Zeitschritte extrahiert wurden, ließ sich das Problem mit einem klassischen linearen Regressor lösen, da der Prozess

deterministisch ist. Verzögerungsmerkmale und zusätzliche Merkmale wie Feiertage, Sonderangebotszeiträume für spezielle Produkte und Werbung durch verschiedene Kampagnen veränderten jedoch die Natur der Aufgabe. Verzögerungsmerkmale implizieren eine zeitliche Verschiebung in der Zeitreihe, was in der Praxis bedeutet, dass nur eine begrenzte Anzahl an Schritten in die Zukunft prognostiziert werden kann. Ebenso war es entscheidend, dass die zusätzlichen Merkmale wie Feiertage und Sonderangebotszeiträume auch für den zu prognostizierenden Zeitraum bekannt und verfügbar sind, um das Modell mit der nötigen Genauigkeit zu schätzen.

Um diese Herausforderung zu lösen, gab es eine Reihe von Strategien, die gewünschten Prognosen in mehreren Schritten zu erstellen. Zu diesen Strategien gehören die direkte Strategie, die rekursive Strategie sowie die mehrstufige Strategie [VO15]. Eine direkte Prognosestrategie bei ML bezieht sich hierbei auf den Ansatz, bei dem das Modell direkt auf die zukünftigen Werte trainiert wird, was auch als One-Step-Forecast bezeichnet wird. Im vorliegenden Kontext bezog sich die direkte Prognosestrategie auf den Ansatz, bei dem ein Modell für jede zu prognostizierende Zeitperiode oder Sequenz einzeln erstellt wurde. Dies bedeutet, dass bei der Prognose der nächsten zehn Tage auch zehn verschiedene Modelle erstellt wurden, und zwar eines für jeden zukünftigen Tag. Dieser Ansatz ist relativ einfach zu implementieren und kann durchaus effektiv sein, wenn die Datenmuster relativ stabil sind.

Die rekursive Prognosestrategie bezeichnet einen ML-Ansatz, bei dem im vorliegenden Fall ein einziges Modell verwendet wurde, um mehrere zukünftige Zeitschritte vorherzusagen. Im Gegensatz zur direkten Prognosestrategie, bei der für jede Vorhersage ein separates Modell erstellt wird, wurde bei der rekursiven Prognosestrategie ein Modell trainiert, bei dem der nächste Zeitschritt auf Basis der vorherigen Daten vorhergesagt wird. Dieses Modell wurde dann wiederum rekursiv verwendet, um weitere zukünftige Schritte vorherzusagen zu können, indem die zuletzt vorhergesagten Werte als Eingabe für die nächsten Vorhersagen verwendet wurden [WA22].

Ein mehrstufiges Modell auch als Multi-Step-Forecasting bekannt, erstellt eine vollständige Prognose für jede als Eingabe verwendete Instanz und kann mit der rekursiven oder der direkten Prognosestrategie kombiniert werden [WA22]. Der final eingesetzte ML-Algorithmus für die einzelnen Dimensionen ist in der nachfolgenden Abbildung 3 dargestellt.

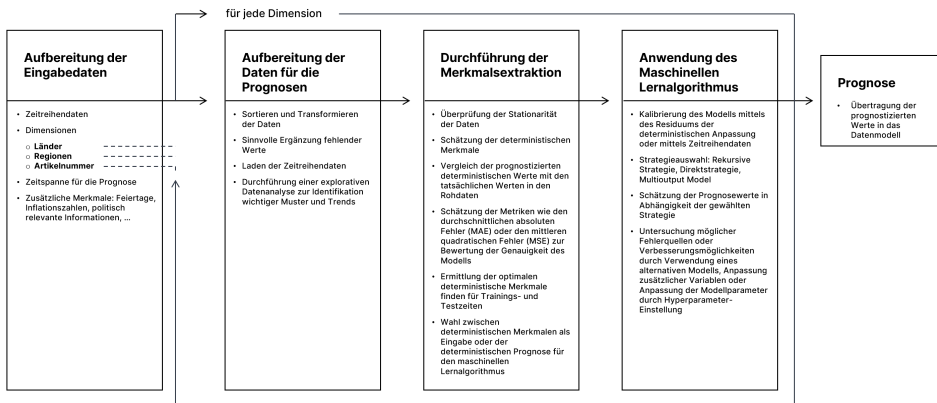


Abbildung 3 Vorgehen bei der Zeitreihenprognose mit maschinellem Lernen

In diesem speziellen Fall wurde ein mehrstufiges Ausgabemodell verwendet, das auf dem XGBRegressor basiert. Der XGBRegressor ist ein Teil der XGBoost-Bibliothek und steht für eXtreme Gradient Boosting Regressor. Es ist ein effektives ML-Modell, das auf dem Prinzip des Gradienten-Boosting aufbaut und zur Vorhersage von kontinuierlichen Zielvariablen verwendet wird.

Zusätzlich wurde zu diesem leistungsstarken Regressionsmodell auf das Modul `statsmodels.tsa.deterministic` und insbesondere auf dessen Funktion `CalendarFourier` zurückgegriffen. Diese Python-Bibliothek dient der statistischen Modellierung und beinhaltet verschiedene Funktionen zur Analyse und Modellierung von Zeitreihen. `CalendarFourier` ist dabei ein Werkzeug dieser Bibliothek, um trendbasierte, zyklische und saisonale Muster aus den zugrundeliegenden Daten zu extrahieren. Diese Methode konnte erfolgreich angewandt werden, um die Auftragswerte und Positionsmenge für verschiedene Materialgruppen vorherzusagen.

Insgesamt war die Durchführung eine sehr komplexe Aufgabe, die eine sorgfältige Vorbereitung erforderte. Ebenso war es wichtig, neben einer gründlichen Planung eine fundierte Implementierung sicherzustellen. Ein multidisziplinäres Team aus Data Scientisten, Prozessmanagern und Fachexperten konnte dazu beitragen, den gewählten Ansatz erfolgreich umzusetzen. Darüber hinaus war es notwendig, die Auswirkungen der Lösung auf die Unternehmenskultur und die Mitarbeitenden zu berücksichtigen, um sicherzustellen, dass die neuen Methoden erfolgreich in die bestehenden Prozesse integriert werden.

6 Fazit & Ausblick

Die Bedeutung der betrieblichen Beschaffung für den Unternehmenserfolg nimmt in den letzten Jahren immer mehr zu. Ursache dafür sind die Internationalisierung der

Beschaffungsmärkte sowie das Aufkommen neuer Technologien, die eine Neuausrichtung und Optimierung des Beschaffungsprozesses ermöglichen, wodurch Effizienzsteigerungen und Wettbewerbsvorteile realisiert werden können. Unternehmen, die diese Entwicklung verpassen, könnten in der Zukunft mit gravierenden Wettbewerbsnachteilen zu kämpfen haben. Die Beschaffung ist mit einer Vielzahl aktueller Herausforderungen konfrontiert, deren Bewältigung für den nachhaltigen Unternehmenserfolg von herausragender Bedeutung ist. Zu den größten Herausforderungen zählen mangelnde Transparenz und Komplexität in den Lieferketten und im Beschaffungsprozess. Erschwerend kommt hinzu, dass in vielen Unternehmen veraltete Technologien eingesetzt werden oder der Bestellungsprozess teilweise noch immer manuell durchgeführt wird. Diese Probleme können dazu führen, dass Ineffizienzen im Beschaffungsprozess entstehen, die sich zwangsläufig negativ auf die Unternehmensleistung auswirken.

Moderne Technologien wie Process Mining und KI können gewinnbringend eingesetzt werden, um diese Probleme zu adressieren. Mit Hilfe von Process Mining können die tatsächlich stattfindenden Prozesse erkannt und visualisiert werden, um darauf aufbauend Ineffizienzen und Schwachstellen im Prozessablauf zu identifizieren und zu beheben. Die Analyseergebnisse dienen als Grundlage zur Optimierung und Automatisierung des Beschaffungsprozesses mit Hilfe von KI und ML. Durch maschinelle Lernverfahren können große Mengen an Daten schnell und präzise analysiert und verarbeitet werden. Die ML-Technologie kann die Effizienz verbessern und Unternehmen dabei unterstützen, bessere Beschaffungsentscheidungen zu treffen. Durch die Analyse von Daten durch KI können überdies Muster erkannt werden, die auf Probleme oder auch Chancen hinweisen und damit zielführend verwendet werden können, notwendige Entscheidungen zu treffen oder Prozesse zu optimieren.

Ebenso wurde erarbeitet, dass Automatisierung eine der dringlichsten Aufgaben in der Beschaffungsfunktion ist. Dadurch wird der Beschaffungsprozess effizienter ausgeführt und dessen Ergebnisse verbessert. Ebenso werden die Mitarbeitenden von langweiligen Routineaufgaben entbunden, wodurch Kapazitäten geschaffen werden, wichtige strategische Initiativen, wie beispielsweise Digitalisierung und Automatisierung, weiter voranzubringen. Process Mining in Verbindung mit KI ermöglicht die Integration und Analyse aller Unternehmensdaten in Echtzeit und schafft dadurch eine aktuelle und präzise Datengrundlage als Basis wichtiger Unternehmensentscheidungen. Die automatische Ermittlung von Leistungskennzahlen versetzt Unternehmen außerdem in die Lage, Risiken frühzeitig zu erkennen und Mitigationsmaßnahmen zu ergreifen, bevor Schäden tatsächlich entstehen. Durch den kontinuierlichen Abgleich der stattfindenden Prozesse mit dem idealen Prozessmodell, können auftretende Abweichungen in Echtzeit erkannt und dadurch das Risikomanagement verbessert werden.

Die Kombination von Prozess Mining und KI ermöglicht Unternehmen somit, tiefe Einblicke in ihren Beschaffungsprozess zu erhalten und potenzielle Engpässe und Verbesserungsbereiche zu identifizieren. Durch die sorgfältige Anwendung maschineller Lernmethoden auf Zeitreihendaten kann ganz konkret eine präzise Prognose erstellt

werden, die zusätzlich hilft, den Beschaffungsprozess effizienter zu gestalten. Trotz der Komplexität der Methoden bietet dieser Ansatz ein enormes Potenzial zur Optimierung von Geschäftsprozessen. Kosten können hierdurch gesenkt werden, während gleichzeitig die Mitarbeiterproduktivität erhöht, die Prozesstransparenz und Prozessqualität gesteigert und das Risikomanagement verbessert wird. Alle diese Faktoren führen zu einer Steigerung der Unternehmensperformance. Insofern sollte die Implementierung einer Process Mining Lösung in der Beschaffung für jedes Unternehmen höchste Priorität besitzen. Sollten die dafür erforderlichen Kenntnisse und Fähigkeiten bei den Mitarbeitenden der Beschaffungsabteilung fehlen, ist die Inanspruchnahme eines externen Dienstleisters angebracht, um die digitale Transformation nicht zu gefährden, Wettbewerbsnachteile zu erleiden und Kunden an Mitbewerber zu verlieren.

Literaturverzeichnis

- [AA16] van der Aalst, Wil M.P (2016).: Process Mining – Data Science in Action. 2nd edn., Springer, Heidelberg
- [AC12] Accorsi, R., Ullrich, M., & van der Aalst, W. (2012). *Aktuelles Schlagwort: Process Mining*. Informatik Spektrum, 35(5), S. 354-359. <https://doi.org/10.1007/s00287-012-0641-4>
- [BA22] Barenkamp, M. (2022): Künstliche Intelligenz als Unterstützungsfunktion der Vorhersage und Prozessexzellenz im Process Mining. *Wirtschaftsinformatik & Management*, 14(3), S. 160-170. <https://doi.org/10.1365/s35764-022-00404-8>
- [BA23] Barenkamp, M. (2023): Künstliche Intelligenz im Process Mining: Anwendung und Potenziale. *Wirtschaftsinformatik & Management*, 15(2), S. 134-140. <https://doi.org/10.1365/s35764-023-00468-0>
- [BS23] Barenkamp, M. & Schnier, T. (2023). *Künstliche Intelligenz im Process Mining: Anwendung und Potenziale*. *Wirtschaftsinformatik & Management*, 15(2), S. 134-140. <https://doi.org/10.1365/s35764-023-00468-0>
- [BO15] Box, George EP, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung (2015): *Time series analysis: forecasting and control*. John Wiley & Sons
- [CE23] Celonis (2023): *Combat Inflation Through Operational Efficiency in Procurement*. https://assets.ctfassets.net/zmrtlfup12q3/1pPzK2akj0S9T0g3FAtHLk/174a3eb27db7d51122be3345fea6d158/Celonis_CPG_for_Procurement_Datasheet_Final.pdf
- [DH20] Dharmawan, Y. S. (2020): Procurement Process Analysis Using Process Mining in Cement Manufacturing Company. *IPTEK Journal of Proceedings Series*, 5, S. 39-44. <https://doi.org/10.12962/j23546026.y2020i5.7929>
- [FI22] Fischer, N., Roux, A., Haas, F., Platsch, A., Carradine, I., Tué, F., Schäfer, L. & Meintrup, J. (2022). *PwC Digital Procurement Survey 2022*. <https://www.pwc.com/gx/en/services/consulting/digital-procurement/pdf/PwC-digital-procurement-survey-4th-edition-2022.pdf>, Zugriff am 20. Juni 2023.

- [FL19] Flood, C. & Devoll, J. (2019). Reimagining Procurement with a Strategic Lens. Harvard Business Review. <https://resources.softwaretrends.com/resources/91154/reimagining-procurement-with-a-strategic-lens>, Zugriff am 20. Juni 2023.
- [JO17] Jones, S. (2017). Process Mining und interne Revision: ein Traumpaar? <https://www.celonis.com/de/blog/process-mining-and-internal-audit-a-match-made-in-heaven/>, Zugriff am 20. Juni 2023.
- [KA21] Karumsi, D., Clements, C., Foster, C., Joos, J., Rose, M., & Thirion, J. (2021). Future of Procurement. <https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2021/05/future-of-procurement.pdf>, Zugriff am 20. Juni 2023.
- [KR23] Krieger, W. (2023): Beschaffung. <https://wirtschaftslexikon.gabler.de/definition/beschaffung-30913>, Zugriff am 18. Juni 2023.
- [KD21] Krishnan, D. (2021): The Power of Process Mining for Procurement. <https://www.celonis.com/blog/the-power-of-process-mining-for-procurement/>, Zugriff am 18. Juni 2023.
- [LF22] Lobe, J. Z.; Busse, M.; Frankenberger, M.; Sander, D.; Fürst, F. M.; Vallée, F. & M'Barek, R. (2022): Process Mining in der Logistik. https://assets.ctfassets.net/zmrtlfup12q3/515gQkwcKDOmvyglzPccVF/0fa7a540c7177781ac18a10c0bd1bc9b/Whitepaper_Process_Mining_in_der_Logistik_-_Celonis.pdf, Zugriff am 18. Juni 2023.
- [SA22] SAP (2022): Argumente für Veränderungen: Fünf Fragen, die jeder CPO stellen sollte. <https://www.sap.com/germany/cmp/dg/de-make-a-case-for-change/index.html>, Zugriff am 18. Juni 2023.
- [SM13] Schultz, M.; Müller-Wickop, N.; Werner, M. & Nüttgens, M. (2013): Geschäftsprozessorientierte Prüfung von IT-Systemen. HMD Praxis der Wirtschaftsinformatik, 50, S. 41-51. <https://doi.org/10.1007/BF03340775>
- [SA13] Stocker, T.; Accorsi, R. & Rother, T. (2013): Computergestützte Prozessauditierung mit Process Mining. HMD Praxis der Wirtschaftsinformatik, 50, S. 92-103. <https://doi.org/10.1007/BF03340838>
- [VE16] Veit, F., Geyer-Klingeberg, J., Madrzak, J., Haug, M., & Thomson, J. (2017): The Proactive Insights Engine: Process Mining meets Machine Learning and Artificial Intelligence. In BPM (Demos)
- [VE21] Verdonck, T., Baesens, B., Óskarsdóttir, M. et al. Special issue on feature engineering editorial. Mach Learn (2021). <https://doi.org/10.1007/s10994-021-06042-2>
- [VO15] Vogel, Jurgen. Prognose von Zeitreihen. Springer Fachmedien Wiesbaden, 2015
- [WA22] Wang, C., Baratchi, M., Bäck, T., Hoos, H. H., Limmer, S., & Olhofer, M. (2022). Towards Time-Series Feature Engineering in Automated Machine Learning for Multi-Step-Ahead Forecasting. Engineering Proceedings, 18(1), 17.

Assessing the performance of Neural Networks in Recognizing Manual Labor Actions in a Production Environment

Untertitel des Beitrages

Maximilian Höfinghoff¹, Ralf Buschermöhle¹, Goy Hinrich Korn², Marcel Schumacher¹, Arne Seipolt¹

Abstract: Action recognition technology has gained significant traction in recent years. This paper focuses on evaluating neural network architectures for action recognition in the production industry. By utilizing datasets tailored for production or assembly tasks, various architectures are assessed for their accuracy and performance. The findings of this study provide some insights and guidance for researchers and practitioners to select an appropriate architecture or pretrained models for action recognition in the production industry.

Keywords: Action Recognition, Production, Benchmark, Machine Learning

1 Introduction

The digital transformation in the production industry is experiencing a rapid transformation, leading to a growing interest in leveraging technologies like Digital Twins. This technology has great potential to revolutionize production activities. “Industry 4.0 has enabled technological advancements in sensing, monitoring, and decision-making tools. These advancements helped the precise implementation of Digital Twins for the real-time monitoring and optimization of the process“ [AC23]. By 2028, approximately 60 percent of executives spanning various industries intend to integrate Digital Twins into their operational strategies, as stated in a report published in 2022 [lt23]. Action Recognition can support this endeavor by analyzing processes in the production chain. The objective of this paper is to assess suitable neural network architectures for action recognition in production settings. All architectures should be capable of accurately classifying common manual labor actions. Additionally, this paper aims to provide guidance regarding the selection of pretrained models, datasets, and frameworks.

1.1 Action Recognition

In recent years, significant process has been made in the field of computer vision, particularly in the development of neural network architectures. These advancements have led to increased interest in action recognition, a specific field that focuses on action

classification in video data. “Human action recognition is an important application domain in computer vision. Its primary aim is to accurately describe human actions and their interactions from a previously unseen data sequence acquired by sensors. The ability to recognize, understand and predict complex human actions enables the construction of many important applications such as intelligent surveillance systems, human-computer interfaces, health care, security, and military applications.” [Ph22] This paper focuses on action recognition that relies on the analysis of video data with neural networks.

To achieve accurate action recognition, researchers have explored and created various architectural approaches tailored to extract meaningful information from input data and classify actions precisely. The development of different architectures reflects ongoing efforts to optimize action recognition systems and address the inherent complexities associated with perceiving and understanding human actions by a machine.

1.2 Motivation

Many production steps are already automated in the production facilities, but there are still numerous tasks that require human workers due to their complexity, adaptability, and precision requirements. Manual labor actions encompass a wide range of activities, such as assembling components, operating machinery, inspecting products, packaging items, and conducting quality control checks. Action recognition can be used to automatically capture these actions and provide real-time information on action timing and status. An application that requires this data is a digital twin. Action recognition as a tool to support the creation of a digital twin is necessary to enhance the data that is used to accurately map the digital production process to the physical process. By analyzing the whole production process and recognizing manual labor actions, production companies can gain valuable insights into the efficiency of their processes. This knowledge allows them to identify bottlenecks, streamline workflows, and optimize resource allocation. By eliminating inefficiencies and optimizing the sequence and timing, production companies can reduce production cycle times, minimize machine downtime, and set up time, and enhance overall productivity [AC23]. Overall, capturing and understanding manual labor actions in production processes through action recognition techniques unlocks great possibilities. By leveraging neural networks and automation technologies, companies can enhance productivity, reduce costs, and improve the overall efficiency and quality of their operations.

1.3 Video Understanding through deep temporal modeling

Given the potential benefits of capturing and understanding manual labor actions in production processes, it is essential to explore the importance of specific architectures and datasets for the task of action recognition. Unlike image classification [LRR21], which focuses on recognizing objects in still images, action recognition involves a more complex task, understanding and identifying dynamic actions or activities in videos.

Primarily, the size of video data is inherently larger compared to image data, posing a difficulty for the architecture to maintain efficiency during inference. Furthermore, when the architecture operates on local edge devices in real-world settings, the challenge of being compact and efficient becomes even more pronounced.

Before the use of neural networks, features had to be handcrafted by practitioners to capture specific visual patterns and characteristics in images. CNNs were the first neural networks that have been highly successful in image recognition tasks by capturing spatial features in two-dimensional images. However, when it comes to videos, an additional temporal dimension needs to be considered. Traditional CNNs are not naturally designed to handle this temporal information, which is crucial for understanding and recognizing actions or activities in videos. To address this limitation, 2D CNNs are introduced in video understanding. These models extend the concept of CNNs by incorporating deep temporal modeling. By sequentially processing frames over time, 2D CNNs can learn to recognize temporal patterns and variations in videos. Another important factor is that the architectures need to model long-term dependencies across frames. Otherwise, they may fail to capture crucial temporal context and thereby lose important information for action recognition or analysis. Similarly important to the neural network architecture is the pre-training. It is a critical step in action recognition that involves training models on large-scale datasets that are oftentimes unrelated to the specific domain at hand. By utilizing diverse and extensive datasets, models can learn general visual representations that capture a broad range of visual features and patterns. Pretraining on such datasets allows the models to develop a foundational understanding of visual cues, which can be transferred and fine-tuned for the domain of choice e.g., the production industry. A diverse selection of pretrained models exists, each with architectures designed and trained on specific datasets. By leveraging suitable architectures and datasets, practitioners can make use of action recognition, enabling more precise identification and classification of manual labor actions in the production process.

2 Related Work

2.1 Previous studies on action recognition in production

The initial stage involves conducting a comprehensive literature review. Unfortunately, only a limited number of papers specifically address the domain of action recognition within the production industry or similar practical applications. [Hi20] developed algorithms for construction worker detection, tracking, and worker action recognition. The action recognition algorithm utilized an I3D architecture with RGB images and optical flow dynamics streams. It was trained to differentiate between eight actions such as walking, crouching, picking up and carrying. The system achieved a mean accuracy of

60.2%. Another paper [An21] focused on how action recognition can be utilized by a robot to understand human activities. The study explored multiple architectures fused with

sensor data to recognize activities like walking, kneeling, and bending. Recurrent Neural Networks, particularly LSTMs, were found to be the most effective architecture, achieving an accuracy score of 85%. Two other studies, [A119, A122] introduced a multimodal sensor system for assembly action recognition using electromyography (EMG) and inertial measurement unit (IMU) data collected from wrist bands. These studies employed fusion methods based on CNNs to combine the data effectively. [A119] demonstrated the effectiveness of the data in identifying precise finger motions, while skeletal data performed better in recognizing broader arm motions. Both studies emphasized the significance of fusing information from different modalities at various levels to enhance the system's reliability. Additionally, [A122] provided evidence that combining two separate sensor data-based action recognition models significantly improved the accuracy of predicting actions involving coordinated effort from both hands.

2.2 Action Recognition Benchmarks

The following section describes datasets that suit the purpose of this study and provided annotated videos with adequate actions in a simulated production environment. The four datasets are called InHARD [Me20], HA4M [Ci22], Ikea_ASM [Be20] and HRI30 [FEA22].

InHARD "Industrial Human Action Recognition Dataset" (InHARD), a dataset specifically designed for industrial human action recognition in real-world settings. InHARD comprises over 2 million RGB frames and skeleton data collected from 16 subjects, providing a rich and diverse source of data. With 13 different industrial action classes and more than 4800 action samples, this dataset offers ample opportunities for studying and developing a wide range of learning techniques for analyzing human actions within industrial environments, particularly those involving human-robot collaborations.

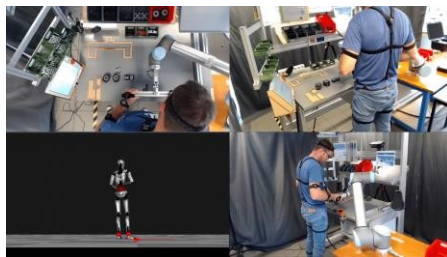


Figure 1: Data sample of InHARD dataset showing the four different modalities

HA4M The HA4M dataset encompasses a diverse range of multi-modal data capturing actions performed by various subjects in an assembly scenario within the production domain. The HA4M dataset offers 217 videos of the assembly task performed by 41 subjects. It includes six types of concurrent data: RGB frames, Depth maps, IR frames, RGB-Depth-Aligned frames, Point Clouds, and Skeleton data.



Figure 2: Two data samples from the HA4M dataset

Ikea ASM The IKEA ASM dataset is an extensive collection of multi-modal and multi-view videos, specifically designed to facilitate in-depth analysis and comprehension of human activities during assembly tasks. It comprises 371 samples of furniture assemblies, each accompanied by ground-truth annotations. The dataset encompasses various crucial elements, including 3 RGB views, a depth stream, atomic actions, human poses, object segments, object tracking, and extrinsic camera calibration.



Figure 3: Multiple data samples from the IKEA ASM dataset

HRI30 HRI30 addresses the need for human action recognition datasets specifically tailored for real-world applications in Human-Robot Interaction (HRI) contexts, particularly within production and service industries. HRI30 is designed to reflect actions commonly performed within industrial tasks. The HRI30 dataset consists of 2,940 video clips representing 30 action classes categorized into Human-Object Perception, Body-Motion Only, and Human-Robot Collaboration.



Figure 4: Two data samples from the HRI30 dataset

Dataset	Action Description	Type of Data	No. Samples	Actions	Camera Position	Video length
InHARD	Production tasks	RGB frames, skeleton data	2 mio.	74	Multiple views	-
HA4M	Product Assembly	Multiple (RGB videos, skeleton sensor data)	217	27	Frontal view(1-2m)	Seconds to several minutes
Ikea ASM	Furniture Assembly	Multiple (RGB videos, Skeleton, sensor data)	371	33	High Frontal view (2-3m)	a few seconds to tens of seconds
HRI30	Production tasks	RGB videos	2,940	30	Frontal view (2,30 m)	few seconds

Tab. 1: Overview of the different dataset in production domain

3 Methodology

3.1 Selection of the production dataset

Both the HRI30 and InHARD datasets encompass a wide range of actions commonly performed within production and industrial environments. They include actions such as delivering objects, drilling, polishing, picking up and putting down tools, and various human-robot collaborative tasks. This action variety aligns with the specific actions and tasks that occur in production settings, making these datasets more relevant and applicable for action recognition in the production domain. In contrast, the Ikea ASM and H4MI datasets have a specific emphasis on assembly tasks.

Industrial Context The HRI30 and InHARD datasets are specifically designed to capture actions within a collaborative industrial environment between humans and robots. They focus on simulating real-world industrial tasks, ensuring that the actions and scenarios presented in the datasets closely resemble the conditions and challenges faced in production settings. This contextual relevance makes these datasets more suitable for training and evaluating action recognition models for production applications.

In terms of action recognition in production environments, the HRI30 and InHARD datasets offer a combination of action variety, industrial context, dataset size that make them better suited in this context compared to the HA4M dataset and the Ikea ASM. These datasets provide researchers and developers with more realistic and relevant data for training and evaluating action recognition models in the context of industrial applications. However, we were only able to evaluate the architectures on a single dataset in our paper. This limitation arose due to the data format of the training data of the InHARD dataset. It consists of videos with multiple actions, each described with a start and end frame. The framework we utilized for the training and testing is better suited for analyzing videos with a single action or raw frames.

3.2 Evaluation environment: MMaction2 Framework

To test the neural network architectures on the datasets the mmaction2 library [Oa23] was used. The mmaction2 library is a module of the OpenMM Lab framework M [Op23], designed specifically for action recognition tasks. It provides a comprehensive set of tools and functionalities to facilitate the development and evaluation of action recognition models. The mmaction2 module offers a wide range of state-of-the-art action recognition models, including both two-stream and one-stream architectures. These models incorporate advanced techniques such as 3D convolutional neural networks (CNNs), temporal segment networks (TSN), and non-local neural networks (NLNet), among others. This variety allows researchers and developers to test the most suitable architecture for their specific needs and datasets. The mmaction2 module also includes various data pre-processing and augmentation techniques. It supports popular video datasets such as

Kinetics, Something-Something, and Moments in Time, and provides convenient data loaders and utilities for efficient data handling of individual data. Additionally, it offers functionalities for spatial and temporal data augmentation, which helps improve the robustness and generalization of the trained models. In terms of evaluation, the `mmaction2` module provides comprehensive tools for model assessment and comparison. It supports various evaluation metrics, including top-k accuracy and mean average precision (mAP). It also includes utilities for visualizing model predictions, enabling users to gain insights into the performance of their action recognition models. The framework also offers slurm cluster support, making it possible to train and test the architectures efficiently on multiple gpu-machines.

3.3 Selection of neural network architecture for evaluation

In this section, we discuss the selection of the neural network architecture for evaluating the performance of our proposed dataset. We consider the proposed architectures by Paper [Yi20] that were available in the `mmaction2` framework.

Since it is very resource consuming in terms of time and cost to do the pre-training the architectures by oneself, pre-trained models must be used. The `mmaction2` framework provides a collection of pretrained models with the proposed architectures trained and tested on different data sets for action recognition tasks. These models have been trained on large-scale datasets and have demonstrated strong performance in various action recognition benchmarks. We leverage these pretrained models to evaluate the action recognition capabilities of the proposed production dataset. By using these established models, we can compare the performance of the architectures and draw conclusions about the pre-trained models. This evaluation will help us identify the most suitable neural network architecture and pre-train dataset for action recognition in production environments. After examining the `mmaction2` framework, seven architectures were identified as suitable: TSN, TPN, I3D, R2+1D, TSM, CSN and UniFormer V2. These architectures were recommended by the previously addressed paper and were also pretrained on human-centric datasets. The datasets feature the actor in a frontal view, which was necessary to align with the benchmark dataset. Additionally, this dataset performed well on the typical benchmark datasets.

3.4 Training and validation procedure

The train/test split was around 90/10 (2946/294 videos) as proposed by the creators of the dataset. All architectures were trained 100 epochs with all other configurations set as proposed by `MMaction2` framework. The video underwent multiple transformations before the training, these were also suggested by the framework. The optimization parameters for learning were also taken from the preset configuration files of the framework and can be seen in Table 2.

Model	Type	Learning Rate	Momentum	Weight Decay	Nesterov
TSM	SGD	0.02	0.9	0.0001	-
I3D	SGD	0.01	0.9	0.0001	-
TPN	SGD	0.01	0.9	0.0001	True
CSN	SGD	0.0005	0.9	0.0001	-
UniformerV2	AdamW	1e-05	-	0.05	-
TSN	SGD	0.01	0.9	0.0001	-
R2+1D	SGD	0.01	0.9	0.0001	-

Tab. 2: Overview of the optimization parameters of the different architectures

3.5 Evaluation metrics for performance assessment

The primary metrics to evaluate the performance of the model are the top1-accuracy and the top5-accuracy. The top1-accuracy provides a measure of how well the model performs when considering only the single most likely prediction predicted by the model. Top-5 accuracy, on the other hand, considers the top five predictions made by the model. It measures the percentage of correctly predicted labels when considering any of the top five predictions as the correct class label. Number of parameters of a model refers to the trainable variables in a machine learning model. These variables are learned during the training process and contribute to the model's ability to capture and represent the underlying patterns in the data. In action recognition, models with a larger number of parameters can potentially capture more complex features and exhibit higher capacity to learn intricate action representations. However, having too many parameters can also lead to overfitting, where the model becomes too specialized to the training data and performs poorly on unseen data. Therefore, optimizing the number of parameters is crucial to strike the right balance between model complexity and generalization.

Based on these metrics, the evaluation of the action recognition model should consider the trade-off between accuracy and model complexity (parameters). Ideally, a good action recognition model should achieve high accuracy (both top-1 and top-5) while having less parameters to ensure computational efficiency and avoid overfitting.

4 Experimental Results

The evaluation of the architectures on the HRI30 dataset had the following results:

Model	Pretrain-Dataset	acc/top1	acc/top5	# of Param.
CSN(ResNet50 IR+BNFrozen)	IG65Ml	0,9898	1	13,13M
TSM(ResNet50 NonLocalDotProduct)	ImageNet	0,9898	0,9966	31,68M
TPN (ResNet50)	ImageNet	0,9796	1	-
I3D (ResNet50)	ImageNet	0,9762	1	28,0M
TSN (ResNet50)	ImageNet	0,9694	1	24,32M
R2+1D (ResNet34)	None	0,8469	1	63,8M
UniFormerV2 (UniFormer-S)	None	0,381	0,9388	21,4M

Tab. 3: Overview of the results of the study

5 Discussion

5.1 Interpretation of the experimental findings

The models achieved varying levels of accuracy (acc/top1 and acc/top5) on their respective datasets. TSM (ResNet 50 NonLocalDotProduct) and CSN (ResNet50 IR+BNFrozen) achieved high accuracy with values of 0.9898 and 1 (CSN) and 0,9966 (TSM) for acc/top5, respectively. This indicates their effectiveness in classifying videos. On the other hand, R2+1D (ResNet34) had significantly lower accuracy, with 0.8469 for acc/top1. The UniFormerV2 (UniFormer-S) performed poorly and could not do the task at all with a 0.381 acc/top1. With regards to the pretrain dataset, TSM, TPN, I3D, and TSN models were pretrained on the ImageNet dataset, while CSN was trained on the IG65Ml dataset. R2+1D and UniFormerV2 models did not have a pretrain dataset. Since both top scoring models were pretrained on a different dataset, one might argue that both pretrained datasets provide the model with a good general understanding of human actions. This claim can be supported by the fact that the lowest scoring models in this experiment did not use a pretrained dataset. The number of parameters (# of Param.) in the models varies significantly. R2+1D (ResNet34) has the highest number of parameters with 63.8M, indicating a relatively larger and more complex model architecture. CSN (ResNet50

IR+BNFrozen) has the lowest number of parameters with 13.13M, suggesting a more compact model design. The analysis of the number of parameters indicates that, for this dataset, a smaller model may be sufficient for video understanding einzustellen.

5.2 Discussion of the implications for real-world production applications

As seen before, the choice of architecture and pretrain dataset have significantly impacted the model's performance and generalizability to real-world scenarios.

Model Selection: The choice of model depends on the specific requirements of the production application. Models like TSM (ResNet 50 NonLocalDotProduct) and CSN (ResNet50 IR+BNFrozen) exhibit high accuracy on the respective datasets and can be suitable for applications of video recognition of typical production type actions.

Dataset Compatibility: Understanding the dataset on which the models were trained is essential for selecting a suitable model for a production application. Models trained on ImageNet may perform well in industrial contexts, since the learned broad actions seem transferable to this domain. The model can also be further pretrained on the previously introduced datasets. For general production tasks, InHARD and HRI30 seem reasonable. For assembly tasks Ikea Assembly and HA4M should be considered.

Computational Resources: Consideration should be given to the computational complexity and model size of the selected models. Models with higher parameter counts and FLOPS (Floating Point Operation per second), such as TSN (ResNet50) and R2+1D (ResNet34), require more computational resources for training and inference. In production environments, where efficiency and real-time processing are crucial, it may be necessary to opt for models with lower computational requirements, such as CSN (ResNet50 IR+BNFrozen) with fewer parameters. Nevertheless, OpenMMLab also supports deployment of the models which significantly reduces their size on the end devices.

In summary, there is often a trade-off between model performance (accuracy) and computational efficiency. Real-world production applications need to strike a balance between these factors. While higher accuracy models may provide more precise results, they often require more computational resources. It is important to evaluate the requirements and constraints of the specific application to select a model that meets the desired performance level while operating within the available resources.

6 Conclusion

In this paper seven preselected neural network architectures specialized for the task of action recognition in a production environment were evaluated. The architectures were compared based on their top1-accuracy, top5-accuracy. The initial selection of models was determined based on their performance on established benchmarks in the field of action

recognition. Pretrained models were obtained from the mmaction2 framework, which was also utilized for subsequent training and evaluation purposes. For the proposed dataset HRI30, which included 30 actions from the production context, many of the architectures achieved very good results, with CSN and TSM achieving the highest top1-accuracy. Additionally, four essential datasets were introduced specifically tailored to assembly and production line tasks, catering to the unique requirements of a project.

Overall, this evaluation instills hope that the architectures perform well in real-world scenarios, indicating that action recognition is highly applicable in the production environment. However, it is important to acknowledge that the study's results are limited by the fact that the laboratory environment does not fully mirror the conditions of an actual production facility. As a result, there are plans to conduct an additional evaluation using the more demanding industrial environment dataset InHARD.

Acknowledgement

This work is supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) under grant No. 01MD22001C as part of the "EDGE Data Economy" funding program.

References

- [AC23] Attaran, M.; Celik, B. G.: Digital Twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal* 6, p. 100-165, 2023.
- [AI19] Al-Amin, M. et al.: Action Recognition in Manufacturing Assembly using Multimodal Sensor Fusion. *Procedia Manufacturing* 39, p. 158-167, 2019.
- [AI22] Al-Amin, M. et al.: Fusing and refining convolutional neural network models for assembly action recognition in smart manufacturing. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 4/236, p. 2046-2059, 2022.
- [An21] Anagnostis, A. et al.: Human Activity Recognition through Recurrent Neural Networks for Human-Robot Interaction in Agriculture. *Applied Sciences* 5/11, p. 2188, 2021.
- [Be20] Ben-Shabat, Y. et al.: The IKEA ASM Dataset: Understanding People Assembling Furniture through Actions, Objects and Pose, 2020.
- [Ci22] Cicirelli, G. et al.: The HA4M dataset: Multi-Modal Monitoring of an assembly task for Human Action recognition in Manufacturing. *Scientific Data* 1/9, p. 745, 2022.
- [FEA22] Francesco Iodice; Elena De Momi; Arash Ajoudani: HRI30: An Action Recognition Dataset for Industrial Human-Robot Interaction. Zenodo, 2022.
- [Hi20] Hiroaki Ishioka et al.: Single Camera Worker Detection, Tracking and Action Recognition in Construction Site, 2020.

- [LRR21] Lorente, Ó.; Riera, I.; Rana, A.: Image Classification with Classic and Deep Learning Techniques, 2021.
- [It23] ltd, R. a. M.: Digital Twins Market by Technology, Twinning Type, Cyber- to- Physical Solutions, Use Cases and Applications in Industry Verticals 2023 - 2028. <https://www.researchandmarkets.com/reports/5308850/digital-twins-market-by-technology-twinning>, Accessed: 27.05.2023.
- [Me20] Mejdí DALLEL et al.: InHARD - Industrial Human Action Recognition Dataset in the Context of Industrial Collaborative Robotics: 2020 IEEE International Conference on Human-Machine Systems (ICHMS), p. 1–6, 2020.
- [NSC21] Nain, M.; Sharma, S.; Chaurasia, S.: Safety and Compliance Management System Using Computer Vision and Deep Learning. IOP Conference Series: Materials Science and Engineering 1/1099, p. 12013, 2021.
- [Op23] OpenAction: Welcome to MMAction2's documentation! – MMAction2 1.0.0 documentation. <https://mmaction2.readthedocs.io/en/latest/>, Accessed: 26.05.2023.
- [Op23] OpenMMLab: OpenMMLab. <https://openmmlab.com/>, Accessed: 26.05.2023.
- [Ph22] Pham, H. H. et al.: Video-based Human Action Recognition using Deep Learning: A Review, 2022.
- [Yi20] Yi Zhu et al.: A Comprehensive Study of Deep Video Action Recognition. ArXiv, 2020.

Features of Event-Driven Message Queuing Architectures in Manufacturing: A Reference Model for Comparison

Stefan Gudenkauf ¹, Javier Franke ² and Janek Behrens ³


Abstract: With increasing digitization and the use of cost-effective, ever more intelligent sensor and actuator systems at the edges of classic IT networks (edge computing), ever-increasing amounts of data are continuously being generated and sent from a wide variety of data sources. At the same time, manufacturing processes are subject to dynamic adjustments such as staff shortages, material shortages and fluctuations in energy costs. In addition, there is the requirement to demonstrate the sustainability of produced goods within the supply chain and to the outside world, which requires the collection of key figures across all levels of the manufacturing pyramid – and ideally also across the entire life cycle of the product. What they all have in common is that a very large amount of continuously and simultaneously operating systems can process sent data quickly and can react to relevant events in near real time. To build these systems, increasingly distributed applications based on the principle of message queuing (MQ) and event handling are attracting increasing interest in the manufacturing industry. In this paper⁴, we surveyed recent models and architectures for such event-driven systems. Based on this survey, we propose a consolidated feature model to uniformly describe and evaluate event-driven manufacturing systems, regardless of whether an organization’s own architectural needs or the offers of external providers are evaluated.

Keywords: Event-driven Architectures, Manufacturing Systems, Message Queuing, Feature Modeling.


1 Introduction

The automation systems established in the manufacturing industry are often a reflection of hierarchically organized processes and decision-making structures: they are hierarchically organized and communication is often *point-to-point*, see Fig. 1. Disadvantages of this hierarchy are the high communication effort and the limited flexibility due to rigid interfaces. In the context of current trends such as


¹ Jade University of Applied Sciences, 26389 Wilhelmshaven, Germany, stefan.gudenkauf@jade-hs.de,

 <https://orcid.org/0000-0002-1813-3448>

² Jade University of Applied Sciences, 26389 Wilhelmshaven, Germany, javier.franke@jade-hs.de,

 <https://orcid.org/0009-0005-0413-7076>

³ Jade University of Applied Sciences, 26389 Wilhelmshaven, Germany, janek.behrens@student.jade-hs.de,

 <https://orcid.org/0000-0002-4825-1752>

⁴ This work is supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) under grant No. 01MD22001B as part of the "Edge data economy" technology programme.

- **Internet of Things and Services (IoTS) and Industry 4.0** [SG16], in which parts of the automation hierarchy are regarded as cyber-physical production systems (CPPS),
- **Increasing service orientation** of the individual automation functions through cloud services and open service-oriented standards such as OPC UA [Di21], as well as
- **The need for ever-increasing optimization** of automation as a whole through methods of artificial intelligence (AI) and machine learning (ML) [Fr21],

these disadvantages are increasingly understood as competitive disadvantages. Therefore, manufacturing companies are striving to dissolve the classic automation hierarchy in favour of a service-oriented architecture whose services have to be defined and orchestrated. This requires suitable data-intensive applications [K117] that can record, store and distribute data across all levels of the automation hierarchy.

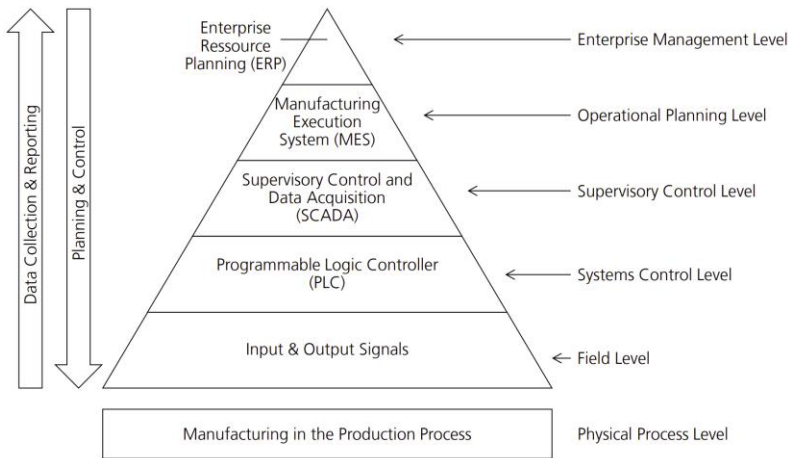


Fig. 1: Example of an automation hierarchy

1.1 Event-driven Architectures in the Manufacturing Industry

Traditional distributed applications such as distributed file systems such as HDFS [Bo08], cloud storage such as Amazon S3 [Pa08], and key value stores such as Apache Cassandra [LM10] are usually not sufficient for near real-time processing of such data streams [SJ14], [Fu21]. Instead, distributed applications based on the principle of message queuing (MQ)⁵ are considered suitable [Fu21], [So18], [Ja18], [Ja20], [Yo19]. Common examples include Apache Kafka [Kr11], [Sh22], Apache Pulsar [Ra19], [Ta23], [Jo21], RabbitMQ [MS19], RocketMQ [Yo19], ZeroMQ [Ze23] and several

⁵ Also often denoted as message-oriented middleware (MOM).

systems based on the Message Queuing Telemetry Transport (MQTT) protocol [MK20], [Mi18]. These technologies are currently being used successfully in software development in the context of microservices architectures [LF22], [Wo18] and the domain-driven design approach [Ev04]. Conversely, the use of microservices is also attracting increasing interest in the manufacturing industry, as individual services with a clearly defined function (single task) are being operated as separately distributable units (containers) – ideally as independently as possible of the underlying hardware infrastructure [Ce21]. Fig. 2 illustrates the commonalities and differences of such systems to traditional approaches on data processing.

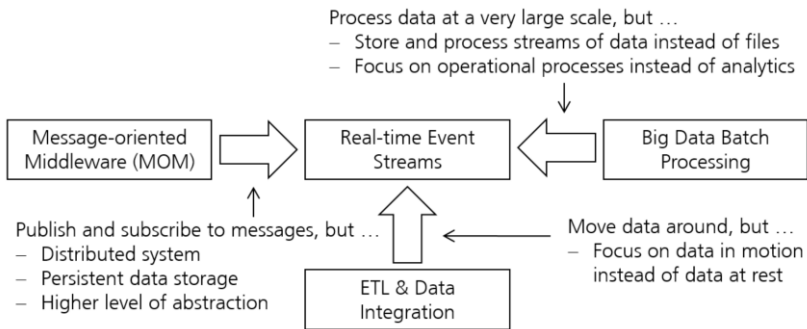


Fig. 2: Relations of broker architectures to other data processing approaches, cf. [Sh22]

The ongoing trend towards service orientation of the entire automation hierarchy now generates the opportunity to transfer the advantages of such event-based software architectures developed in operational software technology to a holistic, data-stream-based production in Industry 4.0 [SG16]. From the viewpoint of the data-providing and consuming services in production, these architectures then play the role of event-based integration platforms or simply as IoT brokers. However, potential users of such systems must find a way to uniformly describe and evaluate event-driven manufacturing systems, regardless of whether they want to formulate their own architectural needs or to evaluate the product offers of external providers.

The contribution of this paper is two-fold: (1) a survey of recent models and systems in manufacturing based on the notion of event-driven architecture (EDA) and in particular on message queuing (MQ), and (2) a consolidated feature model to uniformly describe and evaluate event-driven manufacturing systems, regardless of whether an organization's own architectural needs or the offers of external providers are evaluated. Section 2 presents the survey along with the description of the survey method, and Section 3 then discusses the features of event-driven MQ architectures in manufacturing. We discuss future work in Section 4 and summarize our findings in Section 5.

2 Subject of Research

According to the GQM approach of Basili et al. [Ba94] we formulate the goal of our work as the improvement (purpose) of the comparability (issue) of event-driven message-oriented middleware systems (object/process) from the viewpoint of the manufacturing industry (viewpoint). To do so, we examine the following questions:

1. What features do such event-driven systems of the recent past have?
2. Can we provide a reference model, with which one can describe and compare these systems?

To address the first question, this section presents an overview on several models and architectures of event-driven systems in the context of manufacturing, and a description of the survey method. We also summarize the differences between the surveyed approaches.

2.1 Survey Completeness

The works considered represent samples from academia and industry. We focused on approaches that (1) are based on the Event-Driven Architecture (EDA) software architecture paradigm and are (2) especially based on Message Queuing (MQ). Also (3), we considered approaches that are either generic or specifically oriented towards manufacturing/industrial production. Approaches that discuss specific aspects of MQ systems and architectures in-depth, such as performance and security, are excluded from the survey but may be added in future work. The considered approaches are:

- **Systems:** Apache Kafka [St18], Confluent Platform [Co20], Apache Pulsar [Ta23], RabbitMQ [MD19], LISA [Th15], Cybus Connectware [Cy22], and ZeroMQ [Pf22]
- **Models:** Bruns and Dunkel [BD10], Fu et al. [Fu21], Sommer et al. [So18], Yongguao et al. [Yo19], Dobbelaere and Esmaili [DE17], and Mishra and Kertesz [MK20]

2.2 Survey Method

The survey method is based on [Gu13] and is conducted in several steps. First, we began with a breadth-first search to identify relevant works. Although we did not apply a rigorous research method such as proposed by Kitchenham et al. [Ki07] [KB13], we used Google/Google Scholar, and SpringerLink as primary search platforms. Keywords included “event-driven”, “broker”, “message queuing” and “message-oriented” in conjunction with “manufacturing”. With the exception of Bruns and Dunkel [BD10], which we consider as a reference to EDA in general, the works considered were no older than 2015. We then selected works that comply with the selection criteria presented in

subsection 2.1. Next, we described the concepts discussed in these works using feature models [Ba05] [CE00] [CH06], since they combine accessibility, formality and structure discovery, as discussed by Gudenkauf et al. [Gu13]:

- **Accessibility:** “Similar to mind maps, feature models are simple and can be clearly arranged, represented, and compared to each other.”
- **Formality:** “Feature models are formal models that can be validated. Also, they allow modeling exclusions, multiple selections, and cardinality.”
- **Structure discovery:** “Feature models do not impose an organizational structure beforehand and even allow to model incomplete concepts. This is in contrast to tables, for example, in which columns and rows typically have to be named.”

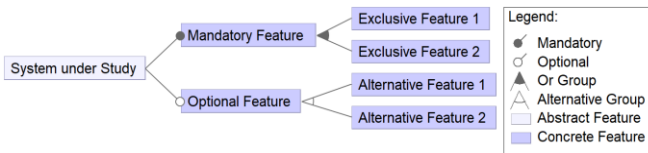


Fig. 3: An example of a feature model

Fig. 3 shows an example of a feature model. The root element represents a *system under study* (SUS). This element has two sub-features, of which one is mandatory (black-filled circle) and the other is optional (white-filled circle). Again, both features have sub-features. In general, the connection of a feature and its sub-features is described as an *and*-, *or*-, or *alternative*-group [Ba05] [Th14]. If a feature is selected, the selection of its parent feature is implied and all mandatory sub-features of an *and*-group must also be selected. An *or*-group implies that at least one sub-feature must be selected, and an *alternative*-groups implies that exactly one sub-feature has to be selected [Ba05] [Th14]. Abstract features denote features not mapped to concrete artefacts. Additionally, feature models support constraints for the selection of features in different branches and allow an arbitrary model depth. But for the sake of readability and ease of understanding, the depths of individual branches should be as uniform as possible [Ba05].

Finally, we compared the surveyed approaches by analysing the feature models, and synthesized a consolidated feature model to uniformly describe message queuing systems in the context of industrial production, see Section 3. For feature modeling, we used FeatureIDE [Me17], as shown in Fig. 4.⁶

2.3 Survey Results

In this section, we present the results of the feature model analysis of the surveyed models and architectures. For the sake of brevity, the section represents a summary of

⁶ See <https://featureide.github.io/>

the analysis.

Systems: In his book, Stopford specifically discusses the application of the open source event streaming platform Apache Kafka to the problem of tying back shared facts in distributed business information systems to a “single thread of irrevocable truth” [St18]. To do so, the author introduces streaming and Kafka, discusses patterns and techniques for development, and presents a system example using Kafka Streams and KSQL. Shapira et al. also provide an extensive discussion of Apache Kafka [Sh21], beginning with an overview of typical use cases for Kafka in a data ecosystem. Based on Apache Kafka, Confluent [Co20] provides an overview of the architecture of the Confluent Platform. The white paper is intended to serve as a reference and guideline for data architects and system administrators who are planning to deploy the platform in production.

Apache Pulsar is a distributed publish-subscribe messaging system originally created at Yahoo and provided by the Apache Software Foundation [Ta23]. In his Master’s thesis, Nuikka considers it a significant competitor to Apache Kafka for its performance characteristics, but notes that Pulsar lacks both community support and technological maturity [Nu21].

Madhu and Sunanda [MS19] give a brief overview of the main features of the RabbitMQ messaging system and discuss its use of the Advanced Message Queuing Protocol (AMQP). They also present the different types of exchange used to propagate messages to queues.

Theorin et al. present an event-driven service oriented architecture named Line Information System Architecture (LISA) [Th15]. LISA consists of a message bus, a custom message exchange format, and communication and service endpoints. As a message bus, LISA uses the middleware ActiveMQ. Communication endpoints represent low-level adapters between devices and the message bus, and service endpoints manage transformations of low-level events to standardized structures for connectivity to manufacturing execution systems (MES) and calculating key performance indicators (KPIs). Fill transformations append a set of static attribute-value pairs to a given event, Map transformations append a set of attribute-value pairs based on the current state of the system, and Fold transformations aggregate a finite sequence of events into a single event. In their paper, the authors emphasize the need for retrofitting devices in the production plant as they are based on different technologies and come from different eras.

Cybus Connectware is a commercial on-premise data integration platform specially developed for industrial production at the shop floor [Cy22]. It integrates a variety of shop floor data sources, protocols and formats such as Modbus, Message Queuing Telemetry Transport (MQTT), MTConnect, OPC UA and Profinet, thus providing a scalable factory-wide data inventory. At its core, its architecture consists of a MQTT broker. Connectware supports several data consumers including Cloud providers (e.g., Microsoft Azure and Amazon Web Services), IoT Applications (e.g., Siemens Mindsphere), data analytics applications (Grafana, Tableau and Elasticsearch) and local

IT systems such as SAP ERP and other enterprise resource planning (ERP) and manufacturing execution systems (MES). Cybus connectivity strategy is based on changeless machine connection that aims to avoid configuration changes to Programmable Logic Controllers (PLCs).

ZeroMQ [Ze23] is an asynchronous programming library that uses Sockets for high-throughput distributed systems. Although based on the message queuing paradigm, it does not require a dedicated message broker but can be used to implement a broker-based message exchange pattern. In his thesis [Pf22], Pfefferkorn finds evidence that ZeroMQ is suitable as a communication basis for distributed control and automation systems (DCS). To do so, he implemented the individual system nodes using Raspberry Pi single-board computers, on which the IEC 61499 standard for distributed automation is supported using the open source framework 4DIAC.

Models: Bruns and Dunkel provide an in-depth discussion of Event-Driven Architecture (EDA) in general, and Complex Event Processing (CEP) in particular [BD10]. The topics are discussed from three different perspectives, the technical benefits for companies, software design, and practical implementation. The EDA introduced by Bruns and Dunkel can be regarded as a high-level reference architecture. Additionally, depending on the use case, they recommend modifications for sensor networks and analysis systems.

Fu et al. [Fu21] compare five popular MQ systems that are representative in multiple aspects. The work includes a summary of the main features of MQ systems in general. These are discussed in terms of system production, quality of service, and performance.

In [So18], Sommer et al. present a comparison of four well-established message-oriented middleware systems in terms of qualitative and quantitative criteria, with the focus on performance and their suitability for industrial production.

Yongguo et al. [Yo19] discuss the main features of message-oriented middleware (MOM) based on a literature research and compares the main features and application scenarios of several common MOM products, namely JMS, Kafka, ZeroMQ, MQTT, AMQP, and RocketMQ. The table presented by Yongguo et al. that summarizes their comparison can be regarded as a simple feature hierarchy for MOM.

Dobbelaere and Esmaili [DE17] conduct a qualitative and quantitative/empirical comparison of the two popular and commercially supported open source pub/sub systems Apache Kafka and RabbitMQ. To do so, they establish a common comparison framework based on the core functionalities of pub/sub systems, highlight the distinct features of and best-suited use cases for the two systems. They conclude their survey with a decision table for the two systems.

Mishra and Kertesz present an extensive survey on the use of the Message Queuing Telemetry Transport (MQTT) protocol for machine-to-machine (M2M) communication and Internet of Things (IoT) systems [MK20]. They also propose a taxonomy to compare the properties of various concrete MQTT broker systems and programming libraries to

make it easier for scientists and end-users to select a specific MQTT implementation.

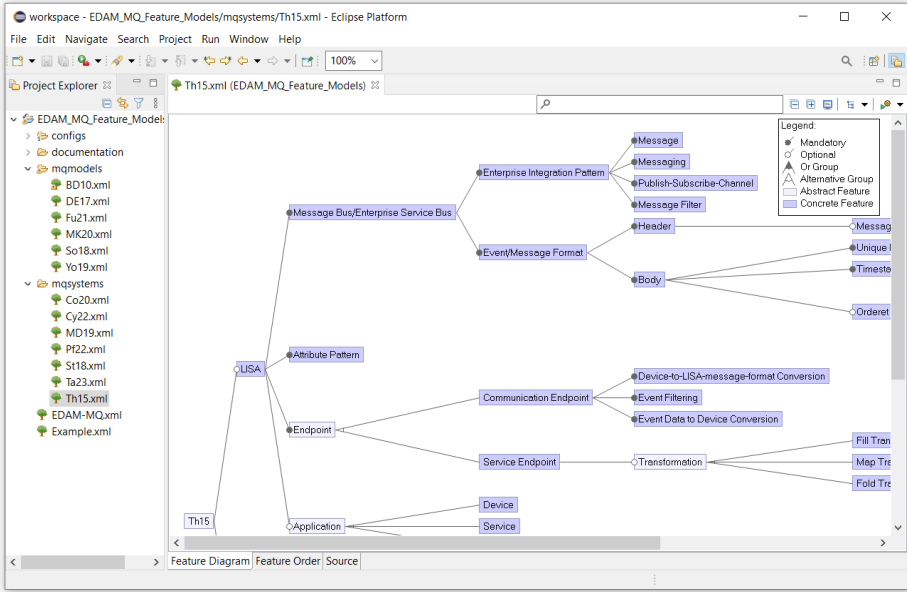


Fig. 4: Modeling with FeatureIDE

Summary: Overall, the work differs in terms of abstraction, as discussed above and shown in Tab. 1. The level of abstraction considers whether an approach describes a concrete architecture (system) or a generic representation or discussion of architectures (model). In addition, domain orientation considers if an approach is oriented towards an application domain or not (i.e., generic vs. manufacturing). Out of the 14 approaches, 5 are oriented towards manufacturing. We also consider 5 of the work as stemming from industry, while most of the work are of academic origin.

Formality shows how the authors describe their respected approach (e.g., using a formal taxonomy or model language, block diagrams, informal lists and/or plain text). We regard all of the work considered as informal (text, tables and diagrams), which limits its comparability and allows for unwanted interpretations. Thus, in the following section, we provide an aggregated feature model to fill the lack of a formal and understandable model for event-driven broker architectures in manufacturing. However, although the work considered is relatively recent, the number is small and can therefore only be considered a good start to more rigorous research.

Work	Abstraction	Domain	Type of Work	Formality	Provenance
[BD10]	model	generic	book	informal	academia
[Fu21]	model	generic	paper	informal	academia

[So18]	model	manufacturing	paper	informal	academia
[Yo19]	model	generic	paper	informal	academia
[DE17]	model, system	generic	paper	informal	academia
[MK20]	model	manufacturing	paper	informal	academia
[St18]	system	generic	book	informal	industry
[Sh21]	system	generic	book	informal	industry
[Co20]	system	generic	whitepaper	informal	industry
[Ta23]	system	generic	online	informal	industry
[MS19]	system	generic	paper	informal ⁷	academia
[Th15]	system	manufacturing	paper	informal	academia
[Cy22]	system	manufacturing	whitepaper	informal	industry
[Pf22]	system	manufacturing	thesis	informal	academia

Tab. 1: Overview of survey results

3 Features of Event-Driven MQ Architectures in Manufacturing

In principle, event-based or event-driven architectures (EDA) use information on events in order to communicate between services that are decoupled from one another. Such architectures are usually push-oriented, with all events processed only when needed [BD10]. If middleware systems based on the principle of message queuing (message-oriented middleware, MOM) are used as integration platforms, various criteria can be considered, see [Fu21], [Yo19]. Based on the analysis of the surveyed approaches in Subsection 2.3, we developed a consolidated feature model to describe and evaluate *event-driven message queuing architectures in manufacturing* (EDAM-MQ).

It is important to note that, as a model for comparison, a feature of the EDAM-MQ represents the description of a part or aspect of a concrete system, not the system aspect itself. Therefore, the selection/deselection of a feature represents the presence or absence of the respective description. The feature model consists of four main features: use case features, architecture-related features, system realization features, and quality of service-related features, see Fig. 5.

3.1 Use Case Features

Regarding the use cases of an event-driven message queuing architecture, we adopted the general use cases discussed by Shapira et al. [Sh21], namely *activity tracking*, *user notification*, *metrics logging*, *commit logging* and *stream processing*. We also added the specific use case of *industrial edge device connectivity* [Cy22], as it represents a fundamental challenge to manufacturing that employs a variety of devices “based on different technologies from different eras” [Th15]. Fig. 5 shows the features for use case descriptions.

⁷ However, the Advanced Message Queuing Protocol (AMQP) discussed in [MD19] is approved as an international Standard, see [Oa14].

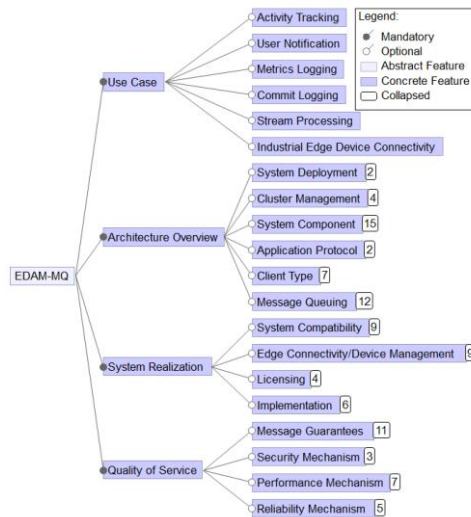


Fig. 5: Main features of the EDAM-MQ feature model; third-level sub-features collapsed

3.2 Architecture-related Features

Architecture-related features provide an overview of the architecture of an event-driven message queuing system. We consider *system deployment*, *cluster management*, *system components*, the native *application protocols*, *client types*, and the *message queuing model* as principal sub-features, see Fig. 6. First, system deployment can occur *cloud-hosted* or *on-premise* [Co20], and descriptions of cluster management includes mechanisms for *load balancing* [Yo19], *failure prevention* [Yo19] [St18], *local network server clustering* [MD19] and *replication* [St18].

Typical system components include *brokers*, a *management user interface* (UI), and storage systems for the overall system *configuration*, for *metadata*, and for *queues/messages*. Regarding the brokers, either a *master-slave* or a *peer-to-peer* operation model is usually applied [Fu21]. Moreover, regarding the Management UI, typical features include *status monitoring* [Yo19], *system configuration* [Yo19], *user management* [MD19], *permission management* [MD19], *queue management* [MD19] and *message exchange management* [MD19]. Thereby, status monitoring can provide *machine-centred* or *order-centred* views [Th15].

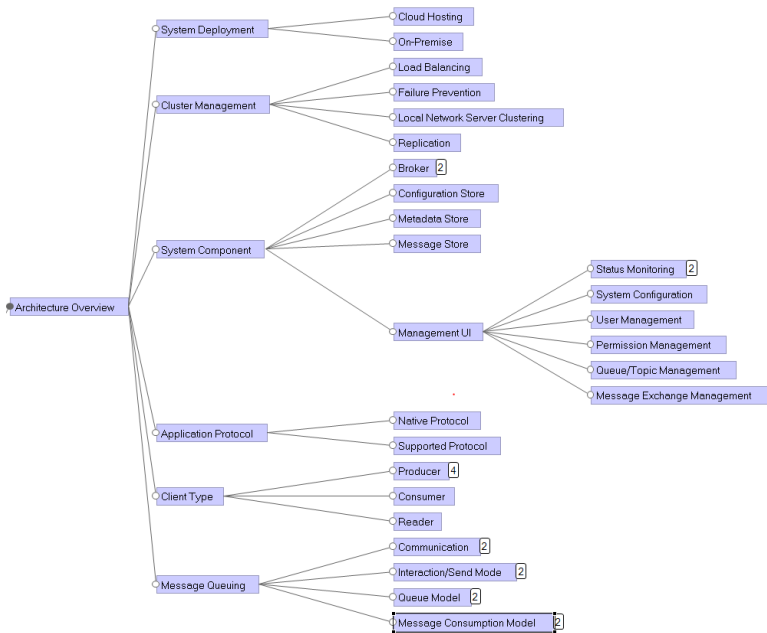


Fig. 6: Architecture-related description features

Message queuing systems usually apply a specific application protocol for messaging *natively*, and *support* several protocols additionally. Typical protocols include AMQP, STOMP, MQTT, XMPP, and HTTP [Yo19] [Fu21] [DE17]. Clients of such systems are distinguished in *producers* and *consumers*. In addition, we also consider *readers* that read but do not consume messages [Ta23]. Producers can *share access* to message queues; can have *exclusive access*, *exclusive access with fencing* or *wait for exclusive access* [Ta23].

Message queuing is discussed in terms of *communication*, *interaction*, the *queue model* applied, and the *message consumption* model. The communication model distinguishes between *point-to-point* communication and *publish-subscribe* communication [Yo19] [Fu21] [So18], while interaction mode differs between *request-reply* and *fire-and-forget* client interaction [Yo19]. Queue models can be *hierarchical* [Cy22], such as supported by MQTT, or *non-hierarchical*. Finally, message queuing either employs *push-* or *pull-* based message consumption [Fu21].

3.3 System Realization Features

Regarding the realization of the system under study, we consider the feature descriptions for *system compatibility*, *device management*, *licensing*, and *implementation*. Fig. 7 presents an overview of these features.

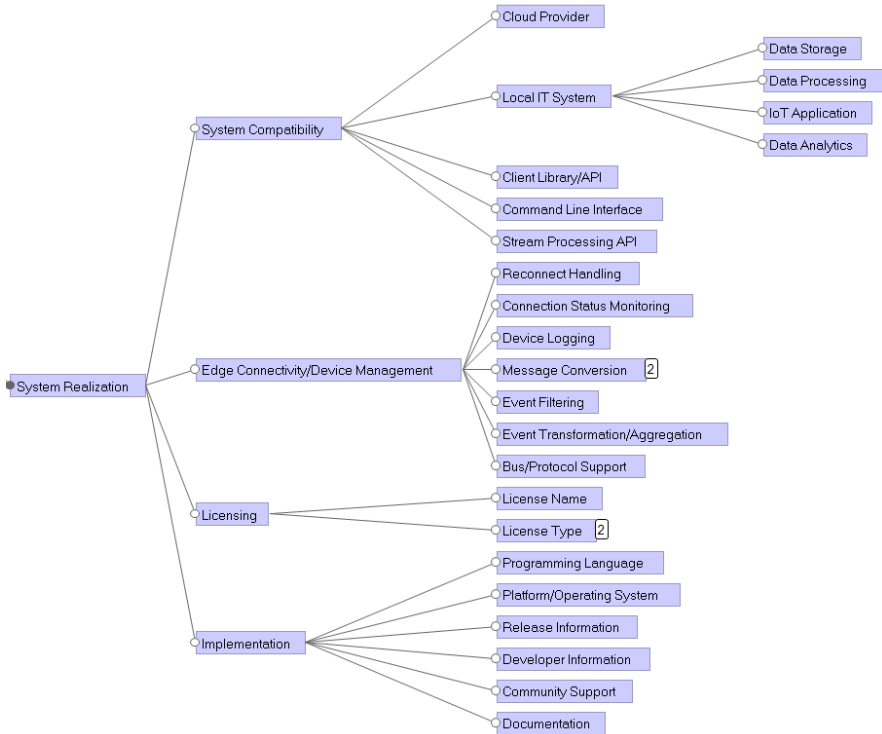


Fig. 7: System realization features

System compatibility [Cy22] can discuss *compatible cloud providers* such as Microsoft Azure and Amazon Web Services (AWS), a variety of *local IT systems* such as SAP ERP [Cy22] and HDFS [Fu21], *client libraries* and *application programming interfaces* (APIs), *command line interfaces* (CLIs) and *stream processing APIs* such as Kafka Streams and KSQL [St18].

Device management [Cy22], on the other hand, has to guarantee connectivity to a wide variety of Edge devices [Th15]. Typical features include *reconnect handling* [Cy22], *connection status monitoring* [Cy22], *device logging* [Cy22], *message conversion* (*message-to-device* and *device-to-message*) [Th15], and *event filtering* and *event transformation/aggregation* [Th15]. To do so, device management must support a variety of *bus systems/protocols* such as Modbus, OPC UA and Profinet [La21] [Cy22].

Licensing discusses the *name* and the *type* (*closed* or *open source*) of the system's software license [MK20] [So18]. Implementation names the primary *programming language* with which the system was realized and specifies the required *platform/operating system* [Fu21] [MK20] [MD19] [So18]. In addition, we consider information on the *software release*, the *developer*, on *community support* and on *documentation* as relevant for comparison [MK20] [So18].

3.4 Quality of Service-related Features

For the long-term and sustainable operation of a system, mechanisms to ensure service quality are of particular importance. Regarding event-driven message queuing systems, we consider guarantees for *message delivery* and *message ordering*, and the mechanisms to promote *security*, *performance*, and *reliability* as paramount. Quality of Service-related features are shown in Fig. 8

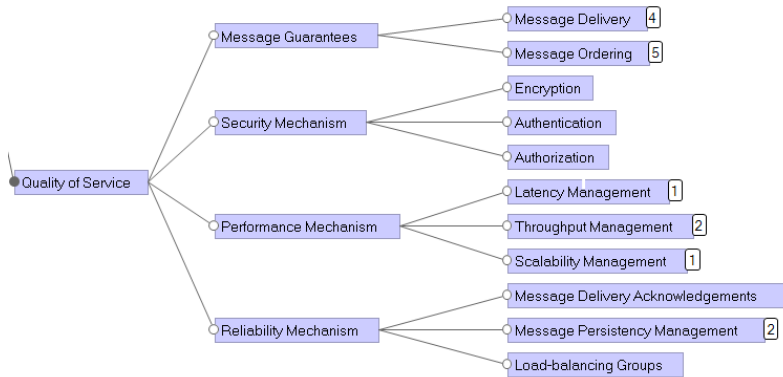


Fig. 8: Quality of Service-related feature descriptions

Message delivery typically considers the guarantees *at-most-once* (interval $[0..1]$), *at-least-once* (interval $[1..n]$), and *exactly-once* (interval $[1..1]$) [Fu21] [DE17]. If none is supported, a guarantee of *maybe* (interval $[0..n]$) is assumed.⁸ Message ordering can guarantee *partition-/queue-ordering*, *global ordering* of messages, or *no ordering* at all [Fu21] [Yo19] [DE17].

Regarding security, we include mechanisms for *encryption*, *authentication* and *authorization*. Encryption is often provided by SSL/TLS [Yo19]; authentication includes SASL [Yo19] [So18] and x.509 certificates [Cy22], amongst others. Authorization can be provided by access control list (ACL) [St18] [So18] and role-based access control (RBAC) [Co20]. Performance-related features include *latency management*, *throughput management* and *scalability management*. Latency can be addressed by optimizing *memory access* [Fu21], throughput by *message batching* [Fu21] [St18] and *partitioning* [St18] and scalability by using *quotas* [St18], for example. Finally, reliability mechanisms include *message delivery acknowledgements* [MD19], *message persistence management* (*in-memory* or *on-disk*) [MD19] and *load-balancing groups* [St18].

⁸ Note that the message delivery guarantee of *at-least-once* requires message persistence and that *exactly-once* requires transactional behavior for consuming and processing a message.

4 Future Work

Regarding the description of concrete event-driven systems for manufacturing that are based on the message queuing paradigm, the description of a feature according to the model presented in Section 3 can be said to be present or not. Gudenkauf et al. describe a concept for the visual representation of the state of a feature model instance using radial space-filling sunburst diagrams, coloring and interactive visualization using cloud services as an example [Gu13]. Their so-called Cloud Service Navigators (CSN) also support additional user-defined attributes, as shown in Fig. 9 (b). Adapting the CSN concept to the field of event-driven manufacturing can help companies improve long-term viable decisions about mission-critical middleware. We also advise a more rigorous literature review to increase the validity of the model, see Section 2.3.

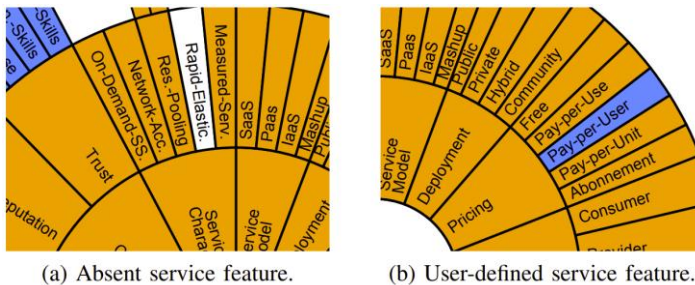


Fig. 9: (a) Section of a CSN that does not describe the feature „Rapid-Elasticity“; (b) Section of a CSN that extends the “pricing” feature with a user-defined “pay-per-user” pricing model

In Addition to the features of event-driven architectures, scenario-based methods have proven themselves for the qualitative evaluation of architectures, since they also take into account the risks of design decisions [Gu23]. Gudenkauf et al. describe an interactive approach and proper tooling to architecture evaluation based on the well-known Architecture Tradeoff Analysis Method (ATAM). This approach, Interactive Software Architecture Analysis (ISAA) can possibly be combined with CSN's interactive feature modeling to form a holistic decision system.

5 Conclusion

Distributed applications based on the principle of message queuing (MQ) and event handling are attracting increasing interest in the manufacturing industry. However, the selection of a concrete system architecture represents a significant decision with long-term consequences for a company. To improve the overview of the features of such systems, we surveyed different recent models and architectures for event-driven MQ systems. Based on this survey, we propose a consolidated feature model EDAM-MQ to uniformly describe and evaluate event-driven manufacturing systems, regardless of

whether an organization's own architectural needs or the offers of external providers are evaluated. As future work, we also envision the integration of interactive feature modeling for EAM-MQ systems and interactive architecture analysis based on quality scenarios as a promising experiment to form a holistic decision system.

Bibliography

- [Ba94] Basili, V.; Caldiera, G.; Rombach, H. D.: Goal Question Metric Approach: Encyclopedia of Software Engineering. John Wiley & Sons, Inc., pp. 528-532, 1994.
- [Ba05] Batory, D.: Feature Models, Grammars, and Propositional Formulas. In (Hutchison, D. et al. Eds.): Software Product Lines. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 7–20, 2005.
- [BD10] Bruns, R.; Dunkel, J.: Event-driven architecture. Softwarearchitektur für ereignisgesteuerte Geschäftsprozesse. Springer, Berlin, Heidelberg, 2010.
- [Bo08] Borthakur, D.: HDFS Architecture Guide. Hadoop apache project 53, 2008.
- [CE00] Czarnecki, K.; Eisenecker, U. W.: Generative programming. Methods, tools, and applications. Addison-Wesley, Boston, 2000.
- [Ce21] Cerquitelli, T. et al.: Manufacturing as a Data-Driven Practice: Methodologies, Technologies, and Tools. Proceedings of the IEEE 4/109, pp. 399–422, 2021.
- [CH06] Czarnecki, K.; Helsen, S.: Feature-based survey of model transformation approaches. IBM Systems Journal 3/45, pp. 621–645, 2006.
- [Co20] Confluent, Inc.: Confluent Platform Reference Architecture, 2020.
- [Cy22] Cybus: Cybus Connectware Technical Specification. <https://www.cybus.io/wp-content/uploads/2022/02/cybus-connectware-technical-specifications-siemens-industrial-edge.pdf>, accessed 22 May 2023.
- [DE17] Dobbelaere, P.; Esmaili, K. S.: Kafka versus RabbitMQ. The first public release of this report appeared in: P. Dobbelaere and K. S. Esmaili. Kafka versus RabbitMQ: A comparative Study of two Industry Reference Publish/Subscribe Implementations. In Proceedings of the 11th ACM International Conference on Distributed and Event-based Systems, pages 227–238. ACM, 2017. <http://arxiv.org/pdf/1709.00333v1>.
- [Di21] DIN EN IEC 62541-4:2021-08:2021-08-01, OPC Unified Architecture – Teil 4: Dienste (IEC 62541-4:2020); Englische Fassung EN IEC 62541-4:2020.
- [Ev04] Evans, E.: Domain-driven design. Tackling complexity in the heart of software. Addison-Wesley, Boston, Munich, 2004.
- [Fr21] Frochte, J.: Maschinelles Lernen. Grundlagen und Algorithmen in Python. Hanser, München, 2021.
- [Fu21] Fu, G.; Zhang, Y.; Yu, G.: A Fair Comparison of Message Queuing Systems. IEEE Access 9, pp. 421–432, 2021.
- [Gu13] Gudenkauf, S. et al.: A Reference Architecture for Cloud Service Offers. In (Gašević, D.

- et al. Eds.): Proceedings of the 17th IEEE International EDOC Conference. IEEE Computer Press, Vancouver, Canada, pp. 227–236, 2013.
- [Gu23] Gudenkauf, S.; Bachmann, U.; Hartmann, N.: A Concept and a Multitenant Web Application for Interactive Software Architecture Analysis. In (Sales, T. P. et al. Eds.): Enterprise Design, Operations, and Computing. EDOC 2022 Workshops. Springer International Publishing, Cham, pp. 268–283, 2023.
- [Jo21] Joseph, J.: Mastering Apache Pulsar. O'Reilly Media, Inc, 2021.
- [KB13] Kitchenham, B.; Brereton, P.: A systematic review of systematic review process research in software engineering. *Information and Software Technology* 12/55, pp. 2049–2075, 2013.
- [Ki07] Kitchenham, B. et al.: Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3. EBSE Technical Report, 2007.
- [Kl17] Kleppmann, M.: Designing data-intensive applications. The big ideas behind reliable, scalable, and maintainable systems. O'Reilly, Beijing, Boston, Farnham, Sebastopol, Tokyo, 2017.
- [Kr11] Kreps, J.; Narkhede, N.; Rao, J.: Kafka: a Distributed Messaging System for Log Processing: NetDB'11, 2011.
- [La21] Langmann, R.: Vernetzte Systeme für die Automatisierung 4.0. Bussysteme - Industrial Ethernet - Mobile Kommunikation - Cyber-Physical Systems. Hanser, München, 2021.
- [LF22] Lewis, J.; Fowler, M.: Microservices. a definition of this new architectural term. <https://martinfowler.com/articles/microservices.html>, accessed 14 Jun 2022.
- [LM10] Lakshman, A.; Malik, P.: Cassandra: a decentralized structured storage system. *ACM SIGOPS Operating Systems Review* 2/44, pp. 35–40, 2010.
- [Me17] Meinicke, J. et al.: Mastering software variability with FeatureIDE. Springer, Cham, 2017.
- [Mi18] Mishra, B.: Performance Evaluation of MQTT Broker Servers. In (Gervasi, O. et al. Eds.): Computational Science and Its Applications – ICCSA 2018. Springer International Publishing, Cham, pp. 599–609, 2018.
- [MK20] Mishra, B.; Kertesz, A.: The Use of MQTT in M2M and IoT Systems: A Survey. *IEEE Access* 8, pp. 201071–201086, 2020.
- [MS19] Madhu, M. P.; Sunanda, D.: Distributing Messages Using Rabbitmq with Advanced Message Exchanges. *International Journal of Research Studies in Computer Science and Engineering* 2/6, 2019.
- [Nu21] Nuikka, J.: Comparison of Cloud Native messaging technologies. Master's thesis. Tampere University, 2021.
- [Oa14] OASIS Advanced Message Queuing Protocol (AMQP) TC: ISO/IEC 19464:2014(E). Information technology — Advanced Message Queuing Protocol (AMQP) v1.0 specification. ISO copyright office, Switzerland, 2014.
- [Pa08] Palankar, M. R. et al.: Amazon S3 for science grids. In (Kosar, T. Ed.): Proceedings of the 2008 international workshop on Data-aware distributed computing - DADC '08.

ACM Press, New York, New York, USA, pp. 55–64, 2008.

- [Pf22] Pfefferkorn, D.: Verwendung von ZeroMQ in verteilter Automatisierung, Voralberg, 2022.
- [Ra19] Ramasamy, K.: Unifying Messaging, Queuing, Streaming and Light Weight Compute for Online Event Processing: Proceedings of the 13th ACM International Conference on Distributed and Event-based Systems. ACM, New York, NY, USA, p. 5, 2019.
- [SG16] Siepmann, D.; Graef, N.: Industrie 4.0 – Grundlagen und Gesamtzusammenhang. In (Roth, A. Ed.): Einführung und Umsetzung von Industrie 4.0. Grundlagen, Vorgehensmodell und Use Cases aus der Praxis. Springer Gabler, Berlin, Heidelberg, pp. 17–82, 2016.
- [Sh22] Shapira, G. et al.: Kafka. The Definitive Guide Real-Time Data and Stream Processing at Scale. O'Reilly, Sebastopol CA, 2022.
- [SJ14] Shahrivari, S.; Jalili, S.: Beyond Batch Processing: Towards Real-Time and Streaming Big Data. arXiv, 2014.
- [So18] Sommer, P. et al.: Message-oriented Middleware for Industrial Production Systems: 2018 IEEE 14th International Conference on Automation Science and Engineering (CASE). IEEE, pp. 1217–1223, 2018.
- [St18] Stopford, B.: Designing Event-Driven Systems. Concepts and Patterns for Streaming Services with Apache Kafka. O'Reilly, Sebastopol, 2018.
- [Ta23] The Apache Software Foundation: Apache Pulsar Docs Version 3.0.x. Concepts and Architecture. <https://pulsar.apache.org/docs/3.0.x/concepts-overview/>, accessed 28 May 2023.
- [Th15] Theorin, A. et al.: An Event-Driven Manufacturing Information System Architecture. IFAC-PapersOnLine 3/48, pp. 547–554, 2015.
- [Th14] Thüm, T. et al.: FeatureIDE: An extensible framework for feature-oriented software development. Science of Computer Programming 79, pp. 70–85, 2014.
- [Wo18] Wolff, E.: Microservices. Grundlagen flexibler Softwarearchitekturen. dpunkt.verlag, Heidelberg, 2018.
- [Yo19] Yongguo, Jiang and Qiang, Liu and Changshuai, Qin and Jian, Su and Qianqian, Liu: Message-oriented Middleware: A Review: 2019 5th International Conference on Big Data Computing and Communications (BIGCOM). IEEE, pp. 88–97, 2019.
- [Ze23] The ZeroMQ authors: ZeroMQ. An open-source universal messaging library. <https://zeromq.org/>, accessed 22 May 2023.

Ökologische Nachhaltigkeit - KI und Kreislaufwirtschaft

Zukunftsstrategien partizipativ entwickeln

KI-basierte Lösungen für die Kreislaufwirtschaft mit 3D Mapping gestalten

Anna Hoffmann¹

Abstract: Die Transformation zu KI-basierten Lösungen nimmt zu. Gleichzeitig nimmt die Kompetenz der Menschen ab, hinter den unsichtbaren Strom von Machine Learning Anwendungen zu blicken und die Zusammenhänge noch zu erkennen. KI erfordert viele Ressourcen, wird aber oft nur für den kurzfristigen kommerziellen Erfolg eingesetzt, statt auf die großen globalen Herausforderungen zu zielen. In dem Beitrag wird ein Lern- und Strategieprozess vorgestellt, der die Kompetenz des Menschen zu mündigen Entscheidungen über den Machine Learning Einsatz in komplexen Kontexten nachhaltig stärkt und das Wissen über neue Lösungsmöglichkeiten für die Kreislaufwirtschaft vertieft.

Keywords: KI, Zukünfte, Machine Learning, Design Thinking, Serious Games, Partizipation, Co:Creation, Strategie

1 KI als abstraktes Vorstellungsfeld und die Gefahr der menschlichen Unmündigkeit

Die unsichtbaren Vorgänge, die sich innerhalb von digitalen Prozessen abspielen, die auf Künstlicher Intelligenz basieren, sind für die meisten Menschen kaum greifbar und nur schwer vorstellbar. Der Begriff „Intelligenz“, der zunächst nur aus Marketingzwecken mit dieser Form der Technologie assoziiert wurde, trägt ein Übriges dazu bei, die innere Vorstellungswelt weit auszudehnen, worum es bei „KI“ denn nun geht. Viele Menschen halten sich nicht für unbegrenzt intelligent, und jetzt kommt eine Form der Maschinen auf, die genau das für sich in Anspruch nimmt: unbegrenzte Intelligenz auf Knopfdruck.

Dabei muss es gar nicht um das eher unwahrscheinliche Szenario gehen, das sich die KI verselbstständigt und wir als Menschheit unter die Herrschaft einer künstlichen Superintelligenz geraten. Das wird noch lange Zeit Science Fiction bleiben, denn Künstliche Intelligenz ist eben keine echte Intelligenz, sondern nur eine neue Art der Berechnung und Auswertung von Wahrscheinlichkeiten. Das allerdings in oft atemberaubender Geschwindigkeit, dank der Architektur und Exzellenz moderner Hardware.

Zur echten Intelligenz benötigen wir nach wie vor den Menschen und seinen schöpferischen Geist, denn echte Intelligenz bezieht eine ganze Reihe von

¹ Hoffmann Consulting, Hermann-Scheidemann-Weg 7, 14532 Stahnsdorf, anna@anna-hoffmann-consulting.de, <https://www.co-decision.de/>

Wahrnehmungen und Schlussfolgerungen mit ein, die kontextsensitiv und damit höchst variabel sind und sich deshalb einer strukturierten Programmierung entziehen. Doch das ist vielen Menschen unbekannt, denn die Propaganda für die übermenschlichen Fähigkeiten des Machine Learning und der KI läuft auf vollen Touren. Wir Menschen sollen durch die Fähigkeiten der KI unterstützt werden, und viele erhoffen sich Erleichterung für die von ihnen als leidvoll empfundenen Routinetätigkeiten wie das Formulieren von Texten.

Der Ausdruck von Gedanken in Schriftform ist eine grundlegende Kulturtechnik des Menschen. Warum das Schreiben von Text also nicht mehr als originärer schöpferischer, und damit befriedigender intellektueller Akt begriffen wird, sondern als Mühsal und Qual, die der KI-Roboter bitte abnehmen soll, hat mehr mit den zu hinterfragenden Kulturformen unserer modernen Arbeitswelt zu tun, als mit den Optionen auf Erleichterung, die durch die KI Einzug halten sollen.

Doch die Menschen hinterfragen die Nützlichkeit und Sinnhaftigkeit des Technologieeinsatzes immer weniger. Erst recht nicht, wenn diese neue Technologie ihnen als überlegen und omnipräsent verkauft wird. Der moderne Mensch verhält sich gegenüber den Technologien der Künstlichen Intelligenz unmündig, als hätten die Jahrzehnte der humanistischen Bildung den Geist der Aufklärung verblässen lassen.

Mit den ausufernden Möglichkeiten der Generative AI dürfte sich das nochmal potenzieren, den auch wenn die Buchstaben GPT in aller Munde sind, hat sich kaum jemand damit beschäftigt, wofür sie stehen. Selbst Entwickler und Berater, die Large Language Models (LLMs) nutzen wollen, um Software mit völlig neuen Funktionalitäten der Textgeneration auf den Markt zu bringen, wissen oft nicht, was eine Generative (= G) KI von einer anderen Form der KI unterscheidet, wie durch das Attribut Pretrained (= P) die Funktionalität eines LLMs eingeschränkt wird und das die Architektur des Transformers (= T) von Google als Open Source auf den Markt gebracht wurde, um jetzt via OpenAI (an dem mittlerweile nichts mehr „open“, also offen und frei zugänglich ist) bei Microsoft für neue Umsätze und Marktanteile zu sorgen.

Das wirklich Neue an dieser Form der Technik ist zum einen der direkte Bezug zum Menschen und seinem „Alleinstellungsmerkmal“, der Befähigung zur Erkenntnis und zum schöpferischen Denken, und zum anderen die Unsichtbarkeit gekoppelt mit der Allgegenwärtigkeit der Technologien des Machine Learning. Gefühlt sind Algorithmen des Machine Learning aus unserem Leben schon lange nicht mehr wegzudenken, sie sind in unseren Smartphones zu Hause, in der Zusammensetzung der „Google Search“, in jedem Sprachassistenten, in den Funktionalitäten von Netflix und Spotify.

Doch wann genau haben die KI-gestützten Anwendungen Einzug gehalten? Das bleibt für die Mehrheit ein Geheimnis, denn anders als bei Software aus früheren Jahrzehnten müssen Anwendungen der Künstlichen Intelligenz nicht extra installiert werden. Wir geben als Anwender keinen Startschuss in Form einer Installation oder Aktivierung, denn innerhalb des Machine Learning sind wir als mündige Menschen mit eigener

Entscheidungskraft selten vorgesehen. KI-Lösungen werden von wenigen für die Mehrheit entschieden und dann im Hintergrund innerhalb von bestehenden Anwendungen integriert. Deshalb sind sie allgegenwärtig, ohne dass wir als Gesellschaft eine aktive Entscheidung für ein Leben mit Machine Algorithmen getroffen haben. KI ist gekommen, um zu bleiben, und das als wirkmächtige Kraft, die in der Regel von außerhalb in unser Leben gespült wird, und ohne dem mündigen Menschen eine Wahlmöglichkeit offen zu lassen.

„>>Big Data<< birgt die Gefahr des reduktionistischen Denkens und der Ablösung der eigenen geistigen Kreativität durch Automaten“ [Go19].

2 Komplexität begreifbar machen

2.1 Gaming als Standardanwendung des Reinforcement Learning

„Die Digitalisierung und die fortschreitende Wissensgesellschaft werden die meisten Routinearbeiten überflüssig machen. Dem Menschen bleibt – in einer Welt automatisierter Fabriken und Dienstleistungen – gar keine andere Option als originelles Denken oder, wie Humanisten es gesagt hätten, die Suche nach Erkenntnis“ [Lo20].

In kaum einer anderen Branche hat sich der Einsatz von Machine Learning so flächendeckend durchgesetzt wie in der Spieleindustrie, der elektronischen wohlgemerkt. Kein Action Shooter, keine virtuelle Spielwelt kommt ohne Machine Learning in Form des Reinforcement Learning, des verstärkenden Lernens aus. Das ist nur den meisten Nutzern nicht so bekannt. Sie wissen, dass sich im Hintergrund vieles „automatisch“ tut, aber was genau?

Der Suchtcharakter vieler Spiele kommt genau daher, dass es einen künstlich erzeugten digitalen „Widerstand“ in der virtuellen Welt gibt, einen Gegenspieler in Form eines Agenten oder einer Spielaufgabe, die die eigene Intelligenz herausfordert und die eigene Vorstellungskraft beflügelt. Im Reinforcement Learning legt der Mensch die Regeln fest: Menschen kreieren einen realen (Robotik) oder virtuellen (Gaming) abgegrenzten Raum, in dem bestimmte Regeln gelten.

Diese Regeln können durch die Umgebung oder den Kontext gesetzt sein, sie können auch durch begrenzte Ressourcen bestimmt, oder einfach von Menschen definiert sein. Innerhalb dieses Raums müssen diese Regeln durch Vorgaben eingehalten werden, jede Aktion wird daraufhin überprüft, ob das geschieht oder nicht. Werden die Vorgaben eingehalten, kommt ein positives Signal zurück (positive Bestärkung), wenn nicht, erfolgt ein negatives Feedbacksignal (negative Verstärkung).

Hier wird also nicht eine „Künstliche Superintelligenz“ zum Einsatz gebracht, die durch eigenes Verstehen und Erkennen den Spielenden vor dem Bildschirm in Verzückung versetzt, sondern der Geist anderer Menschen hat Regeln und Spielräume definiert, die sich im Zusammenspiel mit dem Nutzer zu einem einzigartigen Erleben verzahnen.

Moderne Spielwelten leben von einem hohen Grad an Komplexität. Gaming eröffnet somit eine den meisten Menschen vertraute Form, um neues Verstehen zu schaffen. Warum also die Formen des Spiels nicht nutzen, um so abstrakte und verwirrende Denkräume wie die der Künstlichen Intelligenz den Menschen näher zu bringen? Und damit aus dem „Game“ ein Serious Game, ein ernsthaftes Spiel wird, geben wir dem Spiel und dem Einsatz des Machine Learning eine wichtige Note: ein relevantes Einsatzszenario, einen Use Case mit einem hohen „Need“, einen Anwendungsfall mit einem wichtigen Bedarf.

2.2 KI-Lösungen sichtbar machen und in relevante Kontexte einbringen

Noch immer werden KI-Lösungen nach dem Prinzip „machen, was machbar ist“ entwickelt, ohne zu sehen, wie die vorhandenen Ressourcen am sinnvollsten eingesetzt werden sollten und müssten. „When framing AI initiatives, at a minimum, you should answer: >What pain point does the AI solution solve?<. Even though this phase is the most crucial of any Business AI initiative, what I´ve observed is that it´s one of the most disregarded pahses“ [Ga22].

Auch beim aktuellen Rausch über die Möglichkeiten der Generative AI, und insbesondere der LLMs, ist wieder zu beobachten, dass selten langfristige und nachhaltige Ziele mit den neuen Use Cases anvisiert werden, sondern Anwendungen im Fokus stehen, die schnell auf den Markt gebracht werden können und eine kurzfristige Perspektive sowie ein unterkomplexes Verständnis der zugrundeliegenden Technologie zum Ausdruck bringen. Angesichts der weltweiten multiplen Krisenlagen sollte eine solche kurzfristige Sicht der Vergangenheit angehören.

Dabei gibt es mehr als genug Anwendungsfälle, bei denen wir die Skalierbarkeit und Geschwindigkeit von Machine Learning Applikationen dringend gebrauchen könnten: „Im März 2020 schlug die Europäische Kommission das erste Maßnahmenpaket zur Beschleunigung des Übergangs zu einer Kreislaufwirtschaft vor“ [EU23].

Im Gegensatz zu den vielen Anwendungen der Generativen KI, in denen weitreichende Risiken durch die Verwendungen von Daten mit ungeklärten Urheberrechten oder durch den Einsatz von personenbezogenen Daten entstehen können, arbeitet die überwiegende Anzahl der Machine Learning Anwendungen im Dienst der Kreislaufwirtschaft „Machine to Machine“ – hier werden also nur Maschinendaten oder Objektdaten übertragen. Solche Szenarien können also wesentlich einfacher implementiert und auf den Markt gebracht werden als die zahlreichen Anwendungen, die im Moment im Generative AI-Hype für den Einsatz für den Menschen konzipiert werden – etwa im Bereich der Medizin oder der Bildung.

Laut dem „Global Transboundary E-waste Flows Monitor 2022“ [UN22] von unitar, dem Trainings- und Forschungsinstitut der Vereinten Nationen, ist zum Beispiel bei 83% der globalen Ströme für Elektroschrott unklar, was mit ihnen passiert. Sicher ist nur, dass diese Wertströme weder für die Kreislaufwirtschaft aufbereitet werden, noch ökologisch angemessen entsorgt werden. Allein bei diesem Thema tut sich ein breites Anwendungsgebiet für die moderne KI auf. Jede Lösung, die hier zu einer Verringerung des Problems beiträgt, wird dringend gebraucht. Und das ist nur ein Aspekt der Kreislaufwirtschaft, der in den nächsten Jahren vorangebracht werden muss.

Die Ellen MacArthur Foundation macht sich seit längerem international für die Kreislaufwirtschaft stark. 2019 hat sie in Zusammenarbeit mit Google und McKinsey eine Studie zu den Möglichkeiten von „Artificial Intelligence and the circular economy- AI as a tool to accelerate the transition“ [EL19] veröffentlicht, die konkrete Potentiale des Machine Learning zur Umsetzung der Kreislaufwirtschaft sowie schon umgesetzte Best Practices beschreibt. „AI capabilities can help build a circular economy, at a faster rate than would be possible without AI. AI can boost development and design of completely new circular products and businesses. Equally, it can help traditional players in their transition to become more circular“ [EL19].

Alles, was es dafür braucht, ist Vorstellungskraft. Vorstellungskraft, wie neue Lösungen, neue globale Wertstoffströme und neue Ökosysteme für eine vernetzte globale Kreislaufwirtschaft aussehen könnten: „Neue Werkzeuge wie künstliche Intelligenz können dafür sorgen, dass die einzigen Grenzen, die wir erfahren, die unserer Vorstellungskraft und der Endlichkeit der Ressourcen der Welt sind“ [EL19].

Doch genau das könnte sich als Problem erweisen: Nicht die Technik bremst uns aus, sondern unser Unvermögen, die Zukunft anders zu imaginieren als die Vergangenheit.

2.3 Geistige Kreativität und zukunftsgerichtete Vorstellungskraft als Schlüssel für komplexe Probleme

Um nachhaltige und umsetzbare Lösungen für die aktuellen Probleme zu finden, können wir uns nicht von der Vergangenheit leiten lassen, denn unser Handeln in der Vergangenheit hat zu den heutigen Problemen geführt. Darum sollten wir uns auch nicht von den inneren Bildern leiten lassen, die wir in uns tragen, denn auch diese wurden von unseren Erfahrungen der Vergangenheit geformt.

Was nötig ist, sind neue Visionen von dem, was möglich ist und sein sollte: „Changing the way we see the world also changes the way we are in the world – and how we imagine promising development paths and their governance“ [Gö16].

Doch wie können wir zu neuen inneren Bildern über die Möglichkeiten, die Welt in der Zukunft zu organisieren, kommen? Indem wir auf das zurückgreifen, was schon jetzt eine große Auswirkung auf die inneren Bilder von Menschen hat, das Gaming. Spielen hilft nicht nur Kindern, sich ein Bild von Ihrer Umwelt und den Interaktionsmöglichkeiten

darin zu machen. Durch den spielerischen Umgang mit den Schnittstellen einer komplexen Problemlage können wir im Kopf Ideen ausprobieren, verwerfen oder weiterführen. Wir dürfen experimentieren, Fehler machen und von vorn beginnen, ohne das Schaden entsteht. „Experimenting is the best technique für generating data on what might work in the future“ [Ch11].

„Das ist die Phase des Erkundens und Experimentierens, in der sich der sich der Möglichkeitsraum öffnet, eine Vielfalt an unterschiedlichen Ansätzen entsteht, die sich miteinander austauschen und im Idealfall kombinieren lassen, um Mehrfachlösungen, also Multisolving, zu ermöglichen“ [Gö22]. Statt echter Risiken entsteht beim experimentellen Spielen ein freier Denkraum, in dem alte Vorstellungen ungestraft als überholt über Bord geworfen, und neue, gewagte Denkmuster ausprobiert und etabliert werden können. Das Spiel wird zum offenen Raum für neue Erkenntnisse und aktives Lernen, und damit zum ernsthaften Spiel, zum Serious Game.

Der Denkraum des Serious Games kann so zum Lösungsraum für komplexe Probleme werden, denn komplexe Probleme können nicht durch klare lineare Prozesse bewältigt werden. Eine vordefinierte Schrittabfolge kann nur verlässlich zum Ziel führen, wenn das Ziel bereits bekannt ist. Komplexe Probleme haben als Aufgabenstellung bewusst unscharf formulierte Ziele, um auch unbeabsichtigte, aber gute Lösungen erzielen zu können [Gr11].

2.4 Konstruktionismus als Grundlage für Serious Games und haptisches Denken

Der amerikanische Mathematiker Seymour Papert prägte den Begriff des Konstruktionismus [MI16]: Danach können Menschen durch aktives Konstruieren leichter lernen und neues Wissen generieren. Das konkrete Gestalten erzeugt gleichzeitig Theorien und Kenntnisse im Denken, und dieses neue Wissen ermöglicht den Umgang mit noch komplexeren realen Sachverhalten, was wiederum zusätzliches Wissen generiert: „Modelle machen Wissen nicht nur sichtbar, sondern durch das Bauen wird auch Wissen konstruiert“ [Kr14]. Seymour Paper, der sich am MIT auch der Erforschung von Künstlicher Intelligenz zugewandt hatte, beeinflusste Apple mit seinen Ideen und hat durch seine Arbeit Serious Games wie LEGO® Mindstorms mit angeregt.

Auf den Gedanken des Konstruktionismus beruht auch LEGO® Serious Play®, das 1996 von Per Kristiansen und Robert Rasmussen in Kooperation mit LEGO veröffentlicht wurde. Ihre Forschungen haben ergeben, dass Denkprozesse, die in Kombination mit dem Einsatz der Hände durchgeführt werden, durch die besonders zahlreichen neuronalen Hand-Hirn-Verbindungen zu einem tieferen Verständnis führen.

„Für das Verstehen verwenden wir im Deutschen das Wort >begreifen<. Es zeigt, wie sehr wir von der materiellen Welt abhängig sind, wie sehr wir, auch wenn es um reine

Geistesprodukte geht, uns etwas >vorstellen< müssen, was so viel heißt wie: einen Gedanken gegenständlich machen“ [Lo22].

Unter dem Begriff „Mapping“ wird das Kartografieren von Informationen in einer visuellen Übersicht verstanden: Mapping ermöglicht eine bessere Orientierung über den dargestellten Wissensraum und stellt somit die Grundlage für eine vertiefte Analyse einer komplexen Problematik dar. 3D Mapping als Methode des partizipativen Gestaltens ist eine interaktive Form des Wissensmanagements, in der kollektive Informationen und implizites (= verinnerlichtes) Wissen zu einer komplexen, multidimensionalen Problematik auf einer Fläche durch 3D Objekte und modellierte Strukturen abgebildet werden. Mit den Arbeitsformen von LEGO Serious Play ist ein 3D Mapping möglich, das einer heterogenen Gruppe von interessierten Personen den Austausch über eine komplexe Problematik auf neuartige Weise eröffnet.



Abb. 1: 3D Mapping zu IoT-Schnittstellen in der Bioökonomie, © Anna Hoffmann

3D Mapping kann auch als Form des Wissenstransfers in Veränderungsprozessen oder für partizipative Entscheidungsprozesse genutzt werden. Die Methoden dazu wurden durch das Presencing Institut [PI16] der Sloan School of Management am MIT im Rahmen der Entwicklung der „Theory U“ [Sc19] von Otto Scharmer evaluiert. Haptische Systemmodelle wie das 3D Mapping verschaffen ein gemeinsames, sichtbares Verständnis von der aktuellen Ausgangssituation, möglichen neuen Denk- und Veränderungsprozessen und dem erwünschten Zielzustand.

Gleichzeitig wird durch diese Herangehensweise eine vertiefte Reflexion angeregt, sowohl im eigenen Denken, als auch durch den Austausch mit der Gruppe, ohne dass die Thematik zu stark reduziert oder vereinfacht werden muss. „Reduzieren wir die Komplexität eines Problems, in das wir innerhalb eines Systems eingreifen, zu stark, verlieren wir nicht nur wichtige Ursachen, sondern genauso wichtige Effekte aus den

Augen. Deshalb ist es hilfreich, die Auswahl dessen, was wir für unser Problemverständnis als relevant erachten, immer wieder zu hinterfragen“ [Gö22].

3 Machine Learning basierte Zukunftslösungen für die Kreislaufwirtschaft durch 3D Mapping partizipativ gestalten

3.1 Arbeitskarten zu den Möglichkeiten des Machine Learning für die Kreislaufwirtschaft

Im Weiteren wird in Theorie und Praxis eine in vielen Workshops im Kontext meines Forschungsprojekts „Co:Decision“ [Ho23] erarbeitete Mixform von LEGO Serious Play, 3D Mapping und Serious Game vorgestellt, mit der eine gemeinsame Verständnisebene von den Möglichkeiten des Machine Learning zur Etablierung und Skalierung der Kreislaufwirtschaft erarbeitet werden kann. Die Wissensbasis dafür bildet die bereits erwähnte Studie der Ellen MacArthur Foundation in Zusammenarbeit mit Google und McKinsey „Artificial Intelligence and the circular economy- AI as a tool to accelerate the transition“. Um die damit einhergehende Informationsflut für Personen mit geringem Vorwissen verarbeitbar zu halten, ohne die Inhalte zu verflachen oder zu stark zu reduzieren, wurden für die wichtigsten Machine Learning Funktionalitäten im Kontext der Kreislaufwirtschaft Spielkarten entwickelt.



Abb. 2: Spielkarte zum Serious Game „Machine Learning und Kreislaufwirtschaft“, © Anna Hoffmann

3.2 Spielkarten als haptische Knoten in der Wissensarbeit

Die Spielkarten können an die Stelle von Machine Learning Schnittstellen auf der Spielfläche platziert werden und symbolisieren so eine Integrationschnittstelle im Datenstrom. Die Nichtgreifbarkeit von KI-Lösungen und ihre Unsichtbarkeit im Hintergrund von digitalen Anwendungen bekommt so Sichtbarkeit und plastische Handhabbarkeit, die allen Teilnehmenden als Orientierung dient.

Die Outcomes der Machine Learning Funktionen sind dabei bewusst knappgehalten: In dem Workshop steht die Gestaltung eines ganzen Ökosystems für die Kreislaufwirtschaft im Fokus, nicht die Möglichkeiten der einzelnen Machine Learning Algorithmen an sich. „Mehr und detailliertere Daten zu den einzelnen Teilen führen nicht zwangsläufig zu einem besseren Verständnis des Ganzen“ [Go19].



Abb. 3: Spielkarte zum Serious Game „Machine Learning und Kreislaufwirtschaft“, © Anna Hoffmann

„Es ist nachweisbar, dass Managemententscheidungen langfristig qualitativ besser ausfallen, wenn die Entscheider eventuell nur wenige, dabei aber die wirklich relevanten Informationen berücksichtigt haben im Vergleich zu Managern, die eine Maximierung ihres Informationsstands anstreben“ [Sc11].

Die Machine Learning Spielkarten erhalten damit im 3D Mapping die Funktionalität von „Wissensatomen“ oder „Knoten“: Bewegliche Objekte wie Spielkarten oder Bausteine sind praktische Hilfsmittel in der Wissensarbeit, da man mit ihnen komplexe Themen in

Einheiten zerlegen kann: in Wissensatome oder Knoten. „Jedes beliebige Objekt kann als Knoten bezeichnet werden, also als etwas, das Teil eines größeren Ganzen oder Systems ist“ [Br11].

Knoten ermöglichen zwei wichtige Funktionen in der Analyse komplexer Themen:

- Beziehungen können leichter ermittelt, analysiert und dargestellt werden und
- die Vielfalt in komplexen Systemen kann durch Knoten optisch so verdichtet werden, dass die Diversität und Fülle im Sichtfeld bleiben, ohne unzulässig reduziert oder vereinfacht zu werden.

„Indem man Objekte in der Umgebung platziert, setzt man Ideen auf dynamische Weise in einen Kontext. Kombinationen mehrerer Informationen können so lange bestehen, wie man will, und man kann sie im Handumdrehen mischen oder neu anordnen“ [Br11].



Abb. 4: Serious Game zu „Machine Learning und Kreislaufwirtschaft“, © Anna Hoffmann

Das Durchlesen, Platzieren und spätere Kommentieren der einzelnen Spielkarten hilft zudem, das Wissen um Machine Learning Funktionalitäten zu vertiefen und zu festigen: „Lernen umfasst sowohl die vorausgehenden Erfahrungen als auch die Integration und Verarbeitung von Informationen. Daraus werden neues Wissen abgeleitet und neue Handlungsmöglichkeiten eröffnet“ [Ed17].

Im Workshopverlauf folgt nach einem generellen Wissensinput zu den Möglichkeiten des Machine Learnings zur Kreislaufwirtschaft sowie aktuellen Herausforderungen wie der globalen E-Waste Problematik eine stille Phase der Einzelarbeit, in der die Teilnehmenden

sich orientieren und ihre eigenen Ansätze zu einem Kreislaufsystem gestalten können. Einzige Rahmenbedingung dabei ist, in dem eigenen Kreislaufmodell Schnittstellen für die Kreislaufsysteme der anderen Teilnehmenden einzubauen. „Jeder ist permanent dabei; gebaut wird in Metaphern und die Teilnehmer geben ihrem Modell eine Bedeutung“ [Be18].



Abb. 5: Haptisches Gestalten eines Kreislaufwirtschaftsmodells, © Anna Hoffmann

Durch die plastische Darstellung von Gesichtspunkten und Ideen werden die Teilnehmenden im Verlauf des Workshops dabei unterstützt, sich verbal klarer auszudrücken und sie entwickeln ein besseres gemeinsames Verständnis von der Thematik. Dafür stellt jeder Teilnehmende nach der Bauphase in Eigenarbeit sein Kreislaufsystem in der Gruppe vor, so dass in der nächsten Arbeitsphase ein gemeinsames Model aller Teilnehmenden entstehen kann. „Wer vernetzt denkt, tauscht Wissen, handelt sozial, ohne seine Interessen zu verleugnen“ [Lo20]. Damit das erfolgreich gelingen kann, sollte eine wesentliche Voraussetzung erfüllt sein: Psychologische Sicherheit.

3.3 Lernen erleichtern durch die Etablierung von „Psychologischer Sicherheit“

Psychologische Sicherheit umschreibt ein Arbeitsklima, indem es sicher ist, persönliche und zwischenmenschliche Risiken in der Kommunikation auf sich zu nehmen. Dieser Faktor ist immer dann gegeben, wenn Personen ohne persönlichen Schaden negative Kommentare oder schlechte Nachrichten direkt äußern können. Amy C. Edmondson, Professorin für Leadership and Management an der Harvard Business School, hat zu den

Auswirkungen von Psychologischer Sicherheit ausgiebig geforscht [Ed12]. Nach ihren Erkenntnissen ist Psychologische Sicherheit der einflussreichste Aspekt, um die Innovationskraft und Lernfähigkeit in Teams zu stärken.

„Wenn eine Gruppe per Kommunikationsregeln gezwungen ist, konstruktiv zu denken, ist plötzlich kein Platz mehr für Dinge, die mit der Idee, um die es geht, gar nichts zu tun haben“ [Er13]. Der Moderation kommt für einen gelingenden Lernprozess innerhalb des Workshops eine große Bedeutung zu. Die Facilitation führt die Gruppe durch den gesamten Prozess und ermöglicht für alle ein Grundgefühl von Sicherheit. Bewertungen in der Darstellung und Ausführung wie „falsch“ oder „richtig“ werden deshalb vom Facilitator verhindert oder unterbunden, ebenso Äußerungen der Teilnehmenden, die beleidigend, verletzend oder verunsichernd wirken können. Die Facilitation steuert das emotionale Erleben der Teilnehmenden bewusst mit, weil ohne ein Grundgefühl von Psychologischer Sicherheit kein wirksamer Co-kreativer Prozess entstehen kann.

„Es gibt eine weitverbreitete Angst, bloßgestellt zu werden oder zu scheitern. Die Vorstellung bevorstehender Schmach und Scham führt dazu, unseren Platz lieber in der zweiten Reihe zu suchen und den Schritt nach vorne zu vermeiden. Solange wir uns aber nicht würdig fühlen, unsere Perspektive, unsere Ideen, uns als Person hörbar werden zu lassen, schließen wir tatsächlich ein, unsere eigenen intellektuellen Fähigkeiten zu sabotieren“ [Er13].

3.4 Förderung des dialogischen Austauschs

Wenn Menschen ihrer Vorstellungskraft Form und Gestalt geben, indem sie Ideen, Konzepte oder Strategien konstruieren und externalisieren – sie greifbar und teilbar machen – können sie sie nicht nur selbst besser reflektieren, sondern andere dazu einladen, mit ihnen gemeinsam zu reflektieren.

Das partizipative Bauen von haptischen Modellen sowie der dialogorientierte Austausch darüber führen deshalb häufig zu Erkenntnissen, die sich in bloßen Diskussionen über das Thema nicht gezeigt hätten. „Die Sprache ist jedoch unser mächtigstes Werkzeug – auch zur Weiterentwicklung von Ideen“ [Ep14].

Deshalb ist das Sprechen über die gestalteten Kreislaufsysteme so wichtig wie das eigentliche Bauen. Ziel des Workshops ist es, einen gemeinsamen Gedankenraum zu erzeugen. Das geschieht ebenso über die Bilder wie über die Sprache. Beides trägt zu einer gemeinsamen Verständnisebene bei. „Die tatsächliche Darstellung sorgt dafür, dass die Geschichten bei den Teilnehmern sehr gut haften bleiben“ [Be18].

3.5 Informationen werden zu anwendbarem Wissen und eröffnen neue Handlungsräume

Während des Workshops haben die Teilnehmenden die Gelegenheit, sich frei im Raum zu bewegen. Es gibt genügend Freiflächen auf Tischen oder dem Boden, um die zu gestaltenden Kreislaufsysteme groß denken und umsetzen zu können. „Wir wissen, dass es uns leichter fällt, kreativ zu sein, wenn wir uns bewegen können, nicht am Tisch sitzen müssen und viel freie Fläche zum Beschreiben und Skizzieren haben“ [Ep14].

Der gemeinsame Austausch, das eigene Überlegen und Experimentieren, die Möglichkeit und vielleicht auch die Notwendigkeit, durch ein gemeinsames Model das eigene Kreislaufsysteme zu adaptieren, schafft optimale Bedingungen für iterative Lernprozesse mit nachhaltiger Wirkung. „Der Geist arbeitet besser, wenn er ein umfassenderes und weniger einseitiges Bild der Wirklichkeit hat“ [Mi18].

Durch das haptische Gestalten und den Austausch in der Gruppe wird aus den aufgenommenen Informationen anwendbares Wissen und aus den gemachten Erfahrungen entfalten sich neue Handlungsräume. Dafür muss das Gehirn vom Modus des konvergenten Denkens in den Modus des divergenten Denkens überwechseln, der durch eine große Offenheit für neue Erfahrungen und Ideen geprägt ist. Der Neurowissenschaftler und Kreativitätsexperte Henning Beck formuliert es so: „Das Gehirn zerlegt das Problem in Einzelteile, vergleicht es mit schon bekannten Lösungsmustern und kann dann neue Lösungen erzeugen“ [Be20].

Das hier vorgestellte Workshopformat unterstützt das agile Arbeiten in Teams und fördert eine innovationsfreundliche Organisationskultur. Durch den spielerischen Umgang mit den großen wirtschaftlichen und technologischen Herausforderungen der Transformation zur Kreislaufwirtschaft trägt der Workshopprozess zu einer partizipativen Lern- und Kollaborationsform in Unternehmen bei, die nachhaltige Spuren hinterlässt. „Grundsätzlich sind in fast allen agilen Praktiken die Aspekte des Experiments und des Feedbacks, also des schnellen Lernens enthalten, um komplexe Zusammenhänge zu erfassen und diese optimal zu vereinfachen“ [Go19].

4 Ausblick: Menschliche Kreativität in Zeiten von Generative AI

„Die Befähigung und Ermächtigung der Mitarbeiter in den wertschöpfenden Bereichen ist die zentrale Grundlage für Erfolg“ [Ed17]. Zu dieser Befähigung gehört vor allem auch die Befähigung zum Denken, Entscheiden und Handeln in komplexen Zusammenhängen. Mitarbeitende wollen zunehmend nicht nur Einblick in die Kontexte ihres eigenen Arbeitsplatzes haben, sondern den größeren Gesamtkontext ihres Wirkens einsortieren können.

„Agilisten akzeptieren nicht nur die Komplexität, sie begrüßen sie sogar, weil dies zu kreativen Höchstleistungen anspornen kann“ [Go19]. Die vielen Ermahnungen zu „Sinn“ und „Purpose“ in der Arbeitswelt zielen ja genau darauf ab: die eigene Wertschöpfung in einem größeren Zusammenhang sehen zu können, der das eigene Handeln sinnvoll macht. Genau das stärkt die individuell erlebte Selbstwirksamkeit, und das fördert wiederum die Resilienz des einzelnen Mitarbeitenden und in Summe die Resilienz der Organisation. „Wer von der Welt um sich herum nichts weiß, läuft immer mehr Gefahr, ins Leere zu arbeiten und zu leben“ [Lo20].

Zum neuen Weltwissen gehören Grundkenntnisse von der Funktionsweise und den Chancen und Risiken der Methoden der Künstlichen Intelligenz, insbesondere des Machine Learning. Seit dem geräuschvollen Aufschlag der Generative AI, der schöpferischen Künstlichen Intelligenz, Anfang 2023 ist der Mehrheit der Menschen klar, dass eine neue Form der digitalen Technik in ihre Arbeits- und Alltagswelt Einzug gehalten hat, und dass diese Technologie das Potential hat, viele vertraute Handlungs- aber auch Denkweisen zu verändern.

Maja Göpel listet auf, welche Fähigkeiten in der Zukunft benötigt werden, damit sich in der Gesellschaft insgesamt die Befähigung zum Umgang mit völlig neuen Situationen erhöht: „Neugierde und kritisches Denken gehören genauso dazu wie Selbstwirksamkeit, Multiperspektivität, Kooperationsvermögen, kulturelle Sensitivität, digitale Resilienz und Kommunikationsvermögen“ [Gö22]. All diese Fähigkeiten können durch das Format von haptischen 3D Mappings zu komplexen Problemlagen gestärkt werden.

E-Learning allein wird uns nicht die Befähigung zum mündigen Umgang mit einer KI-basierten technologischen Welt zurückbringen, dafür braucht es einen partizipativen Lern- und Entscheidungsprozess, der den Menschen mit allen seinen Sinnen anspricht, abholt und einlädt. Denn: „Mitdenkende Arbeitskräfte werden an Bedeutung gewinnen, und deshalb wird ihre gezielte Integration in die durch künstliche Intelligenz geprägte Zukunft von entscheidender Bedeutung sein“ [Go19].

Literaturverzeichnis

- [Go19] Gomez, P.; Lambertz, M; Meynhardt, T: Verantwortungsvoll führen in einer komplexen Welt, Hauptverlag, 2019.
- [Gö23] Göpel, M: Wir können auch anders, Ullstein Verlag, 2022.
- [Gö16] Göpel, M: The great mindshift, Wuppertal Institut, Springer Open, 2016.
- [Lo20] Lotter, W: Zusammenhänge – Wie wir lernen, die Welt wieder zu verstehen, Edition Körber, 2020.
- [EU23] EU, Europäisches Parlament, <https://www.europarl.europa.eu/news/de/headlines/society/20210128STO96607/wie-will-die-eu-bis-2050-eine-kreislaufwirtschaft-erreichen>, Stand; 14.7.2023.

- [Ga22] Ganesan, K: The Business Case for AI, Opinois Analytics Publishing, 2022.
- [UN22] UN, unitar, <https://ewastemonitor.info/gtf-2022/>, Stand: 14.7.2023.
- [El19] Ellen Macarthur Foundatain, <https://ellenmacarthurfoundation.org/artificial-intelligence-and-the-circular-economy>, Stand: 14.7.2023.
- [Kr14] Kristiansen, P.; Rasmussen, R.: Building a Better Business Using the Lego Serious Play Method, Wiley, 2014.
- [Ed12] Edmondson, A.C.: Teaming – How Organizations learn, innovate and compete in the knowledge economy, Harvard Business School Press/Wiley, 2012.
- [MI16] MIT, <https://news.mit.edu/2016/seymour-papert-pioneer-of-constructionist-learning-dies-0801>, Stand: 14.7.2023.
- [Ch11] Christensen, C.M.; Dyer, J.; Gregersen, H.: The Innovator`s DNA, Harvard Business Review Press, 2011.
- [PI16] Presencing Institute, MIT, <https://www.u-school.org/>, Stand: 14.7.2023.
- [Ho23] Forschungsprojekt Co.Decision, <https://www.co-decision.de/>, Stand: 14.7.2023.
- [Br11] Brown, S.; Gray, D; Macanufo, J.: Gamestorming, O`Reilly 2011.
- [Sc19] Scharmer, C.O.: Essentials der Theorie U: Grundprinzipien und Anwendungen, Carl-Auer Verlag, 2019.
- [Sc11] Schmidt, G.: Organisation – Aufbauorganisatorische Strukturen, ibo Schriftenreihe, 2011.
- [Be20] Beck, H: Das neue Lernen heißt Verstehen, Ullstein Verlag, 2020.
- [Be18] Berndt, C.; Nickel, S.: Let`s change mit innovativen Tools, Haufe Verlag, 2018.
- [Ed17] Edelkraut, F.; Gramß, D.; Graf, N.: Agiles Lernen, Haufe-Lexware, 2017.
- [Ep14] Eppler, M.; Hoffmann, F; Pfister, R.: Creability – Gemeinsam Kreativ, Schäfer-Poeschel, 2014.
- [Er13] Erbdinger, J; Ramge, T.: Durch die Decke denken, Redline Verlag, 2013.
- [Mi18] Miketta, M.: Thinking Environment, Jungfermann, 2018.

Ökologische Nachhaltigkeit - KIU-2023

4. Workshop Künstliche Intelligenz in der Umweltinformatik

Andreas Abecker¹, Julian Bruns² und Stefan Naumann³

Abstract: Im Rahmen des INFORMATIK FESTIVAL 2023 der Gesellschaft für Informatik (GI) e.V. im Herbst 2023 in Berlin findet die vierte Auflage des Workshops KIU zur Nutzung von Methoden der Künstlichen Intelligenz in der Umweltinformatik statt. In der KIU-Workshopreihe werden seit 2020 anwendungsorientiert und interdisziplinär innovative Beiträge der KI für wichtige Fragen von Umweltschutz und Nachhaltigkeit vorgestellt und diskutiert. Auch der vierte Workshop soll dabei helfen, eine deutschsprachige Wissenschafts- und Anwendungscommunity zu diesen Themen zu etablieren, um langfristig die Kreativität und die Wirkung dieses wichtigen Aufgabenfelds zu unterstützen.

Keywords: Umweltinformatik, Künstliche Intelligenz, Maschinelles Lernen, Umweltmonitoring, Umweltschutz, Ressourceneffizienz, Nachhaltigkeit

1 Motivation

Im aktuellen Sommer 2023 werden die Folgen des von Menschen verursachten Klimawandels überdeutlich: Rekordtemperaturen werden weltweit gemessen; Hitze- und Trockenperioden inklusive Waldbränden wechseln sich ab mit Starkregenereignissen und Unwettern. In den großen Städten und Ballungsräumen wird immer deutlicher, dass es für Vermeidungsstrategien bereits zu spät ist, dass man auch mit Hochdruck Begrenzungsstrategien zum Klimawandel und Anpassungsmaßnahmen zur Resilienzsteigerung durchführen muss. Viele genauso erschreckende Bedrohungen für die Umwelt und die natürlichen Ressourcen bleiben weiterhin auf hohem Niveau ernst, vom Artensterben bis zum Mikroplastik in den Meeren und der Nahrungskette. Wenn die Menschheit auf dem gegenwärtigen zivilisatorischen und Wohlstandsniveau überleben will, sind schnelle und umfassende Transformationsprozesse von Wirtschaft und Gesellschaft unabdingbar. Dabei zeigt schon der aktuelle Streit um das Gebäudeenergiegesetz, dass diese Transformationsprozesse (Energie, Produktion, Transport und Mobilität, Freizeit, Ernährung, Wasserversorgung, ...) sehr schmerzhaft und sehr konfliktbelastet sein werden.

Neuen Technologien und intelligenten Lösungen muss dabei eine zentrale Rolle zukommen, um die erforderlichen Transformationsprozesse effektiver und effizienter zu gestalten, ihre Wirkung zu erhöhen und ihre Kosten zu reduzieren. Dabei hat die Informatik nach dem Big

¹ Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe andreas.abecker@disy.net

² dmTECH GmbH, Am dm-Platz 1, 76227 Karlsruhe

³ Hochschule Trier, Umwelt-Campus Birkenfeld, Institut für Softwaresysteme, Postfach 1380, 55761 Birkenfeld
s.naumann@umwelt-campus.de

Data Hype in den frühen 2000er Jahren und dem Deep Learning Hype seit den 2010er Jahren in der jüngsten Vergangenheit mit dem ChatGPT Hype um generative Sprachmodelle enorme Aufmerksamkeit bekommen, weil das Potential von KI-Lösungen dramatisch wächst und auch zunehmend bekannt und verstanden wird. Insofern ist die Zeit jetzt reif, dass Methoden der Informatik und der Künstlichen Intelligenz eine zentralere Rolle spielen, wenn es darum geht, komplexe natur-, umwelt-, erd- und lebenswissenschaftliche Zusammenhänge besser zu erfassen, zu überwachen, zu verstehen, mit den ökonomischen und gesellschaftlichen Prozessen zu verknüpfen und auf dieser Basis bessere technische Lösungsansätze zum Umgang mit den großen Umweltproblemen zu entwickeln.

In der KIU-Workshopreihe versuchen wir in diesem sehr großen Themenraum seit 2020, anwendungsorientiert und interdisziplinär neue innovative Ideen zu finden, offene Forschungsfragen und Lösungsansätze zu diskutieren und konkrete Beispiele aus der Kombination von KI und Umwelthanwendungen zu präsentieren.

2 Hintergründe zur Umweltinformatik und verwandte Gebiete

Die Umweltinformatik befasst sich interdisziplinär mit der Analyse und Bewertung von Umweltsachverhalten [Kr18]. Aus IKT-Sicht spielen dabei traditionell beispielsweise Simulationen komplexer Systeme, Geographische Informationssysteme und räumliche Datenanalyse, Messnetze und Sensordatenverarbeitung eine große Rolle [Fi21], die Themen Fernerkundung und Bildverarbeitung sind in jüngerer Zeit mit enormem Potential hinzugetreten.

Die Umweltinformatik findet ihre wichtigsten Anwendungen traditionell noch in der Wissenschaft und in der öffentlichen Verwaltung (Natur- und Umweltschutz, Umweltdatenportale, Katastrophenschutz, Verbraucherschutz, Wassermanagement etc.). Mit Blick auf die oben erwähnten notwendigen Transformationen in vielen Bereich unseres gegenwärtigen Lebens und Arbeitens, kommen in den vergangenen Jahren auch zunehmend Lösungsansätze aus Wirtschaft und Gesellschaft hinzu, um dort ressourcenschonendere Lebens-, Arbeits- und Produktionsweisen zu schaffen. Denn dort sind natürlich die großen Hebel für wirkmächtige Veränderungen zu finden. Der Staat kann Rahmenbedingungen vorgeben, bestimmte Phänomene monitoren, vielleicht auch incentivieren; aber am Ende sind bspw. die staatlichen Investitionen nur ein Bruchteil der privatwirtschaftlichen Investitionen und Ausgaben. Private oder bestenfalls „halbstaatliche“ Strukturen (wie kommunale Betriebe) dominieren überwiegend oder vollständig solche wichtigen Sektoren und Systeme wie:

- **Bauen und Wohnen**
 - mit enormem Anteil am Primärenergieverbrauch sowie Treibhausgasemissionen durch mineralische Bauweisen

- wo bspw. Auswertungen von Wärmebildkameras oder von Luftbildern die Planung von Sanierungsmaßnahmen oder EE-Installationen unterstützen können oder kommunale Informationssysteme die Nutzung von Holz als Baustoff fördern können [Ze23]
- Mobilität und Transport
 - mit großen Auswirkungen auf die Nutzung fossiler Energieträger (Verbrennungsmotoren), aber auch auf Luftqualität in Ballungsräumen oder auf Mikroplastik in der Natur (Reifenabrieb)
 - wo bspw. intelligent vernetzte multimodale Verkehrssysteme den Individualverkehr mit ÖPNV erleichtern⁴ oder besser abgestimmte Logistiksysteme einen effizienteren Ressourceneinsatz ermöglichen
- Ernährung, Land- und Forstwirtschaft
 - mit großen Auswirkungen auf Boden und Wasser (z.B. Nitratbelastung des Grundwassers durch Dünger), die Artenvielfalt (Insektensterben durch Pflanzenschutzmittel, PSM, und Monokulturen) u.v.m.
 - wo bspw. teilflächenspezifische Düngung aufgrund intelligenter Sensordatenauswertung Düngemittel einsparen kann [He23], Agrarroboter die Schädlingsbekämpfung ohne PSM ermöglichen u.v.m.

Dies sind nur einige wichtige Beispiele, wie große Sektoren unseres Arbeitens und Wirtschaftens sowohl die Umwelt stark beeinflussen (mit erheblichem Schadenspotential) als auch enorme Chancen für Verbesserungen mit KI-basierten oder KI-unterstützten neuen Lösungen bieten. In diesen Bereich fällt natürlich insbesondere auch der traditionelle Bereich Abfall und Entsorgung mit der moderneren Herangehensweise der Kreislaufwirtschaft. Hier finden sich bspw. KI-Ansätze für die bessere Trennung wiederverwendbarer Bestandteile von Müll durch Bilderkennung, bis hin zu Produktpässen für komplexe Produkte, um die spätere Trennung und Wiederverwendung wertvoller Materialien zu ermöglichen.

Da, wie bisher dargestellt, fast alle Lebensbereiche relevante Umweltbezüge besitzen, ist der Ansatz unserer „Umwelt-KI“ offensichtlich eng verbunden und teilweise überlappend mit vielerlei anderen anwendungsorientierten Informatik-Unterdisziplinen. So gibt es zum Beispiel Verbindungen in die Agrarinformatik (reduzierter Einsatz von Pflanzenschutzmitteln, Dünger oder Wasser durch intelligente Sensorik und Steuerung), die Hydroinformatik (datengetriebene, ML-basierte Methoden zum Verständnis komplexer hydrologischer Prozesse), die Energieinformatik (Steuerung virtueller Kraftwerke auf Basis von EE-Technologien und Abstimmung mit Netz und Nutzern) und viele mehr.

Außerdem versucht sich in den vergangenen Jahren, insbesondere außerhalb Deutschlands, ein neues interdisziplinäres Forschungsfeld unter Schlagworten wie „ICT for Sustainability

⁴ vgl. das 2023 gestartete BMUV-Leuchtturmprojekt „GreenTwin: Grüner Digitaler Zwilling mit Künstlicher Intelligenz für CO2-sparende Kooperative Mobilität und Logistik im ländlichen Raum“

(ICT4S)“ oder „Computational Sustainability“ zu konstituieren. Hier geht es einerseits um sehr konkrete Fragestellungen des „Green by IT“, wie wir sie oben bereits diskutiert haben, z.B. im Bereich Smart Grid oder Vernetzte Mobilitätslösungen. Andererseits adressiert man aber auch die KI-basierte Unterstützung originärer Umwelt-, Naturschutz- und Nachhaltigkeitsfragestellungen, wie z.B. mit Bilderkennungsverfahren für die drohnenbasierte Wildtierbeobachtung oder mit auf Reinforcement Learning basierender Entscheidungsunterstützung für die Priorisierung von Schutzgebieten [Si22].

Außerdem stellt der Ressourcenverbrauch (Strom, Kühlung, Materialien für die Hardware) von IKT-Systemen in den letzten Jahren einen stark wachsenden, bereits beträchtlichen Teil der gesamtgesellschaftlichen Ressourcenverbräuche dar. Da gerade die sehr leistungsfähigen modernen Verfahren der KI, wie Convolutional Neural Networks oder Large Language Models, in der Regel echte „Ressourcenfresser“ sind, muss die Community sich darum kümmern, dass die entsprechenden Systeme noch Teil der Lösung bleiben und nicht ein relevanter Teil des Problems werden. Im Gegensatz zu den bisher diskutierten Themen des Ansatzes „Green by IT“, gehört zu einem wohlverstandenen Umwelt-KI Ansatz also auch der Aspekt des „Green in IT“, also der ressourceneffizienten Erbringung der KI-Dienste. Hier geht es zunächst einmal um ein gutes Verständnis der Ressourcenverbräuche⁵, optimierte Algorithmik, Green Coding, unter Umständen auch um grundlegende Architekturüberlegungen zur Ressourcenschonung, z.B. durch Ansätze wie Edge Computing oder Fog Computing, bis hin zur Zertifizierung von Software unter Aspekten des Umweltschutzes⁶.

3 Warum KI in der Umweltinformatik?

Die Umweltinformatik betrachtet zumeist komplexe ökologische Prozesse, die (noch) nicht vollständig bekannt, verstanden oder verstehbar sind (chaotische Systeme), deren Verhalten nur approximativ oder vereinfachend modelliert, simuliert oder vorhergesagt werden kann oder soll (z.B. aus Aufwandsgründen) und bei deren Beobachtung zuweilen wichtige Größen unbekannt sind oder nur geschätzt werden können. Zusätzlich müssen bei der Betrachtung von Ökosystemen noch weitere Systeme, die aus Informatiksicht nicht einfach zu behandeln sind, wie z.B. das Wetter, betrachtet werden, da sie direkte oder indirekte Einflüsse auf die betrachteten Systeme ausüben. Aus Sicht der Anwendung müssen häufig Entscheidungen unter Unsicherheit getroffen und schwierige Abwägungen in Zielkonflikten getroffen werden, die Fach- und Erfahrungswissen erfordern (z.B. in Planungsverfahren, im Notfallmanagement, bei der Politikgestaltung). Die entsprechenden Fragestellungen haben vielfältige Wechselwirkungen zu schwierigen Themen wie z.B. Klimawandel, Energiewende, Biodiversität und Nachhaltigkeit, aber auch mit großen Technologietrends wie Smart Cities, Smart Agriculture, Smart Mobility oder Smart Grids.

⁵ vgl. das 2023 gestartete BMUV Leuchtturmprojekt „KIRA: KI-Referenzmodell für Energie- und Ressourceneffizienz und dessen industrielle Anwendung“; <https://www.umwelt-campus.de/iss/projekte/laufende-projekte/kira>

⁶ vgl. hierzu den „Blauen Engel Software“, siehe z.B. <https://www.umwelt-campus.de/green-software-engineering/refoplan-20>

Die Anwendung von Methoden und Technologien der KI drängt sich also auf, für den Umgang mit unvollständigem Wissen, für die Entscheidungsunterstützung bei komplexen Fragestellungen, zum datengetriebenen Verstehen komplexer Systeme bis zur Prognose.

Im Rahmen der INFORMATIK 2020 fand der Workshop KIU daher das erste Mal statt, mit der Idee, eine neue Community zum Themenfeld „KI und Umweltinformatik“ zu begründen. Teilnahme und Verlauf von KIU-2020 waren ermutigend. Es gab einen eingeladenen und 8 eingereichte und begutachtete Vorträge. Bei der Online-Durchführung waren permanent etwa 30 Zuhörer:innen präsent, in Spitzenzeiten waren es sogar bis zu 50 – bei insgesamt über 100 unterschiedlichen Tagungsteilnehmern im Workshop. Auch bei der zweiten Durchführung KIU-2021 gab es einen eingeladenen sowie 8 begutachtete Fachbeiträge und gute Teilnehmerzahlen. KIU-2022 fand in Hamburg erstmals als Präsenzveranstaltung statt, jedoch vor Ende der Corona-Pandemie noch vor kleinem Publikum – was jedoch der Tiefe der fachlichen Diskussionen und des persönlichen Austauschs durchaus zuträglich war.

In den vergangenen wenigen Jahren hat sich, zumindest in der Wissenschaft, aufgrund der KI-orientierteren Förderlandschaft, die Szene massiv weiterentwickelt. Ein Programm wie die „KI-Leuchttürme“ des BMUV umfasst inzwischen vielfältigste FuE-Projekte auf den meisten relevanten Umwelt- und Nachhaltigkeitsgebieten und mit einer gewissen Methodenvielfalt.⁷ Aber auch in anderen Fachgebieten mit großem Umweltbezug setzen sich KI-Ansätze durch. Zum Beispiel verfolgen mehrere der aktuell vom BMBF in der Fördermaßnahme „Wasser-Extrem-Ereignisse (WAX)“ geförderten Verbundprojekte zumindest auch KI-basierte Methoden, zum Beispiel für die Ausweisung von Notabflusswegen bei Starkregenereignissen.⁸

Die Szene der mit Umwelt und KI befassten Forschenden und Anwenderinnen wächst also. Dennoch sehen wir zurzeit noch keine Konstituierung oder Konsolidierung der Community. Es sollte unser Ziel sein, dies etwas voranzutreiben.

4 Thematische Schwerpunkte des Call-for-Papers

Auch wenn klar ist, dass zurzeit der KIU-Workshop noch ein überschaubares Event bleibt, das nur wenige technische und fachliche Fragestellungen konkret ansprechen wird, ist der Call-for-Papers jedes Jahr bewusst breitestmöglich gehalten, um alle potentiell an der KIU-Community Interessierten auch zu erreichen. Daher hatte die nicht-ausschließliche Themenliste beträchtliche Ausmaße:

KI-Technologien für die Umwelt (nicht ausschließliche Liste):

- Big / Smart / Linked / Open Data
- Bildverarbeitung und Fernerkundung

⁷ vgl. <https://www.z-u-g.org/foerderung/ki-leuchttuerme-fuer-umwelt-klima-natur-und-ressourcen/>

⁸ vgl. <https://www.fona.de/de/massnahmen/foerdermassnahmen/wasserextremereignisse.php>

- Data Mining, Machine Learning, Deep Learning
- Digital Twins
- ELSI-Aspekte zur Umwelt-KI, Responsible AI
- Explainable AI, Human-Centered AI
- Multiagentensysteme
- Multimodale Interfaces, AR, VR und KI
- Robotik
- Semantische Technologien
- Soft Computing / Computational Intelligence
- Spatial Data Mining, Location Intelligence
- Sprachverarbeitung, Chatbots, Generative AI
- Unsicherheit und Vagheit
- Case-Based Reasoning, Wissensbasierte Systeme und Wissensmanagement

Umweltanwendungen mit KI-Unterstützung (nicht ausschließliche Liste):

- Betriebliche und behördliche Umweltinformationssysteme, Umweltmonitoring
- Biodiversität und Artenschutz
- Computational Sustainability
- Erneuerbare Energien und Energiewende
- Green IT und Energiemanagement
- Katastrophenschutz und -management (aus Umweltsicht)
- Klimawandel und Klimaanpassung
- Nachhaltige Produktion, Kreislaufwirtschaft
- Natur- und Umweltschutz inkl. Schutz der Meere
- Ressourcenschutz und Landmanagement
- Smart Agriculture, Smart Forestry (aus Umweltsicht)
- Smart City, Smart Mobility (Umweltaspekte)
- Umweltbildung
- Verbraucherschutz (Umweltaspekte)

- Wasser 4.0

5 Eingereichte Fachbeiträge

Aus Sicht der Organisatoren ist es schade, dass wir 2023 nur vier Vorträge zur Präsentation haben. Dies mag auch daran liegen, dass das Umwelt- und Nachhaltigkeitsthema so allgegenwärtig geworden ist, dass man sich teilweise selbst Konkurrenz macht, wie hier z.B. mit einem ganzen eigenen Workshop zur Kreislaufwirtschaft, die wir inhaltlich als Untermenge unserer Thematik sehen. Wir werden versuchen, bei der Zeitplanung des Workshops zumindest zu ermöglichen, dass man KIU-2023 mit einem dieser eng verwandten Workshops quasi als „Doppelpack“ zeitlich nacheinander besuchen kann.

Positiv ist dennoch zu erwähnen, das auch diese nur vier Beiträge zusammen ein durchaus noch interessantes und breites Themen- und Methodenfeld abdecken. Teilweise in Ergänzung und Fortführung früherer KIU-Beiträge, teilweise ganz neu. Wir hoffen, dass wir die geringe Anzahl von Vorträgen eine umso intensivere fachliche Diskussion ermöglicht. Konkret sind im Workshop KIU-2023 vertreten:

- Der Beitrag „Sustainability in Artificial Intelligence - Towards a Green AI Reference Model“ von Sebastian Weber, Achim Guldner, Stefan Naumann, Lejla Begic Fazlic und Guido Dartmann befasst sich mit Grundlagen zur Übertragung der Idee des Green Software Engineering auf KI-Systeme.
- Der Beitrag „3D Printing and Sustainability: A Web Analysis Approach“ von Robert Dehghan, Sebastian Schmidt und Julian Schwierzy setzt Methoden der Web Science und des Text Mining ein, um die Nachhaltigkeitsorientierung von Firmen im Bereich 3D-Druck zu bewerten.
- Der Beitrag „Landwirtschaftliche Ertragsvorhersage im Kontext begrenzter realer Trainingsdatensätze: ein Transfer-Learning-Ansatz unter Verwendung tiefer neuronaler Netze“ von Alexander Münzberg, Christian Troost, Nils Reinosch, Daniel Martini, Liv Seuring, Alexander Niehus, Rajiv Srivastava, Thilo Streck, Thomas Berger und Ansgar Bernardi rundet frühere Beiträge der Autoren im Bereich Smart Farming ab, wo es darum geht, durch Transferlernen intelligente Optimierungsmethoden in der Landwirtschaft auch bei dünner Datenlage effizient einzusetzen.
- Der Beitrag „Nutzung offener Standards für die Integration von Pegel- und KI-basierte Prognose von Pegel- und Abflussdaten zur Verbesserung der Frühwarnung bei Sturzfluten“ von Jonathan Vogl, Desiree Hilbring und Divas Karimanzira setzt ebenfalls frühere Arbeiten fort und nutzt den OGC Standard SensorThings API, um neuronale Netze für die Sturzflutfrühwarnung in Smart City Kontexten einzusetzen.

6 Programmkomitee

- Dr. Ansgar Bernardi; Deutsches Forschungszentrum für Künstliche Intelligenz, Kaiserslautern; <https://www.dfki.de/>
- Dr. Matthias Budde; Disy Informationssysteme GmbH, Karlsruhe; <http://www.disy.net/>
- Prof. Dr. Frank Fuchs-Kittowski; Hochschule für Technik und Wirtschaft, Berlin; <https://www.htw-berlin.de/>
- Dr. Katharina Glock; FZI Forschungszentrum Informatik, Karlsruhe; <https://www.fzi.de/>
- Dr. Desirée Hilbring; Fraunhofer IOSB, Karlsruhe; <https://www.iosb.fraunhofer.de/>
- Dr. Julian Huber; MCI The Entrepreneurial School, Innsbruck; <https://www.mci.edu/de/>
- Prof. Dr. Christian Jolk; Technische Hochschule Ostwestfalen-Lippe, Höxter; <https://www.th-owl.de/>
- apl. Prof. Dr. Sven Lautenbach; Universität Heidelberg, Heidelberg; <https://www.geog.uni-heidelberg.de/>
- Dr. Tanja Liesch; Karlsruhe Institut für Technologie, Karlsruhe; <http://www.kit.edu/>
- Prof. Dr. Jens Nimis; Hochschule Karlsruhe Technik und Wirtschaft, Karlsruhe; <https://www.hs-karlsruhe.de/>
- Dr. Steffen Thoma; FZI Forschungszentrum Informatik, Karlsruhe; <https://www.fzi.de/>
- Dr. Marc Wieland; Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Oberpfaffenhofen; <https://www.dlr.de/eoc>

Literatur

- [Fi21] Fischer-Stabel, P.: Umweltinformationssysteme: Grundlagen einer angewandten Geoinformatik und Geo-IT. 3., vollständig neu bearbeitete und erweiterte Auflage. Wichmann, 2021, ISBN: 9783879077007.
- [He23] Heiß, A.; Paraforos, D. S.; Sharipov, G. M.; Ullrich, P.; Bruns, J.; Abecker, A.; Griepentrog, H. W.: Versatile and user-centered concept for temporally and spatially adapted nitrogen application based on multiple parameters. *European Journal of Agronomy* 145/Article 126792, 2023, ISSN: 1161-0301.
- [Kr18] Krallmann, H.: *Umweltinformatik: Informatikmethoden für Umweltschutz und Umweltforschung*. De Gruyter, 2018, ISBN: 9783486789348.

- [Si22] Silvestro, D.; Gorla, S.; Sterner, T.; Antonelli, A.: Improving biodiversity protection through artificial intelligence. *Nature Sustainability* 5/5, <https://doi.org/10.1038/s41893-022-00851-6>, 2022, ISSN: 1161-0301.
- [Ze23] Zernicke, C.; Hafner, A.; Abecker, A.; Stolpe, H.: WEB-GIS-TOOL: Estimation of greenhouse gas savings due timber use in the urban built environment. In (Nyrud, A. Q.; Malo, K. A.; et al., Hrsg.): *World Conference on Timber Engineering 2023 (WCTE2023), Timber for a Livable Future*. <https://doi.org/10.52202/069179-0521>, Curran Associates, Inc., S. 4010–4019, 2023.

3D Printing and Sustainability: A Web Analysis Approach

Robert Dehghan ¹, Sebastian Schmidt ², and Julian Schwierzy ³

Abstract: Besides manifold advantages in production, 3D printing is also often touted as a green technology. However, the quantitative connection between its use and sustainability is unclear so far. In our study, we examined all companies in the DACH region applying a web analytics approach. Using web scraping and deep learning, we assessed the 3D printing and sustainability intensity for each company. Our results suggest that 3D printing companies are significantly more likely to present themselves as sustainable than other companies. However, as our analysis is purely descriptive, correlations and causality still need to be clarified in future research. Our web mining methodology is a promising approach to study economic phenomena with an environmental dimension.

Keywords: 3D printing, Sustainability, Web Text Mining, Deep Learning

1 Introduction

In the influential Brundtland Report, sustainability is defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs"[Wo87]. Today, sustainability has become one of the most pressing, current global topics. In addition to being a discipline in its own right in the scientific community [KT06], many companies, governments, and institutions have begun to make sustainability a top priority. As stakeholders expect them to be socially responsible and environmentally conscious, many companies have launched initiatives such as reducing CO₂ emissions and eliminating waste. In this context, novel technologies are often promoted as a means to improve environmental sustainability. This also applies to 3D printing.

3D printing describes a group of manufacturing technologies converting digital data into physical products in a layer by layer process. The adoption of 3D printing entails considerable benefits compared to traditional manufacturing technologies, e.g. cost-efficient production of small volumes, development and fabrication of products with complex geometries, accelerating product development by rapid prototyping, and product customisation [At17; GRS14; WKP15]. Besides various economical advantages, this technology holds the potential to transform industries towards a more sustainable future, as it helps companies

¹ ISTAR.LAI, Julius-Hatry-Straße 1, 68163 Mannheim, Germany & University of Mannheim, Chair of SME Research and Entrepreneurship, L9 1-2, 68161 Mannheim, Germany robert.dehghan@istari.ai

² ISTAR.LAI, Julius-Hatry-Straße 1, 68163 Mannheim, Germany & University of Salzburg, Department of Geoinformatics - Z_GIS, Schillerstraße 30, 5020 Salzburg, Austria sebastian.schmidt@istari.ai

³ Technical University of Munich, School of Management julian.schwierzy@tum.de

to save resources and facilitate a circular economy [De17]. While manufacturing with conventional technologies, such as milling, produces a substantial amount of material as waste, unused material can often even be reused in a process with 3D printing [NGR20]. Another advantage of 3D printing is its capability to manufacture on-demand spare parts, components, and products [At17]. Such a production enables companies to produce their goods close to their consumers, thus reducing the transportation distance and the associated CO₂ emissions. Furthermore, its utilisation provides the possibility to develop and cost-efficiently fabricate products with a lightweight design, which e.g. in the case of transportation vehicles lowers fuel requirements and in turn emissions [GUV14]. However, the quantitative relationship between 3D printing application and sustainability has not yet been properly investigated in research. In this paper, we present a methodology that enables this analysis at the company level for the entire DACH region (Germany, Austria, Switzerland). In our study, we aim to answer the following research question: How is the relationship between 3D printing adoption and sustainability engagement characterised in the DACH region?

2 Methodology

Traditionally, economic research on topics such as innovation, technology diffusion or sustainability engagement is conducted on the basis of qualitative expert interviews, surveys or patent analyses. A central disadvantage of these methods is that they are very time-consuming and costly, and only small company samples can thus be considered. Web text mining from corporate websites is a very fruitful alternative, which has already been successfully demonstrated in several studies [Dö22; KL21; Sc22a]. For this purpose, texts can be extracted from corporate websites by web scraping and then classified by machine learning models. Since the current generation of natural language processing models can only operate on rather small window sizes, feeding entire websites is usually not feasible. Therefore, the splitting and classification of texts from websites into shorter sections is often resorted to. Through a subsequent aggregation of the results at website level, a classification of the entire website (e.g. technology adoption yes/no) and thus of the company under consideration can be made. The final classification can then e.g. correspond to the highest ranking label that was identified on a subpage.

We performed such an analysis for the entire active firm population in the DACH region drawing data from the ORBIS database (including the URL to the firm website and location of the respective firm). Since we expected the most important information to be included on the top-level websites, we ranked all subpage URLs by length and scraped only the shortest 25 [KA20]. In this process, we downloaded the respective HTML content and metadata for each subpage in June 2022.

For our study, we fine-tuned two Transformer models [Va17] for classification tasks: Firstly, a model that analysed the 3D printing intensity of a company, i.e. how often a company reports on 3D printing related activities on their website in relation to the website length. Secondly, we created a model that took a closer look at the commitment to sustainability,

i.e. how often aspects of environmental sustainability are communicated in the corporate website texts. Based on the model outputs, we calculated firm-level intensities as the ratio of relevant paragraphs, i.e. related to 3D printing or sustainability, to the website size. For the 3D printing model, we created an extensive, multi-lingual keyword list and carried out manual labelling to create a large training dataset (n=3,000). For this, we distinguished four categories (cf. Table 1): *manufacturer*, *service*, *retail*, and *information* [Sc22b]. Although this allowed our model to make even more precise statements, for the sake of clarity we decided to merge these categories into a single score for 3D printing adoption for this analysis. We employed a multi-lingual sentence-transformer MiniLM model which maps the input textual data to a dense vector space with 384 dimensions [RG19]. This enables our analysis for regions where different languages are spoken, although this is of little relevance to this analysis as most of the websites studied were in German or English. Due to the Sentence-BERT implementation, our model architecture already had a basic knowledge of natural language and, thus, only required small training data to be fine-tuned for our specific purposes [MR19; WRG21]. In addition to the labelled texts, we also added company and website meta-information (such as title and description) to this multi-dimensional input vector. Each paragraph that was identified in the keyword search was encoded with the domain adapted sentence transformer, which then output a value on a continuous scale between 0 and 1 using a logistic regression model in its final layer. This value represented the probability that a paragraph could be assigned to the respective category. Finally, the class with the highest probability was assigned to each text paragraph [RG19]. The same model was already used in [Sc22b].

Category	Definition
manufacturer	manufacturer of 3D printers and/or equipment (e.g. software, material)
service	use of 3D printing in own production, consulting, training, and/or personalised printing services
retail	seller of 3D printers, material, spare parts, etc.
information	disseminator of information about 3D printing technologies

Tab. 1: Labelling categories for 3D printing model.

Our model for sustainability engagement classification was employed similarly. For labelling purposes, we assigned five categories with the following ordinal ranking: *frontrunner*, *enabler*, *engaged*, *information* and *not engaged*. However, for this analysis, we grouped the first four categories into one. For the creation of our keyword list, we restricted ourselves to terms concerning the ecological dimension of sustainability. Consequently, paragraphs referring to e.g. 'working environments' were discarded as irrelevant in the labelling process, i.e. labelled as *not engaged*. We considered a sustainability intensity > 0 as indicative for a firm with sustainability engagement. The same method was also employed in [Sc22a].

3 Results

Since we performed our analysis on firm level, this approach allowed us to make quite precise statements about characteristics of 3D printing companies, such as age or location. In total, we looked at a population of 1,782,172 firms in Germany, Austria, and Switzerland. 207,368 (11.6%) of these companies had a sustainability intensity > 0 , thus, being classified as 'sustainability-engaged companies'. In the overall population, we identified 8,170 3D printing companies (6,473 in Germany, 713 in Austria, 984 in Switzerland). This corresponded to 0.46% of the firm population (0.48% in Germany, 0.40% in Austria, 0.38% in Switzerland). 2,500 (30.6%) of the 3D printing companies were sustainability-engaged, which was a much higher number than the 11.6% in the overall firm population. The distribution of sustainability scores was highly skewed to the right for all sustainability-engaged companies (median: 0.275, max: 6.935), as well as for all sustainability-engaged 3D printing companies (median: 0.383, max: 4.154).

The vast majority of 3D printing firms was classified in the *service* category (71.4%, cf. Table 2). Only a small fraction of firms acted as *manufacturers*, while there were barely any in the *retail* category. The distribution among the categories did not differ significantly between all 3D printing companies and sustainable 3D printing companies. However, the share of *manufacturers* was slightly higher for the sustainable 3D printing firms.

Category	All	Sustainable	Share
manufacturer	643 (7.87%)	232 (9.28%)	36.1%
service	5,831 (71.37%)	1,664 (66.56%)	28.5%
retail	35 (0.43%)	9 (0.36%)	25.7%
information	1,661 (20.33%)	595 (23.80%)	35.8%

Tab. 2: 3D printing firms per category.

An important variable in economic considerations is the firm size. Therefore, we split our firm population into different categories based on the number of employees. For this, we used the staff headcount category of the SME definition of the European Commission [Th03]: Small (0-9 employees), medium (10-49 employees), large (50-499 employees), and very large firms (> 500 employees). Most of the 3D printing companies were actually small firms, while only 247 were classified as very large. However, for these very large firms 64.8% were indeed sustainability-engaged companies. Only 23.8% of the small companies were sustainability-engaged. However, in the overall firm population, merely 9.4% of small companies and 42.5% of very large firms were sustainability-engaged.

Firm Size	All	Sustainable	Share
small	3785 (0.101)	902 (0.000)	23.8%
medium	2161 (0.095)	700 (0.000)	32.4%
large	1180 (0.087)	573 (0.150)	48.6%
very large	247 (0.085)	160 (0.568)	64.8%

Tab. 3: 3D printing firm count per firm size category. The median intensity is represented in brackets.

Most of the identified 3D printing firms were rather young, as 61% were founded after the turn of the millennium (cf. Table 4). Consequently, the percentage of firms founded in the 1990s or earlier was higher for the subset. For some companies, there was no information on their founding date in the ORBIS database. Overall, an earlier founding year was associated with a higher relative share of sustainability-engaged companies and higher intensity. This trend, however, was in line with the population of all companies in the DACH region. Therefore, we could not conclude a link between sustainability and age of 3D printing companies.

Decade	All	Sustainable	Share
2020s	66 (0.8%)	15 (0.6%)	22.7%
2010s	2,658 (33.5%)	730 (29.5%)	27.5%
2000s	2,121 (26.7%)	617 (24.9%)	29.1%
1990s	1,550 (19.5%)	509 (20.6%)	32.8%
1980s	652 (8.2%)	230 (9.3%)	35.3%
before 1980	896 (11.3%)	399 (16.1%)	44.5%
Total	7,943 (100%)	2,473 (100%)	

Tab. 4: 3D printing firm count per decade of founding.

In a further step, we looked at the spatial distribution of sustainable 3D printing companies. The 10 NUTS3 regions with the highest number of sustainable 3D printing firms were Berlin (113), Munich (city) (90), Hamburg (81), Zurich (67), Bern (57), Vienna (55), Dusseldorf (50), Stuttgart (48), Munich (district) (42) and Frankfurt (36). This mostly corresponded with the NUTS3 regions with the highest overall number of 3D printing companies and economic centres in the DACH region. Figure 1 aggregates the number of firms on NUTS3 level and normalises them with the overall number of 3D printing companies in the respective region. Only NUTS3 areas with more than 10 3D printing firms were considered for this analysis. Accordingly, the map mainly reflects established economic and demographic centres of the research area.

We also calculated the Moran's I for the ratio, employing a Queen contiguity approach for the underlying spatial weights matrix. The resulting value of 0.00029 was not statistically

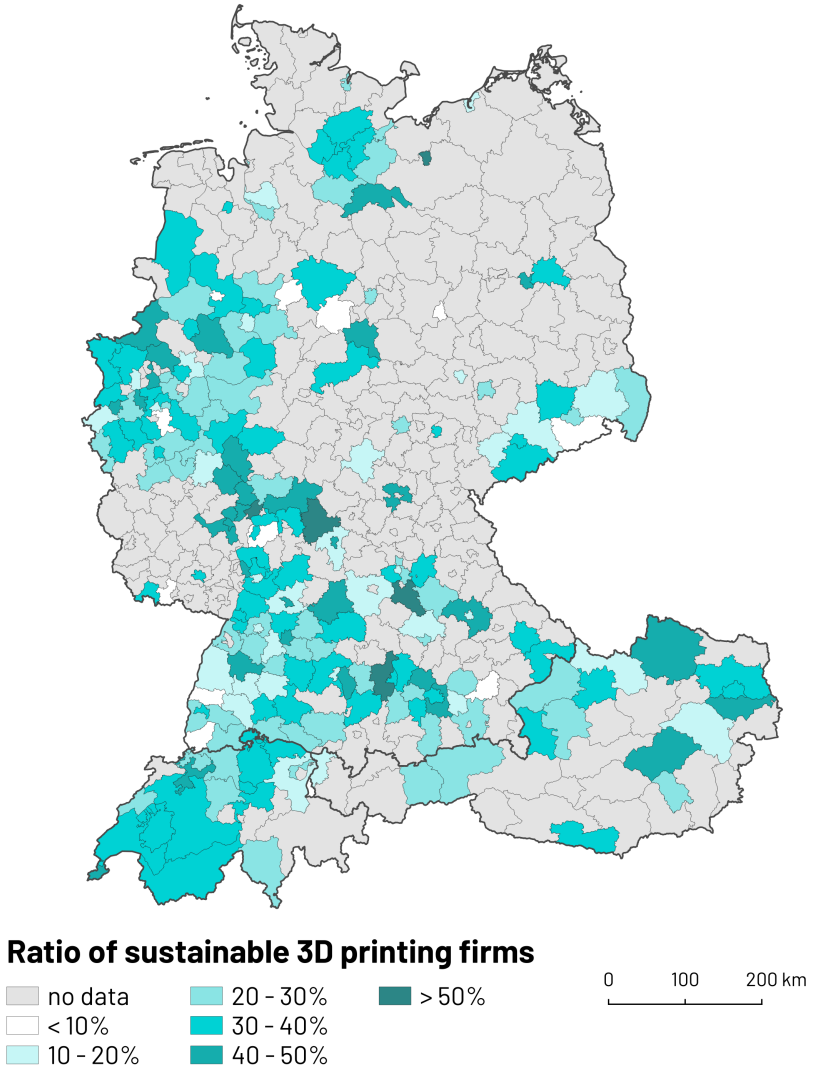


Abb. 1: Ratio of sustainable 3D printing firms on all 3D printing firms.

significant ($p=0.458$), indicating no spatial autocorrelation, i.e. the absence of spatial clustering. However, this might have been influenced by the low number of the remaining regions.

4 Discussion & conclusion

Our study was able to make first precise statements about the exact location and company characteristics (e.g. age and size) of each 3D printing firm. In total, we looked at a population of 1,782,172 firms in Germany, Austria and Switzerland of which we identified 0.46% as 3D printing companies. According to our results, 3D printing companies were overall almost three times more inclined to present themselves as sustainability-engaged in comparison to the overall firm location in the DACH region. We found that larger and older companies were more inclined to be sustainability-engaged, which was consistent with the overall firm population. Furthermore, we observed that companies classified as *manufacturer* and *information* contained a higher share of sustainability-engaged companies than *service*. This may indicate that manufacturers and information providers understand the technological possibilities for sustainable activities. Service firms, on the other hand, do not seem to fully exploit their ecological potential, as their share of sustainability-engaged companies was considerably smaller.

We did not find any indications for a spatial clustering of sustainable 3D printing companies. However, since all of our analyses were solely descriptive, there might be statistical relationships we were not able to uncover.

Therefore, a more extensive statistical analysis should be performed. This should also account for control variables, such as the firm sector or location factors (e.g. infrastructural quality). Our finding that larger and older companies are more likely to engage in sustainability could also be related to a potential bias in our indicator. Nevertheless, such a bias would not impede further economic analyses since it is possible to control for company age. As described above, our model would also make it possible to differentiate the companies into more precise categories (e.g. sustainability *frontrunner* or 3D printing *manufacturer*). This information could therefore also be included in further statistical evaluations.

It is important to note that our methodology can only capture the self-representation of the companies. Against the backdrop of greenwashing, it is therefore quite conceivable that many companies use formulations in the context of sustainability merely as 'empty words'. The aforementioned statistical analysis or linking with other data (e.g. ESG reports, surveys, patents) could provide deeper insights into the reliability of the website text content.

Nevertheless, this paper presents web text mining based on corporate websites as a promising alternative or complement to traditional approaches to economic, ecological and geographical research. Only through the text-based classification of the companies into different categories, we were able to make statements on the topics of technology diffusion and sustainability engagement as well as their interplay. We deliberately chose

3D printing for our study because it is a technology that has a particularly high potential to make production more sustainable. However, the data sources and basic methodology could also be applied to other technologies or topics in a business context. Here, it is important to emphasise that the approach is applicable to different scales, both for a rough quantification at the region level and a more detailed description at company level. A number of environment-related questions can thus be addressed, such as an international comparison of the discourse on sustainability or an assessment of the success of regional policy measures.

Acknowledgement

We would like to thank Jan Kinne, David Lenz and Jannik Reißfelder (ISTARI.AI) for their development efforts and outstanding support.

All authors contributed equally to the study.

Literatur

- [At17] Attaran, M.: The rise of 3-D printing: The advantages of additive manufacturing over traditional manufacturing. *Business horizons* 60/5, S. 677–688, 2017.
- [De17] Despeisse, M.; Baumers, M.; Brown, P.; Charnley, F.; Ford, S. J.; Garmulewicz, A.; Knowles, S.; Minshall, T.; Mortara, L.; Reed-Tsochas, F. et al.: Unlocking value for a circular economy through 3D printing: A research agenda. *Technological Forecasting and Social Change* 115/, S. 75–84, 2017.
- [Dö22] Dörr, J. O.; Kinne, J.; Lenz, D.; Licht, G.; Winker, P.: An integrated data framework for policy guidance during the coronavirus pandemic: Towards real-time decision support for economic policymakers. *PLOS ONE* 17/2, 2022.
- [GRS14] Gibson, I.; Rosen, D.; Stucker, B.: *Additive Manufacturing Technologies: 3D Printing, Rapid Prototyping, and Direct Digital Manufacturing*. Springer, Berlin, 2014.
- [GUV14] Gebler, M.; Uiterkamp, A. J. S.; Visser, C.: A global sustainability perspective on 3D printing technologies. *Energy policy* 74/, S. 158–167, 2014.
- [KA20] Kinne, J.; Axenbeck, J.: Web mining for innovation ecosystem mapping: A framework and a large-scale pilot study. *Scientometrics* 125/, S. 2011–2041, 2020.
- [KL21] Kinne, J.; Lenz, D.: Predicting innovative firms using web mining and deep learning. *PLOS ONE* 16/4, 2021.
- [KT06] Komiyama, H.; Takeuchi, K.: Sustainability science: building a new discipline. *Sustainability science* 1/, S. 1–6, 2006.

- [MR19] Malte, A.; Ratadiya, P.: Evolution of transfer learning in natural language processing, 2019, arXiv: 1910.07370.
- [NGR20] Nadagouda, M. N.; Ginn, M.; Rastogi, V.: A review of 3D printing techniques for environmental applications. *Current opinion in chemical engineering* 28/, S. 173–178, 2020.
- [RG19] Reimers, N.; Gurevych, I.: Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Nov. 2019, URL: <http://arxiv.org/abs/1908.10084>.
- [Sc22a] Schmidt, S.; Kinne, J.; Lautenbach, S.; Blaschke, T.; Lenz, D.; Resch, B.: Greenwashing in the US metal industry? A novel approach combining SO₂ concentrations from satellite data, a plant-level firm database and web text mining. *Science of The Total Environment* 835/, S. 155512, 2022.
- [Sc22b] Schwierzy, J.; Dehghan, R.; Schmidt, S.; Rodepeter, E.; Stoemmer, A.; Uctum, K.; Kinne, J.; Lenz, D.; Hottenrott, H.: Technology Mapping Using WebAI: The Case of 3D Printing, 2022, arXiv: 2201.01125.
- [Th03] The Commission of the European Communities: Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. 2003.
- [Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I.: Attention Is All You Need, 2017, arXiv: 1706.03762.
- [WKP15] Weller, C.; Kleer, R.; Piller, F. T.: Economic implications of 3D printing: Market structure models in light of additive manufacturing revisited. *International Journal of Production Economics* 164/C, S. 43–56, 2015.
- [Wo87] World Commission on Environment and Development: *Our Common Future* Oxford. 1987.
- [WRG21] Wang, K.; Reimers, N.; Gurevych, I.: TSDAE: Using Transformer-based Sequential Denoising Auto-Encoder for Unsupervised Sentence Embedding Learning./, 2021, eprint: 2104.06979.

Nutzung offener Standards für die Integration von Pegeldaten und KI-basierte Prognose von Pegel- und Abflussdaten zur Verbesserung der Frühwarnung bei Sturzfluten

Jonathan Vogl¹, Désirée Hilbring² und Divas Karimanzira³

Abstract: Im Kontext von Frühwarnsystemen, deren Bedeutung in Zeiten mit steigender Anzahl an Extremwetterereignissen wächst, spielt die Verfügbarkeit von Messdaten in Kombination mit einem schnellen Datenaustausch zwischen KI-Algorithmen zur Berechnung und Plattformen zur Visualisierung von Prognoseergebnissen eine entscheidende Rolle. Im Rahmen des Eigenforschungsprojekts PrognoSF des Fraunhofer IOSB werden Niederschlagsprognosedaten des DWD mit lokalen Sensordaten aus Smart City Plattformen kombiniert und Pegelprognosen auf Basis von KI durchgeführt, die später auch Frühwarnungen, gerade bei Gefährdung von kritischer Infrastruktur, gewährleisten sollen. Für die flexible Integration von Daten und Modulen kommen offene Standards zum Einsatz. Für die Pegelprognose wird ein neuronales maschinelles Übersetzungsmodell trainiert, das aus tiefen rekurrenten Netzen besteht.

Keywords: Sensor Things API, Pegelprognose, Abflussprognose, Sturzflutfrühwarnung

1 Motivation

Der Klimawandel und die wachsende Urbanisierung verbunden mit der voranschreitenden Versiegelung von Flächen führen zum verstärkten Auftreten von Überschwemmungen ausgelöst durch Starkregenereignisse. Frühwarnsysteme sind seit langem Thema der Forschungslandschaft, trotzdem ist die Trefferquote bei den Prognosen vor allem bei lokalen Ereignissen immer noch nicht ausreichend. Ziel des Eigenforschungsprojektes PrognoSF des Fraunhofer IOSB ist es diesem Problem zu begegnen, indem Niederschlagsprognosedaten des DWD mit lokalen Daten (Niederschlagsdaten, Pegeldaten) aus Smart City Plattformen kombiniert werden, um die Prognosen zu verbessern und dadurch die Vorwarnzeit zu verlängern. Dafür sollen verschiedene, mit offenen Standards versehene KI-Prognosemodule, die für die Verbesserung der Frühwarnung eingesetzt werden können, entwickelt werden.

¹ Fraunhofer IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland jonathan.vogl@iosb.fraunhofer.de

² Fraunhofer IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland desiree.hilbring@iosb.fraunhofer.de

³ Fraunhofer IOSB-AST, Am Vogelherd 90, 98693 Ilemnau, Deutschland divas.karimanzira@iosb.fraunhofer.de

2 Stand der Technik

Im Bereich der Vorhersage von Hochwassergefahren und der Risikokartierung gibt es mehrere Entwicklungen und Ansätze. Die Modelle können hauptsächlich in empirische, physikalische und konzeptionelle Modelle unterteilt werden und unterscheiden sich entsprechend in ihrer Geschwindigkeit, ihren Ressourcen und ihrem Datenbedarf. Vollständige hydrodynamische 1D- und 2D-Modelle sind für die Simulation von Überschwemmungen gut etabliert und in einer Reihe kommerzieller Software verfügbar, darunter SOBEK, XPSWNN2D, MIKE FLOOD und InfoWorks ICM. Die meisten verfügbaren Modelle basieren auf der Lösung einer Reihe von partiellen Differentialgleichungen, die einen enormen Rechenaufwand erfordern.

Es gibt mehrere Entwicklungen und Ansätze, die Deep-Learning-Modelle verwenden, aber ihre Genauigkeit ist oft auf eine kurze Vorlaufzeit beschränkt und vielen Modellen mangelt es an Generalisierung, d. h. für jeden Standort werden separate Modelle trainiert z.B. [Kr19] und [XYD20]. In diesem Beitrag wird ein Ansatz für ein tiefes rekurrentes neuronales Netz für eine standortübergreifende, mehrstufige Vorhersage des Abflusses oder des Pegels vorgeschlagen.

Das Projekt PrognoSf arbeitet auch an diesen Themenbereichen, aber auch an der Fragestellung, wie Basisdaten und KI-Methoden für die Frühwarnung von Sturzfluten in ein Frühwarnsystem integriert werden können. Beispielhaft wird dafür im Folgenden die Integration einer KI-Pegel- und Abflussprognose auf Basis von sensorbasierten Pegelmessungen diskutiert.

Im Bereich der Algorithmenintegration setzen wir auf den Tasking Core der SensorThings API. Die Anbindung von (KI-)Prognosemodulen, die auf dem SensorThings API Standard des Open Geospatial Consortium basieren, ist beispielsweise im Forschungsprojekt NiMo 4.0 (Nitrat Monitoring), das Prognosen für den erwarteten Nitratgehalt des Grundwassers berechnet, geschehen [Em22][He21]. Weitere ähnliche Ansätze sind uns nicht bekannt.

Bezüglich der Verfügbarkeit von Pegeldata existiert das Portal Pegelonline der Wasserstraßen- und Schifffahrtsverwaltung des Bundes (WSV). Dies stellt verschiedene Webservice bereit, wie eine eigene REST API, eine SOAP API sowie eine auf dem OGC Sensor Observation Service Standard basierende Schnittstelle.

3 Bereitstellung von Pegeldata via SensorThings API

Eine wichtige dynamische Datengrundlage für die Vorhersage von Sturzfluten bilden die Echtzeitdaten verschiedener Pegelmessstellen in Flüssen. Damit diese Daten flexibel verschiedenen KI-basierten Prognosealgorithmen zur Verfügung gestellt werden können, setzen wir den offenen Standard SensorThings API des Open Geospatial Consortiums ein, einen leichtgewichtigen Standard für das Datenmanagement von zeitreihenbasierten Daten aus dem IoT Kontext. Während der erste Teil der SensorThings API („Sensing“) Messdaten verwaltet, kann mit dem zweiten Teil („Tasking Core“) die flexible Anbindung von KI-Algorithmen erfolgen, und zwar unabhängig von der notwendigen Datengrundlage

des Algorithmus [LHK16] [LK19]. Im Projekt wird die Open-Source Implementierung FROST[®]-Server der SensorThings API eingesetzt. Zur Erfassung von Messdaten wurden eigene Sensoren in der Pilotregion Steinheim (NRW) aufgestellt. Diese senden regelmäßig aktuelle Pegelstände sowie weitere Messdaten (Temperatur, Luftdruck und -feuchtigkeit). Über eine MQTT-Schnittstelle werden diese Zeitreihendaten gemäß des Sensing-Standards kontinuierlich im FROST[®]-Server ergänzt. Jede Messstelle (Sensor) wird als *thing* modelliert, dazu hat ein Sensor für jede Messgröße einen *datastream*, der die jeweilige Messreihe darstellt. Einzelne Messwerte sind dabei als *observations* gespeichert und neue Werte lassen sich auch durch Anlegen neuer *observation*-Entitäten, die dem entsprechenden *datastream* zugeordnet ist, hinzufügen.

Auf die Messdaten kann jederzeit mittels REST-Abfragen über die http-Schnittstelle zugegriffen werden. Dadurch stehen die aktuellsten Daten sowohl Plattformen zur Visualisierung bereit (siehe Abbildung 1) als auch als Input für KI-Algorithmen. Gleichzeitig können jene Algorithmen ihre Ergebnisse wiederum als *datastreams* speichern und diese so für andere Komponenten verfügbar machen, etwa Plattformen zur Visualisierung oder weiteren Algorithmen als Eingabe. Um Mess- von Prognosedaten klar zu trennen, wird eine separate FROST-Instanz zum Speichern von Resultaten verwendet.

4 Ausführung von Tasks via SensorThings API

Ein weiteres Ziel ist die Integration der Algorithmen sowie die Möglichkeit, Berechnungen von einer Plattform aus anzustoßen. Dabei soll einerseits automatisch in regelmäßigen Zeitabschnitten eine Pegelprognose für die nächsten Stunden erfolgen, deren Ergebnisse maschinenlesbar für externe Systeme als auch für Endanwender in einer Plattform sichtbar sind. Zusätzlich soll auch die Möglichkeit bestehen, die Prognose manuell, mit Auswahl weniger, leicht verständlicher Parameter (bspw. Dauer der Vorhersage) für Analysen zur Ausführung zu bringen. Hierfür wird der zweite Teil der SensorThings API verwendet. Er ist über die Entität *thing* mit dem ersten Teil verknüpft und enthält die Entitäten



Abb. 1: Die von den eigenen Sensoren gemessenen Zeitreihen werden über den FROST-Server abgefragt und im Grafana Dashboard visualisiert

taskingCapability und *task*. Damit kann ein Algorithmus als *taskingCapability* angelegt werden. Eine konkrete Ausführung von diesem wird durch Erzeugen eines dieser *taskingCapability* zugeordneten *tasks* mit Spezifikation von Konfigurationsparametern modelliert. Das Festlegen von Konfigurationsparametern und Anlegen des *tasks* kann auch über eine Plattform anwendergerecht integriert sein.

Die Ausführung der Algorithmen wird dann durch die Komponente PERMA gehandhabt. Dazu müssen die Algorithmen z.B. als Docker-Container zur Verfügung stehen. Da PERMA auf Basis von Java läuft, bietet die in Python geschriebene Komponente PyPerma eine einfache Möglichkeit zur Anbindung insbesondere von KI-Algorithmen, die mit Python implementiert sind. Wird im FROST-Server ein entsprechender *task* einer *taskingCapability* erzeugt, sorgt PERMA dafür, dass der entsprechende (KI-)Algorithmus auf einer Containerplattform wie Kubernetes ausgeführt wird. Ergebnisse können dann wie in Abschnitt 3 beschrieben als separate *datastreams* gespeichert werden oder über andere Entitäten, beispielsweise als *properties* von *things*, wenn die Ergebnisse in Bezug zu einer konkreten Messstelle stehen (etwa deren Wichtigkeit für die Prognose).

Bereits trainierte Modelle zu speichern, abrufbar und damit Ergebnisse reproduzierbar zu machen, ist über einen Cloudspeicher möglich.

5 Pegel- und Abflussprognose

Ein Beispiel für einen mittels der Sensor Things API integrierten Algorithmus ist eine KI-basierte Pegelprognose, die den Pegelstand und den Abfluss abhängig von den Ergebnissen der Niederschlagsprognose und weiterer Eingangsdaten prognostiziert. In einem zugehörigen *task* werden dem integrierten Algorithmus die folgenden Parameter übermittelt:

- Pegelprognose oder Abfluss: was soll prognostiziert werden?
- Training und Messstellen/Gebiete: soll ein Modell für die angegebenen Messstellen/Gebiete trainiert werden oder auf Basis eines bereits vorhandenen Modells eine Prognose durchgeführt werden?
- Einsatzgebiet: die Pegelmessstellen, für welche die Prognose im Einzelnen berechnet werden soll.
- Prognosezeitraum: damit können je nach Anwendungsfall beliebige, verschiedene Vorwarnzeiten (z.B 1h, 2h, 6h) unterschiedlicher Prognosequalität berechnet werden.

Man beachte, dass die Messstellen zum Training und zur Prognose nicht identisch sind. Insbesondere müssen die zu prognostizierenden Messstellen keine Teilmenge derer des Trainings sein. Die Qualität der Prognose hängt dann von der Übertragbarkeit des Modells auf andere Gebiete ab.

5.1 Methode

Die Architektur des Modells ist in der Abbildung 2 dargestellt. Es handelt sich um ein neuronales maschinelles Übersetzungsmodell, das auf tiefen rekurrenten neuronalen Netzen mit übereinander liegenden Schichten für alle Zeitschritte basiert, anstatt einen einzelnen Vektor als Vermittler zwischen Eingabe und Ausgabe zu verwenden. In [LBZ17] wurde gezeigt, dass das maschinelle Übersetzungsmodell von Google mit dieser Struktur kein Engpassproblem hat und bessere Ergebnisse als Encoder-Decoder-Modelle bei der Sprach­übersetzung zeigt. Als Knoten können verschiedene tiefe rekurrente neuronale Netzwerke verwendet werden, z. B. Long Short Term Memory-Netzwerke (LSTM), Gated Recurrent Units (GRU) etc. Neben den auf historischen Zeitreihen basierenden Daten (Niederschlag, Evapotranspiration usw.) und prädiktiven Eingangsdaten (Niederschlag) berücksichtigt das Modell Merkmale auf der Ebene des Wassereinzugsgebiets, die Konzentrationszeit auf der Grundlage des längsten Pfads, das Gefälle und den Anteil der Bodentypen. Alle diese physikalischen Merkmale werden im Laufe der Zeit als konstant angenommen, enthalten aber Informationen, die eine Unterscheidung zwischen verschiedenen Teileinzugsgebieten ermöglichen. Im Idealfall soll das Netz die Verarbeitung der dynamischen Eingaben von einer Reihe statischer Einzugsgebietsmerkmale abhängig machen. Das heißt, das Netz soll eine Abbildung von meteorologischen Zeitreihen auf den Abfluss erlernen, die selbst von einer Reihe statischer Einzugsgebietsmerkmale abhängt, die im Prinzip überall in unserem Modellierungsgebiet gemessen werden könnten. Eine Möglichkeit, dies zu tun, wäre, die statischen Merkmale als zusätzliche Eingaben bei jedem Zeitschritt hinzuzufügen. Das heißt, wir könnten die Vektoren $x[t]$ bei jedem Zeitschritt einfach um eine Reihe von Einzugsgebietsmerkmalen ergänzen, die sich im Laufe der Zeit nicht ändern. Dieser

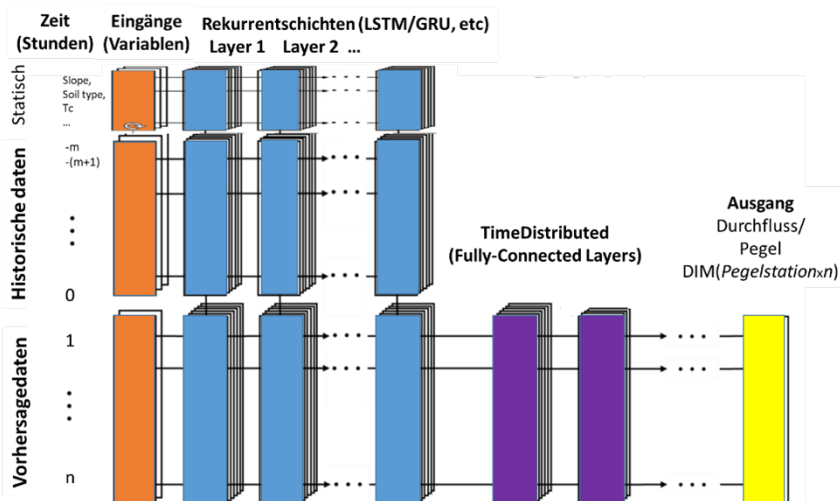


Abb. 2: Architektur für die mehrschrittige Pegel-/Abflussprognose von mehreren Pegelstationen

Ansatz erlaubt es uns jedoch nicht, direkt zu überprüfen, was das rekurrente neuronale Netz aus diesen statischen Einzugsgebietsmerkmalen wirklich lernt. Daher wird das rekurrente neuronale Netz leicht verändert, indem eine eingebettete Schicht hinzugefügt wird, wie in [Kr19] beschrieben. Bei einem LSTM als rekurrentes Netzwerkmodell wird beispielsweise die Eingabe wie in Gleichung 1 und 2 geändert, während der Rest des LSTM unverändert bleibt.

$$i = \sigma(W_i x_s + b_i), \quad (1)$$

$$f[t] = \sigma(W_f x_d[t] + U_f h[t-1] + b_f), \quad (2)$$

$$g[t] = \tanh(W_g x_d[t] + U_g h[t-1] + b_g), \quad (3)$$

$$o[t] = \sigma(W_o x_d[t] + U_o h[t-1] + b_o). \quad (4)$$

wobei i ein Eingangsgatter ist, das nun über die Zeit konstant ist. x_s sind die statischen Eingaben (z. B. Einzugsgebietsattribute) und $x_d[t]$ sind die dynamischen Eingaben (z. B. meteorologische Einflüsse) bei Zeitschritt (t).

Gleichung 1 und 2 ermöglichen es dem Netz, die statischen Eingänge x_s und die dynamischen Eingänge $x_d[t]$ innerhalb der Architektur explizit getrennt zu verarbeiten und ihnen spezielle Aufgaben zuzuweisen. Die statischen Merkmale steuern über das Eingangsgatter (i), welche Teile des LSTM für jedes einzelne Einzugsgebiet aktiviert werden, während die dynamischen und rekurrenten Eingänge steuern, welche Informationen in den Speicher geschrieben werden ($g[t]$), welche vergessen werden ($f[t]$) und welche der gespeicherten Informationen zum aktuellen Zeitschritt t ausgegeben werden ($o[t]$).

Die Einbettungsschicht ermöglicht einen nicht-naiven Informationsaustausch zwischen den Einzugsgebieten und erlaubt komplexe Wechselwirkungen zwischen den Einzugsgebietsmerkmalen, und - was besonders wichtig ist - sie macht es möglich, dass diese Wechselwirkungen direkt von den Niederschlag-Abfluss-Daten aller für das Training verwendeten Einzugsgebiete informiert werden.

Um den Abfluss besser zu modellieren, wird eine halbverteilte Struktur verwendet, bei der jedes Einzugsgebiet als Teileinzugsgebiet betrachtet wird und alle Zeitreihendaten und physikalischen Merkmale nur auf dem Teileinzugsgebiet basieren. In den meisten früheren Studien (z. B. [Kr19]) werden verteilte Modelle (mit halbverteilter Modellstruktur) einzeln trainiert, und zwar sequenziell vom Oberlauf zum Unterlauf. Dies ist jedoch bei einem verallgemeinerten Modell nicht möglich, da alle Wassereinzugsgebiete gleichzeitig in einem Modell trainiert werden müssen. Daher wird die Beziehung zwischen Oberlauf und Unterlauf konstruiert, die auch als Input für das Training dient.

5.2 Experimentelle Analyse

Für die Studien wurden Daten aus einer Modellregion in Deutschland verwendet, um die Performanz der Methoden zur Abfluss-/Pegelvorschau zu demonstrieren. Die Daten umfassten 15-Minuten-Beobachtungen von Niederschlag, Evapotranspiration und Abfluss von verschiedenen Stationen in dem Gebiet. Darüber hinaus waren auch einige konstante

charakteristische Merkmale der Standorte verfügbar. Dazu gehörten Informationen über das Einzugsgebiet, die Konzentrationszeit, das aus einem DEM berechnete Gefälle und Bodeninformationen, die nach verschiedenen Typen (Lehm, undurchlässige Oberfläche usw.) kodiert waren. Da die Daten in unterschiedliche Größenordnungen fallen, wurden sie vor der Trainingsphase mit Hilfe des Min-Max-Operators normalisiert. Die Beobachtungsdaten wurden in eine Trainings-, eine Validierungs- und eine Teststichprobe unterteilt. Die Beobachtungen von Januar 2003 bis Juli 2011 wurden für das Training verwendet, die Daten von Juli 2011 bis Juni 2016 für die Validierung und die restlichen Daten von Juli 2016 bis Februar 2022 für die Tests. Zusätzlich wurden einige frühe Aufzeichnungen von Extremereignissen aus den Jahren 1804 und 1910 zum Trainingssatz hinzugefügt. Die Modelle verwenden die Daten der letzten 72 Stunden zu Niederschlag und Abfluss, um den Abfluss der nächsten Stunden vorherzusagen. Dazu gehen die Niederschlags- und Temperaturdaten (für die Evapotranspiration) des Prognosezeitraums mit ein, die später bei einer Echtzeitanbindung durch Vorhersagen ersetzt werden müssen. Als Ziel haben wir uns gesetzt, mindestens 6 Stunden im Voraus zu prognostizieren, da in dieser Zeit nach starken Regenfällen Sturzfluten auftreten können. In mehreren Experimenten und auch in [GI22] wurde festgestellt, dass das Wissen über die letzten drei Tage eines Einzugsgebiets für zukünftige Vorhersagen entscheidend ist.

5.3 Ergebnisse und Diskussionen

Zur Bewertung der Ähnlichkeit zwischen dem beobachteten und dem simulierten Abfluss der Vorhersagemodelle für die verschiedenen Zeitschritte werden zwei Leistungskennzahlen verwendet. Die erste ist die traditionelle, in hydrologischen Modellen übliche Nash-Sutcliffe Efficiency (NSE), da sie die Modellleistung in eine interpretierbare Skala normalisiert, wie in (5) ausgedrückt.

$$\text{NSE} = 1 - \frac{\sum_{t=1}^T (\hat{Q}(t) - Q(t))^2}{\sum_{t=1}^T Q(t) - \bar{Q}(t)^2}, \quad (5)$$

wobei T die Gesamtzahl der Zeitschritte, $\hat{Q}(t)$ der simulierte Abfluss zum Zeitpunkt t , $Q(t)$ der beobachtete Abfluss zum Zeitpunkt t und $\bar{Q}(t)$ der mittlere beobachtete Abfluss ist. $\text{NSE} = 1$ bedeutet eine perfekte Übereinstimmung zwischen Simulationen und Beobachtungen; $\text{NSE} = 0$ bedeutet, dass die Modellsimulationen die gleiche Erklärungskraft haben wie der Mittelwert der Beobachtungen.

Die zweite Metrik ist die Kling-Gupta Efficiency (KGE) (6). Diese Metrik behebt mehrere Unzulänglichkeiten der NSE und wird zunehmend für die Modellkalibrierung und -bewertung von hydrologischen Modellen verwendet.

$$\text{KGE} = 1 - \sqrt{(r - 1)^2 + \left(\frac{\hat{\sigma}}{\sigma} - 1\right)^2 + \left(\frac{\hat{\mu}}{\mu} - 1\right)^2}, \quad (6)$$

wobei r die lineare Korrelation zwischen Beobachtungen und Simulationen ist, $\alpha = \frac{\hat{\sigma}}{\sigma}$ ein Maß für den Fehler der Flussvariabilität und $\beta = \frac{\hat{\mu}}{\mu}$ ein Verzerrungsterm ist. σ ist die Standardabweichung in den Beobachtungen, $\hat{\sigma}$ die Standardabweichung in den Simulationen, $\hat{\mu}$ ist der Simulationsmittelwert und μ der Beobachtungsmittelwert. Wie NSE bedeutet $\text{KGE} = 1$ eine perfekte Übereinstimmung zwischen Simulationen und Beobachtungen.

Zur Veranschaulichung sind Ergebnisse der Abflussprognose für zwei verschiedene Pegel (Stationen 505 und 108) des Ahrtals aus historischen Daten des Zeitraum der Ahrtalkatastrophe mit Prognosen für 6 Stunden im Voraus berechnet worden. Die Ergebnisse sind in Abbildung 3 dargestellt. Es ist zu erkennen, dass die Methode ein sehr gutes Ergebnis mit einem NSE für 6 Stunden im Voraus von 0,965 und einem KGE von 0,947 für Station 505 und einem NSE für 6 Stunden im Voraus von 0,772 und einem KGE von 0,723 für Station 108 zeigt. Besonders bei dem Extremereignis von 2021 zeigt das Modell deutlich bessere Ergebnisse mit einem Spitzenwert von $800\text{m}^3/\text{s}$ (gemessen $1000\text{m}^3/\text{s}$) im Vergleich zu früheren Studien in [GI22] mit nur $400\text{m}^3/\text{s}$.

Aktuell arbeiten wir im Projekt PrognosSF daran, das trainierte Modell nun für die Echtzeitanwendung nutzen zu können. Dazu müssen die lokalen Pegelsensoren angebunden werden. Diese wurden in den FROST[®]-Server integriert und sind nun für die Pegelprognose verfügbar, die mittels PyPerma integriert wird. Zusätzlich benötigen wir Prognosen (6 Stunden) für Niederschlag und Evapotranspiration. Dazu verwenden wir zunächst Daten des DWD und entwickeln aber auch ein eigenes Niederschlagsprognosemodul, das mittels der SensorThingsAPI leicht an die Pegelprognose angebunden werden kann.

6 Plug and Play Vision

Diese Infrastruktur vernetzt die einzelnen Schritte einer Sturzflutprognose über eine einfache, einheitliche Schnittstelle. Der modulare Aufbau ermöglicht Anpassungen einzelner Komponenten weitestgehend unabhängig der anderen. Dazu wird die Verfügbarkeit von Daten bereits bei der Entwicklung von KI-Algorithmen gewährleistet. Für die Entwicklung eines Frühwarnsystems für Sturzfluten können mit der Plug and Play Methode neben der Pegel- und Sturzflutprognose weitere Module in das System integriert werden, um einerseits Eingangsdaten zu verbessern und andererseits die Ergebnisse verschiedenen Endanwenderplattformen zur Verfügung zu stellen. Exemplarisch findet sich modulare Aufbau zu PrognosSF in Abbildung 4. Die Messdaten können dort entweder zunächst vorverarbeitet werden (bspw. Niederschlagsprognose für Gebiet) oder direkt in die Algorithmen eingehen wie beispielsweise bei der Pegel- und Abflussprognose. Dessen Ergebnisse stehen dann über die offenen Schnittstellen für die Nutzung in verschiedenen Endanwenderplattformen gleichermaßen zur Verfügung. Der geringe Aufwand bezüglich Austauschbarkeit bzw. des Hinzufügens einzelner Komponenten ist ersichtlich.

Dieses Vorgehen kann in ähnlicher Weise für andere Anwendungsfälle übertragen und angepasst werden. Insbesondere erlaubt es, verschiedene Anwendungsfälle parallel zu integrieren, ohne dass diese alle zu Beginn spezifiziert oder gar bekannt sein müssen. Diese Flexibilität kann im Zuge von künftigen Arten von Extremwetterereignissen, die

zum Entwicklungszeitpunkt noch nicht relevant erschienen oder überhaupt bewusst waren, besonders bedeutsam sein.

Literaturverzeichnis

- [Em22] Emde, K.; Budde, M.; Fischer, T.; Martin, T.; Hilbring, D.: Interaktive Steuerung der Ausführung von KI-Algorithmen in Umweltinformationssystemen über OGC SensorThings. In (Demmler, D.; Krupka, D.; Federrath, H., Hrsg.): INFORMATIK 2022. Gesellschaft für Informatik, Bonn, S. 1517–1527, 2022.
- [Gl22] Glaser, R.; Kahle, M.; Kempf, M.; Martin, B.: Classifying the 2021 “Ahrtal” flood event using hermeneutic interpretation, natural language processing, and instrumental data analyses. Environmental Research Communications, 4, 2022.
- [He21] Hertweck, P.; van der Schaaf, H.; Hilbring, D.; Weis, J.; Liesch, T.; Budde, M.: Integration von KI-Algorithmen in Umweltinformationssysteme mittels SensorThings API. In: INFORMATIK 2021. Gesellschaft für Informatik, Bonn, S. 301–312, 2021.
- [Kr19] Kratzert, F.; Klotz, D.; Shalev, G.; Klambauer, G.; Hochreiter, S.; Nearing, G.: Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrology and Earth System Sciences, 23(12):5089–5110, 2019.

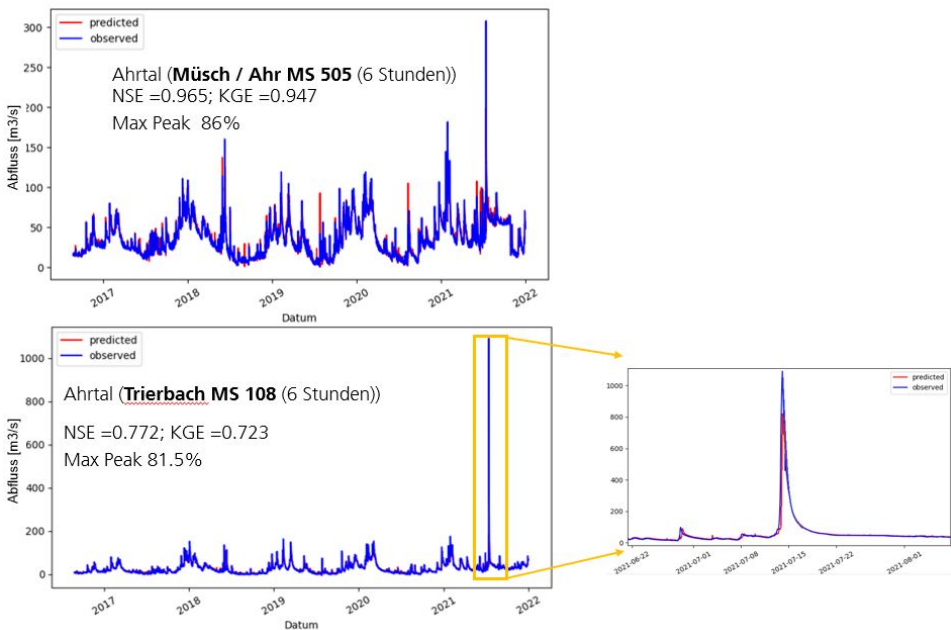


Abb. 3: 6-Stunden-Vorhersage für die Stationen 505 und 108, mit Vergrößerung des Ahrtalereignisses

- [LBZ17] Luong, M.-T.; Brevdo, E.; Zhao, R.: Neural Machine Translation (seq2seq) Tutorial. <https://github.com/tensorflow/nmt>, 2017.
- [LHK16] Liang, S.; Huang, C.-Y.; Khalafbeigi, T.: OGC SensorThings API Part 1: Sensing, Version 1.0. 2016.
- [LK19] Liang, S.; Khalafbeigi, T.: OGC SensorThings API Part 2–Tasking Core, Version 1.0. 2019.
- [XYD20] Xiang, Z.; Yan, J.; Demir, I.: A Rainfall-Runoff Model with LSTM-based Sequence-to-Sequence Learning. *Water Resources Research*, 56, 01 2020.

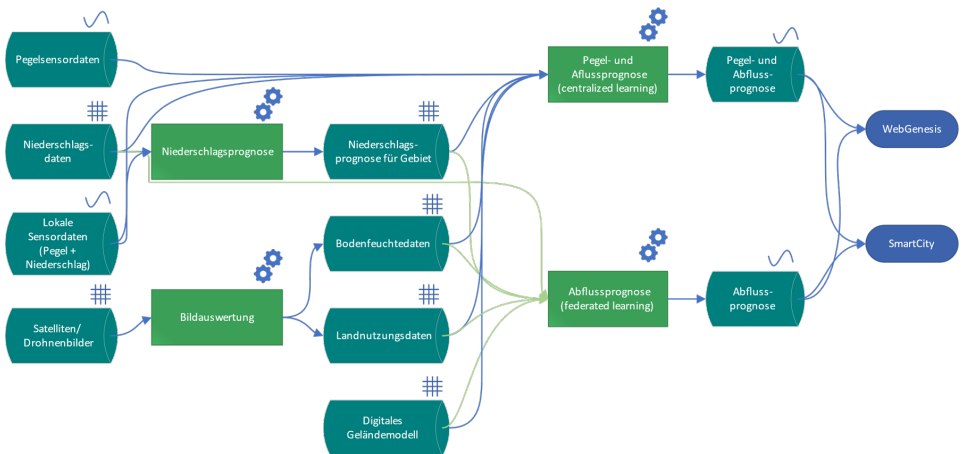


Abb. 4: Zusammenspiel von Daten und Modulen für Sturzflut-Frühwarnung

Sustainability in Artificial Intelligence - Towards a Green AI Reference Model

Sebastian Weber,¹ Achim Guldner,¹ Lejla Begic Fazlic,¹ Guido Dartmann,¹ Stefan Naumann¹

Abstract: The interest in Green Artificial Intelligence (AI) is growing as AI research is increasingly focusing on and taking into account environmental sustainability. This paper aims to clarify and emphasize the distinction between terms like sustainable AI, Green AI, Green by AI, and Green in AI, highlighting their importance in the context of environmentally responsible AI practices. We find that existing Green Software reference models are insufficient for meeting the unique requirements of Green AI. Thus, we argue that a tailored Green AI reference model is needed to guide and promote environmentally responsible practices in the field of AI, addressing the special considerations associated with Green AI.

Keywords: Green AI, Sustainable AI, Reference Model, Resource and Energy Consumption

1 Introduction

The use of Artificial Intelligence (AI), and especially its sub-area Machine Learning (ML), increasingly plays a central role in all areas of life and brings many advantages to various application domains [Ma23]. Given the accelerating effects of climate change, it has also become increasingly urgent for every sector to undertake measures to reduce its greenhouse gas emissions, both to mitigate its contribution to this global challenge and to work towards a sustainable future. While it is true that enhancing efficiency and sustainability in domain-specific application areas can be achieved through the optimization with AI and ML-based systems [Ro22; Vi20], equal importance should be placed on constructing these systems in a manner that they adhere to sustainability specifications, as described by Naumann et al. [Ke18; Na11; Na20], without becoming resource drivers.

Schwartz et al. [Sc20] already pointed out that there are two research areas of AI, one that focuses on improving the outcome in the form of accuracy (which they call 'Red AI') and one that tries to balance the outcome and the impact on the environment (which they call 'Green AI'). There is evidence that the last percentage points in a 'Red AI' application use a significant amount of energy [Se22], e. g., using twice the energy to gain the last percentage point in accuracy. Current work by Verdecchia et al. [Ve23] showed that the interest in Green AI is rising, but mostly focused on the training phase of ML algorithms (neglecting

¹ Institute for Software Systems, Environmental Campus Birkenfeld, Trier University of Applied Sciences, 55765 Birkenfeld, Germany {seb.weber,a.guldner,l.begic,g.dartmann,s.naumann}@umwelt-campus.de

the rest of the AI pipeline). Despite resulting in promising energy savings of over 50 %, these findings are only rarely transferred to the industry and important stakeholders.

Besides the energy consumption necessary to provide the high compute power, several other challenge arise in the field of Green AI along the life-cycle of an AI-based system. The systems usually require large amounts of data for (re-)training that, in turn, need an infrastructure for collection, transfer, management, and storage, all of which require hardware resources and power to operate (e. g., as explored in Wu et al. [Wu21]). Furthermore, the rapid evolution of the systems often requires the replacement of hardware components, which also requires resources. This aspect was addressed, e. g., by Luccioni et al. [Lu23], who include the embodied emissions associated with the materials and processes involved in the hardware production. However, as pointed out by Bashir et al. [Ba23], integrating embodied emissions into the calculation of the overall footprint can overstate the possible carbon reductions in the hardware production chain and thus lower the incentives for directly optimizing the operational carbon.

To address these issues of AI-based systems, it is necessary to enable the stakeholders of such systems to design, create, operate, and use them in a way that their impact on the environment is as small as possible. To do so, a reference model is required that structures research, helps the industry to understand the issues and make them transparent in their systems, and provides recommendations, processes, and tools to reach the goals of assessing and minimizing the environmental impact of AI-based systems. The objective of this paper is to establish the context and provide a definition for a *Green AI Reference Model*. The section on related work outlines the contemporary understanding of Green and Sustainable AI, Green AI, and Green Software Engineering. It also presents the findings of an exploratory literature review that we undertook with the aim of identifying and categorizing existing models from research, focused on the eco-sustainability of AI-based systems. In the “Approach and Definitions” section, we clarify the concept of “Green AI”, explain our understanding of a reference model, and define the term “Green AI Reference Model”. Subsequently, we categorize the Green AI reference model. To conclude, we summarize our findings and offer projections on the potential form of this model. The overarching goal of our work is the creation and continuous improvement of the reference model.

2 Related Work

Comparing the research areas of Green and Sustainable Software and Green and Sustainable AI with each other, the following Fig. 1 juxtaposes the amounts of research papers found in the context of the two fields from the first contributions in 2009. While there are papers containing the term “Sustainable Software” before 2009, like [HL08], these sources usually use the term sustainable in the sense of “lasting” or “durable” and refer to issues like modular, interoperable, and reusable software. Of course, this is also one aspect of sustainability, but before 2009, no papers were found that considered ecological sustainability. Therefore, 2009

was set as the start of the interval. For 2023, the number was extrapolated by multiplying the results already available at the end of April by three (dashed lines).

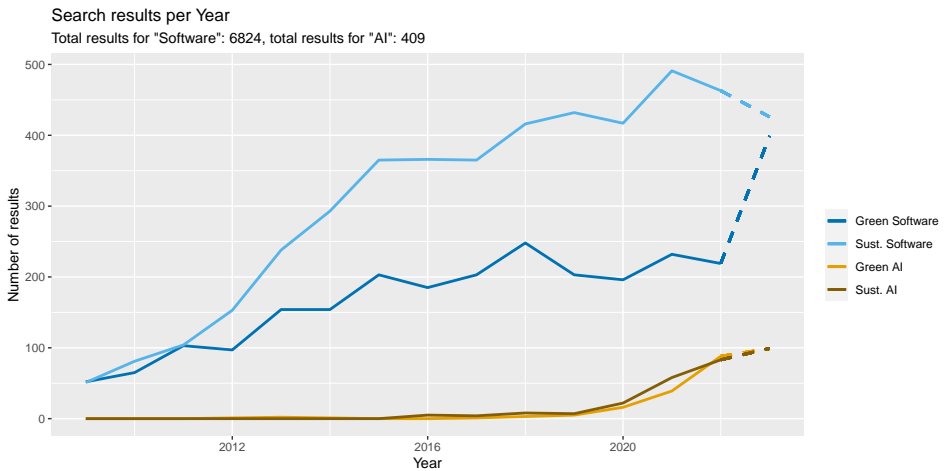


Fig. 1: Google scholar results for green and sustainable software and green and sustainable AI

2.1 Green and Sustainable AI

Sustainability is a complex concept that can be organized into different aspects, commonly referred to as dimensions or pillars, depending on the definition. These dimensions, including ecological, social, and economic, help categorize and understand the various components of sustainability, which was already summarized for the field of software, e. g., by Calero et al. [Ca21] and by van Wynsberghe [Va21] for the field of AI. From this, we derive the concept of Sustainable AI to entail achieving balance across all dimensions of sustainability, whereas Green AI – as a part of Sustainable AI – focuses on ecological sustainability.

There are several different approaches to make AI “greener”, i. e., to optimize AI-based software systems, such that their environmental impact is as low as possible. As Verdecchia et al. [Ve23] point out, there are different approaches towards Green AI, for example measuring the energy consumption during training or inference. This not only makes the energy requirements transparent for developers, users, or customers, but also aides developers if they want to optimize their systems. The AI community *Huggingface* added a CO₂ emission field to its models². Although attempts have been made to demonstrate and compare the energy usage and CO₂ footprint of AI models (e. g., [Ge22b; Gu21; He20]), it remains challenging to make meaningful comparisons in actual practice. For instance, the process of incorporating the CO₂ equivalent emitted by a model during its training into a meaningful comparison with another system requires a comprehensive understanding of the specific hardware used, the energy mix during that time, and other pertinent factors.

² <https://huggingface.co/docs/hub/model-cards-co2> [2023-05-01]

Looking deeper in the current discussion around large language model (LLM), e. g., authors in [Ge22a; RA22] analyzed the performance and carbon footprint for BERT-based approaches, light models, and hybrid (symbolic and ML approaches) when they are used in very specific and vertical domains. In a series of experiments to examine how LLM lose value, they document clear disadvantages in terms of the necessary computational sources and energy consumption. The classification results in particular domains show that using LLMs does not demonstrate a substantial difference compared to other approaches even when they are used in a specific language domain. Implementation of a symbolic approach that combines domain specific linguistic knowledge into the feature representation enables achieving good performance. The carbon footprint and energy consumption are measured for different models during testing, training, and validation. The results by above mentioned authors show that the "light" models and symbolic AI actually achieved energy efficiency savings of around 75 % when compared to BERT-based models. This corresponds to a comparable reduction in both cost and carbon footprint. It is possible to have NLP systems that are fast, have a high performance, and a smaller ecological footprint. These results show that it is necessary to systematically examine optimization possibilities, which can be assisted with a structured model. They also show that measured values (like 75 % savings in this case) are valuable but can also be deceiving when they cannot be interpreted well, without comparisons, metrics, etc.

2.2 Green AI and Green Software Engineering

AI-based systems are, of course, also software systems. Hence, Green AI is a sub-field that belongs to the broader category of Green Software Engineering (GSE). These two fields exhibit certain similarities, in particular, when it comes to the need for a platform that delivers the AI-based service (e. g., a website for the user interface). GSE also faces challenges in standardizing its terminology, which has resulted in the interchangeable use of terms like Green Coding and Sustainable Software Engineering (SSE). This was, e. g., addressed by Calero and Piattini [CP17]. Nonetheless, due to the unique characteristics of AI, such as the ability to “learn” rather than being strictly algorithm-based and its black-box nature, it necessitates a distinctive approach.

2.3 Green AI Reference Models in the Literature

In our research, we encountered difficulties finding a (reference) model that provides a comprehensive overview of the Green AI topic. Therefore, we scanned the literature, to assess if (1) any approaches for a reference model in the area of Green AI already exist and (2) which reference models are available for GSE that could be applied or transferred and adapted to the area of Green AI. We began by formulating search queries that took into account the various

terms commonly used in the field and searched the Google Scholar³ and dblp⁴ databases. To ensure a comprehensive search, we included acronyms associated with AI and GSE in our queries. This helped us capture a broader range of relevant publications. One of our key objectives was to identify papers that discussed or proposed a reference model in the context of our research topic. We specifically looked for works that provided a structured framework or model for understanding and implementing concepts related to AI and GSE. Since the comprehensive results and search queries would exceed the scope of this publication, we provide the lists as supplementary material at <https://gitlab.rlp.net/green-software-engineering/towards-a-green-ai-reference-model-supplementary-material>.

The scan shows that several, differently named models (like process, life-cycle, or measurement models), frameworks, etc. exist. However, especially one model (i. e., [Na11]) can be viewed as a state-of-the-art reference model in the area of Green Software. Despite its continued development [Ke13; Ke15; Ke18; Na15], it is only partially applicable for Green AI due to several reasons: (1) It lacks an extension to address the unique requirements of AI applications due to their learning nature and (usually) black-box characteristic. (2) It does not incorporate current research in the sub-field of Green AI.

3 Approach and Definitions

Based on the exploratory literature scan and previously known papers, we compiled an overview of the state of research in the area of Green AI, and especially models, recommendations, and tools that aid stakeholders to analyze, evaluate, and optimize the ecological sustainability of their AI-based systems. The overarching goal is to create a reference model for Green AI, which we delineate in the following. To approach this goal, we have compiled definitions for the most important terms, a model hierarchy, and initial considerations about the possible contents of the model.

3.1 Green By AI vs. Green AI vs. Green in AI

Ongoing discussions revolve around the terms “Sustainable AI”, “Ecologically Sustainable AI”, and “Green AI”. In addition to this, there is another conversation surrounding the term “Green (by/in) AI”. In literature and research, the term “Green AI” encompasses various interpretations depending on the context it is used in: One interpretation of “Green AI” involves the endeavor to make AI technology itself more environmentally friendly and energy-efficient (e. g., as discussed by [Sc20]). This perspective focuses on reducing the carbon footprint and resource consumption associated with AI systems throughout their life cycle, encompassing stages such as design, training, deployment, and operation.

³ <https://scholar.google.com/>

⁴ <https://dblp.uni-trier.de/db/>

Alternatively, “Green AI” is also used to describe applications that leverage AI⁵ to promote environmentally friendly practices and make processes more ecologically sustainable. This aspect is sometimes also referred to as “Green by AI” (e. g., in [Wa21]). These applications utilize AI algorithms and techniques to optimize resource allocation, minimize waste, or enhance efficiency across diverse domains such as energy production, transportation, agriculture, and more (Rolnick et al. [Ro22] produced an extensive list, with ways to apply AI/ML to tackle climate change).

In this paper, we aim to establish a shared understanding and facilitate effective communication regarding the various dimensions of “Green AI”. To promote a comprehensive understanding of the term, similar to van Wynsberghe’s [Va21] approach in defining sustainable AI, we propose the following definition, aligning it with the definition of “Green Software” by Naumann et al. [Na11].

Definition 1 (Green AI) *Green AI refers to AI-based systems, whose direct and indirect negative impacts on the environment that result from development (e. g. training), deployment (e. g. MLOps), usage (e. g. inference) and end-of-life of the systems are minimal “Green in AI” and/or which have a positive effect on ecologically sustainable development “Green by AI”.*

Acknowledging that the term “AI for Green” is used interchangeably with “Green by AI” as employed, e. g., in [Wy22]. It should be pointed out that “Green in AI” and “Green by AI” approaches define a spectrum in which the AI-based systems can be ranked. This means that a “Green in AI” system can be applied onto a “Green by AI” problem. However, the AI use case should be considered regarding its ecological impact as well. Similarly, this principle applies to “Green by AI”. While it is not obligatory for “Green by AI” applications to adhere to environmentally friendly AI best practices, it is crucial to recognize that the overall conservation of resources should not be compromised by utilizing AI. Consequently, it is always necessary to consider whether a non-AI solution exists that can effectively address the problem while requiring fewer resources.

3.2 Reference Model

In software engineering, reference models are important tools to structure and guide research as well as help practitioners to quickly understand an area. According to Fettke and vom Brocke [FB19] however, there is still no generally accepted definition for reference models in computer science. We therefore first define what we understand as a reference model in the context of this work and give some general examples, before we define the term *Green AI Reference Model*.

⁵ e. g., in <https://www.green-ai-hub.de/> [2023-04-30]

3.2.1 Definition of 'Reference Model'

For this paper, we build upon a modified definition from MacKenzie et al. [Ma06].

Definition 2 (Reference model) *A Reference Model serves as an abstract framework for comprehending the primary entities and their significant interrelationships within a particular environment or domain. It facilitates the development of specific reference architectures or concrete architectures by adhering to consistent standards or specifications relevant to that environment. More specifically, it should comply with the following requirements: A reference model*

- *helps stakeholders (researchers, developers, users, etc.) to understand the big picture of a topic by structuring its main components, e. g., by classifying common terminologies,*
- *provides guidance to create new, or categorize more specific, models or to compare frameworks, and*
- *helps to identify potential and gaps in research.*

But it is independent of specific standards, technologies, implementations, or other concrete details.

Since at present, the field of Sustainable AI is rapidly evolving, in addition to these requirements, the reference model should be able to adapt to these changes and be updated to include emerging research results.

3.2.2 Green AI Reference model

Accordingly, we define the “Green Artificial Intelligence Reference Model” as follows.

Definition 3 (Green AI Reference Model) *The Green AI Reference Model is an abstract framework for understanding the principal entities and their significant interrelationships that impact the ecological sustainability of AI-based software systems. It enables the development of specific models, methods, procedures, or architectures, employing unified terms, definitions, standards, and specifications. These collectively ensure that the life-cycle processes of AI-based systems, from creation over usage to disposal, are carried out in a manner that aligns with the principles of Green AI (see Def. 1).*

Furthermore, the reference model should address the following aspects:

Supporting stakeholders: *It should classify and sort common definitions and the partially ambiguous terminology (such as Green AI, Green By AI, Sustainable AI, etc.), e. g., in form of a glossary or taxonomy. Furthermore, the stakeholders' perspectives should be included*

into each model component.

Modeling: It should provide guidance to create new or categorize existing, yet more specific models and allow the comparison of frameworks. In this regard, it should allow state-of-the-art frameworks to be adapted to support Green AI.

Enable research: It should allow scientists to identify gaps in research, e. g., by outlining the critical elements (like measurements, transparency of the resource consumption, etc.), so that areas that require more investigation become obvious. Hence, it should allow the creation of more effective and efficient Green AI models.

Ecological sustainability: In addition to the general goals of the reference model, it must enable the analysis and evaluation of the AI-based systems in terms of issues that cause resource and energy consumption, which triggers impact the consumption, and how the systems can be optimized to reduce these impacts.

4 Reference Modeling and Discussion

Fig. 2 shows a conceptual categorization of exemplary and existing state-of-the-art models that can be derived from or integrated into the Green AI Reference Model. This illustrates where the reference model should be placed in the model hierarchy of Green Software.

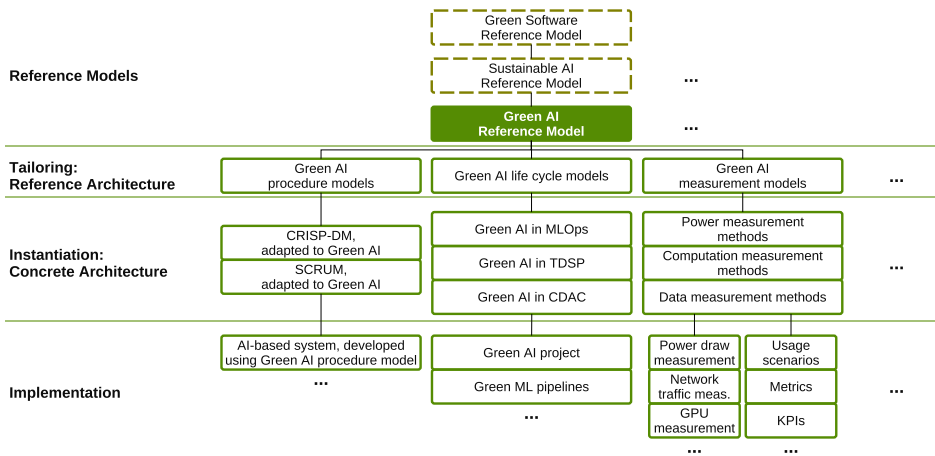


Fig. 2: Reference model hierarchy, instantiation/tailoring, and examples

The model itself can be seen as a tailored Sustainable AI Reference Model (which should also contain models for the other pillars of sustainability), which, in turn, can be seen as a tailored Green Software Reference Model, as described by [Na11]. The Green AI Reference Model can then be tailored to include individual reference architectures, like procedure models, life cycle models, measurement models, etc. that each contain more and more concrete models, with the leafs of the hierarchy representing the specific implementations, measurements, ML models, etc.

To elaborate on the model in the future, we propose a two-pronged approach: Through the vertical expansion of further reference architectures, the addition of stakeholders' perspectives, the addition of measurement models, criteria, and indicators, etc. we aim at broadening the hierarchy to include new and existing architectures (top-down). By integrating existing Green Software and Green AI approaches into state-of-the-art models like CRISP-DM [WH00], CRISP-ML(Q) [St21], MLOps [Kr23], TDSP [Mi17], or CDAC [DA22], or AI projects in development (bottom-up), we want to tailor and verify the reference model, detect influencing factors, research gaps, or missing reference architectures. To assess our Green AI Reference Model, we plan to apply it to existing AI-based applications as well as completely new projects and to create surveys for industry and researchers.

5 Conclusion and outlook

The concept of sustainable AI encompasses several dimensions that need consideration during the development of AI-based applications. We highlighted the significance of a reference model for enhancing our comprehension of Sustainable AI and guiding research. Consequently, we conducted an investigation to identify existing models that adhere to the specified criteria for a Green AI Reference Model. However, our search yielded no suitable options, and also the extension to Green Software did not yield a satisfactory model. The absence of a Green AI Reference Model underscores the pressing need for its development.

In the future, we will focus on the creation of the Green AI Reference Model that serves as a framework for guiding the design and implementation of environmentally conscious AI applications. By incorporating the specific requirements of AI-based systems, including their learning nature and the evolving field of Green Artificial Intelligence, a comprehensive Green AI Reference Model could effectively address the sustainability challenges inherent in AI development. The development of such a reference model should provide researchers, developers, and stakeholders with a standardized approach to ensure the integration of sustainable practices into AI systems, thereby promoting environmentally friendly and socially responsible AI innovation. The reference model should be modular, generic, and expandable to include tailored sub-models and ideas like criteria, metrics, measurement methods, etc. We plan the development in a bottom-up and top-down approach, so that it supports and builds upon already established models, like MLOps, CRISP-DM, etc.

Acknowledgements

This work was funded by the Ministry for Science and Health of Rhineland-Palatinate as part of the research training group (Forschungskolleg) "AI-CPPS", the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety, and Consumer Protection project "KIRA" under Grant 67KI32013B, and the German Federal Ministry for Economic Affairs and Climate Action project "EASY" under Grant 01MD22002D. (<https://www.umwelt-campus.de/en/acknowledgements-towards-green-ai-model>).

References

- [Ba23] Bashir, N. et al.: On the Promise and Pitfalls of Optimizing Embodied Carbon. In: 2. Workshop on Sustainable Computer Systems. ACM, Boston, USA, 2023, URL: <https://doi.org/10.1145/3604930.3605710>.
- [Ca21] Calero, C. et al.: Introduction to software sustainability. In: Software Sustainability. Springer, pp. 1–15, 2021, URL: https://doi.org/10.1007/978-3-030-69970-3_1.
- [CP17] Calero, C.; Piattini, M.: Puzzling out Software Sustainability. Sustainable Computing: Informatics and Systems 16/, pp. 117–124, 2017, URL: <https://doi.org/10.1016/j.suscom.2017.10.011>.
- [DA22] De Silva, D.; Alahakoon, D.: An artificial intelligence life cycle: From conception to production. Patterns 3/6, 2022, URL: <https://doi.org/10.1016/j.patter.2022.100489>.
- [FB19] Fettke, P.; vom Brocke, J.: Referenzmodell. In: Enzyklopädie der Wirtschaftsinformatik. GITO, Berlin, 2019, URL: <https://wi-lex.de/index.php/lexikon/entwicklung-und-management-von-informationssystemen/systementwicklung/softwarearchitektur/wiederverwendung-von-softwarebausteinen/referenzmodell/>.
- [Ge22a] Gee, L. et al.: Fast Vocabulary Transfer for Language Model Compression. In: Proceedings of EMNLP 2022. ACL, Abu Dhabi, UAE, pp. 409–416, 2022, URL: <http://dx.doi.org/10.18653/v1/2022.emnlp-industry.41>.
- [Ge22b] Georgiou, S. et al.: Green AI: Do Deep Learning Frameworks Have Different Costs? In: Proceedings of ICSE 2022. Pittsburgh, USA, 2022, URL: <https://doi.org/10.1145/3510003.3510221>.
- [Gu21] Guldner, A. et al.: Exploration and systematic assessment of the resource efficiency of Machine Learning. In: INFORMATIK2021: Computer Science and Sustainability. Gesellschaft für Informatik, pp. 287–299, 2021, URL: <https://doi.org/10.18420/informatik2021-023>.
- [He20] Henderson, P. et al.: Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning. Journal of Machine Learning Research 21/1, pp. 1–43, 2020, URL: <https://dl.acm.org/doi/abs/10.5555/3455716.3455964>.
- [HL08] Hülse, M.; Lee, M.: Aspects of sustainable software design for complex robot platforms in multi-disciplinary research projects on embodied cognition. In: Intelligent Robots and Systems. Sept. 2008, URL: <http://hdl.handle.net/2160/1868>.
- [Ke13] Kern, E. et al.: Green Software and Green Software Engineering - Definitions, Measurements, and Quality Aspects. In: First International Conference on Information and Communication Technologies for Sustainability (ICT4S). 2013, URL: <https://api.semanticscholar.org/CorpusID:40610952>.

- [Ke15] Kern, E. et al.: Processes for Green and Sustainable Software Engineering. In: *Green in Software Engineering*. Springer, Cham, pp. 61–81, 2015, URL: https://doi.org/10.1007/978-3-319-08581-4_3.
- [Ke18] Kern, E. et al.: Sustainable software products - Towards assessment criteria for resource and energy efficiency. *Future Generation Computer Systems* 86/, pp. 199–210, 2018, URL: <https://doi.org/10.1016/j.future.2018.02.044>.
- [Kr23] Kreuzberger, D. et al.: Machine Learning Operations (MLOps): Overview, Definition, and Architecture. *IEEE Access* 11/, 2023, URL: <https://doi.org/10.1109/ACCESS.2023.3262138>.
- [Lu23] Luccioni, A. S. et al.: Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model. *Journal of Machine Learning Research*/, 2023, URL: <https://www.jmlr.org/papers/volume24/23-0069/23-0069.pdf>.
- [Ma06] MacKenzie, C. M. et al.: Reference model for service oriented architecture 1.0, OASIS standard, <https://www.oasis-open.org/committees/download.php/19679>, 2006.
- [Ma23] Maslej, N. et al.: The AI Index 2023 Annual Report, tech. rep., [https://aiindex.stanford.edu/report/\[2023-05-23\]](https://aiindex.stanford.edu/report/[2023-05-23]), AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, 2023.
- [Mi17] Microsoft, <https://learn.microsoft.com/en-us/azure/architecture/data-science-process/overview> [2023-05-10], 2017.
- [Na11] Naumann, S. et al.: The GREENSOFT Model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 1/4, 2011, URL: <https://doi.org/10.1016/j.suscom.2011.06.004>.
- [Na15] Naumann, S. et al.: Sustainable Software Engineering: Process and Quality Models, Life Cycle, and Social Aspects. In: *ICT Innovations for Sustainability*. Springer, Cham, 2015, URL: https://doi.org/10.1007/978-3-319-09228-7_11.
- [Na20] Naumann, S. et al.: The Eco-label Blue Angel for Software—Development and Components. In: *Progress in IS*. Springer, Cham, pp. 79–89, 2020, URL: https://doi.org/10.1007/978-3-030-61969-5_6.
- [RA22] Rigutini, L.; Algherini, S.: Towards sustainable technology: “green” approaches to NLP, <https://towardsdatascience.com/toward-sustainable-technology-green-approaches-to-nlp-81a9fb2cf521> [2023-05-12], 2022.
- [Ro22] Rolnick, D. et al.: Tackling Climate Change with Machine Learning. *ACM Comput. Surv.* 55/2, 2022, URL: <https://doi.org/10.1145/3485128>.
- [Sc20] Schwartz, R. et al.: Green AI. *Communications of the ACM* 63/12, pp. 54–63, 2020, URL: <https://doi.org/10.1145/3381831>.
- [Se22] Serizel, R. et al.: Performance above all? Energy Consumption vs. Performance for Machine Listening, a Study on DCASE Task 4 Baseline, Preprint, Nov. 2022, URL: <https://hal.inria.fr/hal-03850797>.

- [St21] Studer, S. et al.: Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology. *Machine Learning and Knowledge Extraction* 3/2, pp. 392–413, 2021, URL: <https://doi.org/10.3390/make3020020>.
- [Va21] Van Wynsberghe, A.: Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics* 1/3, 2021, URL: <https://doi.org/10.1007/s43681-021-00043-6>.
- [Ve23] Verdecchia, R. et al.: A Systematic Review of Green AI. *WIREs Data Mining and Knowledge Discovery* 13/4, 2023, URL: <https://doi.org/10.1002/widm.1507>.
- [Vi20] Vinuesa, R. et al.: The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications* 11/1, 2020, URL: <https://doi.org/10.1038/s41467-019-14108-y>.
- [Wa21] Waltersmann, L. et al.: Artificial Intelligence Applications for Increasing Resource Efficiency in Manufacturing Companies - A Comprehensive Review. *Sustainability* 13/12, 2021, URL: <https://doi.org/10.3390/su13126689>.
- [WH00] Wirth, R.; Hipp, J.: CRISP-DM: Towards a standard process model for data mining. In: *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. Vol. 1, 2000, URL: <https://api.semanticscholar.org/CorpusID:1211505>.
- [Wu21] Wu, C.-J. et al.: Sustainable AI: Environmental Implications, Challenges and Opportunities, arXiv: 2111.00364, 2021.
- [Wy22] van Wynsberghe, A. et al.: Special Issue “Towards the Sustainability of AI; Multi-Disciplinary Approaches to Investigate the Hidden Costs of AI”. *Sustainability* 14/24, 2022, URL: <https://doi.org/10.3390/su142416352>.

Landwirtschaftliche Ertragsvorhersage im Kontext begrenzter realer Trainingsdatensätze: ein Transfer-Learning-Ansatz unter Verwendung tiefer neuronaler Netze

Alexander Münzberg¹, Christian Troost², Nils Reinosch³, Daniel Martini⁴, Liv Seuring⁵, Alexander Niehus⁶, Rajiv Srivastava⁷, Thilo Streck⁸, Thomas Berger⁹, Ansgar Bernardi¹⁰

Abstract: Anhand von KI-gestützten Entscheidungshilfen in der Landwirtschaft, beispielsweise durch Anpassung von Düngeapplikationen oder des zeitlichen Feldarbeits-Managements, kann die Produktivität auf einer ökologischen und nachhaltigen Sicht gesteigert werden. Wir beschreiben eine Lösung, um mit Hilfe von neuronalen Netzen Ertrags- und Wachstumsprognosen in realen landwirtschaftlichen Daten zu erzielen. Das Problem geringer Trainingsdatenmengen wird dadurch gelöst, dass zunächst ein System anhand von Simulationsdaten antrainiert und mittels Transfer-Learning an spezifische reale Betriebsbedingungen anhand einiger weniger Realdaten angepasst wird. Die Ergebnisse der Realprognose werden anhand einer Kreuzvalidierungsstrategie evaluiert.

Keywords: Smart Farming; Entscheidungsunterstützende Systeme; Deep Learning; Transfer Learning; Zeitreihenvorhersage; Convolutional Neural Networks; Autoencoder

1 Einleitung

Die Anforderungen an Menge und Qualität von Lebensmitteln und anderen organischen Materialien steigt immer mehr an durch eine wachsende Weltbevölkerung, neue Trends von Ernährungsgewohnheiten der Menschheit und eine steigende Nachfrage an Rohstoffen. Dies erforderte in vergangenen Jahren eine Steigerung der Produktivität in

¹ DFKI GmbH, Smarte Daten und Wissensdienste, Trippstadter Str. 122, Kaiserslautern, 67663, alexander.muenzberg@dfki.de

² Universität Hohenheim, Schloss Hohenheim 1, Stuttgart, 70599, christian.troost@uni-hohenheim.de

³ Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V., Bartningstr. 49, Darmstadt, 64289, n.reinosch@ktbl.de

⁴ Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V., Bartningstr. 49, Darmstadt, 64289, d.martini@ktbl.de

⁵ Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V., Bartningstr. 49, Darmstadt, 64289, l.seuring@ktbl.de

⁶ Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V., Bartningstr. 49, Darmstadt, 64289, a.niehus@ktbl.de

⁷ Universität Hohenheim, Schloss Hohenheim 1, Stuttgart, 70599, rajiv.srivastava@uni-hohenheim.de

⁸ Universität Hohenheim, Schloss Hohenheim 1, Stuttgart, 70599, thilo.streck@uni-hohenheim.de

⁹ Universität Hohenheim, Schloss Hohenheim 1, Stuttgart, 70599, thomas.berger@uni-hohenheim.de

¹⁰ DFKI GmbH, Smarte Daten und Wissensdienste, Trippstadter Str. 122, Kaiserslautern, 67663, ansgar.bernardi@dfki.de

der Landwirtschaft, was aber durch den Einsatz größerer Mengen an Düngemitteln, Tierdung und Pestiziden zu immer höheren Umweltbelastungen führt. Rückstände haben hierbei erhebliche Auswirkungen auf die Tierwelt und Insekten sowie die Bodenqualität und das Grundwasser [MTB22]. Die Anwendung technologischer intelligenter Ansätze verspricht hier Abhilfe. KI-basierte Entscheidungshilfen und die Digitalisierung der Aspekte landwirtschaftlicher Arbeit versprechen individuelle und präzise Optimierungen der Aspekte des Ressourceneinsatzes in der Landwirtschaft [MTB22]. Somit hilft beispielsweise die Möglichkeit einer gesteuerten Ertrags- und Wachstumsprognose dabei, einzelne Parameter der Düngeapplikationen und der Feldarbeitsplanung so zu optimieren, damit mit geringstem Aufwand und Düngemengen größtmögliche Erträge erzielt werden können und der ökologische Fußabdruck dabei deutlich reduziert wird. Eine valide Entscheidungshilfe muss hierbei eine umfassende Sicht auf das gesamte komplexe biologische System berücksichtigen und alle Konsequenzen möglicher Handlungsalternativen prognostizieren. Über die letzten Jahrhunderte hat sich sehr viel Wissen über landwirtschaftliche Prozesse angesammelt. In Simulationssystemen kodifiziert und programmiert lässt sich solches Expertenwissen nutzen, um die Folgen landwirtschaftlicher Entscheidungen (in alternativen Variationen) zu berechnen und einen wertvollen Beitrag bei der Entscheidungsfindung zu leisten. Leider bleiben jedoch bei solchen Berechnungen die spezifischen Details einzelner Standorte oder Situationen bestimmter Betriebe unberücksichtigt bzw. sie erfordern enorm aufwändige Systemanpassungen [MTB22]. Ansätze des maschinellen Lernens, wie tiefe neuronale Netze, haben bereits in vielen Anwendungsbereichen gute Ergebnisse geliefert, bei der Prognose von Daten mit vielen Parametern und komplexen Abhängigkeiten. Solche Methoden erfordern aber viele Trainingsdaten in ausreichender Qualität, um gute Ergebnisse zu erzielen [HT19]. Die gewünschten Realdaten in ausreichender Quantität und Qualität sind in der Landwirtschaft schwer zu beschaffen. Deshalb untersucht unser Ansatz, ob es möglich ist die Vorzüge vom Simulationssystem und tiefen neuronalen Netzen zu vereinen. Somit wird ein sogenanntes tiefes Faltungsnetz, ein Convolutional Neural Network (CNN) trainiert, um die Vorhersagefähigkeit des Simulationssystems zu imitieren (mit gleichem gegebenem Input den gleichen Output zu erzielen). Anschließend wird das CNN durch Training einiger weniger Daten von Realbetrieben an die spezifischen Bedingungen dieser Betriebe angepasst. Es wird erprobt, ob das CNN anschließend in der Lage ist bessere Prognosen für den spezifischen Betrieb zu liefern als das Simulationssystem [Be20]. Für die Generierung der synthetischen Simulationsdaten nutzen wir das, von der Universität Hohenheim entwickelte, Simulationssystem MPMAS_XN. Das System ermöglicht eine umfassende Simulation des Ressourcenverbrauchs, des biologischen Outputs und der ökonomischen Konsequenzen unter verschiedenen Annahmen über Arbeitsprozesse, den Einsatz von Düngemitteln und Pestiziden sowie der Wetterentwicklung. Um eine sinnvolle Breite des Parameterraums abzudecken, wurden vom Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) Daten zu landwirtschaftlichen Prozessen und technischen Möglichkeiten zur Verfügung gestellt [MTB22]. Die genannten Ansätze wurden im Rahmen des vom Bundesministerium für Wirtschaft und Forschung (BMBF, Förderkennzeichen FKZ 01|S19073A) geförderten SimLearn-Projektes erprobt, wobei

zwei der drei zentralen Fragestellungen bereits in vorangegangenen Publikationen beantwortet wurden.

In [Mu22] haben wir gezeigt, dass ein CNN in der Lage ist die simulierten Zeitreihen des Pflanzenwachstums im Feld zu imitieren und korrekt vorherzusagen, wenn es mit Daten aus dem MPMAS_XN-Simulationssystem trainiert wurde. Somit wurde die Eignung des Ansatzes der tiefen neuronalen Netze für diese Aufgabe gezeigt.

In [MTB22] haben wir die Fähigkeit zur Generalisierbarkeit des CNN aufgezeigt und damit, dass das System auch gute Ergebnisse nach Veränderung verschiedener Kombinationsvariationen im Bereich Boden- und Pflanzenprofil, Feldarbeit, Wetter und Fruchtfolgen liefert.

Kernthema dieser Arbeit ist die Untersuchung des Systems im realen Betrieb: Kann das resultierende erlernte System an die Gegebenheiten des einzelnen realen landwirtschaftlichen Betriebes durch Transfer-Learning angepasst werden und somit in diesem Bereich qualitativ hochwertige Ergebnisse liefern?

2 Datenbasis

2.1 Simulationsdaten

Das MPMAS_XN-Modell ist eine Kombination aus dem Mathematischen Programmierbasierten Multiagentensystem-Modell (MPMAS) und Expert-N (XN) Boden- und Pflanzen-Simulator. MPMAS errechnet optimierte Produktions- und Investmentpläne für Farmbetriebe und simuliert die gelieferten wirtschaftlichen Ergebnisse dynamisch im Zeitverlauf. XN simuliert Pflanzenwachstum und Bodenprozesse über die Zeit, einschließlich tatsächlicher Erträge, die mit den von MPMAS ausgewählten Produktionsplänen erzielt wurden [Tr20] [TB16]. Als externen Input stützt sich MPMAS_XN auf einen regelmäßig aktualisierten Datensatz des KTBL mit technischen Merkmalen von tausenden Feldbetrieben und Maschinen- und Gerätetypen sowie täglichen Zeitreihen von beobachteten oder simulierten Wetterdaten über mehrere Jahrzehnte und mehrdimensionalen Bodenprofildaten [Mu22]. Für die systematische Auswahl von Datenproben für das Training der verwendeten Modelle lieferte die Uni Hohenheim eine große Anzahl synthetisch generierter Zeitreihendatensätze über viele Jahre. Die Daten der jeweiligen Zeitreihendatensätze liefern nachfolgend aufgelistete Informationen.

- Bodenprofil-Charakterisierung bzw. -Zusammensetzung
- Pflanzencharakterisierung
- Feldarbeitsprozesse (geplant und durchgeführt)

- Wetter und Klima
- Boden- und Pflanzenwachstumzustände (lassen sich noch weiter unterteilen in Stickstoff-, Kohlenstoff- und Wassergehalt in Boden und Pflanze, Wachstumsmesswerte, etc.)
- Erträge

2.2 Realdaten

Die Realdaten für das Anpassen der entwickelten neuronalen Netze stammen von einem experimentellen Betrieb der Universität Hohenheim und wurden semi-strukturiert als Exceltabelle zusammengefasst. In den Datensätzen Informationen über die jeweiligen Felder und Parzellen sowie Feldarbeitsprozessdaten einhalten. Des Weiteren wurden Daten mit Informationen über Bodenprofil, Pflanzencharakterisierung sowie Wetter- und Klimadaten verschiedener Wetterstationen hinzugefügt. Ein Realdatensatz enthält die Daten einer Parzelle für jeweils ein bestimmtes Anbaujahr. Um diese in eine standardisierte, einheitliche und strukturierte Form zu überführen und für Training und Vorhersage durch Deep-Learning Verfahren anwendbar zu machen, werden die Daten analysiert und neu gelabelt und durch sogenanntes R2RML-Mapping in einen Wissensgraphen im Resource Description Format (RDF) überführt. In dieser Struktur konnten die benötigten Datenkomponenten mit Hilfe der SQL-ähnlichen Abfragesprache SPARQL selektiert werden (speziell für Wissensgraphen). Zusätzlich können durch Hinzufügen von weiterem Expertenwissen aus dem KTBL-Bestand Inkonsistenzen aufgedeckt und fehlende Werte ermittelt werden [Re23]. Die Realdaten werden nach dem Selektieren in die gleiche Form wie die synthetischen Daten überführt. Einzelne Werte wie die Stickstoffzusammensetzung von Düngemitteln oder die Saatgutmenge (Körneranzahl) pro Kilogramm werden mit Hilfe des KTBL-Datenbestandes errechnet. Fehlende Boden- und Pflanzenwachstumzustände werden durch die Verwendung sogenannter Autoencoder interpoliert.

3 Technologien und Methoden

3.1 Convolutional Neural Networks

Mit modernen Deep-Learning Verfahren können Regressionen auf die Datenwerte innerhalb von Zeitreihendatensätzen angewandt werden. CNN sind ursprünglich für die Analyse von Bilddaten entwickelt worden, da sie sich durch effektive Zusammenfassung der Merkmale für die Analyse in mehrdimensionalen Datensätzen eignen [ON15]. Des Weiteren eignen sie sich aber auch sehr gut für die Vorhersage sehr großer Mengen an zweidimensionalen Zeitreihendatensätzen, da durch die Faltungsschicht feste Zeitabschnitte abgetastet und dimensioniert werden können, die schließlich der

Vorhersage mittels Regression unterzogen werden [Sa19]. Durch die Möglichkeiten der Merkmalszusammenfassung und des Downsamplings sind CNN bei der Vorhersage von Big-Data-Zeitreihendatensätzen leistungsfähiger als herkömmliche neuronale Netze oder Rekurrente Neuronale Netze (RNN) (z. B. Long Short Term Memories (LSTM)) [Br19] [Sa19].

Wir verwenden das CNN für die mehrstufige Zeitreihenvorhersage (rekursive Vorhersage) als sogenanntes Multistep-Ahead-CNN um die einzelnen Zeitschritte und somit den zeitlichen Verlauf anhand einer gegebenen Anzahl an Tagen als Startbedingung (in unserem Falle 4 Tage) vorhersagen zu können. Die genaue Funktionsweise und Ergebnisse der generellen Vorhersagefähigkeit wurden bereits in [Mu22] und [MTB22] beschrieben. In [MTB22] wurde zusätzlich die Anpassungsfähigkeit des CNN evaluiert. Dabei zeigte sich, dass es generalisierbar ist, hinsichtlich der Veränderungen von allgemeinen Eigenschaften, wie Bodenprofil, Fruchtart, Wetter und Daten des Feldarbeitsmanagements.

3.2 Autoencoder

Autoencoder besitzen die Fähigkeit, eine Anzahl von Eingabedaten zu komprimieren (Encoding) und somit die Größe an Daten bzw. die Anzahl von Datenattributen (Features) zu reduzieren. Dabei sollen die wichtigen Merkmale erhalten bleiben. Durch das Decoding werden anschließend die dimensionierten Daten auf ihre ursprüngliche Größe der Eingabedaten zurückgeführt bzw. die Daten werden so dekomprimiert, dass sie den ursprünglichen Eingabedaten sehr ähnlich sind. Dies erlaubt es diesen Modellen Rauschen bzw. Ausreißer in Daten zu erkennen. Zusätzlich können Autoencoder dazu verwendet werden, um durch das Encoding/Decoding fehlende Werte zu interpolieren.

Da uns Datenwerte der Boden- und Wachstumszustandes innerhalb dieser ersten 4 Vorhersagetage (die Startbedingung des CNN) in den Realdaten fehlen, wird zusätzlich ein Autoencoder angewendet, der mit den großen Mengen an Simulationsdaten trainiert wurde und in der Lage ist, diese fehlenden Startwerte der Realdaten zu interpolieren. Wir setzen somit die fehlenden Werte auf 0 und der Encoding-/Decoding-Effekt des Autoencoders interpoliert den jeweiligen Wert neu. In [Bh21] wird dieses Verfahren im Detail beschrieben.

4 Verfahren zur Evaluierung

Zu Beginn wurden Autoencoder und CNN anhand 70 unterschiedlicher Variationen von simulierten Parzellendatensätzen (mit Unterschieden in Fruchtfolge, Bodenprofil, Feldarbeitsprozess und Wetter), des MPMAS_XN-Simulationssystem der Universität Hohenheim trainiert. Des Weiteren lieferte die Universität Hohenheim die oben beschriebenen Realdaten als Excel-Dateien.

Für die Evaluierung der Transfer-Learning Fähigkeit des CNN wurden die Realdaten aus den Excel-Tabellen selektiert und sowohl Autoencoder als auch CNN, zur Vorhersage der Zeitreihen (für ein jeweiliges Anbaujahr) zugefügt.

Für die Evaluierungsschritte werden die gewählten Datensätze in Gruppen unterschiedlicher Variationen der Bereiche Fruchtart und Bodentyp zugeordnet, aber auch Datensätze, die Teile eines gesamten Feldes sind, werden zusammen gruppiert. Jeder Datensatz bezieht sich auf ein Anbaujahr im Bereich von 2016 bis 2021. Insgesamt gibt es in den 32 ermittelten Datensätzen die 4 verschiedenen Bodenprofile (Q14, Q21, Q53 und Q59) mit unterschiedlichen Bodenarten (eingeordnet nach Schluff, Ton und Sandanteil nach [Du07]) und weiteren unterschiedlichen Charakterisierungswerten (wie beispielsweise Schüttdichte, Gesteinsanteil und PH-Wert) enthalten. Des Weiteren sind in den Datensätzen jeweils eine der Fruchtarten Mais (MS), Winterweizen (WW), Wintergerste (WG) und Sommergerste (SG) enthalten. Folgende Gruppierungen der Datensätze ergeben sich durch Unterteilungen:

- Gruppe 1: Fruchtart MS und Bodentyp Q21, 3 Datensätze
- Gruppe 2: Fruchtart MS und Bodentyp Q59, 3 Datensätze
- Gruppe 3: Fruchtart MS, 6 Datensätze
- Gruppe 4: Feld Nr. 4 (MS, Bodentyp Q21), 2 Datensätze
- Gruppe 5: Feld Nr. 32 (MS, Bodentyp Q59), 2 Datensätze
- Gruppe 6: Fruchtart WW und Bodentyp Q21, 10 Datensätze
- Gruppe 7: Fruchtart WW und Bodentyp Q53, 8 Datensätze
- Gruppe 8: Fruchtart WW, 18 Datensätze
- Gruppe 9: Fruchtart WG und Bodentyp Q14, 3 Datensätze
- Gruppe 10: Fruchtart WG und Bodentyp Q21, 4 Datensätze
- Gruppe 11: Fruchtart WG, 7 Datensätze
- Gruppe 12: Feld Nr. 400 (WW und WG, Bodentyp Q21), 7 Datensätze
- Gruppe 13: Feld Nr. 500 (WW und WG, Bodentyp Q21), 14 Datensätze
- Gruppe 14: Feld Nr. 700 (WW, Bodentyp Q53), 5 Datensätze
- Gruppe 15: Bodentyp Q14 (SG und WG), 4 Datensätze
- Gruppe 16: Bodentyp Q21 (MS, WW und WG), 18 Datensätze
- Gruppe 17: Bodentyp Q53 (WW), 8 Datensätze
- Gruppe 18: Bodentyp Q59 (MS), 3 Datensätze

Die Evaluierung der Anpassungsfähigkeit des CNN mittels Transfer-Learning wurde in zwei Schritten durchgeführt.

4.1 Evaluierungsschritt 1

Zunächst wurden die Gesamterträge der Realdatensätze der unterschiedlichen Gruppen vor der Modellanpassung vorhergesagt, um dann den Vorhersagefehler zu berechnen. Als Vorhersagefehler für jeden vorhergesagten Wert wurde ein Distanzwert *yieldDist* aus dem jeweils tatsächlichen, durch den Landwirt bestimmten Ertragswert *actYield* und dem jeweils vorhergesagten Ertragswert *predYield* berechnet (1).

$$yieldDist = abs(actYield - predYield) \quad (1)$$

Für jeden Parzellendatensatz einer Gruppe sind ursprünglich 9 mögliche Teildatensätze für das Wetter (von 9 verschiedenen Stationen) vorhanden. Die Distanzwerte der Vorhersagefehler der jeweiligen Parzellendatensätze unterscheiden sich für jeden dieser Teildatensätze (nach Tests) nur sehr gering (im 2. bzw. 3. Nachkommastellenbereich), wobei für die Evaluierung (in beiden zu vergleichenden Evaluierungsschritten) jeweils der Teildatensatz für Wetter gewählt wurde, mit dem der geringste Distanzwert für jeden Parzellendatensatz erzielt wurde.

4.2 Evaluierungsschritt 2

In Evaluierungsschritt 1 haben wir die Vorhersagegüte des nicht re-trainierten Modells festgestellt.

Für jeden Einzelfall der Parzellenbewirtschaftung aus einer der Gruppen mit jeweils mehreren ähnlichen solcher Fälle wird eine Ertragsvorhersage nach Re-Training des Modells durchgeführt. Wir re-trainieren vor jeder Vorhersage das vortrainierte Modell mit jedem, dem Einzelfall ähnlichen Fall derselben Gruppe und nutzen dieses Modell zur Vorhersage des Einzelfalls. Somit machen wir eine Kreuzvalidierung, genauer bedeutet dies wir nehmen innerhalb einer Gruppe alle möglichen Kombinationen, um die vorhandenen Fälle in Re-Trainings und vorherzusagende Fälle aufzuteilen. Als Anpassungsgüte wird für jede Einzelfallvorhersage der Distanzwert *tlPredYield* als Vorhersagefehler berechnet (2).

$$tlYieldDist = abs(actYield - tlPredYield) \quad (2)$$

5 Ergebnisse

Die Streuungsmaße der jeweiligen Distanzwerte *yieldDist*, des ersten Evaluierungsschrittes, werden in Abb. 1 für die jeweilige Gruppe als Boxplots dargestellt.

Ebenso werden die Streuungsmaße der jeweiligen Distanzwerte *tYieldDist*, des zweiten Evaluierungsschrittes in Abb. 2, für die jeweilige Gruppe, als Boxplots dargestellt.

Abb. 3 zeigt die beiden Boxplots der gesamten Distanzwerte (*yieldDist* und *tYieldDist*), aus Evaluierungsschritt 1 sowie Evaluierungsschritt 2.

Vergleicht man die Boxplots aus Abb. 1 und 2 miteinander, wird sichtbar, dass sich die mittleren Distanzwerte nach dem Transfer-Learning in Evaluierungsschritt 2 deutlich verringerten und sich somit die Vorhersagen verbesserten. Eine Anpassung durch Transfer-Learning an die einzelnen Betriebe durch Training von Datensätzen mit ähnlichen Parzelleneigenschaften war somit erfolgreich. In Abbildung 3 wird dies noch einmal an den mittleren Werten aller Distanzwerte (aus allen Gruppierungen) deutlich.

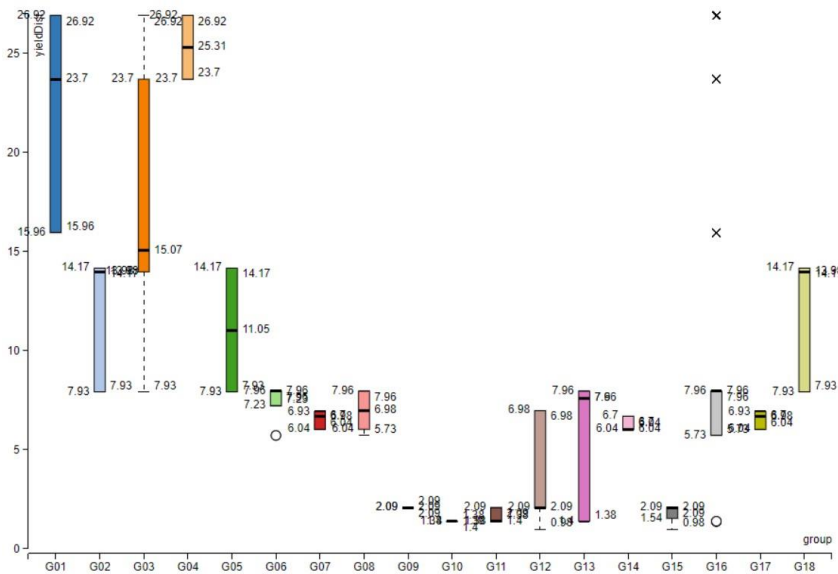


Abb. 1: Evaluierungsschritt 1; Boxplots mit den statistischen Streuungsmaßen der Distanzwerte *yieldDist* für die jeweiligen Gruppierungen

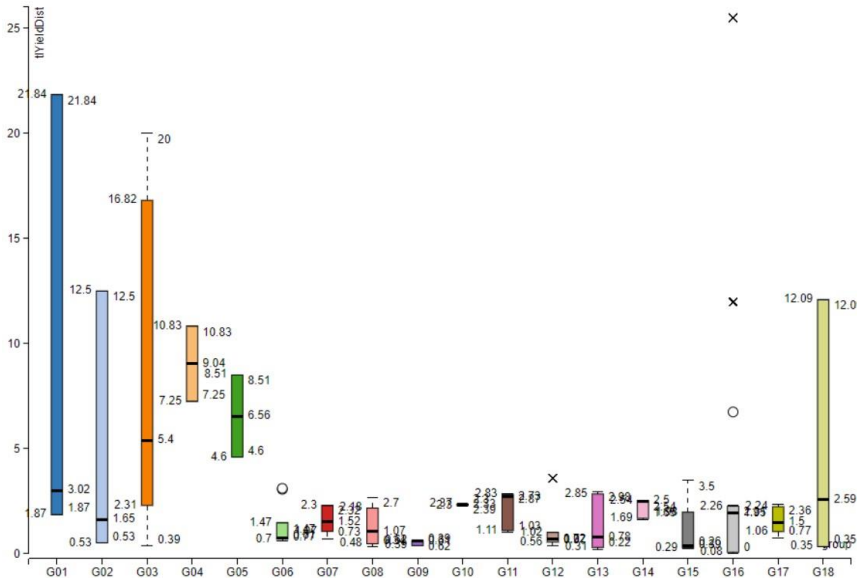


Abb. 2: Evaluierungsschritt 2: Boxplots mit den statistischen Streuungsmaßen der Distanzwerte *tYieldDist* für die jeweiligen Gruppierungen

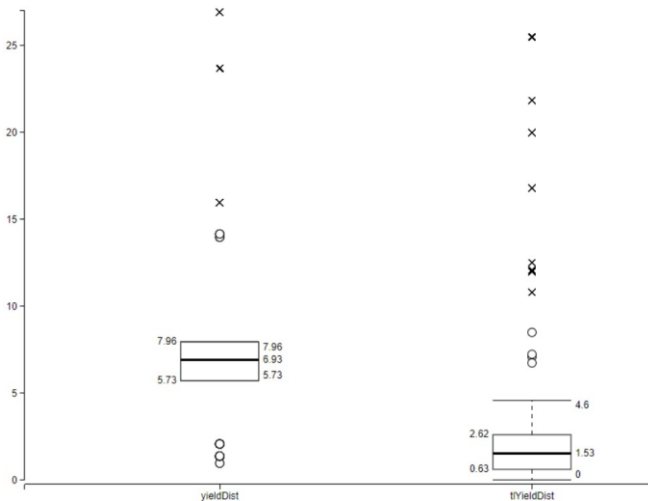


Abb. 3: Boxplots mit den statistischen Streuungsmaßen aller Gesamtdistanzwerte von Evaluierungsschritt 1 und 2

Sichtbar wird auch, dass Werte von Parzellen in Gruppen mit wenigen vorhandenen Datensätzen deutliche Ausreißer enthalten. Dies zeigt, dass eine bestimmte Menge von realen Datensätzen für das Transfer-Learning notwendig ist, um qualitativ hochwertige Ergebnisse zu erzielen.

6 Zusammenfassung und Ausblick

Wir haben das, in vorherigen Arbeiten entwickelte und dargestellte CNN um Autoencoding erweitert, um es für Realdaten-Training und -Vorhersage anwendbar zu machen, auch wenn diese Daten Lücken enthalten. Durch Transfer-Learning ist dieses CNN nun in der Lage, gute Vorhersagen in den limitierten Mengen an gegebenen Realdaten zu treffen. Die These, dass das CNN zu Anpassungen an reale Betriebsbedingungen durch das Transfer-Learning in der Lage ist, haben wir in dieser Arbeit bewiesen. Zukünftige Schritte sind nun zum einen die Beweisführung, dass das angepasste Modell im Realbetrieb besser vorhersagt als das Simulationsmodell (durch Verwendung der gleichen Eingabeparameter der Realbedingungen). Zum anderen soll überprüft werden, ob die Vorhersagequalität noch weiter optimiert werden kann, durch erweiterte Anpassung anhand von Realdaten sowie durch die Anpassung und Optimierung von Hyperparametern der Modelle. Gegebenenfalls können bei der Anwendung des CNN für weitere reale Betriebsumgebungen mit abweichenden naturräumlichen Gegebenheiten zusätzliche Datensimulierungen des MPMAS_XN-Systems, zur Generation zusätzlicher Trainingsdaten für die Anpassung, zwingend notwendig sein. Insbesondere wenn sich Gegebenheiten wie Bodenzusammensetzung, Ackergröße oder Klima und Wetter stark von den bereits trainierten simulierten Daten abweichen und keine diesbezüglichen Realdaten für eine Anpassung vorhanden sind. Zukünftige weitere Forschungsfragen wären, ob und wie sich die Qualität der vorhandenen Realdaten noch weiter durch Hinzunahme von Expertenwissen (z. B. in Form von Ontologien) verbessern lässt und ob sich quantifizieren lässt, welche Menge solcher Realdaten mit bestimmter Qualität für ausreichend qualitative Prognosen vorhanden sein muss. Eine Anwendung der beschriebenen technischen Verfahren in weiteren Anwendungsfelder, in denen Prognosen durch Regression numerischer Daten einen Mehrwert liefern können, ist denkbar. Dies wäre beispielsweise im Bereich weiterer wasserbezogenen Prognosen, z. B. für Grundwassermodelle eine Überlegung wert.

Literaturverzeichnis

- [Be20] Berger, T.; u. a.: Combining Machine Learning and Simulation Modelling for Better Predictions of Crop Yield and Farmer Income. In: A. van Griensven, J. Nossent, D. Ames (Eds.), Proceedings 10th International Congress on Environmental Modelling and Software, Brussels, Belgium, 2020.

- [Bh21] Bhadani, R.: AutoEncoder for Interpolation. arXiv: 2101.00853, 2021.
- [Br19] Brunel, A.; u. a.: A CNN adapted to time series for the classification of Supernovae. In: IS&T International Symposium on Electronic Imaging, 090, 1-4, 2019.
- [Du07] Duewel, O.; u. a.: Bodenarten der Böden Deutschlands. FISBo BGR, https://www.bgr.bund.de/DE/Themen/Boden/Produkte/Schriften/Downloads/Bodenarten%5C_Bericht.pdf, 2007.
- [HT19] van Harmelen, F.; ten Teije, A.: Boxology of Design Patterns for Hybrid Learning and Reasoning Systems“. In: In: Journal of Web Engineering, Vol. 18 1-3, 97–124, River Publishers, 2019.
- [MTB22] Muenzberg, A.; Troost, C.; Bernardi, A.: Using Transfer Learning for Quality Improved Forecasting of Temporal Agricultural Processes by Adapting Convolutional Neural Networks. In: INFORMATIK2022, Gesellschaft für Informatik (GI), Bonn. https://dl.gi.de/bitstream/handle/20.500.12116/39502/kiu_05.pdf, 2022.
- [Mu22] Muenzberg, A.; u. a.: Machine Learning on Simulated and Real Farm Data Based on an Ontology-Controlled Data Infrastructure. In: Proceedings of the AAAI 2022 Spring Symposium on Machine Learning and Knowledge Engineering for Hybrid Intelligence (AAAI-MAKE 2022). Hrsg. von et al. Martin A. Stanford University, Palo Alto, California, USA: <http://ceur-ws.org/Vol-3121/paper18.pdf>, 2022.
- [ON15] O’Shea, K.; Nash, R.: An introduction to convolutional neural networks. arXiv:1511.08458, 2015.
- [Re23] Reinosch, N.; u. a.: SIMLEARN–Ontologiegestützte Integration von Simulationsmodellen, Systemen für maschinelles Lernen und Planungsdaten. In: 43. GIL-Jahrestagung, Resiliente Agri-Food-Systeme, Gesellschaft für Informatik (GI), Bonn, https://dl.gi.de/bitstream/handle/20.500.12116/40295/GIL_2023_Reinosch_477-482.pdf, 2023.
- [Sa19] Sadouk, L.: CNN Approaches for Time Series Classification. In: In: C.-K. Ngan (edt.), Time Series Analysis - Data, Methods, and Applications. IntechOpen, 2019.
- [TB16] Troost, C.; Berger, T.: Advances in probabilistic and parallel agent-based simulation: Modelling climate change adaptation in agriculture. In: In: Sauvage, S., Sánchez Pérez, J.-M., Rizzoli, A. E. (Eds.): Proceedings of the 8th International Congress on Environmental Modelling and Software, July 10-14, Toulouse, France, 2016.
- [Tr20] Troost, C.; u. a.: The Bioeconomic Modelling System MPMAS_XN: Simulating Short and Long-term Feedback Between Climate, Crop growth, Crop Management and Farm Management. In: van Griensven, A., Nossent, J., Ames, D.P. (Eds.) 10th International Congress on Environmental Modelling and Software. Brussels, Belgium, 2020.

Ökologische Nachhaltigkeit -
Resilient Smart Farming Lab
(RSFLab) 2.0 - Edge Computing mit
der HofBox erleben

RSF-Lab'23: Konzepte und Anwendungen zur resilienten digitalen Landwirtschaft

Franz Kuntke¹, Daniel Eberz-Eder², Matthias Trapp³ und Christian Reuter¹

Abstract: Neben positiven Aspekten wie der Produktivitätssteigerung bringt die Digitalisierung auch neue Gefahren mit sich. Entsprechend muss der Prozess gerade in Bereichen von gesellschaftlich enormer Bedeutung kritisch begleitet werden, um eine fundierte Entscheidung bei Auswahl und Entwicklung neuer Technologien zu treffen. Die Vision ist hierbei ein resilientes Smart Farming (RSF), bei dem die Fortschritte der Digitalisierung in der Landwirtschaft genutzt werden, ohne dabei die Ausfallsicherheit der landwirtschaftlichen Primärproduktion und somit die Lebensmittelversorgung der Verbraucher zu gefährden. Dieser Workshop konzentriert sich auf die Bewältigung dieser Forschungsherausforderungen und liefert Beiträge zu verschiedenen Themenbereichen. Dazu gehören (1) ein Hofbox-Ansatz basierend auf etablierten Open-Source Werkzeugen, (2) ein mobiles Assistenzsystem für den Transport von künstlichen Besamungsportionen, (3) die historische Perspektive auf kritische Infrastrukturen in der Region Rhein/Main, und (4) eine Messenger-Applikation zur Notfallkommunikation mittels LoRaWAN-basierten IoT-Setups.

Keywords: Resilienz, Edge Computing, Landwirtschaft, Smart Farming, RSF

1 Einleitung

In der landwirtschaftlichen Praxis gilt es sich zugleich auf die schnell ändernden klimatischen Bedingungen einzustellen und zugleich einen steigenden Bedarf hochwertiger Nahrungsmittel zu erfüllen. Die zunehmende Vernetzung und Digitalisierung innerhalb der Landwirtschaft sollen hierbei einen Teil der notwendigen Effizienzgewinne ermöglichen [DN22]. Wichtig ist dabei allerdings diesen Prozess kritisch zu begleiten, um mögliche Probleme und Gefahren frühzeitig erkennen und lösen zu können. Wenn beispielsweise in der Domäne verbreitete IT-Systeme eine hohe Vulnerabilität aufweisen, wären bei Angriffen auf diese Systeme sehr viele Betriebe in ihrer Betriebsfähigkeit gestört, was wiederum massive Auswirkungen auf die Ernährungssicherheit haben könnte. Hier gilt es durch Analysen bestehender Systeme und Entwicklung neuer, *resilienter* Konzepte für neue Impulse und Innovationen zu sorgen.

¹ Technische Universität Darmstadt (TUDA), Wissenschaft und Technik für Frieden und Sicherheit (PEASEC), Pankratiusstr. 2, 64298 Darmstadt, <nachname>@peasec.tu-darmstadt.de, <https://orcid.org/0000-0002-7656-5919>, <https://orcid.org/0000-0003-1920-038X>

² Land Rheinland-Pfalz, DLR Rheinhessen-Nahe-Hunsrück, Rüdeshheimer Str. 60 - 68, 55545 Bad Kreuznach, daniel.eberz@dlr.rlp.de

³ RLP AgroScience, Breitenweg 71, 67435 Neustadt an der Weinstraße, matthias.trapp@agrosience.rlp.de

In Zukunft strebt man an, dass *Smart Farming* mithilfe der Erfassung und Analyse von Prozess- und Sensordaten sowie die Einbindung georeferenzierter Daten eine präzise und nachhaltige Bewirtschaftung landwirtschaftlicher Flächen ermöglicht. Gegenwärtig machen sich verfügbare Dienstleistungen und Produkte auf dem Markt hauptsächlich das Cloud Computing zunutze, bei dem betriebliche Daten – also freiwillig bereitgestellte Betriebsgeheimnisse – extern auf Servern in Rechenzentren statt lokal gespeichert werden. Sollte ein Cloud-Anbieter diese Daten für illegitime Zwecke nutzen, wäre es nicht nur schwierig, dies nachzuverfolgen, da es sich um seine eigenen Rechnersysteme handelt, sondern die Nutzung wäre auch schwer zu sanktionieren. Ebenso ist für Außenstehende nur schwer ersichtlich, wie die entsprechenden Daten geschützt sind – sowohl gegenüber mutwilligen Eindringens Fremder⁴, als auch gegenüber technischem Versagen⁵.

Eine mögliche Lösung besteht darin, zumindest für essenzielle Dienste eigene dezentrale Systeme einzurichten (z.B. die „Hofbox“), durch die Programme auch ohne Internetverbindung genutzt werden können [Re18, Eb23]. Solche "Offline-First"-Systeme sollten die gewohnten Online-Funktionen bieten, um beispielsweise eine optionale Steuerung über das Smartphone zu ermöglichen und Online-Dienste abrufen zu können. Um die Widerstandsfähigkeit zu erhöhen ist es wichtig, dass die Architektur von Software auch Offline-Szenarien berücksichtigt.

Die exponentiell zunehmende Vernetzung und Digitalisierung bringen große Veränderungen und auch potenzielle neue Verwundbarkeiten auf allen Ebenen mit sich, die den Akteuren selbst teils noch unklar sind [Ku22]. Insbesondere in kritischen Infrastrukturen wie der Landwirtschaft muss zunehmend Aufmerksamkeit auf diese Aspekte gelegt werden. Der Sicherheitsaspekt erfordert eine kritische Auseinandersetzung, um eine Infrastruktur für ein resilientes Smart Farming (RSF) zu schaffen. Diese Infrastruktur sollte die Fortschritte der Digitalisierung in der Landwirtschaft nutzbar machen, ohne die Ausfallsicherheit der landwirtschaftlichen Primärproduktion und somit die Lebensmittelversorgung der Verbraucher zu gefährden.

⁴ „Hackerangriffe“ treffen auch große IT-Firmen, z.B. *Adesso* Anfang 2023 <https://www.heise.de/news/Cyberattacke-auf-IT-Dienstleister-Adesso-Systeme-kompromittiert-Daten-kopiert-7477544.html> (abgerufen am 05.07.2023)

⁵ Auch bei den größten Cloud-Anbietern kann es zu Ausfällen durch technisches Versagen kommen, wie es z.B. 2021 bei Amazons *AWS* <https://www.heise.de/hintergrund/Die-technischen-Hintergruende-von-Amazons-AWS-Ausfall-6293942.html> (abgerufen am 04.07.2023), und *OVH* <https://www.heise.de/news/OVH-Feuer-zerstoert-Rechenzentrum-in-Strassburg-ein-weiteres-beschaedigt-5076320.html> (abgerufen am 04.07.2023) geschah.

2 Aktuelle Beiträge zum Thema

Nach einem erfolgreichen ersten Workshop im Jahr 2022 [Re22] geht das Format nun im Jahr 2023 unter dem Namen „RSFLab 2.0: Konzepte und Anwendungen zur resilienten digitalen Landwirtschaft“ in die zweite Runde. Ziel ist es Forschung zu Resilient Smart Farming zu fördern und sichtbar zu machen, und die Vernetzung in diesem Themengebiet zu stärken. Das Resilient Smart Farming Lab (RSFLab) ist im Projekt GeoBox entstanden und bietet als offenes Laboratory die Möglichkeit zur Vernetzung, Kooperation und Weiterentwicklung. Es wurde dazu erneut ein Aufruf gestartet, wissenschaftliche Beiträge zu diesem Thema einzureichen:

- Technologien für eine widerstandsfähige digitale Infrastruktur
- Cybersicherheit in kritischen Infrastrukturen
- Edge Computing
- Wissenstransfer im Kontext digitaler Landwirtschaft
- Digitale Technologien in der Landwirtschaft
- Intelligente Landwirtschaft und Präzisionslandwirtschaft
- Umweltsensoren und Umweltinformatik (Umweltinformationssysteme)
- Erdbeobachtung, Umweltmodellierung und -simulation
- Anwendungen von geographischen Informationssystemen (GIS)
- Robotik in der Landwirtschaft
- Semantisches Web und Ontologien im Kontext der Landwirtschaft

Die auf Basis eines Peer-Reviews selektierten Beiträge adressieren dieses Thema in vielfältiger Weise.

In ihrem Beitrag „The Hofbox as a decentralised solution for agricultural operations“ stellen Martin Weis (Landwirtschaftliches Technologiezentrum Augustenberg), Sebastian Bökle (Universität Hohenheim, Institut für Agrartechnik) und Christian Bauer (Landwirtschaftliches Technologiezentrum Augustenberg) das Konzept der Hofbox als Alternative zu cloudbasierten Softwareangeboten vor, in Anlehnung an vorausgehende Konzipierungen und Entwicklungen zum Resilient Smart Farming und dem Hofbox-Konzept [Re18, Ku20, Eb21, Eb23]. Dabei werden offen verfügbare Softwarekomponenten auf lokaler Computerhardware bereitstellt. Dieses Konzept ermöglicht die lokale Datenspeicherung und -verarbeitung, sowie den gezielten Datenaustausch über dezentrale Systeme. Der Beitrag stellt eine mögliche Umsetzung des Konzeptes vor, mit Fokus auf der Systemarchitektur und den eingesetzten Softwarekomponenten, welche die Nutzung von Geodaten mit standardisierten Schnittstellen ermöglichen.

Die Autoren Paul Schulze, Frank Fuchs-Kittowski (Hochschule für Technik und Wirtschaft Berlin), Tim Hafemeister (Institute for Reproduction of Farm Animals Schönow), Mario Berndl, Christian Simmet (Minitüb GmbH) und Martin Schulze (Institute for Reproduction of Farm Animals Schönow) untersuchen in ihrem Beitrag

„Mobiles Assistenzsystem zum Monitoring und Bewertung des Zustandes von Ebersperma“ ein mobiles Assistenzsystem, das Teil eines Gesamtsystems zur Überwachung und Dokumentation des gesamten Transportprozesses von Besamungsportionen zur künstlichen Besamung von Nutztieren ist. Mit dem mobilen Assistenzsystem können relevante Faktoren, die die Qualität des empfindlichen Ebersperma während des Transportes beeinflussen können, erfasst, ausgewertet und visualisiert werden. Zu den wichtigsten Funktionen der vorgestellten App gehören die Überwachung des Zustands der Besamungsportionen während des Transports, Benachrichtigungen über den Straßenzustand und typische Logistikfunktionen.




Der Beitrag „Das Netz hat Geschichte: Historisch-technische Analyse der kritischen Infrastrukturen in der Region Rhein/Main“ von Jonas Franken, Marco Zivkovic, Nadja Thiessen, Jens Ivo Engels, und Christian Reuter (Institut für Geschichte und PEASEC, Technische Universität Darmstadt) zeigt wie wichtig es ist, die historische Entwicklung und den Wandel der kritischen Infrastrukturen in der Region Rhein/Main zu verstehen. Diese Infrastrukturen sind oft komplex und entwickeln sich im Laufe der Zeit. Es fehlt allerdings eine historische Perspektive auf die Trends der Technologien in Sektoren, die wesentliche Dienstleistungen für die Gesellschaft erbringen. Im Beitrag wird die Konzeption und Fragestellungen eines interdisziplinären Forschungsprojekts zur Klärung dieser Forschungslücke dargestellt.

Im Beitrag „Optimierte Messenger-Applikation zur Notfallkommunikation via LoRaWAN-DTN“ stellen Denis Orlov, Franz Kuntke, und Christian Reuter (PEASEC, Technische Universität Darmstadt) die Entwicklung einer benutzerfreundlichen Messaging-Anwendung vor, die sich auf die Kommunikation mit einem bestehenden LoRaWAN-basierten Backend konzentriert. Die App soll dabei eine effektive Kommunikation zwischen Helfern und Betroffenen während und nach Krisenereignissen ermöglichen und bietet grundlegende Messaging-Funktionen wie Kontaktverwaltung, Speicherung des Chatverlaufs und Benachrichtigungen. Die auf Smartphone-Nutzung angepasste App bietet dabei auch erweiterte Funktionen wie eine leicht zugängliche SOS-Schaltfläche zum schnellen Senden von Notfallnachrichten.

Literaturverzeichnis

- [DN22] Dörr, J.; Nachtmann, M.: Handbook Digital Farming: Digital Transformation for Sustainable Agriculture. Springer Berlin Heidelberg, Berlin, Heidelberg, 2022.
- [Eb21] Eberz-Eder, D.; Kuntke, F.; Schneider, W.; Reuter, C.: Technologische Umsetzung des Resilient Smart Farming (RSF) durch den Einsatz von Edge-Computing. Gesellschaft für Informatik, Bonn, 2021.
- [Eb23] Eberz-Eder, D.; Kuntke, F.; Brill, G.; Bernardi, A.; Wied, C.; Nuderscher P.: Prototypische Entwicklungen zur Umsetzung des Resilient Smart Farming (RSF) mittels Edge Computing. Gesellschaft für Informatik, Bonn, 2023.
- [Ku20] Kuntke, F.; Reuter, C.; Schneider, W.; Eberz, D.; Bernardi, A.: Die GeoBox-Vision: Resiliente Interaktion und Kooperation in der Landwirtschaft durch dezentrale Systeme. Gesellschaft für Informatik. 2020.
- [Ku22] Kuntke, F.; Linsner, S.; Steinbrink, E.; Franken, J.; Reuter, C.: Resilience in Agriculture: Communication and Energy Infrastructure Dependencies of German Farmers. International Journal of Disaster Risk Science, Bd. 13, Nr. 2, S. 214–229. 2022
- [Re18] Reuter, C.; Schneider, W.; Eberz, D.; Bayer, M.; Hartung, D.; Kaygusuz, C.: Resiliente Digitalisierung der kritischen Infrastruktur Landwirtschaft - mobil, dezentral, ausfallsicher. Gesellschaft für Informatik, Bonn, 2018.
- [Re22] Reuter, C.; Eberz-Eder, D.; Kuntke, F.; Trapp, M.: RSF-Lab'22: Resilient Smart Farming Laboratory: Für eine widerstandsfähige und intelligente Landwirtschaft. Gesellschaft für Informatik, Bonn, 2022.

Mobiles Assistenzsystem zum Monitoring und Bewertung des Zustandes von Ebersperma


Paul Schulze ¹, Frank Fuchs-Kittowski ², Tim Hafemeister³, Mario Berndl⁴,
Christan Simmet⁴ und Martin Schulze⁵ 


Abstract: In diesem Beitrag wird ein mobiles Assistenzsystem zum Monitoring und Bewertung des Transports von Ebersperma präsentiert. Das mobile Assistenzsystem ist Teil eines Gesamtsystems zur Überwachung und Dokumentation des gesamten Transportprozesses von Besamungsportionen für die künstliche Besamung von Nutztieren. Mit dem mobilen Assistenzsystem lassen sich relevante Faktoren, die die Qualität des fragilen Produkts Ebersperma während der Lieferung beeinflussen, erfassen, bewerten und visualisieren, um letztlich das Fahrverhalten der Liefernden und somit die Produktqualität des Eberspermas zu beeinflussen. Im vorliegenden Artikel werden die Anforderungen, das Konzept und die Implementierung des mobilen Assistenzsystems beschrieben. Zu den zentralen Funktionen der App gehören: die Überwachung des Zustands der Besamungsportionen während des Transports, Benachrichtigungen über den Straßenzustand und typische Logistikfunktionen. Die Entwicklung des mobilen Assistenzsystems ermöglicht eine vollständige Digitalisierung des Produktionsprozesses von Besamungsportionen sowie eine umfassende Überwachung und Dokumentation des gesamten Transportprozesses, was zur Verbesserung der Produktqualität beim Kunden führt.

Keywords: mobiles Assistenzsystem, Fahrassistenzsystem, App, Monitoring, Logistik, Ebersperma

1 Einleitung


In der heutigen Nutztierhaltung ist die Künstliche Besamung (KB) das etablierte biotechnologische Verfahren zur Erzielung guter Zuchterfolge. Hierbei werden Besamungsportionen über Straßen an nachgelagerte landwirtschaftliche Betriebe geliefert. Relevante Einflussfaktoren, die während des Transports auf die Produktqualität wirken sind bspw. Erschütterungen [Sc18, Ha22] und Temperatur [Jo00].

¹ Hochschule für Technik und Wirtschaft, Fachbereich 2, Umweltinformatik, Wilhelminenhofstraße 75A, 12459 Berlin, Germany, Paul.Schulze@htw-berlin.de,  <https://orcid.org/0000-0003-4997-7931>

² Hochschule für Technik und Wirtschaft, Fachbereich 2, Umweltinformatik, Wilhelminenhofstraße 75A, 12459 Berlin, Germany, Frank.Fuchs-Kittowski@htw-berlin.de,  <https://orcid.org/0000-0002-5445-3764>

³ Institute for Reproduction of Farm Animals Schönow, 16321 Bernau, Germany, t.hafemeister@ifn-schoe-now.de

⁴ Minitüb GmbH, Hauptstraße 41, 84184 Tiefenbach, Germany, minitube@minitube.de

⁵ Institute for Reproduction of Farm Animals Schönow, 16321 Bernau, Germany, m.schulze@ifn-schoe-now.de,  <https://orcid.org/0000-0001-7932-9412>

Für das fragile Produkt „Ebersperma“ fehlt aktuell ein System zur Überwachung und Dokumentation des gesamten Transportprozesses. Hierzu wurden bereits Vorarbeiten durch die Autoren dieses Beitrags (im Forschungsprojekt „IQ-TranS“⁶) zur Aufzeichnung und Bewertung des Transportstresses während einer Liefertour geleistet [Sc22b, Ha22]. Mit einem mobilen Assistenzsystem für die Liefernden ist eine aktive Einflussnahme zur Reduktion des Transportstresses möglich. In der Literatur finden sich Beiträge, in denen Fahrassistenzsysteme zur Minimierung von Vibrationen verwendet werden [AL22, KL19]. Die Lösungen sind jedoch ausschließlich für jeweiligen speziellen Anwendungsfall entwickelt worden. Es existiert bisher kein mobiles Assistenzsystem für den Transport von Besamungsportionen. Daher ist die Entwicklung eines angepassten Systems notwendig.

Durch ein mobiles Assistenzsystems sollen daher relevante Einflussfaktoren, die sich auf die Produktqualität der Besamungsportionen während der Lieferung auswirken, erfasst, bewertet und visualisiert werden. Mit diesem System soll letztlich das Fahrverhalten der Liefernden beeinflusst und somit die Produktqualität der Besamungsportionen erhöht werden. Ziel dieses Beitrages ist Entwicklung eines mobilen Assistenzsystems zur Unterstützung des gesamten Transportprozesses von Besamungsportionen für die Liefernden.

Dieser Artikel gliedert sich wie folgt: Im Abschnitt 2 werden verwandte Arbeiten mit Bezug zur mobilen Assistenzsystemen beschrieben sowie der Hintergrund der bisherigen Arbeiten am Transportmonitoringsystem „IQ-TranS“ und die in diesem Beitrag verwendeten Methodik vorgestellt. Nachfolgend werden im Abschnitt 3 die im Rahmen einer Anforderungsanalyse erhobenen Anforderungen an das mobile Assistenzsystem dargestellt. Darauf aufbauend wird im Abschnitt 4 das mobile Assistenzsystem konzipiert. Hierfür wurde ein Visualisierungskonzept zur Erfüllung der Anforderungen erarbeitet. Im Abschnitt 5 wird die Implementierung des mobilen Assistenzsystems als App dargestellt. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick im Abschnitt 6.

2 Hintergrund

Einführend werden hier zunächst mobile Assistenzsysteme von intelligente Transportsystemen abgegrenzt und in Beziehung zu Fahrassistenzsystemen gestellt. Es werden aktuelle Arbeiten zur Fahrassistenzsystemen, die dem Monitoring und Bewertung des Zustandes der Ladung dienen, näher beschrieben. Nachfolgend werden allgemeine Anforderungen an mobile Assistenzsysteme dargestellt. Anschließend wird der Hintergrund und die bisherigen Arbeiten am Transportmonitoringsystem „IQ-TranS“ vorgestellt.

⁶ www.iqtrans-projekt.de

2.1 Verwandte Arbeiten

Assistenzsysteme sind technische Systeme, die bei der Durchführung einer Handlung in einer bestimmten Situation unterstützen. Solche Systeme analysieren die gegenwärtige Situation und stellen den analysierten Sachverhalt für den Nutzer verständlich dar. [Te17, Ge14] Eine wesentliche Eigenschaft mobiler Assistenzsysteme ist, dass diese in mobile Geräte eingebunden sind. Ein solches System umfasst alle Soft- und Hardwarekomponenten, um die Assistenzfunktionen autonom bereitzustellen [Te17]. Im Güterverkehr werden intelligente Transportsysteme u. a. eingesetzt, um die Ladung zu überwachen [Mi13]. Fahrassistenzsysteme unterstützen Fahrende bei der Bedienung eines Fahrzeuges [Ge14]. Alle drei Begriffe: mobile Assistenzsysteme, intelligente Transportsysteme und Fahrassistenzsysteme haben eine gemeinsame Schnittmenge – Assistenzfunktionen autonom bereitzustellen.

Sowohl mobile Assistenzsysteme als auch intelligente Transportsysteme bezeichnen im Kern IT-Systeme, die eine gegenwärtige Situation analysieren und dem Nutzenden autonom Fakten liefern sowie Unterstützung bieten. Fahrassistenzsysteme stellen die Schnittmenge zwischen mobilen Assistenzsystemen und intelligente Transportsystemen dar, indem sie eine gegenwärtige verkehrsrelevante Situation analysieren und dem Fahrenden in verschiedenen Situationen oder bei Handlungen im Bezug zur Bedienung des Fahrzeuges unterstützen [KL19, Au08]. Es gibt eine Vielzahl an verfügbaren Fahrassistenzsystemen und jedes dieser Systeme hat einen speziellen, für den bestimmten Anwendungsfall zugeschnittenen Einsatzzweck [Lo14]. Die Spezifik der Anforderungen in den speziellen Anwendungsfällen erfordert spezielle Handlungs- und Lösungsanforderungen. Exemplarisch werden nachfolgend einzelne Artikel beschrieben.

Für die Überwachung der Temperatur gibt es zahlreiche Lösungen im Bereich der Kühlkettenüberwachung bei Lebensmitteln [Wo19]. Es gibt bisher wenig Systeme, die Erschütterungen überwachen und ein direktes sowie visuelles Feedback an den Fahrenden geben [KL19]. Für Rettungswagen wurde ein System entwickelt, welches erschütterungsarme Routen für den Transport verletzter Personen bestimmt [AL22]. Für den Transport von empfindlichen Gütern, die innerhalb eines engen Temperaturbereiches gekühlt werden müssen und die eine Überwachung der Erschütterungen erfordern, gibt es kein bestehendes System. Die Entwicklung eines angepassten Systems ist daher notwendig.

Für mobile Assistenzsysteme können allgemeine Anforderungen definiert werden. So soll die Darstellung der Informationen auf das Wesentlichste zur Unterstützung des Nutzenden reduziert werden [KL19, GSL14]. Die Bedienung ist intuitiv zu gestalten und in einer Art und Weise, dass diese nicht vom Fahrgeschehen oder eigentlichen Tätigkeit ablenkt [Lo14]. Die dargestellten Informationen müssen aktuell und nicht widersprüchlich zu anderen verfügbaren Informationsquellen sein [KL19, GSL14].

2.2 Hintergrund zum Transportmonitoringsystem „IQ-TranS“

Das in diesem Beitrag konzipierte und vorgestellte mobile Assistenzsystem ist Teil des Transportmonitoringsystems zur Qualitätssicherung beim Transport von Ebersperma „IQ-TranS“. Nachfolgend ist es zunächst notwendig die Besonderheiten beim Transport von Ebersperma und der Hintergrund zur Entwicklung des Transportmonitoringsystems vorzustellen.

Die Beschaffenheit von Ebersperma stellt hohe Anforderungen an den Transport dieser leicht verderblichen Ware. Es sind zahlreiche Einflussfaktoren wie die Art und Weise der Verdünnung, Lagerung, Temperatur, Licht u. a., die die Qualität (negativ) beeinflussen, bekannt [HGK16, Ms17]. Letztere Arbeiten zeigen erstmalig, dass Erschütterungen während des Transports einen negativen Einfluss auf die Qualität von Ebersperma haben [Ht22a, Sc18, Ha23]. Damit kann der Transport einen Einfluss auf die Qualität der frisch produzierten Besamungsportionen haben, da diese bis zu 12 h über die Straße vom Produzenten (Besamungsstation) an nachgelagerte landwirtschaftliche Betriebe geliefert werden. Die Qualitätsminderung während des Transports führt zu einer wirtschaftlichen Wertminderung [FBS18, SRW15]. Bei Ebersperma handelt es sich um ein hochempfindliches Gut, welches sehr spezielle Anforderungen an den Transport hat (konstante Kühlkette bei 16–18 °C, erschütterungsarm sowie Licht- und UV-geschützter Transport), um in seiner Beschaffenheit und Qualität erhalten zu bleiben.

Bisher ist eine effiziente und lückenlose Überwachung und Dokumentation des gesamten Transportprozesses von Ebersperma in Echtzeit nicht möglich. Dies liegt bisher im fehlenden Wissen über die genauen Transportparameter begründet. Um ein solches Transportmonitoringsystem zu entwickeln, wurden zunächst im Rahmen des Forschungsprojektes „IQ-TranS“ basierend auf einer Nutzerbefragung die Anforderungen identifiziert und darauf aufbauend eine Grobarchitektur des gesamten Transportmonitoringsystems erstellt, welche sich aus den Komponenten „mobiles Assistenzsystem“, „Logistiksystem“, „Bestellsystem“ und „Transportboxen mit Sensorik“ zusammensetzt. Das entwickelte und in diesem Beitrag vorgestellte mobile Assistenzsystem ist Teil dieses Transportmonitoringsystems (vgl. Abb. 1).

Das entwickelte Grobkonzept des Transportmonitoringsystems stellt sich folgendermaßen dar. Die mit Sensoren ausgestatteten Transportboxen werden verwendet, um die Temperatur der Besamungsportionen zu regulieren und gleichzeitig Echtzeitdaten über den Transportprozess zu erfassen. Die Datenübertragung von den Transportboxen zum mobilen Assistenzsystem erfolgt über Bluetooth Low Energy. Das mobile Assistenzsystem stellt für die Liefernden Assistenzfunktionen zur Lieferung von Besamungsportionen zur Verfügung und wird in den nachfolgenden Abschnitten dieses Beitrages konzipiert. Der aktuelle Zustand jeder Transportbox sowie der Fortschritt der Lieferfahrt werden über Mobilfunk an das Logistiksystem übertragen. Das Logistiksystem generiert automatisch elektronische Ladelisten und Lieferscheine, die an das Assistenzsystem übertragen werden. Die benötigten Daten erhält das Logistiksystem aus dem Bestellsystem und der Pro-

duktionsdatenbank. Das Bestellsystem ermöglicht die Aufgabe, Überprüfung und Bearbeitung von Bestellungen.

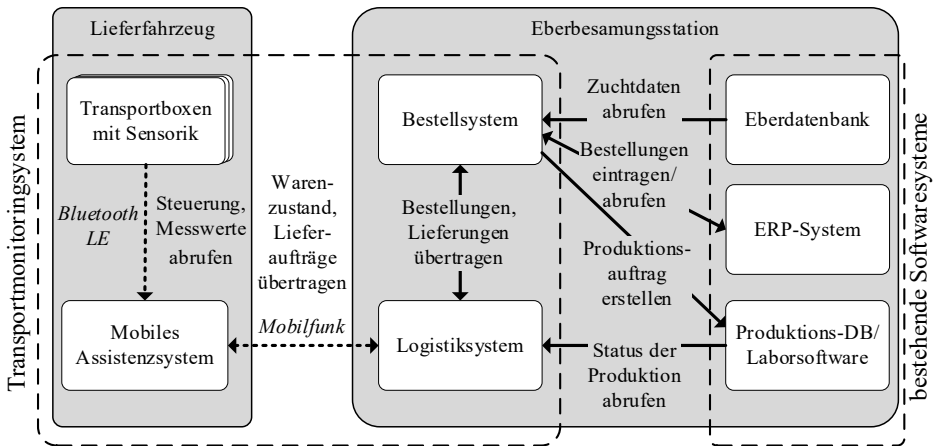


Abb. 1: Architektur des Transportmonitoringsystems

2.3 Methodisches Vorgehen

Es wurde zunächst eine qualitative Befragung von sechs ausgewählten Besamungsorganisationen (Produktionsunternehmen für Besamungsportionen von Ebersperma) mit insgesamt etwa 10.000 Ebern in Brasilien, Deutschland und den USA im Forschungsprojekt „IQ-Trans“ durchgeführt. Hierfür wurde ein Interviewleitfaden entwickelt, der sowohl die aktuelle Situation beim Ablauf der Eberspermalieferung, als auch die angrenzenden betrieblichen Prozesse thematisiert. Die Auswertung wurde qualitativ und ohne statistische Methoden (n = 6). Durch die Befragung war es möglich Schwachstellen zu identifizieren und Anforderungen an ein mobiles Assistenzsystem zur Verbesserung der Produktqualität während des Transports abzuleiten. Im Ergebnis der Befragung wurde eine Dokumentation der Ergebnisse sowie Prozessbeschreibungen und -modelle des IST-Zustands des Transportvorgangs erstellt. Darauf aufbauend wurde der SOLL-Zustand mit Prozessbeschreibungen und -modellen dargestellt. Zur Umsetzung der Anforderungen wurden Visualisierungskonzepte und Assistenzfunktionen erarbeitet. Das mobile Assistenzsystem wurde als Android-App implementiert und die Funktionalitäten des Assistenzsystem evaluiert.

3 Anforderungen

Im Folgenden werden die Anforderungen an das mobile Assistenzsystem beschrieben. Hierfür werden zunächst die Ergebnisse der Expertenbefragung skizziert.

Die Auswertung der Expertenbefragung mit Blick auf den Transportprozesses ergab, dass es einen Bedarf gibt, den Transportvorgang vom Produzenten (Eberstation) zum Kunden (Ferkelerzeugungsbetrieb) vollständig und automatisiert zu dokumentieren sowie die Liefernden beim Transport zu unterstützen. Bei den befragten Unternehmen stellt sich der Transportvorgang als heterogen dar. Der Transport von Besamungsportionen wird üblicherweise von eingewiesenen Mitarbeitenden durchgeführt oder es werden Subunternehmen beauftragt. Die eigenen Mitarbeitenden sind im Umgang mit den empfindlichen Besamungsportionen geschult. Als Fahrzeuge werden Pkw bzw. kleine Lkw verwendet. Die Lieferung erfolgt i. d. R. termingenau zu garantieren Lieferzeiten. Es werden bis zu 30 Betriebe während einer Tour beliefert. Die Besamungsportionen sollen zukünftig in temperierten Transportboxen mit integriertem Messsystem transportiert werden. Hierzu wurde von den Autoren bereits ein Messsystem entwickelt, mit dem es möglich ist, alle relevanten Einflussfaktoren beim Transport von Besamungsportionen automatisch zu erfassen (vgl. [Sc22b]). Für die innerbetriebliche Verwaltung werden derzeit selbstentwickelte bzw. kommerzielle ERP-Systeme (z. B. *EasyKB*) ERP-Systeme von den befragten Unternehmen verwendet.

Für das benötigte spezielle Assistenzsystem wurden folgende Akteure identifiziert: die Eberstation und liefernde Personen. Um den unterschiedlichen Anforderungen der befragten Eberstationen zu entsprechen, muss das geplante mobile Assistenzsystem in den bestehenden Lieferprozess integriert werden. Nachfolgend werden die abgeleiteten Anforderungen und Use-Cases für das Assistenzsystem beschrieben:

- **Kritische Warenzustände und Handlungsempfehlungen:** Während der Fahrt zum Kundenbetrieb sollen der Liefernde über kritische Zustände der Besamungsportionen informiert werden. Es sollen situationsabhängige Handlungsempfehlungen zum Warenzustand ausgegeben werden.
- **Transportboxen und Identifikation der Warensendungen:** Es sind mehrere temperierte Transportboxen mit integriertem Messsystem zu berücksichtigen. Die zu transportierenden Besamungsportionen sind den Transportboxen eindeutig zuzuordnen.
- **Warenzustand:** Weiterhin sollen der aktuelle und der voraussichtliche Zustand der Besamungsportionen bei Lieferung als Motilitätsverlust auf Grundlage der kumulativ gemessenen Erschütterungen berechnet werden. Als Berechnungsgrundlage ist die Publikation aus dem Forschungsprojekt „IQ-Trans“ [Ha23] heranzuziehen.
- **Dokumentation der Warenübergabe:** Die Warenübergabe (teilw. kontaktlos) soll dokumentiert werden. Es sind Abstellort und die Temperatur der Ware zu erfassen sowie eine Fotodokumentation zu erstellen. Die Vollständigkeit der Lieferung ist sicherzustellen.
- **Routing:** Das Routing des Fahrzeuges stellt während der Fahrt zum Kundenbetrieb eine Hauptfunktion dar und soll jederzeit verfügbar sein. Hierzu sind typische Rou-

tingfunktionen zu integrieren. Die Liefernden sollen durch die oben genannten Assistenzfunktionen nicht von der Routenführung abgelenkt werden.

- **Logistikfunktionen:** Es sollen typische Logistikfunktionen bereitgestellt werden. Hierzu zählen automatische Ladelisten, digitale Lieferscheine und Unterstützung des Liefernden bei der Be- und Endladung des Fahrzeuges.
- **Assistenzfunktionen:** Als weiterführende Assistenzfunktionen sollen die Liefernden vor Gefahrstellen, die sich unmittelbar auf der Fahrtstrecke befinden, gewarnt werden. Als Gefahrstellen sind bspw. unebene Fahrbahn, Bahnübergang u. ä. zu berücksichtigen.
- **Installation:** Das mobile Assistenzsystem soll in den Lieferfahrzeugen zum Einsatz kommen. Ein Einbau oder Nachrüsten soll ermöglicht werden.
- **Weitere Anforderungen:** Es sind Qualitätsanforderungen einzuhalten. Hierzu zählen die Fehlertoleranz des Assistenzsystems: Kein Ausfall des Assistenzsystems bei einem Ausfall von einzelnen Komponenten sowie die Austauschbarkeit der Komponenten des Assistenzsystems.

4 Konzeption

Zur Umsetzung der Anforderungen werden Visualisierungskonzepte und die Assistenzfunktionen zur Bewertung der Erschütterungen vorgestellt.

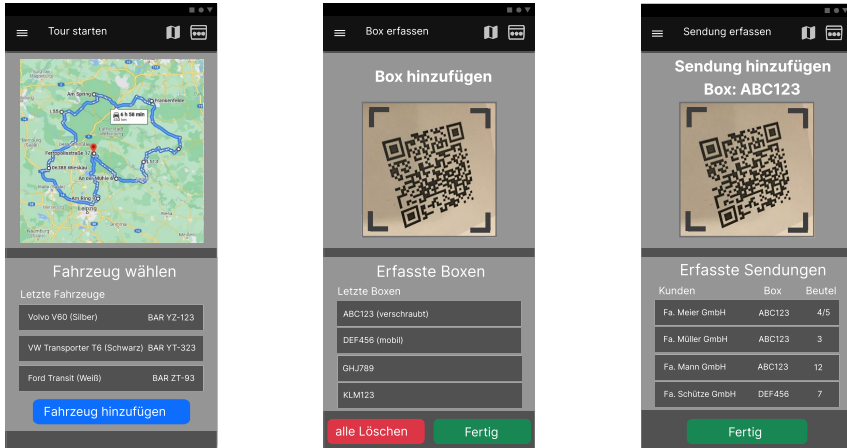
4.1 Visualisierungskonzept

Nachfolgend wird das Visualisierungskonzept zur Umsetzung der oben aufgeführten Anforderungen und Use-Cases für die Interaktion mit den Nutzenden dargestellt. Es wird die wesentliche Funktionsweise des geplanten mobilen Assistenzsystems anhand von Mockups der Benutzungsoberfläche beschrieben. Die dargestellten Mockups entsprechend in der Reihenfolge einer typischen Lieferung, beginnend mit der **Versandvorbereitung und Routenplanung**, welche in der Abb. 2 dargestellt ist.

Tour starten: Der Liefernde erhält eine Übersicht, über die für ihn vorgemerkte und bereits disponierte Liefertour. Auf der Kartenansicht ist die gesamte Lieferroute zu allen Kundenbetrieben incl. der voraussichtlichen Gesamtfahrtzeit dargestellt. Zum Starten der Tour ist das zugewiesene Fahrzeug zu wählen.

Box einbuchen: Der Liefernde erhält eine Übersicht über die bereits befüllten (und erfassten) und mit Sensoren ausgestatteten Transportboxen. Zum Befüllen einer Box ist ein QR-Code, welcher sich an der jeweiligen Box befindet, zu scannen. Das Scannen erfolgt vollautomatisiert durch die aktive Kameraansicht.

Sendung einbuchen: Der Liefernde erhält eine Übersicht, der bereits erfassten Sendungen, in der aktuell ausgewählten Transportbox. Zum Befüllen dieser Box ist der QR-Code, welcher sich an jeder Sendung befindet, zu scannen. Das Scannen erfolgt vollautomatisiert durch die aktive Kameraansicht.



Tour starten

Box einbuchen

Sendung einbuchen

Abb. 2: Visualisierungskonzept: Starten einer Liefertour

Nach erfolgreichem Einbuchen aller Sendungen startet die Liefertour, siehe Abb. 3.

Navigation: Der Liefernde fährt zum nächsten Kundenbetrieb und wird mit einer Navigationsfunktion bei der Streckenführung unterstützt. Die Navigation kann jederzeit unterbrochen werden, um sich nähere Informationen über den Zustand der einzelnen Boxen anzeigen zu lassen, eine alternative Route auszuwählen oder ein neues Ziel in die Routenavigation einzugeben. Bei Annäherung an einen Kundenstandort wird die Übergabefunktion bereitgestellt.

Gefahrenstellen: Nähert sich der Liefernde einer bekannten Gefahrenstelle (z. B. unebener Fahrbahn, einem Unfall oder einer sonstigen Gefahrenstelle), so erhält der Liefernde eine Warnmeldung. Zu der dargestellten Warnung wird zusätzlich noch eine Audiowarnung ausgegeben. Die Navigationsfunktion wird nicht verdeckt und bleibt weiterhin nutzbar.

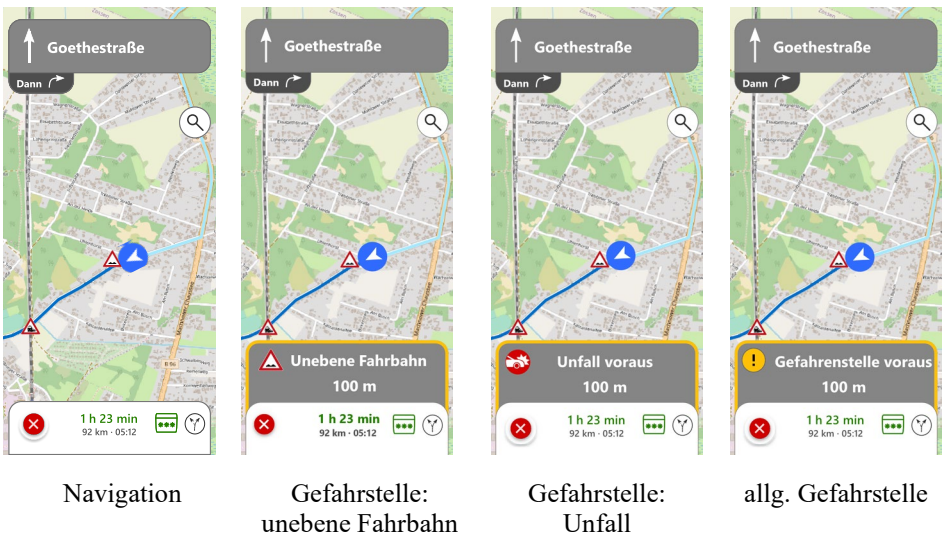


Abb. 3: Visualisierungskonzept: Navigation und erkannte Gefahrenstellen während der Liefertour
 Werden während der Fahrt kritische Zustände der Besamungsportionen erkannt, werden dem Liefernden situationsabhängige Warnungen dargestellt, siehe Abb. 4.

Navigation mit kritischen Warnungen: Wird für mindestens eine der Sendungen eine gefährliche Situation erkannt, so erhält der Liefernde eine Warnmeldung. Zu der dargestellten Warnung wird zusätzlich noch eine Audiowarnung ausgegeben. In der Warnmeldung wird über die Art der gefährlichen Situation informiert. Die Farbgebung impliziert die Gefährlichkeit der Situation. Bei einer roten Warnmeldung, für bspw. eine ausgefallene Kühlung oder starken Erschütterungen und damit einhergehend einer imminenten Verschlechterung des Produktzustandes, ist eine zeitnahe Abhilfe durch den Liefernden unabdingbar.

Temperaturwarnung: Dem Liefernden werden Detailinformationen zur aktuellen Warnmeldung über eine erkannte Temperaturabweichung dargestellt. Hierzu zählen die betroffene Box und die Art der Abweichung. Zusätzlich wird eine Handlungsanweisung dargestellt und eine Audiowarnung ausgegeben.

Kritische Erschütterungswarnung: Es werden Detailinformationen zur Erschütterungswarnung dargestellt. Hierzu zählen die betroffene Box und die Art der Situation. Es wird eine Handlungsanweisung dargestellt und ein Audiowarnung ausgegeben.



Navigation mit kritischen Warnungen

Temperaturwarnung (in zwei Schweregraden)

Kritische Erschütterungswarnung

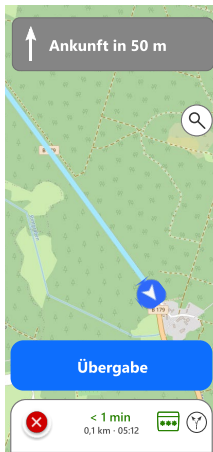
Abb. 4: Visualisierungskonzept: Warnungen an den Liefernden

Bei Ankunft am Kundenbetrieb kann die Warenlieferung übergeben und dokumentiert werden, siehe Abb. 5.

Ankunft beim Kunden: Nähert sich der Liefernde einem Kundenstandort, so wird die Übergabefunktion bereitgestellt. Die Übergabefunktion ist auch manuell zu starten.

Übergabe vorbereiten: Der Liefernde initiiert die Übergabe durch Scannen der jeweiligen Sendung. Das Scannen erfolgt vollautomatisiert durch die aktive Kameraansicht. Steht die Kameraansicht nicht zur Verfügung, kann die Übergabe auch manuell durchgeführt werden.

Übergabe dokumentieren: Zur Übergabe erhält der Liefernde Informationen über den Übergabeort und kann die Übergabe mit Fotos und der aktuellen Temperatur der Sendung dokumentieren. Mit der Übergabe werden die Messungen der Box während des Transports mit der übergebenen Sendung verknüpft.



Ankunft beim Kunden



Übergabe vorbereiten



Übergabe dokumentieren

Abb. 5: Visualisierungskonzept: Warenübergabe am Kundenbetrieb

4.2 Bewertung der Erschütterungen und Berechnung Motilitätsverlust

Zur Bewertung der Erschütterungen wird der Displacement-Index herangezogen. Der Displacement-Index korreliert mit der Straßenqualität [Ha22]. Es konnte durch die Autoren in [Ha23] gezeigt werden, dass ab einem Displacement-Index von $\approx > 3,0$ sich der Zustand der Besamungsportionen bei langer anhaltender Einwirkung von Erschütterungen signifikant verschlechtert. Daher wird dieser Wert für den Displacement-Index als Schwellenwert verwendet, ab dem eine Warnung für den Liefernden ausgegeben werden soll. Mittels des in [Ha23] experimentell bestimmten Regressionsmodells zum Motilitätsverlust von Ebersperma durch die Einwirkung von Erschütterungen über die Zeit, kann der aktuelle und prognostische Zustand der Besamungsportionen berechnet werden. Die Berechnung des Motilitätsverlustes (TSM für *total sperm motility*) wird nach der folgenden Gleichung durchgeführt:

$$TSM [\%] = c * \sum_0^t D_{it}$$

wo c der Motilitätsverlustfaktor, D_i der Displacement-Index und t die Einwirkzeit der Sekunden ist.

5 Implementation und Praxistest

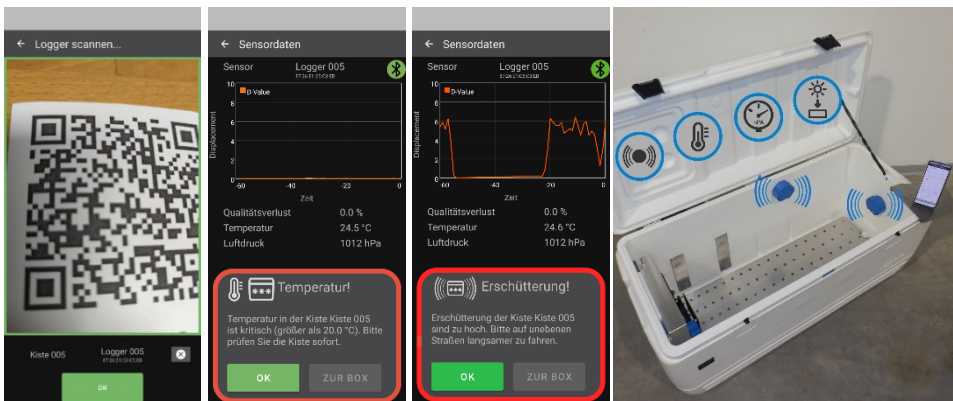
Das entwickelte mobile Assistenzsystem zum Monitoring und Bewertung des Transports von Ebersperma wurde als Demonstrator implementiert. Die Implementierung erfolgte als

native Android-Anwendung unter Verwendung einer Schichtenarchitektur und Anwendung des MVVM-Entwurfsmusters. Für den Praxistest wurden ein Teil der Abschnitt 4.1 aufgeführten Visualisierungen im Demonstrator umgesetzt. In der Abb. 6 sind Screenshots der relevanten Views für den Praxistest dargestellt. Die Maps-Ansicht wurde explizit nicht implementiert, da dies im Wesentlichen ein Google-Maps-Klone darstellt und dieses Feature für den Praxistest nicht relevant war.

Ziel des Praxistest war es, die Assistenzfunktionen zu evaluieren. Mit dem mobilen Assistenzsystem ist es möglich, den Transportstress einzelner Sendungen zu überwachen. Die Messungen der Transportboxen werden sendungsbezogen zugeordnet und gespeichert. Dadurch ist es möglich, den Transportstress für jede Sendung einzeln zu bewerten. Während einer Lieferfahrt wird der Liefernde über kritische Zustände der Spermaportionen informiert und es werden Handlungsanweisungen ausgegeben. Weiterhin wird der aktuelle Zustand der Spermaportionen jeder Box auf Grundlage der kumulierten Erschütterungen berechnet und als „Qualitätsverlust“ (für den Motilitätsverlust) ausgegeben. Durch die entwickelten Visualisierungen erhält der Liefernde einen umfassenden Überblick über den Zustand der transportierten Sendungen. Dadurch lassen sich mögliche Problemfälle während des Transports erkennen und durch frühzeitiges Eingreifen des Liefernden beheben. Dies führt letztlich zu einer Verbesserung der Produktqualität.

Durch die Implementierung als App, kann das mobile Assistenzsystem in Lieferfahrzeugen einfach eingebaut oder nachgerüstet werden. Es können mehrere Transportboxen des Messsystems mittels QR-Codes eingebucht werden. Zur Evaluation wurden 10 Transportboxen erfolgreich eingebucht.

Die Fehlertoleranz des Systems konnte erfolgreich im Test evaluiert werden. Der Ausfall einzelner Messkomponenten wird erkannt und beeinträchtigt andere Messkomponenten nicht. Durch die Modularisierung des Messsystems können Komponenten leicht ausgetauscht werden (siehe Abb. 6). Weiterführenden Assistenzfunktionen wie Warnung vor Gefahrstellen wurden konzeptionell erarbeitet, fanden aber nicht mehr Einzug in die implementierte App.



Box einbuchen

Temperatur-
warnungErschütterungs-
warnungtemperierte Transportbox mit
modularen Sensoren und
Assistenzsystem

Abb. 6: Screenshots des mobilen Assistenzsystems für den Praxistest und Prototyp der Transportbox

6 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein mobiles Assistenzsystem zum Monitoring und Bewertung des Transports von Ebersperma konzeptionell entwickelt und in Teilen als prototypische App implementiert. Das mobile Assistenzsystem komplettiert das Gesamtsystem [Sc22a] bestehend aus einem Echtzeitmonitoringsystem zur Digitalisierung des Spermatransports [Sc22b], einem Dashboard zur Überwachung des Transportvorganges [Sc23] und speziell an den Transport von Besamungsportionen angepassten temperierten Transportboxen mit Sensorik (vgl. Abb. 6).

Die Entwicklung wurde auf Grundlage einer Anforderungsanalyse nach einer Befragung von Besamungsbeherstationen durchgeführt. Mit dem mobilen Assistenzsystem kann der Zustand der Besamungsportionen während des Transports überwacht werden. Der Liefernde wird auf kritische Zustände der transportierten Besamungsportionen hingewiesen und kann diese durch befolgen der jeweiligen Handlungsempfehlungen beheben. Das entwickelte mobile Assistenzsystem kann dazu beitragen, das fragile Produkt „Ebersperma“ während des Transports zu schützen und das Risiko von Schäden zu verringern und damit zur Verbesserung der Produktqualität beitragen.

Das zum Gesamtsystem gehörende Dashboard kann Echtzeit-Updates über den Status der Lieferungen bereitstellen, so dass auftretende Probleme schnell erkannt und behoben wer-

den können. Zusammenfassend kann das entwickelte Gesamtsystem helfen, die Lieferkette zu optimieren, die Kundenzufriedenheit zu verbessern sowie Geschäftsprozesse zu rationalisieren.

Nachfolgende Arbeiten werden sich einer Nutzerevaluation widmen. Es ist geplant, den Prototyp während mehrerer Lieferfahrten zu nutzen und die Nutzererfahrungen anschließend auszuwerten. Auf Grundlage dieser Ergebnisse ist eine Weiterentwicklung des mobile Assistenzsystems geplant. Zudem ist es möglich, das mobile Assistenzsystem auf andere Arten fragiler Güter (bspw. Kunstwerke etc.) zu transferieren. Hier ist allerdings weitere Forschung erforderlich, um die einzuhaltenden Grenzwerte zu bestimmen.

Acknowledgements

Dieses Projekt wird vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) aufgrund eines Beschlusses des Deutschen Bundestages gefördert. Referenznummer: 16KN077342 (IQ-TranS).

Literatur

- [AL22] Aldegheishem, A.; Alrajeh, N.; Parra, L.; Romero, O.; Lloret, J.: Driving Assistance System for Ambulances to Minimise the Vibrations in Patient Cabin. In: *Electronics* Bd. 11, Nr. 23, 2022.
- [Au08] Auckland, R.A.; Manning, W.J.; Carsten, O.M.J.; Jamson, A.H.: Advanced driver assistance systems: Objective and subjective performance evaluation. In: *Vehicle System Dynamics* Bd. 46, Taylor & Francis, Nr. sup1, S. 883–897, 2008.
- [FBS18] Fuchs-Kittowski, F.; Bortfeldt, R.; Schulze, M.: *TransportLog 1.0 – Mobile Sensing-App zur Analyse des Einflusses von Transportstress auf die Eberspermaqualität*: Gesellschaft für Informatik e.V., 2018.
- [Ge14] Gerke, W.: *Technische Assistenzsysteme: vom Industrieroboter zum Roboterassistenten*: DE GRUYTER, 2014.
- [GSL14] Gorecky, D.; Schmitt, M.; Loskyll, M.: Mensch-Maschine-Interaktion im Industrie 4.0-Zeitalter. In: Bauernhansl, T.; ten Hompel, M.; Vogel-Heuser, B. (Hrsg.): *Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung Technologien Migration*. Wiesbaden: Springer Fachmedien, S. 525–542, 2014.
- [Ha22] Hafemeister, T.; Schulze, P.; Bortfeldt, R.; Simmet, C.; Jung, M.; Fuchs-Kittowski, F.; Schulze, M.: Boar Semen Shipping for Artificial Insemination: Current Status and Analysis of Transport Conditions with a Major Focus on Vibration Emissions. In: *Animals* Bd. 12, Nr. 10, S. 1331, 2022.

- [Ha23] Hafemeister, T.; Schulze, P.; Simmet, C.; Jung, M.; Fuchs-Kittowski, F.; Schulze, M.: Intensity and Duration of Vibration Emissions during Shipping as Interacting Factors on the Quality of Boar Semen Extended in Beltsville Thawing Solution. In: *Animals* Bd. 13, Multidisciplinary Digital Publishing Institute, Nr. 5, S. 952, 2023.
- [HGK16] Hoy, S.; Gauly, M.; Krieter, J.: *Nutztierhaltung und -hygiene*. Stuttgart: Eugen Ulmer KG, 2016.
- [Ht22a] Hafemeister, T.; Schulze, P.; Bortfeldt, R.; Grossfeld, R.; Simmet, C.; Fuchs-Kittowski, F.; Schulze, M.: Detection of vibration emissions during the transport of boar semen. In: *Reproduction in Domestic Animals*. Bd. 57, S. 1–30, 2022.
- [Jo00] Johnson, L.; Weitze, K.F.; Fiser, P.; Maxwell, W.M.C.: Storage of boar semen. In: *Animal Reproduction Science* Bd. 62, Nr. 1–3, 2000.
- [KL19] Khan, M.Q.; Lee, S.: A Comprehensive Survey of Driving Monitoring and Assistance Systems. In: *Sensors* Bd. 19, Multidisciplinary Digital Publishing Institute, Nr. 11, S. 2574, 2019.
- [Lo14] Lotz, C.; Herb, T.; Schindhelm, R.; Vierkötter, M.: *Matrix von Lösungsvarianten Intelligenter Verkehrssysteme (IVS) im Straßenverkehr*, Fahrzeugtechnik (Nr. F 97). Bergisch Gladbach: Bundesanstalt für Straßenwesen, 2014.
- [Mi13] Mirzabeiki, V.: An overview of freight intelligent transportation systems. In: *International Journal of Logistics Systems and Management*, Inderscience Publishers Ltd, 2013.
- [Ms17] Schulze, M.; Schäfer, J.; Jung, M.; Waberski, D.: Reproduktionsforschung – Neue Wege zur Sicherung der Spermaqualität in der Schweinebesamung. In: *Züchtungskunde* Bd. 89, S. 22–28, 2017.
- [Sc18] Schulze, M.; Bortfeldt, R.; Schäfer, J.; Jung, M.; Fuchs-Kittowski, F.: Effect of vibration emissions during shipping of artificial insemination doses on boar semen quality. In: *Animal Reproduction Science* Bd. 192, S. 328–334, 2018.
- [Sc22a] Schulze, P.; Fuchs-Kittowski, F.; Hafemeister, T.; Schulze, M.: *Digitalisierung des Spermatransports – Anforderungen und Softwarearchitektur*: Gesellschaft für Informatik e.V., 2022.
- [Sc22b] Schulze, P.; Fuchs-Kittowski, F.; Hafemeister, T.; Urban, A.; Berndl, M.; Simmet, C.; Schulze, M.: Development of a measuring system for monitoring transport of boar semen from artificial insemination centers to sow farms: Gesellschaft für Informatik, Bonn, 2022.
- [Sc23] Schulze, P.; Schütze, F.; Fuchs-Kittowski, F.; Hafemeister, T.; Schulze, M.: Logistics Dashboard for transport of fragile goods. In: *Proceedings of Eighth International Congress on Information and Communication Technology*. Bd. Volume 4. London:

Springer Singapore, 2023.

- [SRW15] Schulze, M.; Rüdiger, K.; Waberski, D.: Rotation of Boar Semen Doses During Storage Affects Sperm Quality. In: *Reproduction in Domestic Animals* Bd. 50, 2015.
- [Te17] Teucke, M.; Werthmann, D.; Lewandowski, M.; Thoben, K.-D.: Einsatz mobiler Computersysteme im Rahmen von Industrie 4.0 zur Bewältigung des demografischen Wandels. In: Vogel-Heuser, B.; Bauernhansl, T.; ten Hompel, M. (Hrsg.): *Handbuch Industrie 4.0 Bd.2: Automatisierung*, Springer Reference Technik. Berlin, Heidelberg: Springer, S. 575–604, 2017.
- [Wo19] Albrecht, W.: Aufgabenbereiche von IT-Systemen in der Logistik. In: Wehking, K.-H. (Hrsg.): *Technisches Handbuch Logistik 2: Fördertechnik, Materialfluss, Intralogistik*. Berlin, Heidelberg: Springer, S. 11–70, 2020.

The Hofbox as a decentralised solution for agricultural operations

Martin Weis,¹ Sebastian Bökle,² Christian Bauer³

Abstract: In digital farming applications cloud-based software offers are predominantly used, which simplifies software and data handling, but lacks transparency of data storage and usage. Internet access becomes essential, which makes time-critical and weather-dependant applications dependent on internet connectivity posing risks for timely execution. To address these issues, a Hofbox concept is being developed that provides openly available software components on local computer hardware. It is based on a modular structure with automatically installable and configurable microservices. The Hofbox thus enables local data storage and processing as well as targeted data exchange via cloud systems that are organised in a decentralised manner. Machinery rings and farmers are targeted decentralised entities. The connected farm boxes are maintained according to an edge computing approach. All components are open source thus ensuring adaptability and expandability, focusing on the use of geodata for small and medium-sized farms.

Keywords: resilience, decentral cloud systems, geodata, software, edge computing, smart farming, Hofbox

1 Introduction

Technology for farmers improves constantly and the digitalisation of processes requires improved data handling and analysis capabilities. The farmer can select from various service providers implementing solutions for specific digital processes. Many digital processes require data, such that the user is required to provide basic data sets for each provider, which leads to a situation, where the farmer has to provide his data in multiple systems. Changes have to be propagated and data updated in all systems, leaving the burden for the user to take care of synchronisation. This often can only be done manually and therefore is a cumbersome and error-prone process, which needs to be repeated regularly and each time has to be prioritised in presence of all other management tasks.

Most farmers already have digital processes in their domain, ranging from applying for government subsidies to bookkeeping of material and management information, often in

¹ Landwirtschaftliches Technologiezentrum Augustenberg, Kutschenweg 20, 76287 Rheinstetten-Forchheim, Germany martin.weis@ltz.bwl.de

² Universität Hohenheim, Institut für Agrartechnik, Fg. Verfahrenstechnik in der Pflanzenproduktion (440d), Garbenstraße 9, 70599 Stuttgart, Germany seboekle@uni-hohenheim.de

³ Landwirtschaftliches Technologiezentrum Augustenberg, Kutschenweg 20, 76287 Rheinstetten-Forchheim, Germany christian.bauer@ltz.bwl.de

exchange with suppliers and contractors. Digitalisation of processes can help to reduce the amount of manual work needed. Most benefits of a digital transformation are achieved by seamless integration of systems. Systems can inter-operate based on defined interfaces existing in hardware and software components. The developments towards systems connected via internet has led to various standardisation efforts for data exchange. Systems can expose APIs (application programming interfaces) to provide interoperability and functionality with access to data.

Unfortunately there are only a few standards for farm data and agricultural processes, existing ones often lack openness and interoperability — data and access to systems are often handled like a trade secret. This may be one of the reasons for the lack of wider adoption by users, since many practical experiences show improvable integration of systems, before new systems inter-operate with existing farm equipment [GG23].

Many processes in the field are supported by location information: GPS systems allow a precise measurement of position and support location based control of machinery with electronic control units (ECU). Such control systems are used to avoid application overlap and precise navigation based on sensor measurements and track information gathered during operation. Additionally, some processes make use of external geodata to optimise the operation, like precision farming techniques using management zones with information about in-field variation [MEA18].

With the advent of remote sensing data becoming widely available for farming operations and processing providers offering products targeting end users specifically, such data can be integrated into workflows on the farm. Such products are available via cloud systems in the internet, due to the necessary data access, storage and computing power needed.

Farmers incorporating multiple providers need to provide at least parts of their master data for each service to benefit from their products. Services providing farm management information systems (FMIS) keep track of assets and support management planning and field operations. Most FMIS optionally include spatial modules to handle geodata of field boundaries, fewer have modules to allow map storage for soil maps, yield measurements, application maps or track logging.

For a site-specific fertiliser application the farmer has to inter-operate with multiple systems and partners: an application map provider may create management zones based on selected satellite data. The map data has to be converted into a task, which can be done within the system, by importing into a capable FMIS or by creating a job on the agronomic terminal and importing the map there. Contractors may be involved to conduct the operation in the field, suitable digital task files can provide all necessary information: ISO 11783-10 defines a standardised data exchange format named ISO-XML [IS15] for task controllers and may contain logged data about a task after finishing it [Sc22]. A typical site-specific task contains information about the management zones with varying amounts of agricultural resources, a field boundary, involved personnel, machinery assets and type of resources used in the field.

The handling of geodata is a relatively new task for the farmer, and requires some knowledge and specific systems, since the geospatial domain has special file formats and metadata requirements to operate. Geodata visualisation and calculations based on the data require capable software being aware of GIS operations. Systems found in the agricultural domains often only implement subsets and may fail to handle the data correctly. Typical issues relate to coordinate system definitions (CRS) and transformation, data types and encoding of attributes as well as geometric type and properties validity. The typical geodata exchange format in the agricultural domain is based on Shapefile, a multi-file format which is difficult to handle for laymen and often lacks metadata in machinery context [Ce]. For FMIS and machinery often undocumented requirements need to be fulfilled, for example the CRS being expected to be WGS84 with geographic coordinates, and only subsets of geometry and attribute types may be recognised, resulting in errors difficult to identify and circumvent.

The full potential of precision farming unlocks with the the ability to continuously gather and combine data sets over time — a typical cycle for a vegetation period starts before sowing and ends after harvest. Combining yield maps with application maps and other data improves the evaluation of decisions and builds knowledge for an optimised management in the future [Mi22].

While software running in a cloud has some advantages like full control of provider on development cycles, version updates and scalability of the systems it runs on, drawbacks arise if these are not available to the user. This can temporarily be caused by connectivity issues or down-times or the end of service.

The goals of edge cloud implementations were elaborated by project initiatives like iGreen [Be14] and Geobox [Eb23]: resilience, offline functionality [RSE19], data security and open standards are the main aspects and the presented approach is guided by these goals.

As agriculture is part of a critical infrastructure and digitisation in agricultural practices further increases, technical and organisational requirements must be addressed. The goal is to involve any kind of farmers, avoid dependencies on hardware and software providers as well as internet connectivity. Ensuring farmer's data sovereignty and safety as a basis to unlock advantages of digital farming and foster participation of farms.

The presented approach addresses the following requirements:

1. Cost-effective infrastructure components based on open source software to avoid lock-in effects and enable adaptation and development
2. Functional components description as IaC (infrastructure as code) to be published under FLOSS licenses
3. Inter-operation via standardised interfaces and APIs (application programming interfaces) to allow extension and exchange of components
4. Farmers must be able to use the solutions on- and offline likewise

5. Cheap, commercial off-the-shelf (COTS) equipment for the hardware that can grow based on future demands
6. Network and server infrastructure for secure network connections, run by trustworthy entities (including decentral, custom deployment possibilities)

2 The Hofbox concept: System design and implemented components

The research aims to provide decentral systems located at farmers offices to provide a means for local data storage and management. This approach introduces resilience against network connectivity issues. It requires a robust software setup with remote management to reduce the administration workload for the user, but still requires additional data handling skills to profit from the software. The software setup is based on modern IT concepts and FLOSS (free/libre open source) software where possible to address requirement 1. Based on previous concepts [WBM22] the implementation design will be elaborated in the following.

The decentral devices are part of an edge computing configuration, where hardware, software and data are available near to the point of gathering and usage of data. The software concept is based on microservices providing functionality and interoperability with standardised interfaces and APIs to address requirement 3. The microservices provide services and can depend on each other [HAF17]. They should not be running in a redundant way: if multiple services depend on a database, they can use the same centrally provided database service and do not need to run their own instance. Fig. 1 depicts dependencies between services: there is a PostgreSQL database spatially enabled with PostGIS as a dependency for other components that need a database like FarmOS, an open source FMIS. In Fig. 1 on the left GIS enabled components are found, on the right there are IoT components that can ingest, store and process data from sensors. Sensors are part of weather stations, machinery or building information systems.

The components in Fig. 1 from left to right are:

- PostGIS: spatially enabled PostgreSQL database, providing an SQL API [HO21]
- GeoServer: provides access to raster, vector and map data via OGC compliant standards and internal APIs [Tu23]
- GeoNode: a geospatial content management system providing geodata handling, web map creation and sharing via Geoserver's and internal APIs [FG20]
- QGIS: open source desktop GIS to organise, visualise and process geodata [QG22], representative for other GIS software and a full desktop environment with abilities to access file based geodata and services
- File server: file based data storage with remote access APIs

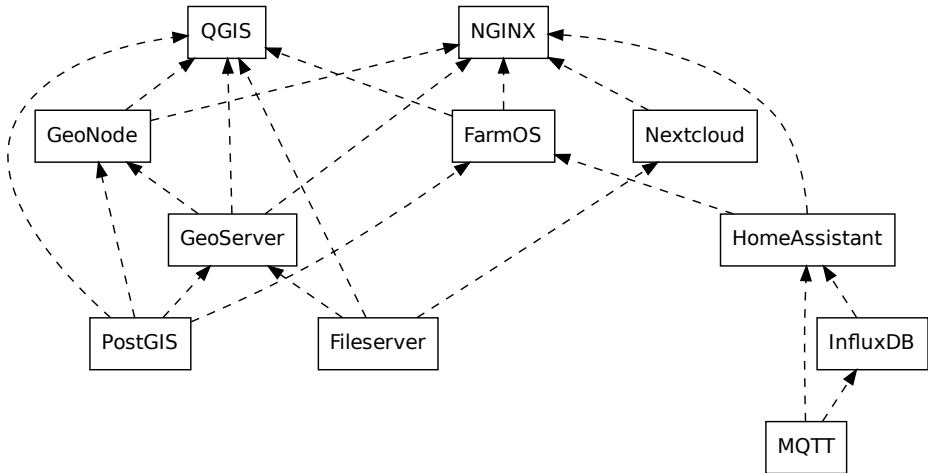


Fig. 1: Microservices and their dependencies graph: arrows link prerequisite services to the ones depending on them.

- FarmOS: an open source FMIS that provides APIs and interfaces for farm related data (management, planning, record keeping information), integrating sensor data and geodata [fa23]
- NGINX: a web server that can act as reverse proxy to access the web interface of components via hypertext protocol HTTP[S] [NG23]
- Nextcloud: provides a cloud solution for file access and synchronisation, includes office components and extended functionality via apps, providing APIs via HTTP[S] [Ne23]
- HomeAssistant: a smart home component with various integrations and APIs for sensors and devices, allows monitoring, control and automation [WCS22]
- InfluxDB: time series database for the storage of measurements with HTTP[S] APIs [Sa23]
- MQTT: a pub/sub protocol used by services and many IoT devices like sensors, they can connect locally or through internet (e.g. via TheThingsNetwork if they are LoRaWAN enabled) [Ba19]

These dependencies illustrate some of the initially implemented components and can be extended at any time with additional services. For interoperability and requirement 3, GIS components provide standardised service interfaces as defined by the open geospatial consortium (OGC). These services provide either rendered map data and raw image data via map service or raw vector data via feature service, recently unified as OGC API [He23].

The setup and management of these software components on edge devices are organised as an IaC approach addressing requirement 2, where the system and software configurations are available as code. This allows replicated and customised installations on an unlimited number of devices. The current hardware platform for deployments are small sized barebone-PCs (Intel NUC, generation 11), addressing requirement 5. The booted system on these devices is Proxmox VE (Virtual Environment), a professional and open sourced virtualisation platform. This platform is the basis for the virtual systems implementing the microservices. The configuration can be automated: Ansible is used as tool to define virtual machines and their individual configuration. Most virtual systems are implemented as LXC containers, a widely adopted method that uses control group and name space features included in the hosts Linux kernel and therefore avoids the overhead for full virtualisation [SST23]. The latter may only be necessary for certain systems, for example if a virtual machine with Windows OS is to be installed that cannot use shared kernel resources. The default LXC container OS is Debian, a widely used Linux distribution for servers. The minimal Debian system needs to be extended to enable a Docker environment, since most microservices are defined and implemented as Docker containers. Docker containers themselves define software setups as code. The development systems rely on publicly available Docker infrastructure and images, for production this can be changed to self-hosted infrastructure to precisely control the shipped software (with custom Dockerfiles and a Docker hub, “Hofbox container registry” in Fig. 2).

The networking needs special attention, since each Hofbox is considered to be located in local home networks and not directly reachable from the internet. Then firewalls keep ports closed and often network address translation (NAT) takes place. Only outgoing connections can then be initiated, and reverse tunnel endpoints or VPN (virtual private network) solutions need to be implemented to address requirement 6. Since only the edge devices of partners should be connected, specialised VPN approaches like Tailscale [Pe20] or the open source implementation Headscale [DF23] are considered, allowing to configure and initiate peer-to-peer connections based on Wireguard, a modern and lightweight VPN protocol [Do20]. Additionally, DNS (domain name resolution) can be configured and services can be reached by their DNS name locally and if a VPN connection exists to address requirement 4.

Fig. 2 illustrates the setup and encapsulation based on virtualisation as well as network connections: The Hofbox is physically located in a local network together with other devices and provides access within that network. The microservices are configured with customised defaults, the dependencies are managed via configuration for database access URIs, credentials or geospatial API endpoints. In Fig. 2 the virtualisation is generalised and depicts two of the containers with more details: the leftmost container runs the Open Horizon framework and the Hofbox software components developed by the Geobox projects [Eb23]. The edge device management is implemented differently in the way the control servers are contacted. Open Horizon uses a client software that actively gets instructions on configuration, software and updates from the IBM Edge Application Manager (IEAM),

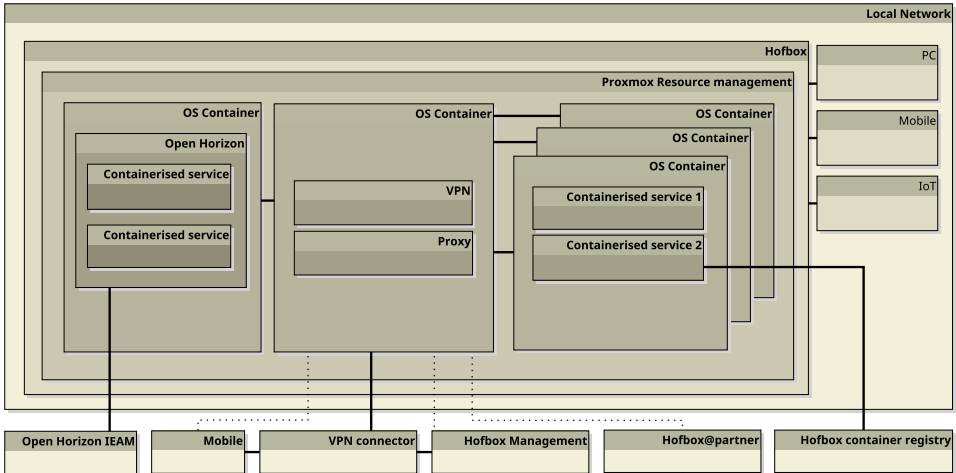


Fig. 2: Hofbox based on virtualisation and networking, dotted connectors denote possible peer-to-peer VPN connections.

whereas in our setup the “Hofbox Management” component actively connects to the Hofbox and applies changes without an installed client. In both cases containers are installed and updated from a corresponding container registry. The installation results are comparable and for the microservices it does not matter how exactly they get installed, their integration is possible via APIs. The container with the “VPN” and “Proxy” components in the centre of Fig. 2 is responsible for remote access and provides unified access to all services via a reverse proxy. All other microservices run parallel to these and communicate via the internal network. One of the virtualised systems provides a desktop environment with a preconfigured GIS software stack (see QGIS in Fig. 1) that users can connect to and operate remotely.

A detailed backup concept is under development, the Proxmox platform provides means for complete backups of the infrastructure and data, additionally database backups are needed that can easily be restored. Backups will be planned off-site depending on the availability of bandwidth.

3 Summary and outlook

The presented approach for the organisation and management of edge devices allows for a high degree of automation. System definition and configuration are available as code and support modern and open software development practices. The implementation with microservices provides defined interfaces and scalability by design. Cloud software setups on edge devices are resilient against internet failures. The same microservices can additionally be installed and interacted with in the cloud. Interactions are possible with any service,

no matter if running locally or in the cloud. Multiple geodata providers can be integrated: governments, companies or open data initiatives like OpenStreetMap provide map services and reusable data. The Hofbox provides a solid basis for geodata-driven processes and use cases on farms, supporting precision farming data and their life cycles.

Original data sources can be used directly wherever services provide suitable interfaces. For a resilient system, remote data sources should be cached locally - at least the necessary data for the operation area of the farm. Concepts for regular data updates exist, but they need to be automated. Additionally, synchronisation between cloud systems and Hofbox may be needed, if the leading system is a cloud system. For a fine-grained access control authentication and authorisation methods have to be developed and set up for the components, preferably with a single sign-on (SSO) approach.

The technical development requires constant man power, even if automation eases development and roll-out. Not only the system needs development and updates, a professional backing with tutorials and support is necessary before such systems can become a product on the market and get widely adopted. Use cases based on existing project data sets will show the applicability of the setup. Working use cases do not solve the problem of missing implementation of standards and interfaces in proprietary and third party software, but may help to foster adoption of open standards. Education and tutorial materials are finally needed to teach farmers specific tasks and the integration into their existing workflows. Due to the open development model with open source software the approach can be taken up and driven by an interested community.

References

- [Ba19] Banks, A.; Briggs, E.; Borgendale, K.; Gupta, R.: MQTT Version 5.0, OASIS Standard, OASIS, Mar. 2019, URL: <https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html>, visited on: 06/17/2023.
- [Be14] Bernardi, A.: Intelligente Wissenstechnologien für das öffentlich-private Wissensmanagement im Agrarbereich - iGreen, Schlussbericht, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI GmbH), 2014, URL: http://www.igreen-projekt.de/iGreen/fileadmin/Download/iGreen_Schlussbericht_Verbund_final.pdf, visited on: 02/17/2017.
- [Ce] Cepicky, J.: Switch from Shapefile, website, URL: <http://switchfromshapefile.org/>, visited on: 06/13/2023.
- [DF23] Dalby, K.; Font, J.: Headscale source code repository on github, 2023, URL: <https://github.com/juanfont/headscale>, visited on: 06/15/2023.
- [Do20] Donenfeld, J. A.: WireGuard: Next Generation Kernel Network Tunnel, tech. rep. draft revision e2da747, June 2020, URL: <https://www.wireguard.com/papers/wireguard.pdf>, visited on: 06/15/2023.





- [Eb23] Eberz-Eder, D.; Kuntke, F.; Brill, G.; Bernardi, A.; Wied, C.; Nuderscher, P.; Reuter, C.: Prototypische Entwicklungen zur Umsetzung des Resilient Smart Farming (RSF) mittels Edge Computing. In (Hoffmann, C.; Stein, A.; Ruckelshausen, A.; Müller, H.; Steckel, T.; Floto, H., eds.): Referate der 43. GIL-Jahrestagung: Resiliente Agri-Food-Systeme. Vol. P-330, Gesellschaft für Informatik e.V. (GI), Osnabrück, Germany, pp. 309–314, 2023, ISBN: 978-3-88579-724-1, URL: <https://dl.gi.de/handle/20.500.12116/40264>, visited on: 06/15/2023.
- [fa23] farmOS contributors: farmOS: a web-based farm record keeping application, ed. by Stenta, M., version 2.1.1, 2023, URL: <https://farmos.org/>, visited on: 06/15/2023.
- [FG20] Fabiani, A.; Giannecchini, S.: GeoNode, il CMS geospaziale Open Source, alessio.fabiani@geo-solutions.it, Mar. 2020, URL: <https://doi.org/10.5281/zenodo.3723241>.
- [GG23] Gabriel, A.; Gandorfer, M.: Adoption of digital technologies in agriculture—an inventory in a european small-scale farming region. *Precision Agriculture* 24/1, pp. 68–91, Feb. 2023, ISSN: 1573-1618, URL: <https://doi.org/10.1007/s11119-022-09931-1>, visited on: 06/17/2023.
- [HAF17] Hewage, P.; Anderson, M.; Fang, H.: An Agile Farm Management Information System Framework for Precision Agriculture. In: *Proceedings of the 9th International Conference on Information Management and Engineering. ICIME 2017*, Association for Computing Machinery, Barcelona, Spain, pp. 75–80, Oct. 2017, ISBN: 978-1-4503-5337-3, URL: <https://dl.acm.org/doi/10.1145/3149572.3149583>, visited on: 06/15/2023.
- [He23] OGC API - Common - Part 1: Core, tech. rep., version 1.0.0, Open Geospatial Consortium, Mar. 2023, URL: <http://www.opengis.net/doc/is/ogcapi-common-1/1.0>, visited on: 06/17/2023.
- [HO21] Hsu, L.S.; Obe, R.O.: *PostGIS in Action*. Manning, 2021, ISBN: 9781617296697.
- [IS15] ISO: ISO 11783-10:2015 Tractors and machinery for agriculture and forestry — Serial control and communications data network — Part 10: Task controller and management information system data interchange. Beuth Verlag, Berlin, 2015.
- [MEA18] Martínez-Casasnovas, J. A.; Escolà, A.; Arnó, J.: Use of Farmer Knowledge in the Delineation of Potential Management Zones in Precision Agriculture: A Case Study in Maize (*Zea mays* L.) *Agriculture* 8/6, p. 84, 2018, ISSN: 2077-0472, URL: <http://www.mdpi.com/2077-0472/8/6/84>, visited on: 06/25/2018.

- [Mi22] Mittermayer, M.; Maidl, F.-X.; Nätscher, L.; Hülsbergen, K.-J.: Analysis of site-specific N balances in heterogeneous croplands using digital methods. *European Journal of Agronomy* 133/, p. 126442, 2022, ISSN: 1161-0301, URL: <https://www.sciencedirect.com/science/article/pii/S1161030121002136>, visited on: 04/26/2022.
- [Ne23] Nextcloud GmbH: Nextcloud, website, 2023, URL: <https://nextcloud.com/>, visited on: 06/17/2023.
- [NG23] NGINX contributors: NGINX, 2023, URL: <https://www.nginx.com/>, visited on: 06/15/2023.
- [Pe20] Pennarun, A.: How Tailscale works, blog article, Mar. 2020, URL: <https://tailscale.com/blog/how-tailscale-works/>, visited on: 06/15/2023.
- [QG22] QGIS Development Team: QGIS Geographic Information System, QGIS Association, 2022, URL: <https://www.qgis.org>, visited on: 10/11/2022.
- [RSE19] Reuter, C.; Schneider, W.; Eberz, D.: Resilient Smart Farming (RSF) – Nutzung digitaler Technologien in krisensicherer Infrastruktur. In: 39. GIL-Jahrestagung, Digitalisierung für landwirtschaftliche Betriebe in kleinstrukturierten Regionen - ein Widerspruch in sich? Gesellschaft für Informatik e.V., Bonn, pp. 177–182, 2019, ISBN: 978-3-88579-681-7, URL: <https://dl.gi.de/bitstreams/b487cbad-c1ac-4609-8745-46d2edf540cd/download>, visited on: 07/01/2023.
- [Sa23] San Emeterio de la Parte, M.; Lana Serrano, S.; Muriel Elduayen, M.; Martínez-Ortega, J.-F.: Spatio-Temporal Semantic Data Model for Precision Agriculture IoT Networks. *Agriculture* 13/2, 2023, ISSN: 2077-0472, URL: <https://www.mdpi.com/2077-0472/13/2/360>, visited on: 06/17/2023.
- [Sc22] Schlingmann, N.: Chapter 3 - ISOBUS – standards and uses for data from farm machinery. In (Mottram, T., ed.): *Digital Agritechnology*. Academic Press, pp. 49–67, 2022, ISBN: 978-0-12-817634-4, URL: <https://www.sciencedirect.com/science/article/pii/B978012817634400001X>.
- [SST23] Satra, R.; Syafie, L.; Tubagus, M.: Comparison of server technologies using Kernel-based virtual machine and container virtualization. *AIP Conference Proceedings* 2595/1, 070003, May 2023, ISSN: 0094-243X, eprint: https://pubs.aip.org/aip/acp/article-pdf/doi/10.1063/5.0123841/17570422/070003_1_5.0123841.pdf, URL: <https://doi.org/10.1063/5.0123841>, visited on: 06/17/2023.
- [Tu23] Turton, I.: geoserver/geoserver: Release notes - GeoServer - 2.23.1. Version 2.23.1/, May 2023, URL: <https://doi.org/10.5281/zenodo.7986420>, visited on: 06/15/2023.

- [WBM22] Weis, M.; Bökle, S.; Möller, K.: Die digitale Hofbox: ein Konzept für die Datenhaltung auf dem Hof mit freier Software. In (VDLUFA, ed.): *Sensorsysteme in der Landwirtschaft - Chancen und Herausforderungen*. Vol. 2022. VDLUFA-Schriftenreihe 78, VDLUFA, VDLUFA-Verlag, Darmstadt, pp. 301–307, Sept. 2022, ISBN: 978-3-941273-34-4, URL: <https://www.vdlufa.de/wp-content/uploads/2023/01/Kongressband-2022-Low-res.pdf>, visited on: 02/01/2023.
- [WCS22] Wee, B. S.; Chin, C. S.; Sharma, A.: Artificial Intelligence of Things Enabled Fungiculture in Shipping Container. In: *2022 IEEE/ACIS 23rd International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*. Pp. 115–119, Dec. 2022, URL: <https://ieeexplore.ieee.org/abstract/document/10051780>, visited on: 06/17/2023.

Das Netz hat Geschichte:

Historisch-technische Analyse der kritischen Infrastrukturen in der Region Rhein/Main


Jonas Franken ¹, Marco Zivkovic², Nadja Thiessen ³, Jens Ivo Engels ⁴ und Christian Reuter ⁵

Abstract: Kritische Infrastrukturen sind häufig über Jahrzehnte gewachsene, komplexe Netze. Dennoch fehlt derzeit die historische Perspektive auf die Aufschichtungstendenzen von Technologien in den Sektoren, die für die Gesellschaft essenzielle Dienste bereitstellen. Ein besseres Verständnis von Ausbreitungs-, Ausbau-, Ersatz- und Ausmusterungsprozessen kann Entscheidungshilfe und Orientierung für resilientere Versorgungsnetzarchitekturen in der Zukunft geben. Kompatibilitätsprobleme mit Legacy-Soft- und Hardware sind bekannte Phänomene in vielen KRITIS-Einrichtungen. Entsprechend gewinnen Wissens- und Erfahrungstransfers bei zunehmend komplexen, dennoch über Jahrzehnte verwendete Technologien in landwirtschaftlichen Betrieben enorm an Bedeutung. Der Beitrag vollzieht die Konzeption und Fragestellungen eines interdisziplinären Forschungsprojekts nach, in welchem die Verwundbarkeit der kritischen Infrastrukturektoren Verkehr und Kommunikation im Rhein-Main-Gebiet analysiert wird. Von den Leistungen beider Sektoren hängt die digitale Landwirtschaft stark ab. Insbesondere rurale, beim digitalen und Schienennetzausbau häufig vernachlässigte Gebiete werden dabei mittels explorativer Interviewstudie und anschließender archivbasierter, quantitativer Überprüfung der zuvor generierten Hypothesen aus einer raum-zeitlichen und technischen Perspektive untersucht.

Keywords: Infrastrukturelle Abhängigkeiten, Wissenstransfer in KRITIS, Vulnerabilitätsanalyse

1 Einleitung

Wohl kaum eine Infrastruktur gilt moderner als IT und elektronische Kommunikationssysteme [La18]. Dennoch hat Deutschland den Anschluss an die digitale Moderne wegen

¹ Technische Universität Darmstadt, Wissenschaft und Technik für Frieden und Sicherheit (PEASEC), Pankratiusstraße 2, 64289 Darmstadt, franken@peasec.tu-darmstadt.de,  <https://orcid.org/0000-0003-0650-0308>


² Technische Universität Darmstadt, Neuere und Neueste Geschichte, Residenzschloss 1, 64283 Darmstadt, zivkovic@pg.tu-darmstadt.de

³ Technische Universität Darmstadt, Neuere und Neueste Geschichte, Residenzschloss 1,

64283 Darmstadt, nadja.thiessen@tu-darmstadt.de,  <https://orcid.org/0000-0002-9916-7112>

⁴ Technische Universität Darmstadt, Neuere und Neueste Geschichte, Residenzschloss 1,

64283 Darmstadt, jens_ivo_engels@tu-darmstadt.de,  <https://orcid.org/0000-0002-8680-1882>

⁵ Technische Universität Darmstadt, Wissenschaft und Technik für Frieden und Sicherheit (PEASEC), Pankratiusstraße 2, 64289 Darmstadt, reuter@peasec.tu-darmstadt.de,  <https://orcid.org/0000-0003-1920-038X>

des lückenhaften Ausbaus der Digitalisierung in der Fläche bislang nicht vollendet. Die Nutzung von Faxgeräten in Gesundheitsämtern in der COVID-19 Krise ist nur eines von vielen öffentlich skandalisierten Beispielen, in denen die Verheißungen zukunftsgestaltender Digitalität den Einschränkungen durch ‚alte Technik‘ gegenübergestellt wurden [KG20]. Bei genauerem Hinsehen erweist sich dieses Bild vereinfacht, denn auch digitale Kommunikationsmittel haben eine Geschichte. Mit der Betrachtung dieser Geschichte, so unsere Annahme, lassen sich bestehende Probleme, insbesondere in denjenigen Gebieten, die schwächerer angebunden sind, besser verstehen und Optimierungen zielgerichteter umsetzen.

Die Ursprünge der Digitalisierung gehen in Deutschland bis auf die 1960er Jahre zurück. Vor allem in ländlichen Regionen basieren Internetanschlüsse bis heute auf historischen Telefonkabeln aus Kupfer [La18]. Auch Softwares können aufgrund der ständigen Updates als Schichtmodelle verstanden werden, in denen laufend neue Komponenten implementiert werden [SB06]. In ihren Grundstrukturen reichen sie oft Jahrzehnte zurück, wie etwa in der Software der deutschen Bankhäuser die Programmiersprache COBOL seit den 1980er Jahren [We19]. Allerdings wirken die Auswirkungen historischer Strukturen noch weiter in die Vergangenheit zurück und überschreiten die Grenzen zwischen Infrastruktursektoren. Die Historizität von Infrastrukturen hat enorme Auswirkungen auf ihre aktuelle Gestalt und auf ihre künftige Entwicklung. Die Vulnerabilität von Infrastrukturen ist entsprechend stark von historischen Entwicklungen geprägt. Immer dort, wo technische Komponenten unterschiedlichen Alters miteinander verbunden sind, bestehen technische Risiken durch fehlerhafte Konfiguration und die Notwendigkeit von Kompatibilität. Die räumliche Ausdehnung der Netzwerke aufgrund historischer Strukturen reproduziert außerdem potenziell ungleiche Vulnerabilitätsniveaus, welche bei Ausfällen, die auch durch bewusste Angriffe entstehen können, zutage treten. In Kontexten, in denen Anbindung ans und Partizipation im Internet zur Norm wird, sind Gruppen bereits dann benachteiligt, wenn sie höheren Ausfallrisiken ihrer Kommunikation ausgesetzt sind [St06]. Wird beispielsweise Konnektivität fundamental für die smarte, digitalisierte Landwirtschaft zunehmend vorausgesetzt, müssen die oftmals gering versorgten Regionen besser angebunden sein, um ihre Produktion aufrechtzuerhalten. Entscheidungen in Investitions-, Planungs- und Ausbauphasen von Infrastrukturen sollten daher mit historischem Wissen und auch im Hinblick auf Exklusions- und Ausfallszenarien getroffen werden.

Ausgangspunkt unserer Überlegungen ist die Annahme, dass Betreiberfirmen und technische Praktiker:innen über umfassendes, aber *implizites* Wissen über das Alter ihrer Systeme und die damit verbundenen Vulnerabilitäten verfügen. Damit ist Folgendes gemeint: Praktiker:innen wissen oft sehr genau, mit welchen technischen Komponenten sie es in ihren Systemen zu tun haben, welche technischen Standards verbaut sind, welche Formen von Schäden wann zu erwarten sind. Was jedoch in der Regel fehlt, ist 1. ein modellhaftes Verständnis für die über Zeit und Raum hinweg zustande gekommene Komplexität des Ist-Zustandes, 2. ein Verständnis dafür, aufgrund welcher sozialen, politischen, technologischen Kontexte die in der Vergangenheit getroffenen Entscheidungen zustande gekommen sind, wozu 3. insbesondere sektorenübergreifende Zusammenhänge gehören (hier: zwischen IKT-, Verkehrsinfrastruktur und der Landwirtschaft). Mit anderen Worten: Das

vorhandene Wissen wird nicht explizit als historisches reflektiert und sein Transfer nicht koordiniert angeleitet.

Im Kern wollen wir damit einerseits das implizite Wissen um historisch bedingte Vulnerabilität bei Infrastruktur-Praktiker:innen untersuchen. Andererseits stellen wir diesem die Befunde empirisch-quantitativer Analyse entgegen: Wir analysieren die ‚tatsächliche‘ Gestalt der Infrastrukturen hinsichtlich ihrer historisch-räumlichen Genese und wenden dabei Methoden der Netzwerkanalyse und der historischen Infrastrukturforschung an. So können Wege aufgezeigt werden, Wissen über historische Zusammenhänge, Pfadabhängigkeiten und Strukturen für die künftige Entwicklung resilienter Infrastrukturen nutzbar zu machen. Wir fokussieren uns auf die Rhein-Main-Region, da hier eine hohe Dichte von technischen Infrastrukturen und sowohl urbane Zentren (Frankfurt am Main, Darmstadt) als auch große ländliche Räume (Odenwald, Taunus, Wetterau) gegeben sind. Aufgrund der Vielseitigkeit der Untersuchungsregion erhoffen wir uns generalisierbare Ergebnisse für den gesamtdeutschen Kontext, wo ebenfalls groß Stadt-Land-Kluffen prägend für den Infrastrukturausbau sind.

Zudem konzentrieren wir uns auf die beiden Sektoren Kommunikation und Verkehr als wichtige Upstream-Sektoren [Ku22]. Das initiale Interesse gilt der IKT-Infrastruktur. Wir gehen davon aus, dass die historischen Strukturen der aktuellen Kommunikationsnetze nicht ohne Betrachtung der - meist älteren - Verkehrsnetze verständlich sind. Insbesondere die klassischen Verkehrsinfrastrukturen beeinflussen mit ihrem Trassenverlauf den Verlauf von Kommunikationsinfrastrukturen. Internetkabel sind bspw. oft nahe Straßen und Gleisen verlegt [LKZ16, Un18]. Ländliche Regionen mit einem losen Schienennetz tendieren entsprechend dazu, weniger redundant an das Internet angebunden zu sein.

Unser interdisziplinär geprägtes Vorgehen kombiniert technik- und geschichtswissenschaftliches Wissen und Erfahrung, sowie dazugehörige Theorien und Methoden. Im Forschungsprozess bieten sich viele Möglichkeiten, die Grenzen der Disziplinen und Felder zu überwinden und daraus Lehren für Kollaboration in der technischen Friedens- und Sicherheitsforschung abzuleiten. In einer Metabetrachtung der Digital-Divide-Forschung stellte van Dijk [Di18] kürzlich fest: „*The theory of the digital divide needs to be multi-disciplinary and preferably interdisciplinary*“. Entsprechend setzt das Vorhaben seine Interdisziplinarität nicht als Option, sondern als neuen Standard für hochwertige Forschung im Themenfeld der IKT-Infrastrukturen voraus. Mit dem Fokus auf Vulnerabilität von KRITIS, insbesondere in Ausfallszenarien, knüpfen wir auch inhaltlich an Betrachtungen benachteiligter, eher landwirtschaftlich geprägter Regionen [RSE19]. Denn Konflikte, Kriege und Naturkatastrophen bilden einen Kontext von großflächigen KRITIS-Ausfällen und daraus folgender Verschlechterung der Bereitstellung essenzieller Güter (insb. Nahrungsmittel), wie der Krieg in der Ukraine und die Starkregenereignisse im Westen Deutschlands 2021 jüngst verdeutlichten.

2 Forschungsstand

Trotz einer in den letzten Jahren boomenden historischen Infrastrukturforschung gibt es so gut wie keine Versuche, geschichtswissenschaftliche Perspektiven für die aktuelle Infrastrukturentwicklung nutzbar zu machen. Uns sind insbesondere keine Studien bekannt, die das oben skizzierte implizite historische Wissen von Betreiber:innen überhaupt exploriert hätten. Es gibt zwar seit rund 20 Jahren ethnologische Arbeiten über Wissen und Praktiken von Infrastrukturbetreiber:innen, beispielsweise in Kontrollräumen und Lagezentren [LM16], doch geht es hier nicht um historische Aspekte. Am nächsten kommen unserem Ansatz die Studien von Timothy Moss, der aber im Wesentlichen archivbasiert arbeitet [Mo20]. Gleichwohl gibt es eine Vielzahl an Forschungssträngen, auf denen unsere Forschung aufbaut bzw. an die sie anknüpft:

Bereits in den 1960er Jahren wurde festgestellt, dass der Zugang zu Kommunikation ein zentraler Faktor für die Entwicklungsmöglichkeiten von Individuen und Gruppen ist [Ke67]. Mit der Proliferation des Internets Anfang der 1990er Jahre und der zunehmenden Digitalisierung des Alltags (Privatkommunikation, Handel, Produktion, eGovernment, etc.) wurde die Theorie des „*Digital Divides*“ als Rahmen entworfen, um soziale Ursachen und Folgen der Digitalisierung systematisch zu analysieren [Wa01]. Deren Erforschung gliedert van Dijk in drei Phasen [Di18]: In der ersten Phase wurden allein die physischen Anbindungen an Infrastrukturen, die Konnektivität ermöglichen, und die daraus erwachsenen Möglichkeiten analysiert. Zugang wurde in dieser ursprünglichen Phase als binäre abhängige Variable wahrgenommen, differenzierte in der zweiten Phase zunehmend aus, indem um soziale und demographische Faktoren, die die Nutzung von Onlinetechnologien bedingen, unter dem multikausalen Konzept der „*Internet literacy*“ einbezogen wurden [RD05]. In der dritten Phase erreichte die Forschung jüngst die Reife, die die Anwendung der Digitalen Kluft als unabhängige Co-Variable für die Analyse anderer Phänomene ermöglichte, beispielsweise die Wirtschaftskraft eines Staates oder die Spaltung einer Gesellschaft [Di18].

Der von uns verfolgte Fokus auf die (daten-)transportierenden Infrastrukturen knüpft folglich primär an die erste Phase der Digital-Divide-Forschung an. Ungleichheiten der Resilienz besitzen weiterhin eine hohe Bedeutung, was jüngst Studien zu Internet-Infrastrukturen nachweisen [Fr22, Th19]. Die Forschungslandschaft der Digitalen Kluft zeichnen sich dabei durch eine hohe Variation der Analyseebenen aus. Während der häufigste Fall nationale Einzelfallstudien sind [Kh12, NIS17, RH17, ST16], existieren ebenfalls einige makroregionale [Ca19, FH08, MKH21, SKS21, Sz18] und globale [ACL10, CW04, COB18, Fr22, RC10] Analysen. Gleichwohl tritt neben die räumliche Betrachtung von Infrastrukturen die historische Komponente, wodurch Effekte technischer Aufschichtung der Infrastrukturen abeleitet werden können. Zwar existieren ebenfalls einige Werke zur zeitlichen Entwicklung der IKT-Infrastrukturen [Ki96, Ma02, Sa13]. Diese gehen wiederum selten auf die räumlichen Aspekte der Ausbreitung und den dahinterliegenden Motivationen und Pfadabhängigkeiten von Entscheidungsträger:innen ein. Wir wollen damit die Forschungslücke der historisch-technischen Betrachtungen zweier KRITIS, deren

Verhältnis sich im Laufe durch einen Wechsel von gegenseitiger Abhängigkeit zur vermeintlichen technischen Entkopplung ausgezeichnet hat, adressieren. Als Motivation fungieren die heutigen Interaktionen von KRITIS-Sektoren sowie die gegenseitige Gefährdung der Strukturen aber auch die Chancen der räumlichen Nähe. Somit kann die Analyse zur Debatte der abnehmenden Wichtigkeit von Distanzen beitragen, die bereits 1996 von Couclelis [Co96: p. 388] mit Blick auf das Internet zusammengefasst wurde: „*Distance is not dead, but its significance and effects on the geographies of the information society have become a lot more complex and subtle.*“

3 Konzeption

Demnach stellt sich die Frage, wie eine Analyse vorgenommen werden kann, die historisches, meist implizites Wissen mit zeitlich-räumlicher Ausbau- und Aufschichtungstendenzen kombiniert. Die beiden Infrastrukturen Verkehr und IKT sind als komplexe Netze entworfen, die auf mehreren Ebenen und mittels mehrerer parallel existierenden Strukturen die Güter und Personen bzw. Daten transportieren. Die offensichtlichen Unterschiede im Hinblick auf Transportgeschwindigkeit, Anzahl paralleler Netzstrukturen und deren alltäglicher Sichtbarkeit können im gemeinsamen Analyserahmen der interdisziplinären KRITIS-Forschung überwunden werden. Um die Historizität technischer Infrastruktur konzeptionell zu fassen, gehen wir vom Modell der Zeitschichten aus, worin sich im Verlauf der Jahrzehnte oder gar Jahrhunderte in den Infrastruktursystemen immer neue Komponenten ‚anreichern‘ und sedimentieren. Das Modell der Zeitschichten wurde für die historische Forschung erstmals von Koselleck entworfen [Ko00]. Es ist in der historischen Technikforschung aufgegriffen und weiterentwickelt worden [We19, En20]. Die Zeitschichten kann man in der Gegenwart identifizieren und nach ihrer jeweiligen zeitlichen Herkunft bestimmen. Im Ergebnis entsteht das Bild eines heterogenen, aber integrierten Systems. Wie bereits dargelegt interessieren uns vor allem heterogene technische Komponenten und gewachsene Trassenstrukturen. Die Vereinigung von historischen und technischen Methoden ermöglicht eine abstrahierte Betrachtung der Sektoren und ihrer Versorgungsrollen in Vergangenheit, Gegenwart und Zukunft. Bereits bestehende methodische Überschneidungen können dabei einen gegenseitigen Mehrwert schaffen.

4 Fragestellungen und Diskussion des Vorgehens

Aus den obigen Vorüberlegungen haben wir folgende Fragen für den ersten Teil empirischen Teil herausgearbeitet:

- Welches Wissen über das Alter einzelner Komponenten im jeweiligen Infrastrukturnetz ist vorhanden? Wird dieses Wissen in Verbindung mit Vulnerabilitätsfaktoren gebracht?

- Gibt es Strategien im Umgang mit zeitlich heterogenen Komponenten? Wie hoch wird die räumliche Nähe von IKT- und Verkehrsinfrastrukturen eingeschätzt und welche Gefahren und Chancen können aus ihr entstehen?
- Wird historisches Wissen in Planungen von Infrastrukturen einbezogen?
- Gibt es Pfadabhängigkeiten bei Entscheidungen und wie können diese aufgebrochen werden?
- Welche Rolle spielt die Geschichte für die Gestaltung der Zukunft?

Zur Beantwortung dieser Unterfragen dient eine Interviewstudie mit Betreiber:innen, Eigentümer:innen, Planer:innen, Auftraggeber:innen, Regulator:innen und weiterer Expert:innen beider Infrastruktursektoren als erster empirische Schritt. Beim selektiven Sampling wird auf eine Ausgewogenheit der mindestens 20 Interviews im Hinblick auf die beiden Sektoren, die Betriebsgrößen (mittelständische und große Unternehmen), Planungs- und Betriebsebene, sowie die ausgeglichene Verortung in urbanen Ballungszentren bzw. ruralen Räumen als Referenzen geachtet. Für das Sample wird zunächst mittels Onlinerecherche ein Pool potenzieller Interviewees gebildet, die den o.g. Profilen entsprechen. An Personen dieses Pools werden in mehrstufigen Verfahren Interviewanfragen geschickt, um Unausgewogenheit der Antwortraten vorheriger Akquiserunden auszugleichen. Diese Kaltakquise wird durch eine Schneeballmethodik mittels einer Anfrage um weitere Kontaktvermittlung am Ende des Leitfadens ergänzt, um auszuschließen, dass „Schlüsselakteure“ [Dö21] der Sektoren übersehen werden. Da in Teilbereichen beider Infrastruktursektoren teils große Akteure (z.B. DB, Telekom) dominieren können, wollen wir nicht ausschließen, mehrere Personen derselben (Dach-)Organisation zu interviewen. Wir orientieren uns explizit nicht an den zugegebenermaßen willkürlichen Grenzen der KRITIS-Verordnungen, um auch kleinere in den Sektoren tätige Akteure in die Untersuchung mitaufnehmen zu können [ST22].

Methodisch erfolgt auch hier eine enge Kombination beider Disziplinen. Die Teilfragen sollen durch offene, theoriegenerierende Interviews [BM09, Kü19, Po18] beantwortet werden. Von historischer Seite geht es hier um eine Art ‚Oral History‘ des impliziten Wissens um das, was wir konzeptionell als Zeitschichten auffassen. Auf technischer Seite wird der gegenwärtige Zustand und die Vulnerabilität beider KRITIS Sektoren erfragt. Durch die Auswahl von Expert:innen als Sample nehmen wir bewusst von der klassischen Oral History-Methodik ab, die üblicherweise einen expliziten Fokus auf Bevölkerungsteile legt, deren Geschichte aus strukturellen Gründen geringer dokumentiert wird. Da wir diese Gefahr bei Betreiberunternehmen niedriger einschätzen, werden semi-strukturierte Leitfadeninterviews gewählt, die weiterhin genügend Freiraum für explorative und problemorientierte Nachfragen ermöglichen, ohne zu zeitaufwändig zu sein [Dö21].

Die Ergebnisse der Interviewstudie können durchaus für den deutschen Kontext generalisiert werden; zum einen aufgrund der Diversität der Bevölkerungsdichte der Rhein-Main-Region und zum anderen, weil zumindest ein Teil der interviewten Akteur:innen auch national bzw. in weiteren Regionen tätig sind. Bisher durchgeführte Interviews sind in dieser Hinsicht bereits vielversprechend. Als Zwischenergebnis können aus der Interviewstudie Hypothesen abgeleitet werden, die im weiteren Verlauf quantitativ und qualitativ

überprüft werden. Alle folgenden Arbeitspakete bleiben demnach thematisch offen, um flexibel ggf. unerwartete Hypothesen diesem Schritt in die Analyse aufnehmen zu können.

Im nächsten Schritt sollen die Hypothesen per Regionalstudie mittels quantitativer Netzwerkanalysen überprüft werden. Dort kann geprüft werden, wie stark sich der Ausbau und die Dichte der IKT- und Verkehrsnetze ähneln und entwickelt hat. Die Betrachtung verschiedener Technologien, ihrer Proliferations- und ggf. Stilllegungsentwicklungen erfordert den Zugang zu Archivquellen für (zeit-)geschichtliche Daten, die eine temporale Netzwerkanalyse ermöglichen [Sa22], die die folgenden Fragen beantworten kann:

- Wie haben sich aufkommende IKT- und Verkehrsinfrastrukturen über Zeitschichten räumlich verbreitet?
- Wie stark korrelieren physische IKT- mit Verkehrsinfrastrukturen im Rhein-Main Gebiet miteinander?

Auf Grundlage einer archivbasierten Studie (Planungsunterlagen, Betriebsunterlagen, historische Karten) wird für ausgewählte Beispiele eine historisch fundierte Darstellung der Zeitschichten der jeweiligen Infrastruktur vorgenommen.⁶ Auf dieser Basis wiederum wird der Detailgrad und die Akkuratess des impliziten historischen Wissens aus den Interviews überprüft und bewertet. Entsprechend kann abgeschätzt werden, inwieweit das Wissen über Vulnerabilitäten umfassend oder lückenhaft ist. Die zweite Frage wird durch eine vergleichenden Netzwerkanalyse beantwortet, womit strukturelle Ähnlichkeiten und Unterschiede von Graphen, die in diesem Fall technische Netzinfrastrukturen modellieren, identifiziert werden [Ta19, Su18]. Während wir in der Interviewstudie offenlassen, welche konkreten Infrastrukturen und Technologien von den Teilnehmenden erwähnt werden, werden wir uns aus Ressourcengründen bei der Hypothesenprüfung mittels Netzwerkanalysen einerseits auf den Kabelnetzaufbau (Kupfer und Glasfaser) fokussieren, weil es die zentrale Infrastruktur für den Datenverkehr ist. Andererseits werden wir uns im Verkehrssektor auf das Schienennetz konzentrieren, da hier bereits historische Datensätze zumindest in Ansätzen vorhanden sind. Innerhalb dieser Infrastrukturen sehen wir bislang die Güte der Archivbasis als den einzigen limitierenden Faktor. Spätere Untersuchungen können allerdings auch nachträglich um das für die Landwirtschaft besonders wichtige Straßenverkehrsnetz ergänzt werden.

Durch die Kombination von Archivstudie und quantitativer Netzwerkanalyse wird ermittelt, in welcher historischen Phase sich die heute feststellbaren Strukturen ausbildeten und ggf. verfestigten bzw. dynamisierten, also welche historischen Ursachen z.B. für bestimmte Vulnerabilitäten festzustellen sind. Die beiden Netzwerkanalysedimensionen – temporal und vergleichend – sind Ansätze, die nach unserem Wissen bislang nicht

⁶ Grundlage für die historische Überprüfung, inkl. der historischen Netzwerkanalyse, sind Dokumente von Betreibern und Behörden, z.B. das Bundesarchiv Koblenz: BMV, Eisenbahnbundesamt, Bundesanstalt für Post und Telekommunikation sowie Fernmeldetechnisches Zentralamt, das Hessische Hauptstaatsarchiv: Bundesbahn- und Oberpostdirektion Frankfurt das Hessische Staatsarchiv Darmstadt: Straßenbauämter, das Institut für Stadtgeschichte Frankfurt: Verkehrswesen und Post sowie das Frankfurter Wirtschaftsarchiv: Sammlungen zum Verkehrsverbund und Informationstechnologie.

kombiniert wurden. Ein Ziel der quantitativen Empirie wird daher sein, die Machbarkeit einer kombinierten Netzwerkanalyse zu erörtern, durch die Vulnerabilität in unterschiedlichen Netzwerkstrukturen und Zeitschichten quantifiziert werden könnten. Da sich die Datenbasis der Netzstrukturen spezifisch auf die Rhein-Main Region bezieht, sind die Schlussfolgerungen nur bedingt auf andere Regionen bzw. national übertragbar und können allenfalls Hypothesen für weitere Untersuchungskontexte anbieten. Der methodische Ansatz hingegen kann bei Erfolg auch auf weitere zeit-räumliche Infrastrukturentwicklungen weiterer Kontexte und Sektoren übertragen werden, wovon die quantitative Infrastrukturforschung insgesamt einen Nutzen hat.

Die von beiden untersuchten Sektoren abhängige Landwirtschaft kann von den Ergebnissen unserer Untersuchung gleich mehrfach profitieren. Zum einen werden wir Methoden zur Identifikation vulnerabler Gebiete erarbeiten, wodurch die Priorisierung von ruralen Kontexten beim resilienzfördernden Infrastrukturausbau begründen kann [Re22]. Zum anderen ist durch die häufig lange Nutzung von Technologien die Landwirtschaft ein Feld, in dem Wissenstransfer zur Aufschichtung und Kompatibilität mit fortschreitender Digitalisierung des Sektors eine zunehmend zentrale Rolle spielen werden. Entsprechende Erkenntnisse aus anderen KRITIS-Sektoren haben sicherlich auch für die Agrarwirtschaft einen Mehrwert.

Danksagung

NetzGeschichte wird durch die IANUS-Förderlinie des Forums interdisziplinäre Forschung (FiF) an der Technischen Universität Darmstadt gefördert. Diese Arbeit wurde zudem teilweise durch die LOEWE Initiative des Landes Hessen im Rahmen des LOEWE-Zentrums emergencITY und aus Mitteln des Zweckvermögens des Bundes bei der Landwirtschaftlichen Rentenbank im Rahmen des Projekts AgriRegio gefördert.

Literaturverzeichnis




- [ACL10] Ayanso, A.; Cho, D. I.; Lertwachara, K.: The digital divide: global and regional ICT leaders and followers. *Information Technology for Development*. 16/10, S. 304–319, 2010.
- [BM09] Bogner, A.; Menz, W.: The Theory-Generating Expert Interview: Epistemological Interest, Forms of Knowledge, Interaction. In: (Bogner, A., Littig, B., Menz, W. Hrsg.): *Interviewing Experts*. Palgrave Macmillan, London,; S. 43–80, 2009. https://doi.org/10.1057/9780230244276_3.
- [Ca19] Cariolle, J.: *Telecommunication Submarine Cable Deployment and the Digital Divide in Sub-Saharan Africa (Revised Version)*., 2019. <https://doi.org/10.2139/ssrn.3202941>.
- [Co96] Couclelis, H.: The death of distance. *Environment and planning B: Planning and Design* 23(4)/96, S. 387-389, 1996.

- [COB18] Cruz-Jesus, F.; Oliveira, T.; Bacao, F.: The Global Digital Divide. *Journal of Global Information Management* 26/18, S. 1–26, 2018. <https://doi.org/10.4018/JGIM.2018040101>.
- [CW04] Chen, W.; Wellman, B.: The Global Digital Divide – Within and Between Countries. *IT & Society*. 1/04, S. 39–45, 2004.
- [Dö21] Döringer, S.: The problem-centred expert interview: Combining qualitative interviewing approaches for investigating implicit expert knowledge. *International Journal of Social Research Methodology* 24/21, S. 265–278, 2021. <https://doi.org/10.1080/13645579.2020.1766777>.
- [De20] Van Deursen, A.J.: Digital Inequality During a Pandemic: Quantitative Study of Differences in COVID-19-Related Internet Uses and Outcomes Among the General Population. *J Med Internet Res*. 22/20, S. 1–13, 2020. <https://doi.org/10.2196/20073>.
- [Di18] Van Dijk, J. Afterword: the state of digital divide theory. In (Ragnedda, M.; Muschert, G. Hrsg.): *Theorizing Digital Divides*. Routledge, United Kingdom, S. 199–206, 2018.
- [En20] Engels, J.I.: Infrastrukturen als Produkte und Produzenten von Zeit. *NTM Zeitschrift Für Geschichte Der Wissenschaften, Technik Und Medizin* 28/20 S. 69–90, 2020. <https://doi.org/10.1007/s00048-019-00234-7>.
- [FH08] Fuchs, C.; Horak, E.: Africa and the digital divide. *Telematics and Informatics* 25/08 S. 99–116, 2008. <https://doi.org/https://doi.org/10.1016/j.tele.2006.06.004>.
- [Fr22] Franken, J.; Reinhold, T.; Reichert, L.; Reuter, C.: The Digital Divide in State Vulnerability to Submarine Communications Cable Failure. *International Journal of Critical Infrastructure Protection (IJCIP)* 38/22, 2022. <https://doi.org/https://doi.org/10.1016/j.ijcip.2022.100522>.
- [Ke67] Keesing, D.B.: Outward-Looking Policies and Economic Development. *The Economic Journal* 77/67 S.303–320, 1967. <https://doi.org/10.2307/2229306>.
- [KG20] Kersting, N.; Graubner, D.: Die digitale Transformation der deutschen Verwaltung Analysen zu Marktversagen und Daseinsvorsorge in Zeiten der Covid-19-Pandemie. In (Roters, W., Gräf, H., Wollmann, H. Hrsg.): *Zukunft Denken Und Verantworten*. Springer Fachmedien Wiesbaden, Wiesbaden, S. 231–252, 2020. https://doi.org/10.1007/978-3-658-31703-4_16.
- [Kh12] Khan, G.F.; Moon, J.; Swar, B.; Zo, H.; Rho, J.J.: E-government service use intentions in Afghanistan: technology adoption and the digital divide in a war-torn country. *Information Development* 28/12, S. 281–299, 2012. <https://doi.org/10.1177/0266666912438879>.
- [Ki96] Kittler, F.: The history of communication media. *Ctheory* S. 7–30, 1996.
- [Ko00] Koselleck, K.: *Zeitschichten. Studien zur Historik*. Suhrkamp, Berlin, 2000.
- [Kü19] Küsters, I.: Narratives Interview. In: (Baur, N.; Blasius J. Hrsg.) *Handbuch Methoden Der Empirischen Sozialforschung*. Springer Fachmedien Wiesbaden, Wiesbaden S. 687–693, 2019. https://doi.org/10.1007/978-3-658-21308-4_45.
- [Ku22] Kuntke, F.; Linsner, S.; Steinbrink, E.; Franken, J.; Reuter, C.: Resilience in Agriculture: Communication and Energy Infrastructure Dependencies of German Farmers.

- International Journal of Disaster Risk Science (IJDRS) 13/22, S. 214–229, 2022. Doi:10.1007/s13753-022-00404-7.
- [La18] Van Laak, D.: *Alles im Fluss: Die Lebensadern unserer Gesellschaft – Geschichte und Zukunft der Infrastruktur*. S. Fischer Verlage, Frankfurt am Main, 2018.
- [LKZ16] Limbach, F.; Kuebel, H.; Zarnekow, R.: Improving rural broadband deployment with synergistic effects between multiple fixed infrastructures. *Australasian Journal of Information Systems* 20/16, S. 1-17, 2016. <https://doi.org/10.3127/ajis.v20i0.1191>.
- [LM16] Luque-Ayala, A.; Marvin, S.: The maintenance of urban circulation: An operational logic of infrastructural control. *Environment and Planning D: Society and Space* 34/16, S. 191-208, 2016.
- [Ma02] Malecki, E.J.: The economic geography of the Internet’s infrastructure. *Economic Geography* 78/22, S. 399–424, 2022.
- [Me99] Melosi, M. V.: *The Sanitary City: Urban Infrastructure in America from Colonial Times to the Present*. The Johns Hopkins University Press, Baltimore, 1999. <https://doi.org/http://dx.doi.org/10.2307/2651437>.
- [MKH21] Myovella, G.; Karacuka, M.; Haucap, J.: Determinants of digitalization and digital divide in Sub-Saharan African economies: A spatial Durbin analysis. *Telecommunications Policy* 45/21, 2021. <https://doi.org/10.1016/j.telpol.2021.102224>.
- [Mo16] T. Moss, T.: Discarded surrogates, modified traditions, welcome complements: The chequered careers of alternative technologies in Berlin’s infrastructure systems. *Social Studies Science* 46/16 S. 559–582, 2016. <https://doi.org/10.1177/0306312716657205>.
- [Mo20] Moss, T.: *Remaking Berlin: A History of the City through Infrastructure, 1920–2020*. MIT Press, 2020.
- [NIS17] Nishijima, M.; Ivanauskas, T.M.; Sarti, T.M.: Evolution and determinants of digital divide in Brazil (2005–2013). *Telecommunications Policy* 41/17, S. 12–24, 2017. <https://doi.org/10.1016/j.telpol.2016.10.004>.
- [Po18] Portelli, A.: Living Voices: The Oral History Interview as Dialogue and Experience. *Oral History Review* 45/18, S. 239–248, 2018. <https://doi.org/10.1093/ohr/ohy030>.
- [RC10] Robison, K.K.; Crenshaw, E.M.: Reevaluating the Global Digital Divide: Socio-demographic and conflict barriers to the Internet Revolution., *Sociological Inquiry* 80/10, S. 34–62, 2010. <https://doi.org/10.1111/j.1475-682X.2009.00315.x>.
- [RD05] Riggins, F.; Dewan, S.: The Digital Divide: Current and Future Research Directions. *Journal of the Association for Information Systems* 6/05, S. 298–337, 2005. <https://doi.org/10.17705/1jais.00074>.
- [Re22] Reuter, C.; Kuntke, F.; Trapp, M.; Wied, C.; Brill, G.; Müller, G.; Steinbrink, E.; Franken, J.; Eberz-Eder, D.; Schneider, W.: *AgriRegio: Infrastruktur zur Förderung von digitaler Resilienz und Klimaresilienz im ländlichen Raum am Beispiel der Pilotregion Nahe-Donnersberg*. 52. Jahrestagung der Gesellschaft für Informatik (Workshop-Beiträge), Lecture Notes in Informatics (LNI) Hamburg, Germany (2022).
- [RH17] Radoll, P.; Hunter, B.: *Dynamics of the Digital Divide*. Australian National University Canberra, 2017.

- [RSE19] Reuter, C.; Schneider, W.; Eberz, D.: Resilient Smart Farming (RSF) – Nutzung digitaler Technologien in krisensicherer Infrastruktur, 39. GIL-Jahrestagung, Lecture Notes in Informatics (LNI) Vienna, Austria, S. 178-183, 2019.
- [Sa13] Sandvig, C.: The Internet as Infrastructure. Oxford University Press 2013. <https://doi.org/10.1093/oxfordhb/9780199589074.013.0005>.
- [Sa22] Salama, M.; Ezzeldin, M.; El-Dakhkhni, W.; Tait, M.: Temporal networks: a review and opportunities for infrastructure simulation. *Sustain Resilient Infrastruct.* 7/22, S. 40–55, 2022. <https://doi.org/10.1080/23789689.2019.1708175>.
- [SB06] Star, S.L.; Bowker, G.C.: How to Infrastructure. In: Handbook of New Media: Social Shaping and Consequences of ICTs. SAGE Publications, Ltd, London, 151–162, 2006. <https://doi.org/10.4135/9781848608245.n12>.
- [SKS21] Saunavaara, J.; Kylli, R.; Salminen, M.: Telecommunication line infrastructure and the Arctic environment: past, present and future. *Polar Record* 57/21, S. 1 – 12, 2021.
- [ST16] Sujarwoto, S.; Tampubolon, G.: Spatial inequality and the Internet divide in Indonesia 2010–2012. *Telecomm Policy* 40/16, S. 602–616, 2016. <https://doi.org/10.1016/j.telpol.2015.08.008>.
- [ST22] Schläger, U.; Thode, J.C.: Rechtliche Grundlagen der Informationssicherheit. In: Handbuch Datenschutz und IT-Sicherheit. Erich Schmidt Verlag GmbH & Co. KG, Berlin, S. 541-569, 2022.
- [Su18] Sugiyama, M.; Ghisu, M.E.; Llinares-López, F.; Borgwardt, K.: graphkernels: R and Python packages for graph comparison. *Bioinformatics* 34/18, S. 530–532, 2018. <https://doi.org/10.1093/bioinformatics/btx602>.
- [Sz18] Szeles, M.R.: New insights from a multilevel approach to the regional digital divide in the European Union. *Telecomm Policy* 42/18, S. 452–463, 2018. <https://doi.org/10.1016/j.telpol.2018.03.007>.
- [Ta19] Tantardini, D.; Ieva, F.; Tajoli, L.; Piccardi, C.: Comparing methods for comparing networks. *Scientific Reports* 9/19, 17557, 2019. <https://doi.org/10.1038/s41598-019-53708-y>.
- [Th19] Thorat, D.: Colonial Topographies of Internet Infrastructure: The Sedimented and Linked Networks of the Telegraph and Submarine Fiber Optic Internet. *South Asian Review* 40/19, S. 252–267, 2019. <https://doi.org/10.1080/02759527.2019.1599563>.
- [Un18] UNESCAP, Co-Deployment of Fibre Optic Cables along Transport Infrastructure for SDGs, Including Cross Border, Bangkok, 2018.
- [Wa01] Warf, B.: Segueways into Cyberspace: Multiple Geographies of the Digital Divide. *Environment and Planning B Planning Design* 28/01, S. 3–19, 2001. <https://doi.org/10.1068/b2691>.
- [We19] Weber, H.: Zeitschichten des Technischen: Zum Momentum, Alter(n) und Verschwinden von Technik. In: (Heßler, M.; Weber, H. Hrsg): *Provokationen Der Technikgeschichte: Zum Reflexionszwang Historischer Forschung*, Ferdinand Schöningh, Paderborn S.. 107–150, 2019.

Optimierte Messenger-Applikation zur Notfallkommunikation via LoRaWAN-DTN

Denis Orlov ¹, Franz Kuntke ², Christian Reuter ²

Abstract: Die vorliegende Arbeit präsentiert die Entwicklung einer Messenger-App mit Schwerpunkt auf Benutzerfreundlichkeit, für die Nutzung mit einem bestehenden LoRaWAN-DTN-Backend. Die App ermöglicht den Austausch von Nachrichten mit anderen Personen über ein vorhandenes Kommunikationssystem auf LoRaWAN-Basis. Das grundlegende Softwaregerüst wurde mithilfe agiler Softwareentwicklungsmethoden als Progressive-Web-App entwickelt und iterativ verbessert. Das Ergebnis ist eine plattformübergreifende App für Desktop-PCs und Android-Smartphones. Die App bietet grundlegende Messenger-Funktionen wie Kontaktverwaltung, Chatverlauf-Speicher und Benachrichtigungen. Zusätzlich enthält die App erweiterte Funktionen wie einen leicht zugänglichen SOS-Button, um Notfallnachrichten schnell absetzen zu können. Ziel der Entwicklung war es, die Gebrauchstauglichkeit gegenüber einem ersten Prototyp zu verbessern. Die App soll effektive Kommunikation zwischen Helfern und Betroffenen ermöglichen, während und nach Krisenereignissen wie beispielsweise der europäischen Flutkatastrophe 2021. In folgenden Arbeiten soll das System unter Nutzung dieser App im Einsatz getestet werden.

Keywords: Notfall-Messenger, Usability, LoRaWAN Multi-Hop, Disruption Tolerant Networking (DTN), Bundle Protocol 7 (BP7)

1 Einleitung

Internetbasierte Kommunikation hat sich zu einem unverzichtbaren Bestandteil der modernen Gesellschaft entwickelt. Allerdings haben in der Vergangenheit Krisenereignisse wie die europäische Flutkatastrophe 2021 gezeigt, dass Kommunikationsinfrastrukturen auch verletzlich sind und deren großflächige Reparatur besonders in dünn besiedelten Regionen einige Zeit in Anspruch nehmen kann. Gerade während solcher Krisen, aber auch im Nachgang besteht jedoch ein Bedarf an Kommunikation zwischen Helfern und Betroffenen. Als ein Lösungsansatz wurde bereits in einer vorhergehenden Arbeit [KBR23] der Ansatz eines Internet-unabhängigen Kommunikationsnetzwerkes speziell für landwirtschaftliche Betriebe vorgestellt, das auf herkömmliche LoRaWAN-Gateways setzt und autark betrieben werden kann. Es wurde dadurch gezeigt, dass die Sterntopologie von LoRaWAN mit Hilfe eines Disruption

¹ Technische Universität Darmstadt (TUDa), Wissenschaft und Technik für Frieden und Sicherheit (PEASEC), Pankratiusstr. 2, 64298 Darmstadt, <https://orcid.org/0000-0003-4134-5630>

² Technische Universität Darmstadt (TUDa), Wissenschaft und Technik für Frieden und Sicherheit (PEASEC), Pankratiusstr. 2, 64298 Darmstadt, <name>@peasec.tu-darmstadt.de, <https://orcid.org/0000-0002-7656-5919>, <https://orcid.org/0000-0003-1920-038X>

Tolerant Network (DTN)-Ansatzes in ein Multi-Hop-fähiges Netzwerk umgewandelt werden kann. Die Gebrauchstauglichkeit des ersten Prototypen stand dabei nicht im Vordergrund und lässt entsprechend zu wünschen übrig.

In der vorliegenden Arbeit wird die Entwicklung einer neuen Messenger-App vorgestellt, mit einem Fokus auf hohe Gebrauchstauglichkeit. Die entstandene App ist in der Lage, sich mit dem bestehenden LoRaWAN-DTN-Backend zu verbinden und Nachrichten im Notfallkommunikationsnetz auszutauschen, durch Senden und Empfangen. Ziel war es, die grundlegenden Funktionen eines Messengers, wie z.B. eine Kontaktverwaltung, ein Chatverlauf-Speicher und Benachrichtigungen eingegangener Nachrichten umzusetzen und dabei eine hohe Zugänglichkeit durch eine einfache und ansprechende Oberfläche zu erhalten. Ebenso wurden Zusatzfunktionen eingebaut, wie ein einfach erreichbarer SOS-Button um schnell Notfallnachrichten absetzen zu können.

Im Folgenden werden zunächst Grundlagen über die Backend-Infrastruktur dargestellt, und anschließend auf das Konzept und Umsetzung der App eingegangen. Es folgt eine erste heuristische Usability-Analyse der Entwicklung nach den 10 Kriterien von Nielson.

2 Grundlagen: Notfallkommunikation via LoRaWAN-Backend

Im geplanten Einsatzszenario besteht ein Kommunikations-Backend aus den folgenden Komponenten (siehe Abbildung 1): Ein LoRaWAN-Gateway und ein verbundener (Mini-)Server, auf dem eine ChirpStack-Instanz mit der Erweiterung *Spatz* läuft. Das Interface von *Spatz* ist im lokalen Netzwerk beispielsweise per Wi-Fi erreichbar und ist entsprechend per Client Applikation auf einem Smartphone nutzbar. Das Backend ist für das Routing der Nachrichtenpakete zuständig und nutzt dabei einen Disruption Tolerant Networking (DTN) Ansatz. Hierbei werden Pakete zu logischen Einheiten gebündelt, und in einer *Store-Carry-Forward* Methode weitergeleitet. Die Routing-Algorithmen und Clients wurden unter Annahme entwickelt, dass Geräte auch zeitweise oder permanent nicht verfügbar sein können und müssen entsprechend mit solchen Situationen umgehen. Details zum Backend sind in einer Vorarbeit veröffentlicht [KBR23].

Für die Messenger-App ist das clientseitige Interface von *Spatz* selbst wichtig. Dieses ist per WebSockets erreichbar und nimmt Nachrichten entgegen. Die Nachrichtenpakete

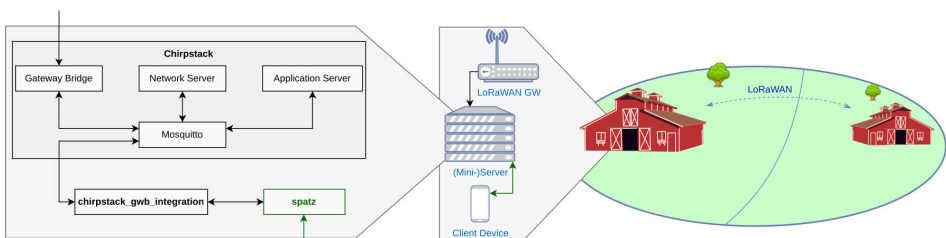


Abbildung 1: Schaubild der zugrundeliegenden Architektur (Quelle: [KBR23])

entsprechen dabei dem Standard Bundle Protocol 7 (BP7) [BFB22] und werden an das Backend in serialisierter Form übertragen, als Concise Binary Object Bundle (CBOR) [BH20]. Ebenso werden eingehende Nachrichten bei aktiver WebSocket-Verbindung an den Client in Form von CBOR-kodierten BP7 Paketen geschickt. Eine bestehende Referenzimplementierung für den Client basierend auf VueJS/JavaScript existiert bereits. Da das zugrundeliegende Konzept des Notfallkommunikationssystem auf DTN basiert, ist prinzipiell eine spätere Erweiterung denkbar, dass Smartphones um die Funktion von *Data Mules* erweitert werden. Dies würde ermöglichen, dass *Netzwerk-/Funklücken* über den Transport von Daten auf Smartphones manuell geschlossen werden können, insofern die Informationen eine hohe Latenz verkraften und entsprechend nicht zeitkritisch sind.

3 Konzept & Umsetzung

Das Konzept orientierte sich im Aufbau an bestehenden Messengern (siehe Abbildung 2) und wurde als Progressive Web App mit Ionic/Angular [WU18] umgesetzt, um mit einer Code-Basis³ die beiden Ziel-Plattformen Android und PC (Browser) abzudecken – das

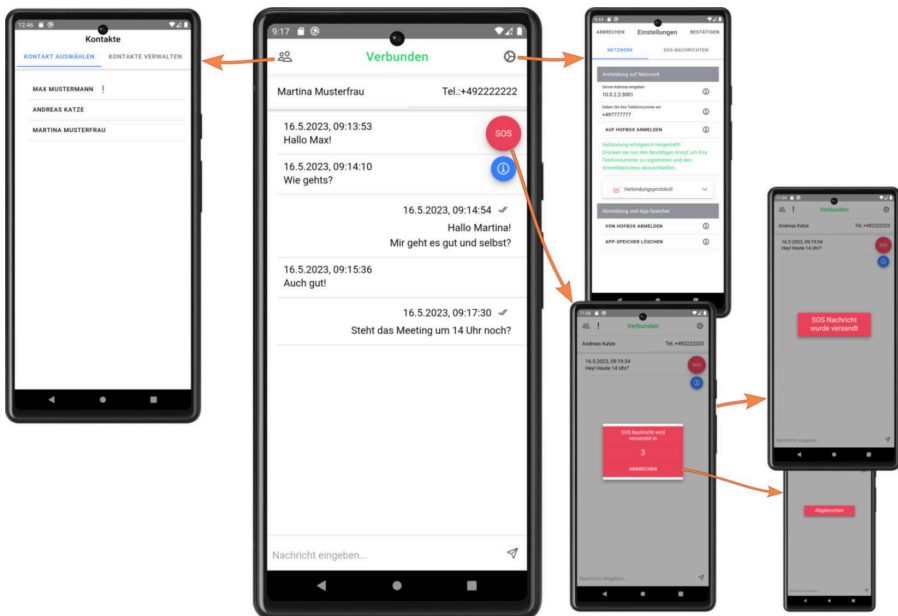


Abbildung 2: Die Oberfläche besteht aus 3 Bereichen: Eine Kontaktübersicht (links), eine Chatansicht (mittig) und eine Einstellungsansicht (oben rechts). In der Chatansicht wird ein SOS-Button angezeigt, der bei Betätigung nach Ablauf von 5 Sekunden (unten rechts) definierten Kontakten eine Notfallnachricht schickt.

³ <https://github.com/PEASEC/LoRaWAN-DTN-Messenger>

Framework verspricht hier später auch native Clients für weitere Plattformen wie Apple iOS und Desktop-Systeme (Apple macOS, Microsoft Windows und GNU/Linux) zu erstellen. Das zentrale Element stellt die Chatansicht dar, in dem der Nachrichtenverlauf mit einem einzelnen Kontakt chronologisch aufgelistet wird. Ein Eingabefeld ermöglicht die Eingabe und das Abschicken einer neuen Nachricht. Ein textueller Indikator am oberen Rand zeigt dabei stets den aktuellen Verbindungsstatus zum Backend an („Verbunden“, „Nicht verbunden“). Symbole an den Nachrichten selbst zeigen dabei den Status an, ob das Backend die jeweilige (eigene) Nachricht in Empfang genommen hat. Somit kann die Nutzer:in sehen, ob das Backend die Nachrichten erhalten hat. Eine dedizierte Bestätigung des Empfangs einer Nachricht beim Gegenüber ist aufgrund von System-Limitierungen im ersten Schritt nicht vorgesehen, könnte jedoch nachträglich hinzugefügt werden. Weiterhin ist in der Chat-Ansicht ein SOS-Button enthalten, der einen Countdown startet, nach dessen Ablauf eine vorgefertigte Nachricht inkl. (bei Verfügbarkeit) aktueller Geo-Koordinate (Längen-/Breitengrad) an einen hinterlegten Notfall-Kontakt versandt wird. Beim Empfang einer Nachricht mit Geo-Koordinate wird daraus automatisch ein Link generiert, der die Position in OpenStreetMap im Browser darstellt, was beim Empfänger Internetverbindung voraussetzt. Alternativ kann die Koordinate in ein offline-fähiges Kartenprogramm kopiert werden.

Zwei weitere Ansichten ermöglichen das Management von Kontakten, sowie die Konfiguration von Parameter, wie Adresse des Backend im lokalen Netzwerk und Text der SOS-Nachricht.

4 Heuristische Analyse der Gebrauchstauglichkeit

Eine erste Evaluation der Applikation ist mit Hilfe einer heuristischen Evaluation nach den 10 Design-Prinzipien von Nielsen [Ni93] erfolgt, um eine erste Einschätzung der Oberfläche und dessen Limitierungen zu erhalten. In Tabelle 1 sind die einzelnen Kriterien gelistet mit einer Zusammenfassung der Auswertung

Design-Prinzip	Zusammenfassung der Auswertung
1. Sichtbarer System-Status	Benutzer:in wird immer über notwendige, kontextspezifische Statusinformationen und dessen Änderungen informiert. Limitation des Systems: Benutzer:in wird nicht benachrichtigt, ob eine Nachricht das Ziel erreicht hat, oder gar gelesen wurde.
2. Übereinstimmung zwischen System und Realität	Der Begriff „Hofbox“ und das dahinterstehende Konzept ist der einzig sichere Fall von Fachjargon, der für nicht mit dem Projekt vertrauten Personen unklar ist.
3. Möglichkeit der freien Navigation	Das Versenden von Nachrichten, das Löschen des App-Speichers und das Entfernen eines Kontakts kann nicht widerrufen werden. Beim Löschen des App-Speichers und

Design-Prinzip	Zusammenfassung der Auswertung
	beim Entfernen eines Kontakts ist eine erneute Bestätigung erforderlich.
4. Konsistenz und Standardisierung	Die Beschreibung von Aktionen weist in der App eine hohe Konsistenz auf.
5. Fehlervermeidung	Mit dem Versenden von Nachrichten existiert eine Aktion die unwiderrufliche Konsequenzen hat, jedoch nicht von der Benutzer:in erneut bestätigt werden muss. Ansonsten werden mögliche fehlerhafte Eingaben abgefangen und Aktionen die zu Fehlern führen könnten, entsprechend verhindert und die Benutzer:in darüber in Kenntnis gesetzt.
6. Wiedererkennung geht vor Erinnerung	Im Chat ist der ausgewählte Kontakt direkt sichtbar. Darüber hinaus ist keine Erinnerung notwendig.
7. Flexibilität und Effizienz	Personalisierung der Benutzeroberfläche und die Abkürzung von Aktionen sind bisher nicht vorgesehen. Einzige Effizienzmaßnahme: Die Web-Oberfläche erlaubt es mit Strg+Enter eine Nachricht zu verschicken, um nicht den Senden-Knopf per Mausclick betätigen zu müssen.
8. Ästhetisches und minimalistisches Design	Die App besitzt ein minimalistisches Design. Eine logische Unstimmigkeit besteht in der Einstellungsansicht: Der „App-Speicher löschen“-Button wird im Netzwerksegment der Einstellungs-Ansicht angeführt, beeinflusst jedoch mehr als die Netzwerk-Einstellungen.
9. Information und Korrekturangebote bei Fehlern	Für fehlerhafte Eingaben wird jeweils der Fehler als auch ein Lösungsvorschlag präsentiert. Bei Aktionen die zu Fehlern führen könnten wird zwar der Fehler angemerkt, jedoch nur ein indirekter Lösungsvorschlag gegeben.
10. Hilfe und Dokumentation	Eine separate Dokumentation ist für die App nicht vorgesehen. Die App selbst verfügt an vielen Stellen über Informationssymbole, welche die Funktionalitäten der App erklären. Beispielsweise ist unterhalb des SOS Knopfes ein Informationssymbol zu sehen, welches bei Betätigung die SOS Funktionalität in einer Textbox beschreibt. Weiterhin kann in der Einstellungs-Ansicht beobachtet werden, dass für jedes Eingabefeld und jede Aktion ein Informationssymbol zur Verfügung steht, welche eine Textbox öffnet mit den jeweils notwendigen Informationen.

Tabelle 1: Ergebnisse der App-Analyse hinsichtlich der 10 Prinzipien nach Nielsen [Ni93].

5 Fazit

Die entwickelte App ermöglicht die Kommunikation mit einem modifizierten LoRaWAN-Netzwerk basierend auf Disruption Tolerant Networking. Sie kann mit Bundle Protocol 7 Pakete umgehen und bietet eine einfache Oberfläche, die sich an geläufigen Messengern orientiert. Das Ziel war es den Zugang zu dem Kommunikationsbackend möglichst einfach und nutzerfreundlich zu gestalten. Eine heuristische Analyse hinsichtlich der 10 Prinzipien nach Nielsen [Ni93] bescheinigt eine gute Gebrauchstauglichkeit. Die Oberfläche der App orientiert sich an gängigen Messengern und bietet Hilfstexte für besondere Funktionen und Einstellungen an. Wichtige Komfortfeatures sind soweit bekannt eingebaut, beispielsweise das automatische Neu-Verbinden mit dem Backend, sobald das Gerät wieder im korrekten Netzwerk (lokales Wi-Fi) ausgewählt ist. Ebenso werden Nachrichten zwischengespeichert, solange noch keine Verbindung besteht. Die Integration von SOS-Nachrichten soll eine schnelle und einfache Kommunikation eines Notfalls ermöglichen. Mit Hilfe der Kriterien von Nielsen wurden einige Design-Probleme frühzeitig erkannt und umgebaut, so dass derzeit die heuristische Evaluation wenig Verbesserungspotenzial offenbart. Tests mit potenziellen Anwendern sollen in Zukunft die Gebrauchstauglichkeit der Applikation evaluieren und Änderungswünsche am gesamten System erarbeiten.

Bibliografie

- [BFB22] Burleigh, S; Fall, K; Birrane, E. J.: Bundle Protocol Version 7 (Request for Comments Nr. RFC 9171): Internet Engineering Task Force, 2022
- [BH20] Bormann, C.; Hoffman, P. E.: Concise Binary Object Representation (CBOR) (Request for Comments Nr. RFC 8949): Internet Engineering Task Force, 2020
- [KBR23] Kuntke, F.; Baumgartner, L.; Reuter, C.: Rural Communication in Outage Scenarios: Disruption-Tolerant Networking via LoRaWAN Setups, In: Proc. of 20th Global Information Systems for Crisis Response and Management Conf. (ISCRAM), 2023.
- [Ni93] Nielsen, J.: Chapter 5: Usability Heuristics. In: Usability engineering. Academic Press, Boston, ISBN 978-0-12-518405-2, S. 115–164, 1993.
- [WU18] Waranashiwar, J.; Ukey, M.: Ionic Framework with Angular for Hybrid App Development. In: International Journal of New Technology and Research, Bd. 4, Nr. 5, 2018.

Ökologische Nachhaltigkeit -
Kolloquium Landwirtschaft der
Zukunft - Ist KI ein wesentlicher
Schlüssel zur nachhaltigeren
Landwirtschaft?

AI in the Wild: Challenges of Remote Deployments

Jens Dede,¹ David Wewetzer,² Anna Förster³

Abstract: The effect of humanity on the earth becomes more and more apparent. Besides the publicly discussed climate change and overpopulation, also the number of conflicts with wildlife increases. The technological progress of the past years helped to understand these challenges better. Monitoring solutions, known to the public as the Internet of Things (IoT), increase the amount of collected data, whereas artificial intelligence (AI) supports analyzing and gathering a deeper understanding. Most projects in the area of wildlife try to achieve a more sustainable usage of natural resources and a better coexistence with our environment.

The mAIInZaun project focuses on the conflict between wolves and livestock. It aims to introduce these new technologies into grazing management and foster non-lethally coexistence between livestock and predators. Artificial intelligence (AI) analyzes images and videos of the areas surrounding the pasture. The algorithms detect possible attackers or predators, such as wolves, stray dogs, bobcats, etc. In the second step, these animals are scared away using adaptive technologies. These can be sound, ultrasound, scent, light, etc.

These systems are usually operated in remote environments, raising challenges like hardware design, power requirements, and maintenance. This paper will discuss these challenges and how we address them in the mAIInZaun project.

Keywords: Artificial Intelligence, Deployment, Edge Computing

1 State of the Art

Council Directive 98/58/EC of 20 July 1998 concerning protecting animals kept for farming purposes [Co98] requires farm animals not kept inside barns to be protected from hazards whenever possible. This includes bad weather conditions and attacks by predators. The farm animals must also be protected from breaking out and unwanted animals accessing the pasture. This requires adapted fences to defend against predators like wolves[Bu, La22]. As of now, no fence meets the financial and ecological requirements of farmers and society at the same time. In some scenarios, like on dykes, integral nature reserves, and stony grounds, fixed fences with undermining protection are impossible or not allowed.

However, herd protection fences are the most promising animal protection approach[Pr16, DL22]. On meadows, the protection is mainly done using different types of fences. Mesh,

¹ Universität Bremen, Nachhaltige Kommunikationsnetze, Otto-Hahn-Allee 1, 28359 Bremen, Deutschland
jd@comnets.uni-bremen.de

² Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Deutschland wewetzer@uni-bremen.de

³ Universität Bremen, Nachhaltige Kommunikationsnetze, Otto-Hahn-Allee 1, 28359 Bremen, Deutschland
anna.foerster@comnets.uni-bremen.de

wire, or fixed bars are used for sheep and goats, whereas wire or fixed fences are preferred for cattle and horses. Most fences can also be implemented as electric ones. A generator creates pulses up to 10.000 volts leading to a safe but painful shock: The animal accepts the fence as a barrier. However, in some cases, these fences can be lethal for small wild animals and also horses. Those physical fences also have several drawbacks: They increase the number of barriers, cut the landscape, and interrupt the natural paths of wild animals reducing the genetic variety. An alternative is virtual fencing[An07, Ca19]: The animals are equipped with necklaces that contain sensors (motion, GNSS) and actuators (sound, electric shocks). Approaching the (virtual) boundary of the meadow leads first to an audio warning and later to electric shocks. This teaches the animal where the boundaries are. Another solution is the scent fence: the smell of a repellent should prevent the animals from crossing a street or leaving the meadow. One common challenge for fences is the financial effort for building and maintenance. The only option might be to change from grazing livestock to stall husbandry.

Besides fencing, other options like livestock guardian dogs, bonuses, hunting, or compulsory stabling exist. Also, the classical approach with a shepherd and herding dogs is an option.

As an alternative or extension to all these classical approaches, our mAIInZaun⁴ project supposes an AI-based system. It detects predators and starts deterrents to protect the livestock.

2 The mAIInZaun Project

The mAIInZaun project aims to strengthen traditional fences by adding AI-based technology. It consists mainly of two parts: The central part is detection: Sensors like video cameras continuously monitor the fence's environment. A connected embedded system processes these images locally at the perimeter by running AI, i.e., an object detector. If the system detects a predator, it activates the second part: the deterrents. Ultrasound, flickering light, and scent are three possible deterrents. They chase away the predator, ideally preventing an attack on the livestock. The challenges, especially of the sensing part for such a high-tech system, are entirely different from the established fences and are discussed in this work.

3 Challenges in Image Recognition in the Wild

Operating an AI in remote environments raises different challenges than classical systems. We focus on the two main parts: Hardware (Section 3.1) and detection (Section 3.2).

⁴ <https://intelligenter-herdenschutz.de/>

3.1 Hardware Challenges

The (physical) requirements of the hardware are often underestimated. This section discusses the main challenges in hardware design: What must be considered when building a new system? An additional challenge is the fault detection and error handling of the running system[ABF23], which is not discussed in this work due to length constraints.

3.1.1 Protect Against the Environment: The Case

Deploying a high-tech system on a field yields several challenges. The obvious ones are in the hardware itself: The devices have to survive in harsh environments, i.e., dust, rain, snow, dropping, etc. For that, a proper casing is required. The protection standard is defined using the so-called IP Codes, which consist of two numbers: The first stands for the protection against solid particles, the second for the water protection[In13, Eu13]. For our system, no dust should enter, i.e., the first digit is 6. Regarding water, we aim to achieve 7 or 8 (devices can survive underwater), which is challenging. So realistically, we reach 5 or 6 (waterjets from all directions). The resulting protection is called IP 65 or IP 66. It is essential to ensure high IP levels for the involved parts of the system, especially plugs, switches, cable glands, lights, etc.

Another neglected challenge is the humidity and temperature differences, which continuously change in outdoor environments. The challenge is that hot air contains more water than cold air. Due to the changes between day and night, the humidity from the air condensates inside the case resulting in a notable amount of water – even in completely tight cases. Pressurization valves can be used as a countermeasure, which lets air but not water pass.

The complete system must be protected against shock damage during handling. Here, several boxes are available on the market, and selecting one is not a real challenge. According to our experience, equipment from the automotive sector is a good starting point for the casing challenges: They are protected against harsh conditions on the road, are not very expensive, and are available.

3.1.2 How to Power Everything: The Energy

Another main challenge is the energy supply of the overall system. In this work, we neglect the energy required by the deterrents and only focus on the system's detection (AI) parts. An overview of possible parts and their power consumption is given in Table 1. We focus on two types of systems: One is using an off-the-shelf surveillance camera from Axis connected via Ethernet to a computing device. This solution has the advantage that we can use all the flexibility of the hardware and easily can extend the system. Here, mainly three options for the components are considered: Raspberry PI, NVIDIA Jetson, and Google

Coral. Alternative solutions are the Luxonis cameras, which can directly run AI models on the camera itself. Those are very powerful but restricted regarding additional functionality like additional image processing. For reference, we also added a small WiFi-LTE router from Huawei for Internet access.

Tab. 1: Comparison of different components and their energy consumption.

Component	Functionality	Required Power (W)	Power Source
Raspberry PI 4 ⁵	Processing	12.5-15W	USB, 5V
NVIDIA Jetson Nano ⁶	Processing	10-20W	USB, 5V
CORAL Dev Board ⁷	Processing	10-15W	USB, 5V
AXIS M2026 LE ⁸	Camera	5.1-7.9W	PoE
Luxonis OAK-D Pro ⁹	Camera, Processing	10W	USB, 5V
Huawei E8372 ¹⁰	Internet Router	5W	USB, 5V

Let's now consider a USB power bank with an approximate capacity of 10.000 mAh at 5V, i.e. 50 Wh, and ignore all the losses and side effects, additional converters, etc. We can estimate the resulting lifetime: Assuming all devices operate at 5 V, a 10 W device draws a current of 2 A, and the power bank will be empty after approximately 5 hours. If we consider the surveillance camera with a raspberry pi and an internet router, we draw 30 W or 6 A. Here, the power bank will be empty after roughly 100 minutes.

One can also consider a more powerful battery which can be compared to a car battery: higher capacity (55 Ah) at a higher voltage (12.8 V, resulting in 704 Wh) and higher weight (10 kg). Assuming a perfect voltage regulator again, we could get 140.000 mAh at 5 V leading to an estimated lifetime of 24 h. With the 10 W load, the battery will last 70 hours.

Using a solar panel might be an option to extend the lifetime. As the system also has to work during nighttime, cloudy days, and wintertime with only a little sunlight, the solar panel should offer at least 4-6 times the power which is required for the system. The fact that the panels could be shadowed by trees and bushes or get dirty is neglected here.

This shows two main challenges with the power supply: First, the careful selection and fine-tuning of the components regarding energy consumption is vital for a long lifetime in remote environments. Solutions like wakeup mechanisms or using different sensors can significantly increase the system lifetime. Second: If real-time data processing or even collection is planned, the power supply design is a crucial point.

⁵ <https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-datasheet.pdf>

⁶ <https://developer.nvidia.com/downloads/embedded/dlc/jetson-nano-system-module-datasheet>

⁷ <https://coral.ai/static/files/Coral-Dev-Board-datasheet.pdf>

⁸ <https://www.axis.com/dam/public/9a/35/3c/datasheet-axis-m2026-le-mk-ii-network-camera-en-US-289362.pdf>

⁹ https://github.com/luxonis/depthai-hardware/blob/master/DM9098_OAK-D-Pro/Datasheet/OAK-D-Pro_Datasheet.pdf

¹⁰ <https://consumer.huawei.com/in/routers/e8372/specs/>

As of now, we are in the state of data collection. Hence, the power supply is considered less critical. For the productive system, we plan to use optimized hardware that can hibernate for a more extended time and is woken up in case of activity.

3.1.3 In the Wild: Weight and Deployment

The third challenge is the deployment itself. Non-technical users are usually deploying the systems into sometimes very remote areas, leading to the requirement that everything can be set up quickly. Complex alignments of the devices or pairing should be avoided. The system should perform self-tests to ensure good camera alignment and notify users of errors like dropped cameras or dirty lenses. Also, the weight (and thus the battery size) is limited as the device might be carried by hand. Ideally, no regular maintenance like changing batteries, refilling materials, or just cleaning lenses or solar panels is required.

3.1.4 Connecting Home: Communication

For communication, we distinguish between two types of systems: If internet access is required, one needs a connection based on the infrastructure of third parties, for example, using WiFi, cellular networks (4G, 5G), or satellite (Starlink, Iridium, etc.), SigFox, etc. If only local communication is required, i.e., between the different parts of the system, alternative low-power systems can be used. Here, LoRa and IEEE 802.15.4 (ZigBee) are the two leading players. Here, the bandwidth and the energy requirements highly vary between less than a Watt (LoRa, IEEE 802.15.2) up to 50-75W (150 W in high performance) for Starlink. The latter power requirements can not be met over a longer time using batteries.

3.2 Finding Wolves: Detection Challenges

Besides the hardware setup, also the software has several challenges. Using AI requires high-quality input data – at least for training. Most algorithms focus on professionally taken images with the object of interest in focus. In the wild, animals should be detected which are only partly visible. Vegetation, weather, dirt, and lighting conditions hamper the detection and require unique algorithms and methods. The training and tuning of those require a significant number of training images from different scenarios and thus time to collect those.

Figure 1 shows four exemplary photos from our dataset. Besides the changing weather- and light conditions, as can be seen in Figure 1a, 1b and 1c, also other animals and insects like spiders (c.f. Figure 1d) can affect the images. In addition, the objects of interest – in our cases, the wolves – do not appear directly in front of the camera but pass by sometimes close, sometimes far away at different speeds. This rises two main challenges: Firstly, the system needs to be able to detect those images with different camera resolutions, light sensitivity,



(a) A bright day with back-light. The camera needs a high dynamic range to be able to offer good images. (b) Foggy weather. The range of visual images is limited. (c) rain and snow in combination with (infrared) light blind the camera and thus also reduce the range. (d) Animals like spiders can also affect the range of the camera.

Fig. 1: Example images: What can go wrong with camera images in the wild?

frame rate, etc.: The further away objects should be seen, the higher the resolution needs to be. Secondly, a training dataset for this kind of images: Most available datasets offer nice ready-to-print photos. The use of those is only limited to our camera trap images. Therefore, we collect our own training dataset.

Especially for the latter, several hundred up to thousands have to be labeled due to the wide variety of the composition of images. Even though we have implemented a tool reducing the workload for a single person[DF23], it is still a time-consuming task.

We are implementing and evaluating a workflow to optimize the labeling and detection of this kind of images captured in the wild. These first results look promising but also show that traditional approaches need to be fixed on this task.

Also, data privacy is an essential part of the overall system. Especially with imaging technologies, special measures have to be taken to ensure the rights of passing by people. In our case, the detection runs on the fence without transmitting data to the internet or storing them permanently. Thus, we consider this a minor issue.

4 Conclusion: Is AI an essential key to more sustainable agriculture?

The application of AI in our project differs significantly from others in agriculture. As discussed in this work, we are facing two main challenges: power requirements for the in-field system and gathering suitable (training) data. Especially considering the achievements in the area of AI and the corresponding toolchains in the last years, we are very convenient to solve those with an optimized hardware approach and workflow for the training data generation. In the end, the AI outcome of this project can be an essential key to more sustainable agriculture.

Bibliography

- [ABF23] Attarha, S.; Band, S.; Förster, A.: Automated Fault Detection Framework for Reliable Provision of IoT Applications in Agriculture. In: 19th International Conference on the Design of Reliable Communication Networks (DRCN). IEEE, pp. 1–8, 2023.
- [An07] Anderson, D.M.: Virtual fencing—past, present and future. *The Rangeland Journal*, 29(1):65–78, 2007.
- [Bu] Bundesministerium der Justiz und für Verbraucherschutz: Verordnung zum Schutz landwirtschaftlicher Nutztiere und anderer zur Erzeugung tierischer Produkte gehaltener Tiere bei ihrer Haltung (Tierschutz-Nutztierhaltungsverordnung - TierSchNutztV), § 3 Allgemeine Anforderungen an Haltungseinrichtungen. Available online: https://www.gesetze-im-internet.de/tierschnutztv/__3.html, Retrieved: 2023-05-09.
- [Ca19] Campbell, D.L.M.; Lea, J.M.; Keshavarzi, H.; Lee, C.: Virtual fencing is comparable to electric tape fencing for cattle behavior and welfare. *Frontiers in Veterinary Science*, 6:445, 2019.
- [Co98] Council of European Union: , Council regulation (EU) no 58/1998: concerning the protection of animals kept for farming purposes, 1998. <http://data.europa.eu/eli/dir/1998/58/2019-12-14>, Retrieved: 2023-05-09.
- [DF23] Dede, J.; Förster, A.: Animals in the Wild: Using Crowdsourcing to Enhance the Labelling of Camera Trap Images. In: 2nd DISCOLI Workshop on DIStributed COLlective Intelligence (DISCOLI 2023). IEEE, 2023. accepted paper.
- [DL22] DLG e.V., Fachzentrum Landwirtschaft, Eschborner Landstraße 122, 60489 Frankfurt am Main: Herdenschutz gegen den Wolf: Leitfaden Elektrozaune. 2022. Available online: https://www.dlg.org/fileadmin/downloads/landwirtschaft/themen/publikationen/merkblaetter/dlg-merkblatt_455.pdf, Retrieved: 2023-05-15.
- [Eu13] European Committee for Electrotechnical Standardization: , Degrees of protection provided by enclosures (IP Code), 10 2013.
- [In13] International Organization for Standardization: , ISO 20653:2013-02: Road vehicles – Degrees of protection (IP code) – Protection of electrical equipment against foreign objects, water and access, 02 2013.
- [La22] Landwirtschaftskammer Nordrhein-Westfalen: Checkliste Cross Compliance 2022 für landwirtschaftliche Unternehmen in Nordrhein-Westfalen. 2022. Available online: <https://www.landwirtschaftskammer.de/landwirtschaft/beratung/pdf/cross-compliance-checkliste-nrw.pdf>, Retrieved: 2023-05-05.
- [Pr16] Priebe, R.; Leitner, P.J.; Hasselfeldt, K.; Kulmann, J.: Sichere Weidezäune. Bundesanstalt für Landwirtschaft und Ernährung, Bonn, Germany, 6 edition, April 2016.

Textual Descriptions Used for Classification of Oaked vs Unoaked Wines

Ronald Böck,¹ Siddarth Venkateswaran,¹ Thi Nguyen,² and Dominik Durner²

Abstract: Winemaking and grapegrowing are sciences with a long tradition dealing with one of the most complex beverages in the world. This complexity stems from the winemaking process itself as well as the characteristics of the final product. Wine's aroma is often described through scalar assessments, though here we are focusing on textual descriptions, transferring methods from the natural language processing (NLP) community to the wine domain, in particular to analyse the statements of human panellists. Textual descriptions were used for the classification of oaked versus unoaked wines as an initial demonstration of NLP in the wine domain. We achieved significant discrimination results of 0.79 F1-score comparing BERT and Naïve Bayes classifiers. This shows that more natural textual (and potentially spoken) descriptions of wine, being later combined with classical scalar assessments, can provide more flexibility to human panellists.

Keywords: Wine Descriptors, Classification, BERT, NLP

1 Introduction

Viticulture and oenology, the sciences of grapegrowing and winemaking, have a long tradition dealing with one of the most chemically and sensorially complex beverages in the world. The many volatile compounds contributing to wine aroma can be measured by instrumental means, such as gas chromatography, though an accurate prediction of how wine will actually smell based on its chemical composition is not yet possible [FdS22]. Therefore, the responsibility of describing wine aroma currently lies with human panellists taking part in formal sensory evaluations [BGT21].

For the PINOT³ project, we are developing a multimodal, artificial intelligence (AI)-based approach to make assessments of wine aroma more objective. As can be seen in Figure 1, there are many points along the wine production and supply chain where AI can potentially assist (also in terms of sustainability), ranging from the prediction of wine quality based on the starting grape material, to matching wines to consumers' tastes and expectations. Here, we are focussing on the application of AI toward describing the finished product, assisting in the sensory evaluation of wine aroma. As a first step, we sought to determine whether AI

¹ Genie Enterprise, Research Division, Donnersbergweg 1, 67059 Ludwigshafen, Germany {rboeck,venkat}@genie-enterprise.com

² Weincampus Neustadt, Breitenweg 71, 67345 Neustadt an der Weinstraße, Germany {thi.nguyen,dominik.durner}@dlr.rlp.de

³ <https://pinot-ai.com>

can be used to extract a common “understanding” of wine’s properties based on textual descriptions generated by panellists during sensory evaluation.

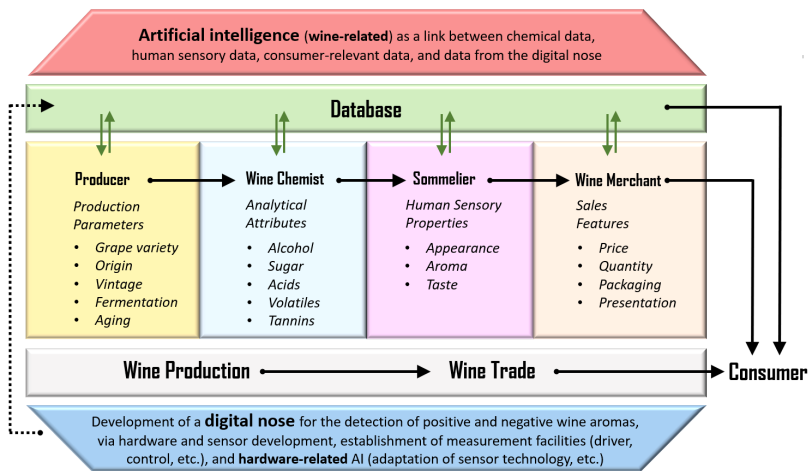


Fig. 1: Scope graphic highlighting the production chain in wine making. Additionally indicating the particular process’ stakeholders. This current core is being framed by the options of AI support on an either hardware- or wine-related perspective.

1.1 Motivation and Research Aspect

The purpose of this investigation was to transfer Natural language processing (NLP) techniques to the winemaking domain and to evaluate the performance of two approaches to text-based wine assessment, relying on large neural networks (Bidirectional Encoder Representation from Transformers (BERT)) and on classical machine learning (Naïve Bayes). Their ability to distinguish between oaked and unoaked wines based on textual descriptions collected during sensory evaluation by wine professionals was compared, aiming for a possible combination with classical scalar assessments, thereby providing more flexibility to the human panellists.

1.2 Related Work

Until recently, sensory evaluation has conventionally relied on quantitative methods resulting in numerical descriptions of flavour; considered by many as “gold standard” in sensory evaluation, descriptive analysis yields intensity values for various aroma, taste, and mouthfeel properties, as rated by trained panellists on scales. Alternatively, check-all-that-apply methodology generates the counts/frequencies at which certain descriptors are chosen by panellists to describe the product [BGT21]. While such data is relatively simple to

process and analyse, panel training (to reach consensual understanding of descriptors) is often time-consuming and costly. There is also the risk of forgetting to include important attributes, and the risk of panellists feeling imposed/limited by the list of attributes [Vi23]. However, this landscape is changing, with more rapid, cost-effective, and consumer-oriented methods being developed, involving open questions and allowing free responses [Pi15]. Such methodologies have only recently been implemented in the field of oenology for wine analysis [La13; Ma20; Vi23]. Panellists are asked to describe wines in their own words, and while the text generated is much more natural and richer in information, the manner in which it must be pre-processed for analysis is far from straightforward [Ma21]. Advancements in technology, particularly in machine learning, have facilitated the automation of a once tedious process [Vi20].

Classical text classification techniques involve multiple pre-processing steps, creating an embedding on these cleaned texts using approaches like Term Frequency Inverse Document Frequency (TF-IDF), and feeding them to a classification algorithm (e.g. [IKT05]). One drawback of this method involves the curse of dimensionality, which arises due to an increase in the amount of data and features over a period of time. To handle this issue, feature reduction techniques like Global Vectors for Word Representation were implemented (e.g. [Si22]). Further issues with classical Machine Learning (ML) approaches involve the lack of consideration of context in which a phrase is used (e.g. [SMH22]). Recently, the ability to train large neural networks led to the development of models like BERT [De18] which takes into account the semantic nature of texts, thereby attaining better classification performances as compared to traditional ML techniques [GG20]. In this paper, we compare performances of large neural networks [De18] and classical ML techniques [Le07], distinguishing oaked and unoaked wines, based on textual descriptors, collected from sensory evaluation by trained wine professionals.

2 Data Collection and Preparation

Classification is based on more than 1,200 textual descriptions generated during a wine tasting conducted in January 2023 at Weincampus Neustadt. Sixteen wines (four white, four red, with and without oak) were served in both black and clear glasses to a panel of 16 wine experts, who were asked to provide textual descriptions of each wine's smell and taste.⁴

For experiments, we used a stratified sample approach, considering multiple data splits naturally given by tasting conditions. These will be referred to as scenarios for the classification experiments conducted in this paper, described in Table 1. Furthermore, we compared classification based on text only to classification based on a combination of text and meta-information giving additional details about the wine, namely its vintage, grape variety, country of origin, and vineyard. The baseline in general, but also specifically for the scenarios 2 to 7, is scenario 1 where all texts were taken into consideration.

⁴ We are only able to share the data on request.

Scenarios	Data Selection	Train/Test Breakup per Class
1	Taking All Texts Into Consideration	600 - 40
2 & 3	Considering texts specific to black or clear glass sessions respectively	300 - 20
4 - 7	Considering texts fine-grained to either smell or taste sessions, based on black or clear glasses	150 - 10

Tab. 1: Data breakup across scenarios

3 Methods and Experimental Setup

3.1 Classification Algorithms and Data Embedding Techniques

Since the manuscript’s aim is to demonstrate capabilities of NLP-techniques in the wine domain, we selected two approaches for the classification experiments: a BERT-based Sentence Classifier [De18] and the Naïve Bayes Classifier [Le07]. BERT (cf. [De18]) is a neural network applied commonly in the field of NLP. It is trained in an unsupervised manner to predict texts surrounding a specific word, making them bi-directional in nature. Considering the limitations of preceding NLP-based, recurrent-based neural networks which could process textual representations in a fixed sequence (either left to right, or vice-versa), the bi-directional nature of BERT made it a ground-breaking implementation for all downstream NLP tasks (e.g. sentence classification). The Naïve Bayes Classifier is based on the Bayes’ Theorem that assumes the effect that particularly the presence of a word in a text is independent of the value of other words predicting a class (cf. [Le07]). Given the fact that the Naïve Bayes Classifier is considered as a baseline classification algorithm, and given the ground-breaking abilities of the BERT model for tasks related to NLP, these methods were chosen for all experiments in this research.

Furthermore, two different textual embedding techniques were employed, namely German-based BERT embeddings as well as TF-IDF embeddings. The German BERT models were used in a pre-trained version of [De20], being adapted to the current needs in the wine domain. Considering their bi-directional nature, the BERT embeddings are contextual by definition. In contrast, TF-IDF embeddings measures the frequency of occurrence of a word in a given set of sentences, neither considering the position of a word in a sentence, nor its surrounding words (cf. [SMH22]), this being non-contextual.

In our experiments, the BERT classifier was fine-tuned on BERT embedded texts, while the Naïve Bayes Classifier was trained from scratch on the TF-IDF embedded texts. These techniques were chosen to compare the model performances of large neural networks, which have attained state-of-the-art performances in the field of NLP, against classical ML techniques (cf. e.g. [GG20; SMH22]), in the wine domain.

3.2 Experimental Setup

The BERT Sentence Classifier experiments were based, in particular, on the Flair framework [Ak19]. For matter of reproducibility, the network and training parameters were set as follows, given the predefined structure of BERT according to [Ak19; De18]: a learning rate of $5e^{-6}$, a mini batch size of 4, each for a total of 10 epochs.

Experiments using the Naïve Bayes approach were implemented using the default setting in the Scikit-Learn package [Pe11].

In addition, we conducted separate experiments per classifier type and scenario, considering either purely textual content or textual content merged with wine-metadata. To account for generalisation, at least 5-folds for each condition were trained. For presentation of results, we used the F1-Scores per scenario, being averaged across respective folds.

Additionally, for a better rating of the achievements, non-parametric statistical evaluations were performed to check if the differences between the textual descriptors across different scenarios, mentioned in Section 2, were significant. Further, intra-scenario significance values were computed to check if adding meta-data to the textual descriptors yielded any differences in the model performances. For both investigations, we applied the non-parametric Kruskal-Wallis-Test (cf. [KW52]), using two significance levels of $p < 0.05$ (significant) and $p < 0.01$ (highly significant), usually compared to the baseline.

4 Results

Table 2 compares the model performances across the scenarios. Considering the baseline (scenario 1), an F1-score ≥ 0.73 was achieved using the BERT-classifier with and without the addition of wine metadata. It can also be seen across all scenarios that BERT-classifiers outperformed the Naïve Bayes counterparts.

Regarding the other scenarios, the best performance was achieved on the black glass setting (highlighted in Table 2). A first interpretation might be that the human participants were rather focused to the wine's characteristics. However, this is a matter of further investigations. Using wine-metadata improved the classification performance just slightly (cf. Table 2).

With respect to the motivation in Section 1.1, we can state that a textual description-based classification of wine characteristics, especially considering oaked vs unoaked wines, is possible. Furthermore, we saw that there is some discriminative power in the textual descriptions provided by the human panel to estimate wine characteristics.

On delving into the non-parametric statistical evaluations at an inter-scenario level, BERT-classifiers observed smaller p -values, although no statistical significance was reached (mean⁵ p -value of 0.15) as compared to Naïve Bayes classifiers (mean p -value of 0.46). Further, at an intra-scenario level we found the following: Considering the BERT results, we can state that significant differences are given between baseline (scenario 1) and multiple

⁵ Mean value was calculated as average across all folds and scenarios.

Scenario	Only Textual Descriptions		Textual Descriptions + Wine Metadata	
	BERT	Naïve Bayes	BERT	Naïve Bayes
1	0.73	0.59	0.78	0.58
2	0.79*	0.60	0.78	0.59
3	0.77	0.58	0.77*	0.60
4	0.68	0.61	0.74	0.61
5	0.65**	0.50	0.67**	0.48
6	0.74*	0.64	0.75**	0.64
7	0.63**	0.53	0.64**	0.54

Tab. 2: Comparison of BERT Sentence Classifier fine-tuned on German BERT embeddings, and Naïve Bayes Classifier trained on TF-IDF embeddings across different scenarios in which the wine descriptions were collected. Significance is indicated by * ($p < 0.05$) and ** ($p < 0.01$).

other scenarios, especially scenario 2 ($p = 0.04$), 5 ($p = 0.00$), 6 ($p = 0.01$), and 7 ($p = 0.00$). Same observations are holding true in the setting of metadata, where scenario 1 has significant difference to scenario 3 ($p = 0.02$), 5 ($p = 0.00$), 6 ($p = 0.00$), and 7 ($p = 0.00$). However, there was no significant difference observed across scenarios utilising the Naïve Bayes classifier (with an mean p -value of 0.92).

In a broader sense, the usage of large neural networks outperforms classical ML-techniques, irrespective of the size of the corpus. They have the ability to learn the semantic and the syntactic nature of the texts, due to which they tend to find differences across classes. This benefit can be used also in the wine domain to estimate particular characteristics more objectively, aiming for a support in the entire production process as mentioned in Section 1.

5 Discussion and Conclusion

This paper transferred NLP-techniques to the wine domain and evaluated the performance of BERT against Naïve Bayes models for the task of classifying wines (oaked/unoaked), based on the textual descriptions provided by trained human panellists. While BERT-classifiers outperformed Naïve Bayes classifiers across all scenarios (cf. Table 2), they tended to perform best when trained with texts related to the glass' colour (i.e. black/white, sessions 2 and 3). Also regarding the limited availability of data, and a 50% chance level of a 2-class task, F1-scores greater than 0.63 (worst performance) were already achieved across all scenarios, with and without the addition of metadata. This is showing the usability of NLP methods also in the wine domain.

As mentioned in Section 1, our work is also related to the support of wine makers and stakeholders in oenology. Using a combination of sensor-based measures of wines and trained NLP models on panel-based assessments, we are aiming in future work for an AI-based assistance in quality assessment of wines.

Acknowledgement



We acknowledge support by the PINOT project funded by the German Federal Ministry of Food and Agriculture (BMEL) under grant number 28DK107A20 and 28DK107C20.

Bibliography

- [Ak19] Akbik, A.; Bergmann, T.; Blythe, D.; Rasul, K.; Schweter, S.; Vollgraf, R.: FLAIR: An easy-to-use framework for state-of-the-art NLP. In: NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations). Pp. 54–59, 2019.
- [BGT21] Barbe, J.-C.; Garbay, J.; Tempère, S.: The Sensory Space of Wines: From Concept to Evaluation and Description. A Review. *Foods* 10/6, 2021, ISSN: 2304-8158.
- [De18] Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805/, 2018.
- [De20] Deepset: Open Sourcing German BERT Model, Accessed on May 4, 2023, 2020, URL: <https://www.deepset.ai/german-bert>.
- [FdS22] Ferreira, V.; de la Fuente, A.; Sáenz-Navajas, M. P.: 1 - Wine aroma vectors and sensory attributes. In (Reynolds, A. G., ed.): *Managing Wine Quality* (Second Edition). Second Edition, Woodhead Publishing Series in Food Science, Technology and Nutrition, Woodhead Publishing, pp. 3–39, 2022, ISBN: 978-0-08-102067-8.
- [GG20] González-Carvajal, S.; Garrido-Merchán, E. C.: Comparing BERT against traditional machine learning text classification. arXiv preprint arXiv:2005.13012/, 2020.
- [IKT05] Ikonomakis, M.; Kotsiantis, S.; Tampakas, V.: Text classification using machine learning techniques. *WSEAS transactions on computers* 4/8, pp. 966–974, 2005.
- [KW52] Kruskal, W. H.; Wallis, W. A.: Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association* 47/260, pp. 583–621, 1952, visited on: 01/12/2023.
- [La13] Lawrence, G.; Symoneaux, R.; Maitre, I.; Brossaud, F.; Maestrojuan, M.; Mehinagic, E.: Using the free comments method for sensory characterisation of Cabernet Franc wines: Comparison with classical profiling in a professional context. *English, Food Quality and Preference* 30/2, pp. 145–155, 2013.
- [Le07] Leung, K. M. et al.: Naive bayesian classifier. Polytechnic University Department of Computer Science/Finance and Risk Engineering 2007/, pp. 123–156, 2007.

- [Ma20] Mahieu, B.; Visalli, M.; Thomas, A.; Schlich, P.: Free-comment outperformed check-all-that-apply in the sensory characterisation of wines with consumers at home. *Food Quality and Preference* 84/, ed. by Ltd., E. S., p. 103937, Sept. 2020.
- [Ma21] Mahieu, B.; Schlich, P.; Visalli, M.; Cardot, H.: A multiple-response chi-square framework for the analysis of Free-Comment and Check-All-That-Apply data. *Food Quality and Preference* 93/, p. 104256, Oct. 2021.
- [Pe11] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E.: Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12/, pp. 2825–2830, 2011.
- [Pi15] Piqueras-Fiszman, B.: 12 - Open-ended questions in sensory testing practice. In (Delarue, J.; Lawlor, J. B.; Rogeaux, M., eds.): *Rapid Sensory Profiling Techniques*. Woodhead Publishing Series in Food Science, Technology and Nutrition, Woodhead Publishing, pp. 247–267, 2015, ISBN: 978-1-78242-248-8.
- [Si22] Singh, K. N.; Devi, S. D.; Devi, H. M.; Mahanta, A. K.: A novel approach for dimension reduction using word embedding: An enhanced text classification approach. *International Journal of Information Management Data Insights* 2/1, p. 100061, 2022.
- [SMH22] Subakti, A.; Murfi, H.; Hariadi, N.: The performance of BERT as data representation of text clustering. *Journal of big Data* 9/1, pp. 1–21, 2022.
- [Vi20] Visalli, M.; Mahieu, B.; Thomas, A.; Schlich, P.: Automated sentiment analysis of Free-Comment: An indirect liking measurement? *Food Quality and Preference* 82/, p. 103888, 2020, ISSN: 0950-3293.
- [Vi23] Visalli, M.; Dubois, M.; Schlich, P.; Ric, F.; Cardebat, J.-M.; Georgantzis, N.: Relevance of free-comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise. *Food Quality and Preference* 105/, p. 104785, 2023, ISSN: 0950-3293.

Emission-Reducing Vehicle Routing in Food Logistics

Armin Wolf ¹ and Silke Cuno ²


Abstract: A web service for emission-reducing vehicle routing in food logistics is presented. This service computes emission-reduced round trips for the transportation of food products from (and to) a depot to (and from) regional locations like food stores or food hubs. Further, this service considers the travelling times between locations and the produced greenhouse gas emissions of the according drives as well as the time windows needed for pickup and delivery. The loading capacities of the vehicles and an efficient “last-in-first-out” packaging order of the vehicles are respected, with the consequence that goods picked-up last should be delivered and unloaded first. This is to avoid unnecessary work.


Keywords: Artificial Intelligence; Vehicle Routing; Food Logistics; Sustainability; Operations Research; Graph Theory; Resource Scheduling Problem; Constraint Optimisation; Travelling Salesperson Problem

1 Introduction

Within the research and development project “Stadt-Land-Fluss” (SLF) in the field of regional food systems, funded by the German Federal Ministry of Food and Agriculture, an AI application for transport optimisation in food logistics was developed: a web service that computes emission-reduced round trips for pickup and delivery tasks of goods, mainly food.

The objective of the SLF emission-reducing vehicle routing service is to minimise the total emissions of greenhouse gases caused by the daily journeys of a regional wholesaler’s fleet of vehicles, who distributes his daily product pallets of different goods. These journeys start either from a central warehouse or from rural areas and lead to various delivery points in and around a city, where the food orders are delivered. The underlying vehicle routing problem for the transport of food orders defines a time and resource planning problem where the most optimal schedule respecting time, space and capacity constraints is created: Each vehicle has a start time at its origin location and an end time at its target destination. For each food order the transport vehicle must be determined as well as its pickup time at its origin and its delivery time at its destination. The task is to select a sequence for visiting several locations in such a way that no location apart from the first is visited more than once. The whole journey should be at low-emission as possible and a round trip, i. e., the first stop must be the same as the last. Time windows for the collection or delivery of goods are taken into account as well as the loading capacities of the delivery vehicles used. Goods to be delivered first have to be loaded last. The emissions of the different vehicles in the fleet must

¹ Fraunhofer FOKUS, Kaiserin-Augusta Allee 31, 10589 Berlin, Deutschland armin.wolf@fokus.fraunhofer.de 

² Fraunhofer FOKUS, Kaiserin-Augusta Allee 31, 10589 Berlin, Deutschland silke.cuno@fokus.fraunhofer.de 

be taken into account. The routing service optimises these routes to reduce emissions and save fuel. The problem is characterised by the fact that in addition to the travel time between the stations to be served, the emissions for these journeys are also reflected on. Collecting emissions data is a considerable challenge, such that in most cases the driving distances multiplied by a vehicle-typical emission value (e. g., in [gCO₂equiv/km]) were used.³

2 Related Work

A state-of-the-art approximate solver for Travelling Salesperson Problems (TSPs) is the so-called Lin-Kernighan-Helsgaun (LKH) solver [He00; He09] implementing variable depth local search [LK73]. However, in vehicle routing solving the underlying TSP is not sufficient because additional constraints such as limited resources, time windows for pickup and delivery and precedence constraints must be satisfied, too. To overcome these limitations there is an extension of the LKH [He17] transforming such routing problems into standard symmetric travelling salesperson problems and handling constraints by means of penalty functions within the approximate local search. Beyond these approximate solution approaches, there are also exact solution approaches.

Recently, tour scheduling, similar to vehicle routing, was considered and solved based on Constraint Programming (CP) [DV21]. The main differences are that trailers are transported by autonomous trucks without consideration of load capacities nor consideration of target destinations of the trucks after a one-day period, which is, however, important for human drivers. Furthermore, the objective there is the reduction of empty tours while we focus on the minimisation of the emissions of greenhouse gases. Finally, a (mixed integer) linear solver is used in [DV21] while we use a finite integer domain constraint solver based on recent research results (cf. [Be12]). There filtering methods for the *WeightedCircuit* constraint are presented. This constraint and even the underlying filtering methods are highly important for modelling and solving round trip problems like the TSP or the vehicle routing problems considered here. The presented methods are based on a relaxed problem considering so-called *1-trees*. In [IR19; IR21b] filtering based on *k-cuts* is presented, where necessary conditions of the graph to be connected by a *Hamiltonian* cycle, i. e., a round trip, are considered to determine necessary mandatory edges or to remove edges not contained in any solution or to detect inconsistencies. A similar approach based on the idea of swapping *k* edges on a tour used within LKH solvers to improve (e. g., the distance or emission of) the tour is considered in the *k-opt* constraint [IR21a]. These filtering methods can be combined with the filtering on the *WeightedCircuit* constraint. However, their power depends on the “density”/“connectivity” of the considered graph. If the graph is strongly connected filtering will less effective. This can be the case when the upper bound of the total weight of the round trip resp. of the *Hamiltonian* cycle will be rather weak. Thus this kind of filtering is effective when the upper bound is close to the optimum such that the remaining graph is rather sparse.

³ see also Sec. 5 on dealing with this challenge.

3 Emission-Reducing Vehicle Routing in Food Logistics

3.1 The Considered Routing Problem

We are looking for a solution to an *emission-reducing vehicle routing problem*, namely a schedule for a round trip of the vehicle starting at and ending at a depot, visiting all other destinations exactly once within their allowed time windows. This schedule consists of picking up and delivering the orders within the defined time windows, satisfying the required travel times between the locations while minimising the total greenhouse gas emission for the trip. Specifically, an *emission-reducing vehicle routing problem* is defined by

- A *depot* which is a place where the round trip starts and ends.
- A *vehicle* with a limited capacity, an earliest time and a latest time of availability.
- A list of *orders*, i. e., a list of goods to be delivered from an origin location, usually the depot, to another destination. Each order requires that
 - The collection and delivery are work tasks of a certain duration.
 - The goods must be collected from their origin within a given time window.
 - The goods are delivered to the destination within a given time window.
 - The goods are loaded into the vehicle in such a way that the last loaded goods are unloaded first, if possible.
 - The capacity required by the goods never exceeds the capacity of the vehicle.
- a *distance matrix* where the journey times between any two locations are given.
- an *emission matrix* giving the greenhouse gas emissions produced by the vehicle when driving between any two locations.

The main diagonals of both matrices are obviously zero and currently the matrices must be *symmetric*, i. e., the data for driving from A to B is the same as for driving from B to A .

3.2 Solution Approach and Innovation

Our vehicle routing problem is modelled and solved as a “Constraint Optimisation Problem” (COP). Emission-reducing vehicle routing is a specialisation of the “Travelling Salesperson Problem with Time Windows” (TSPTW) belonging to the class of NP-hard problems. Therefore Artificial Intelligence (AI) methods (in particular heuristic search methods) are used in combination with Operations Research (OR) filtering methods for exact solutions.⁴

⁴ approximate solution approaches are referred in Section 2.

Thus, our solution approach is essentially based on a constraint optimisation model adopted from [DCP16]. For the *WeightedCircuit* constraint, we use graph-theoretical filtering methods adopted from [Be12] and further OR filtering methods, e. g., for the *AllDifferent* constraint the methods presented in [Ré94], to reduce the search space.

The implementation and application of recent research results and the deployment of a vehicle routing service that can help to reduce greenhouse gas emissions in food logistics which is available over the web using open standards is the main innovation of our approach.

3.3 Implementation and Deployment as Web Service

For emission-reducing vehicle routing we use the Java constraint solver library “firstCS” [Wo12]. This library is expanded to satisfy the application specific constraints in the problem model (cf. Section 3.2). Therefore, we implemented the according filtering methods. For improved vehicle routing schedules, a special heuristic search is implemented that determines loading and unloading activities for each vehicle mainly in a “last-in-first-out” order. In this way rearranging activities of the loads/pallets stored in the loading space during the delivery trip should generally be avoided. For optimisation we have chosen a monotonous “Branch & Bound” strategy. In this way, step-wise improved solutions, at least 5 % better than the previous solution, are found until there is not any better solution.

The emission-reducing vehicle routing was realised as a web service. It can be used by sending planning requests in a predefined JSON format via http(s) to the URL of the web service. Increasingly better planning results are then sent to the enquirer by e-mail, in JSON and CSV format, for his/her enquiry. Alternatively, using a uniquely generated and returned reference number of the request, the current best planning result can be queried via http as well. These asynchronous approaches were chosen because the calculation of round trips can be very time-consuming. An example of a planning request is given in Listing 1.

```
{
  "timeUnit": "min",
  "capacityUnit": "pallettes",
  "emissionUnit": "g (CO0-equiv)",
  "locations": ["Berlin-Depot", ..., "Blankenfelde-Dahlewitzer"],
  "vehicles": [
    {
      "vehicle": "v#0815",
      "sourceLocation": "Berlin-Depot",
      "earliestDepartureTime": "2022-02-08T04:00:00Z",
      "targetLocation": "Berlin-Depot",
      "latestArrivalTime": "2022-02-08T13:00:00Z",
      "availableLoadCapacity": 33,
      "transitionTimes": [[0, ...29], ..., [29, ..., 0]],
      "transitionEmissions": [[0, ..., 14000], [14000, ..., 0]],
    }
  ],
}
```

```

"orders": [
  {
    "order": "1170",
    "origination": "Berlin-Depot",
    "earliestPickupTime": "2022-02-08T01:00:00Z",
    "latestPickupTime": "2022-02-08T04:30:00Z",
    "destination": "Berlin-Drakestrasse",
    "earliestDeliveryTime": "2022-02-08T04:30:00Z",
    "latestDeliveryTime": "2022-02-08T13:00:00Z",
    "requiredLoadCapacity": 2,
    "loadingDuration": 1,
    "unloadingDuration": 1
  }, ...
]
}

```

List. 1: Sample planning request in JSON format

Such a JSON request (cf. List. 1) consists of data on physical dimensions, locations, technical data on the used vehicle (e. g., load capacity, availability, etc.), transition times and emissions of the vehicle for the route sections between the defined locations as well as data on the orders of a daily delivery tour. The durations for loading and unloading goods are ideally based on measurements; here they are empirical values. Loading must start within the earliest and latest pickup time, unloading within the earliest and latest delivery time. Using these data sets, planning requests could be made to optimise the round trips of the vehicles and the according pickup and delivery tasks.

4 First Results

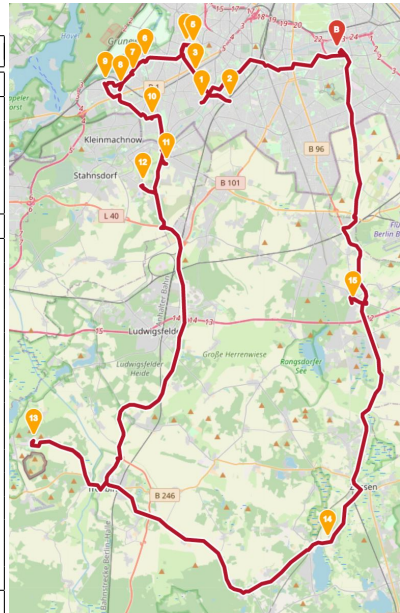
First results are based on real-world data. There, the round trip and the schedule of the pickup and delivery tasks of the vehicle routing problem consist of 15 orders. This is presented in Table 1 where orders from a depot located in Berlin must be delivered to 15 different locations in and around Berlin. The journeys from the depot to the destinations and back to the depot produce a total emission of about ≈ 103.75 kg of greenhouse gas, which is minimal and about $\approx 10\%$ less than the emission of the initial solution, which is about ≈ 116.5 kg of greenhouse gas. While the computation of the initial solution requires ≈ 60 secs,⁵ the computation of the optimal solution takes about ≈ 150 secs. Similar results were obtained for another vehicle routing problem defined by 12 orders. There, the initial solution (≈ 100.8 kg emission) is computed within ≈ 20 secs. and the optimal solution (≈ 94.0 kg emission; $\approx 7\%$ improvement) is computed within ≈ 21 secs. Table 1 contains for each order: its number, i. e., its identifier (first column), according tasks, i. e., pickup and delivery (second column), the locations where these tasks are performed (third column) and the earliest and latest start times of these tasks (last two columns). All tasks in Table 1 must be carried out in this chronological order. The differences between the earliest and latest start times define some buffer times, which become important when a task or a drive takes

⁵ Run-time measured on Intel Core i7-6600 CPU, 2.6 GHz, 12 GByte memory, Windows 10 Pro, OpenJDK 8u372.

longer than assumed. If there is enough buffer time then there will be no delay to the overall timetable and it is expected that all tasks can be completed within the given time slots.

[visualisation generated with openrouteservice.org]

Task schedule of a pickup-and-delivery round trip				
order	task	location	earliest	latest
8517	pickup	Berlin-Depot	04:10	04:11
.
.
1158	pickup	Berlin-Depot	04:23	04:29
1430	pickup	Berlin-Depot	04:24	04:30
	depart.	Berlin-Depot	04:25	04:31
1430	delivery	Berlin-Ferdinandstraße	04:47	08:11
1158	delivery	Berlin-Glärner	04:55	08:19
1170	delivery	Berlin-Drakestraße	05:01	08:25
2645	delivery	Berlin-Königin-Luise-Str.	05:08	08:32
6198	delivery	Berlin-Hechtgraben	05:12	08:36
6353	delivery	Berlin-Onkel-Tom-Straße	05:22	08:46
6355	delivery	Berlin-Fischerhüttenstraße	05:26	08:50
8455	delivery	Berlin-Lindenthaler	05:31	08:55
1849	delivery	Berlin-Breisgauer	05:36	09:00
2790	delivery	Berlin-Claszeile	05:47	09:11
5576	delivery	Teltow-Kanada	05:57	09:21
8045	delivery	Teltow-Schenkendorfer	07:34	09:33
8517	delivery	Trebbin-Bismarckstraße	08:06	10:05
5045	delivery	Mellensee-Hauptstraße	08:37	10:36
3309	delivery	Blankenfelde-Dahlewitzer	09:02	11:01
	arrival	Berlin-Depot	09:32	13:00



Tab. 1: The task schedule and the optimal round trip of a vehicle routing problem with 15 orders

5 Conclusion, Discussion and Future Work

We developed a web service for emission-reducing vehicle routing in food logistics based on recent research results. First experiments show that the implemented functionalities meet the requirements. Next, there will be an evaluation in practice to check the feasibility of the routes and schedules as well as the validity of the data, e. g., are the journey times and emissions realistic, do they need to be adjusted?

Vehicle routing problems are inherently intractable, so we focus on problems faced by small and medium-sized enterprises to avoid long searches for solutions. In the future we aim to reduce the effort of collecting all data by using other services like openrouteservice.org for duration and distance matrices and use vehicle specific emission factors to compute the emissions from this received data if the real values are unknown. We also plan to generalise our solution approach so that *asymmetric* transition matrices can be considered, too. This is important, e. g., in hilly areas where emissions are different on routes in opposite directions (downhill vs. uphill). Furthermore, we like to examine whether the results in [IR21a; IR21b] or the alternative approach in [DCP16] can help to improve the performance of solution finding in our application context.

Acknowledgement

This work was carried out within the framework of the research and development project Stadt-Land-Fluss. The project is funded by the Federal Ministry of Food and Agriculture based on a resolution of the German Bundestag. Funding Reference Number: 2821KI002.

References

- [Be12] Benchimol, P.; van Hoesel, W.-J.; Régin, J.-C.; Rousseau, L.-M.; Rueher, M.: Improved Filtering for Weighted Circuit Constraints. *Constraints* 17/3, pp. 205–233, July 2012, visited on: 08/09/2022.
- [DCP16] Ducomman, S.; Cambazard, H.; Penz, B.: Alternative Filtering for the Weighted Circuit Constraint: Comparing Lower Bounds for the TSP and Solving TSPTW. *Proceedings of the AAAI Conference on Artificial Intelligence* 30/1, Mar. 2016, visited on: 08/09/2022.
- [DV21] Dalmeijer, K.; Van Hentenryck, P.: Optimizing Freight Operations for Autonomous Transfer Hub Networks. arXiv:2110.12327 [math], Comment: 16 pages, 5 figures, 5 tables, Oct. 2021, arXiv: 2110.12327 [math], visited on: 01/25/2022.
- [He00] Helsgaun, K.: An Effective Implementation of the Lin–Kernighan Traveling Salesman Heuristic. *European Journal of Operational Research* 126/1, pp. 106–130, Oct. 2000, visited on: 05/15/2023.
- [He09] Helsgaun, K.: General K-Opt Submoves for the Lin–Kernighan TSP Heuristic. *Mathematical Programming Computation* 1/2, pp. 119–163, Oct. 2009, visited on: 05/15/2023.
- [He17] Helsgaun, K.: An Extension of the Lin–Kernighan–Helsgaun TSP Solver for Constrained Traveling Salesman and Vehicle Routing Problems: Technical Report. In. Dec. 2017, visited on: 05/15/2023.
- [IR19] Isoart, N.; Régin, J.-C.: Integration of Structural Constraints into TSP Models. In (Schiex, T.; de Givry, S., eds.): *Principles and Practice of Constraint Programming*. Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 284–299, 2019, ISBN: 978-3-030-30048-7.
- [IR21a] Isoart, N.; Régin, J.-C.: A K-Opt Based Constraint for the TSP. In (Michel, L. D., ed.): *27th International Conference on Principles and Practice of Constraint Programming (CP 2021)*. Vol. 210. Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 30:1–30:16, 2021, ISBN: 978-3-95977-211-2, visited on: 08/10/2022.

- [IR21b] Isoart, N.; Régin, J.-C.: A Linear Time Algorithm for the K-Cutset Constraint. In (Michel, L. D., ed.): 27th International Conference on Principles and Practice of Constraint Programming (CP 2021). Vol. 210. Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 29:1–29:16, 2021, ISBN: 978-3-95977-211-2, visited on: 05/18/2023.
- [LK73] Lin, S.; Kernighan, B. W.: An Effective Heuristic Algorithm for the Traveling-Salesman Problem. *Operations Research* 21/2, pp. 498–516, 1973, JSTOR: 169020.
- [Ré94] Régin, J.-C.: A Filtering Algorithm for Constraints of Difference in CSPs. In: AAAI Conference on Artificial Intelligence. [TLDR] A new filtering algorithm is presented that achieves the generalized arc-consistency condition for these non-binary constraints and has been successfully used in the system RESYN, to solve the subgraph isomorphism problem., Aug. 1994, visited on: 05/17/2023.
- [Wo12] Wolf, A.: firstCS—New Aspects on Combining Constraint Programming with Object-Oriented Programming in Java. en, *KI - Künstliche Intelligenz* 26/1, pp. 55–60, Feb. 2012.

“HalloBzar”: A German chatbot for accessing the regional digital marketplace


Vanshika Bawa ¹, Ibrahim Baroud², and Stefan Schaffer³

Abstract: As part of the “Stadt-Land-Fluss” project, we investigate the usage of an AI-powered chatbot to support three key players of the regional food supply chain: producers, food processors, and buyers, in selling or buying their products digitally. A Minimal Viable Prototype was developed with input from domain specialists and includes two scenarios with concrete use cases; namely, “see crop plannings”, “create sales item”, “looking for a food-processor” in the first and, “create delivery series” or “see offers” in the second. Additionally, the chatbot provides personalized follow-up questions in the case where an offer or crop planning is due. We identify 17 intents and 9 entities and evaluate the performance of NLU components using 5-fold cross validation. This work contributes to the field by curating domain-specific use cases and data based on expert insights. We also extend the data using large language models. Additionally, we show that using DIET with spaCy outperforms using DIET with BERT-based tokenizers and featurizers in Intent Detection (ID) and Entity Recognition (ER) on our data when keeping other parts of the chatbot constant.

Keywords: Chatbot; Conversational Design; Pre-trained Large Language Models; Prototyping; Food Supply Market; Regional Food; Deep Learning.

1 Introduction

With the rise in smart food systems [To23, Jo23] and an increase in climate-crisis awareness among agriculturists [RM22], there is a need for innovative solutions that support communities in using new technologies. In the context of regional farming, such efforts can improve technology acceptance, leading to denser farm-to-regional-market linkages and higher efficiency for stakeholders along the food supply chain. This indirectly strengthens the presence of regional economies and bio-practices, factors also found highly important by food shoppers [De20]. At present, digital marketing for wholesale sellers and buyers, such as farmers or school canteens, is becoming more relevant, replacing traditional modes of communication such as fax, email, and telephones [Bi23, Ne23]. To make the transition to digitalized infrastructure more intuitive, natural language processing tools have increasingly gained importance, as they bridge the gap between natural language and software systems. Moreover, chatbot agents can automate the

¹ German Research Centre of Artificial Intelligence (DFKI), Cognitive Assistants, Alt-Moabit 91c, 10559 Berlin, vanshika.bawa@dfki.de,  <https://orcid.org/0009-0008-9585-8902>,

² Ibid, ibrahim.baroud@dfki.de,

³ Ibid, stefan.schaffer@dfki.de,

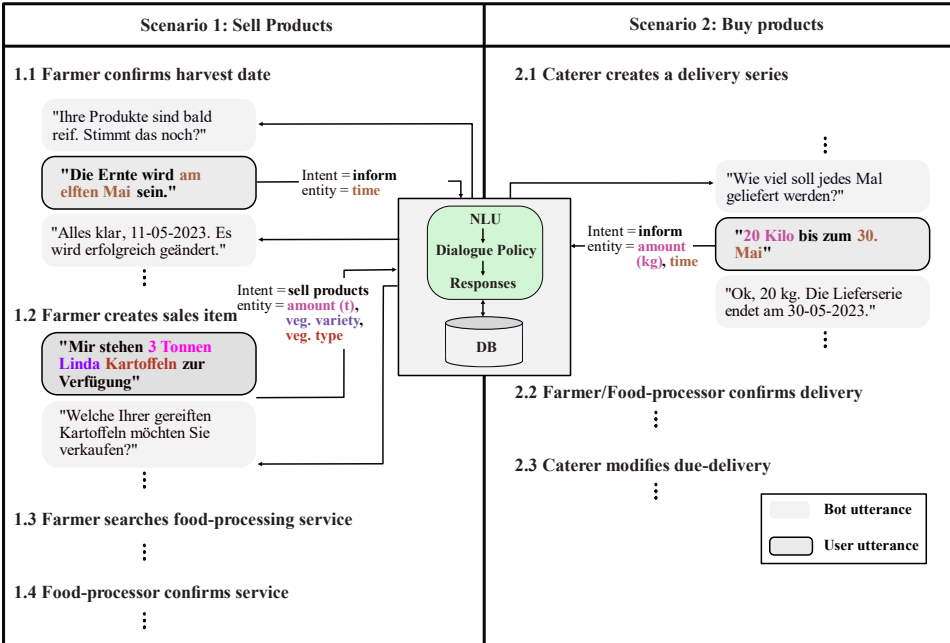
communication process and interact with complex database systems. Other benefits include around-the-clock availability with real-time notifications in a multi-stakeholder environment and low requirement of technological competence from users.

2 Developing the chatbot

Designing a chatbot for the regional food supply system is a multi-level process. It starts with identifying abstract scenarios with potential benefits and conceptualizing conversational designs. Then a Natural Language Processing (NLP) stack, with a suitable configuration for the target language, such as German, is created. The NLP stack consists of various pre-trained components that can often be further fine-tuned for domain-specific terminologies, such as “Kolli” (a unit), “NAPF” (an acronym for Natural Food Deposit System) etc. in a German-language food market. Relevant terms are extracted from conversational data collected from domain experts and potential users through co-designing workshops. The natural language is then decoded in the Natural Language Understanding (NLU) part of the pipeline-based dialogue system, trained component-wise sequentially for respective tasks. Dense featurization refers to the extraction of high-dimensional word representations from words using a mapping provided by pre-trained transformers like spaCy [Ho20] or BERT [De18]. Such word-embeddings are then used for downstream tasks like classifying sentence-level Intent, known as ID, and predicting word-level entity, known as ER. This is followed by a dialogue policy wherein a classifier, such as the transformer based DIETClassifier [Bu20], predicts the next action based on dialogue history and the current user-utterance. Lastly, template-based safe responses can be selected using this policy, circumventing the hallucination problem [Ji23]. Other open-source regex-based systems [Gi21] can parse entities like time, which is expressed heterogeneously in natural language, into a standard format. In task-oriented chatbots, ER—that is, mapping natural language to pre-defined slot values—is vital for success in tasks, e.g., SQL-queries. Further, understanding of dialects becomes crucial with speech input [Hu22]; however, open-source conversational data with annotations for this domain are largely missing, limiting in-context learning and the human-centric experience of chatbots.

To this end, we describe concrete conversational designs called use cases (UC), which are described as concrete dialogue snippets. The database is filled with curated domain-specific dummy data, produced as a part of the Minimal Viable Prototype (MVP) development. In this study, we identified two specific scenarios for the MVP (see Fig. 1). In the first, a farmer has registered their crop plans in the chatbot system and would eventually like to market their raw or processed food products in a digital marketplace. In the second, a catering service that has a contract wants to create a delivery series for the product. Some bot–user interactions and their processing by the Chatbot are shown in Fig. 1. Finally, using template-based responses and constant dialogue policy, we compare the influence of BERT and spaCy NLU components on intent detection and entity recognition for our use cases.

A Two MVP Scenarios with use cases



B Examples of some domain-specific and general intents used

Intent name	Example utterances
Remind later	"Ich möchte zuerst etwas anderes tun."
See crop plans	"Gib mir einen Überblick über meine Anbaupläne."
See offers	"Ich möchte gerne meine Angebote einsehen."
Greet	"Moin"
Change something	"Ich möchte das abändern."
Help	"Ich benötige Hilfe."
Stop	"Habe gerade keine Zeit mehr, mache später weiter."
Goodbye	"tschau"
Deny	"das stimmt nicht"
Affirm	"Ja, das ist Korrekt!"

Fig. 1: A high-level schematic representation of implemented scenarios for different user roles and processing of utterances for intents and entities. **A** shows sequential UCs in each scenario.

Scenario 1 begins with registration of crop plans in the system which enables UC 1.1, where a farmer is notified of the planned harvest of their crop. A ripened crop can be processed via a third party in UC1.3 and registered as a sales item on the marketplace in 1.2. If a processor was chosen, they are notified in UC1.4. In scenario 2, we assume that a caterer has an active contract, which enables them to create a delivery series of an available sales item in UC 2.1 Once confirmed by the contract partner in UC2.2, the caterer is notified when a delivery is due, in case they want to change the amount delivered. **B**: The default paths a user follows without errors, exceptions, or unexpected behaviour are termed happy paths, covered in **A**. In practice, however, users change their mind, want to be reminded later, make mistakes and corrections, etc. Some intents used to capture this behaviour are shown.

3 Methodology

3.1 Data

The NLU dataset used for training and evaluation was created by the authors with feedback from domain experts. More example utterances were added to the NLU dataset by potential users for relevant intents and entities. Furthermore, synthetic data was generated using ChatGPT [In23] and T5 [Ch22] to add more examples involving the challenging “time” entity. We found that this approach improves the performance in extracting the “time” entity. The NLU data include annotated examples about 17 intents and 9 entities. Fig. 2 presents the number of training examples per intent and entity.

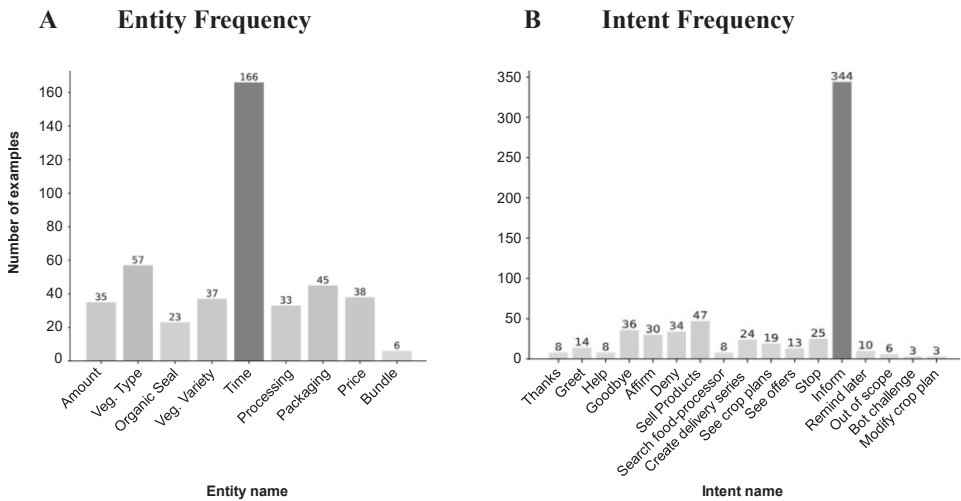


Fig. 2: Number of entity (A) and intent (B) training examples in the NLU training data. More training examples were added to improve the recognition of the time entity, which comes in various forms. The examples for the intent “inform” include all entity annotations and therefore the highest intent frequency.

3.2 Models

For the Intent Detection and Entity Extraction, the Dual Intent Entity Transformer (DIET) is used with RASA 2.8.17 [Bo17]. When provided with pretrained word embeddings, e.g. from pre-trained language models such as BERT, DIET is able to use sparse and dense features to solve the tasks of intent recognition and entity extraction simultaneously. In this work, two NLU configurations are evaluated and compared based on the tokenizers and featurizers. Both configurations use sparse word embeddings, namely; Count Vectors features, Lexical and Syntactic features, and Regex features. For Count Vectors Featurizers, word and character n-grams are used to create the n-gram representations.

Contrastively, the first configuration uses the BERT tokenizer and featurizer while the second configuration uses the spaCy tokenizer and featurizer as dense embeddings. The uncased `de_core_news_sm` model and the `bert-base-uncased` model are used in the spaCy and BERT respectively to obtain the pretrained dense features.

3.3 Evaluation

To study the influence of tokenizers and featurizers in each configuration, we keep the rest of the NLU pipeline components fixed. We optimize the epochs hyperparameter of each configuration by training on 300 epochs and evaluating the performance on ID and ER every 10 epochs. The data was split in train (80%) and test (20%) using random sampling. The optimal number of epochs is selected after considering accuracy and f1 metrics. We finally evaluate the BERT and spaCy configurations using the respective optimal epochs, and cross-validation over five folds.

4 Results

The two introduced configurations using BERT and spaCy tokenization and featurization were used to train on our data using 300 epochs. It was found that the DIET classifier has the best performance on the evaluation dataset when trained for 120 epochs using BERT word-embeddings and 250 epochs using those from spaCy. Next, we evaluated the performance of the DIET classifier using cross validation with five folds and each configuration with the tuned epochs parameters. Tab. 1 shows the achieved results as weighted averages across all intents or entities.

Metric	BERT (epochs=120)		SpaCy (epochs=250)	
	ID	ER	ID	ER
Precision	0.8425	0.8716	0.8290	0.8700
Recall	0.8494	0.8344	0.8393	0.8921
F1-score	0.8437	0.8490	0.8321	0.8794

Tab. 1: DIET performance on evaluation set when using BERT and spaCy embeddings in the NLU pipeline. All values were calculated as weighted averages across the performance of the model on each intent or entity.

The results in Tab. 1 show that the DIET classifier achieves better performance in the ER task when using spaCy. However, it achieves a slightly better performance in ID when using BERT in the NLU pipeline as dense features. To gain more insight about the model performance for each intent and entity in the dataset, Fig. 3 shows the DIET classifier’s confusion matrices for both tasks.

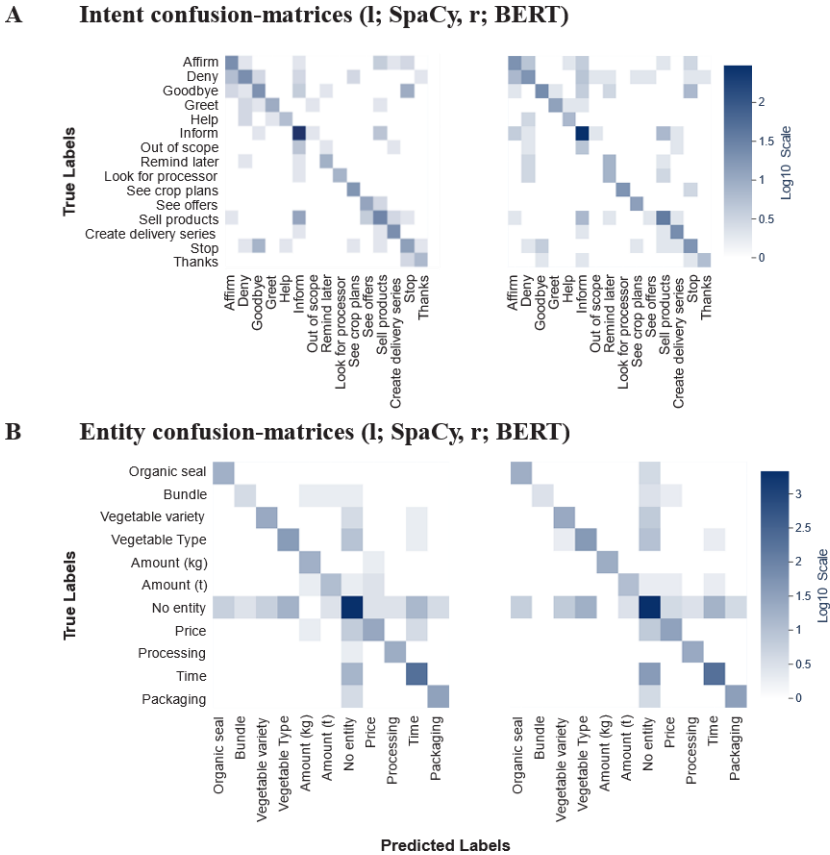


Fig. 3: Y-axis shows true intent (A) and entity (B) labels, and x-axis shows the respective predicted labels. Each element in the matrix is Log_{10} of the number of entities. The diagonals, therefore, represent correctly predicted labels in each confusion matrix.

5 Conclusion and future work

Early testers, who had limited exposure to the chatbot development, were able to successfully use the prototype. The pipeline-based system trained on domain data in this study performs comparably to end-to-end LLMs [Wa23] with in-context learning. Despite the recent success of LLMs, their performance in information extraction tasks remains poor [Ma23] without supervised learning on domain-data. However, as digitalization advances, data sources available for agriculture are also growing. In the future, chatbots can play a vital role in making agriculture more sustainable. In our subsequent work, the chatbot's overall performance will be evaluated in a multi-user setting, and sustainable food recommendations and efficient delivery planning will be integrated.

Bibliography

- [Bi23] bitkom, <https://www.bitkom.org/Presse/Presseinformation/Digitale-Landwirtschaft-soziale-Medien-Hof>, accessed: 24/05/2023.
- [Bo17] Bocklisch, T. et al.: Rasa: Open Source Language Understanding and Dialogue Management. CoRR 2017. arXiv preprint <http://arxiv.org/abs/1712.05181> accessed: 24/05/2023.
- [Bu20] Bunk, T. et al.: DIET: Lightweight Language Understanding for Dialogue Systems. arXiv preprint <https://arxiv.org/abs/2004.09936>, 2020, accessed: 24/05/2023.
- [Ch22] Chung, H. W. et al.: Scaling Instruction-Finetuned Language Models. arXiv preprint <https://arxiv.org/abs/2210.11416>, 2022, accessed: 24/05/2023.
- [De18] Devlin, J. et al.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint <https://arxiv.org/abs/1810.04805>, 2018, accessed: 24/05/2023.
- [De20] German Federal Ministry of Food and Agriculture: Deutschland, Wie Isst: Der BMEL-Ernährungsreport 2020. German Federal Ministry of Food and Agriculture: Berlin, Germany, 2020. https://www.bmel.de/SharedDocs/Downloads/DE/Broschueren/ernaehrungsreport-2020.pdf?__blob=publicationFile&v=19_ accessed: 24/05/2023.
- [Du23] Duckling, <https://github.com/facebook/duckling>, accessed: 24/05/2023.
- [Ho20] Honnibal, M. et al.: spaCy: Industrial-strength Natural Language Processing in Python, 2020. URL: <https://doi.org/10.5281/zenodo.1212303>.
- [Hu22] Hundertmark, S.; Schreiber, F.: IFZ Conversational Insurance Studie 2022. <https://hub.hslu.ch/insuranceinsights/ifz-conversational-insurance-studie-2022/>, accessed: 24/05/2023.
- [In23] Introducing ChatGPT, <https://openai.com/blog/chatgpt>, accessed: 24/05/2023.
- [Ji23] Ji, Z. et al.: Survey of Hallucination in Natural Language Generation. ACM Computing Surveys 12/23, pp. 1-38, 2023.
- [Jo23] John Deere, <https://www.deere.co.uk/en/agriculture/future-of-farming/>, accessed: 24/05/2023.
- [Ma23] Ma, Y. et al.: Large Language Model Is Not a Good Few-shot Information Extractor, but a Good Reranker for Hard Samples!. arXiv preprint <https://arxiv.org/abs/2303.08559>, 2023, accessed: 24/05/2023.
- [Ne23] nearby food, <https://www.nearbuy-food.de>, accessed: 24/05/2023.
- [RM22] Rohleder, B.; Meinel, T.: Die Digitalisierung der Landwirtschaft. Presentation held in Berlin, 12. Mai 2022. <https://www.bitkom.org/sites/main/files/2022-05/Bitkom-Charts%20Landwirtschaft.pdf>, accessed: 24/05/2023.
- [To23] top agrar online, www.topagrar.com/technik/news/farmmanagementsysteme-im-ueberblick-12404106.html, accessed: 24/05/2023.

- [Wa23] Wang, S. et al.: Gpt-ner: Named entity recognition via large language models. arXiv preprint <https://arxiv.org/abs/2304.10428>, 2023, accessed: 24/05/2023.

Drohnenbasiertes Verfahren zur Detektion geschädigter Obstbäume in Obstbaumpflanzungen

Vorstellung von Ergebnissen des Projekts „Digitaler Obstbau“

Bonito Thielert¹, Patrick Menz¹, Sebastian Warnemünde¹, Katharina Holstein², Lauritz Klein¹, David Kiliyas¹, Miriam Runne³, Wolfgang Jarausch³ und Uwe Knauer⁴

Abstract: Phytoplasmen-induzierte Erkrankungen von Obstbäumen stellen eine große Herausforderung im europäischen Obstanbau dar. Apfeltriebsucht und Birnenverfall zählen zu den wirtschaftlich relevantesten Obstkrankheiten. Im Rahmen des Projekts „Digitaler Obstbau“ wurden verschiedene Technologien zur Diagnose der Krankheiten untersucht und weiterentwickelt. Die drohnenbasierte Bonitur der genannten Krankheiten ermöglicht die flächendeckende räumlich hochauflösende Erfassung von Pflanzungen und die Symptomerkennung für einzelne Obstbäume. Untersuchungen mit der Hyperspektralkamera Cubert UHD-185 Firefly und dem integrierten Phantom 4 Multispectral Aufnahmesystem zeigen die Durchführbarkeit der drohnenbasierten digitalen Bonitur. Die Daten zeigen mit 76 % eine gute Klassifikationsrate zur Erkennbarkeit der Krankheitssymptome.

Keywords: Hyperspektral, Multispektral, Digitale Bonitur, Apfeltriebsucht, Birnenverfall

1. Einleitung

Obstbaumpflanzungen sind in Europa bedeutende landwirtschaftliche Flächen und schützenswerte Kulturgüter. Sie sind durch die veränderten klimatischen Bedingungen und die Verbreitung neuer Krankheiten erhöhtem Stress ausgesetzt. Zudem wird gesamtgesellschaftlich neuen und umweltschonenden Methoden zum Pflanzenschutz ein zunehmender Wert beigemessen. Ein schnelles Erkennen von geschädigten oder erkrankten Bäumen ermöglicht frühzeitige Handlungsmaßnahmen. Für ein flächendeckendes und regelmäßiges Monitoring bieten sich Drohnenbefliegungen an [SBI19], [BHS14]. Die entstehenden Aufnahmen erlauben ein großflächiges und systematisches Gesamtbild. Im Obstbau wirtschaftlich relevante Erkrankungen sind die beiden Phytoplasmosen Apfeltriebsucht und Birnenverfall. Beide führen zu einer Kleinfrüchtigkeit und im Verlauf der Krankheit zum Absterben des gesamten Baumes. Daher sind beide Phytoplasmosen von hoher wirtschaftlicher Signifikanz. Zum Schutz der

¹ Fraunhofer IFF, Kognitive Prozesse und Systeme, Sandtorstr. 22, 39106 Magdeburg, bonito.thielert@iff.fraunhofer.de

² Hochschule Anhalt, Fachbereich Informatik und Sprachen, Lohmannstr. 23, 06366 Köthen

³ RLP AgroScience GmbH, Breitenweg 71, D-67435 Neustadt/W

⁴ Hochschule Anhalt, Fachbereich Landwirtschaft, Ökotrophologie, Landschaftsentwicklung, Strenzfelder Allee 28, 06406 Bernburg

Anlagen vor einer weiteren Ausbreitung der Phytoplasmosen können allen voran die befallenen Bäume gerodet oder die relevanten Insektenvektoren bekämpft werden [We22a], [We22b], [TA07], [BD09], [Ba20]. Um den wirtschaftlichen Verlust möglichst gering zu halten, ist eine schnelle Erkennung von befallenen Arealen wünschenswert [Ja13], [BD09], [NV16]. Ein relevantes Symptom der Phytoplasmosen sind spektrale Veränderungen, welche im Spätsommer bis hin in den Herbst auftreten und für spektrale Untersuchungen sehr vielversprechend sind [BGN02], [Mi17]. Korrelierende Symptome wurden bereits bei anderen Phytoplasmosen beobachtet [Be20]. Das drohnengestützte Monitoring kann daher eine wesentliche Unterstützung für die bonitierenden Fachkräfte zur schnellen Phytoplasmaerkennung darstellen.

2. Methoden

1.1 Durchführung der Befliegungskampagne

Voruntersuchungen der Blätter von Apfel- und Birnbäumen im Labor mit Hyperspektralkameras Hyspex VNIR-1600 und Hyspex SWIR-320m-e lieferten unter kontrollierten Bedingungen Hinweise auf relevante Wellenlängenbereiche [Me23], [Ho21]. Diese Untersuchungen bestätigten signifikante Veränderungen des Reflexionsverhalten von Blättern infizierter Bäume, insbesondere im Bereich des Red-Edge und im Bereich des roten Lichts. Dies ist in Abb. 1 am Beispiel der mittleren Reflexionsspektren gesunder und an Birnenverfall erkrankter Birnbäume dargestellt.

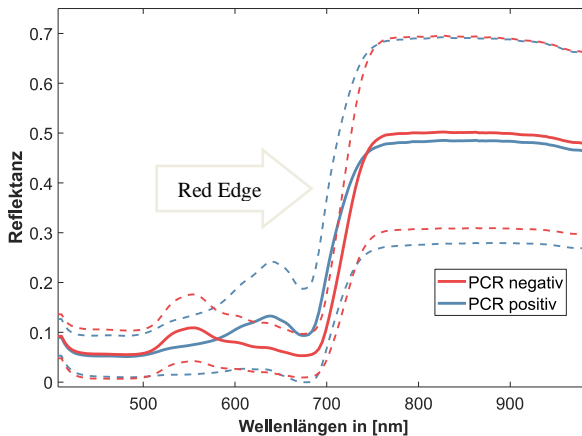


Abb. 1: Visualisierung der Symptomausprägung anhand der gemittelten spektralen Signaturen (durchgezogene Linie) und der zugehörigen Standardabweichung (gestrichelte Linie) von gesunden (blau) und an Birnenverfall erkrankter Birnbäumen (rot). Die Spektren repräsentieren die unter Laborbedingungen aufgenommene Reflexion des Lichts im Bereich 400-1000 nm

Die Drohnenbefliegungen wurden im Spätsommer und Herbst durchgeführt. Es wurden Plantagen mit unterschiedlichen Sorten und einem unterschiedlichen Befallsgrad untersucht. Für die Befliegungen wurden zwei Kamerasysteme ausgewählt, welche eine Untersuchung des relevanten Wellenlängenbereiches erlauben und für den drohnengestützten Einsatz marktverfügbar sind. Tab. 1 stellt die Eigenschaften dieser Kamerasysteme gegenüber. Bei der Kamera UHD-185 des Herstellers Cubert handelt es sich um ein leichtes Hyperspektralkamerasystem. Es wird in Kombination mit einem miniaturisierten PC-System betrieben. Zur radiometrischen Kalibrierung wird vor dem Flug eine Weiß-Referenz und eine Aufnahme mit angebrachtem Objektivdeckel als Dunkelstrom-Messung aufgenommen.

Bei der Multispektralkamera des Drohnensystems P4 Multispectral des Herstellers DJI handelt es sich um eine Kamera mit 6 Sensoren. Neben einem RGB-Sensor mit Bayer-Mosaik dessen Daten im JPEG-Format gespeichert werden, sind 5 als TIFF-speichernde CMOS-Bildsensoren verbaut. Die Beschränkung auf einen bestimmten Spektralbereich, wie in der Tab. 1 dargestellt, erfolgt über fest installierte Filter vor der Kameralinse.

Kamerasysteme	Cubert UHD-185 FireflEYE	DJI P4 Multispectral
	125 Bänder:	6 Bänder:
	450-950 nm	<ul style="list-style-type: none"> • 450 nm ± 16 nm • 560 nm ± 16 nm
Spektrale Auflösung	FWHM 8 nm	<ul style="list-style-type: none"> • 650 nm ± 16 nm • 730 nm ± 16 nm • 840 nm ± 26 nm • RGB-Sensor
	Panchromatischer Sensor mit 1 MP	
Räumliche Auflösung	50x50 Spektren	2 MP
	1 MP interpoliert	
Radiometrische Auflösung	12 Bit	16 Bit

Tab. 1: Beschreibung der eingesetzten Sensoren

Die Kamera UHD-185 entspricht hinsichtlich des Spektralbereiches annähernd der in den Voruntersuchungen eingesetzten Hyperspektralkamera Hypspx VNIR-1600. Die Zahl der spektralen Bänder ist mit 125 geringer und die Halbwertsbreite (FWHM) ist mit 8 nm höher, so dass sich benachbarte Kanäle stärker überlappen. Deutlich geringer ist die räumliche Auflösung mit 2.500 Spektren pro Bild. Durch Pan-Sharping kann die räumliche Auflösung zwar erhöht werden. Dies erreicht jedoch nicht die Qualität eines Systems mit höherer Auflösung.

Das Aufnahmesystem P4 Multispectral hat eine hohe Auflösung von 2 Megapixel pro Kanal. Zudem korrespondiert die spektrale Charakteristik mit den in den

Voruntersuchungen ermittelten relevanten Wellenlängenbereichen des roten Lichts und des Red Edge, in welchen sich Intensitätsunterschiede zwischen gesunden und Phytoplasma-infizierten Blättern als Symptom zeigen [ASC19], [Ba19]. Das Aufnahmesystem ist in das unbemannte Luftfahrtsystem vollständig integriert. Die aufgenommenen Daten werden georeferenziert auf einer SD-Karte gespeichert. Die Flugplanung, die Einstellung des Sensors und die Durchführung des Fluges erfolgten mit der Software Ground Station Pro, die vom Hersteller zum Zeitpunkt der Durchführung kostenfrei für das Apple iPad angeboten wurde. Die radiometrische Kalibrierung erfolgt durch einen auf der Oberseite des Fluggeräts verbauten Helligkeitssensor.

1.2 Erhebung von Referenzdaten

Zeitnah zu den Drohnenflügen erfolgte die Bonitur der Obstplantagen durch Experten. Dabei wurden sowohl gesunde als auch infizierte Bäume und ihre Positionen erfasst. Das Vorliegen einer Infektion wurde durch eine PCR-Untersuchung anhand von Blattproben im Labor überprüft [Lo95], [Ja04]. Somit standen für die weiteren Untersuchungen qualitativ hochwertige Referenzdaten basierend auf einer visueller Bonitur und PCR-Laboruntersuchungen zur Verfügung. Die Daten der visuellen Bonitur und der PCR-Testungen wurden als Excel-Datei neben weiteren Merkmalen wie der ID der Plantage, der Baumreihe und des Baumes zur Verfügung gestellt. Auf diese Weise konnten die Referenzdaten den einzelnen Bäumen in der Spektralaufnahme zugeordnet werden. Weiterhin wurden Bodenkontrollpunkte in Form von 1x1m großen Standardplatten in schwarz/weißer Färbung verwendet und mit RTK-GPS eingemessen. Mit diesen Referenzdaten kann später die Georeferenzierung korrigiert oder überprüft werden.

1.3 Prozessierung der Daten

Für die Daten der Cubert UHD-185 wird mittels der Herstellersoftware ein Datenexport durchgeführt, bei dem die radiometrische Kalibrierung durch Auswertung der Dunkelstrom- und Weißreferenzmessung erfolgt. Die Daten der P4 Multispectral können sowohl als TIFF, als auch als JPEG-Dateien direkt verwendet werden. Die Daten des Umgebungslichtsensors sind in den EXIF-Metadaten enthalten und werden nachträglich zur radiometrischen Korrektur verwendet.

Die Verarbeitung der hyper -und multispektralen Einzelbildern zu orthorektifizierten, georeferenzierten Orthofotos und Höhenmodellen erfolgt mittels Photogrammetrie. Dazu wurde die Software Agisoft Metashape verwendet. Die Einzelbilddaten werden eingeladen. Für die P4 Multispectral Daten wird eine Reflektanzkalibrierung mit dem Helligkeitssensor durchgeführt. Die Reflektanzwerte entsprechen dem Anteil, des von Objekten in der Szene reflektierten Lichts je Wellenlängen. Die wesentlichen Verarbeitungsschritte sind die Identifikation charakteristischer Punkte in den Einzelbildern (Key Point Detection), die Zuordnung zu korrespondierenden Punkten in überlappenden Bildern (Tie Point Detection), die Ableitung von Punktwolken (Sparse und

Dense Point Cloud) und die orthografische Projektion zur Erstellung des georeferenzierten Orthofotos. Eine vollständige Automatisierung ist über Skripte möglich. Durch interaktive Schritte kann das Ergebnis weiter verbessert werden oder eine Georeferenzierung auf Basis der Bodenkontrollpunkte vorgenommen werden. Die orthorektifizierten Karten haben eine Auflösung von bis zu 5 cm pro Pixel für die hyperspektrale S185 Kamera und 2 cm pro Pixel für die multispektrale DJI P4 Multispectral Drohne.

1.4 Maschinelles Lernen

Die Verarbeitung der Spektraldaten erfolgt innerhalb des hauseigenen HawkSpex®Flow ML-Frameworks des Fraunhofer IFF auf Basis von MatlabR2020a. Dieses erlaubt die skriptbasierte Durchführung von Datenimport, Datentransformationen und die Anwendung verschiedener Lernverfahren sowie die Modellevaluation. Zunächst erfolgt eine Vordergrund-Hintergrund-Segmentierung auf Basis des Höhenmodells. Im Vergleich zum Boden höher gelegenen Punkte werden als Baum gelabelt. Das Label dient als Maske für die weitere Prozessierung. Anschließend erfolgt das Labeling der Bäume entsprechend der Referenzdaten und ein entsprechendes Sampling von spektralen Signaturen.

Der so entstandene Datensatz wurde mittels N-facher Kreuzvalidierung und verschiedenen parametrisierten Lernverfahren zur Berechnung von Klassifikationsmodellen genutzt. Die N-fache Kreuzvalidierung diente der Schätzung der Modellgüte und Generalisierungsleistung. Das Vorgehen entspricht einem Auto-ML-Ansatz und erlaubt die adaptive Wahl des jeweils bestgeeigneten Modells. Als Teil der Untersuchungen wurden die spektralen Signaturen unter Zuhilfenahme von spektralen Indizes in verschiedene symptomrelevante Merkmalsräume transformiert.

3. Ergebnisse

Im Rahmen der Klassifikation der hyperspektralen Signatur und daraus abgeleiteter Merkmale wurden die in Tab. 2 dargestellten Ergebnisse erzielt. Dabei konnten anhand der hyperspektralen Daten sowohl für Apfel als auch für Birnenplantagen gute Korrektklassifikationsraten von 76 % erzielt werden.

Korrekt-klassifikationsraten	Apfelplantagen	Birnenplantagen
Spektrale Signatur	60 % ± 2 %	74 % ± 1 %
Mit ℓ^1 -Normierung	76 % ± 7 %	76 % ± 1 %
Mit ℓ^2 -Normierung	76 % ± 7 %	76 % ± 2 %

Tab. 2: Klassifikationsergebnisse für Hyperspektraldaten

Die Abbildung 2 zeigt das Luftbild einer Birnenplantage. In der Darstellung ist das Infektionsgeschehen anhand eines spektralen Index an den P4 Multispectral dargestellt. Spektrale Indizes erlauben eine Transformation der spektralen Signatur in einen neuen

Merkmalsraum, in dem das Infektionsgeschehen besser detektierbar ist. Stark symptomatische Bäume, im Bild als blau dargestellt, werden deutlich sichtbar.

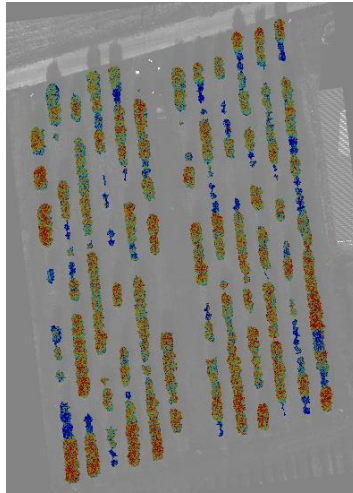


Abb. 2: Visualisierung der Symptomerkennung anhand eines charakteristischen Index einer Birnenplantage. Der Wert des Indizes wird mit einem Farbverlauf von Blau, einem stark symptomatischen Bereich, bis hin zu Rot, mit einer geringen Symptomausprägung, dargestellt.

4. Zusammenfassung und Ausblick

Es konnte ein Vorgehensmodell und eine Auswertepipeline für die drohnengestützte Erfassung von Symptomen wichtiger Obstkrankheiten entwickelt und getestet werden. Es zeigte sich, dass sowohl hyperspektral als auch multispektral erfasste Bilddaten zur Unterscheidung kranker und gesunder Bäume in den Obstplantagen geeignet sind. Der erreichte Stand ist Ausgangspunkt weiterer Untersuchungen. Insbesondere muss geklärt werden, in wie weit parallel auftretende biotische und abiotische Stressfaktoren zu ähnlichen oder gleichen spektralen Signaturen führen. Die weitere Automatisierung der Datenerfassung und -auswertung auf Basis der erstellten Klassifikationsmodelle ist ebenfalls Gegenstand aufbauender Arbeiten. Die Datenbasis soll in Folgeprojekten kontinuierlich ausgebaut werden, um die Anwendung komplexerer Verfahren des maschinellen Lernens zu erlauben.

Acknowledgment

Die Förderung des Vorhabens erfolgte aus Mitteln der landwirtschaftlichen Rentenbank im Rahmen des Projektes digitaler Obstbau.

Literaturverzeichnis

- [SBI19] Stefas, N.; Bayram H.; Isler, V.: Vision-based monitoring of orchards with UAVs. *Computer and Electronics in Agriculture*, 2019.
- [BHS14] Bulanon, D.; Horton, M.; Salvatore, P.; Fallahi, E.: *Apple Orchard Monitoring Using Aerial Multispectral Imaging Environmental Science*, 2014.
- [We22a] Weibel, J.; Buchmann, B.; Bünter, M.; Debonneville, C.; Egger B.; Kehrli, P.: *Apfeltriebsucht: Candidatus Phytoplasma mali. Agroscope Merkblatt 142/2022*, 2022.
- [We22b] Weibel, J.; Buchmann, B.; Bünter, M.; Debonneville, C.; Perren, S.; Kehrli, P.: *Birnenverfall: Candidatus Phytoplasma pyri. Agroscope Merkblatt 156/2022*, 2022.
- [TA07] Tedeschi, R.; Alma, A.: 'Candidatus Phytoplasma mali': the current situation of insect vectors in northwestern Italy. *Bulletin of Insectology*, Bd. 2, pp. 187-199, 2007.
- [BD09] Bertaccini A.; Duduk, B.: *Phytoplasma and phytoplasma diseases: a review of recent research. Phytopathologia Mediterranea*, pp. 355-378, 2009.
- [Ba20] Barthel, D.; Fischnaller, S.; Letschka, T.; Janik, J.; Mittelberger, C.; Öttl, S.; Panassiti, B.; Angeli, G.; Baldessari, M.; Bianchedi, P. L.; Campisano, A.; Covelli, L. T.; Dallago, G.; Ioriatti, C.; Mazzoni, V.; Moser, M.; Pedrazzoli, F.; Rota-Stabelli, O.; Weil, T.; Oppedisano, T.; Anfora, G.; Jarausch, W.; Östereicher, J.; Unterthurner, M.; Schweigkofler, W.; Tedeschi, R.; Schuler, H.: *Scopazzi del melo: stato attuale della ricerca - Apfeltriebsucht: aktueller Stand der Forschung. San Michele all'Adige (TN): Fondazione Edmund Mach; Laimburg: Ora (BZ): Centro di Sperimentazione Laimburg*, 2020.
- [Ja13] Jarausch, W.; Angelini, E.; Eveillard S.; Malembic-Maher, S.: *Management of European fruit tree and grapevine phytoplasma diseases through genetic resistance. Phytopathogenic Mollicutes*, pp. 16-24, 2013.
- [BD09] Bertaccini A.; Duduk, B.: *Phytoplasma and phytoplasma diseases: a review of recent research. Phytopathologia Mediterranea*, pp. 355-378, 2009.
- [NV16] Naghmeh N.; Vadamalai, G.: *Diagnostic techniques for detection of phytoplasma diseases: past and present. Journal of Plant Diseases and Protection*, pp. 16-25, 2016.
- [BGN02] Bertamini, M.; Grando, M.-S.; Nedunchezian, N.: *Effects of Phytoplasma Infection on Pigments, Chlorophyll-Protein Complex and Photosynthetic Activities in Field Grown Apple Leaves. In Biol Plant 47 (2)*, 2002.

- [Mi17] Mittelberger, C.; Pichler, C.; Yalcinkaya, H.; Erhart, T.; Gasser, J.; Schumacher, S.; Janik, K.; Robatscher, P.; Kräutler, B.; Oberhuber, M.: Pathogen-induced chlorosis: Non-fluorescent chlorophyll catabolites (NCCs) related to the pheophorbide a oxygenase (PaO) pathway found in degreened leaves of phytoplasma-infected apple and apricot trees. In *J. Agric. Food Chem.* 65, 2017.
- [Be20] Bendel, N.; Backhaus, A.; Kicherer, A.; Köckerling, J.; Maixner, M.; Jarausch, B.; Biancu, S.; Klück, H.-C.; Seiffert, U.; Voegelé, R.T.; et al. Detection of Two Different Grapevine Yellows in *Vitis vinifera* Using Hyperspectral Imaging. *Remote Sens.*, 2020.
- [Me23] Menz, P.; Warnemünde, S.; Thielert, B.; Kiliyas, D.; Holstein, K.; Runne, M.; Jarausch, W.; Knauer, U.: Hyperspectral Imaging of Leaf Samples and Analysis by Random Forest Classifiers and Artificial Neural Networks. *Plant Methods*, p. draft, 2023.
- [Ho21] Holstein, K.; Knauer, U.; Kiliyas, D.; Runne, M.; Kohler K.; Jarausch, W.: Früherkennung von Apfeltriebsucht und Birnenverfall durch Remote Sensing und Auswertung mit Methoden des maschinellen Lernens In: *Mitteldeutsche Digitaltage im Wein- und Obstbau, Meißen*, 2021.
- [ASC19] Al-Saddik, H.; Simon J.-C.; Cointault, F.: Assessment of the optimal spectral bands for designing a sensor for vineyard disease detection: The case of "Flavescence dorée". *Precisions in Agriculture*, pp. 398-422, 2019.
- [Ba19] Barthel, D.; Fischnaller, S.; Eisenstecken, D.; Kerschbamer, C., Messner, M.; Dordevic, N.; Robatscher, P.; Janik, K.: Near-infrared spectroscopy analysis - A useful tool to detect apple proliferation diseased trees?. *Phytopathog. Mollicutes*, pp. 79-80, 2019.
- [Lo95] Lorenz, K.-H.; Schneider, B.; Ahrens, U.; Seemüller, E.: Detection of the Apple Proliferation and Pear Decline Phytoplasmas by PCR Amplification of Ribosomal and Nonribosomal DNA. *Phytopathology*, pp. 771-776, April 1995.
- [Ja04] Jarausch, W.; Peccerella, T.; Schwind, N.; Jarausch B.; Krczal, G.: Establishment of a quantitative real-time PCR assay for the quantification of apple proliferation phytoplasmas in plants and insects. *Acta-Horticulturae*, pp. 415-420, 2004.

KI-basierte Detektion von Feldhamsterbauen auf landwirtschaftlichen Nutzflächen mittels multi-sensoraler UAS-Daten

F. Thürkow,¹ C. Lorenz,² J. Ramstetter,² I. Hoppe³ und M. Haase³

Abstract: Der zunehmende Druck auf Landwirtschaftsbetriebe, vitale Feldhamsterpopulationen zu erhalten, steht dem regional variierenden Risiko von Ernteaussfällen durch Feldmausfraßschäden gegenüber. Diese Herausforderung ist eng mit den Zielen des Sustainable Development Goals 2 der Vereinten Nationen verknüpft, die Ernährungssicherheit und eine nachhaltige Landwirtschaft anstreben. In dieser Studie wird die Erschließung einer Methode zur automatisierten Detektion von Feldhamsterbauen anhand von Drohnen-basierten RGB- und Thermaldaten untersucht. Zu diesem Zweck werden Faster / Mask R-CNN Objektdetektionsmodelle trainiert. Die Auswertung konzentriert sich auf die Güte dieser, welche anhand unterschiedlicher Flugparameter und einer Feinkartierung validiert werden. Für die RGB-Sensoren Weitwinkel (Brennweite = 4.5 mm) und Zoom (Brennweite = 21-75 mm) kann Faster R-CNN 56 % der Baue erfassen. Mask R-CNN erzielt auf Basis der RGB-Sensoren, sowie des Thermalsensors Trefferquoten von über 80 %. Mit Blick auf die aktuelle und zukünftige Rolle von Naturschutz und Landwirtschaft deuten die Ergebnisse dieser Studie auf einen signifikanten Mehrwert KI-basierter UAS-Ansätze hin.

Keywords: UAS, Künstliche Intelligenz, Objektdetektion, R-CNN, Landwirtschaft, Feldhamster

1 Einleitung

Die Ernährung einer wachsenden Weltbevölkerung bei gleichzeitigem Anspruch an eine umweltschonende und nachhaltige Landwirtschaft stellt Landwirtschaftsbetriebe vor zunehmende Herausforderungen. Die 2015 von den Vereinten Nationen beschlossenen *Sustainable Development Goals 2* (SDG2) beinhalten sowohl eine Null-Hunger-Strategie als auch die Förderung einer nachhaltigen Landwirtschaft [Gi19; Un]. Auf regionaler Ebene sind diese innerhalb der Europäischen Union in der *Gemeinsamen Agrarpolitik* (GAP) verankert und leiten politische Entscheidungen [Di]. Dem Ziel, eine ertragreiche und nachhaltige landwirtschaftliche Produktion zu gewährleisten, können insbesondere Tierfraß bedingte Schäden entgegenwirken. In dem landwirtschaftlich geprägten Flächenland Sachsen-Anhalt gehen diese vornehmlich auf die Feldmaus (*Microtus arvalis*) zurück. In Gradationsjahren belaufen sich diese Ernteverluste auf bis zu 80 % [EWR11; JT10; Lu13]. Eine Besonderheit des Bundeslandes ist, dass sich die Risikoregionen der Feldmausschäden nahezu vollständig

¹ Bern University of Applied Sciences, TI, Biel, Schweiz florian.thuerkow@bfh.ch

² Umwelt- und GeodatenManagement GbR, Halle (Saale), Deutschland info@umgeodat.de

³ Landschaftspflegeverband "Grüne Umwelt" e. V. info@lpv-gruene-umwelt.de

mit dem regionalen Vorkommensgebiet des Feldhamsters (*Cricetus cricetus*) decken [Bua; Wo]. Der Feldhamster ist weltweit vom Aussterben bedroht und gilt als streng geschützte Art, die gemäß § 44 BNatschG weder gestört noch getötet werden darf [Ba19; Bub]. Dieser Schutzstatus des Feldhamsters macht es erforderlich, dass bei massiven Feldmausbefall nur dann gegen diese mit Rodentiziden vorgegangen werden kann, wenn das Vorkommen des Feldhamsters nachweislich ausgeschlossen ist [Ba19; Bua; Mia]. Zu den gängigsten Erfassungsmethoden gehören Feinkartierungen, die eine vollständige Begehung des betroffenen Schlages erfordern [KKW01]. Die Anwendung unbemannter Luftfahrtsysteme (UAS) in Kombination mit Künstlicher Intelligenz (KI) könnte als vielversprechende Technologien Abhilfe schaffen [EI19; PR18]. Der Einsatz von UAS (Drohnen) erfolgt zunehmend auch in der Landwirtschaft [Ch18; Gr22; Ju21]. Aktuelle Anwendungsgebiete sind die Überwachung der Ernte, die Präzisionslandwirtschaft, die Schädlings- und Krankheitsbekämpfung, die Aussaat, sowie der Natur- und Artenschutz [Ch18; Ta20]. Darüber hinaus ermöglichen Thermalmessungen die Überwachung der Boden- und Pflanzentemperatur sowie des Pflanzenstresses [Pa17]. Die Integration solcher Ansätze in KI-Modelle ermöglicht eine präzise und zeitnahe Überwachung und Analyse landwirtschaftlicher Flächen. Dies trägt wiederum zu einer verbesserten Entscheidungsfindung und letztlich zu einer nachhaltigeren und produktiveren Landwirtschaft bei [BMEL]. Im Sinne der Feldhamsterdetektion existiert jedoch noch keine hinreichende Lösung. Eine automatisierte, KI-gestützte Erkennung von Feldhamsterbauen auf landwirtschaftlichen Nutzflächen unter Verwendung von UAS-RGB- und Thermaldaten könnte dazu einen wichtigen Beitrag leisten. Hier knüpft unsere Arbeit im Projekt CRIFORA an [CR].

In der hier vorgestellten Studie verwenden und vergleichen wir zwei Ansätze des maschinellen Lernens (ML): Faster R-CNN und Mask R-CNN, die für die KI-gesteuerte, bildbasierte Objekterkennung genutzt werden [Gi14; Re15]. Ziel unserer Arbeit ist die automatisierte Detektion von Feldhamsterbauen anhand dieser ML-Ansätze auf Grundlage von RGB- und Thermalorthomosaik im Kontext einer nachhaltigen und ressourcenschonenden Landwirtschaft.

2 Material und Methoden

2.1 Untersuchungsgebiet und Datengrundlage

Das Untersuchungsgebiet (Abb. 1a) konzentriert sich auf eine Fläche von insgesamt 0,6 ha in der Nähe von Prosigk (Landkreis Anhalt-Bitterfeld, Sachsen-Anhalt). Der dort ansässige landwirtschaftliche Betrieb wirtschaftet pfluglos bzw. im Direktsaatverfahren. Der Boden des Standortes wird der Tschernosem-Lössbörde zugeordnet [Mib; Mi99]. Auf dieser nährstoffreichen Schwarzerde wurde im Untersuchungszeitraum Winterweizen angebaut – die anbaustärkste Kultur Sachsens-Anhalts [La]. Das Gebiet ist aufgrund abiotischer und biotischer Faktoren ein geeignetes Habitat für Feldhamster und Feldmäuse. Feldhamsterbaue lassen sich in ihrer Form und ihrem Verteilungsmuster von Feldmausbauen gut differenzieren.

Im Vergleich sind die Eingänge von Feldhamsterbauen meßbar größer (bspw. 8 statt 3 cm im Durchmesser) und tiefer angelegt.



Abb. 1: a) Räumliche Verortung des Untersuchungsgebiets. b) Loch mit Fraßkreis (Zoomfoto: 120 m)

Die Drohnenbefliegungen mit parallel stattfindender Feinkartierung des betroffenen Schlages erfolgten am 24.04. und 05.05.2023. Zu diesem Zeitpunkt waren bereits Feldhamster aktiv. Geöffnete Fallröhren zeichneten sich durch Fraßkreise aus (Abb. 1b). Für die Drohnenbefliegung nutzten wir die *DJI Matrice 30 T* (M30T) in Höhen von 12, 30, 50, 100 und 120 m bei einer Fluggeschwindigkeit von 1 m/s. Die M30T verfügt über einen Weitwinkel- (W), einen Zoom- (Z) und einen Thermalsensor (T). Bei dem W-Sensor handelt es sich um einen 12 MP 1/2" CMOS-Scanner, mit einer Brennweite von 4.5 mm, einer f/2.8-Blende und einer maximalen Bildauflösung von 4000x4000 px [DJ]. Der Z-Sensor setzt auf einen 48 MP 1/2" CMOS-Scanner, mit einer variablen Brennweite zwischen 21 und 75 mm, f/2.8 bis f/4.2 für die Blende und einer maximalen Auflösung von 8000x6000 px [DJ]. Der T-Sensor ist ein ungekühltes VO_x-Mikrobolometer mit einer Brennweite von 9.1 mm, einer f/1.0-Blende und einer maximalen Auflösung von 640x512 px mit Messgenauigkeiten von ± 2 °C [DJ]. Die Feinkartierung wurde unmittelbar im Anschluss an die Befliegungen durchgeführt. Die gefundenen Feldhamsterbaue maßen wir mit einem differentiellen GNSS (ppm 10xx GNSS-Sensor) zentimetergenau ein.

Höhe [m]	Flugdauer [min]	Gebiet [ha]	GSD T [cm/px]	GSD W [cm/px]	GSD Z [cm/px]	W	Z	T
12	88,27	0,6	1,58	0,43	-	x	-	x
30	28,29	0,6	3,96	1,07	0,11	x	x	x
50	17,43	0,6	6,59	1,78	0,19	x	x	-
100	9,20	0,6	13,19	3,56	0,38	x	x	-
120	11,8	0,6	15,82	4,27	0,46	x	x	-

Tab. 1: Überblick der Flugparameter und eingesetzter Sensoren. *GSD* Ground Sampling Distance, *W* Weitwinkel, *Z* Zoom, *T* Thermal

Die Validierung der Güte der KI-Ansätze auf Grundlage dieser Befliegungen und der Feinkartierung erfordert bereits trainierte Modelle. Die in dieser Studie eingesetzten KI-Modelle trainierten wir anhand von über 2.200 Bildern, die zwischen *Juni 2022* und *März*

2023 an drei Standorten unterschiedlicher Landschaftseinheiten Sachsen-Anhalts aggregiert wurden (Prosigk, Eilsleben, Wanzleben). Zirka 2.000 dieser Luftbildaufnahmen basierten auf den Sensoren W und Z und gingen in das Training der RGB-KI-Modelle (Faster und Mask R-CNN) ein. Auf das Thermalmodell (Mask R-CNN) entfielen ca. 200 Bilder.

2.2 Gewählter Ansatz des maschinellen Lernens

Für das Vorhaben der automatisierten Erfassung von Feldhamsterstrukturen wurden bildbasierte ML-Ansätze zur Objekterkennung gewählt: Faster R-CNN und Mask R-CNN (ResNet-50), deren Implementierung mit Tensorflow⁴ (TF2) und Detectron2⁵ (D2) erfolgte. Beide stellen Weiterentwicklungen der Ansätze CNN, R-CNN und Fast R-CNN dar, die in Kontexten der Gesichtserkennung und des autonomen Fahrens entwickelt wurden [Gi13; Gi15; ON15]. Das Ziel dieser ist es, anhand von Bildmaterial festgelegte Objektklassen automatisiert zu erfassen [Gi14; Re15]. Faster R-CNN und Mask R-CNN setzen auf zwei zentrale Verfahren: einem Regional Proposal Network (RPN) und einem Convolutional Neural Network (CNN) [Re15]. Das RPN ist dem CNN vorgeschaltet. Auf Grundlage eines eingegangenen Bildes werden durch das RPN grobe rechteckige Objektvorschläge, auch Regions of Interest (ROI) genannt, bestimmt und diese an das CNN übermittelt [He17; Re15]. Das CNN extrahiert Merkmale dieser Objektvorschläge und führt Klassifikationen und Regressionen basierend auf diesen aus. Die Ausgabe ist ein Bild mit Bounding Boxes, die das detektierte Objekt umrahmen und die Objektklasse samt Klassifikationsgüte ausgeben [Gi14; He17; Re15]. Mask R-CNN erweitert diesen Ansatz um ein Fully Convolutional Network (FCN), das parallel zum CNN geschaltet ist (Abb. 2) [He17]. Im Gegensatz zum CNN lokalisiert und klassifiziert das FCN Objekte auf Basis der ROIs pixelbasiert. So ist eine exakte Größenbestimmung der detektierten Objekte möglich [WZL19]. Um eine objektbasierte Detektion von Feldhamsterbauen zu erzielen, müssen die ML-Ansätze auf Basis solcher Datensätze trainiert werden. Für Faster R-CNN wurden die auf den Trainingsbildern sichtbaren Hamsterbaue mit Bounding Boxes gelabelt, wohingegen wir für Mask R-CNN pixelgenaue Polygonmasken verwendeten. Selbiges gilt für das Training des Thermal Mask R-CNN Modells. Hier generierten wir aus den Rohbilddaten zunächst Orthomosaik, um die in den Randbereichen ausgeprägten Temperaturverzerrungen zu umgehen.

2.3 Methode

Das Verfahren zur Anwendung und Auswertung der KI-Modelle begann mit der Datenerhebung. Diese setzte sich aus Feldkartierungen zur Sammlung von Ground Truth Daten und parallelen UAS-Flügen zusammen. Die gesammelten UAS-Daten wurden mit der Software

⁴ <https://www.tensorflow.org/>

⁵ <https://github.com/facebookresearch/detectron2>

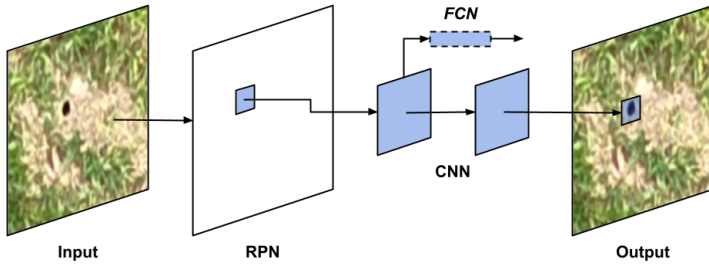


Abb. 2: Grundaufbau von FASTER R-CNN (ohne FCN) und Mask R-CNN (mit FCN)

Pix4DMapper zu Orthomosaiken verarbeitet (vgl. Tab. 1). Diese boten eine vollständige und detaillierte Darstellung des untersuchten Gebiets und bildeten die Grundlage für die spätere räumliche Verortung der erfassten Objekte sowie für die nachfolgenden KI-basierten Analysen. Um die Berechnungsintensität des Detektionsprozesses zu verringern und die KI-Algorithmen effizienter zu gestalten, wurden die Orthomosaik in individuelle Kacheln zerteilt. Hierbei betrug die Größe der Kacheln für Weitwinkel- und Thermaldaten 500x500 Pixel, während für Zoom-Kamera-Daten Kacheln mit einer Größe von 1000x1000 Pixel verwendet wurden. Diese Kacheln wurden an die trainierten Modelle übergeben, welche anhand dieser eine Reihe von Bounding-Boxen bzw. Polygonmasken erstellten, die potenzielle Feldhamsterbaue darstellten, und deren Genauigkeitsbewertung bestimmten. Für die Detektionen wurden spezifische Grenzwerte⁶ für die Genauigkeitsbewertung festgelegt. Detektionen, die diese Grenzwerte nicht erreichen konnten, wurden nicht ausgegeben. Das Festlegen dieser spezifischen Grenzwerte erfolgte auf Grundlage der Datenqualität und Bodenauflösung (GSD).

3 Ergebnisse und Diskussion

Die Feinkartierung ermittelte die Standorte von 16 Feldhamsterbauen. Tab. 2 fasst die Detektionsergebnisse aller Sensoren und Flughöhen zusammen. Die höchste Anzahl an Detektionen, die mit der Feinkartierung übereinstimmen, gehen auf Mask R-CNN mit 14 (Weitwinkel: 12 m), 12 (Thermal: 12 m) und 11 (Zoom: 100 m) zurück. FASTER R-CNN ermittelte im Vergleich für *Weitwinkel: 12 m* und *Zoom: 100 m* 9 Baue. Falsch-positive Detektionen traten für FASTER R-CNN nur marginal auf (Maximalwert = 5 bei *Weitwinkel: 12 m*). Demgegenüber stehen die Werte von Mask R-CNN, die bei niedrigen Flughöhen zwischen 13 - 39 betragen.

Abb. 3 stellt die Detektionen von *Thermal: 12 m* und *Zoom: 100 m* gegenüber. Dargestellt werden die tatsächlich detektierten Löcher, um einen besseren Vergleich zwischen diesen und der Feinkartierung zu ermöglichen. Beide Sensoren erzeugen vergleichbare Detektionen.

⁶ RGB-Weitwinkel = 50 % | RGB-Zoom = 60 % | Thermal = 25 %

Höhe [m]	Weitwinkel [RGB]						Zoom [RGB]						Thermal		
	Mask R-CNN			Faster R-CNN			Mask R-CNN			Faster R-CNN			Mask R-CNN		
	r	z	f	r	z	f	r	z	f	r	z	f	r	z	f
12	14	6	39	9	5	5	-	-	-	-	-	-	12	4	13
30	8	1	1	4	0	0	-	-	-	-	-	-	0	0	0
50	6	1	1	4	1	0	6	1	21	6	0	0	-	-	-
100	0	0	0	0	0	0	11	1	9	9	0	2	-	-	-
120	0	0	0	0	0	1	9	7	2	9	5	0	-	-	-

Tab. 2: Detektionsergebnis der KI-Modelle.

(Legende: r = Detektiertes Loch; von der Feinkartierung erfasst | z = Detektiertes Loch; von der Feinkartierung nicht erfasst | f = Falsch-positiv; ein Objekt wurde fälschlicherweise als Loch detektiert.)

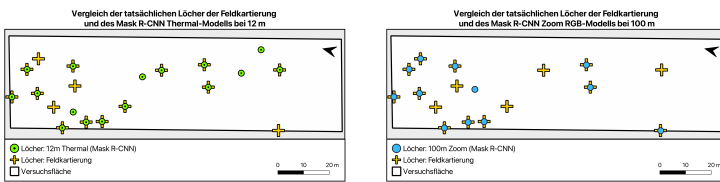


Abb. 3: Gegenüberstellung der tatsächlich kartierten und detektierten Löcher der Modelle Mask R-CNN Thermal für 12 m und Mask R-CNN Zoom (RGB) für 100 m.

Diese Studie konnte zeigen, dass eine KI-basierte, automatisierte Detektion von Feldhamsterbauen anhand von UAS RGB- und Thermalbildern möglich ist und vielversprechende Ergebnisse erzielen kann. Die von uns eingesetzten Modelle, denen die im Rahmen dieser Studie analysierten Daten nicht bekannt waren, sind je nach Flugparameter und Sensor in der Lage, bis zu 87,5 %⁷ der kartierten Baue zu erfassen. Zusätzlich können Baue detektiert werden, die die Feinkartierung nicht erfasste. Diese Erkenntnis deckt sich mit anderen Studien [Co21]. Unsere KI-Modelle wurden größtenteils mit Daten einer anderen Vegetationsperiode trainiert. Dies kann als eine Ursache der aktuell noch unzureichenden Ergebnisqualität betrachtet werden [Hu21].

4 Zusammenfassung und Ausblick

Weitere Ansätze, um die Fehlerquote der Mask R-CNN Ergebnisse zu verringern, liegen in der Anwendung von Denoising-Algorithmen oder der Sensor-Fusion der UAS-Sensoren, die innerhalb des Projektes bereits vorbereitet werden [Ha17; Na20; SZL19]. Die bisherigen Ergebnisse zeigen jedoch schon jetzt das große Potenzial KI-basierter und Drohnen-gestützter Befliegungen landwirtschaftlicher Flächen an - insbesondere im Hinblick auf die Umsetzung der SDG2- und GAP-Ziele, die eine zeitnahe und fallspezifische Festlegung von Strategien benötigen.

⁷ Mask R-CNN: Zoom 100m

Literatur

- [Ba19] Banaszek, A.; Bogomolov, P.; Feoktistova, N.; La Haye, M.; Monecke, S.; Reiners, T.; Rusin, M.; Surov, A.; Weinhold, U.; Ziomek, J.: IUCN Red List of Threatened Species: *Cricetus cricetus*. IUCN Red List of Threatened Species/, Juni 2019, URL: <https://www.iucnredlist.org/en>, Stand: 25. 05. 2023.
- [BMEL] Bundesministerium für Ernährung und Landwirtschaft: BMEL fördert Projekte zum Einsatz von Künstlicher Intelligenz in der Landwirtschaft und den ländlichen Räumen, URL: <https://www.bmel.de/DE/themen/digitalisierung/kuenstliche-intelligenz.html>, Stand: 25. 05. 2023.
- [Bua] Bundesamt für Naturschutz: Kombinierte Vorkommen- und Verbreitungskarte der Pflanzen- und Tierarten der FFH-Richtlinie - 1339 *Cricetus cricetus* (Feldhamster), Stand: 03. 05. 2023.
- [Bub] Bundesministerium für Justiz: § 44 Vorschriften für besonders geschützte und bestimmte andere Tier- und Pflanzenarten, Gesetz über Naturschutz und Landschaftspflege, URL: https://www.gesetze-im-internet.de/bnatschg_2009/_44.html, Stand: 14. 04. 2023.
- [Ch18] Christen, M.; Guillaume, M.; Jablonowski, M.; Moll, K.: Zivile Drohnen-Herausforderungen und Perspektiven. vdf, Zürich, 2018, ISBN: 978-3-7281-3893-4.
- [Co21] Cox, T. E.; Matthews, R.; Halverson, G.; Morris, S.: Hot stuff in the bushes: Thermal imagers and the detection of burrows in vegetated sites. *Ecology and Evolution* 11/11, S. 6406–6414, Juni 2021, ISSN: 2045-7758, 2045-7758, URL: <https://onlinelibrary.wiley.com/doi/10.1002/ece3.7491>, Stand: 27. 05. 2023.
- [CR] CRIFORA: CRICetus FORAMins Radar Detektion, URL: <https://www.crifora.de/>, Stand: 03. 05. 2023.
- [Di] Directorate General for Agriculture and Rural Development: Key policy objectives of the CAP 2023-27, URL: https://agriculture.ec.europa.eu/common-agricultural-policy/cap-overview/cap-2023-27/key-policy-objectives-cap-2023-27_en#:~:text=Key%20objective%3A%20contribute%20to%20climate,well%20as%20promoting%20sustainable%20energy., Stand: 21. 05. 2023.
- [DJ] DJI: Specs - Matrice 30 Series - DJI Enterprise, URL: <https://enterprise.dji.com/matrice-30/specs>, Stand: 25. 05. 2023.
- [EI19] Eli-Chukwu, N. C.: Applications of Artificial Intelligence in Agriculture: A Review. *Engineering, Technology & Applied Science Research* 9/4, S. 4377–4383, Aug. 2019, ISSN: 1792-8036, 2241-4487, URL: <http://etasr.com/index.php/ETASR/article/view/2756>, Stand: 25. 05. 2023.

- [EWR11] Eggert, J.; Wolff, C.; Richter, K.: Searching for alternative methods for a sustainable population management of the common vole (*Microtus arvalis*) in Saxony-Anhalt. en, Julius-Kühn-Archiv; 432; 154-155/, Artwork Size: p. 154-155 ISBN: 9783930037827 Medium: application/pdf Publisher: Julius Kühn-Institut, p. 154–155, Okt. 2011, URL: https://www.openagrar.de/receive/openagrar_mods_00083329, Stand: 25.05.2023.
- [Gi13] Girshick, R. B.; Donahue, J.; Darrell, T.; Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. CoRR abs/1311.2524/, 2013, arXiv: 1311.2524, URL: <http://arxiv.org/abs/1311.2524>.
- [Gi14] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J.: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. ISSN: 1063-6919, S. 580–587, Juni 2014.
- [Gi15] Girshick, R. B.: Fast R-CNN. CoRR abs/1504.08083/, 2015, arXiv: 1504.08083, URL: <http://arxiv.org/abs/1504.08083>.
- [Gi19] Gil, J. D. B.; Reidsma, P.; Giller, K.; Todman, L.; Whitmore, A.; van Itersum, M.: Sustainable development goal 2: Improved targets and indicators for agriculture and food security. en, Ambio 48/7, S. 685–698, Juli 2019, ISSN: 1654-7209, URL: <https://doi.org/10.1007/s13280-018-1101-4>, Stand: 25.05.2023.
- [Gr22] Grenzdörffer, G.: UAV in der Landwirtschaft - Eine Übersicht aktueller Anwendungen und zukünftiger Potenziale./, S. 103–126, März 2022.
- [Ha17] Ha, Q.; Watanabe, K.; Karasawa, T.; Ushiku, Y.; Harada, T.: MFNet: Towards real-time semantic segmentation for autonomous vehicles with multi-spectral scenes./, S. 5108–5115, 2017.
- [He17] He, K.; Gkioxari, G.; Dollar, P.; Girshick, R.: Mask R-CNN. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, Venice, S. 2980–2988, Okt. 2017, ISBN: 978-1-5386-1032-9, URL: <http://ieeexplore.ieee.org/document/8237584/>, Stand: 25.05.2023.
- [Hu21] Hu, C.; Sapkota, B. B.; Thomasson, J. A.; Bagavathiannan, M. V.: Influence of Image Quality and Light Consistency on the Performance of Convolutional Neural Networks for Weed Mapping. Remote Sensing 13/11, S. 2140, 29. Mai 2021, ISSN: 2072-4292, URL: <https://www.mdpi.com/2072-4292/13/11/2140>, Stand: 27.05.2023.
- [JT10] Jacob, J.; Tkadlec, E.: Rodent outbreaks in Europe: dynamics and damage. In: Rodent Outbreaks: Ecology and impacts. Journal Abbreviation: Rodent Outbreaks: Ecology and impacts, S. 207–223, Jan. 2010.

- [Ju21] Jung, J.; Maeda, M.; Chang, A.; Bhandari, M.; Ashapure, A.; Landivar-Bowles, J.: The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. en, *Current Opinion in Biotechnology, Food Biotechnology Plant Biotechnology* 70/, S. 15–22, Aug. 2021, ISSN: 0958-1669, URL: <https://www.sciencedirect.com/science/article/pii/S0958166920301257>, Stand: 25. 05. 2023.
- [KKW01] Köhler, A.; Kayser, M.; Weinhold, U.: Methoden zur Kartierung von Feldhamstern (*Cricetus cricetus*) und empfohlener Zeitbedarf. In: *Jahrbücher des Nassauischen Vereins für Naturkunde* – 122. S. 215–216, 2001, URL: https://www.zobodat.at/pdf/Jahrbuch-Nassauischen-Verein-Naturkunde_122_0215-0216.pdf.
- [La] Landesanstalt für Landwirtschaft und Gartenbau Sachsen-Anhalt: Winterweizen - Hinweise zur Sortenwahl 2022, URL: https://llg.sachsen-anhalt.de/fileadmin/Bibliothek/Politik_und_Verwaltung/MLU/LLFG/Dokumente/04_themen/sortenpruefung/sorten_hinweise/2022/wiwe.pdf, Stand: 04. 05. 2023.
- [Lu13] Luque-Larena, J.; Mougeot, F.; Viñuela, J.; Arroyo, B.; Jareño, D.; Arroyo, L.; Lambin, X.: Recent large-scale range expansion and outbreaks of the common vole (*Microtus Arvalis*) in NW Spain. *Basic and Applied Ecology* in press/, Juli 2013.
- [Mia] Ministerium für Umwelt, Landwirtschaft und Energie Sachsen-Anhalt: Ausschlussgebiete für den Einsatz von Rodentiziden mit dem Wirkstoff Zinkphosphid (Pflanzenschutz) in Sachsen-Anhalt im Zeitraum 1. März bis 31. Oktober, URL: <https://www.isip.de/isip/servlet/resource/blob/305610/5266c8459fe65f0c868d8c3aad6cbb3/rodentizide-nt820-1-data.pdf>, Stand: 12. 05. 2023.
- [Mib] Ministerium für Wissenschaft, Energie, Klimaschutz und Umwelt (MWU) Sachsen-Anhalt: Vorläufige Bodenkarte von Sachsen-Anhalt 1:50000, URL: <https://metaver.de/trefferanzeige?docuuid=841cc2f3-5c07-4f19-a4e8-002e51ab6a39>, Stand: 25. 05. 2023.
- [Mi99] Michael Weller: *Bodenatlas Sachsen-Anhalt*. Geologisches Landesamt Sachsen-Anhalt, Halle (Saale), 1999.
- [Na20] Nayan, A.-A.; Saha, J.; Raqib Mahmud, K.; Kalam Al Azad, A.; Golam Kibria, M.: Detection of Objects from Noisy Images./, S. 1–6, 19. Dez. 2020, URL: <https://ieeexplore.ieee.org/document/9350521/>, Stand: 27. 05. 2023.
- [ON15] O’Shea, K.; Nash, R.: An Introduction to Convolutional Neural Networks. CoRR abs/1511.08458/, 2015, arXiv: 1511.08458, URL: <http://arxiv.org/abs/1511.08458>.

- [Pa17] Park, S.; Ryu, D.; Fuentes, S.; Chung, H.; Hernández-Montes, E.; O'Connell, M.: Adaptive Estimation of Crop Water Stress in Nectarine and Peach Orchards Using High-Resolution Imagery from an Unmanned Aerial Vehicle (UAV). en, *Remote Sensing* 9/8, S. 828, Aug. 2017, ISSN: 2072-4292, URL: <http://www.mdpi.com/2072-4292/9/8/828>, Stand: 25.05.2023.
- [PR18] Patrício, D. I.; Rieder, R.: Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. en, *Computers and Electronics in Agriculture* 153/, S. 69–81, Okt. 2018, ISSN: 0168-1699, URL: <https://www.sciencedirect.com/science/article/pii/S0168169918305829>, Stand: 25.05.2023.
- [Re15] Ren, S.; He, K.; Girshick, R.; Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: *Advances in Neural Information Processing Systems*. Bd. 28, Curran Associates, Inc., 2015, URL: https://papers.nips.cc/paper_files/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abstract.html, Stand: 25.05.2023.
- [SZL19] Sun, Y.; Zuo, W.; Liu, M.: RTFNet: RGB-Thermal Fusion Network for Semantic Segmentation of Urban Scenes. *IEEE Robotics and Automation Letters* 4/3, S. 2576–2583, 2019.
- [Ta20] Talaviya, T.; Shah, D.; Patel, N.; Yagnik, H.; Shah, M.: Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. en, *Artificial Intelligence in Agriculture* 4/, S. 58–73, Jan. 2020, ISSN: 2589-7217, URL: <https://www.sciencedirect.com/science/article/pii/S258972172030012X>, Stand: 25.05.2023.
- [Un] United Nations: Transforming our world: the 2030 Agenda for Sustainable Development | Department of Economic and Social Affairs, URL: <https://sdgs.un.org/2030agenda>, Stand: 25.05.2023.
- [Wo] Wolff, C.: Feldmausschäden im Ackerbau- Erfahrungen aus Sachsen-Anhalt. Präsentation. Feldtag der Gesellschaft für konservierende Bodenbearbeitung, URL: <https://www.gkb-ev.de/publikationen/2015/15-12-02-mause/feldtag-15-12-02-wolff.pdf>, .
- [WZL19] Wei, Z.; Zhang, B.; Liu, P.: Object Dimension Measurement Based on Mask R-CNN. In: *Intelligent Robotics and Applications: 12th International Conference, ICIRA 2019, Shenyang, China, August 8–11, 2019, Proceedings, Part IV*. Springer-Verlag, Berlin, Heidelberg, S. 320–330, 8. Aug. 2019, ISBN: 9783030275372, URL: https://doi.org/10.1007/978-3-030-27538-9_27, Stand: 27.05.2023.

Vertical-Farming-System für die Bilddatengenerierung zur KI-gestützten Identifikation des Wachstumszentrums von Beikräutern

Erkenntnisse aus dem Projekt KIdetect gefördert durch das BMEL

Philipp Flierl¹, Alexander Zimmermann², Michael Niedermeier², Faryal Noori², Erich Fuchs² und Anton Schmailzl¹

Abstract: Ziel des Projekts KIdetect ist der Aufbau eines Vertical-Farming-Systems zur gezielten Anpflanzung von Kulturpflanzen und Beikräutern in einer Laborumgebung. Anhand der Laborumgebung soll verstanden werden, inwiefern eine sichere Identifikation von Beikräutern möglich ist. Neben den Eigenschaften der Pflanzen im menschlich sichtbaren Bereich des elektromagnetischen Spektrums sollen auch Informationen aus dem Kurzwelleninfrarotbereich evaluiert werden. Die Laborumgebung dient neben der Grundlagenforschung von Unterscheidungsmerkmalen schließlich auch der Generierung fotografischer Aufnahmen von Kulturpflanzen und Beikräutern. Es soll eine Bilddatenbank entstehen, anhand welcher die Grundlagen für eine spätere Anwendung der Erkenntnisse auf die Anwendungsumgebung geschaffen werden soll. Der Fokus der Beikräuter liegt auf dem Wachstumszentrum und soll gestützt durch KI zuverlässig identifiziert werden.

Keywords: Vertical-Farming-System, SWIR-Technologie, Unkrautererkennung, Bilddatengenerierung, Laborumgebung

1 Einleitung

Mit dem „European Green Deal“ forcieren die Mitgliedstaaten der Europäischen Union, den Einsatz von Pflanzenschutzmittel bis 2030 zu halbieren [UB22]. Diese Reduktion der Pflanzenschutzmittel kann nur mit einer deutlichen Verbesserung des aktuellen Stands der Pflanzenschutztechnik realisiert werden. Anstatt einer breitflächigen Ausbringung von Herbiziden muss zukünftig eine selektive Beikrautregulierung erfolgen. Für diese selektive Beikrautregulierung muss vollautonom zwischen Beikraut und Kulturpflanze unterschieden werden können. Für das Trainieren entsprechender Algorithmen werden größere Bilddatensätze benötigt, welche nach aktuellem Stand der Technik in der Anwendungsumgebung gewonnen werden. Hierbei lassen sich allerdings keine Bilddatensätzen generieren, welche weder reproduzierbar sind noch kann auf die

¹ Technische Hochschule Deggendorf, Dieter-Görlitz-Platz 1, 94469 Deggendorf, {philipp.flierl, anton.schmailzl}@th-deg.de

² Universität Passau, FORWISS, Innstraße 43, 94032 Passau, {zimmermann, niedermeie, noori, fuchse}@forwiss.uni-passau.de

Umgebungsbedingungen in der Anwendungsumgebung Einfluss genommen werden. Weiter stößt der aktuelle Stand der Technik zur Erkennung von Beikräutern im Bereich des menschlich sichtbaren elektromagnetischen Spektrums (VIS) von 380 nm bis 780 nm an seine Grenzen, wodurch bestehende Systeme fehleranfällig sind.

Um zukünftig die Fehleranfälligkeit bestehender Systeme zur Erkennung von Beikräutern deutlich zu verbessern, müssen neben dem Pflanzenumriss weitere Unterscheidungsmerkmale herangezogen werden. Als bisher unerforschtes Unterscheidungsmerkmal zählt der „spektrale Fingerabdruck“ von Pflanzen im Bereich des Kurzwelleninfrarot (SWIR) von 700 nm bis 1700 nm. Erst durch den Aufbau eines Vertical-Farmings-Systems und die Einbindung von KI kann dieses Unterscheidungsmerkmal genauer untersucht werden. Denn für eine mögliche Validierung muss eine Laborumgebung geschaffen werden, in welcher eine ausreichend hohe Anzahl an reproduzierbaren Bilddaten generiert werden kann.

2 Vertical-Farming-System als Forschungsinfrastruktur

Eine fotografische Aufnahme für eine Bilddatenbank sollte Kulturpflanzen und Beikräuter ähnlich der Realität in der Anwendungsumgebung zeigen, siehe Abbildung 1.

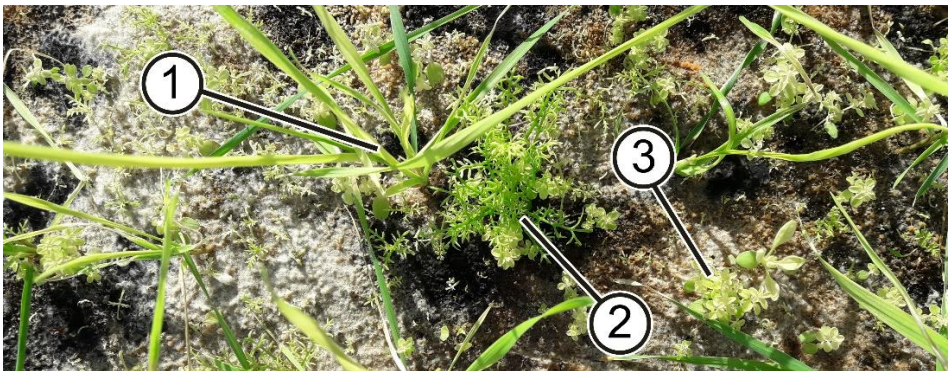


Abb. 1: Angepflanzte Kulturpflanzen und Beikräuter im Vertical-Farming-System: Weizen (1), Kamille (2), Kletten-Labkraut (3)

Die dargestellten Erkenntnisse entstammen aus zahlreichen Vorversuchen und stellen aus wissenschaftlicher Sicht qualitative Aussagen dar. Das vorliegende Paper stellt eine zielgerichtete Erkenntnisdokumentation für die Anpflanzung von Kulturpflanzen und Beikräutern dar. Für aktuelle Versuche aus dem Projekt KIdetect lag der Fokus auf der Kulturpflanze Weizen (1) und den Beikräutern Kamille (2) und Kletten-Labkraut (3). Für den Aufbau der Bilddatenbank werden fotografische Aufnahmen unterschiedlicher Entwicklungsstadien der Pflanzen bei wechselnden Lichtverhältnissen und Perspektiven

benötigt. Aufgrund der Vielzahl der fotografischen Aufnahmen sollte die Generierung der Aufnahmen vollautomatisiert und jahreszeitenunabhängig erfolgen.

Entsprechend den Anforderungen erfolgt der Aufbau des Vertical-Farming Systems als Forschungsinfrastruktur in einem geschlossenen und klimatisierten Überseecontainer. Der aktuelle Stand des Vertical-Farming-Systems ist in Abbildung 2 dargestellt.

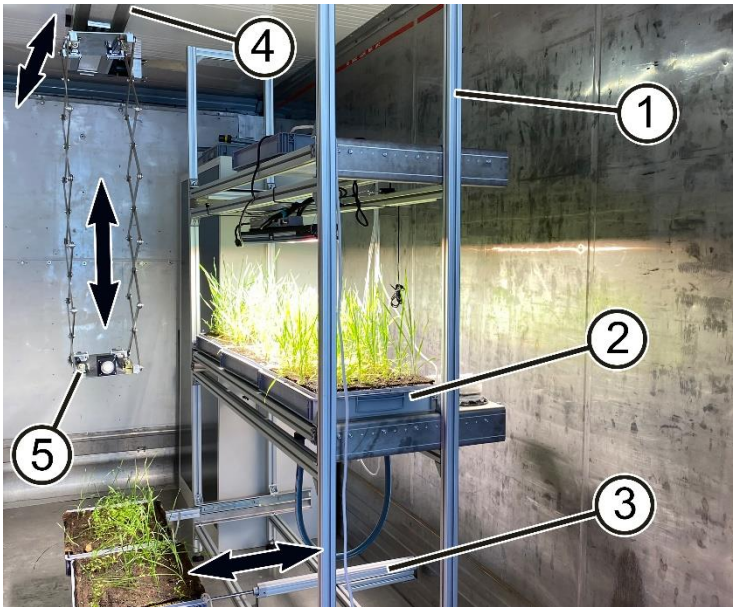


Abb. 2: Aktueller Stand des Vertical-Farming-Systems: Pflanzregal (1), Pflanzkästen (2), Pneumatikzylinder (3), Linearachse (4), Scherenhubtisch (5)

Das Vertical-Farming-System ist 7800 mm lang und besteht aus acht einzelnen Pflanzregalen (1), welche in jeweils drei Pflanzebenen unterteilt sind. Die Kulturpflanzen und Beikräuter wachsen in drei einzelnen Pflanzkästen (2) pro Pflanzebene. Oberhalb einer jeden Pflanzebene befinden sich entsprechende Pflanzlampen mit je 300 Watt der Firma Mars Hydro. Der Abstand der Lampen zum Pflanzbehälter beträgt ca. 450 mm, um alle Pflanzkästen einer Pflanzebene möglichst gleichmäßig auszuleuchten. Die Pflanzebenen sind über Pneumatikzylinder (3) ausziehbar und können dadurch mittig unter eine Linearachse (4) bewegt werden. Durch einen Scherenhubtisch (5), welcher an die Linearachse montiert ist, kann jeder einzelne Pflanzkasten angefahren werden. Der Scherenhubtisch dient als Trägerelement für entsprechende Kameras für die fotografischen Aufnahmen und soll zukünftig auch zur Evaluierung verschiedener Unkrautentfernungsmethoden verwendet werden. Der Fokus des Projekts KIdetect liegt auf der Identifikation des Wachstumszentrums, d. h. für die möglichst genaue Lokalisation des Wachstumszentrums ist eine 3-D-Rekonstruktion erforderlich. Für entsprechende Aufnahmen werden zwei Kameras an dem Scherenhubtisch montiert. Eingesetzt wird zum

einen eine Scientific Kamera mit sCMOS Bildsensor für den Bereich des VIS und eine Kamera mit InGaAs Bildsensor für den Bereich des SWIR. Beide Kameras werden von der Firma Excelitas PCO GmbH bereitgestellt.

Für die Anpflanzung der Kulturpflanzen und Beikräutern wird auf ein hydroponisches System entsprechend Abbildung 3 gesetzt.

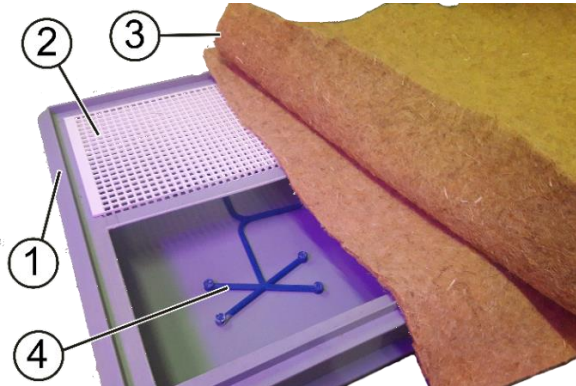


Abb. 3: Hydroponisches System aufgebaut als Pflanzkasten: (1) Euronormbehälter, Trägerplatte (2), Nutzhanfvlies (3), Luftausströmer (4)

Hydroponische Systeme kennzeichnen sich dadurch aus, dass die Pflanzen in eine Wasser-Nährstofflösung wurzeln. Durch die Reduktion an organischem Material bieten sich weniger Nistmöglichkeiten für Insekten. Für das hydroponische System werden Euronormbehälter der Abmessung 60 x 40 x 7,5 cm verwendet und mit einem Volumen von ca. 12 l Trinkwasser gefüllt. Für das verwendete hydroponische System hat sich die Hinzugabe von 10 ml an NPK 3-2-4 pro Pflanzkasten (Euronormbehälter) als zielführend erwiesen, wenn die Kulturpflanze Weizen ein 3- bis 5-Blattstadium erreicht hat. Für die Wasser-Nährstofflösung sollte ein pH-Wert zwischen 6 und 7 eingehalten werden. Die Wasser-Nährstofflösung sollte alle 2 Wochen ausgetauscht werden. Als Auflagefläche für die Samen wird als Substrat Nutzhanfvlies mit einer Materialstärke von ca. 5 mm verwendet. Das Substrat liegt auf vier einzelnen Trägerplatten aus PLA auf, durch welches die Pflanzen hindurchwurzeln können (DE102018132626B4). Durch die jeweils unterhalb einer Trägerplatte befindlichen Luftausströmer (2) wird zum einen die Wasser-Nährstofflösung kontinuierlich mit Sauerstoff angereichert, weiterhin wird durch die Luftausströmer das Substrat insbesondere zur Keimung benässt. Für das Ausströmen der Luft ist ein Mindestvolumenstrom von 35 l/min und 0,3 bar erforderlich. Die Einschaltdauer der Luftdüsen beträgt ebenfalls wie die Einschaltdauer der Lampen acht Stunden pro Tag. Die Luftdüsen sollen möglichst große Luftblasen erzeugen, um das Substrat zu benetzen, aus diesem Grund ist von der Verwendung von Ausströmersteinen aus dem Bereich der Aquaristik abzuraten. In dem gesamten Container befinden sich in Summe 72 Pflanzkästen mit einer Gesamtfläche von ca. 17,3 m² Anpflanzfläche.

3 KI zur Identifikation des Wachstumszentrums

Neben dem Aufbau des Vertical-Farming-Systems als Forschungsinfrastruktur ist das Ziel des Projekts KIdetect die Identifikation des Wachstumszentrums. Sowohl für die Anwendung der verwendeten Kameras als auch für die Auswertung der fotografischen Aufnahmen ist KI zwingend erforderlich.

Die Anwendung von künstlicher Intelligenz erfolgt zum einen in Bezug auf die Registrierung der fotografischen Aufnahmen der beiden verwendeten Kameras, um eine bestmögliche Übereinstimmung der Aufnahmen zu erreichen. Dies erfolgt nach der Einzelkamera-Kalibrierung über die Ermittlung der intrinsischen und extrinsischen Parameter. Mittels einer 3-D-Rekonstruktion (Triangulation) zweier definierter Punkte auf einem Kalibriertarget werden die Kameras mm-genau verifiziert. Für die anschließende Objektzuordnung von Bilddaten wurde ein Pixelmatching durchgeführt. Hierfür wurde sowohl ein Sift-Flann-Matcher als auch ein Epiline-Matcher erprobt, jedoch besteht für beide Algorithmen eine Möglichkeit zur Qualitätsoptimierung. Daher werden die bisherigen Ergebnisse beider Ansätze weiter verbessert und zur Erstellung von Trainingsdaten und die Evaluation eines zukünftigen CNN-Algorithmus genutzt.

Die Anwendung von künstlicher Intelligenz erfolgt zum anderen zur Klassifizierung der Kulturpflanzen und Beikräuter. Diese Klassifizierung gestaltet sich allerdings aufgrund von sich überlappenden Blättern, komplizierten Hintergründen, verschiedenen Pflanzenblättern mit derselben Form, Lichtschwankungen und der Nässe oder Trockenheit des Bodens als herausfordernd. Bisherige Ansätze, die KI-Methoden benutzen, sind z.B. [YZF20], die ein Mask Region-based Convolutional Neural Network (Mask R-CNN) und ein Convolutional Neural Network mit 15 verschiedenen Pflanzenbildern trainieren, um komplexe Pflanzenhintergründe anhand der Blattform zu klassifizieren. In [FC06] wird die Schwellenwertbildung verwendet, um Blattadern vom Hintergrund zu trennen. Hier wird ein Neuronales Netz zur Verbesserung der Klassifizierung eingesetzt. Im Artikel [Ze21] wird ein Fully Convolutional Network mit Bildern von Winterweizen im Hinblick auf ihr Wachstumsstadium, Lichtveränderungen und verschiedene Bodenverhältnisse trainiert. In [Ad20] übertrifft die Schwellenwertbildung andere klassische Methoden der Bildsegmentierung. In der Studie [JCC21] zur Erkennung von Gemüse wurde das CenterNet-Modell so trainiert, dass es Begrenzungsboxen um die Pflanzen zeichnete und die Pflanzen, die außerhalb lagen, als Hintergrund gezählt wurden. Diese Ansätze werden bezüglich ihrer Eignung für das Projekt evaluiert, angepasst und erweitert.

Im Gegensatz zu den vorherigen Ansätzen werden in dem Projekt KIdetect Informationen von zwei Kameras in unterschiedlichen Bereichen des elektromagnetischen Spektrums untersucht. Anhand der gewonnenen Informationen werden Objekte mit Labels annotiert, welche im Folgenden zum Training eines geeigneten ausgewählten CNN genutzt werden. Ziel ist es, Methoden zu bestimmen, mit denen Beikräuter in möglichst vielen Situationen und Konstellationen robust von Kulturpflanzen unterschieden werden können. Dafür erforderlich ist eine iterative Vorgehensweise bestehend aus Lernphasen, Tests und Adaptionen der Modelle.

4 Zusammenfassung

Zusammenfassend lässt sich festhalten, dass durch das vorgestellte Vertical-Farming-System als Forschungsinfrastruktur auf ca. 17 m² gezielt Kulturpflanzen und Beikräuter angepflanzt werden können. Die Anpflanzung erfolgt in 72 einzelnen Pflanzkästen über ein hydroponisches System. Es wurde sich auf die Anpflanzung von Weizen, Kamille und Kletten-Labkraut konzentriert. Für die zielgenaue Identifikation des Wachstumszentrums wird auf eine 3-D-Rekonstruktion (Triangulation) gesetzt, wofür zwei Kameras benötigt werden. Die Anwendung von KI erfolgt zum einen für die Bildregistrierung der fotografischen Aufnahmen der beiden Kameras, zum anderen für die Klassifizierung der Kulturpflanzen und Beikräuter. Verschiedene aktuelle Methoden zur Klassifizierung wurden vorgestellt. Es soll eine Bilddatenbank aus fotografischen Aufnahmen generiert werden, welche sowohl zweidimensionale als auch dreidimensionale Informationen im Bereich VIS und SWIR enthält. Anhand dieser Grundlage soll die zuverlässige Identifikation von Beikräutern in der Laborumgebung sichergestellt werden. Erst bei sicherer Identifikation kann eine Übertragung der gewonnenen Erkenntnisse auf die Anwendungsumgebung erfolgen.

Förderhinweis und Danksagung

Die Erkenntnisse wurden im Projekt KIdetect (Förderkennzeichen: 28DK132) durch das BMEL erarbeitet. Die verwendete Kameratechnologie wurde von der Excelitas PCO GmbH (Projektverbundpartner) zur Verfügung gestellt. Die Fritzmeier Umwelttechnik GmbH & Co. KG (Assoziierter Projektpartner) steht beratend zur Seite.

Literaturverzeichnis

- [Ad20] Adams, J. et.al.: Plant Segmentation by Supervised Machine Learning Methods. The Plant Phenome Journal 03/20, 2020.
- [FC06] Fu, H.; Chi, Z.: Combined Thresholding and Neural Network Approach for Vein Pattern Extraction from Leaf Images. IEE Proc. 153/06, S. 881-892, 2006.
- [JCC21] Jin, X.; Che, J.; Chen, Y.: Plant Segmentation by Supervised Machine Learning Methods. IEEE 09/21, S. 10940-10950, 2021.
- [UB22] Umweltbundesamt, Pressemitteilungen, www.umweltbundesamt.de/presse/pressemitteilungen/neue-eu-verordnung-weniger-pestizide-geht-nur, Stand: 28.05.2023.
- [YZF20] Yang, K.; Zhong, W.; Li, F.: Leaf Segmentation and Classification with a Complicated Background Using Deep Learning. Agronomy 10/20, 2020.
- [Ze21] Zenkl, R. et.al.: Outdoor Plant Segmentation With Deep Learning for High-Throughput Field Phenotyping on a Diverse Wheat Dataset. Front. Plant Sci. 12/21, 2021.

Mobile semantische Dokumentation als Basis für KI-gestützte Beratungsdienste: Das GeoBox Buchungsjournal

Ahmad Kadi¹, Daniel Martini², Ansgar Bernardi¹ und Daniel Eberz-Eder³

Abstract: KI-Systeme zur landwirtschaftlichen Beratung benötigen maschinenlesbare betriebsspezifische Informationen. Das GeoBox Buchungsjournal dokumentiert landwirtschaftliche Beobachtungen und Maßnahmen als formale Instanzen einer Aktivitäts-Ontologie mit Bezug zu etablierten Vokabularen wie AGROVOC. Die Modellierungsprinzipien sichern maximale Flexibilität, Erweiterbarkeit, Austausch und maschinelle Verwertbarkeit der Daten. Der entstehende Wissensgraph wird exemplarisch im Chatbot für die Beratung zum Resistenzmanagement im Pflanzenschutz genutzt.

Keywords: KI-Beratungs-Chatbot im Pflanzenschutz, Mobile Semantische Dokumentation, Ontologie, AGROVOC

1 Einleitung

Entscheidungen in der Landwirtschaft werden zunehmend durch KI-Systeme unterstützt und optimiert. Erfasstes oder erlerntes Expertenwissen wird mit allgemeinen Daten und aktuellen betriebsspezifischen Informationen zusammengeführt, um situationsspezifische Anweisungen zu geben. Traditionelle Datenmodellierung und -erfassung erreicht oft nicht die notwendige Flexibilität, Offenheit und Verknüpfbarkeit der lokal und/oder verteilt gespeicherten Informationen und erfordert häufig lästige Mehrfacheingaben.

Im Innovationsprojekt GeoBox⁴ wurde eine flexibel erweiterbare, mobil nutzbare Lösung für die Erfassung jeglicher feldbezogener Aktivität entwickelt, die betriebliche Datenhoheit und Kontrolle über die eigenen Daten mit flexibler, geteilter Datennutzung verbindet. Daten werden in Graph-Strukturen modelliert und gespeichert. Durch Bezug zu Ontologien und standardisierten Vokabularen entsteht eine formalisierte – und damit maschinell verarbeitbare – Datenhaltung als Basis für Interoperabilität und KI-Einsatz.

Jede Beobachtung oder Maßnahme kann im Feld z.B. per Smartphone erfasst werden; die für die Interaktion notwendigen Eingabemasken werden automatisch aus den

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH - DFKI, Trippstadterstr. 122, 67633 Kaiserslautern, ahmad.kadi@dfki.de / ansgar.bernardi@dfki.de

² Kuratorium für Technologie und Bauwesen in der Landwirtschaft - KTBL, Bartningstraße. 49, 64289 Darmstadt, d.martini@ktbl.de

³ Dienstleistungszentrum ländlicher Raum Rheinhessen-Nahe-Hunsrück DLR-RNH, Rüdesheimer Str. 60-68, 55545 Bad Kreuznach, daniel.eberz@dlr.rlp.de

⁴ GeoBox wird gefördert aus Mitteln des Zweckvermögens des Bundes bei der Landwirtschaftlichen Rentenbank. Siehe <https://geobox-infrastruktur.de>

Beschreibungen der Ontologien generiert. Die Gesamtheit der Einträge, die im Betrieb über die Wirtschaftsjahre entstehen und die sich gegenseitig referenzieren, bildet das sogenannte „Buchungsjournal“. Diese Sammlung betrieblicher Fakten bietet allen Kooperationspartnern eine flexible Schnittstelle zur Abfrage relevanter Daten, die durch den Ontologiebezug automatisch verarbeitbar sind. Explizite Zustimmung zum Datenaustausch sichert die Datenhoheit.

2 Hintergrund

Die Dokumentation landwirtschaftlicher Maßnahmen und Beobachtungen auf einem Feld erfolgt fortlaufend über die Vegetationsperiode. In traditionellen analogen Ackerschlagkarteien werden alle Vorgänge per Hand in Karteikarten erfasst und manuell weiterverwendet. Mittlerweile werden Schlagkarteien digital geführt, die proprietären, geschlossenen Datenmodellierungen der verschiedenen Systeme sind aber aus Sicht der notwendigen Flexibilität, Offenheit und Interoperabilität unzureichend. Neben betrieblichen Datenbeständen benötigen Entscheidungsunterstützungssysteme weitere relevante Datenquellen; z.B. spielen Eigenschaften von Kulturpflanzen, Schaderregern und Pflanzenschutzmitteln oder die Wetterlage eine Rolle bei der Planung von Pflanzenschutzmaßnahmen. Außerdem enthalten Entscheidungsunterstützungssysteme in Form von Modellen oder Entscheidungsbäumen kodifiziertes Expertenwissen. All diese Bestände müssen für die Bereitstellung von auf den jeweiligen Betrieb zugeschnittenen Empfehlungen zusammengeführt werden.

Ontologien und Vokabulare sind wichtig für den Einsatz von KI, da sie eine einheitliche Benennung sichern und gleichzeitig Hintergrundwissen wie etwa Begriffshierarchien oder Anwendungszusammenhänge repräsentieren. Dadurch wird die automatisierte Integration unterschiedlicher Datenquellen ermöglicht. Bei Training und Anwendung von KI-Modellen kann die hierdurch eröffnete Möglichkeit, verknüpfte Kontextinformation einzubinden, genutzt werden, um Labels oder Annotationen mit weiterer Information anzureichern, Zuweisungen von Daten in Eingabevektoren vorzunehmen oder Ausgaben wieder mit standardisierten Begriffen zu belegen. Eine einheitliche Sprache und Interpretation von Daten durch alle Akteure steigern auch das Vertrauen in die Genauigkeit der Daten.

3 Vokabularien, Ontologie und Buchungsjournal

Für eine universell verwendbare, maschinenlesbare Dokumentation nutzt GeoBox das Prinzip, jede Beobachtung oder Maßnahme mit Orts- und Zeitangabe als Instanz einer (allgemeinen oder spezifischen) „Aktivität“ darzustellen. Die konkreten Attribute und Werte dieser Instanz nutzen eindeutige Vokabulare. Hierzu wurde eine Ontologie entwickelt, die die möglichen Feldmaßnahmen und Beobachtungen in eine Klassenhierarchie einordnet Abb. 1. Die Ontologie spezifiziert die Beziehungen zwischen

den Begriffen, legt fest, welche Attribute notwendig sind, und welche Wertebereiche und Datentypen verwendet werden, um eine landwirtschaftliche Maßnahme zu beschreiben. Indem sie Informationen organisiert und strukturiert, erleichtert sie deren Suche, Verständnis und Nutzung. Neben den dargestellten möglichen landwirtschaftlichen Maßnahmen und Beobachtungen kommen auch Stammdaten wie Pflanzenschutzmittel, Maschinen, Saatgut usw. dazu.

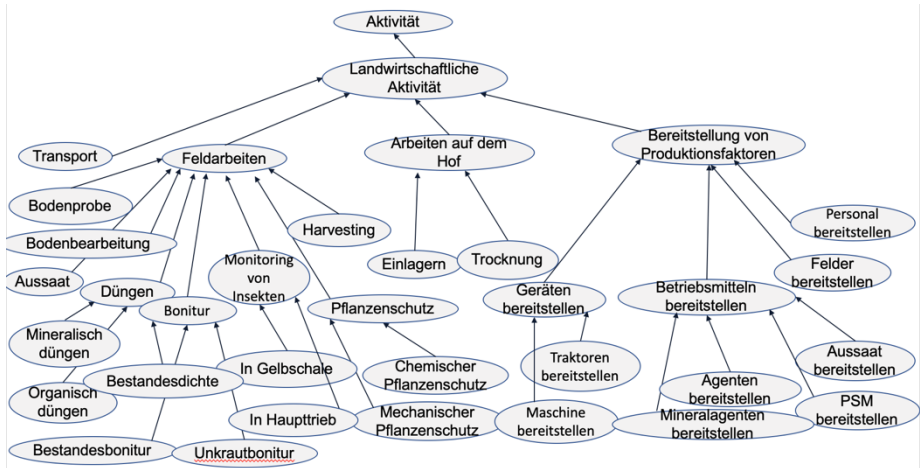


Abb. 1: Klassenhierarchie der Maßnahmen, die mit Import-Funktion, erweitert werden können

Die Ontologie in GeoBox wurde von der PROV-Ontologie [Le13] abgeleitet, die verwendet wird, um Provenance-Informationen zu modellieren, wie die Herkunft, den Ursprung oder die Entstehung von Daten, Dokumenten oder anderen digitalen Objekten. Alle Feldaktivitäten und -beobachtungen sind (indirekte) Unterklassen der Klasse `prov:Activity` und erben damit ihre zentralen Prädikate wie `prov:startedAtTime`, `prov:endedAtTime` und `prov:atLocation`. Stammdaten wie z.B. Pflanzenschutzmittel sind Unterklassen von `prov:Entity`, die Mitarbeiter-Klasse ist eine Unterklasse von `prov:Agent`.

Die Spezifikation, welche Prädikate für welche Klassen angewendet werden sollen, und jeweils erlaubte Wertebereichen erfolgt durch die Verwendung von SHACL-Ausdrücken, [KK17]. Die SHACL-Graphen repräsentieren zusätzlich modellierte Bedingungen, die als gültig akzeptierte RDF-Graphen erfüllen müssen. Beim Speichern der Daten wird gegen die Shapes validiert; ungültige Daten werden abgewiesen. Ein wichtiges Vokabular zur Beschreibung landwirtschaftlicher Objekte und Vorgänge ist der Thesaurus AGROVOC [AG23]. Klassen und Prädikate der GeoBox-Ontologie sind mit den entsprechenden Konzepten aus AGROVOC über Mapping-Relationen verbunden.

Das GeoBox Buchungsjournal entsteht durch das kontinuierliche Hinzufügen von Einträgen, die die Arbeit im Betrieb oder auf dem Feld dokumentieren. Mit dieser universellen, formalen Datenmodellierung und den Vokabularien steht die Grundlage bereit, um standardisierte Aussagen über die jeweiligen Felder zu erstellen.

4 Integration mit KI-Systemen: ChatBot zur Resistenzvermeidung

Der im Buchungsjournal gespeicherte betriebliche Wissensgraph kann durch beliebige SPARQL-Anfragen [SP23] ausgewertet werden. Dadurch können Informationsbedürfnisse externer Systeme nach Inhalt und Form befriedigt werden. Soweit Fremdsysteme und Dritte solche Anfragen stellen, wird die Ausgabe durch Menschen geprüft und freigegeben und so die Datenhoheit gewahrt.

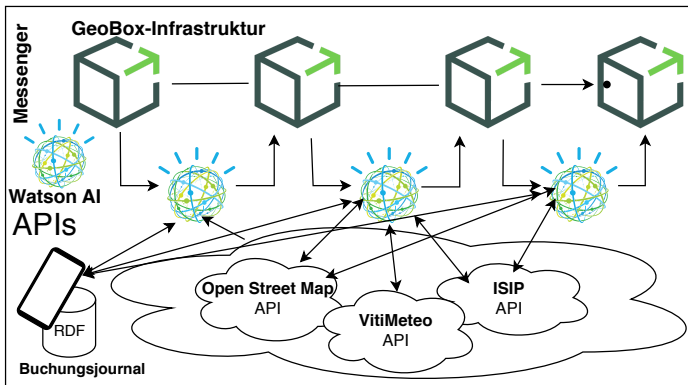


Abb. 2: Chatbot-Architektur: IBM Watson, Buchungsjournal, Datenquellen und Messenger

Im Projekt wurde ein Chatbot für interaktive Beratungsdialoge entwickelt, der Empfehlungen für die Anwendung von Pflanzenschutzmitteln gibt und dabei Resistenzaspekte und die individuelle betriebliche Historie berücksichtigt. Die Architektur Abb. 2 nutzt per API die Entscheidungsbaume des öffentlichen Experten-Beratungsdienstes (ISIP), die wirksame und Resistenzen vermeidende Pflanzenschutzempfehlungen ermitteln. Zusammen mit Wetter- und Lokationsinformationen fließen diese in das auf dem IBM Watson Assistant [WA23] aufbauenden Chatbot-System. Die Interaktion mit den Nutzenden geschieht durch sicheren Austausch über die Kommunikations-App „Messenger“ der Geobox-Infrastruktur.

Im Dialog müssen dann u. a. Fragen zur aktuellen Aussaat und den früher eingesetzten Pflanzenschutzmitteln beantwortet werden. Diese Arbeiten sind im Buchungsjournal unter Nutzung der Vokabulare dokumentiert, so dass der Chatbot unmittelbar SPARQL-Anfragen an das Buchungsjournal stellen kann. Die erhaltene Antwort wird im Chat hervorgehoben und manuell abgeschickt.

Abb. 3 illustriert den Vorteil der ontologiegestützten semantischen Modellierung: Der Entscheidungsbaum erfordert die Angabe der Wirkstoffe der Pflanzenschutzmittel entsprechend der vom Insecticide Resistance Action Committee herausgegebenen IRAC Tabellen [IR23]. Landwirte dokumentieren jedoch vor allem die kommerziellen Bezeichnungen der Produkte. Ein Übersetzungsmechanismus wurde mit Hilfe der Vokabularien und SPARQL-Abfragen umgesetzt, in dem die „Active Ingredient“ der

Pflanzenschutzmittel, mit denen der IRAC-Vokabularien verbunden wurden. Dieses Mapping könnte zwar auch durch einfache Zuordnungstabellen erreicht werden, unter Nutzung von Standardtechnologien des Semantic Web können solche Zuordnungen jedoch auch verteilt entwickelt werden, es können verschiedene Zuordnungssysteme kombiniert werden und es lassen sich feingranulierte Bedeutungsunterschiede semantisch abbilden (z. B. Identität, Äquivalenz, Ähnlichkeit).

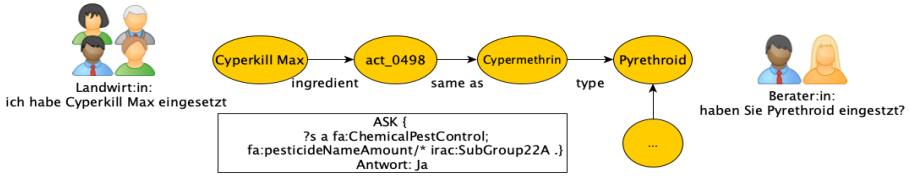


Abb. 3: Übersetzungsmechanismus durch Vokabularien und Abfrage

Die folgenden Screenshots in Abb. 4 illustrieren den Chat-Verlauf einer Resistenzberatung bei der Schädlingsbekämpfung, wobei z.B. die Historie früherer Anwendungen direkt aus dem Buchungsjournal entnommen wird. Zu sehen sind die hervorgehobenen Antworten zwischen den möglichen Optionen.

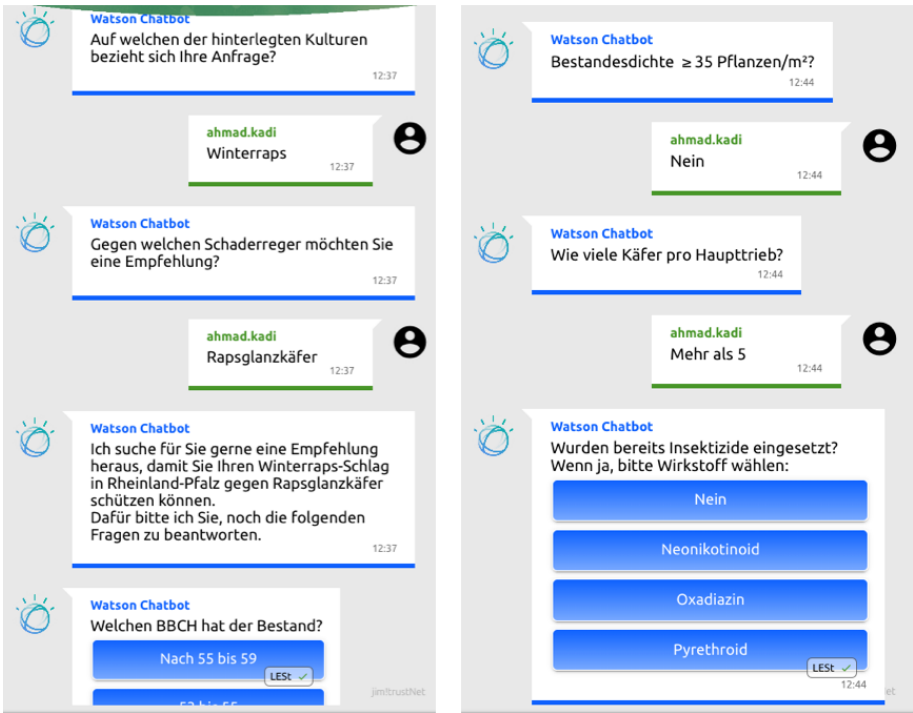


Abb. 4: Beratungsdialog (Chatbot) für die Resistenzberatung.

5 Zusammenfassung

Das Geobox-Buchungsjournal modelliert landwirtschaftliche Aktivitäten in einem ontologiebasierten, formalen Wissensgraph, der beliebig ausgewertet werden kann. Das Beispiel eines Chatbots für Resilienzberatung im Pflanzenschutz illustriert als Proof-Of-Concept die Leistungsfähigkeit einer KI-Unterstützung auf dieser Grundlage. Der Chatbot verbindet den IBM Watson Assistent mit Expertenwissen (Entscheidungsbäumen), Wetter- und Lokationsinformationen und erfragt im Dialog relevante Information aus dem Buchungsjournal. Die Interaktion erfolgt über den sicheren Geobox-Messenger.

Durch die beschriebenen Modellierungsprinzipien realisiert das GeoBox in der ein offenes betriebliches Informationssystem: Die erhobenen Daten sind als Instanzen weithin akzeptierter Ontologie-Konzepte, nach Struktur und Inhalt bekannt, und automatisiert nutzbar. Die referenzierten Vokabulare liefern weiterführende Zusammenhänge. Die Datenstrukturen können jederzeit auch individuell erweitert werden; So sind sie geeignet, sowohl Daten aus verschiedenen Quellen für eine übergreifende Auswertung zu vereinen als auch gezielte, föderierte Abfragen auszuführen, um KI-gestützte Systeme mit notwendigen Daten und ggfs. umfassender Kontextinformation zu versorgen. Außerdem werden die Eingabemasken des Systems automatisch angepasst, sodass für die Erfassung flexibel auf neue Anforderungen reagiert werden kann. Damit illustriert GeoBox die Vorteile einer expliziten semantischen Datenmodellierung für die Realisierung betrieblicher, landwirtschaftlicher Informationssysteme, welche den vielfältigen Anforderungen an Kooperation und Dialogfähigkeit gerecht wird.

Literaturverzeichnis

- [Le13] Lebo, T.; Sahoo, S.; McGuinness, D.; Belhajjame, K.; Cheney, J.; Corsar, D., D.; Garijo, D.; Soiland-Reyes, S.; Zednik, S.; Zhao, J., Prov-o: The prov ontology. W3C recommendation, 30, 2013.
- [KK17] Knublauch, H.; Kontokostas, D., Shapes constraint language (SHACL). W3C Candidate Recommendation, 11(8), 2017.
- [SP23] SPARQL 1.1 Query Language, www.w3.org/TR/2013/REC-sparql11-query-20130321/, Stand 24.04.2023.
- [AG23] AGROVOC multilingual thesaurus, www.fao.org/agrovoc, Stand: 24.04.2023.
- [IR23] IRAC: Insecticide Resistance Action Committee, www.iraac-online.org, Stand: 24.04.2023.
- [WA23] IBM Watson Assistant: KI-Chatbot mit hohem Bedienungskomfort, www.ibm.com/de/de/products/watson-assistant/artificial-intelligence, Stand 27.05.2023.

Canola seed or not? Autoencoder-based Anomaly Detection in Agricultural Seed Production

Maksim Kukushkin,¹ Matthias Enders,² Reinhard Kaschuba,³ Martin Bogdan⁴ and Thomas Schmid⁵

Abstract: Analysing harvested seeds is a time-consuming task in the seed-producing industry. Automating this process has the potential to enhance and expedite agricultural seed production. In our study, we focus on differentiating Canola seeds from visually similar non-Canola seeds using computer vision techniques. Our approach utilises both RGB and hyperspectral images, captured by a specialised camera, to train separate autoencoder neural networks. By leveraging the high spatial resolution of RGB data and the high spectral resolution of hyperspectral data, we develop distinct models for Canola seed analysis, ensuring a comprehensive and robust assessment. The autoencoder networks are trained on a dataset of Canola seeds, allowing for the extraction of latent representations from both RGB and hyperspectral data. This enables efficient compression of input data and effective discrimination between Canola and non-Canola seeds. Our proposed approach demonstrates promising results in detecting non-Canola seeds in unseen test data.

Keywords: anomaly detection, seed production, hyperspectral imaging, autoencoder

1 Introduction

In agricultural seed production, classifying and sorting harvested seeds is ultimately required for quality assurance. At the same time, however, this task is also particularly challenging and requires years of expertise and training for human analysts. Using rapeseed as a sample case, we are developing an AI-supported platform for the classification and sorting of plant seeds and seed purity. To this end, we aim to integrate advanced machine learning techniques with robotic sensors and actors into a unified, continuously learning sorting platform. It will establish sorting as a collaborative process between users and the learning algorithm in order to significantly increase its efficiency in comparison to manual testing.

From a computational point of view, the central task of this approach is the computer vision task of anomaly detection. This task plays a vital role in both human and machine intelligence

¹ Leipzig University, Institute of Computer Science, Augustusplatz 10, 04109 Leipzig, Germany
kukushkin@informatik.uni-leipzig.de

² NPZ Innovation GmbH, Hohenlieth-Hof, 24363 Holtsee, Germany m.enders@npz-innovation.de

³ Hochschule Bielefeld - University of Applied Sciences and Arts, Interaktion 1, 33619 Bielefeld, Germany
reinhard.kaschuba@hsbi.de

⁴ Leipzig University, Institute of Computer Science, Augustusplatz 10, 04109 Leipzig, Germany
bogdan@informatik.uni-leipzig.de

⁵ Martin-Luther-Universität Halle-Wittenberg / Lancaster University in Leipzig / Leipzig University, Germany
thomas.schmid@medizin.uni-halle.de

by identifying unusual or unexpected patterns in data. It holds significant importance across diverse domains such as science, engineering, finance, healthcare, cybersecurity, and more. Anomaly detection is crucial for identifying outliers or deviations from the expected behaviour, enabling timely intervention and decision-making in various applications and disciplines.

Here, we apply anomaly detection techniques in the context of agricultural production of Canola seed (scientifically known as *Brassica napus* L./Rape seed). As Canola seeds play a crucial role in ensuring global food security [AA; He20], the presence of weeds in oilseed rape fields during the harvest presents a significant challenge. It complicates the seed sorting process and compromises the accuracy of regulatory purity testing during seed production. To tackle this issue, we propose an anomaly detection system specifically designed for differentiating Canola from non-Canola seeds using hyperspectral and RGB images. In this paper, we assess this approach for Canola seeds and samples of three commonly encountered weed species by conducting comprehensive anomaly detection analyses.

2 Related Work

Anomaly detection for images is an important area of research with broad applications in various fields [Pi14]. Recently, there has been increasing attention to the development of unsupervised anomaly detection methods that do not rely on labelled data. The three arguably most widely used approaches employed for anomaly detection are autoencoders, generative adversarial networks and transfer learning with convolutional neural networks.

In the context of anomaly detection, an autoencoder (AE) may be trained using a collection of normal images and subsequently applied to analyse new, unseen images. If the AE struggles to reconstruct an input image accurately, it is identified and classified as an anomaly. A vanilla AE may, however, struggle to reconstruct complex image features, such as textures and patterns, which can lead to false positives or false negatives. To address this limitation, researchers have proposed various modifications to the AE architecture, such as using denoising AEs [Lu17] or adding regularisers to the loss function [AC15; Ma15; Ri11; Vi08] to improve their reconstruction accuracy and anomaly detection performance.

Generative adversarial networks (GANs) models [AAB19; Sc17; Sc19] may be used to characterise the normal distribution of the data and subsequently identify anomalies as samples that significantly deviate from this distribution. However, GAN approaches present certain challenges in their training process, such as failure to converge and mode collapse [Me16]. Additionally, distinguishing abnormal samples from the generative distribution poses difficulties, affecting the performance [AAB19]. Another popular approach is transfer learning, where pre-trained CNNs are fine-tuned on anomaly detection tasks [An16; De21; Ro22]. This approach has been shown to be effective in detecting anomalies in new domains with limited labeled data availability. However, due to the disjointed feature extraction and anomaly scoring could lead to suboptimal results [Pa21].

3 Methodology

Due to its simplicity, speed, and established performance, we selected an autoencoder (AE) architecture as the foundation for the anomaly detection system. The availability of sufficient training data, too, was supporting this decision. The AE operates in two stages: Initially, it takes an input image, represented as r_i , and compresses it into a lower-dimensional representation. In the subsequent stage, the AE aims to reconstruct the input image with the highest accuracy. The loss function, which measures the difference between the reconstructed image r'_i and the input image r_i , is defined as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (r_i - r'_i)^2 \quad (1)$$

where r_i and r'_i denote the input and reconstructed images, respectively.

Considering the presence of two types of data, namely hyperspectral data with high spectral resolution but low spatial resolution, and RGB data with low spectral resolution but high spatial resolution, we propose to employ two separate autoencoder-based models tailored to each data type, namely RGB-AE and HS-AE with latent dimension of 4096 (see. Fig 1).

It is important to note that our HS-AE differs from RGB-AE in its architecture. In addition to Conv2D and fully connected layers (similar to RGB-AE), HS-AE also incorporates Conv3D layers to extract spatio-spectral information (refer to Fig. 1b). To enable the use of RGB images as input, we standardised their size by resizing or padding them to a consistent dimension of 192x192 pixels. Furthermore, for the sake of computational efficiency in HS-AE, a hyperspectral region of interest was extracted from the centre of the hyperspectral images, specifically with a volume of 16x16x300.

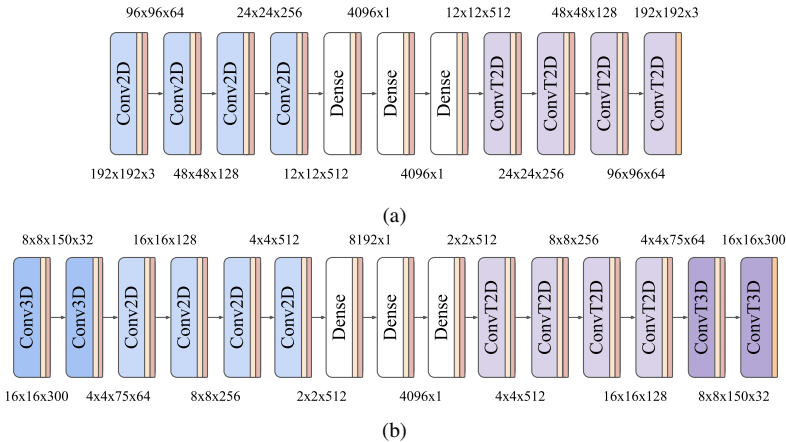


Fig. 1: Architectures of RGB-AE (a) and HS-AE (b). Please note, for RGB-AE and HS-AE use BatchNorm (yellow) and LeakyReLU (red)

4 Experiments

Data. The training dataset utilised in our study comprises a total of 3,156 RGB images along with their corresponding 3,156 hyperspectral images of Canola. For the purpose of testing, we have a separate set containing 370 images, out of which 280 belong to the normal class (*Brassica napus* L.). The remaining 90 images in the test set are equally distributed among three distinct weed species: *Anchusa arvensis* L., *Sinapis alba* L., and *Sinapis arvensis* L. (see Fig. 2). Hence, the primary objective of our anomaly detection approach is to accurately identify these 90 images from the anomalous class. The corresponding hyperspectral images are composed of 300 wavelengths within the visible and near-infrared (VNIR) range of the electromagnetic spectrum (380 - 1000 nm) (see Fig. 3).

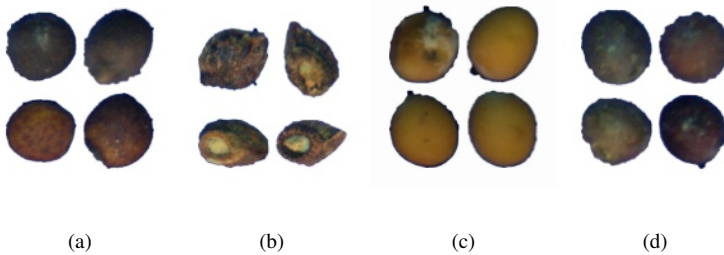


Fig. 2: Examples of images from the test dataset. (a) Represents the normal class, showing images of *B. napus*, (b-d) Depict anomalous classes, displaying images of *Anchusa arvensis* L., *Sinapis alba* L., and *Sinapis arvensis* L. respectively.

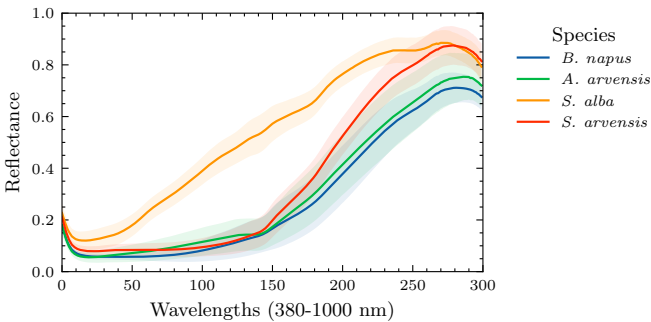


Fig. 3: Mean spectra with standard deviation for each seed species in test dataset

Training setting. Both the RGB-AE and HS-AE models were trained using the same configuration. This configuration involved training with the Adam optimizer, using a learning rate of 0.001 for 100 epochs with a learning rate schedule. A batch size of 128 was employed, and various data augmentation techniques, such as vertical and horizontal flips, as well as rotations, were applied to augment the training dataset and expand its size. To compare the performance of the two models, we employed several evaluation metrics,

including (i) the Area Under the Receiver Operating Characteristic Curve (AUC), (ii) Accuracy, (iii) Sensitivity, (iv) Specificity and (v) weighted F1-Score. In our study, we determined the threshold by using the ROC curve, as the selection of this threshold can lead to significant variations in metrics (ii) to (iv).

Results. As illustrated in Fig. 4a, the HS-AE exhibits an AUC score of 0.96, surpassing the score of 0.90 achieved by the RGB-AE. The optimal thresholds determined by analysing the ROC curve are 7.7191×10^{-3} for the RGB-AE and 8.5624×10^{-4} for the HS-AE. Upon closer inspection of the Table 1, it can be observed that the specificity of the RGB-AE is only slightly higher than that of the HS-AE, reaching 0.857. However, the HS-AE demonstrates a significantly higher sensitivity of 0.941. Furthermore, the F1-Score is also superior for the HS-AE, measuring 0.881 as opposed to 0.845 for the RGB-AE. Similarly, the accuracy, as illustrated in Figure 4b, shows that the HS-AE achieves 0.875, which is 0.35 higher than the accuracy achieved by the RGB-AE.

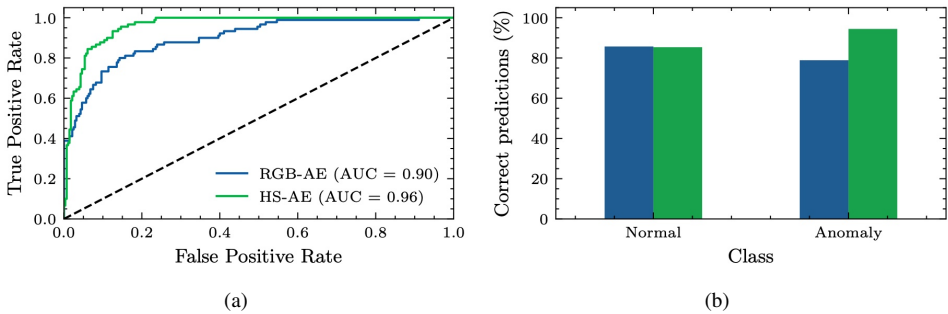


Fig. 4: Comparison of the performance between RGB-AE (blue) and HS-AE (green): (a) ROC curve and Area under the ROC curve, and (b) Number of correctly detected images.

	Accuracy	Sensitivity	Specificity	F1-score	AUC
RGB-AE	0.840	0.788	0.857	0.845	0.897
HS-AE	0.875	0.944	0.853	0.881	0.961

Tab. 1: Evaluation metrics of RGB-AE and HS-AE

5 Discussion

Results from training and testing demonstrate the reliability of our approach for differentiating Canola from non-Canola seeds. In order to gain deeper insights, we conducted a thorough analysis to identify the specific types of anomalies where the HS-AE and RGB-AE made the highest number of incorrect predictions. Our findings, depicted in Fig. 5, reveal that both models performed remarkably well in detecting anomalies associated with the species *S. alba* with the HS-AE model correctly detecting all images of this species.

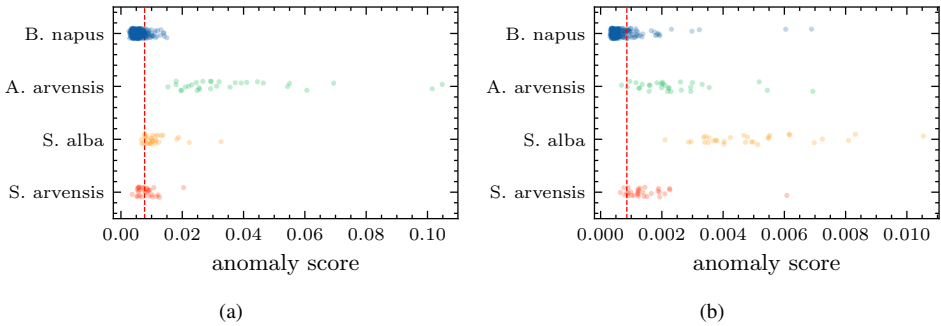


Fig. 5: Reconstruction error distributions (anomaly scores) for test datasets generated by RGB-AE (a) and HS-AE (b), using MSE. The dashed red line indicates the threshold.

On the other hand, the RGB-AE model successfully detected all images associated with the *A. arvensis*, but encountered difficulties in detecting anomalies related to *S. arvensis*. This difficulty can be attributed to the visual similarities between *B. napus* and *S. arvensis*. However, by utilising hyperspectral data, it becomes possible to effectively distinguish between these two species. The aforementioned findings highlight the superiority of the HS-AE, as it outperformed the RGB-AE across almost all evaluation metrics. This indicates that the spectral information of plant seeds holds greater importance than the spatial information for the most cases.

6 Conclusion and Outlook

In conclusion, we strongly believe that AI has a crucial role to play in promoting sustainability in agriculture. By automating routine processes and reducing errors, AI simplifies the lives of agricultural workers. Our system specifically focuses on seed production and aids in ensuring seed purity, a task traditionally performed by humans. By incorporating AI technology, we can enhance the efficiency and accuracy of this process, ultimately contributing to a more sustainable agricultural industry.

In this study, we successfully demonstrated the effectiveness of autoencoders (AE) in differentiating Canola from non-Canola seeds using RGB and hyperspectral images. The results highlight the potential of our approach to optimize and accelerate agricultural seed production.

Moving forward, our future work aims to advance anomaly detection techniques in agriculture. We plan to integrate the strengths of RGB imagery and hyperspectral data, leveraging both approaches to achieve more accurate anomaly detection. Furthermore, we will identify key wavelengths that provide valuable information for distinguishing different types of plant seeds. This will simplify the anomaly detection process, increase efficiency, and potentially reduce costs.

Acknowledgements

This study was carried out as part of the project KIRa funded by the German Ministry of Food and Agriculture (BMEL, FKZ 28DK116A20).

We express our gratitude to NPZ Innovation GmbH for generously providing the dataset, without which this research would not have been possible.

References

- [AA] Agriculture Department, N.; Agriculture, U.: *Agricultural Statistics*, 2002 (Paperback). U.S. Government Printing Office, ISBN: 9780160869273.
- [AAB19] Akcay, S.; Atapour-Abarghouei, A.; Breckon, T.P.: GANomaly: Semi-supervised Anomaly Detection via Adversarial Training. In: *Computer Vision – ACCV 2018*. Springer International Publishing, pp. 622–637, 2019, URL: https://doi.org/10.1007/978-3-030-20893-6_39.
- [AC15] An, J.; Cho, S.: Variational autoencoder based anomaly detection using reconstruction probability. *Special lecture on IE 2/1*, pp. 1–18, 2015, URL: <http://dm.snu.ac.kr/static/docs/TR/SNUDM-TR-2015-03.pdf>.
- [An16] Andrews, J.; Tanay, T.; Morton, E. J.; Griffin, L. D.: Transfer representation-learning for anomaly detection. In: *JMLR*, 2016.
- [De21] Defard, T.; Setkov, A.; Loesch, A.; Audigier, R.: PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization. In: *Pattern Recognition. ICPR International Workshops and Challenges*. Springer International Publishing, pp. 475–489, 2021, URL: https://doi.org/10.1007/978-3-030-68799-1_35.
- [He20] Heuzé, V.; Tran, G.; Sauvant, D.; Lessire, M.; Lebas, F.: Rapeseeds, <https://www.feedipedia.org/node/15617>, [Online; accessed 2023-05-23], Jan. 2020.
- [Lu17] Lu, W.; Cheng, Y.; Xiao, C.; Chang, S.; Huang, S.; Liang, B.; Huang, T.: Unsupervised Sequential Outlier Detection With Deep Architectures. *IEEE Transactions on Image Processing* 26/9, pp. 4321–4330, Sept. 2017, URL: <https://doi.org/10.1109/tip.2017.2713048>.
- [Ma15] Makhzani, A.; Shlens, J.; Jaitly, N.; Goodfellow, I.; Frey, B.: Adversarial Autoencoders, 2015, URL: <https://arxiv.org/abs/1511.05644>.
- [Me16] Metz, L.; Poole, B.; Pfau, D.; Sohl-Dickstein, J.: Unrolled Generative Adversarial Networks, 2016, URL: <https://arxiv.org/abs/1611.02163>.
- [Pa21] Pang, G.; Shen, C.; Cao, L.; Hengel, A. V. D.: Deep Learning for Anomaly Detection. *ACM Computing Surveys* 54/2, pp. 1–38, Mar. 2021, URL: <https://doi.org/10.1145/3439950>.

- [Pi14] Pimentel, M. A.; Clifton, D. A.; Clifton, L.; Tarassenko, L.: A review of novelty detection. *Signal Processing* 99/, pp. 215–249, June 2014, URL: <https://doi.org/10.1016/j.sigpro.2013.12.026>.
- [Ri11] Rifai, S.; Vincent, P.; Muller, X.; Glorot, X.; Bengio, Y.: Contractive Auto-Encoders: Explicit Invariance during Feature Extraction. In: *ICML'11*, Omnipress, Bellevue, Washington, USA, pp. 833–840, 2011, ISBN: 9781450306195.
- [Ro22] Roth, K.; Pemula, L.; Zepeda, J.; Schölkopf, B.; Brox, T.; Gehler, P.: Towards total recall in industrial anomaly detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Pp. 14318–14328, 2022, URL: https://openaccess.thecvf.com/content/CVPR2022/papers/Roth_Towards_Total_Recall_in_Industrial_Anomaly_Detection_CVPR_2022_paper.pdf.
- [Sc17] Schlegl, T.; Seeböck, P.; Waldstein, S. M.; Schmidt-Erfurth, U.; Langs, G.: Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. In: *Lecture Notes in Computer Science*. Springer International Publishing, pp. 146–157, 2017, URL: https://doi.org/10.1007/978-3-319-59050-9_12.
- [Sc19] Schlegl, T.; Seeböck, P.; Waldstein, S. M.; Langs, G.; Schmidt-Erfurth, U.: f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis* 54/, pp. 30–44, May 2019, URL: <https://doi.org/10.1016/j.media.2019.01.010>.
- [Vi08] Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.-A.: Extracting and composing robust features with denoising autoencoders. In: *Proceedings of the 25th international conference on Machine learning - ICML '08*. ACM Press, 2008, URL: <https://doi.org/10.1145/1390156.1390294>.

Federated Learning in Agriculture: Potential and Challenges


Mortesa Hussaini ¹, Anthony Stein ²


Abstract: *Federated learning* is an emerging technique in machine learning that allows multiple parties to *collaboratively* train models without sharing raw data. It has been applied in various fields such as healthcare, finance, and transportation. In this paper, we cast light on the potential of federated learning in the highly relevant social-ecological domain of *agriculture*, a field in which *digitization* is becoming increasingly prevalent. We briefly introduce the collaborative learning concept of federated learning and briefly consider its postulated benefits and open challenges. The potential of federated learning to overcome concerns against digital technology in agriculture, e.g., *data privacy* and *sovereignty* or initial investment and operating costs, is then discussed. We also identify system requirements and stress the necessity of appropriate IT-ecosystems and touch upon specific requirements which can enable federated learning to ensure both a data-sovereign and efficient information and knowledge exchange among multiple parties. Based on that, agricultural use cases where federated learning can unfold its potential by not only improving the quality of machine learning models, but also by alleviating overarching adoption barriers, will be exemplary delineated.

Keywords: Federated Learning; Digital Agriculture; Distributed AI; Data Sovereignty

1 Introduction

Machine Learning (ML) and *Artificial Intelligence* (AI) have already taken an important role in many parts of our lives and there is no sign that this development will stop any time soon. Especially in agriculture, where digitization is increasingly finding its way, farmers can benefit enormously from this technology. It can be used, e.g., as a support for sustainable land management and, thus, to meet the challenges in the area of food security. In a 2020 Bitkom study [Bi23b], 500 German farmers were surveyed on various aspects of digitization in agriculture. The study has shown that farmers value the opportunities of digital technology in agriculture far higher than risks and that about 82% of the farmers already use smart-farming technologies. In contrast to that, only 8% of the surveyed indicated that it is not an actual topic for them. Even though 44% of those who use smart-farming technologies stated that they are planning or at least thinking about applying AI on their farms, only 9% (so only about 7% in total) are actually doing that. Although digitization plays already an important role the use of AI still appears underrepresented in agriculture. What could be

¹ University of Hohenheim, Dept. Artificial Intelligence in Agricultural Engineering & Computational Science Hub (CSH), mortesa.hussaini@uni-hohenheim.de  <https://orcid.org/0000-0002-3621-7776>

² University of Hohenheim, Dept. Artificial Intelligence in Agricultural Engineering & Computational Science Hub (CSH), anthony.stein@uni-hohenheim.de  <https://orcid.org/0000-0002-1808-9758>

the reasons? A newer Bitkom study [Bi23a] shows that data sovereignty plays a major role. Only 1% of the asked farmers said they would make their data available without restriction and even 13% said that they would not provide their data in any case. Farmers have privacy and security concerns [US23] when sharing their data. Right now there are not many laws and regulations on the use of AI which unsettles many farmers and prevents them from using and apply AI on their farms. Another reason could be that farmers cannot connect AI and practical use cases in the real world due to a lack of information and explanation provided by the manufacturers and developers. Thus, it is difficult and not comprehensible for them why and how the investment pays off. ML often requires a lot of data to perform well, preferably from different sources for good generalization. Producing large amounts of data is no longer a problem, as the focus has shifted to the preparation and processing of data. Thereby uploads of large amount of data and subsequent processing is necessary and this process requires sensitive and private data to be handled, which involves high risk and great responsibility to store them centrally. Concerns about the right and responsible handling of this issue as well as the lack of rules and regulations could prevent many users from sharing their data. *Federated Learning* (FL) [Mc17] is a distributed AI paradigm following the working principle of bringing the model to the data and not vice versa. This approach promises great potential, i.e., to secure both data privacy and security, because users do not need to share their private data but only models. Since by this approach the computational effort is shared, it could reduce costs for the individual participating clients. Against this backdrop, research into the integration of FL in productive agricultural contexts is an important step to provide easy access for farmers and facilitate the utilization of AI in agriculture. Few narrowed applications of FL to agriculture for e.g., crop yield prediction and plant disease detection have been proposed in the literature [Kh22], [TMG22], but overall its adoption to this important domain is at its infancy. In this paper, we therefore want to initiate a discussion on FL's potential, unsolved challenges and related aspects regarding its application to digital agriculture.

2 Federated Learning in Agriculture

FL as a *machine learning* (ML) approach focusing on collaborative learning of distributed systems was first introduced in [Mc17]. The authors demonstrated that this method can keep up with state of the art ML methods and even outperform them under certain circumstances. They also applied and are testing this method together with their team from Google Research since 2017 in a practical use case [MR23], the *GBoard*. FL follows the concept of data parallelism [Sh18], where ML-models can benefit from enormous data sources without the need to centralize them. Starting with a pre-trained basis network, which is initially shared to all users (called clients), each of the clients then decentralized trains or optimizes its own individual network with personal data. Instead of their data, the federation of users then share the ML model parameters (called *updates*) of their locally adapted models only. Based on that a new model is being computed by aggregating local client models. This updated 'global' basis model is then be shared within the federation and the process repeats.

More details of the system design can be found in [Bo19]. This approach is improving data sovereignty and protection from scratch and thus oppose privacy and security issues. On top of that FL simplifies it for ML-models to make use of distributed data sources from all over the world (or even beyond, if we include, e.g., satellites), which leads to a bigger overall training set and therefore better performance [Le19] and paves the way to implement Continual Learning [De22] to make the ML-model more robust [Ha20]. Even non-IID (IID: Independent and identically distributed) data, which normally is closer to reality regarding the distribution than prepared data sets, will be easier to handle and thus result to a better (Out-of-Distribution) generalization [Gu22]. Another advantage are the *lower requirements* for single devices due to *shared computational effort* for every user, which implicates lower cost for every participant and therefore better access for everyone who wants to join the federation.

Since FL is not yet another ML model trained to perform well on one specific task, but more a system for these trained models to share their training experiences from different data sets, specific case studies need to be conducted to adapt and apply FL in agricultural contexts. Therefore, we proceed by delineating potential agricultural use cases of FL and then discuss system requirements and options for the realization of FL.

2.1 Potentials and Use Cases

As mentioned above, one of the benefits of FL is the enhancement of data sovereignty and privacy protection. In particular in agriculture, farmers have a strong focus on this aspect, since the leak of personal and farm specific information of the farmers' fields and goods could harm their advantages against their competitors. Since this advantage of FL generally holds for every application, we will not mention it further, but set our focus more on two exemplary agricultural use cases to reveal further advantages of FL to be leveraged.

2.1.1 Plant Disease Detection using Horizontal FL

Plant disease detection is an important task, especially in an early stage of vegetation. Detected early, the farmer can remove the infected plants and thus stop the disease from spreading and creating a greater damage to the yield of the farmer. Considering this, it is comprehensible that there are researches for the development of ML-models to adopt this task. Without consideration of the difficulty of the task, for training a ML-model, which cannot only detect one specific but maybe all diseases, not only a big, but also a diverse data set is needed, which includes all existing diseases. And even if such a data set can be provided, there also exist the challenge of Continual Learning, since there can be an outbreak of a new disease at anytime, which the ML-model needs to adapt on. So we identified three main challenges for this task to be solved: We need a big data set, the data set needs to include enough examples of all known diseases, and the ML-model needs continuously

be fed with new examples to be able to detect also new discovered diseases. Consider a pre-trained ML-model is been distributed to the farmers to train it on their own, we would be confronted with several issues here. First of all, not every farmer is big enough to generate a big enough data set to train such a deep ML-model on its own and even less are willed to pay the costs to generate the needed data set size. Additionally, this data set will not be diverse enough and only include local examples of the individual field of the farmer and thus only some few diseases, if any. And there also lies the next issue: The farmer might be lucky enough to have no outbreak at all, which also means that there are no examples for the ML-model to be trained on detecting the next outbreak. The damage will be greater when the farmer's luck runs out since the ML-model won't be able to detect the disease early enough. The same applies to new discovered diseases, where only few data examples exist which are not shared to others. The issue is that the farmers' local ML-models would only be able to detect a disease after it occurred several times on their fields. So unless it is not possible to centrally collect all data from all farmers over the world, there will always be some lack of information given to the ML-model, so that it cannot perform well under all circumstances. As mentioned before, generating a big amount of data will not be the problem nowadays, but in this case it will be a challenge to convince the farmers to share their sensitive and personal data.

Horizontal FL [SVG23], [IE21] could resolve this challenges. It is a FL approach, where the same ML-model is shared and trained with different training samples, but on the same features for all participants, i.e., the data type and structure will be the same for all farmers, but obviously with different examples. For instance, all farmers use the same drone to generate an RGB image with a specific pixel size from the same perspective. By sharing the ML-model with the farmers and updating it by aggregation of the local ML-models, trained by local individual data sets, we would receive a global ML-model, which includes the trained features of the fields of all farmers. A basic implementation of this approach was proposed in [Kh22]. In this case even smaller farmers could participate and benefit from a well enough trained ML-model, while the general will not lose individual and maybe raw data with specific features. Also farmers without a disease outbreak on their fields do not need to fear that the ML-model could not detect a next outbreak or even new discovered diseases, since the global ML-model will be fed with enough examples from other farmers.

2.1.2 Cross Silo Learning using Vertical FL

The last example showcased some great advantages for the farmers from a Horizontal FL ML-model. But not only farmers could benefit directly from FL, but also companies. Consider a company is providing a ML-model to it's customers for, e.g., crop yield prediction. The ML-model can predict the farmer's crop yield based on, e.g., soil and crop conditions received from, e.g., the company's Farm Management Information System (FMIS). The Prediction would be done more or less well, since it is based only on two parameters, but does not include, e.g., the weather. There could be another ML-model from another

company (or even the government), which predicts the crop yield based on meteorological parameters and maybe also another one predicting the crop yield based on environmental parameters. All this ML-models may make a good prediction, but they would with a high probability do a better job when working together. That may sound simple, but farmers may want to only share their sensitive data in the FMIS of their trust and would not like it if this is also being shared with a third (or even fourth) party. Also the companies may not have any interest in sharing their technology of their ML-models to others.

Here, *Vertical FL* [SVG23], [IE21], [We22] could resolve the issue by aggregating the predictions and sending back a feedback, which includes the information of all participants, to the separate ML-models, which can use this to update themselves individually. The ML-models will share their experiences to each other and thus each one of them should be able to do a better prediction, even though it will lack of some information.

2.2 System Requirements and Options for Effective FL Deployment

As a distributed ML paradigm, FL promises a data sovereign way of knowledge sharing without the need to disclose private data. However, appropriate technical means to foster the adoption of this ML concept are needed. *Digital Agriculture Platforms* offer a unified way for farmers to use digital software services specifically targeted for the agricultural domain. Such platforms (e.g., [Bo23] or [Ag23]) offer farmers the opportunity to connect to so called *agricultural data spaces* [Ka22] complying to international data privacy regulations and therefore preserving data sovereignty to satisfy farmers' privacy concerns. This allows to securely broker data of farmers with specialized digital services offered by (potentially third-party) service providers. When farmers are willing to re-share their models (not their data) in a FL style with other clients sharing similar conditions, local adaptations of AI services become feasible. A central entity as in typical client server architectures can then take charge of fusing the numerously incoming local models into an improved new version of the generalized global model. Enabling FL on such digital service platforms therefore not only allows AI service providers and, more importantly, the end users (farmers) to improve their models without sharing critical holding data. It further enables *continual learning* [Yo21], which allows to incorporate newly generated process data into the local models without forgetting older training cases – in our view an integral aspect for exploiting the full potential of AI in various agricultural contexts. As briefly stated before, FL requires appropriate technical infrastructure on the edge, that is, at the farm holdings to ensure data sovereignty, what induces investment costs. In today's farming practice, farmers already engage in cooperatives sharing orders of magnitude more expensive machinery. This sharing principle can be applied to IT infrastructure, too. Neighbouring farm holdings could engage in joint activities of computer hardware, mobile sensory equipment for local data collection and network infrastructure investment as well as concluding technical consulting contracts for their maintenance. Although the agricultural system exhibits high heterogeneity on larger spatial scales, looking at smaller regional scales, environmental conditions and natural

circumstances (e.g., growing conditions, soil types, use of certain machinery) can also be quite similar. This situation could in turn be exploited by defining *regional representative farm holdings* being highly digitized and well equipped with technical facilities to locally and continually train predictive models using FL. Such representative farm holdings might emerge as intermediate levels on the edge-cloud continuum ranging from individual farms (edge) to the entire digital data and service ecosystems integrating data from participating (client) farms and services from third-party providers on vastly larger spatial and also temporal scales (cloud). In contrast to classical client-server approaches mainly followed by the platform paradigm, setting up reference farm holdings on (maybe even various spatial scales) might call for hybrid peer-to-peer networks, directly connecting the participating farm holdings and thereby accounting for trusted, targeted and, hence, even more sovereign information exchange.

3 Concluding Remarks

One obvious prerequisite for FL to find its way to practical use is the assurance of a reliable network connection between collaborating clients and the facilitating platforms or cloud services. However, there are ways to make FL operational even if the connection for some clients is temporarily not available, e.g., by setting a threshold for the number of participants at each round and skipping the update, if not reached. Another non-FL specific challenge concerns the data labelling task, since the farmers may not be qualified or just have no time to label collected data themselves. A possible solution to this could be unsupervised approaches to FL [Lu22]. The so-called client drift [Ka20] issue, which characterizes the challenge of finding an optimal solution between the global optimum and the local optima, could be resolved by implementing a *Personalized Federated Learning* [Ma20], [Ta22] approach. A possible way to disrupt or even sabotage a FL-system would be given by manipulating clients and their data sets, e.g., by creating exaggerated imbalanced data sets, which would lead to an unrealistic local optimum and thus affect the global model. Nevertheless this scenario can be counteracted by encouraging more clients to collaborate, since this would make any noise of a small group of clients mathematically less relevant. We elaborated on the adoption of the emerging learning paradigm *Federated Learning* to digital agriculture.

Our discussions stressed system requirements which we deem essential to be addressed in order to pave the way for expedient adoption of a distributed and data sovereign ML technology to naturally highly heterogeneous agricultural settings. Potential use cases have been envisioned and mapped to specific FL approaches that allow for their realization, even though in order to demonstrate economic feasibility, business cases to need to be thought and described in more detail, as suggested in [NN11]. To support a more thorough analysis and definition of the potential use cases we only touched upon in this paper, in future work we pursue the development of a theoretical framework that supports a systematic identification and guides practical realization of specifically suitable agricultural use cases that can strongly benefit from FL.

References

- [Ag23] Agri-Gaia: [<https://www.agri-gaia.de/>, accessed 22.05.2023].
- [Bi23a] Bitkom: Die Digitalisierung der Landwirtschaft, [<https://www.bitkom.org/Presse/Presseinformation/Digitalisierung-sichert-Zukunft-der-Landwirtschaft>, 2022, accessed 22.05.2023].
- [Bi23b] Bitkom: Digitalisierung in der Landwirtschaft 2020, [https://www.bitkom.org/sites/main/files/2022-01/27.04.20_digitalisierung-in-der-landwirtschaft-2020.pdf, 2020, accessed 22.05.2023].
- [Bo19] Bonawitz, K.; Eichner, H.; Grieskamp, W.; Huba, D.; Ingerman, A.; Ivanov, V.; Kiddon, C.; Konečný, J.; Mazzocchi, S.; McMahan, H. B.; Overveldt, T. V.; Petrou, D.; Ramage, D.; Roselander, J.: Towards Federated Learning at Scale: System Design, 2019, arXiv: 1902.01046 [cs.LG].
- [Bo23] Bosse, S.; Berns, K.; Bosch, J.; Dörr, J.; Eichhorn, F. C.; Eisert, P.; Fischer, C.; Gassen, E.; Gerstenberger, M.; Gerighausen, H.; Heil, J.; Hilsmann, A.; Hirth, J.; Huber, C.; Hussaini, M.; Kasparick, M.; Kloke, P.; Krause-Edler, H.; Mackle, L.; Magnusson, J.; Möhrle, F.; Möller, M.; Pickel, P.; Rautenberg, C.; Schotten, H. D.; Stanczak, S.; Thiele, L.; Ücdemir, H.; Wania, A.; Stein, A.: Nachhaltige Landwirtschaft mittels Künstlicher Intelligenz – ein plattformbasierter Ansatz für Forschung und Industrie. In (Hoffmann, C.; Stein, A.; Ruckelshausen, A.; Müller, H.; Steckel, T.; Floto, H., eds.): 43. GIL-Jahrestagung, Resiliente Agri-Food-Systeme. Gesellschaft für Informatik e.V., Bonn, pp. 41–52, 2023.
- [De22] De Lange, M.; Aljundi, R.; Masana, M.; Parisot, S.; Jia, X.; Leonardis, A.; Slabaugh, G.; Tuytelaars, T.: A Continual Learning Survey: Defying Forgetting in Classification Tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44/7, pp. 3366–3385, 2022.
- [Gu22] Gupta, S.; Ahuja, K.; Havaei, M.; Chatterjee, N.; Bengio, Y.: FL Games: A federated learning framework for distribution shifts, 2022, arXiv: 2205.11101 [cs.LG].
- [Ha20] Hadsell, R.; Rao, D.; Rusu, A. A.; Pascanu, R.: Embracing Change: Continual Learning in Deep Neural Networks. *Trends in Cognitive Sciences* 24/12, pp. 1028–1040, 2020, ISSN: 1364-6613, URL: <https://www.sciencedirect.com/science/article/pii/S1364661320302199>.
- [IE21] IEEE: IEEE Guide for Architectural Framework and Application of Federated Machine Learning. *IEEE Std 3652.1-2020/*, pp. 1–69, 2021.

- [Ka20] Karimireddy, S. P.; Kale, S.; Mohri, M.; Reddi, S.; Stich, S.; Suresh, A. T.: SCAFFOLD: Stochastic Controlled Averaging for Federated Learning. In (III, H. D.; Singh, A., eds.): Proceedings of the 37th International Conference on Machine Learning. Vol. 119. Proceedings of Machine Learning Research, PMLR, pp. 5132–5143, 13–18 Jul 2020, URL: <https://proceedings.mlr.press/v119/karimireddy20a.html>.
- [Ka22] Kalmar, R.; Rauch, B.; Dörr, J.; Liggesmeyer, P.: Agricultural Data Space. In (Otto, B.; ten Hompel, M.; Wrobel, S., eds.): Designing Data Spaces : The Ecosystem Approach to Competitive Advantage. Springer International Publishing, Cham, pp. 279–290, 2022, ISBN: 978-3-030-93975-5, URL: https://doi.org/10.1007/978-3-030-93975-5_17.
- [Kh22] Khan, F. S.; Khan, S.; Mohd, M. N. H.; Waseem, A.; Khan, M. N. A.; Ali, S.; Ahmed, R.: Federated learning-based UAVs for the diagnosis of Plant Diseases. In: 2022 International Conference on Engineering and Emerging Technologies (ICEET). Pp. 1–6, 2022.
- [Le19] Lei, S.; Zhang, H.; Wang, K.; Su, Z.: How Training Data Affect the Accuracy and Robustness of Neural Networks for Image Classification, 2019.
- [Lu22] Lu, N.; Wang, Z.; Li, X.; Niu, G.; Dou, Q.; Sugiyama, M.: Federated Learning from Only Unlabeled Data with Class-Conditional-Sharing Clients, 2022, arXiv: 2204.03304 [cs.LG].
- [Ma20] Mansour, Y.; Mohri, M.; Ro, J.; Suresh, A. T.: Three Approaches for Personalization with Applications to Federated Learning, 2020, arXiv: 2002.10619 [cs.LG].
- [Mc17] McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B. A. y.: Communication-Efficient Learning of Deep Networks from Decentralized Data. In (Singh, A.; Zhu, J., eds.): Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Vol. 54. Proceedings of Machine Learning Research, PMLR, pp. 1273–1282, 20–22 Apr 2017.
- [MR23] McMahan, B.; Ramage, D.: Federated Learning: Collaborative Machine Learning without Centralized Training Data, [<https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>], 2017, accessed 22.05.2023.
- [NN11] Neef, A.; Neubert, D.: Stakeholder participation in agricultural research projects: a conceptual framework for reflection and decision-making. Agriculture and Human Values 28/, pp. 179–194, 2011.
- [Sh18] Shallue, C. J.; Lee, J.; Antognini, J.; Sohl-Dickstein, J.; Frostig, R.; Dahl, G. E.: Measuring the effects of data parallelism on neural network training. arXiv preprint arXiv:1811.03600/, 2018.
- [SVG23] Silva, P. R.; Vinagre, J.; Gama, J.: Towards federated learning: An overview of methods and applications. WIREs Data Mining and Knowledge Discovery 13/2, e1486, 2023.

- [Ta22] Tan, A. Z.; Yu, H.; Cui, L.; Yang, Q.: Towards Personalized Federated Learning. *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–17, 2022.
- [TMG22] T, M.; Makkithaya, K.; G, N. V.: A Federated Learning-Based Crop Yield Prediction for Agricultural Production Risk Management. In: *2022 IEEE Delhi Section Conference (DELCON)*. Pp. 1–7, 2022.
- [US23] USAID: Data-Driven Agriculture: The Future of Smallholder Farmer Data Management, [<https://www.usaid.gov/digitalag/documents/data-driven-agriculture>, 2018, accessed 22.05.2023].
- [We22] Wei, K.; Li, J.; Ma, C.; Ding, M.; Wei, S.; Wu, F.; Chen, G.; Ranbaduge, T.: Vertical Federated Learning: Challenges, Methodologies and Experiments, 2022, arXiv: 2202.04309 [cs.LG].
- [Yo21] Yoon, J.; Jeong, W.; Lee, G.; Yang, E.; Hwang, S. J.: Federated continual learning with weighted inter-client transfer. In: *International Conference on Machine Learning*. PMLR, pp. 12073–12086, 2021.

Ökologische Nachhaltigkeit -
Workshop on Systems to Support
Renaturation Projects 2023
(SyRePro23)

Semi-automatic extraction of metadata from old geological maps

Kim Bürgl¹, Lydia Müller²

Abstract: Geological maps communicate efficiently geological information. Old geological maps were stored as paper maps and thus need to be digitized when integrating them into digital geographic information systems. Metadata is required to find relevant maps fast. However, metadata is usually created manually with a lot of effort. We present work in progress for a semi-automated approach for extracting metadata from maps. The results show that it lowers the manual effort significantly to extract the location and improves at least the experience of the manual annotation with respect to date metadata.

Keywords: NLP, Metadata, OCR, geological maps

1 Introduction

Geological maps are an efficient way to communicate geographic data. Nowadays, maps are usually stored digitally, allowing arbitrary scaling of the region of interest. Geographic information systems (GIS) are often used to store and display those maps alongside with additional data such as metadata, annotations or further data from different sources for the same geographic location. However, old geographic maps are often stored analogous, i.e. as paper map carrying as well metadata in form of legend and titles and annotations.

Digitizing paper maps allows not only to store them in a digital format and distribute them more easily, it also allows combining them with other maps in a GIS. The process of digitizing paper maps includes not only scanning the maps, but also extracting the relevant metadata from the map and transforming them in the correct format [Lo05].

Metadata extraction often means extracting textual information on the map and classify them into metadata categories. Optical character recognition (OCR) can be used to extract such information. Both, OCR and classification of the extraction data into metadata categories are erroneous and thus, lead to uncertainties in the resulting metadata. Thus, metadata is often entered manually [Lo05] which is a very time-consuming and labor-intensive task.

Since fully automated solutions do not produce reliable data and manual solutions are too costly, semi-automated solutions are developed. Budig et al. [BDW14], for example,

¹ Institut für Angewandte Informatik (InfAI), Goerdelerring 9, 04109 Leipzig, Deutschland; Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Deutschland buergl@informatik.uni-leipzig.de

² Institut für Angewandte Informatik (InfAI), Goerdelerring 9, 04109 Leipzig, Deutschland; Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Deutschland lydia@informatik.uni-leipzig.de

showed with their work on historical maps that an approach, which asks a user in cases with low certainty is beneficial to improve data extraction, but they focused on matching labels and markers. When a user corrects the matching of one label with a marker, their greedy algorithm is rerun and produces corrected pairs of matched pairs. Van Dijk and Wolff [DW17] argue, that algorithmically guided user interaction can be applied when a problem is not solvable fully automatically, with the example of historic map matching.

In our use case, we are developing a semi-automatic approach to extract metadata from old geological maps of closed open pit mines. Those maps will be integrated into a GIS which will make use of the metadata to find and visualize relevant geological information for specific regions (GIS, cf. [An22]) when analyzing the potential reuse of the regions. Therefore, of particular interest is not only the map itself and information on the geographic location, but also the latest date of change of the map.

2 Data Set and its Challenges

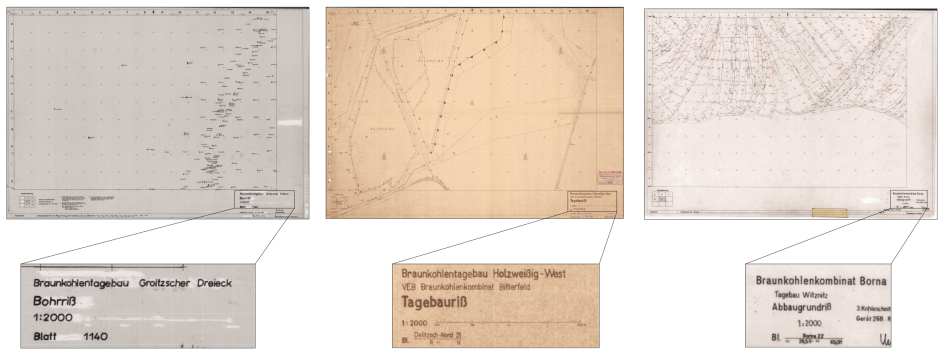


Fig. 1: Three types of geological maps with different metadata layout. One can see the metadata open pit mine name (Groitscher Dreieck, Holzweißig-West, Witznitz, from left to right), the map type (Bohrriß, Tagebauriß, Abbaugrundiß) and the map segment (144, 21, 22)

Our dataset is provided by the Lausitzer und Mitteldeutsche Bergbau-Verwaltungsgesellschaft mbH (LMBV), a state-owned company managing former open pit mining sites in Eastern Germany. It consists of 141 paper maps from the years between 1960 and 1990. Each map has a title expressing the name of the open pit mine shown, the type of the map and the scale. To depict the details on the paper map, mostly, more than one physical paper was required to depict the full open pit mining site. Thus, the number of the paper, i.e. the segment of the map shown, is stated in the title. Example maps can be seen in Figure 1. While the map itself has a date, the marker for the region updated and annotated the maps. Each update and annotation was documented by adding the corresponding date on the map. Thus, we found multiple dates on the same maps. However, we know, that maps date back to the 1960s and that there are no updates or new maps after the 1990s.

Relevant information, which need to be extracted, are the name of the open pit mining site as stated in the title, the type of the map, the number of the paper and at least the latest date. We identified the following challenges of the dataset for the extraction process.

- **Layout:** Figure 1 shows that even though the overall layout of the maps are very similar, the layout of the title varies significantly. The title of the left map consists of the name of the open pit mine (first line), the map type (second line), the scale (third line), and the map segment (fourth line). The title of the map in the middle consists of the name of the open pit mine (first line), the responsible company (second line), the type of the map (third line), the scale (fourth line), and the map segment (fifth line). In the title of the left map, the first and second line are switched when compared to the map in the middle. Thus, it is not clear which lines of the map title contain the relevant information.
- **Layout of map segment:** The map segment states which part of the map is shown. In the simplest case, it is noted as the word „Blatt“ and the map segment as a number. However, different layouts are possible, e.g. Figure 1 middle and left map. The information is found in the last line. It starts with the characters "Bl.". The map segment is preceded by a name and placed above a line. However, sometimes the map segment information is missing in the title, but given only in a graphical way. This variant, even though it is also stated in the title, can be seen in Figure 1 in all three maps in the left lower corner.
- **Scan Quality:** The scan quality sometimes leads to problems, e.g. the information which is only partly visible or low contrast.
- **Fonts:** The different fonts lead to multiple problems. One problem is handwritten information which is sometime even hard to read for a human but also uses frequently abbreviation, e.g. the dates. Another problem is the mix of different fonts (including handwritten fonts). We can observe different common replacements in the recognition of the characters depending on the font. Since we do not know the font when extracting information, we cannot apply an automated error correction based on common replacements.
- **Date Formats:** We find almost every date format on the maps, and it is even inconsistent within a single map. Figure 2 shows some examples. Some dates are given as year, some as month and year, some as day, month and year. The month is given either as Latin numbers, Roman numbers, the full name or as an abbreviation. The year is either specified with four or two digits. The separator sign can be space, dot or slash. Moreover, dates may be handwritten, printed or stamped on the map. Whether the date is extracted completely, is therefor hard to evaluate automatically.
- **Dates Position:** Dates can be found at many positions on the map. They are part of the title or a legend, but may also locate at the bottom of the map. Some are preceded

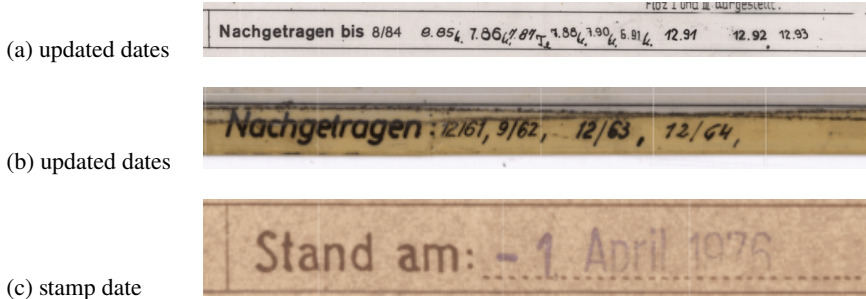


Fig. 2: Dates on maps: (a) and (b) show dates at the bottom of the map documenting the updates of the map. Dates have different formats and are separated with different characters. Sometimes the name of the marker is added. (c) The date may be stamped on the map.

with specific words, some not. Thus, it is hard to evaluate automatically if all dates were found or not.

- **Names of open pit mines:** Sometimes the name of the mining site in the title and the mining site name on the geological map differ (e.g. due to renaming or fusion of mining sites). Thus, it is important to capture the name in the title, not the one on the map. Finding the title position is therefore important.

Given all these challenges, OCR results will not be sufficiently good to extract reliable metadata and an automatic error correction is not possible due to too many uncertainties. However, reliable metadata is required to find the relevant maps in the GIS for a given use case.

As a solution, we propose here a semi-automatic approaches for extracting the metadata from the maps. If the system cannot extract a metadata with high certainty, a human is asked to annotate the metadata. As most maps share some of their layout, metadata such as the name of a pit mine, the type of map and the map section are often written in the same location. Therefore, if a human annotator annotated a map, this information can help to automatically extract this metadata for unseen maps of the same layout. Thus, lowering the manual labor for the whole data set.

Date metadata is more variable, therefore an approach which uses the location of dates on similar maps is not feasible. For this reason, we propose a different approach which focuses on anchor words which sometimes appear before dates and monotony if update dates were used.

Note that the approach is work in progress and still leaks a rigorous evaluation. However, a preliminary evaluation shows promising results.

3 Semi-automatic metadata extraction

We aim at the extraction of the location as open pit mine name, the map type, and the map section, and the date of the last map update. The first three metadata identify maps uniquely within the dataset and are found at similar locations, while the last metadatum identifies the version of the map and can be found anywhere on the map. Due to the differences between location and date, we developed two different approaches.

Our approach for the location consists of two main modules, a similarity measure and the metadata extraction module as shown in figure 3. The similarity measure calculates similarity between maps on the basis of extracted text snippets and their location. The metadata extraction module extracts metadata automatically, given similar maps, which are already annotated with metadata. Our approach for the dates is based on key phrases and date patterns.

Next, we will describe the approach for extracting the location. Afterward, we present the approach for extracting the dates.

3.1 Location Metadata

- **Feature Extraction:** For our proposed similarity measure and the metadata extraction, we use text snippets and their bounding boxes as features. In our prototype, we used the EAST text detection [Zh17] to detect text snippets and their location, and tesseract[Sm07] to extract the text of these snippets.
- **Similarity:** The similarity between two maps (map *a* and map *b*) is calculated as follows: The initial similarity score is set to 0. Each text snippet in *map a* is compared with text snippets in *map b* if they are placed in the same region. The threshold if two text snippets lie in the same region depends on the quality of the scans and has to be defined empirically. If a text snippet of *map a* is in the same region as a text snippet of *map b* and the text is sufficiently similar, the length of the text snippet is added to the similarity score. If the text snippets in *map a* and *map b* are of different length, we choose the shorter one. We currently choose to accept 2 errors for text snippets longer than 5 and one error for text snippets with length 5 or shorter. Choosing to add the length of a found text snippet to the similarity score puts weight on longer sequences, as mostly map description such as the name of the open pit mine, but also some layout description is most helpful for deciding if the layouts of two maps are similar. Note that the map's content, e.g. the drawing of geological conditions and so on is not useful for map similarity, as maps with the same layout but which show a different section of a mine should be similar too.
- **Metadata extraction by region:** For the metadata, which are normally placed in the same region on a map (see figure 1), we propose a semi-automatic approach. If the system is not able to extract the metadata itself due to uncertainty or missing similarity

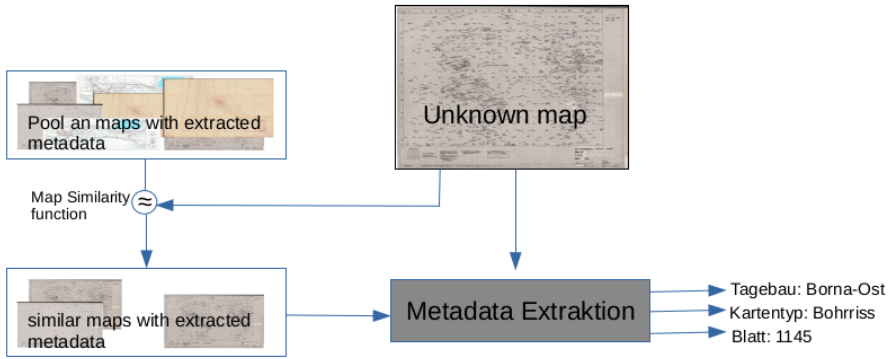


Fig. 3: Schematic figure of the proposed semi-automatic extraction pipeline.

to other maps, a user explicitly selects the region and the value of a metadata. The approach is shown in figure 3, code can be found in Listing 1 and works as follows: For each unseen map, the most similar annotated maps are found with the similarity measure described in section refsec:ml, Similarity. If no similar annotated maps were found, a user annotates the map. That is, for each metadata type, they annotate the bounding box and type in the value of that metadata. Then, the map is added to the annotated maps. If similar annotated maps are found, each extracted feature (see 3.1, Feature Extraction) is compared to the annotated metadata of the similar maps. If a feature in the map to annotate intersects with the metadata in a similar annotated map, the similarity of the text is checked. As the OCR results are error-prone, we use the levensthein distance to decide if a text is sufficiently similar to a verified metadata. For the most similar text according to the levensthein distance, we calculate the ratio of text length to levensthein distance as a threshold (equation 1). If $threshold > 7$, a metadata is accepted without human control, and if $threshold \geq 2$ a human annotator is asked for confirmation. If the threshold is lower, a human annotator has to annotate the metadata manually. These thresholds were found empirically but need to be investigated more formally.

$$threshold = \frac{textlength}{levensthein} \quad (1)$$

For map sections, we propose to use the prefix phrases as anchors which the user needs to annotate. Then, only the extracted numbers need to be checked and corrected. We decided to not accept retrieved metadata automatically in that case, as a single error of a digit leads to wrong results. Nevertheless, the effort of a user is significantly reduced, if they only have to accept a retrieved value instead of annotating it manually.

```

def pitmineExtraction(maps)
    processed = list()

    for map in maps:
        #Extracting text snippets and their localization on the map
        map.Features = OCR(map)}

        for feature in map.Features:
            for processedMap in processed:
                if match(feature, processedMap.pitmine):
                    map.pitmine.text = processedMap.pitmine.text
                    map.pitmine.loc = feature.loc
    if map.pitmine is None:
        map.pitmine = AskUser(map)
    processed.append(map)

```

List. 1: Semi-automatic metadata extraction

3.2 Date Metadata

While most of the metadata is placed at the same location in the map description, dates are often placed as stamps or by hand without a fixed location on the document. Therefore, we propose the following approach to find the most recent date. Again, we try to extract the dates automatically but put the human in the loop if the uncertainty of extraction is too high.

In our dataset are three different types of date markings. For each of them, we defined an extraction routine. On a map, usually only one type of date marking is found. Since we do not know the type beforehand and want to be robust against new layouts, we try to extract dates for all three types of the date markings and afterward select the latest date.

- **Single Dates after key phrase:** The first type is dates, which are found after a special phrase marking it as a single date (see Figure 2 c). If such a date is found, we assume it is the only date after the key phrase and is thus collected as a candidate for the most recent date. To extract the date, we first search in the text snippets containing the key phrase. The region is extended to accommodate the date. The text is extracted optimized for dates and a date pattern searched. On success, the date is extracted. On failure, a user is asked for an annotation.
- **Single Dates without a key phrase:** Some dates were found anywhere in the document, without any key phrase attached. These are mostly stamps. Text from text snippets is extracted optimized for dates. If a date pattern is found with sufficient confidence, the date is accepted as candidate. As we cannot be sure to have found the most recent date in that case, a human annotator is asked for confirmation.
- **Key phrases indicating multiple map updates:** Some maps are updated multiple times. This is indicated with a specific key phrase. The dates are mostly written by

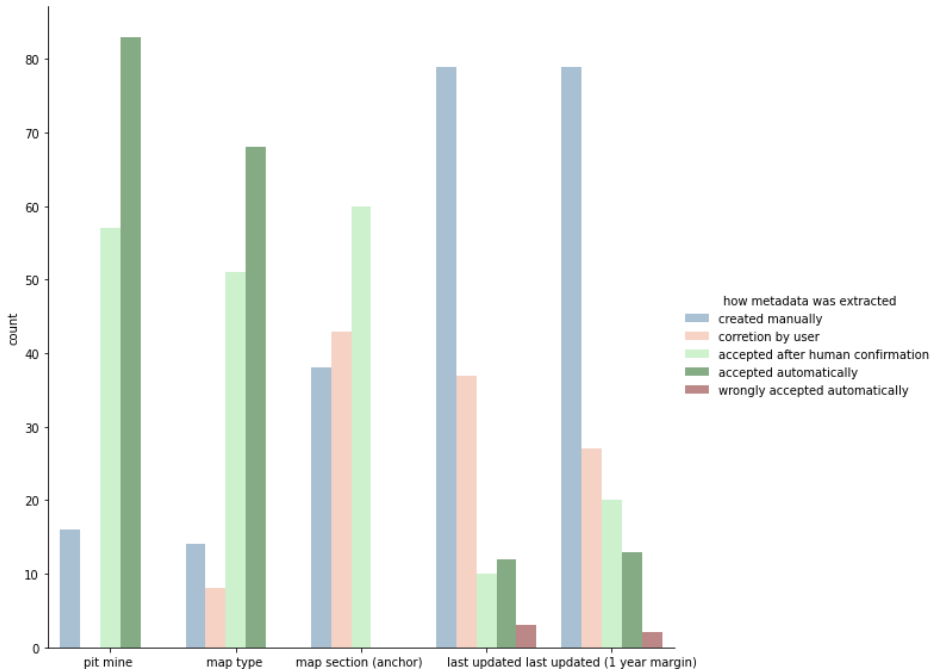


Fig. 4: Bar plot showing the outcome of a metadata extraction session by a user. For each metadata type, the total number of manually created metadata, metadata which was created by the user, metadata which was automatically extracted and confirmed by the user, metadata which was accepted automatically and metadata which was wrongly accepted automatically is shown.

hand (see figure 2 a and b). As these dates are difficult to detect, because they are mostly handwritten, we use multiple OCR pipelines to extract the dates and their position to one another. Again, relying on text snippets located right to the key phrase. If dates are not monotonously ascending from left to right, we have to assume that an error occurred, as the geologists normally wrote the updated date to the right of the existing dates. That, and the self-acclaimed certainty of the OCR method are used to define, if a date can be trusted or not. Depending on the trustworthiness, a user is asked for confirmation.

4 Evaluation

We extracted the metadata of the 141 maps provided by the LMBV manually. This data serves as bases for the evaluation. Note that some thresholds of the approach are chosen based on the maps at hand. In a more, rigorous evaluation of the final approach, maps not seen before should be used.

We evaluated the reduction of manual work and correctness of the extracted metadata. The results are summarized in Figure 4. The first three bar plot groups correspond to the location metadata, i.e. name of the open pit mine, map type, and map section. The last two bar plot groups belong to the date metadata. We evaluated the full date as extracted as well as the date with a 1-year margin to account for smaller errors which often play a minor role in finding the correct version.

The reduction of manual work strongly depends on the type of metadata. Most reduction can be found for the map type and open pit mine name, where 14 and 16 maps have to be annotated manually, respectively. For eight more maps, the map type had to be corrected by the user. However, for the majority of the maps, the open pit mine name and the map type could be inferred with no or minimal effort of a user. The map section needed more manual work. Still, for 60 maps only the user confirmation without any correction was required. Note, that the map section was in any case shown to the user. Most manual effort was required for the date of last update. About 80 maps needed to be annotated manually. About 40 maps had to be corrected by the user. Allowing a one-year margin, reduces the correction to 30 maps. This results mainly depend on the large variation in the layout of dates and the fact, that many dates are handwritten. Thus, many more OCR errors occurred, making it hard to automatically correct them rule based.

We observed no extraction errors for the location metadata. Thus, if we extracted the metadata open pit mine name or map type automatically, it was the correct metadata. Any wrongly identified open pit mine or map type were identified by our approach as cases needing manual confirmation by a user. In the case of dates, this did not work sufficient good. There are three maps with incorrect dates which were automatically accepted, i.e. our uncertainty measure did not identify them as uncertain. Two of the three maps relates to errors in the date larger than 1 year.

In summary, we observe very good result with respect to the extraction of location metadata. Hereby, the map section recognition could be improved further. The results for the date still need improvement. Not only in recognizing the dates, but also the extraction of the dates, need further research to foster the possibilities for improvement. For example, it is an option to improve the OCR step by building a model for handwritten date formats. Also, the measurement of uncertainty in the recognition of the dates need to be improved.

5 Summary

We presented work in progress showcasing the semi-automatic extraction of metadata from legacy maps. We developed and implemented a semi-automatic approach and performed a preliminary evaluation on a dataset of 141 maps. It showed that the approach can drastically reduce the effort needed to tag and extract metadata from maps, which is usually performed manual. Especially string metadata which are existent over multiple maps in the same layout such as the name of an open pit mine or the type of map, such an approach seems to be

quite performant and could reduce the effort needed to tag such maps by a lot. For numeric data such as dates and the map sections, the correction of OCR errors is a lot harder and the reduction of the workload is less impressive. However, it still augments the experience of a human annotator by making the work is less monotonous, i.e. it switches between manual annotation and the examination of extracted metadata.

References

- [An22] Annanias, Y.; Zeckzer, D.; Scheuermann, G.; Wiegrefe, D.: An Interactive Decision Support System for Land Reuse Tasks. *IEEE Computer Graphics and Applications* 42/, pp. 72–83, 2022.
- [BDW14] Budig, B.; van Dijk, T. C.; Wolff, A.: Matching labels and markers in historical maps: an algorithm with interactive postprocessing. In: *MapInteract '14*. 2014.
- [DW17] van Dijk, T. C.; Wolff, A.: Algorithmically-guided user interaction. In: *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. Pp. 1–4, 2017.
- [Lo05] Longley, P. A.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W.: *Geographic information systems and science*. In: John Wiley & Sons, p. 207, 2005.
- [Sm07] Smith, R.: An Overview of the Tesseract OCR Engine. *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)* 2/, pp. 629–633, 2007.
- [Zh17] Zhou, X.; Yao, C.; Wen, H.; Wang, Y.; Zhou, S.; He, W.; Liang, J.: EAST: An Efficient and Accurate Scene Text Detector. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*/, pp. 2642–2651, 2017.

Nachnutzungskonzept für Daten in der Altlastenverwaltung

Lydia Müller,¹ Daniel Wiegrefe²

Abstract: Große Flächen in Deutschland wurden durch Bergbau in Anspruch genommen. Ein Beispiel hierfür sind die zahlreichen Braunkohletagebaue, welche aufgrund des Ausstieges aus der Kohleverstromung bereits jetzt oder bald nicht mehr aktiv betrieben werden. Die Sanierung und Wiedernutzbarmachung dieser Gebiete ist ein sehr aufwendiger Prozess, bei dem zahlreiche Daten durch zahlreiche Akteure erhoben und genutzt werden. Inzwischen gilt die Verwaltung alter Bergbaugebiete als Ewigkeitsaufgabe, sodass die erhobenen Daten auf lange Zeit verfügbar und nutzbar sein müssen. In diesem Paper stellen wir ein Nachnutzungskonzept für die Daten vor, welches auf die Verwaltung der Daten als Ewigkeitsaufgabe auffasst.

Keywords: Nachnutzungskonzept, Altlasten, Ewigkeitslast, Verwaltung, Digitalisierung

1 Einleitung

Alleine in Deutschland sind über 1400km² an Flächen durch den Bergbau in Anspruch genommen worden[Wi23]. Diese Flächen müssen nach Beendigung der Bergbautätigkeiten weiter gesichert werden, damit sie keine Gefahr für die Bevölkerung und die Umwelt darstellen. Meist geht mit dieser Sicherung auch eine umfangreiche Renaturierung einher, damit die Flächen für eine weitere Nutzung wieder verfügbar gemacht werden können. An diesen Maßnahmen sind eine Vielzahl von Experten aus den unterschiedlichen Ingenieurbereichen beteiligt, da es sich meist um komplexe und zeitaufwendige Prozesse handelt. Auch können manche Regionen nach Beendigung der Bergbautätigkeit nicht wiederhergestellt können und bleiben dauerhaft gefährlich, beispielsweise durch die Gefahr von Absenkungen oder massiven Umweltbelastungen im Abraum.

In den letzten Jahrzehnten sorgte die Digitalisierung auch in diesem Bereich für Umstellungen und Veränderungen. So wurde damit begonnen, eine Vielzahl an Daten zur Dokumentation und Planungen für diese Flächen zu erheben. Diese Daten spielen in einer Vielzahl von Prozessen eine wichtige Rolle, zum Beispiel auch in der nachträglichen Nutzung. Neben Geodaten und Sensordaten fallen auch viele textuelle Daten an, die nicht direkt maschinenlesbar[Sc21] verwertet werden können. Zudem werden diese Daten oft in mehrere Datensilos getrennt von einander abgelegt, da sie durch die verschiedenen Akteure erzeugt werden. Jedoch ergeben sich durch die Verschneidung der Daten hohe

¹ Institut für Angewandte Informatik (InfAI), Goerdelerring 9, 04109 Leipzig, Deutschland; Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Deutschland lydia@informatik.uni-leipzig.de

² Abteilung für Bild- und Signalverarbeitung, Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Deutschland daniel@informatik.uni-leipzig.de

Synergieeffekte, da Planungsprozesse effizienter gestaltet werden können und die Daten für spätere Anfrage verfügbar gemacht werden können. Dies ist insbesondere mit Hinblick auf das Ausscheiden von Wissensträgern aus den betreffenden Unternehmen nützlich beziehungsweise bei Übergabe der Daten an die überwachenden Behörden, beispielsweise nach Beendigung der Bergaufsicht.

In diesem Beitrag skizzieren wir ein Nachnutzungskonzept für diese Daten. Dazu beschreiben wir zuerst den Status Quo in Bezug auf Deutschland am Beispiel des Braunkohletagebaus. Darauf aufbauend entwickeln wir die Anforderungen an eine Nachnutzungskonzept für diese Daten aus Sicht der Informatik.

2 Status Quo

In Rahmen von bergbaulichen Maßnahmen wie Braunkohletagebaue werden bereits während des Betriebes viele Daten erhoben. Dabei gibt es je nach Status der bergbaulichen Maßnahme verschiedene Daten und Stakeholder. Wechselt der Status der bergbaulichen Maßnahme, zum Beispiel durch das Auslaufen der Bergaufsicht, kann die Verantwortung und somit auch die Stakeholder über die bergbauliche Maßnahme sich ändern. Teile der Daten werden an die entsprechenden neuen Stakeholder übergeben, jedoch meist nicht der komplette Datenbestand. Ältere Daten, die analog erhoben wurden, werden auch analog, also in Papierform übergeben. Moderne Daten werden zwar zumeist digital erhoben, aber oft noch analog archiviert und weitergegeben.

2.1 Daten

In aktiven bergbaulichen Maßnahmen liegen die Daten zumeist bei dem Bergbauunternehmen. Die Daten dienen des sicheren Betriebs der bergbaulichen Maßnahme, der Planung und Verwaltung. So werden Karten und Risswerke gefertigt, Messungen, z.B. unter Nutzung von Sensoren, erhoben und protokolliert und Gutachten erstellt. Gemäß der Gesetzeslage werden diese Daten nicht nur erhoben, sondern auch archiviert.

Nachdem eine bergbauliche Maßnahme beendet ist, bedarf es einer dauerhaften Sicherung bzw. bei Tagebauen oft eine Wiedernutzbarmachung der Bergbaufolgelandschaft. Eine Nachnutzung ist nicht direkt möglich, sondern die Landschaft muss erst aufwendig saniert werden. Auch während dieser Prozesse gibt es verschiedene Akteure mit unterschiedlichen Datenbeständen.

Zum Einen sind für ein Teil der Sanierungsmaßnahmen die Bergbauunternehmen selber zuständig, zum anderen gibt es auch Nachfolgegesellschaften wie die Lausitzer und Mitteldeutsch Bergbauverwaltungsgesellschaft (LMBV) mbH, deren Aufgabe es ist, Tagebauregionen nachzuverwalten und zu sanieren. Dabei entstehen eine Vielzahl unterschiedlicher Daten. Oft wird aber ein Großteil der Sanierung durch weitere Unternehmen geleistet, die

sich auf die Sanierung spezialisiert haben. Dabei entstehen weitere Daten unterschiedlichster Natur, um die Sanierung zu dokumentieren. Gutachten dokumentieren den Zustand der Regionen und definieren was in diesen Regionen für Maßnahmen durchgeführt werden müssen bzw. für welche Nachnutzung sie geeignet sind. Regelmäßige Dokumentation des Zustandes von Regionen führen zu Berichten in Form von Wort und Bild. Pläne dokumentieren die geplante Nutzung, Kommunikation, also Anfragen und Antworten, dokumentieren wie die Region bereits genutzt werden. Vorschriften bzw. Verordnungen zur Nutzung der Region müssen auch dokumentiert werden.

Die endgültige Aufsicht und Verwaltung von Bergbaufolgelandschaft obliegt zumeist der öffentlichen Verwaltung, in Deutschland ist dies föderal geregelt, z.B. durch das Oberbergamt Sachsen oder die Bezirksregierung Arnsberg. Raumplanungen und Verordnung werden durch diese umgesetzt und überwacht.

2.2 Nutzer und Nutzungsszenarien

Vier große Nutzergruppen der Daten sind die Bergbauunternehmen, die Sanierer, die öffentlichen Verwaltungen und private/gewerbliche Nutzer. Bergbauunternehmen nutzen die Daten während des Betriebs und danach, um ihre Maßnahmen zu planen und durchzuführen. Sanierer brauchen die Daten, um die Maßnahmen zur Renaturierung und Wiedernutzbarmachung zu planen oder Auskunft über den potenziellen Nutzen der Regionen zu geben. Öffentliche Verwaltung sind ebenfalls auskunftverpflichtet und nutzen zudem die Daten in Planungsprozessen. Private und gewerbliche Nutzer benötigen Auskunft über bestimmte Gebiete, um diese zu nutzen: landwirtschaftliche Nutzung, Nutzung als Baugrund, etc. Im Folgenden beschreiben wir vier beispielhafte Nutzungsszenarien für die Daten aus der Verwaltung von Bergbaufolgelandschaften.

- Bergbauflächen muss auch nach dem aktiven Bergbau gesichert werden. Um sie wieder nutzbar zu machen, müssen diese Flächen zudem umfangreich saniert werden. Diese Sanierung erfordert detaillierte Planungsprozesse, welche die Sanierung begleiten. Diese Maßnahmen werden oft von vielen Unternehmen gleichzeitig durchgeführt, so dass ein hoher Grad an Abstimmung und Monitoring notwendig ist. Hier hilft ein digitales System, welches die Informationen und Daten zu den betroffenen Flächen zur Verfügung stellt. Außerdem können so die durchgeführten Maßnahmen dokumentiert werden, so dass diese Informationen später weiterhin zur Verfügung stehen. Dies ist beispielsweise hilfreich, wenn später Schäden in den sanierten Flächen auftreten, so können Experten dann besser die Lage und den Gefährdungsgrad beurteilen.
- Eine zuvor bergbauliche und inzwischen sanierte Fläche wird landwirtschaftlich genutzt. Als bergbauliche Folge ist die Fläche nicht bei jedem Wetter gleich belastbar. Bei starken Regen sollte sie deshalb nicht mit schweren Maschinen über eine bestimmte Gewichtsgrenze befahren werden. Der Landwirt muss vor dem Befahren der Fläche mit einer schweren Maschine den Verwalter der Fläche anfragen, ob ein Befahren möglich ist. Der Verwalter

muss daher schnell und effizient Zugriff auf die entsprechenden Gewichtsgrenzen und Wetterlagen haben, um diese Anfrage zu beantworten und gegebenenfalls zu warnen. Weiterhin kann sich auch während des Befahrens der Status der Betretbarkeit dieser Flächen ändern. In diesem Falle muss der Verwalter die gefährdeten Personen schnell warnen können.

- Ein neuer Radweg soll auf vormals bergbaulich genutzten Gebiet gebaut werden. Bei der Wegeplanung ist es notwendig zu wissen, ob sich die vorgesehenen Flächen dafür eignen und was für die Baumaßnahmen zu beachten ist. Dazu gehört nicht nur, ob die für den Radweg vorgesehene Fläche entsprechend belastbar ist, sondern beispielsweise auch ob die Fläche versiegelt werden darf, welche Ausgleichsmaßnahmen für die Versiegelung erbracht werden müssen und wo dies möglich wäre, welche anderen Umweltauflagen wie Brutzeiten während der Bauzeit beachtet werden müssen oder ob die Baufahrzeuge sicheren Zugang zu der Baumaßnahme haben.

- Das sanierende Bergbauunternehmen muss eine Fläche sichern, welche aufgrund der bergbaulichen Folgen Gefahr bei Betreten aufweist. Zur Sicherung gehört unter anderem entsprechenden Schilder zur Kennzeichnung der Fläche aufzustellen, welche ein Verbot des Betretens ausweisen. Die Aufstellung muss dokumentiert werden. In regelmäßigen Abständen muss die korrekte Platzierung sowie die Lesbarkeit kontrolliert und dokumentiert werden. Im Falle von unbefugtem Betreten und Unfällen währenddessen dient diese Dokumentation als Nachweis der korrekten Sicherung und muss somit auf Dauer des Verbotes gespeichert und einfach abrufbar sein.

3 Nachnutzungskonzept

Aus den Vielfalt an Nutzern und Nutzungsszenarios sowie der Vielfalt der Daten und ihrer Charakteristika ergeben sich die Anforderungen an ein Nachnutzungskonzept. Allen voran ist deutlich zu erkennen, dass ein auf das Problem angepasstes GIS, welches auch Werkzeuge für die Datenaufbereitung beinhaltet, ein hilfreiche Komponente zur Erfüllung dieser Aufgabe sein kann. Im Folgenden ist konzipiert, welche Anforderungen dieses angepasste GIS erfüllen muss.

3.1 Ausrichtung der Datenhaltung auf Ewigkeitslasten

Das Nachnutzungskonzept muss auf Ewigkeitslasten ausgelegt werden. So gestaltet sich zum Beispiel die Renaturierung bzw. Wiedernutzbarmachung von Böden schwierig und kann teilweise nicht kurzfristig gelöst werden. Auch die Folgen des Abpumpen und Verschmutzen des Grundwassers durch Tagebaumaßnahmen werden uns noch sehr lange begleiten. Somit ist die Verwaltung und Sanierung eine Langzeitaufgabe [GMK16], welche nicht nur eine Generation an Sachbearbeitern beschäftigt wird. Daher ist es von entscheidender

Bedeutung, dass das Nutzungskonzept nicht auf der guten Aktenkenntnis einiger weniger Sachbearbeiter beruht. Vielmehr muss durch Digitalisierung und Aufbereitung der Daten in Kombination mit entsprechenden Tools für den Zugriff auf die Daten es möglich sein ohne gute Kenntnis der Aktenlagen in kurzer Zeit an einen umfassenden Informationsstand zu einer bestimmten Region zu kommen. Dies ist beispielsweise von großer Bedeutung bei Havarien oder Katastrophenfällen. Aber auch kleinere Vorfälle wie bei Bränden in den betroffenen Flächen benötigt man schnelle und verlässliche Informationen, unter welchen Bedingungen Flächen betretbar sind.

3.2 Umfangreiche und zielführende Digitalisierung und Aufbereitung der Daten

Es reicht nicht neue Daten digital zu erheben und abzuspeichern. Auch die historischen Daten sind von Bedeutung und müssen umfassend digitalisiert werden. Dabei geht eine Digitalisierung über den reinen Scan analoger Datenbestände hinaus. Je nach Datentyp, z.B. Dokumente mit Fließtext, strukturierte Informationen aus Tabellen und Formulare und Karten, müssen entsprechend digitalisiert werden, um Zugriff auf den Text, die strukturierte Information oder die Geoinformationen aus den Karten zu erhalten. Nur so ist eine weitere Nutzung dieser Daten sinnvoll möglich. Erst durch die Verknüpfung der Daten wird eine effiziente Nutzung durch Kombination der darin enthaltenen Informationen möglich. Aufgrund der Größe und Heterogenität der Datenmenge muss diese Verknüpfung durch KI Verfahren unterstützt werden.

Ob digital-born oder nachträglich digitalisierter Daten, alle Daten müssen für eine effiziente Verwendung aufbereitet werden. Auch sollte geprüft werden, ob bei den Daten das FAIR Prinzip [Wi16] angewendet werden kann. Dazu gehört in jedem Fall eine Aufbereitung und Standardisierung der Metadaten, sodass diese zum Suchen und Filtern verwendet werden können. Gerade bei nachträglich digitalisierten Daten sind die Metadaten entweder fehlend, nicht vollständig oder von unzureichender Qualität vorhanden, da diese erst nach dem Entstehen der eigentlichen Daten erzeugt werden. Aber auch bei digital-born Daten können die Metadaten stark in der Vollständigkeit und Qualität schwanken. Daher gehört zur Aufbereitung die Aufwertung, Korrektur und ggf. Extraktion der Metadaten auf Basis der eigentlichen Daten.

Neben den Metadaten müssen auch die Daten selbst aufbereitet werden. Im einfachsten Fall sind sie bereits in einem geeigneten Format und fehlerfrei. Zumeist müssen die Daten jedoch konvertiert werden und auf Fehler überprüft werden. Dabei sollten behebbare Fehler korrigiert werden und auf nicht behebbare Fehler hingewiesen werden. Informationen in Fließtext müssen so aufbereitet werden, dass zum Einen eine Volltextsuche möglich ist und zum Anderen gezielt nach bestimmten Informationen gefiltert werden kann. Hierbei können vorab Informationen extrahiert werden und mit Verlinkung auf das enthaltende Dokument und der entsprechenden Stelle separat abgespeichert werden. Strukturierte Informationen aus Formularen oder Tabellen müssen digital in geeigneter Form und der entsprechenden Semantik abgespeichert werden. Der reine Text aus der Tabelle ist dabei

nicht ausreichend. Bildgebenden Daten, wie zum Beispiel LiDAR Daten oder Karten, müssen georeferenziert und so umgewandelt werden, dass sie gemeinsam mit den aktuellen Geoinformationen im Vergleich angezeigt und exploriert werden können. Sensordaten müssen den entsprechenden Gebieten durch Verschneidung zugeordnet werden um eine gemeinsame Analyse zu gewährleisten. Je nach Anwendungsfall müssen hierbei historische Daten aus Datenbanken und Livedaten kombinierbar sein. Gegebenenfalls ist es notwendig, vor der Anzeige noch zu analysieren oder zur Modellbildung und für Vorhersagen zu verwenden.

Einen großen Augenmerk bei der Datenaufbereitung muss auf der Durchsuchbarkeit und Filterbarkeit der Daten liegen. Hier spielen die aufbereiteten Metadaten eine besonders große Rolle. Durch sie ist es möglich die Menge an Daten durch entsprechende Filter nach beispielsweise Region und Datum so zu reduzieren, dass die nur noch die relevanten Daten angezeigt werden. Dies funktioniert nur mit einem guten Wechselspiel zwischen den gesammelten Metadaten und verfügbaren Filteroptionen in der Oberfläche des GIS. Zum anderen müssen die Daten auf verschiedene Arten durchsuchbar und facetierbar sein, um für natürlich sprachliche Anfragen die Datenmenge möglichst stark reduzieren zu können, so dass schnell und effizient die relevanten Informationen gefunden werden ohne Informationen zu übersehen.

3.3 Schnittstellen aller Art

Weder ist es zielführend noch realistisch alle Nutzer durch ein Werkzeug in ihren Bedürfnissen zu befriedigen. Genauso wenig zielführend, realistisch und umsetzbar ist die Idee alle Daten in einem Werkzeug zu hinterlegen. Gerade bei rechtlichen Fragen bezüglich Datenschutz und Geheimhaltung würde eine allumfassende Lösung nicht anwendbar sein. Daher ist es umso wichtiger Schnittstellen in und aus dem GIS anzubieten um Interoperabilität zu gewährleisten.

Für einige Nutzer und Nutzungsszenarien ist es wichtig öffentliche Daten wie Wetterberichte oder Pegelstände in ihren Anfragen und Analysen zu integrieren. Daher ist es wichtig, Schnittstellen zu öffentliche Daten zu implementieren und diese so innerhalb des GIS mit nutzen zu können. Dazu müssen die öffentlichen Daten mit den eigenen Daten verschnitten werden und entsprechend angezeigt werden. Aber auch die eigenen, öffentlichen Daten sollten für Nutzer über Schnittstellen so angeboten werden, dass diese die Daten in ihren eigenen GIS oder Werkzeugen nutzen können.

Nicht alle Daten müssen jedoch öffentlich sein. Auch für nicht-öffentliche, sensitive Daten mit zum Beispiel personenbezogene oder geheime Daten ist es sinnvoll diese mit anderen Nutzern unkompliziert teilen zu können. Daher sind Schnittstellen mit Authentifizierung eine sinnvolle Erweiterung des GIS. So können bei Bedarf Daten auch mit anderen Werkzeugen weiterverwendet werden.

Eine nützlich Erweiterung stellt die Möglichkeit dar, eigene, nicht öffentliche Daten mit den Daten des GIS zu verschneiden. Dazu muss die Möglichkeit bestehen, Daten ohne dauerhafte Speicherung in des GIS zu laden und mit den Daten des GIS zu verschneiden. Das Ergebnis sollte dann vom Nutzer analysierbar sein. Nachfolgend sollten die hinzugefügten Daten umgehend gelöscht werden.

3.4 Flexible und robuste Datenhaltung

Nicht nur an die Funktionalität gibt es Anforderung. Da man von Ewigkeitslasten ausgehen muss, ist eine langlebige Datenhaltung entscheidend. Dies bedeutet, dass bei der Wahl der Programme und Dateiformate auf eine langfristige Nutzbarkeit geachtet werden muss. Dabei sollte es vermieden werden, proprietäre Programme und Datenformate zu nutzen, da deren Verfügbarkeit nur durch den Anbieter gesteuert wird. Damit ist man dann abhängig von der dauerhaften Verfügbarkeit dieser Anbieter, wobei gerade im IT Bereich proprietäre Programme und Datenformate schnell veralten, nicht mehr gewartet werden und nicht mehr nutzbar sind. Eine mögliche Lösung könnte die Nutzung von Open Source Technologien sein, da die weitere Pflege dieser Programme und Datenformate auch nach Aufgabe durch einen externen Maintainer weiter betrieben werden kann. Zudem werden Open Source Softwaresysteme und Dateiformate mittlerweile dauerhaft archiviert, so dass eine spätere Nutzung auch in mehreren Jahrzehnten noch möglich wäre. Programme und Dateiformate mit einer großen Nutzerbasis haben durch das Interesse der Nutzer eine hohe Chance auf langfristige Nutzbarkeit. Oftmals werden neue Entwicklung im IT Bereich in sie integriert. Daher eignen sie sich für den Einsatz in Konzepten wie dieses, die langfristig nutzbar sein sollen. Um eine Interoperabilität mit anderen Programmen zu ermöglichen ist es wichtig verschiedene Standardformate zu berücksichtigen. Wenn möglich, sollten diese Standardformate verwendet werden. Dort wo es nicht möglich ist, sollte ein Datenexport in die Standardformate ermöglicht werden.

Anforderungen und Standards können sich ändern. Zudem können neue Typen von Daten hinzukommen. Dies macht es erforderlich, dass Programme und Dateiformate erweiterbar sind, sodass der Aufwand der Integration gering bleibt. Dennoch ist es nicht auszuschließen, dass das gewählte GIS in Zukunft nicht mehr den Anforderungen entspricht oder genutzte Software nicht mehr gewartet wird. Dadurch kann es notwendig werden, die Softwareumgebung zu wechseln. Es muss daher gewährleistet werden, dass alle Daten, Werkzeuge und ggf. das gesamte GIS in eine neue Softwareumgebung portiert und integriert werden kann. Auch diese Anforderungen können Open Source Technologien erfüllen.

4 Zusammenfassung

In diesem Beitrag haben wir ein Nachnutzungskonzept für Daten in der Altlastenverwaltung vorgestellt, welches auf die verschiedenen Nutzergruppen und deren Nutzungsszenarien

eingeht. Hierbei spielen verschiedene Aspekte eine große Rolle. So muss auf Langlebigkeit im Bereich der Software geachtet werden, da die Altlastenverwaltung eine Ewigkeitsaufgabe ist. Eben daraus ergibt sich aber auch eine starke Anforderung an eine flexible und gleichzeitig robuste Datenhaltung, um auf etwaige Änderungen in den Anforderungen gut reagieren zu können. Dabei soll der Änderungsaufwand nur möglichst gering sein und nur Teile des Systems, wenn möglich nicht aber das gesamte System, betreffen. Verschieden Stakeholder für verschiedene Daten lassen eine einfache Integration von neuen Daten und gute Schnittstelle in und aus dem System heraus zu einer wichtigen Anforderung werden. Nicht nur bereits digitale Daten, sondern auch analoge Daten brauchen in gute Integration. Es ist wichtig anzumerken, dass eine Digitalisierung analoger Daten hierbei deutlich über den reinen Scan der analogen Daten hinausgeht. Für alle Daten ist eine gute Datenaufbereitung entscheidend, um als Basis für die verschiedenen Nutzungsszenarien der Altlastenverwaltung zu dienen. Ein System entsprechend dem hier vorgestellte Nachnutzungskonzept hat großes Potenzial jetzige aber auch zukünftige Nutzungsszenarios und Anforderungen an die Altlastenverwaltung stemmen zu können.

Literatur

- [GMK16] Goerke-Mallet, P.; Melchers, C.; Kretschmann, J.: Global post-mining challenges—Experiences gained from the German hard coal mining industry. In: 16th international congress for mine surveying report. S. 287–292, 2016.
- [Sc21] Schröder, C.; Bürgl, K.; Annanias, Y.; Niekler, A.; Müller, L.; Wiegreffe, D.; Bender, C.; Mengers, C.; Scheuermann, G.; Heyer, G.: Supporting Land Reuse of Former Open Pit Mining Sites using Text Classification and Active Learning. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, Online, S. 4141–4152, Aug. 2021, URL: <https://aclanthology.org/2021.acl-long.320>.
- [Wi16] Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E. et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3/1, S. 1–9, 2016.
- [Wi23] Wilke, S.: Flächenverbrauch für Rohstoffabbau, <https://www.umweltbundesamt.de/daten/flaeche-boden-land-oekosysteme/flaeche/flaechenverbrauch-fuer-rohstoffabbau>, letzter Zugriff: 12.05.2023, Jan. 2023.

Classification of Large and Heterogeneous LiDAR Data Sets

Kenn Kaubukowski¹, Baldwin Nsonga^{1,2}, Vanessa Kretzschmar^{1,2}, Yves Annanias²,
Daniel Wiegrefe² and Mario Hlawitschka¹

Abstract:

In the context of the renaturation of discontinued open-cast mines, the interactive visualization analysis of three-dimensional LiDAR provides a comprehensive overview for planning the subsequent use of these areas. When analyzing the measured point clouds, it is beneficial to classify the points, enabling the user to filter objects such as vehicles and buildings. Most current classification techniques for LiDAR data rely on good ground truth and work on specific types of measurements and resolutions. In this work, we present a semantic classification workflow for LiDAR data with highly heterogeneous acquisition and storage parameters. We apply the workflow to a freely available LiDAR data set and showcase the resulting classification. Our method is shown to reliably classify not only large-scale objects, such as buildings, but also small-scale objects, such as power lines and street lights.

Keywords: Geographic Information System, LiDAR, Renaturation

1 Introduction

When raw materials are extracted from open-cast mines, large and unusable crater landscapes remain. To restore these areas for subsequent use, for example, as residential areas, nature reserves, or renewable energy parks, renaturation becomes necessary. For this, a large amount of information must be considered, which makes the renaturation process complex in terms of planning and implementation. This results in the planning of fundamentals and requirements for the implementation of the renaturation. As part of a previous project, an interactive decision tool was developed that extracts restrictions from text documents [Sc21] and links and displays them in a geographic information system (GIS) with the corresponding geographic map data [An22b]. It enables the comparison of data across multiple linked views. Evaluation and interlinking of weather events (such as heavy rain) as a further type of data in combination with validation of the applicability of legal restrictions in relation to such events is also possible [An22a]. Our project partner is the LMBV GmbH³, which is dedicated to renaturalizing former open-cast lignite mines and is responsible for collecting and digitizing historical data on open-cast lignite mines in Saxony.

¹ Leipzig University of Applied Sciences, Chair of Computer Graphics, Karl-Liebknecht-Straße 132, 04277 Leipzig, Germany {kenn.kaubukowski|baldwin.nsonga|vanessa.kretzschmar|mario.hlawitschka}@htwk-leipzig.de

² Leipzig University, Image and Signal Processing Group, Augustusplatz 10, 04109 Leipzig, Germany {nsonga|kretzschmar|annanias|wiegrefe}@informatik.uni-leipzig.de

³ <https://www.lmbv.de>

In addition to the geographic information system, a three-dimensional visualization based on LiDAR data allows for a comprehensive overview of the corresponding areas. A use case in the context of renaturation is the detection of landslides from measurements. LiDAR data, especially in urban areas, include buildings, vehicles, or, in the case of rural areas, vegetation. To assess ground movement reliably, it is necessary to exclude these objects from the analysis. As this can be an elaborate process, automated classification of LiDAR data can streamline the process.

In this work, we propose a classification approach for LiDAR data. Our hybrid classification method balances the strengths and weaknesses of known classification techniques and differentiates between twelve classes and fully classifies data sets with variable results on data sets with highly different point densities. The quality of the results depends on whether certain requirements are met: To achieve optimal results, a data set must have metadata values of *return number*, *number of returns*, and *intensity*. The classified LiDAR data is accordingly colored so that buildings, plants, and other bodies can be distinguished from each other by color coding. The resulting color-coded 3D point cloud is visualized using Potree [SOW20] as an additional map in the tool.

Here, we evaluate our classification method on a publicly available LiDAR data set⁴ provided by the district government of Cologne, North Rhine-Westphalia, Germany. The data set was generated in 2017 using airborne laser scanning and has an average resolution of four to ten points per square meter, where the resolution may be higher in forest areas. In addition to being heterogeneous in a way that includes rural and urban areas, the data set also provides metadata, which we require to perform a reliable classification. Note that no pre-labeled or annotated data is available. As this can be assumed to be the case for most application-driven LiDAR data sets, our proposed method is not based on deep learning but on a handcrafted classification scheme for this kind of data.

2 Related Work

Segmentation and classification of scanned point cloud data have been an active field of research for decades, as they have been shown to be useful in vastly different domains, such as the extraction of geomorphological objects [Ma17], or in the context of architectural heritage [GR20; Ma20]. Motivated by the question of whether the field of robotics and autonomous driving is hungry for 3D LiDAR data, Gao et al. [Ga22] present an extensive review of data sets and methods for point cloud segmentation. Grilli et al. [GMR17] presented another overview of the segmentation and classification algorithms. The surveys indicate that in recent years, machine learning approaches, especially deep learning methods, have been the focal point of point cloud segmentation and classification research.

Qi et al. [Qi17a; Qi17b] presented PointNet and PointNet++, which have become popular methods that utilize deep learning for unstructured 3D point clouds. Various machine learning

⁴ https://www.opengeodata.nrw.de/produkte/geobasis/hm/3dm_1_1as/3dm_1_1as/

approaches have been proposed that are concerned with special convolution kernels for point clouds, e.g. [HTY18; Li18; Li20; Th19; Xu18] or in conjunction with color information [Ya22]. Alternative neural networks, such as recurrent neural networks [HWN18] or self-organizing networks [LCL18], were also explored for point cloud segmentation. Biasutti et al. [Bi18] present an alternative approach to LiDAR point clouds for segmentation and classification. They argue that sensors have a specific topology, which allows the authors to derive 2D depth histograms and leverage segmentation techniques from image processing, which have been extensively studied.

Although machine learning approaches have shown to be suitable for point cloud segmentation and classification, labeled high-quality training data, which is hard to come by, are required to achieve reliable results. These methods also tend to achieve very good results when the input data is similar to the training data in terms of point density and environment. Since we are dealing with heterogeneous data from multiple decades with varying quality and training on all types of data is not feasible, we were looking for a different approach that could be easily adapted to all data sets in our database.

More traditional approaches, which do not rely on neural networks, are usually hand-crafted for specific features. Duillard et al. [Do11] group points into voxels before clustering these voxels in terms of geometric mean and variance values. They employ different strategies for ground segmentation and other features taking sparse data into account. Techniques using k-nearest neighbors have also been shown to produce reliable results for point cloud classification [BJW16; We15]. Using random forests in conjunction with the k-nearest neighbor approach, Jiang et al. [Ji22] outperformed PointNet++ [Qi17b] for the Toronto -3D data set [Ta20], demonstrating the viability of more traditional techniques.

Based on the fact that current techniques could not handle the heterogeneous data well, we present a handcrafted classification approach focused on aerial LiDAR point data of urban and suburban landscapes.

3 Classification

In order to classify the given LiDAR data some pre-processing is required. If a digital terrain model is not provided, it first has to be generated using a region-growing-based approach. Note that the distance of the points to this reference ground is essential for the classification and has to be provided. The general concept of our method is a cascading classification. Specific classes are evaluated one after another in a certain order allowing the use of dependencies. The core techniques used in hybrid classification are filtering procedures, region growing, model fitting, and object segmentation in combination with object analysis. We differentiate between the classes: buildings, vehicles, vegetation, power lines, streets, and ground. In addition, we consider walls, fences, and other linear, slim structures as one class. In the following, for the sake of brevity, we will discuss the classification of buildings,

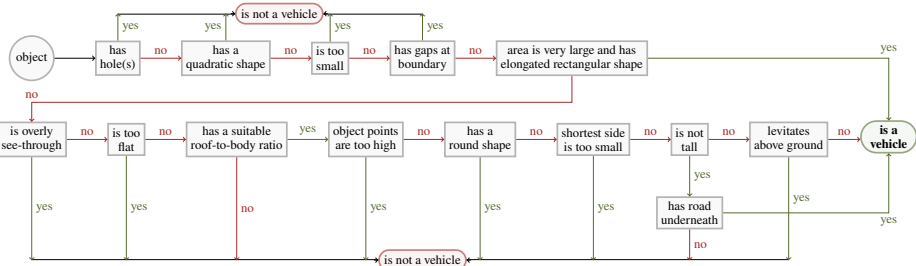


Fig. 1: Decision tree for checking if an object can be referred to as a *vehicle*. Some decision nodes have a specific purpose with a meaning; for example, *is overly see-through* filters out hedges. The sequential nature of this decision tree enables early termination to save up computation time.

vehicles, and vegetation. The workflow for the other classes is similar and is explained in the thesis that this work is based on [Ka23].

- Buildings:** The algorithm is based on finding seed points for building roofs, which are found by eliminating bad candidates. First, the seed points must be the last return. In the next step, additional requirements, namely a minimum distance of 1.8 m to the ground, as well as the absence of any two-dimensional neighboring points that are located noticeably below a potential building point within a small radius, and the recognition of isolated point collections that are too small in number, allow a filtering procedure to remove additional points. For each remaining potential building point, neighboring points within a given radius are used for the RANSAC algorithm [FB81], calculating an optimal plane. Assuming flat surfaces with a certain height to be roofs, points that have a certain number of neighbors with a distance to the optimal plane above a certain threshold can be filtered. The threshold is dynamically chosen so that a high point density reduces the threshold, while a low point density, indicating higher uncertainty, produces a high threshold. The resulting set of points is used as seed points for a region growing approach. Our subsequent region growing approach is based on the point neighborhood (n-nearest neighbors with maximum distance) and the condition if the corresponding points meet certain requirements, implying that they are building points: It should not belong to any already identified class, it should be at least 1 meter above the ground, the return number should be equivalent to the number of returns (or one), and the height difference to the corresponding seed must be below a dynamically chosen threshold (similar to the previous step). To prevent neighboring vegetation from being included in the building class, strict requirements must be chosen. The first step is to verify that the return number and the number of returns of each point have matching values. Additionally, to verify whether a neighbor of a building point is not vegetation, the n-nearest neighbors of the building point, as well as the potential vegetation point, are found. If the two points in question are within a certain radius from each other but not within their n-nearest neighbors, the points are likely part of different objects, and therefore the building class should not

be applied to the candidate. To include chimneys and facades in our building class, we apply region growing and static assignment, respectively. The assignment for the facades is based on the assumption that neighbors below the roof border have little variation in the x and y coordinates. Neighbor points above the roof are evaluated by size avoiding falsely classifying large structures like vegetation as parts of a building.

- **Vehicles:** The classification of vehicle points makes use of object segmentation and object analysis. After a collection of points has been identified and fully classified as vehicle candidates using filtering and region growing, they are segmented using region growing. The filtering procedure for points has the following conditions: It is not classified, the distance to the ground is between 25 cm and 4.25 m, and the neighboring points are within a height range threshold. The remaining points are used for the region growing with the following condition for a valid neighbor: it is not classified, it is at least 5 cm above the ground, and it is not significantly higher than the seed point (depends on point density). Note that we discard regions with too many or too few points. The resulting point sets are then individually analyzed on various characteristics such as the width and height of the oriented bounding box or the rate of points that are on a similar position on the z-axis as the highest point of the collection and are subsequently labeled as vehicles or discarded from the selection of potential vehicles. Figure 1 shows a schematic representation of the decision tree used for our vehicle classification.
- **Vegetation:** Preceding the vegetation, the classification of power lines, street lights, and water is performed. Points that are neither a first nor a last LiDAR return value occur primarily within vegetation. Therefore, filtering procedures are used to find valid seed points based on the following criteria: For low-resolution cases, the return number has to be smaller than the number of returns, and their distance from the ground should be above 50 cm. Using region growing, the vegetation class is then spread over the entire vegetation. Details on the applied conditions are presented by Kaubukowski [Ka23]. The vegetation points are then verified by analyzing their position in relation to their neighboring points. Afterward, points that are not in the vicinity of many other vegetation points are assumed to be noise and, therefore, removed from the classification. Due to strict requirements in terms of return values, the current selection of vegetation points results in isolated unclassified points surrounded by vegetation points. To fill these gaps, the neighboring points of all points previously removed from the classification procedure are analyzed. If more than 25% of their three-dimensional neighbors belong to the vegetation class, they are also classified as vegetation. The ASPRS [Sp19] suggests a division of vegetation into three subclasses according to its height, which we also include. First, all vegetation points are assigned to a new class based solely on their distance to the ground. Here, a suitable division must be chosen based on the regional area. Then region growing is used again to allow the higher vegetation classes to spread over entire structures with only downward growth being allowed.

- **Finalization:** Ultimately, it is not guaranteed that all points are assigned to a class after the initial classification process. To ensure that every point is classified, a final reassessment is performed. The neighboring points of all unclassified points are examined to rank the surrounding classes of said points individually. Depending on a point's distance to the ground and the dominant surrounding class, they are then either classified retroactively or region growing is performed to further spread the dominant class. Different classes are prioritized differently to balance the general classification's weaknesses. Afterwards, all remaining points are classified as noise.

4 Results and Discussion

The resulting data set is fully classified, the only requirement being the availability of the aforementioned metadata. The quality of the results also depends on the given point density.

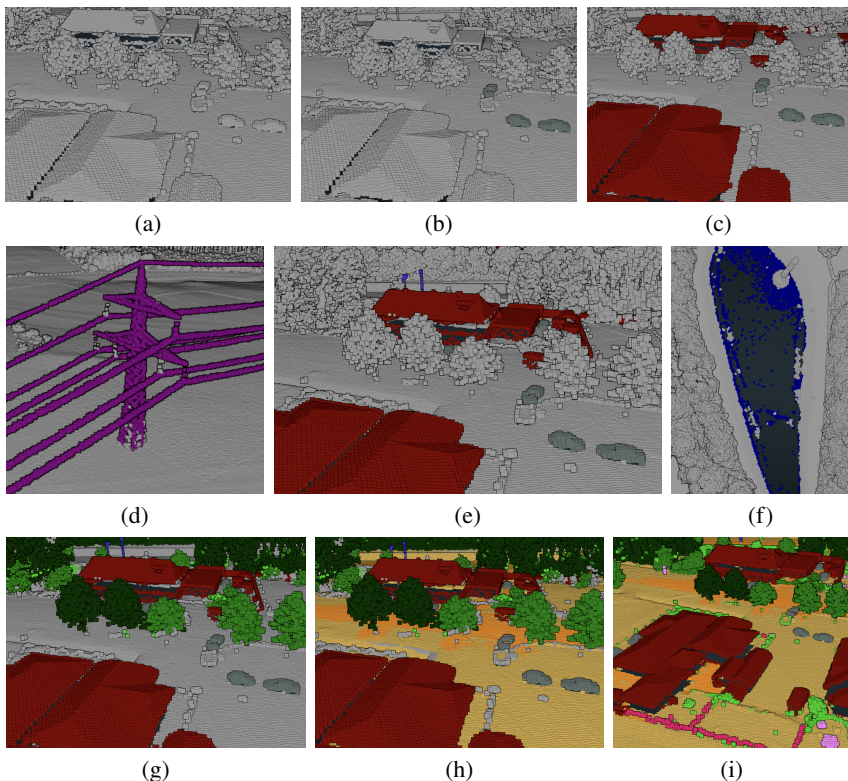
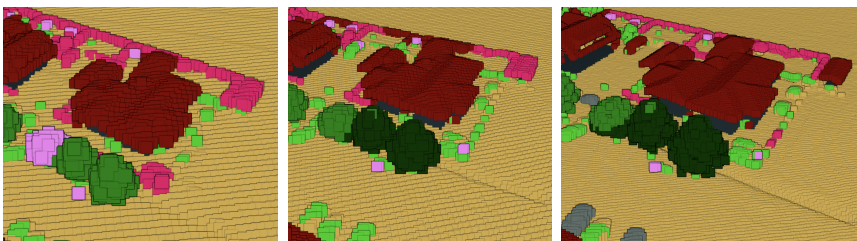


Fig. 2: Result of the presented classification method. Starting with the unclassified LiDAR data (a) various object types are classified: vehicles (b), buildings (c), power lines (d), street lights (e), water (f), vegetation (g), as well as the ground and roads (h). The final classification is shown in (i).

Some classes may not be classifiable in a low-resolution data set, and a potentially unclear partitioning of objects may also lead to inaccuracies. The results generally improve with a higher point density. Since there is no ground truth available, we perform a visual evaluation of our resulting classification (cf. Figure 2). The results indicate that our method performs well in the task of classifying heterogeneous LiDAR data. Furthermore, it shows that we were also able to reliably identify smaller-scale objects such as cars and street lights. To test the stability of our algorithm, we used voxel downsampling and therefore increased the voxel size to 1.5 m. As shown in Figure 3, the quality decreases, while the vast majority of objects remain correctly classified.

A key shortcoming is the heavy reliance on metadata such as the number of returns, intensity, and return number. Although it is possible to apply our algorithms to low-resolution LiDAR data, there is still a lower limit in terms of data quality. In our experience, classification based on position alone without further metadata (e.g. color) is not feasible, especially for roads, which only become visible at very high resolutions. Another key shortcoming is the current runtime of our method, although it could be argued that classification is only necessary when new data is acquired. However, it is highly dependent on the size of the data provided. For example, a square kilometer with approximately 15.6 million points requires up to 60 minutes of computation time to find all of the object classes described above on an average workstation (Apple MacBook Pro, M1 Pro chip, 10 CPU cores, 16 GB RAM). The most computation time is required for the calculation of the terrain model and the classification of street lights, which take 38.6% and 27.6% of the total time, respectively.

In the future, we want to test our algorithms on data sets with less smooth terrain, such as cliffs or fjords. In addition to the optimization of the runtime, we are also working on improving the accuracy of our method. With respect to the described current classification tasks, dynamic instead static thresholds could be beneficial to achieve better stability for varying resolutions. We also want to incorporate external data to provide metadata for the classification workflow. Finally, we are currently using our classification to look at temporal changes in the data over the years. Such information is of great importance to archaeology on the one hand, and to local authorities on the other, in order to prevent disasters such as landslides or subsidence, which are very likely to occur in former mining areas.



(a) 1.5m (89.9% reduction) (b) 1.0m (77.7% reduction) (c) 0.5m (46.1% reduction)

Fig. 3: Voxel sizes and their reduction equivalent compared to the original size in brackets.

5 Conclusion

We proposed a semantic classification procedure for digital surface models of landscapes and how to approach certain object types in the context of the renaturation of former mining areas. This procedure is capable of meeting most LAS specifications [Sp19] in contrast to known methods. We have achieved this classification using low-resolution LiDAR data and without the extensive use of neural networks. Although there is a justified trend towards the use of neural networks, we argue that such algorithmic parts as those presented are still necessary for the semantic classification of different objects, especially for data where no comparable training data are available. However, there are more features to come for an optimal workflow, such as a LiDAR data-based approach to show temporal changes or to reduce the limitations as we discussed before.

Acknowledgment

This research was supported by the Development Bank of Saxony (SAB) under Grant 100400221.

Bibliography

- [An22a] Annanias, Y.; Windolph, J.; Wehlitz, R.; Wiegrefe, D.: An Interactive Decision Support System for Analyzing and Linkage of Weather-Related Restrictions of Opencast Lignite Mines. In: LEVIA'22: Leipzig Symposium on Visualization in Applications. 2022.
- [An22b] Annanias, Y.; Zeckzer, D.; Scheuermann, G.; Wiegrefe, D.: An Interactive Decision Support System for Land Reuse Tasks. *IEEE Computer Graphics and Applications* 42/6, pp. 1–13, 2022.
- [Bi18] Biasutti, P.; Aujol, J. F.; Brédif, M.; Bugeau, A.: Range-image: Incorporating sensor topology for lidar point cloud processing. *Photogrammetric Engineering and Remote Sensing* 84/6, pp. 367–375, 2018.
- [BJW16] Blomley, R.; Jutzi, B.; Weinmann, M.: Classification of airborne laser scanning data using geometric multi-scale features and different neighbourhood types. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3/3, pp. 169–176, 2016.
- [Do11] Douillard, B.; Underwood, J.; Kuntz, N.; Vlaskine, V.; Quadros, A.; Morton, P.; Frenkel, A.: On the segmentation of 3D LIDAR point clouds. In: *IEEE International Conference on Robotics and Automation*. IEEE, pp. 2798–2805, 2011.
- [FB81] Fischler, M. A.; Bolles, R. C.: Random sample consensus. *Communications of the ACM* 24/6, pp. 381–395, 1981.

- [Ga22] Gao, B.; Pan, Y.; Li, C.; Geng, S.; Zhao, H.: Are We Hungry for 3D LiDAR Data for Semantic Segmentation? A Survey of Datasets and Methods. *IEEE Transactions on Intelligent Transportation Systems* 23/7, pp. 6063–6081, 2022.
- [GMR17] Grilli, E.; Menna, F.; Remondino, F.: A review of point clouds segmentation and classification algorithms. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 42/2W3, pp. 339–344, 2017.
- [GR20] Grilli, E.; Remondino, F.: Machine Learning Generalisation across Different 3D Architectural Heritage. *ISPRS International Journal of Geo-Information* 9/6, p. 379, 2020.
- [HTY18] Hua, B.-S.; Tran, M.-K.; Yeung, S.-K.: Pointwise Convolutional Neural Networks. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 984–993, 2018.
- [HWN18] Huang, Q.; Wang, W.; Neumann, U.: Recurrent Slice Networks for 3D Segmentation of Point Clouds. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 2626–2635, 2018.
- [Ji22] Jiang, S.; Guo, W.; Fan, Y.; Fu, H.: Fast Semantic Segmentation of 3D Lidar Point Cloud Based on Random Forest Method. In. Pp. 415–424, 2022.
- [Ka23] Kaubukowski, K. J.: Semantische Klassifizierung digitaler Oberflächenmodelle von Landschaften in Form von massiven Punktwolken über Hybrid-Verfahren, Bachelor Thesis, Hochschule für Technik, Wirtschaft und Kultur Leipzig, Leipzig, 2023.
- [LCL18] Li, J.; Chen, B. M.; Lee, G. H.: SO-Net: Self-Organizing Network for Point Cloud Analysis. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 9397–9406, 2018.
- [Li18] Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B.: PointCNN: Convolution On X-Transformed Points. In (Bengio, S.; Wallach, H.; Larochelle, H.; Grauman, K.; Cesa-Bianchi, N.; Garnett, R., eds.): *Advances in Neural Information Processing Systems*. Vol. 31, Curran Associates, Inc., 2018.
- [Li20] Lin, Y.; Yan, Z.; Huang, H.; Du, D.; Liu, L.; Cui, S.; Han, X.: FPCConv: Learning Local Flattening for Point Convolution. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 4292–4301, 2020.
- [Ma17] Mayr, A.; Rutzinger, M.; Bremer, M.; Oude Elberink, S.; Stumpf, F.; Geitner, C.: Object-based classification of terrestrial laser scanning point clouds for landslide monitoring. *The Photogrammetric Record* 32/160, pp. 377–397, 2017.
- [Ma20] Matrone, F.; Grilli, E.; Martini, M.; Paolanti, M.; Pierdicca, R.; Remondino, F.: Comparing Machine and Deep Learning Methods for Large 3D Heritage Semantic Segmentation. *ISPRS International Journal of Geo-Information* 9/9, p. 535, 2020.

- [Qi17a] Qi, C. R.; Su, H.; Mo, K.; Guibas, L. J.: PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 77–85, 2017.
- [Qi17b] Qi, C. R.; Yi, L.; Su, H.; Guibas, L. J.: PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates Inc., pp. 5105–5114, 2017.
- [Sc21] Schröder, C.; Bürgl, K.; Annanias, Y.; Niekler, A.; Müller, L.; Wiegrefe, D.; Bender, C.; Mengs, C.; Scheuermann, G.; Heyer, G.: Supporting Land Reuse of Former Open Pit Mining Sites using Text Classification and Active Learning. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, pp. 4141–4152, 2021.
- [SOW20] Schütz, M.; Ohrhallinger, S.; Wimmer, M.: Fast Out-of-Core Octree Generation for Massive Point Clouds. In: Computer Graphics Forum. Vol. 39. 7, Wiley Online Library, pp. 155–167, 2020.
- [Sp19] Specification, L.: Specification 1.4—R14, The American Society for Photogrammetry & Remote Sensing, 2019.
- [Ta20] Tan, W.; Qin, N.; Ma, L.; Li, Y.; Du, J.; Cai, G.; Yang, K.; Li, J.: Toronto-3D: A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, pp. 797–806, 2020.
- [Th19] Thomas, H.; Qi, C. R.; Deschaud, J.-E.; Marcotegui, B.; Goulette, F.; Guibas, L.: KPConv: Flexible and Deformable Convolution for Point Clouds. In: IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, pp. 6410–6419, 2019.
- [We15] Weinmann, M.; Schmidt, A.; Mallet, C.; Hinz, S.; Rottensteiner, F.; Jutzi, B.: Contextual classification of point cloud data by exploiting individual 3D neighborhoods. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2/W4, pp. 271–278, 2015.
- [Xu18] Xu, Y.; Fan, T.; Xu, M.; Zeng, L.; Qiao, Y.: SpiderCNN: Deep Learning on Point Sets with Parameterized Convolutional Filters. In (Ferrari, V.; Hebert, M.; Sminchisescu, C.; Weiss, Y., eds.): Computer Vision – ECCV 2018. Springer International Publishing, Cham, pp. 90–105, 2018.
- [Ya22] Yang, J.; Li, H.; Jiang, Z.; Zhang, D.; Yue, X.; Du, S.: Color guided convolutional network for point cloud semantic segmentation. International Journal of Advanced Robotic Systems 19/3, 2022.

Geodaten in der öffentlichen Verwaltung

Status Quo in den sächsischen Kreis- und Landesverwaltungen

Christian Bender¹, Dr. Mario Hesse²

Abstract: Die Digitalisierung der Verwaltung in Deutschland schreitet seit Jahren in der öffentlichen Verwaltung voran. Darauf aufbauend erwachsen neue Möglichkeiten, die Potentiale der Nutzung von raumreferenzierten Fachdaten in der Verwaltungspraxis zu nutzen. Die Ergebnisse einer empirischen Studie auf Ebene der Kreise und kreisfreien Städte sowie der Landesbehörden in Sachsen zeigen auf, unter welchen Rahmenbedingungen die Nutzung von Geodaten erfolgt, wie dessen technische Ausgestaltung beschrieben werden kann und welche künftigen Potenziale und Hindernisse einer Nutzung von Geodaten im öffentlichen Verwaltungshandeln gegenüberstehen.

Keywords: Öffentliche Verwaltung; Geodaten; Georeferenzierung; Digitalisierung; Sachsen

1 Einleitung

Von Seiten der Forschung wird davon ausgegangen, dass bis zu 80 % aller Informationen im öffentlichen Verwaltungshandeln einen direkten oder indirekten Raumbezug aufweisen.³ Bezogen auf die öffentliche Verwaltung erlaubt die Nutzbarmachung georeferenzierter Fachdaten nicht nur eine Verbesserung der bestehenden Fachanwendungen, sondern ermöglicht die Vernetzung bestehender Systeme innerhalb und über die Grenzen der eigenen Verwaltungseinheit hinaus. So können beispielsweise Bauvorhaben durch Informationen aus anderen Fachressorts, wie dem Grünflächenamt, dem Einwohnermeldeamt oder auch der Wirtschaftsförderung den eigentlichen Verwaltungsprozess mit relevanten Zusatzinformationen anreichern und folglich die Planung und Steuerung in der öffentlichen Verwaltung effizienter gestalten.

Digitale Geodaten bilden schon heute eine wichtige Grundlage für viele Fachanwendungen in der Verwaltungspraxis. Dabei wird eine Unterscheidung hinsichtlich der Termini *Geobasisdaten* und *Geofachdaten* getroffen. Erstere sind einheitliche geodätische Daten des amtlichen Vermessungswesens. Hierzu zählen

¹ Universität Leipzig, Institut für öffentliche Finanzen und Public Management, Professur für Finanzwissenschaften, Augustusplatz 10, 04109 Leipzig, bender@wifa.uni-leipzig.de

² Universität Leipzig, Institut für Finanzen und Public Management, Professur für Finanzwissenschaften, Augustusplatz 10, 04109 Leipzig, hesse@wifa.uni-leipzig.de, <https://orcid.org/0000-0002-8298-4175>

³ Die These selbst kann nicht zweifelsfrei belegt werden. So gehen Franklin und Hane von zirka 80 % aus, während beispielsweise Hahmann und Burghardt im Rahmen einer empirischen Studie aufzeigen, dass dieser Wert eher bei 60 % liegt. Vgl. Franklin/Hane (1992), S. 12–15; Hahmann/Burghardt (2013), S. 1171–1189.

beispielsweise Gewässer, Verkehrswege, administrative Grenzen, Grundstücke, Adressen und Gebäude. Letztere sind Daten, die einen Raumbezug aufweisen, demnach eine Zuordnung zu einer räumlichen Lage besitzen (z. B. Nutzungsformen von Flächen, Verkehrsflüsse auf Straßen oder Energieverbräuche bestimmter Gebäude). Die Nutzung digitaler Geodaten in der Verwaltung lässt sich in Deutschland bereits auf die 1970er Jahre zurückdatieren, als erstmals analoge Karten in der Verwaltung systematisch digitalisiert worden sind. Mit der Umsetzung der 2007/2/EG des Europäischen Parlamentes und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE) erfolgte ein weiterer Schub zur Vereinheitlichung der Geodateninfrastruktur (GDI) in Deutschland.⁴ Weitere Neuerungen, wie etwa das Geodatenzugangsgesetz und die damit korrespondierenden Landesgesetzgebungen folgten und bilden den heutigen (Rechts-)Rahmen für die Nutzung digitaler Geodaten.⁵

Wenngleich der Rechtsrahmen für den Einbezug von Geodaten in das Verwaltungshandeln abgesteckt ist, so offenbaren sich in der verwaltungsseitigen Umsetzung mitunter ungenutzte Potenziale. Grund hierfür sind verschiedene Hindernisse, die einer effektiven Einbindung von Geodaten in das öffentliche Verwaltungshandeln entgegenstehen und diese mitunter sogar gänzlich unterbinden. Der vorliegende Beitrag nimmt sich dieser Hemmnisse an und zeigt auf, welche Potenziale sich durch den Einbezug von Geofachdaten in das öffentliche Verwaltungshandeln ergeben. Daraus lassen sich Politikempfehlungen für die Landesgesetzgeber ableiten.

2 Studiendesign

Der Einbezug von Geofachdaten in das Verwaltungshandeln unterlag in den vergangenen Jahrzehnten einer stetigen Entwicklung. So ist deren Nutzen zwar frühzeitig erkannt worden, doch die entsprechende Einbindung in die konkreten Verwaltungsprozesse verlief eher schleppend. Neben den rechtlichen Grundlagen spielt im föderalen Geflecht zwischen Bund, Ländern und Kommunen nicht zuletzt die Frage der Finanzierung eine bedeutende Rolle. Zwar hat der Bund seine finanziellen Bemühungen für die Verwaltungsdigitalisierung im Allgemeinen, beginnend mit der Föderalismusreform II, und für die Einbindung von Geodaten im Besonderen, verstärkt. Jedoch zeigt sich, dass die flächendeckende Digitalisierung des Verwaltungshandelns sowie die Nutzung georeferenzierter Fachinformationen nach wie vor auf eine Verwaltungsstruktur trifft, die sich nur sukzessive den neuen Gegebenheiten und technischen Möglichkeiten anpasst.⁶

Um zu analysieren, inwiefern Geofachdaten in das Verwaltungshandeln bisher integriert worden sind, ist eine Umfrage auf der Ebene der Landkreise und in den Einheiten der

⁴ Vgl. Richtlinie 2007/2/EG.

⁵ Vgl. Verwaltungsvereinbarung GDI-DE und GDI Sachsen.

⁶ Zur Frage der Finanzierung siehe beispielsweise Bender/Mengs/Hesse (2023).

sächsischen Landesverwaltung durchgeführt worden, deren Ergebnisse nachfolgend präsentiert werden.⁷

2.1 Darstellung der gewählten Methodik

Das zugrundeliegende Forschungsdesign kombiniert unterschiedliche Analyseansätze miteinander und ist in Abb. 1 schematisch dargestellt. Zunächst wurde mithilfe einer Literaturrecherche der Verbreitungs- und Nutzungsstand digitaler Geofachdaten im öffentlichen Verwaltungshandeln eruiert. Hierzu sind empirische Studien, wie etwa von Zern-Breuer (2020) oder Bill (2022), analysiert worden.⁸ Aufgrund dieser Vorüberlegungen sind qualitative Expert:innen-Interviews durchgeführt worden, um zu eruiern, inwiefern Geofachdaten im Verwaltungshandeln auf der kommunalen und der Landesebene im Freistaat Sachsen eine Rolle spielen. Die hieraus gezogenen Erkenntnisse sind in die Konzeptionierung des quantitativen Fragebogens eingeflossen. Der Fragebogen fokussierte insbesondere auf die Einbindung von Geofachdaten in das Verwaltungshandeln in verschiedenen fachlichen Kontexten. Dabei sind die Wirksamkeit von Geofachdaten in Bezug zur OZG-Umsetzung, die Relevanz und Frequenz in der täglichen Verwaltungsarbeit, bestehende Rahmenbedingungen und Nutzungsprofile sowie Hindernisse und Effizienzpotenziale näher betrachtet worden.

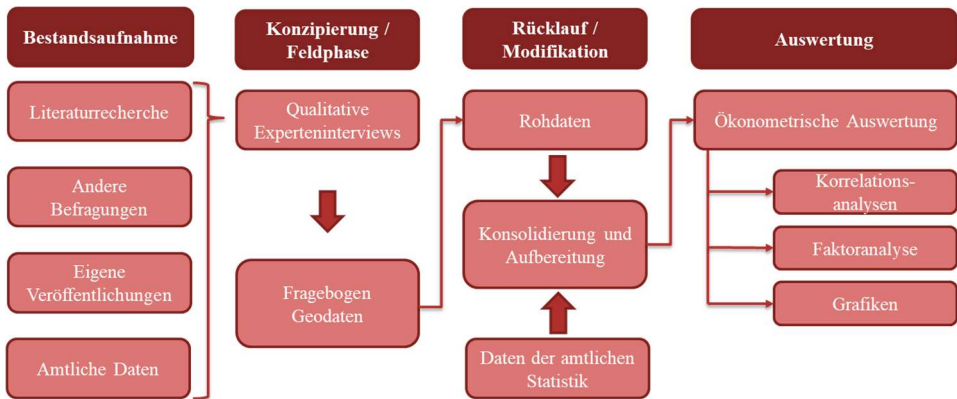


Abb. 1: Mixed-Methods-Forschungsdesign

Die Befragung wurde sowohl online als auch postalisch vorgenommen. Der erzielte Rohdatensatz ist anschließend konsolidiert worden, indem jene Fragebögen ausgeschlossen wurden, die unvollständig ausgefüllt waren. Anschließend sind die Rohdaten um Daten der amtlichen Statistik angereichert worden. Der dadurch entstandene

⁷ Die kreisangehörigen Gemeinden sind hiervon ausgenommen, da diese institutionell keinen direkten Konnex zur Nutzung von Geofachdaten aufweisen. Die Erhebung und Speicherung obliegt der Kreisebene und den landesseitigen Einheiten.

⁸ Siehe hierzu Zern-Breuer et al. (2020) und Bill (2022).

konsolidierte Datensatz ist mithilfe von Korrelationsanalysen sowie einer Faktoranalyse ausgewertet worden.

2.2 Eigenschaften der Stichprobe

Einzelne Abteilungen der Gebietskörperschaften wurden separat angeschrieben. Für die Kreisverwaltungen waren dies die Aufgabengebiete ‚Zentrale Verwaltung‘, ‚Schule und Kultur‘, ‚Soziales und Jugend‘, ‚Gesundheitsdienste‘, ‚Gestaltung der Umwelt‘ und ‚Raumentwicklung‘. Die kontaktierten Abteilungen der Landesbehörden können den Aufgabenbereichen ‚Allgemeinen Dienste‘, ‚Ernährung, Landwirtschaft und Forsten‘, ‚Bildungswesen, Wissenschaft, Forschung, kulturelle Angelegenheiten‘, ‚Energie- und Wasserwirtschaft, Gewerbe und Dienstleistungen‘, ‚Gesundheit, Umwelt, Sport und Erholung‘ sowie ‚Verkehrs- und Nachrichtenwesen‘ zugeordnet werden. Von den 214 kontaktierten Verwaltungseinheiten ist ein (Brutto-)Rücklauf von 102 Fragebögen generiert worden (47,7 %). Leere Fragebögen und eingegangene Fragebögen, die vor dem kritischen Abbruchpunkt⁹ abgebrochen worden sind, konnten nicht berücksichtigt werden, weshalb sich der Stichprobenumfang auf 88 Befragungen (41,1 % Nettorücklauf) reduziert.

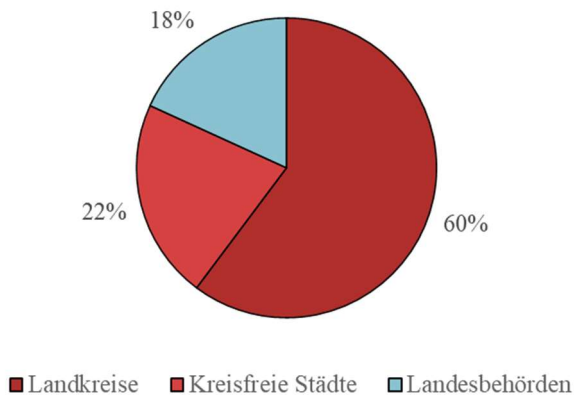


Abb. 2: Struktur des Stichprobenumfangs

⁹ Der Kritische Abbruchpunkt ist in dieser Befragung auf den letzten inhaltlichen Themenschwerpunkt gesetzt worden. Sind folglich nicht alle inhaltlichen Fragestellungen beantwortet worden, ist der Fragebogen nicht in die Betrachtung aufgenommen worden. Nur das Freitextfeld am Ende der Befragung kann somit als optional angesehen werden.

Wie Abb. 2 zeigt, sind rund 60 % des verwertbaren Rücklaufs aus den Landkreisverwaltungen erzielt worden. Weitere 22 % kamen aus den kreisfreien Städten. Rund 18 % der gesamten Stichprobe sind den Landesbehörden zuzuordnen.¹⁰

3 Empirische Erkenntnisse

Im Zuge der Auswertung konnte zuvorderst aufgezeigt werden, dass die Interoperabilität, d. h., die behördenübergreifende Nutzung, in verschiedenen Ausprägungen für die Kreis- und Landesebene von herausgehobener Relevanz ist. Damit verbundene Problemstellungen sind jedoch nicht gleichmäßig verteilt, sondern je nach der Häufigkeit der Nutzung von Geofachdaten im täglichen Verwaltungshandeln durchaus differenziert zu betrachten. So sind beispielsweise die Themenbereiche der Dateiformate und der Metadaten auch ohne die Nutzung eines Raumbezuges in der täglichen (Daten-)Arbeit relevant und fallen arbeitsebenenübergreifend an. Darüber hinaus existieren spezifische Raumproblematiken, die insbesondere die Nutzung von Koordinatensystemen und die notwendige Georeferenzierung umfassen. Erst bei zunehmender Durchdringung und der Einführung eines digitalen Raumbezuges treten diese Problematiken häufiger auf. Damit lässt sich herausstellen, dass die Problemstellungen, die sich durch die Nutzung von georeferenzierten Fachdaten ergeben, durchaus spezifischer Natur sind.

3.1 Rahmenbedingungen in der sächsischen Landes- und Kreisverwaltung

In der Einleitung wurde bereits ein Hinweis darauf gegeben, dass viele der Daten, die in einer Verwaltung täglich genutzt werden, einen Raumbezug haben. Daher wundert es zunächst nicht, dass, wie in Abb. 3 dargestellt, die Häufigkeit der Nutzung von Geofachdaten in der täglichen Arbeit entsprechend hoch ist.

Nur 6 % (Landesbehörden) beziehungsweise 16 % (Kreisverwaltungen) geben an, dass Geofachdaten eine untergeordnete Rolle spielen. Digitale georeferenzierte Fachdaten sind im Verwaltungshandeln präsent und werden von den Befragten zudem als relevant angesehen. Über 75 % der Landesbehörden, beziehungsweise 63 % der Verwaltungseinheiten auf Kreisebene gaben an, dass Geofachdaten für die tägliche Arbeit *sehr relevant* seien. Für weitere 25 % bzw. 27 % sind diese *eher relevant*. Geofachdaten werden somit nicht nur als wichtig erachtet, sondern auch praktisch genutzt. Dabei zeigen die Korrelationsanalysen zudem auf, dass ein positiver Zusammenhang zwischen beiden Fragestellungen besteht: Die regelmäßige Nutzung von Geofachdaten führt dazu, dass die

¹⁰ Angesichts der Strukturen wurde hier kein Bezug zur Einwohnerverteilung der Stichprobe hergestellt. Formal bezieht sich die Tätigkeit der Landesbehörden oftmals auf alle Einwohner im Freistaat Sachsen und auch bei den Landkreisen und kreisfreien Städten ist keine Varianz mit Blick auf die Einwohnerzahlen zu erwarten. Daher konnte keine Einwohnergewichtung der Ergebnisse durchgeführt werden. Zur Thematik der Gewichtung siehe Krone/Scheller (2020) oder Mengers/Bender/Kratzmann et al. (2022).

Relevanz der Geofachdaten höher eingeschätzt wird ($0,74^{***}$)¹¹. Georeferenzierte Fachdaten werden demnach zunehmend als ein unverzichtbarer Baustein im öffentlichen Verwaltungshandeln angesehen.

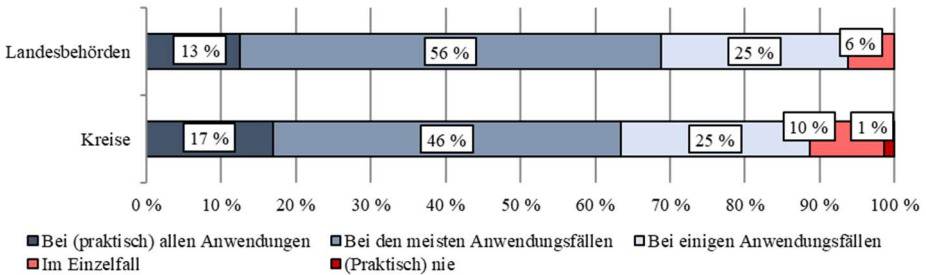


Abb. 3: Häufigkeit der Nutzung digitaler Geodaten

Dagegen wird anhand von Abb. 4 deutlich, wo Probleme liegen: Zwar haben die Landesbehörden, wie auch die Kreisverwaltungen damit begonnen, Fachdaten zu georeferenzieren, doch der Georeferenzierungsprozess ist bei weitem noch nicht abgeschlossen. Erschwerend kommt hinzu, dass, sofern nach dem Tätigkeitsbezug gefragt wird, zwar 58 % aller Befragten angeben, Geofachdaten zu erheben, bereitzustellen und zu speichern, jedoch nur 29 % diesen Datentypus auch für Managementüberlegungen¹² heranziehen.

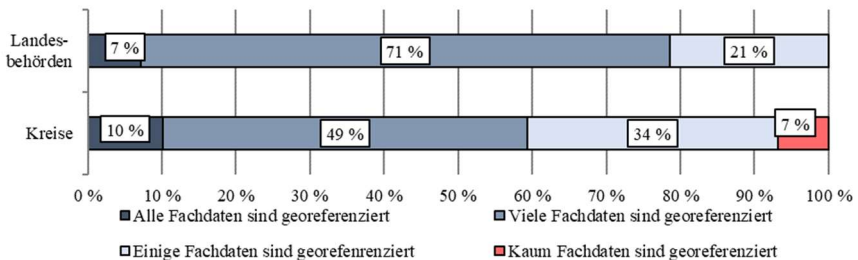


Abb. 4: Anteil bereits georeferenzierter Fachdaten

Einerseits wird somit die Relevanz von Geofachdaten durchaus gesehen und deren Mehrwert in der Implementierung in Fachverfahren erkannt, andererseits schreitet die Georeferenzierung der Fachdaten selbst nur schleppend voran. Strategische Planungs- und darauf fußende Entscheidungsprozesse werden bisher zumeist nicht unter Nutzung von Geofachdaten getroffen. Dies ist insofern interessant, als dass sich ein starker positiver Zusammenhang zwischen dem Umfang der täglich genutzten georeferenzierten Fachdaten

¹¹ Die Sterne geben die Signifikanz zum 1 % (***) , 5 % (**) und 10 % (*) Niveau an. Der Koeffizient gibt die Effektstärke an.

¹² Hierunter werden strategische Planungs- und Entscheidungsprozesse verstanden, die sich in unterschiedlichen Fristigkeiten materialisieren.

und der Nutzungsfrequenz zeigen lässt (0,80***). Sind die Fachdaten erst einmal georeferenziert, werden sie auch sehr regelmäßig in der täglichen Arbeit verwendet. Damit ist die Georeferenzierung kein Selbstzweck, sondern sie unterstützt die vorhandenen Verwaltungsprozesse und bietet in dieser Hinsicht einen Mehrwert. In Bezug zum bisherigen Stand der Georeferenzierung stellt sich daher die Frage, weshalb Potenziale bisher nicht vollumfänglich genutzt werden.

3.2 Potenziale und Hindernisse für künftige Entwicklungen

Mit Blick auf die bisher unvollständige Georeferenzierung in den Verwaltungen lässt sich eine Optimierungspotenzial mit Blick auf die dafür notwendigen Datenstandards aufzeigen (Abb. 5).

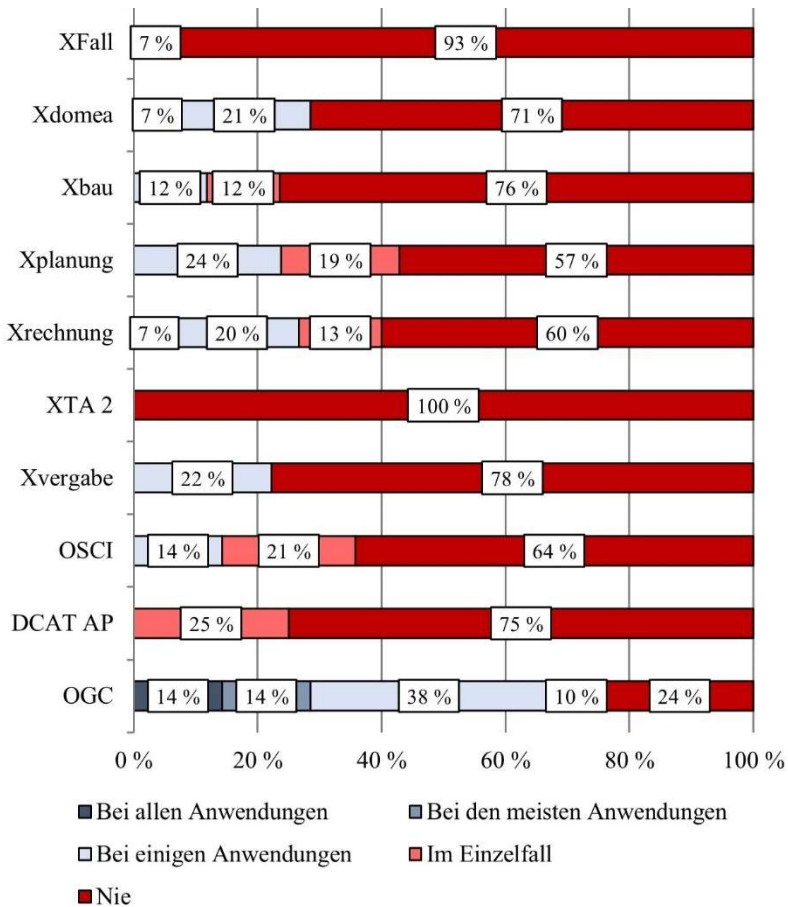


Abb. 5: Standardisierungsmethoden

Weder die teils rechtlich verpflichtenden X-Standards, noch die OGC-Standards werden flächendeckend umgesetzt. Vor dem Hintergrund einer voranschreitenden Digitalisierung der öffentlichen Verwaltung und der zunehmenden Bedeutung raumbezogener Fachinformationen ist dies durchaus problematisch zu sehen. Insbesondere durch die fehlenden OGC-Standards werden Interoperabilitätsprobleme spätestens auftreten, sobald Fachdaten anderer Verwaltungseinheiten in das eigene GIS-System übertragen werden sollen. Auch für den Datenaustausch zwischen den Verwaltungseinheiten werden in Zukunft durch die fehlende Umsetzung der durch den IT-Planungsrat bereitgestellten Standardisierungsmethoden Probleme zu erwarten sein. Die Causa Raumbezug ist somit bei weitem noch nicht abgeschlossen.

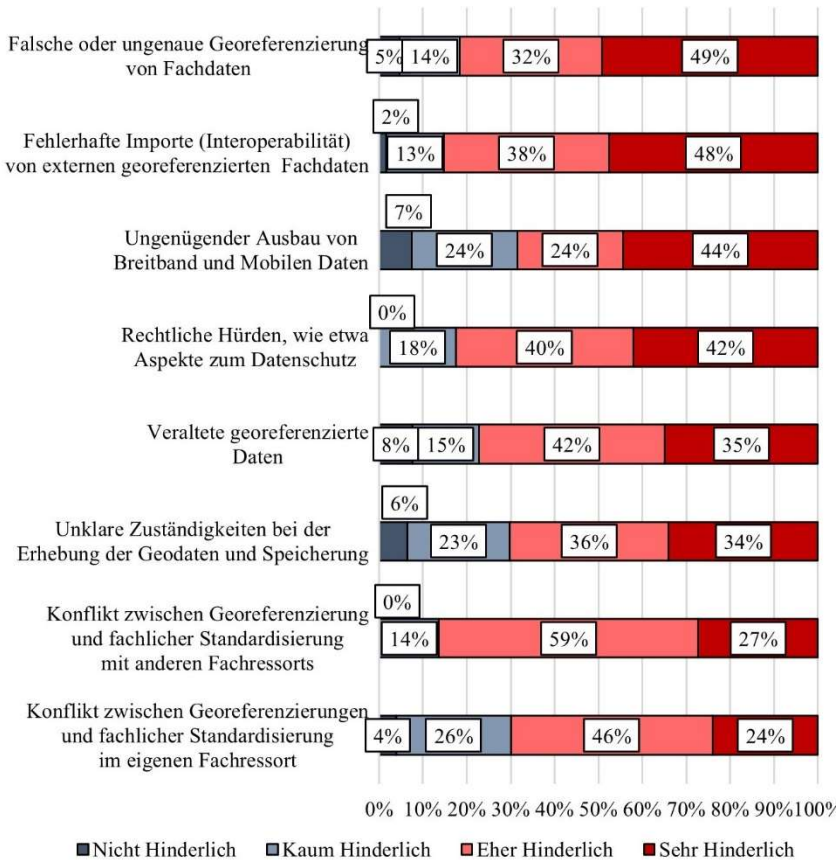


Abb. 6: Hindernisse in der Nutzung georeferenzierter Fachdaten

Dass insbesondere das Problem der falschen oder ungenauen Georeferenzierung und den damit verbundenen Interoperabilitätsproblemen derart präsent sind, hängt mit der fehlenden Umsetzung der beschriebenen Standards zusammen. Ein erster Ansatzpunkt ist

folglich in der konsequenten Umsetzung der bereitgestellten Standards zu sehen. Rechtliche Hürden, Probleme in der Infrastruktur und veraltete Georeferenzierungen sind weitere Problemstellungen. Noch am wenigsten hinderlich sind Fragen der Organisation in der Verwaltungseinheit selbst, obwohl auch hier noch hohe Zustimmungswerte erreicht werden. Werden Geofachdaten von den Mitarbeitern in der Verwaltung im Rahmen der täglichen Arbeit selten verwendet, wird eine unklare Zuständigkeit bei der Erhebung und Speicherung der Geofachdaten häufiger als Hindernis genannt (0,31**). Eine Befähigung und das Nehmen von anfänglichen Berührungängsten im Umgang mit digitalen Geofachdaten können im Umkehrschluss zu einer signifikanten Verringerung der damit im Zusammenhang stehenden Hindernisse führen.

Hinzu kommt, dass insbesondere dann die Interoperabilität vermehrt als problematisch angesehen wird, sobald die *eigenen Fachdaten bereits den Anforderungen genügen* (-0,27*). Aus einer prozessorientierten Sicht treten somit Interoperabilitätsprobleme erst dann auf, wenn die eigenen Fachdaten bereits rege genutzt werden und die Notwendigkeit der Einbindung weiterer Geofachdaten anderer Verwaltungseinheiten gesehen wird. Auch hier zeigt sich, dass der Mehrwert von Standardisierungsmethoden oft erst dann erkannt wird, wenn eine Einbindung von Daten Dritter notwendig wird. Eine sofortige externe *on-demand Nutzung* ist somit zumeist nicht möglich.

Daraus leitet sich die Frage ab, welche Mehrwerte eine Einbindung von Geofachdaten in das Verwaltungshandeln sinnhaft erscheinen lassen und eine intrinsische Motivation zu deren Nutzung befördern. Immerhin scheinen die Hindernisse, die sich durch eine erzwungene Nutzung ergeben, durchaus hoch zu sein. Abb. 7 zeigt auf, welche Effizienzgewinne durch eine vermehrte Einbindung von Geofachdaten in das tägliche Verwaltungshandeln erwartet werden. Hier stechen insbesondere die Planungshilfen, die Nutzung von Fachdaten mehrerer Fachämter und die Qualitätssicherung hervor. Dicht dahinter folgen die verbesserte Interoperabilität, der Aufbruch von Datensilos und die Optimierung interner Verwaltungsprozesse. Anders formuliert ergeben sich Effizienzvorteile in der internen, wie auch externen Verwaltungsperspektive, die eine Nutzensteigerung des vorhandenen Datenbestandes versprechen. Geofachdaten werden als Querschnittsressource verstanden, die die Qualität des Verwaltungshandelns auf breiter Front positiv beeinflussen können.

Diese Effizienzgewinne werden bei einer häufigeren Nutzung digitaler Geodaten auch vermehrt wahrgenommen, insbesondere die Nutzungsintensität bestehender Fachanwendungen (0,49***), die vollständige Digitalisierung aller vorhandener Fachdaten (0,33**) oder auch Planungshilfen (0,26*), die Interoperabilität unterschiedlicher Fachdaten (0,27*) und die Verfügbarkeit extern georeferenzierter Fachdaten (0,26*). Für die Relevanz der täglichen Nutzung sind Fragestellungen der Optimierung bestehender Verwaltungsprozesse somit vordergründig.

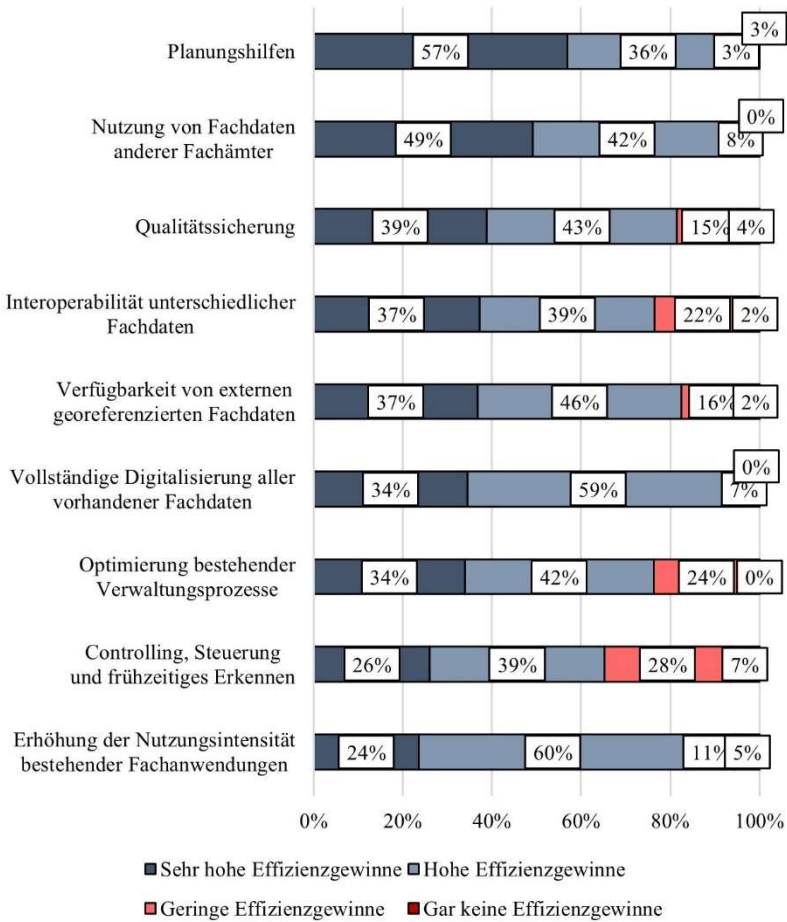


Abb. 7: Potenzielle Effizienzgewinne

Nicht nur hinsichtlich der Relevanz, sondern auch hinsichtlich der Häufigkeit können Rückschlüsse mithilfe der Korrelationsanalyse gezogen werden. Steigt die Häufigkeit der Nutzung von Geodaten in der täglichen Arbeit an, so werden auch deren potenziellen Effizienzgewinne für die öffentliche Leistungserbringung höher eingeschätzt. Signifikante Ergebnisse zeigen sich beispielsweise hinsichtlich der Nutzungsintensität bestehender Fachanwendungen (0,4***), bei Planungshilfen (0,23*), der Erhöhung der Partizipation unterschiedlicher Fachbereiche (0,26*), der Interoperabilität unterschiedlicher Fachdaten (0,25*) und der Verfügbarkeit von externen georeferenzierten Fachdaten (0,24*). Folglich steigen die wahrgenommenen Effizienzgewinne dann an, wenn die Verwaltungseinheit einen ersten Kontakt mit Geofachdaten hatte und eine erste Einschätzung darüber abgeben kann, welchen Mehrwert deren Nutzung hat. Die Akzeptanz gegenüber der vermehrten

Nutzung von Geofachdaten kann folglich gesteigert werden, sofern zunehmendes Wissen darüber besteht. Zusätzlich zeigt sich, dass mit vermehrter Nutzung von Geodaten auch dessen Implementierung in Managementüberlegungen voranschreitet. Hierfür spricht die vermehrte Nutzung von Planungshilfen bei der zunehmenden Verwendung von Geodaten. Nach einem ersten Ausprobieren werden Geofachdaten mehr und mehr zu einem unverzichtbaren Baustein.

4 Politikempfehlungen

Nachdem Rahmenbedingungen, Hindernisse und potenzielle Effizienzgewinne ausgeleuchtet worden sind, stellt sich die Frage, welche Empfehlungen für die Praxis daraus abzuleiten sind. Diese Aufgabe ist jedoch nicht trivial: Landkreise, kreisfreie Städte und Landesbehörden sind in ihrer Ausstattung und Nutzung von Geofachdaten sehr unterschiedlich aufgestellt. Die vorgestellten Politikempfehlungen müssen daher den Spagat zwischen einer begrenzten Allgemeingültigkeit und ihrer Adäquanz für die jeweilige Ebene vollziehen.

In der Betrachtung der Rahmenbedingungen zeigt sich, dass ebenenübergreifend die Abstimmung der Datenaufbereitung als problematisch angesehen wird. Zwar sind Geofachdaten anderer Verwaltungseinheiten für das jeweils eigene Verwaltungshandeln von Relevanz, doch wird eine hierzu förderliche Datenaufbereitung nicht flächendeckend nach den Bedürfnissen der Befragten vorgenommen. Da Datenstandards bereits existieren und bisher nur partiell umgesetzt werden, ist eine weitere Etablierung der existierenden Datenstandards zielführend. Managementüberlegungen werden zudem bisher nur selten in die Nutzungsüberlegungen zu den Geofachdaten mit einbezogen. Vor dem Hintergrund künftiger Herausforderungen der öffentlichen Hand, wie etwa dem demographischen Wandel oder der angespannten Arbeitsmarktlage, können hier Potentiale gehoben werden. Die fehlende Einbindung in Managementprozesse ist ein Schritt, der verwaltungsseitig nicht am Einbezug von Geodaten scheitert, sondern an der prozessualen Integration in bestehende Verwaltungsabläufe. Diese sollten so überarbeitet werden, dass eine leichtere (und standardmäßige) Integration digitaler Geodaten in alltägliche Verwaltungsprozesse möglich ist.

In Bezug auf den Umgang mit Geodaten zeigt sich, dass der Anteil an bereits georeferenzierten Fachdaten und die Häufigkeit der täglich verwendeten Daten miteinander korrelieren. Ein umfangreicheres Angebot von Geofachdaten führt zu einer intensivierten Nutzung, wodurch ein sich selbst verstärkenden Effekt innerhalb der Verwaltungseinheiten entsteht. Hierzu ist es notwendig, etwaige Vorbehalte zu überwinden und den Mehrwert der Datenintegration innerhalb der Verwaltungseinheit hervorzuheben. Deren Limitationen dienen konsekutiv als Anreiz zur Erhebung oder zur Akquisition weiterer Daten. Es bedarf somit eines niedrigschwelligen Initialpunktes für die Geodatennutzung, der durch die Entscheidungsträger:innen in der Verwaltungseinheit initiiert werden muss. Dieser Befund wird durch die Tatsache verstärkt, dass die

Wahrnehmung der Relevanz von Geofachdaten mit einer intensiveren Nutzung in der täglichen Arbeit einhergeht. Geofachdaten werden somit nicht nur genutzt, weil sie vorhanden sind, sondern vielmehr, da während des Nutzungsprozesses deren Relevanz von Seiten der Nutzer:innen anerkannt wird.

Es erscheint folglich nicht nur für Nachzügler förderlich zu sein, niedrighschwellige Angebote schaffen, an denen die Nützlichkeit der Georeferenzierung deutlich wird. Mit der Demonstration von *Good Practices* könnten die positiven Selbstverstärkungseffekte einen ersten Impuls erhalten. Wichtig ist hierbei, dass die Geofachdaten nach einem möglichst einheitlichen Muster bereitgestellt werden, um neue Unsicherheiten und technische Hürden zu vermeiden. Letztlich ist somit auf die Berührungspunkte der Mitarbeitenden abzustellen, wobei Schulungs- und Weiterbildungsprojekte helfen können. Dies zeigt sich auch durch den Befund, dass das Vorhandensein von Standardisierungen, wie beispielsweise durch ein *einheitliches Metadaten-system*, mit einer höheren Reife der allgemeinen Digitalisierung der Verwaltung sowie der spezifischen verwaltungsinternen Verwendung von eigenen Geofachdaten einhergehen. Allgemeine Standardisierungsprojekte können somit der Verwaltungseinheit helfen, Geofachdaten effizienter zu nutzen. Da die hierzu erforderlichen Standards bereits vorliegen, ist eine Implementierung zwar mit einem kurzfristigen verwaltungsinternen Ressourcenverbrauch, jedoch mittel- und langfristig mit erheblichen Effizienzvorteilen verbunden.

Die bisher noch geringe Nutzung von Standardisierungsformaten zeigt zudem, dass die Vorteile, die mit deren Einführung verbunden wären, hervorgehoben werden müssen. Ein Erfahrungsaustausch kann auch hier dabei helfen, initiale Bedenken seitens des Verwaltungspersonals zu begrenzen. In der Standardisierungsdiskussion zeigt sich, dass eine unzureichende Interoperabilität als besonders problematisch wahrgenommen wird, sobald die eigenen Fachdaten den Anforderungen genügen. Dieses Problem wird somit erst erkannt, wenn der datenseitige ‚Wachstumsschmerz‘ bereits eingetreten ist. Um die Phase des Erkennens der eigenen Versäumnisse jedoch effizient zu beheben und Netzwerkeffekte heben zu können, muss das Konzept der Interoperabilität von Anfang an mitgedacht werden. Hier bedarf es demnach eines vorgeschalteten Prozesses, um zu evaluieren, welche verwaltungs-, daten-, und entscheidungsseitigen Vorkehrungen im Vorfeld einer voranschreitenden Geodatenintegration zu treffen sind, um eine effiziente Einbindung von Geodaten über die Grenzen des eigenen Ressorts hinweg zu gewährleisten.

5 Fazit

Der vorliegende Beitrag konnte aufzeigen, wie es um die Nutzung von Geofachdaten der Kreise, kreisfreien Städte und Landeseinrichtungen im Freistaat Sachsen bestellt ist. Die grundlegende Krux besteht letztlich darin, dass zwar Vorteile in der Nutzung von Geofachdaten gesehen werden, sich die Umsetzung jedoch als schwierig gestaltet. Hier

kommen unterschiedliche Hemmnisse zusammen, die eine weitgehende Implementierung von Geofachdaten in querschnittsbezogene Managementüberlegungen erschweren. Die Spannweite reicht von fehlenden Implementierungsstrategien bestehender Standards hin zu internen und externen Datenproblematiken, die eine weiterführende Nutzung georeferenzierter Fachdaten erschweren. Gleichzeitig werden jedoch die allgemeinen Vorteile, die Netzwerk- und Synergieeffekte, durchaus erkannt.

Um diesen gordischen Knoten zu lösen, bedarf es einer Ressource, die aufgrund der Rahmenbedingungen nur spärlich vorhanden ist: Zeit. Der demographische Wandel und die sich verändernden Arbeitsmarktbedingungen, wie auch der technologische Fortschritt machen es unabdingbar, dass Verwaltungen technische Kapazitäten, auch im Hinblick auf die Nutzung von Geofachdaten, optimal nutzen. Gleichzeitig zeigt sich jedoch, dass die Bemühungen zur Georeferenzierung nur schleppend voranschreiten, Vorbehalte hinsichtlich der rechtlichen Rahmensetzung bestehen und nach wie vor infrastrukturelle Nachteile, insbesondere hinsichtlich der Abdeckung mit mobilen Daten sowie der Breitbandversorgung dominieren. Dieser ungünstige Mix wird durch falsche oder ungenaue Referenzierungspraktiken, unzureichende Standardisierungen und personelle Bedenken noch verstärkt.

Um Verwaltungen fit für die Zukunft zu machen, bedarf es daher eines Paradigmenwechsels. Das wird allerdings nur dann gelingen, wenn die Verwaltungen davon überzeugt werden, dass die Einbindung von Geofachdaten in ihrem ureigenen Interesse liegt. Überzeugungsarbeit und Good-Practice-Beispiele können hier als Stein des Anstoßes fungieren. Netzwerk- und Synergieeffekte stellen sich daran anschließend ein. Partizipations- und Beteiligungsmöglichkeiten zu schaffen, sollte daher oberste Priorität der Entscheidungsebene sein.

Literaturverzeichnis

- [Be23] Bender, C.; Mengs, C.; Hesse, M.: Geodaten in der öffentlichen Verwaltung. Eine Analyse am Beispiel des Freistaates Sachsen, Arbeitspapiere des Instituts für öffentliche Finanzen und Public Management Nr. 4, Leipzig, 2023.
- [Bi22] Bill, R.: Offene Geodaten in Raum- und Umweltwissenschaften - Stand und Potenziale. In: (Meinel, G.; Krüger, T.; Behnisch, M.; Erhardt, D. Hrsg.): Flächennutzungsmonitoring XIV: Beiträge zu Flächenmanagement, Daten, Methoden und Analysen. Rhombos-Verlag, Berlin, S. 185-192, 2022.
- [FrH92] Franklin, C.; Hane, P.: An Introduction to Geographic Information Systems: Linking Maps to Databases [and] Maps for the Rest of Us: Affordable and Fun. Database 15, S. 13–21, 1992.
- [GD21] GDI Sachsen, Geodateninfrastruktur Sachsen, <http://www.gdi.sachsen.de/inhalt/devgdi/gdisn/gdisn.html>, Stand: 28.06.2023.

- [Ha13] Hahmann, S.; Burghardt, D.: How much information is geospatially referenced? Networks and cognition. *International Journal of Geographical Information Science* 27/6, S. 1171-1189, 2013.
- [Kr20] Krone, E.; Scheller, H.: KfW Kommunalpanel 2020. Kreditanstalt für Wiederaufbau (KfW), Frankfurt am Main, 2020.
- [Me22] Mens, C.; Bender, C.; Kratzmann, A.; Goldammer, M.; Hesse, M.: Der große Klick. Kommunale Verwaltungsdigitalisierung in Sachsen, KOMKIS Analyse Nr. 20, Leipzig, 2022.
- [Ri07] Das Europäische Parlament und der Rat der Europäischen Union: Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE), in: *Amtsblatt der Europäischen Union* (Hrsg.): L 108/1 vom 25. April 2007.
- [Ve17] Vereinbarung zwischen dem Bund und den Ländern zum gemeinsamen Ausbau und Betrieb der Geodateninfrastruktur Deutschland (Verwaltungsvereinbarung GDI-DE), 2017.
- [Zr20] Zern-Breuer, R.; Seckelmann, M.; Regös, N.; Lorei, H.; Kruse, K. A.; Brunzel, M.: Voruntersuchung zur Einführung eines einheitlichen Geodatenmanagements in Rheinland-Pfalz, Speyrer Arbeitshefte Nr. 245, Speyer, 2020.

Wirtschaft, Management Industrie -
Joint Workshop IntDig 2023 MOC
2023; Intelligente Digitalisierung,
(KI-basiertes) Management und
Optimierung komplexer Systeme

Intelligente Digitalisierung

Ausgewählte Aspekte

Ulrich John¹


Abstract: Im Paper wird der bereits vor einigen Jahren vom Autor belegt und verwendete Begriff *Intelligente Digitalisierung* charakterisiert und durch die Diskussion einiger ausgewählter Aspekte angereichert. Der Inhalt versteht sich als eine Grundlage für (spätere) Erweiterungen, Ergänzungen, weitere Systematisierungen und Fachdiskussionen.

Keywords: Intelligente Digitalisierung, Digitale Transformation, Intelligentes Unternehmen, Intelligente Behörde, KI-basiertes Management

1 Einleitung

Der Begriff „Intelligente Digitalisierung“, der vom Autor bereits seit einigen Jahren verwendet wird (z.B. [JHW19], [Jo22a]), bereits Titel einer Lehrveranstaltungsreihe war und aktuell das Thema des Workshops *IntiDig 2023* ist, kann natürlich durchaus mit verschiedenen Inhalten und Ausrichtungen definiert werden. Ähnlich wie bei den Begriffen „Digitalisierung“ und „Digitale Transformation“ kann man die Meinung entwickeln, dass die einzelnen Inhalte, die sich hinter dem Begriff verbergen, keine grundlegend neuen Dinge sind. Die Bedeutung und Rechtfertigung der Begriffseinführung liegt in der Kombination und Fokussierung der geeignet gewichteten Teilthemen/Aspekte.

Einhergehend mit der Entwicklung der Informationstechnik und Informatik sind wir seit Jahrzehnten mit der Digitalisierung von Wirtschaft und Gesellschaft konfrontiert und beschäftigt. Aufgrund des Entwicklungsstandes und der technologischen Reife/Einsatzbarkeit von IT-Technologien sind wir theoretisch in der Lage, umfassende und umwälzende Digitalisierungen durchzuführen, die unsere Welt in allen Bereichen revolutionieren kann. Auf verschiedenen Betrachtungsebenen und Diskursbereichen stellen sich abgesehen von der technischen Realisierbarkeit die Fragen, was wann tatsächlich digitalisiert werden soll, wie die Digitalisierung hinsichtlich der Ziel-IT-Systemkomponenten und hinsichtlich der Zielprozesse und Zieleigenschaften gestaltet werden soll und wie der eigentliche Prozess der Digitalisierung zu gestalten ist. Blicke in die Praxis zeigen, dass nicht wenige Digitalisierungsprojekte scheitern, deutlich hinter den

¹ IU Internationale Hochschule, Frankfurter Allee 73A, 10247 Berlin, ulrich.john@iu.org/ john@intelligente-digitalisierung.de,  <https://orcid.org/0009-0005-2588-1750>

technischen und prozessualen Möglichkeiten zurückbleiben, unerwünschte Konsequenzen haben, falsch aufgesetzt oder falsch durchgeführt werden.

Zunehmend und maßgeblich bestimmt in diesem Kontext der Begriff „Künstliche Intelligenz“ die Entwicklung – derzeit insbesondere im Zusammenhang mit generativen KI-Systemen. Dabei kann beobachtet werden, dass sich auch diesbezüglich viele „Nicht- oder Halbexperten“ mehr oder weniger einflussstark einbringen, äußern und agieren. Einige, zu ergänzende, Anhaltspunkte zur Identifikation von KI-Experten findet man zum Beispiel auch in [SB21]. Nicht selten werden scheinbar insbesondere Schwerpunktaspekte aufgegriffen, betont und etabliert, dann leider oft unter Vernachlässigung anderer wesentlicher Aspekte. Es besteht die Gefahr, dass Entscheidungsträger auf Basis unrichtiger, unvollständiger oder unausgereifter Informationen „erfolgreich“ beeinflusst werden, vieles falsch gemacht wird/ werden wird und Fehlentscheidungen weitreichende negative – möglicherweise irreversible - Konsequenzen haben werden. Diesbezüglich wären auch systemische Analysen der relevanten Prozesse und Organisationsstrukturen erforderlich, damit auf Basis der dabei gewonnen Erkenntnisse – in Kombination mit den technologischen Kenntnissen und Vorgehenskenntnissen - Restrukturierungen und Reorganisationen durchgeführt werden können.

Auf Basis disruptiver und gereifter IT-Technologien und besonders auch KI-Technologien konnten in den letzten Jahren erhebliche Um- und Neugestaltungen von Prozessen und Strukturen in vielen Bereichen, insbesondere in Unternehmen, durchgeführt werden (siehe z.B. [Ai21], [Mo21], [Ma20], [SB21], [Br22], [WS20]), wobei oft Effizienzsteigerungen, Kosten- und/ oder Ressourcenreduktionen und Erhöhungen von Servicegraden erzielt werden konnten. Darüber hinaus entstanden neuartige Services und Geschäftsmodelle, die teilweise revolutionären Charakter haben. Dennoch kann man auch etliche Problemfälle und Risiken identifizieren: gescheiterte Projekte, falsche Ansätze, falsche oder ungünstige/ riskante IT-/ Softwarebasis, inkompetente (aber überzeugende) Berater, „nicht-optimales“ oder fachlich „suboptimales“ Umsetzungspersonal und so weiter.

Eine vom Autor durchgeführte Analyse ausgewählter Digitalisierungsvorhaben und -initiativen auf verschiedenen Ebenen und Diskursbereichen (vgl. auch oben) zeigt, dass nicht selten Defizite bestanden und bestehen, die Projekte und Initiativen entweder zum Scheitern brachten oder diese unbefriedigende, suboptimale Ergebnisse lieferten. Diese Tatsache zeigt, dass es notwendig ist, systematisch für die jeweiligen Ebenen/ Anwendungsbereiche zu untersuchen, was unter Verwendung der geeigneten (teilweise disruptiven) Technologien digitalisiert werden kann und soll sowie was dabei hinsichtlich des Ergebnisses und hinsichtlich des Weges zum Ergebnis berücksichtigt werden sollte.

In den folgenden Abschnitten des Papers werden einige wesentliche Aspekte genannt und kurz diskutiert. Vor dem Ausblick werden Kerninhalte eines möglichen Studienganges „Intelligente Digitalisierung“ aufgeführt. Eine These ist, dass einige Informatik- und/ oder Wirtschaftsinformatik-Studiengänge entsprechend umgestaltet werden sollten oder dass an deutschen Hochschulen zusätzlich und ergänzend Studiengänge zum Themenfeld „Intelligente Digitalisierung“ eingerichtet werden sollten.

2 Initiale Charakterisierung und Abgrenzung

Bei „Intelligenter Digitalisierung“ geht es – initial intendiert - darum, *erforderliche und sinnvolle Digitalisierungen* in Wirtschaft und Gesellschaft (auf den verschiedenen Ebenen) *zielpassend* und *termingerecht gestaltbar* und *realisierbar* zu machen. Dabei sollen verfügbare IT-Technologien kompetenzbasiert nutzbringend und passend genutzt und eingesetzt werden, wobei ein besonderes Augenmerk auf den Möglichkeiten von KI-Technologien liegt. Die beiden wesentlichen Teile der Intelligenten Digitalisierung, die fallbestimmt verschieden priorisiert werden können sind:

1. intelligentes Vorgehen in allen Phasen von Digitalisierungsvorhaben/ -projekten (einschließlich Projekten der Digitalen Transformation); hier geht es also auch um die konkrete Ausgestaltung der Phasen Analyse, Entscheidungsfindung, Vorbereitung, Planung, Begleitung und Durchführung und auch um die Schaffung von notwendigen/ begünstigenden Rahmenbedingungen
2. geeigneter Einsatz von (zu entwickelnden) KI-basierten Softwarekomponenten – integriert in den resultierenden Prozessen

Zusätzlich läßt sich aufgrund der exponierten Bedeutung ein auf Einhaltung ethischer Prinzipien ausgerichteter Teil aufnehmen:

3. Berücksichtigung ethischer Grundsätze/ ethischer Prinzipien in den Digitalisierungsprozessen, -projekten und in den Resultaten der Digitalisierungen

Ergebnis sollen, grob genannt, IT-basierte Prozesse und/ oder Organisationsstrukturen sein, die die Zielanforderungen bestmöglich (meist unter Berücksichtigung eines geeigneten Kosten-/ Nutzenverhältnisses) erfüllen und ggf. integrierte Entscheidungsunterstützungs- und Beratungsfunktionalitäten optimal realisieren. Neben funktionalen Anforderungen spielen natürlich auch nichtfunktionale Anforderungen eine essentielle Rolle, wobei der Änderbarkeit und Agilität (siehe z.B. auch [EB20], [Jo15], [Jo17], [Jo21]) der IT-basierten Zielprozesse und -strukturen eine besondere Bedeutung zukommt. In diesem Zusammenhang können auch digitale Zwillinge als Bestandteil einer Intelligenten Digitalisierung angesehen werden, die dann zum Beispiel für die Dokumentation, Analyse, Simulation und Planung verwendet werden können.

Der Begriff *Digitalisierung* beinhaltet i.d.R. nicht Anforderungen hinsichtlich der Güte des Resultats der Digitalisierung oder hinsichtlich der Qualität des Prozesses/ Projekts der Digitalisierung. Diese beiden wichtigen Aspekte sind aber mit dem Begriff *Intelligente Digitalisierung* deutlich adressiert. Zusätzlich könnte man ggf. *Digitale Transformation* ebenfalls in *Intelligente Digitalisierung* inkludiert betrachten, da auch hier Qualitätsanforderungen an die Prozesse oder Projekte der Digitalen Transformation an sich ergänzt werden, ebenso an die Zielprozesse und -strukturen, wobei KI-Technologien oft eine zentrale Rolle spielen können.

3 Betrachtungsdimensionen der Intelligenten Digitalisierung

Für die Intelligente Digitalisierung gibt es verschiedene Betrachtungsebenen/ -dimensionen (vgl. auch [Jo22a]) und Diskursbereiche. Eine mögliche Aufteilung in Betrachtungsebenen ist:

- Unternehmen & Organisationen
- Unternehmensverbände & Zuliefernetze
- Behörden
- Länder und Staaten
- Staatenverbände und Staatenorganisationen

Möglicherweise wäre es sinnvoll, feingranularer zu unterteilen – unter anderem abhängig von den Eigenschaften der gewünschten IT-basierten Zielprozesse und -services. Klar ist, dass die Ziele der Digitalisierungen an sich von verschiedenen Stakeholdern abhängen und aus der jeweiligen Sicht dieser auch bewertbar sind. Im folgenden werden zu den Betrachtungsdimensionen einige (nicht vollständig) inhaltliche Aspekte angerissen und weiterführende Quellen exemplarisch benannt.

3.1 Unternehmen und Organisationen

Bei Unternehmen und Organisationen stehen oftmals Effizienz, Erhöhung der Servicegrade, Ressourcen- und Kosteneinsparungen, Ergebnisoptimierung aber auch Agilität, ggf. Transparenz und Resilienz im Fokus der Digitalisierungsprozesse. Kommen Aspekte hinsichtlich der Schaffung zusätzlicher oder gar gänzlich neuer Dienstleistungen hinzu, dann werden Inhalte Digitaler Transformationsprozesse relevant. Werden umfassend und flächendeckend (oder mindestens zu einem hohen Grad) Prozessfunktionalitäten digital realisiert, die menschlich intelligentem Agieren gleichkommen oder solche überflügeln, so bietet es sich an, hier Begriffe, wie *Intelligentes Unternehmen* und *Intelligente Organisation* oder auch *KI-basiertes Unternehmen* und *KI-basierte Organisation* zu verwenden. Für weitere Aspekte, z.B. zu KI-basierten Bausteinen Digitaler Unternehmen, sei auf [Jo15], [Jo17] und [Jo21] verwiesen.

Ist der Typ eines Unternehmens oder einer Organisation beschränkt, so gibt es ggf. Spezifika für diesen Typ. Beispielsweise findet man „KI-orientierte Gedanken und Ideen“ zur Intelligenten Hochschule in [Jo22b].

Neben der bestmöglichen Digitalisierung der Funktionsbereiche und Prozesse, die dann oft auch weitere vorgegebene Eigenschaften, wie Agilität und Resilienz haben sollen, sind die Unternehmens- beziehungsweise Organisationsschnittstellen in die Digitalisierungsprozesse einzubeziehen, wobei beispielsweise Interoperabilitätsaspekte berücksichtigt werden müssen.

3.2 Unternehmensverbünde/ -netze & Zuliefernetze

In der Industrie/ Wirtschaft gibt es oft unternehmensübergreifende Prozesse, die insgesamt Gegenstand von Digitalisierungsprozessen sind. Beispielfelder hierfür sind Zulieferketten/ Zuliefernetze, virtuelle Unternehmen (möglicherweise temporär und relativ kurzlebig), unternehmensübergreifende kollaborative Planungsprozesse, kollaborative Durchführung von Projekten (im Sinne von CSCW). Eine Unterklasse davon sind gemeinsame Produktentwicklungen (concurrent engineering). Ähnliche Strukturen finden wir auch bereits innerhalb größerer Unternehmen – zum Beispiel in Form interner Produktions- und Zuliefernetze. Aus Sicht der einzelnen beteiligten Unternehmen ist neben Effizienz, Resilienz, Ergebnisqualität, Kostenoptimalität usw. von besonderer Bedeutung, dass nicht alle Informationen und nicht das gesamte Unternehmenswissen anderen Beteiligten zur Verfügung gestellt werden darf/ zugreifbar sein darf, sondern nur das freigegebene.

Beispiele, wie KI-basierte Softwarekomponenten, direkt als Resultat von Digitalisierungen in Unternehmensnetzen eingesetzt werden können, sind kommunizierende Planungsagenten, kommunizierende Konfigurations-/ Designagenten und natürlich auch datenbasierte KI-Komponenten, beispielsweise zum Auswerten RFID-basierter Datenmengen in Zuliefernetzen. Ein exemplarisches Beispiel, wie KI in Form von KI-basierten Simulationstools im Zusammenspiel mit Projektprozessen Digitalisierungsprojekte im Zuliefernetz-Umfeld unterstützen kann, findet man in [BJ03] beschrieben.

3.3 Behörden

Behörden als Gegenstand Intelligenter Digitalisierung könnten möglicherweise auch vollständig in den Betrachtungsdimensionen Land, Staat und Staatengemeinschaft angesiedelt werden. Allerdings ist hier der Blickwinkel stark lokal bezogen auf die Behörde, obwohl eine Behörde natürlich eine Landesbehörde, eine Bundesbehörde oder beispielsweise eine EU-Behörde sein kann. Ein Hauptfokus aus Sicht der Bürgerinnen und Bürger sollte darauf liegen, Dienstleistungen der jeweiligen Behörde effizient (auch im Sinne von rasch/ beschleunigt und kostenminimal) zu erbringen – am besten mit einem höheren Servicegrad, zum Beispiel IT-gestützte Ummeldung von Kraftfahrzeugen ohne realen Behördengang. Bezogen auf die konventionellen Prozesse gibt es hier viel Potenzial. Allerdings finden sich in diesem Gebiet einige Negativbeispiele für Digitalisierungsprojekte (z.B. Terminbuchungssystem für die Berliner Bürgerämter) und es stellt sich unter anderem gelegentlich die Frage nach der hinreichenden Setzung systemischer Randbedingungen für eine Intelligente Digitalisierung und nach der kompetenten Gestaltung und Organisation der Digitalisierungsprojekte in dieser Betrachtungsdimension. Neben der eigentlichen Digitalisierung von Behördenprozessen können natürlich auch die Durchführung IT-basierter Projekte und Aktionen (z.B. Vergabe Deutsch-Französischer Freundschaftspässe im Juni 2023, initiiert durch das Bundesministerium für Digitales und Verkehr) unter der Betrachtungsdimension Behörde angesiedelt werden. Die Analyse solcher Projekte und Fragen nach Intelligenter Vorgehen, Intelligenter Realisierung, Transparenz, Gütebewertung und Überwachung/

Steuerung sind es ausdrücklich wert, durchgeführt beziehungsweise thematisiert zu werden, würden aber an dieser Stelle den Rahmen des Papers sprengen.

3.4 Länder und Staaten

Wie bereits im Abschnitt 3.3 thematisiert, finden sich in dieser Betrachtungsdimension auch Fragestellungen der Intelligenten Digitalisierung, die sich ausschließlich auf Behörden von Ländern und Staaten in einer relativ kontextfreien Betrachtung beziehen. Zusätzlich gibt es hier weitere wichtige Fragestellungen der Intelligenten Digitalisierung, die eine globalere Sicht verfolgen. Insbesondere geht es auch um die Gestaltung neuer Funktionalitäten und Prozesse von Ländern und Staaten (ggf. in Form neuer oder erweiterter Behörden, Behördeninteraktionen usw.), um agile Reorganisations- und Dimensionierungsfragestellungen sowie um strategische Planungsprobleme. Einige Informationen und aktivierende Gedanken hierzu findet man in [Jo22a]. Aus der Sicht des IT-basierten Managements gibt es hier deutliche Bezüge zu Rekonfigurations-/ Redesignansätzen, zu KI-basierten Planungsansätzen und zu Simulationssystemen. Anwendungsfälle hierfür könnten z.B. die Reorganisation einer Behördenlandschaft eines Bundeslandes oder des Staates sein. Aus anderem Blickwinkel sind hier auch analoge Entsprechungen zu Unternehmensnetzen/ -verbänden zu finden. So müssen Behörden in behördenübergreifende Prozesse und Projekte eingebunden sein, dürfen allerdings Teile ihrer Informationen und ihres organisationalen Wissens nicht teilen. Auch hier stellt sich die Frage, ob, inwiefern, wofür und in welcher Weise zum Beispiel lernende Verfahren in Kombination mit symbolischen KI-Ansätzen eingesetzt werden könnten und sollten. Offensichtlich sollten in der Betrachtungsdimension „Länder und Staaten“ auch Schnittstellenaspekte berücksichtigt werden. Neben den Prozessen zwischen Behörden eines Landes sind ggf. auch Kommunikationsschnittstellen zur Realisierung von Länder- bzw. auch staatenübergreifenden Prozessen Gegenstand der Betrachtungen.

Generell stellt sich auch die Frage, ob Staaten und Staatengemeinschaften mit ihren staatlichen Instituten, Universitäten und Hochschulen nicht in der Lage sein sollten, unabhängig von Unternehmen zügig eigene Digitalisierungsprojekte erfolgreich durchzuführen. Wenn sie das derzeit nicht sind, sollte man sie dazu nicht befähigen? Warum konnte zum Beispiel nicht durch den Staat zügig ein funktionierendes Terminbuchungssystem für COVID-Impfungen realisiert werden?

3.5 Staatenverbände und Staatenorganisationen

Sicherlich gibt es bzgl. dieser Betrachtungsdimension etliche IT-basierte Prozesse, die für die Öffentlichkeit nicht unbedingt transparent sind. Hintergrund ist, dass bestehende (z.B. NATO, G7) oder auch neu zu schaffende Staatenverbände und -organisationen (z.B. Anti-Putin-Koalition) „optimal arbeitsfähig“ sein sollten, wobei selbstverständlich wieder die Privatheit von einigen Informationen und von organisationalem Wissen gewährleistet sein muss und darüber hinaus rechtliche Randbedingungen eingehalten/ respektiert werden

müssen. Hier identifizierbare Fragestellungen finden sich (in inhaltlicher Verzahnung zu der Betrachtungsdimension „Länder und Staaten“) wiederum in den IT-Anwendungsfeldern Konfiguration, Entwerfen, Design, Planung und haben damit inhaltlich ein hohes Potenzial für den Einsatz von KI-basierten Softwaresystemen/ -komponenten, die integriert in Prozesse (– in Kooperation mit menschlichen Akteuren –) ein intelligentes Gesamtsystem schaffen. Neben dieser technologischen Sicht besteht auch hier die Frage nach dem intelligenten Vorgehen zum Erhalt und zur agilen Adaption eines solchen Gesamtsystems.

4 Ausgewählte Aspekte und Subthemen

In diesem Kapitel sollen in kurzer Weise – ohne Anspruch auf Vollständigkeit – Aspekte und Subthemen der Intelligenten Digitalisierung angesprochen werden, die auf den verschiedenen Betrachtungsebenen (vgl. Kapitel 3) Relevanz besitzen. Bei Projekten der Intelligenten Digitalisierung sollten alle diese Aspekte adäquate Berücksichtigung finden.

4.1 Ethische, juristische und gesellschaftliche/ gesellschaftspolitische Aspekte, Datensicherheit und -schutz

Unbestreitbar lassen sich sehr viele Funktionalitäten und Prozesse auf und zwischen den verschiedensten Betrachtungsebenen digitalisieren – auch wenn derzeit in einigen Digitalisierungsprojekten deutlich Defizite (teilweise massiv) hinsichtlich der Organisation, Durchführung, hinsichtlich des Projektergebnisses und des Projekterfolgs an sich zu erkennen sind. Mit fortschreitenden Entwicklungen der Technologien werden weitere große Bereiche digitalisierbar werden. Im Zuge des Begriffs „Intelligente Digitalisierung“ stellen sich hier Fragen, welche Teildigitalisierungen ethisch vertretbar sind und welche nicht, dabei sind nicht nur die IT-Funktionalitäten zu betrachten, sondern die gesamten IT-basierten Prozesse, die den jeweiligen Gegenstandsbereich prägen und ggf. andere Bereiche tangieren. Ist es zum Beispiel ethisch korrekt, IT-Komponenten zu entwickeln und in Arbeitsprozesse zu integrieren, die es ermöglichen, Mitarbeiterinnen und Mitarbeiter zu überwachen? Ist es ethisch korrekt, potentielle Bewerber IT-basiert zu klassifizieren, wer sichert, dass die Klassifikation korrekt ist? Ist es ethisch korrekt, das Ziel zu verfolgen, durch den Einsatz von KI alle Beschäftigten mittelfristig zu Geringverdienern zu degradieren? Ist es ethisch korrekt, Services anzubieten, die Songs auf der Beliebtheitskala „künstlich“ nach oben bringen? Die Reihe derartiger Fragen kann natürlich noch lang fortgesetzt werden – unter anderem auch Fragen nach Digitalisierungen in militärischen Bereichen – und berührt natürlich auch den gesellschaftsorganisatorischen und gesellschaftspolitischen Raum. Ist es sinnvoll, viele Menschen von der Arbeit „zu befreien“ oder sollte es gezielt verboten werden, einiges zu digitalisieren, soll es Digitalisierungsquoten geben usw.? Offensichtlich sollte und wird diesbezüglich über gesetzliche Regulierung auf verschiedenen Ebenen nachgedacht/ nachgedacht werden. Derzeit wird

zum Beispiel über ein „nationales KI-Register“ und den „Artificial Intelligence Act“ der EU berichtet und debattiert.

Bezüglich gesetzlicher Regulierungen sind komplexe Überlegungen anzustellen, die unter anderem systemische Zusammenhänge zwischen Staaten, Unternehmen usw. berücksichtigen. Wie sollen sinnvolle Regulierungen aussehen? Wer überwacht die Einhaltung der Regulierungen? Wie können im Verletzungsfall Sanktionen durchgesetzt werden? Kann man Riesenkonzerne und Diktatoren tatsächlich bändigen und im Zaum halten? Wie wird gesichert, dass keine Wettbewerbsnachteile entstehen und keine Unternehmensabwanderungen stattfinden?

Ein weiterer zentraler Aspekt der (Intelligenten) Digitalisierung ist der Datenschutz und die Datensicherheit. Auf jeder Betrachtungsebene der Intelligenten Digitalisierung müssen die damit verbundenen Fragestellungen kompetent, konsequent und korrekt berücksichtigt werden. Blicke in die Praxis zeigen auch hier, dass es teilweise erhebliche Probleme und Defizite gibt, die in der Konsequenz erhebliche negative Auswirkungen ergeben können.

Für weitere Gedanken zum Thema „KI, Digitalisierung und Ethik“ sei auf [Ge19] verwiesen. Ausführungen zu rechtlichen Aspekten bei der Anwendung von KI findet man in [BM21].

4.2 Technische und wirtschaftliche Grenzen und Komplexitätsfallen

Zur Intelligenten Digitalisierung gehört es auch, den Sinn von Digitalisierungsprojekten/-vorhaben zu hinterfragen. Neben den ethischen Aspekten spielen dabei auch ökonomische/ wirtschaftliche Fragestellungen sowie technische und theoretische Limitierungen eine Rolle. Den derzeit verbreiteten ChatGPT-Hype dämpfend gibt es hierzu bereits einige ernstzunehmende kompetente Aussagen, die sich zum Beispiel auf die technische und ökonomische Limitierung zukünftiger Sprachmodelle für generative KI beziehen. Ebenso gibt es Arbeiten, die die Fehlerausgaben und (gezielte und unbeabsichtigte) Generierung von Falschaussagen/ Falschinformationen thematisieren.² Werden diese und andere Warnungen nicht ausreichend ernst genommen, kann es zu weitreichenden Fehlentwicklungen in Wirtschaft und Gesellschaft kommen! Hinsichtlich der ökonomischen Aufwände stellen auch große neuronale Netze/ Deep Learning erhebliche Herausforderungen, insbesondere wegen des erheblichen Energiebedarfs für die Lernprozesse aber auch wegen inhaltlicher Beschränkungen, die bei Nichtkenntnis oder/ und Nichtberücksichtigung zu negativen Konsequenzen führen können. Relativ hohe Energiebedarfe werden auch durch die intensive Verwendung von Blockchain-Technologie erzeugt.

² Unbestritten lassen sich Technologien und Ansätze im Kontext generativer Sprachmodelle für einige Digitalisierungen auch sinnvoll und nutzbringend einsetzen. Hinsichtlich der Limitierung und einiger Defizite von ChatGPT sei exemplarisch auch auf [GF22] und [ID23] verwiesen.

Einige Herausforderungen beflügeln natürlich intensive Aktivitäten in Forschung und Entwicklung, beispielsweise bei der Entwicklung verbrauchsärmerer Schaltkreise oder auf dem Sektor des Quantencomputings. Generell stellen die meisten Limitierungen aber – zumindest für die nächsten Jahre – reale Schranken dar.

Problemräume/ adressierte Problemklassen von Digitalisierungsprojekten und -vorhaben beinhalten aufgrund ihrer „natürlichen Struktur“ oftmals kombinatorisch komplexe Teilprobleme. Auch wenn durch den Einsatz (hybrider) KI-Methoden einige kombinatorisch komplexe Probleme der Lösbarkeit zugänglich gemacht werden können, ist dies nur für beschränkte Problemgrößen möglich. Dieser Fakt muss bei Skalierungsbegehren/ Skalierungsbedarfen entsprechend berücksichtigt werden, was wiederum kompetente Entscheidungsträger und/ oder kompetente Berater erfordert. Hinsichtlich der Entwicklungen auf dem Gebiet des Quantencomputings versprechen sich einige das Vordringen in höhere Komplexitätsklassen. Nach Meinung von Experten wird dieses Vordringen – bezogen auf praktische große Probleme – nicht weitreichend und umfassend sein. Interessante und weiterführende Ausführungen zu Grenzen der KI findet man in [MK22].

4.3 Prozesskonformität, das richtige Ziel und der richtige Weg

Die Entwicklung und Einführung von KI-basierten integrierten Systemkomponenten ist nicht losgelöst von den Prozessen zu planen und durchzuführen, in denen diese integriert werden sollen. Beides, Prozesse und IT-Komponenten, bilden eine sich gegenseitig bedingende und beeinflussende Einheit, die in der Analyse, in der Planung und im Ziel holistisch betrachtet werden muss. Blicke in die Praxis zeigen auch hier starke Defizite und fehlerhaftes Vorgehen. So werden teilweise (in nicht wenigen Unternehmen) stolz unter dem PR-Mantel der Modernität und eines hohen Digitalisierungsgrades Funktionalitäten auf Basis von Standardsoftwaresystemen (z.B. für Reisekostenabrechnungen) eingeführt, die zwar für buchhaltungsversierte Angestellte verwendbar sind, für andere jedoch nicht oder nur mit einem enormen Zeitaufwand. Wird dann auch noch vergessen, diesbezügliche Feedbackschleifen im Sinne von Lern- und Änderungsprozessen sowie ein Risikomanagement zu etablieren, so entsteht Chaos, Ineffizienz und die Kontraproduktivität schlägt zu. Die ursprüngliche Intension, Buchhaltungspersonal einzusparen wird dadurch konterkariert, dass lohnkostenintensiveres Personal erheblichen Zeitaufwand hat und ausserdem Frust mit entsprechenden Folgen generiert wird. Es nutzt die beste Software nichts, wenn sie nicht zu den Prozessen paßt und der Akzeptanzgrad (gewollt oder ungewollt) gering ist. Die Prozesskonformität ist ein wesentlicher Aspekt der Zielgestaltung bei der Intelligenten Digitalisierung. Generell, wie bei allen Entwicklungsprojekten, muss eine Zielvorstellung/ Zielspezifikation entwickelt werden. Im Sinne einer agilen Vorgehensweise kann diese natürlich teilweise auch temporären und unvollständigen Charakter haben, wenn dies passend beziehungsweise erforderlich ist. Das Schaffen einer korrekten und vollständigen Zielsystemspezifikation (IT-Komponenten integriert in Zielprozesse) ist Gegenstand des klassischen Requirements Engineerings. Entsprechend abgewandelt gilt dies bei möglicher und notwendiger Anwendung einer agilen Vorgehensweise. Die aus dem klassischen und agilen (je nach

Vorgehensweise) Requirements Engineering bekannten Elemente sind direkt auf die Intelligente Digitalisierung übertragbar, ebenso die anderen aus dem Software Engineering bekannten Phasen. Wird das zu schaffende Gesamtsystem als System betrachtet, das auch in Zukunft noch Änderungen und Reorganisationen benötigen wird, so kann man nach dem entsprechend zu erweiternden Modell des Prozessmanagementlebenszyklus vorgehen, das in der Wirtschaftsinformatik allgemein bekannt ist.

Hinsichtlich des Requirements Engineerings, hinsichtlich eines zyklischen Vorgehens mit Test- und Adaptionsschleifen und damit auch hinsichtlich der Lehre des Software Engineerings scheinen in der Praxis relativ oft Fehler vorzukommen. Unklar bleibt, ob diese verstärkt aus fehlenden Kenntnissen oder aus Kosteneinsparungsgründen erfolgen. Ist bei wichtigen Digitalisierungsprojekten und -vorhaben das ausreichende Vorhandensein von Kompetenzen gesichert? Neben den inhaltlichen Aspekten des Software Engineerings müssen ebenfalls belastbare Kenntnisse und Fähigkeiten auf dem Gebiet des Projektmanagements/ IT-Projektmanagements vorliegen. Diese bestimmen den Weg zum Ziel-Gesamtsystem, einschließlich der Realisierung und Umsetzung von sich dynamisch ergebenden zusätzlichen oder revidierenden Anforderungen. Beispielsweise ist zyklisch eine Stakeholderanalyse durchzuführen, aus der sich möglicherweise Änderungen an der Spezifikation des Zielsystems (IT & Prozesse) ergeben.

Abhängig von der Betrachtungsebene und dem Kompetenzbereich der Projektauftraggeber und/ oder Projektakteure ist es sinnvoll, ein entsprechendes Wissensmanagement bzgl. gesammelter Erfahrungen aufzusetzen und zu etablieren. Dies kann auch sinnvoll für entsprechende Kontrollgremien sein. Ist dies zum Beispiel in Deutschland bereits geschehen? Warum kommt es immer wieder zu peinlichen Systemabbrüchen, kürzlich zum Beispiel bei der Vergabe Deutsch-Französischer Freundschaftspässe im Juni 2023, initiiert durch das Bundesministerium für Digitales und Verkehr. Wußten die Verantwortlichen nicht, dass man bei „erhöhten Zugriffszahlen“ lieber mehrere Server oder/ und Cloud Computing verwenden sollte? Hatte man fachlich kompetente Berater, gab es Systemtests? Zum Glück ging es „nur“ um die Vergabe von Bahntickets und nicht um sicherheitsrelevante Systeme. Dennoch sind viele junge Leute wegen dieser Panne frustriert gewesen. Neben der Etablierung und Nutzung von „Lessons Learned“ kommt natürlich auch „Benchmarking“ und der Analyse von „Best Practices“ als mögliche Elemente einer Intelligenten Digitalisierung eine bereichernde Bedeutung zu.

4.4 Sicherstellung von Kompetenzen/ Technologiekenntnissen und Problemverständnis

Abhängig von der Wichtigkeit des Gelingens von Projekten und Vorhaben der Intelligenten Digitalisierung ist auf belastbare Kompetenz zu achten. Dies umfaßt Technologie- und Vorgehenskompetenz und -kenntnisse und gilt für die im Projekt tätigen Mitarbeiter und Mitarbeiterinnen aber auch für die Mitglieder von Steuerungs- und Beratungsgremien. Was wäre/ ist zum Beispiel, wenn nicht erkannt werden kann, dass Teilprobleme kombinatorisch komplex sind, was wenn nicht bekannt ist, was das für Konsequenzen hat?

Wer berät bei relevanten/ wichtigen Digitalisierungsvorhaben/ -projekten einen Auftraggeber, der nicht vom Fach ist, der Auftragnehmer? Macht es sich nicht Sinn analog zu einer Bauaufsicht bei Hausbauprojekten einen fachlich kompetenten Supervisor anzuhören? Es würden sich hier noch etliche Fragen formulieren lassen, die in der derzeitigen Praxisrealität oft negativ behaftete Grundlagen haben.

Interdisziplinarität ist sicher bei vielen Projekten und Vorhaben der Intelligenten Digitalisierung förderlich oder gar notwendig. Aber in jedem Fall benötigt es eine kritische Masse an kompetenten Fachexperten. Einige Analysen von Digitalisierungs-Beratungsgremien auf der Betrachtungsebene Land und Staat zeigen, dass dies – zumindest teilweise – nicht gegeben ist:

- Beispiel IKT-Lenkungsrat des Landes Berlin (Stand: 5/22): 27 Mitglieder, darunter ein Wirtschaftsinformatiker (FH), ein Ingenieur, Ausbildung einer Person nicht ermittelbar, die restlichen 24 Mitglieder sind fachfremd!
- Beirat Digitalstrategie Deutschland (Stand: 6/23): 19 Mitglieder, darunter 2 Informatiker, 1 Ingenieur, die restlichen 16 Mitglieder sind fachfremd!

Beachtenswert ist, dass die Ausbildung der Gremienmitglieder nicht transparent ist/war, sondern erst relativ aufwendig ermittelt werden musste. Statt der tatsächlichen Transparenz wird auf den entsprechenden Webseiten Kompetenz „vorgegaukelt“.

Hinsichtlich der erfolgsabhängigen Besetzung von Beratungs- und Steuerungsgremien wäre es wünschenswert und notwendig, genaue systemische Analysen durchzuführen und eine diesbezügliche Transparenz herzustellen. Wer bestellt hier wen in Gremien? Warum treffen sich scheinbar in verschiedenen Gremien einige Gruppen immer wieder?

Ist es wirklich förderlich, für die Leitung von großen Digitalisierungsprojekten und den zugehörigen Organisationseinheiten fachfremde Personen zu berufen, die dann in Interviews sinngemäß äußern, dass ja bezüglich der IT eigentlich alles da ist und es jetzt auf die Organisation ankommt? *Vielleicht ja*, aber nur, wenn diese sich mehrere fachkompetente und fähige Berater engagieren. Die belastbare Kompetenz der Berater sollte sich in Summe (möglichst personenredundant) auf alle Teildisziplinen der Intelligenten Digitalisierung erstrecken!

4.5 Modelle/ Digitale Zwillinge

Ein weiterer Aspekt der Intelligenten Digitalisierung ist – abhängig von der Betrachtungsebene und dem Gegenstand und der Relevanz des Digitalisierungsvorhabens/ -projektes – das Thema „Modelle des Zielsystems“. Solche Modelle können bei entsprechenden Abstraktionsniveaus als digitale Zwillinge betrachtet werden, die unter anderem der Dokumentation, Transparenz, der Analyse, Planung und simulativen Absicherung dienen können. Zum Thema Modelle – bezogen auf Unternehmen – sei auf [Jo17] verwiesen.

5 Intelligente Digitalisierung – Inhalte und Notwendigkeit gesonderter oder erweiterter Studiengänge

Eine Analyse gängiger Informatik-/ Wirtschaftsinformatik-Studiengänge sowie einiger Studiengänge zur Digitalen Transformation zeigt, dass in diesen nicht alle für die Durchführung von Projekten oder Vorhaben der Intelligenten Digitalisierung notwendigen/ förderlichen Fächer/ Kurse enthalten sind.³ Nach Meinung des Autors ist im Paper die Besonderheit und Bedeutung der Intelligenten Digitalisierung dargelegt und motiviert worden. Der Autor stellt zur Diskussion, dass es sinnvoll wäre, Studiengänge zum Themengebiet Intelligente Digitalisierung anzubieten. Dies könnten einerseits aufbauende Erweiterungen oder Spezialisierungen von Bachelor-Informatik-/ Wirtschaftsinformatikstudiengängen sein oder separate Bachelor-Studiengänge (unter Nutzung ausgewählter Module anderer Studiengänge) oder konsekutive Master-Studiengänge.

Initial vorzusehende Kurse (als Vorschlag zur Diskussion):

- Einführung in die Intelligente Digitalisierung
- Digitale Transformation
- Programmieren I & II
- Software Engineering
- KI 1 - Einführung in die Künstliche Intelligenz
- KI 2 - Neuronale Netze und Deep Learning
- KI 3 - Data Science
- KI 4 - Logik, Deduktion, Regelverarbeitung
- KI 5 - Deklarative Programmierung/ Constraint Programmierung
- Cybersecurity
- BWL und Management
- VWL
- IT-Recht
- Cloud Computing & BIG Data
- Requirements Engineering
- Projektmanagement
- Prozessmanagement & Digitalisierung
- Webprogrammierung
- Wissensmanagement und -verarbeitung
- Ethik

³ Im vom Autor an seiner vorherigen Hochschule, HWTK/ VICTORIA, entwickelten und erfolgreich akkreditierten und reakkreditierten Studiengang „B.Sc. Informatik und Management“ (Dual & Vollzeit) sind die relevanten Kurse mit Schwerpunktsetzung auf Themen der Künstlicher Intelligenz enthalten. Dieser Studiengang wird derzeit nicht mehr angeboten.

Abhängig von den durch die Studierenden beabsichtigten „Ziel-Einsatzgebieten“ (die könnten im Fall von dual Studierenden initial vorgegeben sein), könnten unter anderem folgende *Wahlkurse* angeboten werden:

- Europäisches Management
- Internationales Management
- Europäisches Recht
- Verwaltungswissenschaften
- Verwaltungsinformatik
- Verwaltungsrecht

Neben der Schaffung neuer Studiengänge für das Themenfeld Intelligente Digitalisierung ist es aus Sicht des Autors ebenso empfehlenswert, bestehende Studiengänge auf themenrelevante Ergänzungsmöglichkeiten zu untersuchen und ggf. zu erweitern/ zu adaptieren. Darüber hinaus sollten ausgewählte themenbezogene Kenntnisse bereits in die Schulausbildung integriert werden.

Fazit und Ausblick

Im Paper wurde der Begriff *Intelligente Digitalisierung* charakterisiert. Hauptfacetten sind einerseits die intelligente Vorgehensweise bei Digitalisierungsprojekten und Projekten der Digitalen Transformation in allen Phasen und Belangen, andererseits der *zielgerechte Einsatz prozessintegrierter KI-basierter Komponenten*, wobei ethische Prinzipien gewahrt werden sollen/ müssen. Verschiedene Betrachtungsebenen/ -dimensionen der Intelligenten Digitalisierung und ausgewählte Aspekte/ Subthemen wurden andiskutiert und durch inhaltliche Vorschläge für gesonderte Studiengänge zum Themengebiet Intelligente Digitalisierung angereichert.

Verschiedene weitere (kollaborative) Weiterentwicklungen und Tätigkeiten bieten sich im Sinne von *Future Work* an. Zu nennen sind hier Ergänzungen, Präzisierungen und Erweiterungen der im Paper genannten Aspekte und eine Überarbeitung der Systematisierung. Zum Beispiel im Kontext eines aufsetzbaren Forschungsprogramms wäre vermutlich auch die Analyse von ausgewählten erfolgreichen und nicht-erfolgreichen Digitalisierungsprojekten der verschiedenen Betrachtungsebenen nutzbringend. Hinsichtlich des Umfelds von Digitalisierungsvorhaben und -projekten auf den Betrachtungsebenen Länder, Staaten, EU, Unternehmen könnten gezielt transparenzbringende Recherchen – möglicherweise teilweise automatisierbar – durchgeführt werden. Systemische Analysen und Betrachtungen für ausgewählte Anwendungsgebiete wären weitere themenrelevante Tätigkeitsfelder.

Literaturverzeichnis

- [Ai21] Aichele, C., Herrmann, J. (Hrsg.): Betriebswirtschaftliche KI-Anwendungen – Digitale Geschäftsmodelle auf Basis Künstlicher Intelligenz. 2. Auflage, Springer Vieweg 2021.
- [BJ03] Baumgärtel, H., John, U.: Combining Agent-Based Supply Net Simulation and Constraint-Technology for Highly Efficient Simulation of Supply Networks using APS Systems. Winter Simulation Conference 2003, New Orleans 2003.
- [BM21] Barton, T., Müller, C. (Hrsg.): Künstliche Intelligenz in der Anwendung – Rechtliche Aspekte, Anwendungspotenziale und Einsatzszenarien. Springer Vieweg 2021.
- [Bi22] Bischoff, M. (Hrsg.): Künstliche Intelligenz – Vom Schachspieler zur Superintelligenz? Springer 2022.
- [Br22] Brune, G.: Künstliche Intelligenz heute – Anwendungen aus Wirtschaft, Medizin, Wissenschaft. Springer Vieweg 2022.
- [EB20] Eckstein, J., Buck, J.: Unternehmensweite Agilität. Vahlen 2020.
- [Ge19] Geske, U.: Künstliche Intelligenz und Digitalisierung – sind wir auf dem richtigen Weg? In [JHW19].
- [GF22] Gene Freuder, Me and ChatGPT, freuder.wordpress.com/author/freuder/, Stand: 19.7.2023.
- [Hw21] hwtk: HWTK Forschung im Fokus, Forschungsbericht Akademisches Jahr 2019/ 20, Berlin, 2021.
- [ID23] Intelligente Digitalisierung (John, U.), www.intelligente-digitalisierung.de/ChatGPT-Cases, Stand: 24.7.2023.
- [JHW19] John, U., Hofstedt, P., Wolf, A. (Hrsg.): Deklarative Ansätze zur Künstlichen Intelligenz – punktuelle Beiträge. Online-Post-Proceedings des 32nd Workshop on (Constraint) Logic Programming, Berlin, 2019.
- [Jo15] John, U.: Digitales Unternehmen – Bausteine für Effizienz, Agilität und Transparenz. INFORMATIK 2015 (Cunningham, D. et al (Ed.)), Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, 2015.
- [Jo17] John, U.: Deklarative Unternehmensmodelle – essentielle Bausteine Digitaler Unternehmen für optimierende Reorganisationen. INFORMATIK 2017 (Maximilian Eibl, Martin Gaedke (Ed.)), Lecture Notes in Informatics (LNI), Gesellschaft für Informatik 2017.
- [Jo21] John, U.: Ein Forschungsfeld stellt sich vor: Agilität, Effizienz und Transparenz für komplexe Prozesse in Unternehmen, Unternehmensnetzen und Organisationen – Digitalisierung und Digitale Transformation. In [Hw21].
- [Jo22a] John, U.: Intelligentes Unternehmen, Intelligenter Staat, Intelligentes Europa – Was ist zu tun? – Anstoß einer kollegialen Experten-Diskussion -. INFORMATIK 2022 (D. Demmler et al (Hrsg.)), Lecture Notes in Informatics (LNI), Vol. P326, Gesellschaft für Informatik 2022.

- [Jo22b] John, U.: Intelligente Hochschule – Digitalisierung und Künstliche Intelligenz im Hochschulkontext. INFORMATIK 2022 (D. Demmler et al (Hrsg.)), Lecture Notes in Informatics (LNI), Vol. P326, Gesellschaft für Informatik 2022.
- [LMR17] Lingau, V., Müller-Seitz, G., Roth, S. (Hrsg.): Management der Digitalen Transformation – Interdisziplinäre theoretische Perspektive und praktische Ansätze. Vahlen 2017.
- [Ma16] Matzler, K. et al: Digital Disruption – Wie Sie Ihr Unternehmen auf das digitale Zeitalter vorbereiten. Vahlen 2016.
- [Mo21] Mockenhaupt, A.: Digitalisierung und Künstliche Intelligenz in der Produktion – Grundlagen und Anwendungen. Springer Vieweg 2021.
- [MK22] Mainzer, K., Kahle, R.: Grenzen der KI – theoretisch, praktisch, ethisch. Springer 2022.
- [Ma20] Marr, B.: Künstliche Intelligenz in Unternehmen – Innovative Anwendungen in 50 erfolgreichen Firmen. Wiley 2020.
- [SB21] Schmid, U., Bruckert, S.: Künstliche Intelligenz in Unternehmen – Zielgruppenspezifische KI-Komponenten identifizieren und vermitteln. Hanser 2021.
- [WS20] Weber, R., Seeberg, P.: KI in der Industrie – Grundlagen, Anwendungen, Perspektiven. Hanser 2020.

Domain-agnostic Intelligent Digital Twins: Merging of Application-near Knowledge Representations with the Proactive Internet of Digital Twins (IoDT)

Joel Lehmann ¹, Andreas Lober ², Tim Häußermann ¹, Alessa Rache ¹, Hartwig Baumgärtel ², Julian Reichwald ¹

Abstract: Flexibility and reconfigurability regarding mass-customized products are the key premises of future industrialization. The amalgamation of Digital Twin (DT) approaches, knowledge-representing technologies, and skill-based engineering methodologies allow such approaches to emerge. This paper deals with the integration of knowledge representations as a driver for holistic system intelligence ranging from physical assets to the Internet of Digital Twins (IoDT) within the Digital Twin Reference Model (DTRM). This is what it takes to spur DTs to intelligence and proactive collaboration behaviors. Despite the domain-agnostic, universal applicability of the concept, exemplary implementation approaches from the field of production and intralogistics illustrate the feasibility. Specific domain knowledge can be dynamically aggregated and used as a basis for negotiation scenarios between DTs.

Keywords: Digital Twin, Internet of Digital Twins, IoDT, Knowledge Graph, Marketplace, Cyber-physical Systems

1 Introduction

The future of industrialization transitions (Industry 4.0 / Industry 5.0) requires flexibility and reconfigurability in the trend toward mass customization [Do20; Lo22]. This requires domain-independent, data-driven Information Technology (IT) instruments and methodologies. To create holistic, cognitive systems, the hurdles of Operational Technology (OT)/IT convergence must therefore be broken through [Le23a; Pa21].

One convenient way to do this is through the concept of the Digital Twin (DT). These are mostly shaped by passive characteristics and need to be composed of active, online, goal-seeking, and anticipatory behaviors in the future [Gr22]. Knowledge-based representation and interaction technologies are therefore not only inevitable for proactive collaboration but also for achieving domain-agnostic deployment features. The Digital Twin Reference Model (DTRM) according to Lehmann et al. [Le23a] combines the perspectives of various established industry standards with future requirements of DTs. The therein-contained approach of a superimposed cross-company Internet of Digital

¹ Mannheim University of Applied Sciences, Center for Mass Spectrometry and Optical Spectroscopy, Paul-Witsack-Straße 10, 68163 Mannheim, Germany, [firstletterfirstname].[lastname]@hs-mannheim.de

² Ulm Technical University of Applied Sciences, Institute for Industrial Engineering and Supply Chain Management, Prittwitzstraße 10, 89075 Ulm, Germany, [firstname].[lastname]@thu.de

Twins (IoDT) inspires the twinning paradigm to be more autonomous and proactive than mere static representations realizing intelligent Digital Twins. Specifically, this should allow negotiation scenarios to take place autonomously in an agent-like manner. The combination of DTs, semantic web technologies, and skill-based engineering can deliver new levels of interoperability [Ma22]. In most negotiation scenarios, a distinction can be made between demand and supply. To realize such scenarios, the concept of modularization through capability-skill-service structures can be applied regardless of the application domain [Kö23]. From the bottom up, a hierarchical skill-based approach offers the greatest possible freedom to map and generate control code and, based on this, the DT of a system [Lo22].

To be able to operate as technology-neutral and domain-agnostic as possible, the basic ideas of semantic representations within the systems will be examined in more detail in this paper. The different knowledge graphs involved in the overall system must be considered holistically, and mediation methods among them need to be investigated. This should allow the technical framework of the DTRM to be universally deployed to fully exploit the added value of the intelligent DTs within the IoDT. Despite the universality of the concept, exemplary implementation strategies from a use case of smart production and intralogistics will be shown subsequently.

This paper contributes an approach to integrating universal knowledge between different operational levels of DTs. The relevant knowledge graphs build up dynamically during operation and aggregate application-specific knowledge. Skill-based engineering methodologies have revealed themselves to be suitable tools for negotiation processes. It is further worked out that the approaches outlined not only inspire DTs to intelligence in the production context but also benefit other fields of application.

The remainder of the paper is structured as follows. First, the related work discusses skill-based engineering methods, followed by related knowledge representing technologies regarding the preliminary work of the DTRM in Section 2. While section 3 deals with the formal analysis of knowledge representations in the overall system, section 4 shows practical implementation approaches. Section 5 discusses the results of the work and concludes with an outlook in section 6.

2 Related Work

To understand the central concepts of the paper, the approaches of skill-based engineering for the realization of modularized structures are outlined first. Then semantic technologies and their mediation procedures are discussed in more detail with respect to the preliminary work of the DTRM and IoDT, which will be further extended hereby.

2.1 Skill-Based Engineering

Skill-based Engineering as a key concept of Industry 4.0 provides the foundation for tackling dynamic and flexible conditions in production as well as highly customized demands [Go16; La14]. The concept focuses on the development of automation systems that rely on suitable orchestration mechanisms. These orchestration mechanisms provide the basis for skill coordination in order to fulfill specific production steps [Zi1]. The concept disassembles products in their atomic components that are manufactured by a set of continuously performed skills. Production resources are decomposed to their offered skills. By describing products and resources based on capabilities and skills, a mapping between them can be enabled. In that case, capabilities are abstract abilities that contain a description of skills, independent of a concrete implementation by a resource. Skills, on the other hand, are concrete abilities of resources that realize capabilities. To ensure interoperability and adjustability the skill description needs to be in a vendor-independent unified way [Do20]. The work plan of a product consists of different capabilities that are fulfilled by skills offered by resources. Thus, the mapping between the products and the resources is done by matching the capabilities with specific skills. Therefore, Köcher et al. [Kö23] propose a Capability-Skill-Service (CSS) model to combine skills, capabilities, and services with the established Product-Process-Resource (PPR) concept. In this model, skills are defined as resource specific functions described by the associated capabilities. A set of capabilities that refer to specific functions independent of their implementation provide services. Dorofeev et al. [DZ18] show the implementation of a skill structure through a common OPC UA interface to realize a service-oriented architecture (SOA) in the context of smart production. This paradigm is characterized by the organization of complex distributed systems by subdivision into services, their interoperability, and loose coupling [DZ18; LCK16]. Therefore, SOA principles coincide with those of intelligent production systems. Jungbluth et al. [Ju23] demonstrate the feasibility of a skill-based approach implemented via OPC UA to control a flexible transport system via state machines. Lober et al. al. [Lo2] show that control logic and knowledge graph based on skills can improve the engineering process of control systems. Therefore, knowledge approaches within the DTRM will be addressed in this context subsequently.

2.2 Knowledge Representation in the Digital Twin Reference Model

In [Le23a], we have shown that a distinction of operational areas with respect to the deployment of DTs is relevant for their understanding and further interaction. By correlating with different already existing reference models of the fourth industrial revolution, such as Reference Architectural Model Industry 4.0 (RAMI 4.0) or Industrial Internet Reference Architecture (IIRA), new specifications for the realization of intelligent, multi-agent driven intelligent DTs have been derived.

The idea of DTs can be traced back to Michael Grieves, who has always been the smallest common denominator in all subsequent concepts and developments. He

separates physical space, in which physical devices (e.g., assets, processes, documents, etc.) are located, digital space, in which the digital images of the physical entities operate, and the digital thread, which connects the two bidirectionally. Through this interconnection to the digital space, the physical devices are provided with the tools of state-of-the-art IT technologies such as artificial intelligence algorithms, etc. [BCF19; Gr22; Sj20].

Our preliminary work on the DTRM was also based on these topologies. As shown in Figure 1, the DTRM is divided into three integration levels. These do not separate as conventionally in hierarchy levels according to IEC 62264 but are completely integrated horizontally and vertically. The differentiation is only made depending on the proximity to the physical process and is therefore divided into *Physical Twin Space*, *Intranet of Digital Twins*, and *IoDT*. To do so, all physical devices located at the field level are represented with their controls in the *Physical Twin Space*. On the higher level, the company-internal replication of the DTs takes place in the *Intranet of Digital Twins*. This is where the aggregation, harmonization, and interface creation of all data and the first possible processing stages (e.g., skill orchestration, predictive maintenance, data analytics, simulation, etc.) take place. Component-specific knowledge graphs describe the DTs semantically and integrate with the overall facility knowledge graph of the *Intranet of Digital Twins* to gain interoperability within the closed facility environment. To achieve cross-company and higher-level interaction and collaboration of intelligent DTs, the *IoDT* is home to a marketplace-oriented collaboration approach, as described in [LE22]. This is also based on a specific marketplace knowledge graph, which is partially funded as a set of the underlying ones.

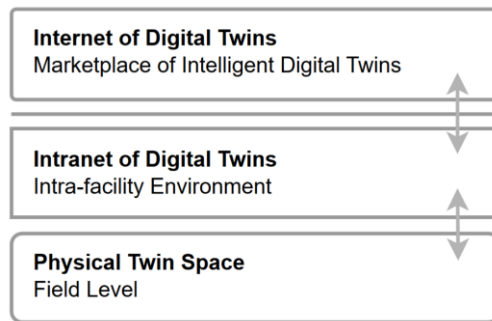


Fig. 1: Topology of the Digital Twin Reference Model by Lehmann et al. al. [Le23]

Agnostic of the application domain, physical devices can be integrated and empowered with active collaboration behavior. In the context of research data management and Internet of Things (IoT) enabled measurement devices, this could be partially shown in in [Le23b], but the focus here is on production processes, as already shown in in [Lo22] and [Le2]. The concept can be adapted to all IoT use cases as long as the resource is sufficiently semantically described. Semantic technologies are the core of the concept and converge for the interaction schemes of the intelligent DTs on the Internet of Digital

Twins. In order to clarify how the various components of knowledge are composed, a specific definition of the respective semantic sources, their structure, further processing, and higher-level aggregation at the superordinate reference levels, in particular the IoDT, is required.

In the sense of the knowledge relationships to be applied here, knowledge graphs consisting of nodes and edges are employed. For this purpose, triples existing of two nodes are connected to each other via a directed edge. These triples are defined by the (Resource Description Framework) RDF [BHL01]. The structure (terminology) of this set of graph objects here is based on specific ontologies which must each be aggregated through mediation approaches. The resulting knowledge graph is, therefore, more holistic, and also includes all necessary assertions [ZLX17]. Ontology mediation or integration comprises three major disciplines: ontology mapping, ontology alignment, and ontology merging. Ontology mapping focuses on the correlations between two ontologies. These correlations are stored independently from the source ontologies and used for querying diverse knowledge bases. The detection of correlations in a semi-automated or automated way is called ontology alignment. A similarity parameter is used to determine the individual confidence of each determined relation. The process of ontology merging describes the operation of creating a new ontology based on at least two source ontologies s [Br06; OYD21]. Therefore, the following section will deal with the formal aggregation and mediation of process knowledge, so that an approach for its implementation can be shown subsequently.

3 Universal Mediation of Knowledge within the DTRM

In order to sharpen the understanding of the amalgamation of the knowledge representations on the different levels of the DTRM, the individual components of the participating knowledge graphs will be formally described. It is assumed that all sets of a general knowledge graph Ω are of the same type and structure. Therefore, the following considerations can be carried out with set annotation at a high abstraction level. For simplicity and exemplification, we also assume the similarity of the different knowledge graph elements, which in turn results in a direct merging without necessary alignment rules. Starting with the description of an asset knowledge graph, the superimposed ones are derived up to the IoDT. According to Köcher et al. [Kö23], parts of the CSS model provide a suitable starting point to adequately describe assets based on their skills and capabilities which they can use in later negotiation scenarios. Therefore, it can be assumed that Ω_{CSS} is a partial subset of Ω_{asset} , respectively is aligned to, cp. Equation 1. In addition, all contained knowledge (e.g., structural hardware/software setup), skills, and their further orchestrated services are mapped in Ω_{asset} .

$$\Omega_{CSS} \subseteq \Omega_{asset} \tag{1}$$

Further, let Φ be a distinct facility and Ω_Φ the instance of the facility knowledge graph. Consequently, it must be composed of the union of the specific domain knowledge Ω_{domain} in which the facility is operating and the knowledge of all assets $\Omega_{asset1}, \Omega_{asset2}, \dots, \Omega_{asset3}$ located at facility Φ as well as the basic knowledge structure of the super classes relevant for a marketplace $\Omega_{marketplace}$ in general, cp. Equation 2.

$$\Omega_\Phi = \Omega_{domain} \cup \bigcup_{i=1}^n \Omega_{asset_i} \Big|_{asset_i \in \Phi} \cup \Omega_{marketplace} \quad (2)$$

It can be assumed that an instantiated marketplace knowledge graph at the IoDT Ω_{IoDT} will aggregate more and more knowledge of its registered assets Ω_{asset_j} of different facilities over its lifetime. This is specifically expressed in the growing number of subclasses, which are not initially part of the structural marketplace knowledge graph $\Omega_{marketplace}$. This leads to the conclusion that the marketplace knowledge graph is a subset of the IoDT knowledge graph Ω_{IoDT} , cp. Equation 3, and that all registered assets from various facilities Ω_{asset_j} at the IoDT are a subset of Ω_{IoDT} , cp. Equation 4.

$$\Omega_{marketplace} \subseteq \Omega_{IoDT} \quad (3)$$

$$\bigcup_{j=1}^m \Omega_{asset_j} \Big|_{asset_j \in IoDT} \subseteq \Omega_{IoDT} \quad (4)$$

To clarify how the universal approach to knowledge management is addressed within the DTRM, the following implementation approaches show practical methods for its realization.

4 Implementation Approach

To illustrate and implement the entire knowledge within the DTRM, a further preliminary work by Lober et al. [Lo23] can be used in which the control logic was re-located from the physical asset on the field level to the intra-facility environment as shown in Figure 2. The *Automated Storage and Retrieval System Digital Twin* (ASRS DT) is composed of different subcomponents and is made available to the DTRM as DT within the *Intranet of Digital Twins*. The adaptation into the DTRM according to Lehmann et al. [Le23a] serves as a framework for the classification of the various layers. In it, the open-source project Eclipse Ditto is utilized as a middleware to directly

represent the DTs. The entire knowledge representations are made available to the rest of the environment using GraphDB, which natively provides an HTTP interface for querying and interaction. The communication between all system components is realized by MQTT. The framework for interaction, negotiation, and collaboration is a specially developed Python environment.

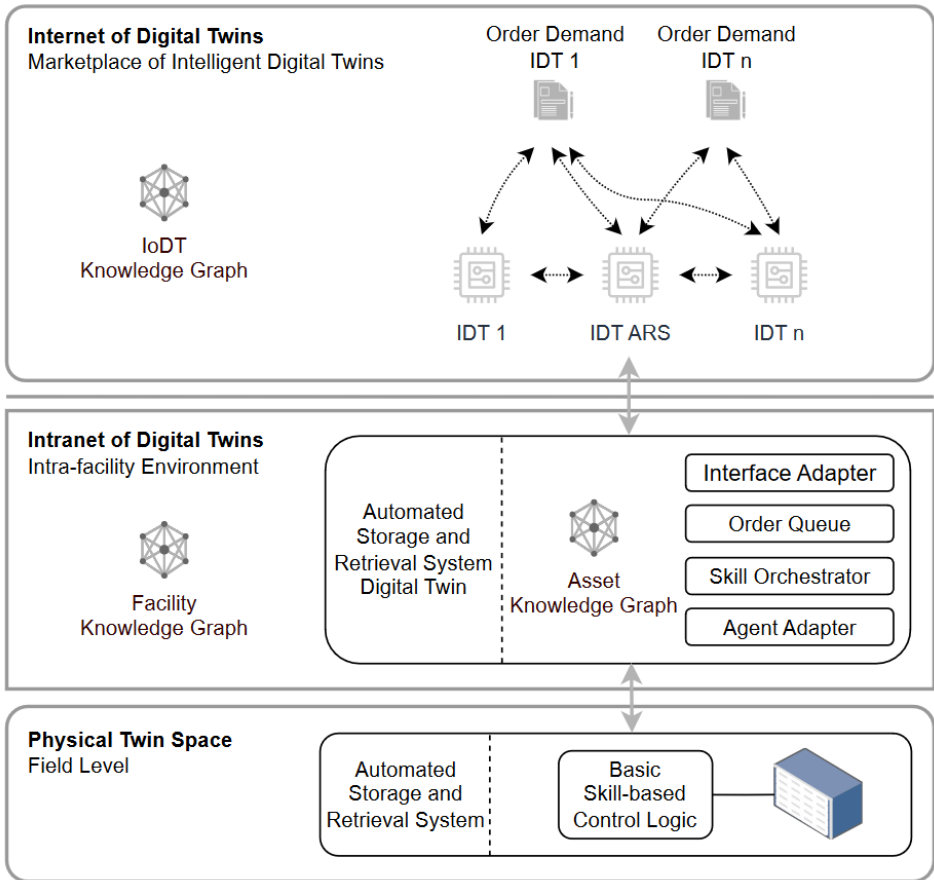


Fig. 2: Architectural Approach for the Practical Amalgamation of Decentralized Process Control Knowledge to the Back-end by Lober et al. [Lo23] through the Application of the Digital Twin Reference Model by Lehmann et al. [Le23a] Refining Process Knowledge within the Internet of Digital Twins in Order to Facilitate Active Collaboration Behavior of Intelligent Digital Twins.

The ASRS DT directly brings a specific knowledge graph describing its skills, based on the CSS model framework. This can consequently be used for merging into the *Facility Knowledge Graph*. Both are then available for the generation and instantiation of the marketplace-oriented *IoDT Knowledge Graph*. The ASRS DT, the also-called

Orchestration Agent according to Lober et al. [Lo23] can be used to enable fast and intelligent networking of different resources. Each of these *Orchestration Agents* have a knowledge graph for the represented resource and can assign to the superordinated IoDT via its offered services. Finally, the intelligent DTs of the cross-company assets registered on the marketplace in the IoDT can then collaborate and negotiate with product demands intelligently. The *IoDT Knowledge Graph* instantiated from the marketplace knowledge graph according to Equations 3 and 4 provides the basis for this.

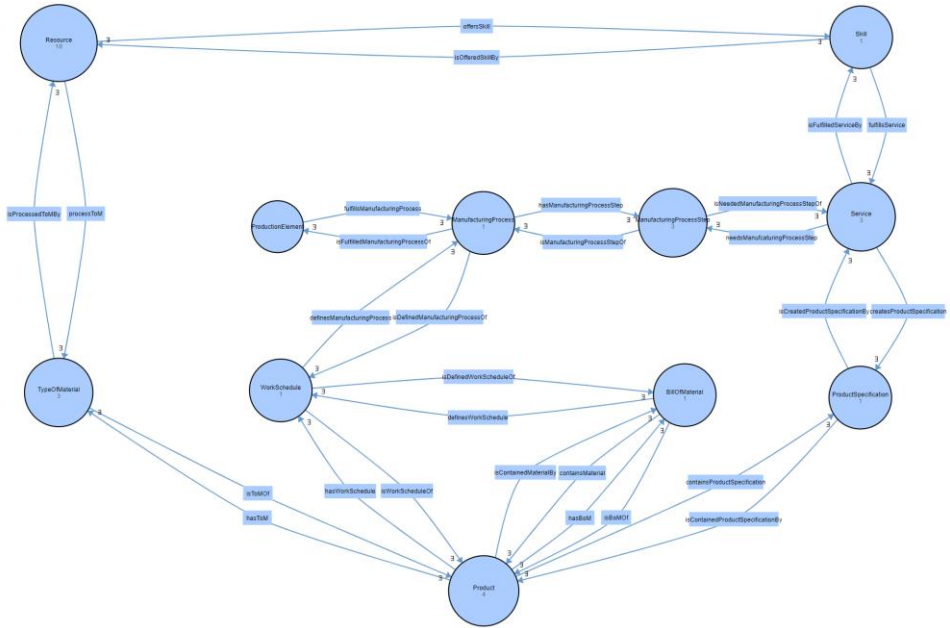


Fig. 3: Merged Facility Knowledge Graph Including Domain-Specific Knowledge Graph, Asset Knowledge Graph, and Marketplace Knowledge Graph.

To support the integration process of the different knowledge graphs, it is necessary that the different semantic representations follow the same structure of class hierarchy. Figure 3 shows the merged knowledge graphs consisting of the domain-specific knowledge graph, the asset knowledge graph, and the marketplace knowledge graph. This merged knowledge graph forms the so-called *Facility Knowledge Graph* and follows the formal structure according to Equation 2 presented in section 3. According to Figure 3 the *Resource*, *Skill*, *Production Element*, *TypeOfMaterial*, and *Service* classes belong to the asset knowledge graph of the represented resource. For the example of the ASRS, these would be the individual components as well as the entire system as a subclass of the *Resource* class, a retrieval and storage process as an example for specific service, the arrangement of the system in a storage or production environment as a subclass of the *ProductionElement* class, and the possible material types that can be “processed” by the ASRS, in this case, retrieved or stored. The *Service* is the last class

that comprises the asset knowledge graph. Here, the connection between skill and service is established so that matching can be performed in a marketplace environment at the IoDT later.

For the marketplace knowledge graph, the classes *WorkSchedule*, *TypeOfMaterial*, *Product*, *BillOfMaterial*, *ProductSpecification*, and *Service* can be identified from Figure 3. In the example of the ASRS, this part of the knowledge graph is used to provide information about products that are not part of the intra-facility environment. The third part of the *Facility Knowledge Graph* is the domain-specific knowledge graph. In this area, the relation to the service request and service provision is defined. In the example of the implementation of this paper, the production lines and routing as well as the manufacturing processes are described here and placed in context with the product and the resources. The class *WorkSchedule* contains the actual work schedule of on specific product and defines its product manufacturing process. This product manufacturing process is represented by the class *manufacturing process*. The manufacturing process of a specific product defines the work sequence in which the product to be manufactured must be produced and when which resource must perform which skill in order to fulfill specific services in a manufacturing process step and to produce specifications of a product.

In the implementation example, the extension of the use case from another preliminary work by Lober et al. [Lo22] is assumed. Therein, a drilling process within a company that is autonomously controlled via a DT is exemplarily registered at the digital marketplace. In the extension of this use case, the logistical difficulty arises that the sheet metal required to manufacture the product is no longer available within the company. The IoDT marketplace should be used to identify a nearby resource of a service provider that could provide this sheet metal. Based on the presented *Facility Knowledge Graph*, queries can be sent to the *IoDT Knowledge Graph*, and various storage systems with the appropriate material will respond. Based on various selection criteria, such as delivery time and availability, the optimal resource for processing the provisioning service can be selected. However, in order for this process to work smoothly, it is necessary to combine the various involved knowledge graphs dynamically into one *Facility Knowledge Graph*. One general approach for integrating the required DTRM knowledge according to Equation 2 is demonstrated through the Algorithm shown in Figure 4. It describes the implementation of the knowledge graph merging process in a language-independent way. The basic prerequisite for this algorithm is the placement of all knowledge graphs to be merged in the intended directory with the simultaneous constraint of semantic similarity to avoid an alignment for the sake of simplicity at this implementation stage.

The process starts by defining the required variables like the *new_namespace* for the merged knowledge graph, *merged_graph* an empty graph object, and an empty list named *graphsList*. Subsequently, all knowledge graphs in the knowledge graphs directory are instantiated as graph objects and added to the *graphsList*. In the next step, this list is iterated over. For each knowledge graph in the list, the namespace is extracted

```
1: new_namespace ← Create a new namespace for merged knowledge graph
2: merged_graph ← Create a new empty graph object
3: graphsList ← an empty list

4: for all knowledgeGraphs in the knowledgeGraphs directory do
5:   emptyGraph ← Create a new empty graph object
6:   Parse the knowledge graph file in XML format and add it to emptyGraph
7:   Append emptyGraph to graphsList

8: for all graphs in graphsList do
9:   namespace_graph ← GETNAMESPACE(knowledgeGraph)
10:  if # in namespace_graph then
11:    Replace # with "" in namespace_graph
12:  graph ← CHANGENAMESPACE(knowledgeGraph, namespace_graph, new_namespace)
13:  Replace knowledgeGraph with graph in graphsList

14: for all graphs in graphsList do
15:   for all triple in graph do
16:    if triple is not in merged_graph then
17:     Add triple to merged_graph

18: Serialize merged_graph as an XML file named "merged_knowledgeGraph.rdf"
```

Fig. 4: Algorithm to Merge Knowledge Graphs.

using a function named *GetNamespace*. This function returns the knowledge graphs' namespace. In the returned namespace “#” is removed if it exists as a safety measure. This is required based on the declaration of *new_namespace* without a “#”. Afterward, an updated graph is created utilizing a function called *ChangeNamespace*. This function changes the namespace to the defined *new_namespace* by iterating over all triples in the input knowledge graph and returning the graph with a changed namespace. It must be emphasized that the first two loops in the algorithm can be combined into one but are presented separately for clarity and comprehensibility. The process continues with the creation of the merged knowledge graph. Therefore, all triples in all graphs from the *graphsList* are iterated over and merged into the *merged_graph*. Thereby it is ensured that no duplicates are inserted into the *merged_graph*. To complete the process the *merged_graph* is serialized as an XML file.

5 Discussion

On the basis of the formal definitions for the structure of the knowledge representations within the DTRM, practical methods for realization could be shown through the implementation approaches. A tangible example from the area of smart production systems and intralogistics for the development of the knowledge graphs as well as the superordinate IoDT knowledge graph provided the basis for the structure and interplay

with the herein derived facility knowledge graph. Oriented on the formal description of the knowledge relationships, an exemplary merging algorithm is shown, which dynamically aggregates the knowledge of the multiple DTRM levels. Thereby it does not matter in which domain the DTRM framework finds its application, as long as a specified domain language is adhered to and thus the defined similarity requirement within the formal description is met. If similarity no longer prevails, the approaches for mediating the knowledge representations must be further expanded and enhanced to include a variable alignment of the ontologies. This would also allow for a transdisciplinary collaboration scenario between DTs from different domains in the future.

Furthermore, the skill-based structure of the asset and their knowledge graphs can be applied to other application domains. As the prerequisite for intelligent negotiation processes skills and services can always be broken down into demand and supply or consumer and provider behavior of e.g., products or services. A holistic negotiation and cross-company interaction require the skill-service concept as shown in this paper. Examples do not just include adoptions in production context as shown in [Le23a; Lo22; Lo23] but also the possible applicability to problems of state-of-the-art research data management as shown in [Le23b], where measuring instruments autonomously document their results and negotiate and correlate measurement outcomes, is given. Scenarios from the IoT and smart city context are also mentioned, where system participants could, for example, decide on traffic situations using a wide range of environmental measurements. Also in the medical context, the DTRM and its knowledge representations could be used for the negotiation and management of all accruing data. In this case, patient data could be used to improve processes during interventions on a role-based principle, together with all the measured values that accumulate. In the future, the goal should be to sharpen the domain openness of the concepts for as many application fields as possible and to have different constellations of knowledge dynamically available within the layers in order to make the interaction of the system components as free as possible. The authors are aware that the highly abstracted use cases must be increased to set up a reliable proof of concept and evaluation.

6 Conclusion and Future Work

In order to meet future increasing requirements for flexibility and reconfigurability, this paper contributes approaches for integrating knowledge representations within the DTRM introduced by [Le23a]. After derivation of the basic approaches of skill-based engineering and the specific knowledge representations of each level of the DTRM, an approach to domain-agnostic mediation of process and collaboration knowledge is formally defined. On this basis, practical implementation approaches for dynamic knowledge aggregation and a vivid example of an ASRS were shown. Skill-based methodologies enable negotiation and collaboration processes beyond the usual manufacturing context in many other areas. Additionally, due to the open concept, the

approaches can be transferred to a wide range of application areas (e.g., research data management, IoT, and medical engineering). This helps to achieve higher levels of automation through vertically integrated intelligent DTs and their specific knowledge at their various operating levels specifically through the integration of the different knowledge graphs. Only the respective domain-specific parts must be exchanged, and a common domain vocabulary has to prevail so that independent new negotiation scenarios between different DTs can take place across company boundaries within the IoDT.

However, there are some limitations that have to be addressed. If different domains (e.g., manufacturing, and medical) with different vocabulary are to collaborate with each other, the semantic similarity deviates and an alignment becomes necessary. This would allow resources that differ from the defined domain standard to be integrated into the framework and participate in the collaboration. Furthermore, the semantic relationships presented here at a high level of abstraction must be properly defined in preparation for independent alignment and the implementation of resulting algorithms. This can further strengthen domain openness and make cross-domain intelligent DT collaboration scenarios possible. Due to the constraints regarding interoperability and standardization, it should be noted that the approaches shown require a strong progress of the paradigms proclaimed by Industry 4.0. Thus, there is still a long way to go until the real adaptation of such open marketplace-based approaches can take place.

Acknowledgements

Parts of the work presented in this paper were supported by a grant from the German Ministry of Education and Research (BMBF), grant number 03IHS264A (TransforMA).

Bibliography

- [BCF19] Barricelli, B. R.; Casiraghi, E.; Fogli, D.: A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* 7, pp. 167653–167671, 2019, issn: 2169-3536, url: <https://ieeexplore.ieee.org/document/8901113/>, visited on: 05/04/2021.
- [BHL01] Berners-Lee, T.; Hendler, J.; Lassila, O.: The Semantic Web. *Scientific American* 284/5, pp. 34–43, 2001, issn: 00368733, 19467087, url: <http://www.jstor.org/stable/26059207>, visited on: 06/10/2023.
- [Br06] de Bruijn, J.; Ehrig, M.; Feier, C.; Martins-Recuerda, F.; Scharffe, F.; Weiten, M.: Ontology Mediation, Merging, and Aligning. In: *Semantic Web Technologies*. John Wiley & Sons, Ltd, pp. 95–113, July 2006, url: <https://doi.org/10.1002/047003033x.ch6>.

- [Do20] Dorofeev, K.: Skill-based engineering in industrial automation domain. In: Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Companion Proceedings. ACM, June 2020, url: <https://doi.org/10.1145/3377812.3381394>.
- [DZ18] Dorofeev, K.; Zoitl, A.: Skill-based Engineering Approach using OPC UA Programs. In: 2018 IEEE 16th International Conference on Industrial Informatics (INDIN). Pp. 1098–1103, 2018.
- [Go16] Gorecky, D.; Weyer, S.; Hennecke, A.; Zühlke, D.: Design and Instantiation of a Modular System Architecture for Smart Factories. IFAC-PapersOnLine 49/31, pp. 79–84, 2016, url: <https://doi.org/10.1016/j.ifacol.2016.12.165>.
- [Gr22] Grieves, M.: Intelligent digital twins and the development and management of complex systems. Digital Twin 2/, p. 8, May 2022, url: <https://doi.org/10.12688/digitaltwin.17574.1>.
- [Ju23] Jungbluth, S.; Barth, T.; Nußbaum, J.; Hermann, J.; Ruskowski, M.: Developing a skill-based flexible transport system using OPC UA. at - Automatisierungstechnik 71/2, pp. 163–175, Feb. 2023, url: <https://doi.org/10.1515/auto-2022-0115>.
- [Kö23] Köcher, A.; Belyaev, A.; Hermann, J.; Bock, J.; Meixner, K.; Volkmann, M.; Winter, M.; Zimmermann, P.; Grimm, S.; Diedrich, C.: A reference model for common understanding of capabilities and skills in manufacturing. at - Automatisierungstechnik 71/2, pp. 94–104, Feb. 2023, url: <https://doi.org/10.1515/auto-2022-0117>.
- [La14] Lasi, H.; Fettke, P.; Kemper, H.- G.; Feld, T.; Hoffmann, M.: Industry 4.0. Business & Information Systems Engineering 6/4, pp. 239–242, June 2014, url: <https://doi.org/10.1007/s12599-014-0334-4>.
- [LCK16] Leitão, P.; Colombo, A. W.; Karnouskos, S.: Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges. Computers in Industry 81/, pp. 11–25, Sept. 2016, url: <https://doi.org/10.1016/j.compind.2015.08.004>.
- [Le22] Lehmann, J.; Lober, A.; Rache, A.; Baumgärtel, H.; Reichwald, J.: Collaboration of Semantically Enriched Digital Twins based on a Marketplace Approach. In: Proceedings of the 19th International Conference on Wireless Networks and Mobile Systems. SCITEPRESS - Science and Technology Publications, 2022, url: <https://doi.org/10.5220/0011141200003286>.
- [Le23a] Lehmann, J.; Lober, A.; Häußermann, T.; Rache, A.; Ollinger, L.; Baumgärtel, H.; Reichwald, J.: The Anatomy of the Internet of Digital Twins: A Symbiosis of Agent and Digital Twin Paradigms Enhancing Resilience (Not Only) in Manufacturing Environments. Machines 11/5, p. 504, Apr. 2023, url: <https://doi.org/10.3390/machines11050504>.
- [Le23b] Lehmann, J.; Schorz, S.; Rache, A.; Häußermann, T.; Rädle, M.; Reichwald, J.: Establishing Reliable Research Data Management by Integrating Measurement Devices Utilizing Intelligent Digital Twins. Sensors 23/1, p. 468, Jan. 2023, url: <https://doi.org/10.3390/s23010468>.

- [Lo22] Lober, A.; Lehmann, J.; Hausermann, T.; Reichwald, J.; Baumgärtel, H.: Improving the Engineering Process of Control Systems Based on Digital Twin Specifications. In: 2022 4th International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM). IEEE, Nov. 2022, url: <https://doi.org/10.1109/elecom54934.2022.9965259>.
- [Lo23] Lober, A.; Lehmann, J.; Reichwald, J.; Ollinger, L.; Baumgärtel, H.: Flexible Skill-based Production Systems through novel OPC UA Design Approaches. IFAC-PapersOnLine, accepted – to be published, pp. 4030–4035, 2023.
- [Ma22] Mayr, M.; Hvarfner, C.; Chatzilygeroudis, K.; Nardi, L.; Krueger, V.: Learning Skill-based Industrial Robot Tasks with User Priors. In: 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE). IEEE, Aug. 2022, url: <https://doi.org/10.1109/case49997.2022.9926713>.
- [OYD21] Osman, I.; Yahia, S. B.; Diallo, G.: Ontology Integration: Approaches and Challenging Issues. *Information Fusion* 71/, pp. 38–63, July 2021, url: <https://doi.org/10.1016/j.inffus.2021.01.007>.
- [Pa21] Patera, L.; Garbugli, A.; Bujari, A.; Scotece, D.; Corradi, A.: A Layered Middleware for OT/IT Convergence to Empower Industry 5.0 Applications. *Sensors* 22/1, p. 190, Dec. 2021, url: <https://doi.org/10.3390/s22010190>.
- [Sj20] Sjarov, M.; Lechler, T.; Fuchs, J.; Brossog, M.; Selmaier, A.; Faltus, F.; Donhauser, T.; Franke, J.: The Digital Twin Concept in Industry – A Review and Systematization. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Vol. 1, pp. 1789–1796, 2020.
- [Zi19] Zimmermann, P.; Axmann, E.; Brandenbourger, B.; Dorofeev, K.; Mankowski, A.; Zanini, P.: Skill-based Engineering and Control on Field-Device-Level with OPC UA. In: 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, Sept. 2019, url: <https://doi.org/10.1109/etfa.2019.8869473>.
- [ZLX17] Zhao, Y.; Liu, Q.; Xu, W.: Open Industrial Knowledge Graph Development for Intelligent Manufacturing Service Matchmaking. In: 2017 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII). Pp. 194–198, 2017.

Planung des Ländlichen On-Demand-Verkehr - Probleme, Analyse und Algorithmen

Sven Löffler,¹ Ilja Becker,¹ Petra Hofstedt,¹ André Nitze,² Silvia Hennig,³ Alexander Klinge²

Abstract: On-Demand-Verkehre können helfen, Lücken im ländlichen öffentlichen Personennahverkehr (ÖPNV) zu schließen. Hohe Lizenzgebühren für proprietäre Software halten jedoch viele Kommunen von der Einführung entsprechender Verkehre ab, insbesondere in strukturschwachen Regionen, die diese besonders nötig hätten. Hinzu kommt, dass entstehende digitale Angebote meist Software-Inseln bleiben und verfügbare Routing-Algorithmen derzeit noch weitgehend ungeeignet sind, bestehendem Linienverkehr zuzuliefern und Parallelverkehre zu vermeiden. In dieser Arbeit werden Ergebnisse einer Machbarkeitsstudie von ländlichem On-Demand-Verkehr im Rahmen eines Projektes vorgestellt und diskutiert. Wir betrachten dazu die aktuelle Situation in der gewählten Modellstadt Spremberg und analysieren die Anwendbarkeit verschiedener Algorithmen zur Planung von On-Demand-Verkehr. Ziel ist eine niedrighschwellige Open Source-Lösung für ein intermodales Routing im ländlichen Raum. Diese soll in die Open-Data-Plattform bbnavi integriert werden können, um die Nutzung bestehender Mobilitätsdaten zu vereinfachen, Interoperabilität herzustellen und die Übertragbarkeit in andere Regionen zu erleichtern. In einem geplanten Folgeprojekt soll auf Basis der Machbarkeitsstudie eine entsprechende Software entwickelt und in einem Modellprojekt in Spremberg erprobt werden. Die zu entwickelnde Open Source-Lösung soll den Kostendeckungsgrad von Letzte-Meile-Verkehren verbessern und es kommunalen Akteuren ermöglichen, selbst zu Mobilitätsdienstleistern zu werden und entsprechende Geschäftsmodelle zu entwickeln.

Keywords: Open Source; ÖPNV; On-Demand-Verkehr; Vehicle Routing; Shortest Path Problem

1 Einleitung

Der öffentliche Nahverkehr ist in großen Städten wie Berlin, Hamburg oder München in der Regel sehr gut ausgebaut und hoch frequentiert, so dass ein schnelles Erreichen beliebiger Zielorte von beliebigen Ausgangsorten in der Regel sowohl kostengünstig als auch zeitnah möglich ist. Bewegt man sich allerdings weg von den großen Städten, hin zu den ländlicheren Gebieten, so lassen der generelle Streckenausbau und die Frequentierung schnell stark nach. Ein Erreichen, selbst von teilweise nahe gelegenen Zielen, ist dann in der Regel nur noch mit dem Privatverkehr oder erheblichem zeitlichen Aufwand möglich.

¹ BTU Cottbus - Senftenberg, FG Programmiersprachen und Compilerbau, Konrad-Wachsmann-Allee 5 03046 Cottbus, Deutschland {SvenLoeffler,Hofstedt,Ilja.Becker}@b-tu.de

² Technische Hochschule Brandenburg, Fachbereich Wirtschaft, Magdeburger Straße 50, 14770 Brandenburg an der Havel, Deutschland andre.nitze@th-brandenburg.de

³ neuland21 e.V., Klein Glien 25, 14806 Bad Belzig, Deutschland, Geschäftsführerin Vereinsvorsitzende silvia.hennig@neuland21.de, Programmabereichsleiter Mobilität alexander.klinge@neuland21.de

Oftmals muss mittels ÖPNV eine Route über den nächst größeren Ort eingeschlagen werden um einen viel näher gelegenen angrenzenden Ort zu erreichen. Dieser, in der Regel im ländlichen Raum sternförmig ausgeprägte Verbindungsausbau, bietet damit oftmals keine ernstzunehmende Alternative zum eigenen Auto.

Einen weitaus dynamischeren Ansatz als den statischen Linienverkehr versprechen On-Demand-Verkehre. Diese könnten durch klassische Busse des Personennahverkehrs, die nicht auf festen Linien fahren sondern sich dynamisch den Bedürfnissen der Bevölkerung anpassen, realisiert werden. Vorteile dieses Vorgehens sind zum einen, dass keine hohen Anschaffungskosten entstehen und zum anderen, dass mit angepasster Streckenführung das Sternproblem umgangen werden und somit der bereits vorhandene ÖPNV unterstützt werden kann. Es entsteht daher keine Konkurrenz zu den bestehenden Anbietern, sondern eine Erweiterung und gleichzeitig eine bessere Erreichbarkeit des bestehenden Angebots.

Die in der Machbarkeitsstudie zu konzipierende Software für den ländlichen On-Demand-Verkehr in der Modellstadt Spremberg soll zusätzlich als Open Source-Software gestaltet werden. Insbesondere soll die bereits existierende Software bbnavi [23a] Verwendung finden. Die ausschließliche Nutzung von Open Source-Software soll eine leichtere Übertragbarkeit auf andere Städte und Kommunen erleichtern.

Das weitere Papier ist wie folgt aufgebaut. In Abschnitt 2 wird auf die Situation des ÖPNV und Besonderheiten der Modellstadt Spremberg als Vertreter für Städte und Gemeinden im ländlichen Bereich eingegangen. Zur Umsetzung des On-Demand-Service als Unterstützung für den ÖPNV wird ein Shortest Path-Algorithmus für die statische Planung des ÖPNV (Abschnitt 3) und ein Vehicle Routing-Algorithmus für die flexible Planung des On-Demand-Services (Abschnitt 4) benötigt. Mittels der Shortest Path-Algorithmen werden dann in der Umsetzung die schnellsten, günstigsten, bestgelegenen ÖPNV-Verbindungen ermittelt, wohingegen die Vehicle Routing-Algorithmen den eigentlichen On-Demand-Service planen und somit die optimalen Touren für die zur Verfügung stehende Fahrzeugflotte berechnen. Im Abschnitt 5 wird die Kombinierbarkeit von Shortest Path-Algorithmen mit Vehicle Routing-Algorithmen analysiert. Abschließend wird in Abschnitt 6 eine Zusammenfassung der bisherigen Arbeit und ein Ausblick auf zukünftige Arbeiten gegeben.

2 Besonderheiten des ÖPNV im ländlichen Raum am Beispiel der Modellstadt Spremberg

Spremberg ist eine Kleinstadt im Süden von Brandenburg an der Grenze zu Sachsen. Die Stadt verfügt über 21.000 Einwohner (stand vom 31.12.2021 [21a]) auf einer Fläche von knapp über 200 km². Das ergibt im Schnitt etwa 105 Einwohner je km². Im Vergleich dazu leben in Berlin ca. 4.100, in Hamburg ca. 2.500 und in München ca. 4.800 Einwohner je km². Aber nicht nur hinsichtlich der Bevölkerungsdichte sondern auch im Ausbau des ÖPNV unterscheidet sich Spremberg klar von den großen Städten. So gibt es z.B. 31 Regionalverkehrsrouten (Bahn), die Berlin als Start, Ziel oder Zwischenhalt bedienen. Für

die Stadt Spremberg ist das eine, die Verbindung von Zittau nach Cottbus Hauptbahnhof (stündlich zwischen 4:30 Uhr und 0:30 Uhr) und wieder (stündlich zwischen 5:22 Uhr und 23:30 Uhr) zurück. Über eine IC-, ICE oder S-Bahn-Anbindung verfügt Spremberg nicht [21b]. Alternativ zur fehlenden Bahnanbindung existiert eine Buslinie (Plus Bus 800), welche die Städte Cottbus, Spremberg, Schwarze Pumpe und Hoyerswerda miteinander verbindet (stündlich zwischen 5:00 Uhr und 21:00 Uhr).

Aufgrund dieser zahlenmäßigen Unterschiede, sowohl in der Bevölkerungsdichte als auch im Streckenausbau und der angebotenen Bus-/Bahnfrequenz, können für ländliche Gebiete, wie die Modellstadt Spremberg, nicht einfach die gleichen Methoden zur Bewältigung des ÖPNV angewendet werden wie in großen Städten oder Ballungsgebieten. In gut ausgebauten Großstädten genügt für das Finden eines schnellsten Transportes von einem Ort zum anderen ein sogenannter Shortest Path-Algorithmus [KV12]. Dies setzt allerdings voraus, dass alle Orte in akzeptabler Zeit mit dem ÖPNV erreichbar sind.

Straßen- und Verkehrsnetze, wie die der Stadt Spremberg und Umgebung, können durch gewichtete Graphen repräsentiert werden. Ein Shortest Path-Algorithmus findet dann den kürzesten Weg zwischen einem Start ($s \in V$) und einem Zielknoten ($z \in V$). Ausgangspunkt dabei ist ein gewichteter Graph $G = (V, E)$ mit Gewichtsfunktion $d : E \rightarrow \mathbb{R}$, wobei V eine Menge von Knoten (englisch vertices) und $E \in V \times V$ eine Menge von Kanten (englisch edges) bezeichnet.

Mit Hilfe der Abstraktion des gewichteten Graphen und einem Algorithmus zum Lösen des Shortest Path-Problems kann sowohl für den privaten Straßenverkehr als auch für den ÖPNV ein kürzester Weg ermittelt werden. Für den Straßenverkehr sind die Knoten Kreuzungen und Orte die durch Straßen (Kanten) miteinander verbunden sind. Die Gewichtsfunktion ordnet dabei jeder Straße eine geschätzte Fahrtdauer zu. Analog kann für den ÖPNV ein Graph verwendet werden, bei denen die Busstationen und Bahnhöfe die Knoten darstellen und die Schienen bzw. Buslinien durch Kanten repräsentiert werden. Auf Algorithmen zum Lösen des Shortest Path-Problems wird in Abschnitt 3 eingegangen. Da in Spremberg (und auch in vielen anderen ländlichen Städten und Gemeinden) kein flächendeckendes Bus- oder Bahn-Netz existiert und damit bestimmte Knotenpunkte nicht oder nur mit großen Umwegen erreichbar sind, müssen die bestehenden ÖPNV-Verbindungen durch flexible On-Demand-Services unterstützt werden. Eine Fahrt vom Friedhof Cantdorf (Spremberg) zum nächstgelegenen Ortsteil Weskow (Spremberg, Luftlinie 3 km, Fahrtzeit Auto 9 min, Fußweg 40 min) kann gegenwärtig z.B. nur mit erheblichem Zeitaufwand mittels ÖPNV bewältigt werden (Fahrzeit ÖPNV 46 min inklusive 24 min Fußweg).

Bei einem On-Demand-Service soll ein Fahrzeug oder eine Flotte von Fahrzeugen dazu genutzt werden, die Transportwünsche möglichst vieler Teilnehmer zu erfüllen und dabei die geringsten Kosten zu verursachen. Die beschriebene Problemstellung wird in der Literatur als Tourenplanung (englisch Vehicle Routing Problem, VRP) bezeichnet. Ziel ist dabei die Gruppierung und optimale Aneinanderreihung verschiedener Transportaufträge [TV02]. Aus den Gegebenheiten in Spremberg leiten sich weitere Einordnungen unseres Problems

in Unterklassen des VRP [Zh22] ab. So handelt es sich um ein Capacitated Vehicle Routing Problem (CVRP), d.h. es müssen also die verschiedenen Kapazitäten der Fahrzeuge (Auto, Kleinbus, Bus) berücksichtigt werden. Es genügt auch nicht, die Transporte einfach nur durchzuführen, sondern dies muss auch innerhalb zulässiger Zeitfenster geschehen (Vehicle Routing Problem with Time Window, VRPTW). Diese Zeitfenster steuern auf der einen Seite, dass die Fahrgäste ihren Anschlusszug rechtzeitig erreichen können und auf der anderen Seite, dass die Gesamtdauer des Transports für den jeweiligen Gast nicht zu groß wird (dann könnte auch der bisherige ÖPNV genutzt werden). Um eine kurzfristige Buchung des On-Demand-Services zu ermöglichen, müssen im späteren Verlauf dynamisch weitere Fahrgäste und Transporte zu bestehenden Routen hinzugefügt werden können (Dynamic Vehicle Routing Problem, DVRP). Auf Lösungsansätze für das VRP und seine Varianten wird in Abschnitt 4 eingegangen.

Für die Unterstützung des bestehenden ÖPNV mittels eines On-Demand-Services ist es nicht nur von Nöten, geeignete einzelne Verfahren zum Finden kürzester Wege und optimierter Routen zu finden, sondern auch diese geeignet zu kombinieren. So besteht zum Beispiel die Möglichkeit von Spremberg nach Bauzen mittels Bussen (Linie 800 und 500) oder Zügen (RB 65 und RB60) zu gelangen. Dem Fahrgast ist dabei die Art der Verbindung egal, was bedeutet, dass er entweder zu einer der Bushaltestellen oder zum Bahnhof gebracht werden muss, ehe er mit dem ÖPNV weiterfahren kann. Die Tourenplanung (im Folgenden auch "Routing") soll dabei gewährleisten, dass der Fahrgast so schnell wie möglich und rechtzeitig zu einem der Haltepunkte gebracht wird, wohingegen ein Shortest Path-Algorithmus zunächst die infrage kommenden Haltepunkte ermitteln muss. Wie die beiden Verfahren kombiniert werden können, wird in Abschnitt 5 diskutiert.

3 Shortest Path-Algorithmen

Das Shortest Path-Problem ist ein sehr gut erforschtes und in verschiedene Unterkategorien differenzierbares Problem. In [Ma17] sind diverse Unterkategorien des Shortest Path-Problems aufgeführt. Die Problemstellung des statischen Ermitteln der schnellsten Bus- oder Bahnverbindung fällt in die Kategorie (static) Single Source Shortest Path (SSSP). In diesem Abschnitt werden verschiedene Shortest Path-Algorithmen zur Lösung des SSSP-Problems vorgestellt und miteinander verglichen.

Der Dijkstra-Algorithmus Der Dijkstra-Algorithmus [Co01] ist ein bekannter Algorithmus zur Berechnung des kürzesten Pfades in einem gewichteten Graphen mit nicht-negativen Kantengewichten. Der Algorithmus arbeitet inkrementell und wählt iterativ den Knoten mit dem geringsten Abstand vom Startknoten aus. Wurde der Knoten mit dem geringsten Abstand ausgewählt, so werden die Abstände aller Knoten, die noch nicht besucht wurden und mit dem aktuellen Knoten verbunden sind, angepasst und der aktuelle Knoten als besucht markiert. Der Abstand eines markierten Knotens kann sich nicht mehr verringern

und ist somit der kürzeste Abstand vom Startknoten zu diesem Knoten. Wird der Abstand zu einem bestimmten Knoten gesucht, so bricht der Algorithmus ab, sobald der gewünschte Knoten besucht wurde.

Der Dijkstra-Algorithmus ist exakt und hat eine Zeitkomplexität von $O(n^2)$, wobei n die Anzahl der Knoten ist. Ein Vorteil des Algorithmus ist, dass er nicht alle Kanten untersuchen muss. Dies ist besonders nützlich, wenn die Gewichte an einigen Kanten teuer sind. Ein Nachteil besteht darin, dass der Algorithmus nur mit nicht-negativ gewichteten Kanten umgehen kann und er nur für statische Graphen anwendbar ist.

Fredman und Tarjan [FT87] verbesserten den Algorithmus von Dijkstra, indem sie einen Fibonacci-Heap (F-Heap) verwenden. Diese Implementierung erreicht eine Laufzeit von $O(n * \log(n) + m)$ (n = Anzahl Knoten, m = Anzahl Kanten), da die gesamte anfallende Zeit für die Heap-Operationen $O(n * \log(n) + m)$ ist und die anderen Operationen $O(n + m)$ kosten. Fredman und Willard [FW90; FW93; FW94] konnten eine Erweiterung unter Verwendung von AF-Heaps (allocation free Heaps) entwickeln, die lediglich einen Aufwand von $O(m + n * \log(n) / \log(\log(n)))$ hat. Der AF-Heap liefert konstante amortisierte Kosten für die meisten Heap-Operationen und $O(\log(n) / \log(\log(n)))$ amortisierte Kosten für das Löschen.

Der Dijkstra-Algorithmus kann sowohl vom Startknoten als auch vom Endknoten aus gestartet werden. Der *bidirektionale Dijkstra-Algorithmus* [Dr69] verbindet beide Ansätze und führt jeweils einen Schritt vorwärtsgehend vom Startknoten und rückwärtsgehend vom Zielknoten aus. Sobald die beiden Vorgehen einen Knoten als besucht von beiden Richtungen ermittelt haben, wurde ein kürzester Pfad zwischen Start- und Zielknoten ermittelt. Ein Vorteil des bidirektionalen Vorgehens ist, dass der Suchraum dadurch verkleinert werden kann.

Der A*-Algorithmus Der A*-Algorithmus [HNR68] ist eine Verallgemeinerung des Dijkstra-Algorithmus, der diesen um eine Kostenschätzfunktion erweitert. Eine gute Kostenschätzfunktion kann die Menge der Knoten, die untersucht werden müssen, bevor die Lösung gefunden wird und somit den Suchraum, signifikant verkleinern. Es wird dabei in jedem Schritt der Knoten als nächster ausgewählt, dessen geschätzte Kosten zum Zielknoten zusammen mit den bisherigen Kosten am geringsten sind. Wenn die verwendete Kostenschätzfunktion die tatsächlichen Kosten nie überschätzt, so ist der A*-Algorithmus optimal, das heißt er findet immer die Lösung mit den niedrigsten Kosten.

Der A*-Algorithmus hat analog zur Breitensuche das Problem, dass viele Knoten gespeichert werden müssen und die Knoten darüber hinaus auch noch nach ihrem Kostenschätzwert sortiert sein müssen. Um diese Probleme zu reduzieren, kann innerhalb der A*-Suche das Vorgehen des iterativen Vertiefens (iterative deepening) angewendet werden. Das iterative Vertiefen entspricht einer Tiefensuche mit festgelegter Tiefenschranke n . Findet die Tiefensuche keine Lösung bis Schranke n , so wird eine neue Tiefensuche mit höherem

n -Wert gestartet. Das Vorgehen des iterativen Vertiefens ist sowohl vollständig als auch optimal und benötigt nur die Rechenzeit der Breitensuche (b^d) und den Speicherbedarf der Tiefensuche ($b * d$, b = Verzweigungsgrad, d = durchsuchte Tiefe des Baumes).

Der Arcflag-Algorithmus Der *Arcflag*-Algorithmus [Hi06] ist eine zielgerichtete Beschleunigungstechnik für den Dijkstra-Algorithmus zur Suche des kürzesten Pfades zwischen zwei Knoten in einem gewichteten Graphen. Die Grundidee besteht darin, die Menge der zu betrachtenden Kanten geschickt auf einen Bruchteil im Vergleich derer zu verringern, welche bei Ausführung des Dijkstra-Algorithmus betrachtet werden müssten. Dabei wird zunächst jede Kante des Graphen um Flaggeninformationen angereichert, welche schließlich bei der Pfadsuche entscheiden, ob diese Kante für die Suche in Betracht gezogen werden muss. Die Vorberechnungsphase des Arcflags-Algorithmus verläuft in zwei Schritten. Zuerst wird das Straßennetz in Regionen eingeteilt. Anschließend wird für jede Region bestimmt, über welche Kanten kürzeste Wege in diese Region führen. Die Information, ob eine Kante Teil eines kürzesten Pfades ist, wird Arcflag genannt.

In der Anfragephase bestimmt der Algorithmus zunächst die Region des Zielknotens. Anschließend wendet er den Dijkstra-Algorithmus an, folgt dabei aber nur den Kanten, die laut Zusatzinformationen in die Zielregion führen. Er lässt also gezielt Kanten aus, die nichts mit der Zielanfrage zu tun haben. Die Komplexität des Dijkstra-Algorithmus kann durch das Verwenden von Arcflags nicht reduziert werden, nichtsdestotrotz verringert sich in der Praxis der Suchraum durch das Weglassen von Kanten, die keine kürzesten Wege zur Zielregion darstellen, signifikant.

Der ALT-Algorithmus ALT steht für *A* Search*, *Landmarks* und *Triangle Inequality*, da dies die Hauptbestandteile des Algorithmus sind. Der ALT-Algorithmus verwendet als Kostenschätzfunktion für den A*-Algorithmus sogenannte Landmarken [GH05]. Landmarken sind eine Teilmenge der Knoten des Graphen, für die die Distanzen aller anderen Knoten zu diesen Landmarken aus der Vorverarbeitungsphase bekannt sind. In der Folge können Entfernungsschätzungen innerhalb des Graphen mittels der Dreiecksungleichung durchgeführt werden. Die Dreiecksungleichung besagt, dass sich für beliebige drei Knoten u , v , l im Graphen der minimale Abstand von u und v kleiner oder gleich dem Abstand von u zu l plus dem Abstand von l zu v sein muss: $d(u, v) \leq d(u, l) + d(l, v)$. Aus dieser Ungleichung lassen sich Grenzen ableiten: Wie in Abbildung 1 dargestellt, ist der Abstand von u nach v kleiner als $dist(l_1, v) - dist(l_1, u)$ sowie $dist(u, l_2) - dist(v, l_2)$. Eine Frage die verbleibt ist: Wie müssen geeignete Orientierungspunkte aus dem Eingabegraphen ausgewählt werden?

Zusammenfassung Shortest Path-Algorithmen. Zusammenfassend lässt sich sagen, dass das Shortest Path-Problem für den statischen ÖPNV mittels Dijkstra-Algorithmus bzw. Weiterentwicklungen davon lösen lässt. Die Eingangsbedingung, dass keine negativen

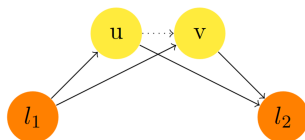


Abb. 1: Eine Visualisierung der Dreiecksungleichung (Triangle Inequality).

Kantengewichte vorliegen, ist sowohl für das Berechnen kürzester Wege als auch für das Berechnen preisgünstigster Wege erfüllt. Nach einer Studie von [Ba22] ist bereits der Dijkstra-Algorithmus in seiner Grundform in der Lage europaweite SSSP-Anfragen in ca. 6s zu beantworten. Die vorgestellten weiterentwickelten Formen benötigten dafür teilweise sogar nur wenige Millisekunden. Eine Anfrage für ländliche Regionen mit deren abgehenden und eingehenden ÖPNV-Verbindungen sollte somit mit allen vorgestellten Algorithmen realisierbar sein.

4 Vehicle Routing-Algorithmen

Ebenso wie das Shortest Path-Problem, ist auch das Vehicle Routing-Problem sehr gut erforscht. Es existiert eine Vielzahl an Algorithmen, die das Problem entweder zeitaufwändig exakt oder schnell heuristisch lösen. Abbildung 2 klassifiziert existierende Lösungsverfahren für Vehicle Routing-Probleme nach Korrektheit und heuristischem Vorgehen. An dieser Stelle soll aufgrund der großen Vielfalt an existierenden Algorithmen nur auf je ein vielversprechendes exaktes und ein heuristisches Verfahren eingegangen werden, die verwendet werden können um das VRP für den On-Demand-Service zu lösen.

4.1 Ein Constraint-Ansatz (Branch and bound)

Die Constraint-Programmierung ist ein mächtiges Werkzeug zum Modellieren und Lösen komplexer Probleme [MS98]. Das generelle Vorgehen bei der Constraint-Programmierung unterteilt sich in zwei Teile: 1. die deklarative Modellierung des Problems als Constraint-Modell, 2. das selbstständige Lösen des Constraint-Modells durch einen Constraint-Solver. Der Constraint-Anwender ist dabei nur dafür verantwortlich, das Constraint-Modell zu erzeugen und den Solver zu konfigurieren. Der Solver selbst fungiert als Blackbox und reduziert schrittweise durch Propagation (Branch and Bound) und einer Backtracking-basierten Tiefensuche den Lösungsraum solange, bis eine (optimale) Lösung gefunden wurde (falls eine solche existiert).

Für das sich aus unserer Anwendung ergebene VRP inklusive der spezifizierten Subklassen Capacitated Vehicle Routing Problem (CVRP) und Vehicle Routing Problem with Time Window (VRPTW) sind in [Zh22] bereits mathematische Beschreibungen angegeben, die sich prinzipiell direkt in Constraints überführen lassen. Einziger Stolperstein, der bei der

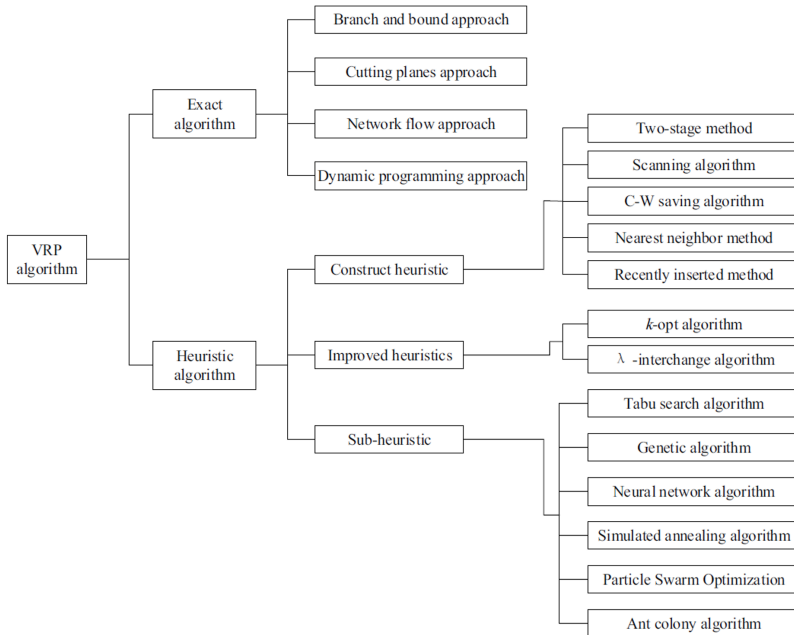


Abb. 2: Existierende Algorithmen zum Lösen von Vehicle Routing-Problemen [Zh22].

Verwendung der Constraint-Programmierung, wie bei jedem anderen exakten Verfahren auch, auftreten kann, ist dass das Problem und damit verbunden dessen Suchraum zu groß wird, als dass in akzeptabler Zeit eine gute oder sogar eine optimale Lösung gefunden werden kann. Da hier allerdings Graphen betrachtet werden, die den ländlichen Raum widerspiegeln (und keine Großstädte), sollten deren Ausmaße handhabbar sein. Sollte die Problemstellung wieder Erwartung trotz dem zu groß sein, um einen Constraint-Solver anzuwenden, so wird auf ein heuristisches Verfahren umgestellt.

4.2 Ein Zwei-Phasen-Ansatz (Two-stage method)

Der Zwei-Phasen-Ansatz probiert in Phase 1 eine initiale Lösung für das Routing-Problem zu finden und in Phase 2 die bestehende Lösung sukzessive zu verbessern [Wa14].

Die erste Phase wird dabei oftmals erneut in zwei Phasen unterteilt. In Phase 1.1 wird dabei zunächst ein Clustering C_1, \dots, C_n der abzufahrenden Knotenpunkte angelegt. In Phase 1.2 wird dann eine initiale Tour innerhalb jedes Clusters von einem Startknoten (Depot) hin über alle Knoten und wieder zum Startknoten zurück angelegt.

Mittels verschiedener lokaler Optimierungen wird in Phase 2 dann versucht die initiale Lösung zu verbessern. Mögliche Optimierungsoperationen sind dabei z.B. *Relocate*-, die

Relocate Split, die *2-opt* Exchange* und die *Swap-Move-Operation*. Bei der *Relocate-Operation* wird die Position eines Knoten innerhalb einer Route geändert. Ein Knoten wird also von seiner aktuellen Position innerhalb der Route zu einer anderen Position in der Route verschoben. Bei der *Relocate Split-Operation* wird ein Knoten von einer Route in eine andere Route verschoben. Die *2-opt* Exchange-Operation* verändert die Reihenfolge von zwei Knoten innerhalb einer Route. Es werden dabei zunächst zwei Knoten ausgewählt, und dann wird die Route zwischen diesen beiden Knoten umgekehrt. Die *Swap-Move-Operation* tauscht die Positionen zweier Knoten innerhalb einer Route.

Im Gegensatz zu den exakten Verfahren handelt es sich bei diesem Vorgehen um eine lokale Suche, die ein lokales Optimum findet. Ob das gefundene lokale Optimum nah am globalen Optimum ist oder dieses sogar erreicht, kann an dieser Stelle nicht garantiert werden.

5 Kombinationsmöglichkeiten

Weil der On-Demand-Service als Zubringer für den weiteren ÖPNV dienen soll, müssen die beiden vorgestellten Problemstellungen SSSP und VRP kombiniert werden. Ein Fahrgast möchte zum Beispiel von Spremberg nach Schwarze Pumpe fahren. Ihm ist es dabei egal, ob er mit dem On-Demand-Service direkt von Zuhause zum Zielort gebracht wird, oder ob er mit dem On-Demand-Service zu einer der regulären Haltestellen der entsprechenden Verbindung innerhalb von Spremberg gebracht wird und von dort aus mit dem ÖPNV weiterfahren kann. Für den Fahrgast ist dabei in der Regel nur die benötigte Zeit von Start bis Ziel interessant. Für die Routenberechnung macht es allerdings einen erheblichen Unterschied, zu welchen Bahnhof der Passagier gebracht werden muss. Befinden sich mehrere solcher Fahrgäste, mit flexiblen Zielen, gleichzeitig in einer Tour, so erschwert das die Tourenplanung zusätzlich erheblich. Nachfolgend werden zwei Vorgehen erläutert, mit solchen Alternativen umzugehen.

- **Single-SSP-VRP:** Es wird zunächst der statische ÖPNV (SSSP) geplant und im Anschluss daran der On-Demand-Service (VRP).
- **Alternative-VRP:** Das VRP wird um Alternativen in der Routenplanung erweitert. Es ergeben sich somit für jeden Fahrgast verschiedene mögliche Start- und Zielorte und -Zeiten.

Single-SSP-VRP In diesem Fall würden zunächst die ÖPNV-Verbindungen der einzelnen Fahrgäste ermittelt werden. Diese würden mit den vorgestellten Algorithmen (siehe Abschnitt 3) berechnet werden. Dabei würde als Start und Ziel jeweils die für den einzelnen Reisenden passendste Haltestelle mit zugehöriger Start- und Ankunftszeit gewählt werden. Dieses Vorgehen würde für alle Reisenden analog erfolgen, so dass alle eine minimale ÖPNV-Reisedauer haben. Im Anschluss daran wird der On-Demand-Service für die zuvor ermittelten Start- und Ankunftsorte und -Zeiten ermittelt. Bei diesem Vorgehen handelt es sich um einen

greedy-Algorithmus, bei dem versucht wird, die lokalen Optima (die ÖPNV-Verbindungen für die einzelnen Fahrgäste) zu einer Gesamtlösung (durch den On-Demand-Service) zu vereinen. Positiv daran ist, dass das VRP dabei so klein wie möglich gehalten wird und auch die einzelnen ÖPNV-Verbindungen nur einmal für jeden Fahrgast mittels SSSP-Algorithmen gelöst werden müssen. Dieses Vorgehen kann sehr schnell eine erste, gute Lösung ermitteln, allerdings kann nicht sichergestellt werden, dass ein globales Optimum gefunden wird. Denkbar ist z.B., dass durch das Ändern des Zielortes einer oder mehrerer ÖPNV-Anbindungen (Wahl eine früheren oder späteren Haltestelle) die Tourenplanung ggf. deutlich bessere Ergebnisse erlaubt.

Eine Erweiterungsmöglichkeit besteht darin, zunächst alle ÖPNV-Anbindungen, die für die einzelnen Fahrgäste infrage kommen, zu berechnen. Das beinhaltet also nicht nur die Haltestelle, die am nächsten am Start- bzw. Zielort ist, sondern auch die, die sich in einer akzeptablen Umgebung für den On-Demand-Service befinden. Anschließend wird für jeden Fahrgast eine Verbindung ausgewählt und das Routing durchgeführt. Nachfolgend wird eine andere Kombination an Fahrgastverbindungen ausgesucht und dafür ein Routing durchgeführt. Sobald ersichtlich ist, dass ein VRP keine bessere Lösung finden kann, kann die Berechnung abgebrochen und die nächste Kombination ausprobiert werden. Mit diesem Vorgehen können globale Optima ermittelt werden, allerdings steigt der Rechenaufwand exponentiell an.

Alternative-VRP Für jeden Fahrgast werden zunächst alle infrage kommenden ÖPNV-Anbindungen ermittelt (SSSP). Anders als zuvor wird nun das VRP nicht auf einer möglichen Kombination dieser ÖPNV-Anbindungen aufgerufen sondern um die möglichen Alternativen erweitert. Das VRP muss dann entweder dafür sorgen, dass Fahrgast F_1 rechtzeitig von Ort A zu Ort B oder zu Ort C gebracht wird. Es ergeben sich also Alternativen für die Start- bzw. Zielorte und -Zeiten. Bei der Constraint-Programmierung können solche Alternativen durch logische Meta-Constraints (*and*, *or* oder *not* die jeweils wieder Constraints als Eingaben erhalten) problemlos eingebaut werden. Auch hier wächst das Problem mit zunehmender Anzahl an Alternativen exponentiell, allerdings können durch den branch and bound-Ansatz deutlich eher solche ÖPNV-Anbindungen ermittelt und ausgeschlossen werden, die in keiner guten Lösung vorkommen können. Im Vergleich zum zuerst vorgestellten Vorgehen Single-SSP-VRP besteht hierbei die Gefahr, dass das Finden einer ersten Lösung länger dauert. Da wir uns aber im ländlichen Raum mit wenigen Alternativen im ÖPNV befinden, scheint eine Umsetzung mittels Constraint-Programmierung realistisch.

6 Zusammenfassung und Ausblick

In diesem Papier wurde das Problem des On-Demand-Verkehrs als Unterstützung für den ÖPNV im ländlichen Bereich diskutiert. Zunächst wurden die Besonderheiten des ländlichen ÖPNV und speziell der Modellstadt Spremberg erörtert, ehe sowohl Verfahren zum Finden

kürzester Wege als auch zur Routenplanung diskutiert wurden. Abschließend wurden zwei Möglichkeiten zur Kombination aus statischer ÖPNV-Planung (SSSP) und flexibler Planung für den On-Demand-Service (VRP) erörtert und Vor- und Nachteile der Verfahren diskutiert.

Im Zukunft müssen die vorgestellten Verfahren für die gegebene Datenlage der Modellstadt Spremberg umgesetzt und praktisch erprobt werden. Wir gehen an dieser Stelle davon aus, dass wir mindestens in der Lage sind mit heuristischen Verfahren eine gute, wenn nicht sogar mit einem exakten Verfahren eine optimale Lösung zu finden.

Danksagung: Das vorgestellte OSLO-Projekt wurde mit Mitteln des „mFUND - Unsere Förderung für die Mobilität der Zukunft“ [23b] (Förderkennzeichen: 19FS1005C) des Bundesministeriums für Digitales und Verkehr gefördert.

Literatur

- [21a] Bevölkerungsentwicklung und Flächen der kreisfreien Städte, Landkreise und Gemeinden im Land Brandenburg 2021, Amt für Statistik Berlin-Brandenburg. Letzmal besucht am 2. Mai 2023, 2021, URL: https://download.statistik-berlin-brandenburg.de/0449b1ed6918ad8c/25590f3b9567/SB_A01-04-00_2021j01_BB.pdf.
- [21b] Linienverzeichnis, Seh-Netz e.V. 2000 — 2023 ÖPNV-Info — Mobilitätsportal für behinderte Reisende. Letzmal besucht am 2. Mai 2023, 2021, URL: <https://www.oepnv-info.de/linienverzeichnis>.
- [23a] bbnavi, DigitalAgentur Brandenburg GmbH, Gefördert durch das Ministerium für Wirtschaft, Arbeit und Energie des Landes Brandenburg. Letzmal besucht am 2. Mai 2023, 2023, URL: <https://bbnavi.de/>.
- [23b] mFUND – Unsere Förderung für die Mobilität der Zukunft, Bundesministerium für Digitales und Verkehr. Letzmal besucht am 8. Mai 2023, 23, URL: <https://bmdv.bund.de/DE/Themen/Digitales/mFund/Ueberblick/ueberblick.html>.
- [Ba22] Bayer, S.; Kluge, R.; Kohl, A.; Messmer, S.: Studentisches Projekt: Gläserner Routenplaner, Betreuer: M. Kobitzsch, S. Meinert, I. Rutter, Prof. P. Sanders, Prof. D. Wagner. KIT - Karlsruhe Institute of Technology. Letzmal besucht am 5. Mai 2023, 2022, URL: https://i11www.iti.kit.edu/_media/projects/rpkit/techniken_poster.pdf.
- [Co01] Cormen, T. H.; Leiserson, C. E.; Rivest, R. L.; Stein, C.: Introduction to Algorithms. The MIT Press, 2001, ISBN: 0262032937.
- [Dr69] Dreyfus, S. E.: An Appraisal of Some Shortest-Path Algorithms. Oper. Res. 17/3, S. 395–412, 1969, URL: <https://doi.org/10.1287/opre.17.3.395>.
- [FT87] Fredman, M. L.; Tarjan, R. E.: Fibonacci heaps and their uses in improved network optimization algorithms. J. ACM 34/3, S. 596–615, 1987, URL: <https://doi.org/10.1145/28869.28874>.

- [FW90] Fredman, M. L.; Willard, D. E.: BLASTING through the Information Theoretic Barrier with FUSION TREES. In (Ortiz, H., Hrsg.): Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA. ACM, S. 1–7, 1990, URL: <https://doi.org/10.1145/100216.100217>.
- [FW93] Fredman, M. L.; Willard, D. E.: Surpassing the Information Theoretic Bound with Fusion Trees. *J. Comput. Syst. Sci.* 47/3, S. 424–436, 1993, URL: [https://doi.org/10.1016/0022-0000\(93\)90040-4](https://doi.org/10.1016/0022-0000(93)90040-4).
- [FW94] Fredman, M. L.; Willard, D. E.: Trans-Dichotomous Algorithms for Minimum Spanning Trees and Shortest Paths. *J. Comput. Syst. Sci.* 48/3, S. 533–551, 1994, URL: [https://doi.org/10.1016/S0022-0000\(05\)80064-9](https://doi.org/10.1016/S0022-0000(05)80064-9).
- [GH05] Goldberg, A. V.; Harrelson, C.: Computing the shortest path: A* search meets graph theory. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, British Columbia, Canada, January 23-25, 2005. SIAM, S. 156–165, 2005, URL: <http://dl.acm.org/citation.cfm?id=1070432.1070455>.
- [Hi06] Hilger, M.; Köhler, E.; Möhring, R. H.; Schilling, H.: Fast Point-to-Point Shortest Path Computations with Arc-Flags. In (Demetrescu, C.; Goldberg, A. V.; Johnson, D. S., Hrsg.): The Shortest Path Problem, Proceedings of a DIMACS Workshop, Piscataway, New Jersey, USA, November, 2006. Bd. 74. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, DIMACS/AMS, S. 41–72, 2006, URL: <https://doi.org/10.1090/dimacs/074/03>.
- [HNR68] Hart, P. E.; Nilsson, N. J.; Raphael, B.: A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Trans. Syst. Sci. Cybern.* 4/2, S. 100–107, 1968, URL: <https://doi.org/10.1109/TSSC.1968.300136>.
- [KV12] Korte, B.; Vygen, J.: Combinatorial Optimization: Theory and Algorithms. Springer Publishing Company, Incorporated, 2012, ISBN: 3642244874.
- [Ma17] Madkour, A.; Aref, W. G.; Rehman, F. U.; Rahman, M. A.; Basalamah, S. M.: A Survey of Shortest-Path Algorithms. *CoRR* abs/1705.02044/, 2017, arXiv: 1705.02044, URL: <http://arxiv.org/abs/1705.02044>.
- [MS98] Marriott, K.; Stuckey, P. J.: Programming with Constraints - An Introduction. MIT Press, Cambridge, 1998, ISBN: 978-0-262-13341-8.
- [TV02] Toth, P.; Vigo, D.: The Vehicle Routing Problem. Society for Industrial and Applied Mathematics, 2002, ISBN: 0-89871-579-2.
- [Wa14] Wang, Y.; Ma, X.; Lao, Y.; Yu, H.; Liu, Y.: A two-stage heuristic method for vehicle routing problem with split deliveries and pickups. *J. Zhejiang Univ. Sci. C* 15/3, S. 200–210, 2014, URL: <https://doi.org/10.1631/jzus.C1300177>.
- [Zh22] Zhang, H.; Ge, H.; Yang, J.; Tong, Y.: Review of Vehicle Routing Problems: Models, Classification and Solving Algorithms. *Arch Computat Methods Eng* 29/, S. 195–221, 2022.

DASU – Transferzentrum für Digitalisierung, Analytics & Data Science Ulm

Intelligente Transferstrategien zur Unterstützung der digitalen Transformation

Reinhold von Schwerin¹ und Alexander Hafner²

Abstract: Dieser Beitrag beleuchtet die Rolle des DASU - Transferzentrum für Digitalisierung, Analytics & Data Science Ulm, einer innovativen Einrichtung, die sich der Förderung der digitalen Transformation und des Wissenstransfers in der Region Ulm und darüber hinaus widmet. Angesichts der vielfältig vorhandenen wirtschaftlichen Potenziale der Bereiche Digitalisierung, Analytics und Data Science, ist das DASU entschlossen, eine Brücke zwischen Wissenschaft, Wirtschaft und Gesellschaft zu schlagen. Der Beitrag beschreibt die Organisationsstruktur des DASU, seine multidirektionalen Transferansätze und seine Rolle als zentrale Anlaufstelle für Unternehmen, insbes. KMU, und Einrichtungen, die eine Transformation zu einer daten- und wissensbasierten Wertschöpfung anstreben. Die Inhalte dieses Beitrags sollen dazu dienen, angereichert durch Fallbeispiele, das Verständnis für die mögliche Rolle solcher intelligenten Institutionen bei der Förderung der digitalen Transformation und der nachhaltigen Entwicklung zu vertiefen und eine Grundlage für weitere Diskussionen zur Notwendigkeit derartiger Einrichtungen zu schaffen.

Keywords: Digitalisierung, Digitale Transformation, Unternehmensinnovation, Industrie 4.0, Analytics & Data Science, Maschinelles Lernen, Künstliche Intelligenz

1 Einleitung

Unternehmen stehen vor wachsenden Herausforderungen, Anforderungen und Fragestellungen in den Bereichen der Digitalisierung, Datenanalyse und Datenwissenschaft, insbesondere dann, wenn es um den Einsatz des maschinellen Lernens und der künstlichen Intelligenz in nachhaltigen, soziotechnischen Unternehmensanwendungen geht. Die Gefahr, hierbei den Anschluss an die internationalen Mitbewerber zu verlieren, ist – insbesondere für kleine und mittlere Unternehmen (KMU) – sehr groß.

Das Transferzentrum für Digitalisierung, Analytics & Data Science Ulm (DASU) ist ein einzigartiges Projekt, welches durch sein besonderes Konzept sowie durch ein Netzwerk unterschiedlichster Akteure Digitalisierung in Wirtschaft, Industrie und Gesellschaft zielgenau ermöglicht. Das multidisziplinäre Transferzentrum, welches in Zusammenarbeit mit der Universität Ulm und der Technischen Hochschule Ulm gegründet wurde, setzt sich zum Ziel, eine zentrale Anlaufstelle bei sämtlichen Fragen der Digitalisierung,

¹ Technische Hochschule Ulm, Albert-Einstein-Allee 53-55, 89081 Ulm, reinhold.vonschwerin@thu.de

² DASU – Transferzentrum für Digitalisierung, Analytics & Data Science Ulm, Ensingerstr. 4, 89073 Ulm, alexander.hafner@dasu.digital

Datenwissenschaft und Datenanalyse zu sein. Durch eine gezielte Bündelung von Kompetenzen aus unterschiedlichen Forschungsbereichen strebt das DASU an, Wissen und Technologien systematisch zu vernetzen und maßgeschneidert in Anwendungsszenarien zu überführen. Im Kern der Aktivitäten stehen hierbei die Entwicklung und Umsetzung von Forschungs- und Transferprojekten in enger Kooperation mit der Wirtschaft. Darüber hinaus baut das DASU ein Digitalisierungslabor auf, das als Plattform für die Erprobung und Visualisierung von Digitalisierungslösungen und datenbasierten Geschäftsmodellen dient. Zudem runden umfangreiche Weiterbildungsangebote sowie ein multidisziplinärer Pool an Experten das Portfolio ab.

All diese Initiativen sind von wesentlicher Bedeutung, um insbesondere KMU bei der Transformation zu einer daten- und wissensbasierten Wertschöpfung zu unterstützen und ihre Wettbewerbsfähigkeit zu stärken. Sie dienen dazu, Hürden im Transformationsprozess abzubauen und das Bewusstsein für die Möglichkeiten der Digitalisierung in allen Ebenen der Gesellschaft zu schärfen.

Dieser Beitrag bietet einen detaillierten Einblick in das DASU [DA21], seine Struktur, seine Aktivitäten und seine Bestrebungen zur intelligenten Digitalisierung. Es wird diskutiert, wie das DASU gezielt Netzwerke aufbaut, um seine Ziele zu erreichen und wie es dazu dient, den Übergang von Unternehmen in den digitalen und datengetriebenen Wettbewerb erfolgreich zu meistern.

2 Das DASU als intelligente Organisation

Die Wirtschaftsregion des Schwabenbunds, bestehend aus den Landkreisen Alb-Donau, Biberach, Heidenheim sowie dem Stadtkreis Ulm, ist geprägt durch eine mittelständische Unternehmenslandschaft, die angesichts internationaler Wettbewerber mit hoher IT-Kompetenz zunehmend unter Anpassungsdruck gerät. Um künftig nicht Wertschöpfungsanteile zu verlieren und eine Entwicklung der zahlreichen Industriebetriebe zu reinen Hardware-Lieferanten zu vermeiden, ist es wichtig, möglichst schnell eigenständiges Wissen aus Daten generieren zu können und neue Wertschöpfungsketten zu etablieren.

Für diese erforderliche Disruption bedarf es jedoch fundierter Transformationshilfen. Denn um die sich bietenden Chancen in den Bereichen der Digitalisierung, Künstlichen Intelligenz (KI), Datenanalyse und Datenwissenschaft in ihrer ganzen Breite auszuschöpfen, müssen deren Methoden, Technologien und Werkzeuge sachgerecht und zielgerichtet angewendet werden. Dazu müssen die Unternehmen die Potentiale in ihrem jeweiligen Geschäftsfeld zunächst einmal erkennen und dann die Bereitschaft für Veränderungen aufbringen. Insbesondere KMU, die sich keine spezialisierten Abteilungen für Digitalisierung, Data Science und Data Analytics leisten können, muss eine zentrale Anlaufstelle geboten werden, welche spezifische Angebote vorhält, um den beschriebenen Transformationsprozess zu begleiten.

Mit dem Ziel, genau diese Anlaufstelle zu schaffen, wurde das DASU 2021 als gemeinnützige Stiftung bürgerlichen Rechts gegründet. Wesentliche Initiatoren waren hierbei die IHK Ulm sowie die Stadt Ulm. Die Stiftung mit derzeit zehn weiteren Stiftern wird zudem ergänzt um das Virtuelle Zentrum (VZ), einer gemeinsamen wissenschaftlichen Einrichtung der Universität Ulm (UUI) und der Technischen Hochschule Ulm (THU). Durch Unterstützung der Wirtschaft, des öffentlichen Sektors und der Forschung wurde somit eine Institution geschaffen, welche intelligent und zielgerichtet Digitalisierungsprojekte vorantreiben und umsetzen kann.

2.1 Organisationsübersicht

Das Transferzentrum DASU gliedert sich in zwei unterschiedliche Organisationseinheiten, welche gemeinsam die zuvor beschriebene Zielsetzung unterstützen und umsetzen. Abbildung 1 zeigt schematisch den Aufbau des DASU mit den jeweils zugehörigen Institutionen und Akteuren.

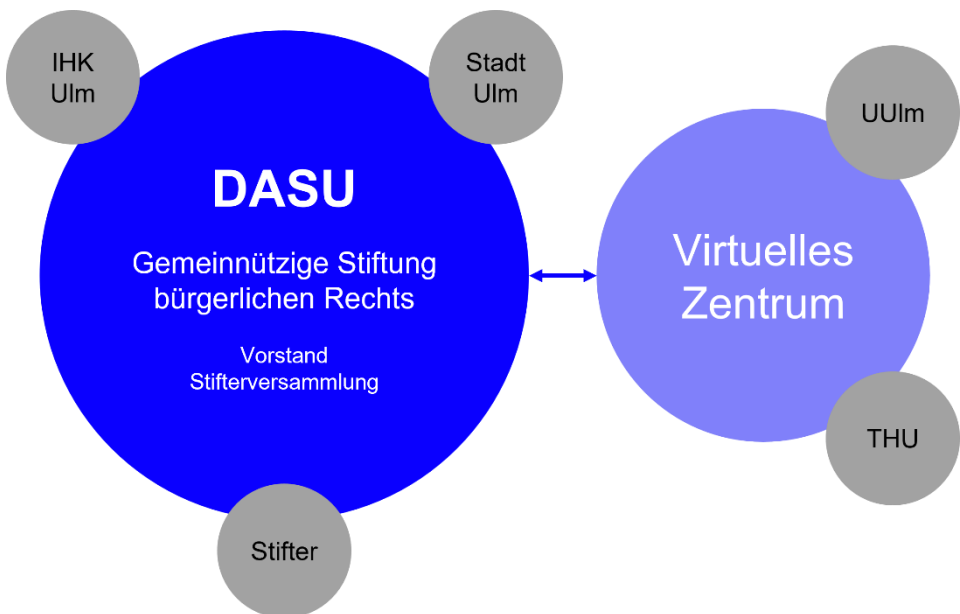


Abb. 1: Schematische Übersicht der DASU-Organisationsstruktur

Im Kern des Transferzentrums steht die 2021 gegründete Stiftung bürgerlichen Rechts. Sie wurde vom Regierungspräsidium Tübingen am 17.08.2021 als rechtsfähige Stiftung gemäß § 80 BGB i.V.m. § 5 StiftG anerkannt. Die wesentlichen Initiatoren des Vorhabens sind die IHK Ulm und die Stadt Ulm, welche das Grundstockvermögen der Stiftung zur

Verfügung stellen. Darüber hinaus ist es gelungen, eine Vielzahl von regionalen Unternehmen und Einrichtungen von der Notwendigkeit des Projekts zu überzeugen und als Stifter zu gewinnen. Zum aktuellen Zeitpunkt (Stand: 18.06.2023) stifteten die folgenden Unternehmen in das Verbrauchsvermögen der Stiftung:

- Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach
- CarByte GmbH, Stuttgart
- Hensoldt Sensors GmbH, Ulm
- Kreissparkasse Biberach, Biberach
- Liebherr-Digital Development Center GmbH, Ulm
- RAM Stiftung, Ulm
- Rentschler Biopharma SE, Laupheim
- Sparkasse Ulm, Ulm
- SWU Stadtwerke Ulm/Neu-Ulm GmbH, Ulm
- Wieland-Werke AG, Ulm

Die breite Unterstützung des DASU aus unterschiedlichsten Wirtschaftsbereichen zeigt eindrucksvoll, dass die Etablierung eines Transferzentrums, welches Ressourcen bündelt und gezielt Netzwerke zwischen Wissenschaft und Wirtschaft stärkt, auf hohe Nachfrage stößt. Erst durch diesen intelligenten, branchenübergreifenden Zusammenschluss lassen sich die Herausforderungen der digitalen Transformation in Unternehmen aller Größe meistern; schlussendlich profitieren von dieser Zustiftung auch kleine Betriebe, die selbst vielleicht nur ein geringes Budget für Digitalisierung aufbringen können. Die Stifter leisten somit einen wesentlichen gemeinwohlorientierten Beitrag und stärken somit eine gesamte Wirtschaftsregion.

Die wichtigsten Steuerungs- und Entscheidungsgremien des DASU sind die Stiferversammlung und der Stiftungsvorstand. Die Stiferversammlung setzt sich zusammen aus Vertretern der stiftenden Unternehmen, der IHK Ulm und der Stadt Ulm. Im Rahmen der Stiferversammlung wird den Stifern die Möglichkeit gegeben, ein Mitbestimmungsrecht über die Aktivitäten des DASU auszuüben und somit die Digitalisierungsangebote maßgeblich zu lenken.

Der Stiftungsvorstand des DASU besteht aus insgesamt drei Vertreter:innen der Professorien der Universität Ulm und der Technischen Hochschule Ulm und untermauert damit den wissenschaftlich fundierten Transfercharakter der Einrichtung.

Neben der Stiftung wurde das Virtuelle Zentrum (VZ) als gemeinsame wissenschaftliche Einrichtung von UULM und THU gegründet, an der sich derzeit mehr als 50 Professor:innen beteiligen. Die Mitglieder des VZ haben sich der Idee des DASU als

intelligentes Transferzentrum verschrieben und unterstützen als fachliche Experten die Umsetzung von multidisziplinären Digitalisierungsprojekten. Multidisziplinarität bedeutet dabei, dass das DASU ganzheitlich aufgestellt ist, da es alle relevanten wissenschaftlichen Bereiche für Datenanalysen und Data Science in sich vereint. Hierzu zählen nicht nur die traditionellen Disziplinen wie Informatik und Mathematik/ Statistik, sondern auch einschlägige Anwendungsdomänen wie Medizin, Wirtschafts- und Ingenieurwissenschaften sowie zahlreiche weitere Fachgebiete wie Human Factors, Mensch-Technik-Interaktion, (Digitales) Recht und Arbeitspsychologie.

Durch die hier gezeigte Organisationsstruktur wird die Grundlage geschaffen, mit aktuellen, modernen Methoden aus den Bereichen Data Science & Analytics nicht nur die Herausforderungen der digitalen Transformation zu meistern, sondern innovative, neue Methoden der Datenauswertung für interne Prozesse, Produktionsabläufe, Kundenakquise, intelligente Dienstleistungen, vorausschauende Wartungen, etc. zu entwickeln und umzusetzen. Die gebündelten Aktivitäten des DASU versetzen Unternehmen aller Größe zudem in die Lage, vorhandenes Innovationspotenzial zu heben und Weiter- sowie Neuentwicklungen von Waren und Dienstleistungen zu initiieren. Die Mitwirkung der beiden Hochschulen hilft dabei das gesamte Spektrum möglicher datenbasierter Innovationen wissenschaftlich zu betreuen und sowohl mit etablierten Methoden als auch neuesten Erkenntnissen der Forschung zu begleiten. Durch diese ganzheitliche, intelligente Verzahnung unterschiedlicher Wissensträger und Akteure positioniert sich das DASU als Innovations-Enabler für die Bereiche Digitalisierung, Analytics & Data Science.

2.2 Multidirektionaler Transfer

Die Idee eines Transferzentrums unterstellt häufig eine traditionelle Interpretation des Wissenstransfers in unilateraler Art, wobei Erkenntnisse aus der Forschung in die jeweiligen Anwendungsbereiche von Unternehmen überführt werden. Dieser Ansatz kann jedoch eine eingeschränkte Perspektive aufweisen, da er in der Regel eine einseitige Bewegung von Wissen und Ressourcen darstellt, was den Erfolg von Digitalisierungsprojekten maßgeblich gefährdet. Das DASU wiederum implementiert daher eine komplexere, mehrdimensionale und inklusive Form des Wissenstransfers, schematisch dargestellt in Abbildung 2.

Die hier gezeigte Art des Transfers bezieht eine Vielzahl von Akteuren ein, darunter nicht nur die Wissenschaft und die Wirtschaft, sondern auch den öffentlichen Sektor, junge Berufstätige und Studierende. In diesem Modell ist der Austausch von Wissen und Technologie wechselseitig und multidirektional. Es erweitert die traditionelle Konzeption des Wissenstransfers und erkennt die gegenseitige Abhängigkeit und Interaktion zwischen verschiedenen Akteuren an.

Innerhalb des multilateralen Transfers kann beispielsweise die Wirtschaft der Wissenschaft helfen, indem sie aufzeigt, welche Themen für sie relevant sind und woran

geforscht werden sollte. Dies stellt sicher, dass die wissenschaftliche Forschung eng mit den Bedürfnissen und Herausforderungen der Wirtschaft verbunden ist und dass die entwickelten Lösungen praxisrelevant und anwendbar sind.

Gleichzeitig haben der öffentliche Sektor, Städte und Kommunen ihren eigenen Bedarf an Digitalisierung, der von den Bedürfnissen der Wirtschaft abweicht. Durch die Einbeziehung dieser Akteure in den multilateralen Transfer kann das DASU sicherstellen, dass auch ihre Bedürfnisse und Interessen berücksichtigt werden, und so zur Entwicklung von Digitalisierungslösungen beitragen, die auf die spezifischen Herausforderungen des öffentlichen Sektors zugeschnitten sind.

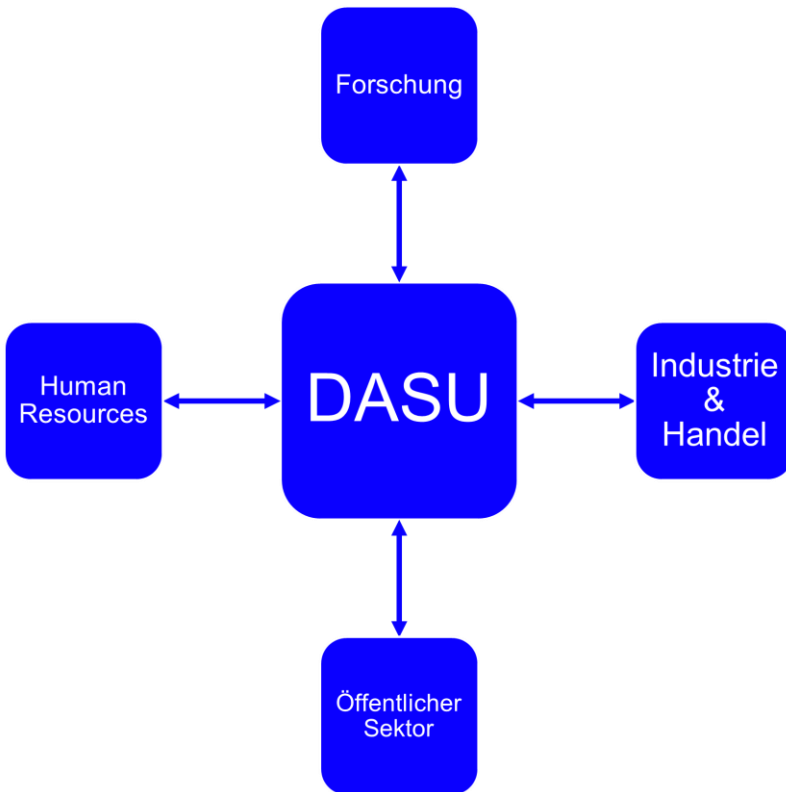


Abb. 2: Übersicht des multilateralen Transfergedankens

Darüber hinaus bietet der multilaterale Transfer Studierenden und jungen Berufstätigen die Möglichkeit, in verschiedene Unternehmen hineinzuschauen und Erfahrungen zu sammeln. Dies ermöglicht ihnen, ihre Fähigkeiten und Kenntnisse in realen Geschäftsumgebungen anzuwenden und zu erweitern, was wiederum zu ihrer beruflichen

Entwicklung und dem Aufbau eines breiten Spektrums an Kompetenzen beiträgt. Auf der anderen Seite profitieren Unternehmen von einem direkten Fachkräftebezug aus den Hochschulen heraus.

Insgesamt bietet der multilaterale Transfer, wie hier vorgeschlagen, eine leistungsfähigere und umfassendere Methode für den Wissens- und Technologietransfer und ermöglicht somit eine bedarfsorientierte Umsetzung von Digitalisierungsprojekten mit deutlich höheren Erfolgsraten. Er erkennt und nutzt die unterschiedlichen Beiträge aller Beteiligten und stellt sicher, dass der Transferprozess nicht nur einseitig, sondern vielmehr ein dynamischer und interaktiver Austausch ist. Er zeigt außerdem, wie Wissenstransfer in der Ära der Datenwissenschaft und Digitalisierung neu gedacht werden kann, um breitere und tiefgreifendere Auswirkungen zu erzielen und eine integrative und nachhaltige digitale Transformation zu fördern.

Das DASU – im Kern der Transferaktivitäten – koordiniert als zentrale Anlaufstelle den angesprochenen Wissens- und Technologietransfer. Es schafft Netzwerke und bringt die richtigen Akteure zueinander, um im gemeinsamen Verbund die erfolgreiche Umsetzung von Digitalisierungsprojekten zu ermöglichen.

3 Der Beitrag des DASU zur intelligenten Digitalisierung

Anhand zweier konkreter Fallbeispiele soll im Folgenden der anwendungsorientierte Einsatz der beschriebenen multilateralen Transferbemühungen verdeutlicht und der Nutzen dieser Herangehensweise für die erfolgreiche Umsetzung von Digitalisierungsprojekten gezeigt werden.

Zudem wird auf die Förderung des DASU im Rahmen des Landeswettbewerbs RegioWIN 2030 eingegangen, mit deren Hilfe der Ansatz auch durch das Wirtschaftsministerium Baden-Württemberg unterstützt wird.

3.1 Projektbeispiel 1: smyl:tp

Im Rahmen des Projekts smyl:tp begleitet das DASU ein regionales Unternehmen bei der Entwicklung einer KI-basierten Plattform für die Kieferorthopädische Diagnose und Behandlungsplanung.

Das Hauptziel dieser webbasierten Plattform besteht darin, Kieferorthopäden bei der Erstellung und Umsetzung von Behandlungsplänen zur Korrektur von Zahnfehlstellungen für ihre Patienten zu unterstützen. Ein wesentlicher Bestandteil dieser Behandlungspläne sind Modelle des Zahnbogens der Patienten, die durch einen Interoral-Scanner (IOS) digitalisiert werden. Dieser Scanner erfasst entweder einen Zahnabdruck oder direkt den Innenraum der Mundhöhle des Patienten. Die Digitalisierung dieser Modelle ermöglicht eine computergestützte Automatisierung der Behandlungsplanung zur Korrektur von

Zahnfehlstellungen. Dazu wurden verschiedene Schritte mithilfe modernster maschineller Lern- und Analyseverfahren automatisiert.

Ein Beispiel für solche Verfahren ist die automatische Segmentierung der einzelnen Zähne aus dem 3D-Modell sowie deren Kennzeichnung mit der entsprechenden Zahnnummer. Zum Einsatz kommt hier ein neuronales Netz, welches speziell für diesen Anwendungsfall eigens entwickelt und trainiert wurde (siehe Abbildung 3).

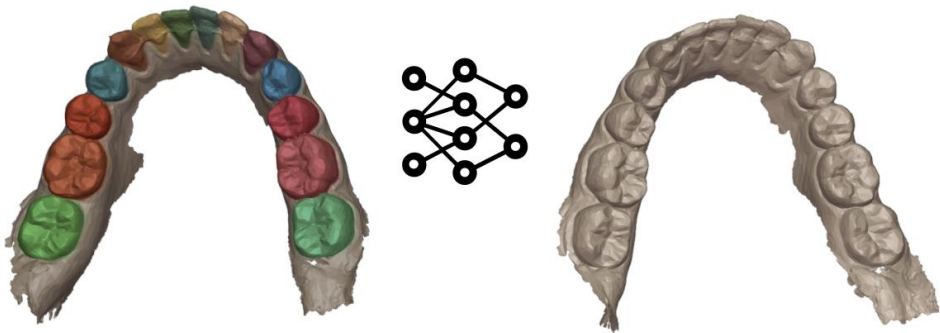


Abb. 3: 3D-Segmentierung von Zähnen durch Deep Learning

In diesem Projekt bringt das DASU durch seine Experten nicht nur Wissen aus den Bereichen des maschinellen Lernens mit ein, sondern bietet darüber hinaus Promovierenden der beteiligten Hochschulen die Möglichkeit, praxisnahe Forschung durchzuführen und die Ergebnisse direkt in der Anwendung umzusetzen – getrieben durch den Einsatz des multidirektionalen Transfergedankens. Das hier vorgestellte Projekt zeigt somit eindrucksvoll, wie die im DASU vorhandenen Ressourcen genutzt werden, um flexibel und zielgerichtet komplexe, technologisch anspruchsvolle und forschungsintensive Digitalisierungsprojekte auch in kleineren Unternehmen durchführen zu können.

Dabei profitieren alle Beteiligten besonders von der Nähe zu den beiden Hochschulen, die in Verbindung mit dem Virtuellen Zentrum DASU auch ein kooperatives Promotionskolleg (KPK) gewährt bekamen. Somit kann der Projektmitarbeiter die innovativen Verfahren sowohl praktisch einsetzen als auch wissenschaftlich mit den übrigen Kollegiat:innen diskutieren und als Anwendungsbeispiel im Rahmen seiner Promotion nutzen.

3.2 Projektbeispiel 2: Vorhersage Parkhausbelegung

Ein weiteres aktives Projekt, welches sich die Vorteile des multidirektionalen Transfergedankens zu Nutze macht, ist die Entwicklung einer Plattform zur Vorhersage von Parkhausbelegungsdaten der Stadt Ulm.

Die Stadt Ulm verwaltet im Stadtgebiet eine Reihe von Tiefgaragen und Parkhäusern und speichert in diesem Zusammenhang Daten zur allgemeinen Parkplatzauslastung. Neben generellen Belegungsdaten der Parkhäuser spielen sogenannte Sonderparkplätze eine wichtige Rolle. Für einige dieser Sonderparkplätze, beispielsweise Frauenparkplätze und Behindertenparkplätze, wurden Sensoren installiert, die eine Belegung des individuellen Parkplatzes feststellen können.

Das Hauptanliegen der Stadt Ulm besteht darin, Bürger über den Status der Parkhausauslastung zu informieren. Dies geschieht bereits teilweise durch das örtlich installierte Parkleitsystem. Darüber hinaus soll eine interaktive Web-Plattform entstehen, mit deren Hilfe Nutzer eine Übersicht über die aktuelle Auslastung der Parkhäuser und deren Sonderparkplätze erhalten.

Zudem kann mithilfe historischer Belegungsdaten und geeigneter KI-Methoden eine Belegung der Sonderparkplätze für spätere Zeitpunkte prognostiziert werden, um Nutzern die Planung einer Anreise in die Ulmer Innenstadt zu erleichtern. Abbildung 4 zeigt die geplante Umsetzung für diese Anwendung.

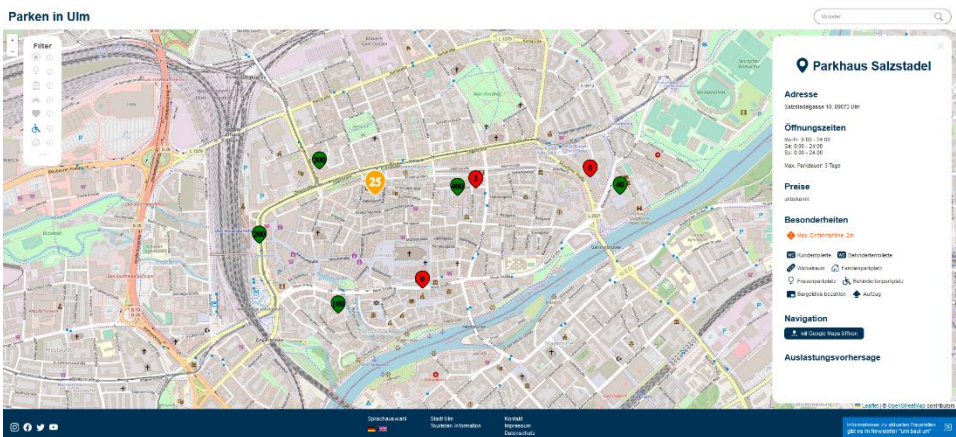


Abb. 4: Prototyp der Web-Plattform mit Filtermöglichkeiten für Sonderparkplätze

Dass auch dieses Vorhaben erfolgreich umgesetzt werden kann, ist der intelligenten Organisation des DASU zu verdanken. So entwickeln Studierende des Masterstudiengangs Intelligent Systems der Technischen Hochschule Ulm einen Prototypen der geplanten Anwendung im Rahmen ihrer Ausbildung – komplett eingebettet in das studentische Curriculum. Auch in der sehr praxisorientierten Ausbildung der THU ist die Zusammenarbeit mit einer Einrichtung des öffentlichen Sektors keine Selbstverständlichkeit und stellt – auch für die Stadt Ulm – ein Novum dar. In regelmäßigen Treffen liefern Studierende Fortschrittsberichte an Professoren, Vertreter der Stadt Ulm und das DASU. Durch diese einzigartige Konstellation stellt das DASU nicht nur das Gelingen des Projekts sicher, sondern beteiligt sich aktiv an der Ausbildung

junger Studierender, indem diese durch die Datenwissenschaftler:innen des DASU gecoach werden.

Der multidisziplinäre Ansatz des DASU ermöglicht außerdem die Einbeziehung sozioökonomischer Fragestellungen in den Entwicklungsprozess und somit eine Fokussierung auf die Bedürfnisse der Digitalisierung im öffentlichen Sektor. So können Bürger durch gezielte Informationsveranstaltungen und Befragungen, beispielsweise im Rahmen des Frauenforums Ulm, den Entwicklungsprozess mit beeinflussen, um von Beginn an eine maximale Akzeptanz der Anwendung zu ermöglichen.

Das hier beschriebene Vorhaben zeigt, wie das Zusammenspiel unterschiedlicher Stakeholder, Interessensgruppen und Projektteilnehmer:innen funktionieren kann, um ein komplexes Digitalisierungsprojekt erfolgreich durchzuführen. Das DASU bringt hierbei nicht nur seine fachliche und methodische Expertise ein, sondern übernimmt eine zentrale Anker-Rolle bei der Projektsteuerung sowie der Vernetzung der beteiligten Parteien.

3.3 Basis für Projekte mit KMU: Förderung RegioWIN 2030

Im Rahmen des Landeswettbewerbs „RegioWIN 2030 – Regionale Wettbewerbsfähigkeit durch Innovation und Nachhaltigkeit“ wurde das DASU von der Region Schwabenbund als Leuchtturmprojekt beantragt. Der Wettbewerb ist eingebettet in die Förderung aus dem Europäischen Fonds für regionale Entwicklung (EFRE) in Baden-Württemberg in der Förderperiode 2021-2027 [EF21]. Die Prämierung des DASU [MW22] verdeutlicht somit die Überzeugung des Wirtschaftsministeriums des Landes Baden-Württemberg, dass das DASU dazu beiträgt, die Wettbewerbsfähigkeit der Region zu verbessern und somit zu einer zukunftsfähigen Regionalentwicklung beitragen wird.

Über einen Zeitraum von 5 Jahren wird das DASU somit dabei unterstützt, KMU bei der digitalen Transformation zu begleiten. In der Region gibt es zahlreiche kleine und mittelständische Produktionsunternehmen, von denen gegenwärtig viele sehr erfolgreich auf dem Weltmarkt agieren, gleichzeitig aber vor großen Herausforderungen insbesondere bei Fragen der Digitalisierung und des Einsatzes von Künstlicher Intelligenz stehen.

Die Umsetzung findet dabei durch (Impuls-)Vorträge, Schulungen und / oder Workshops statt, aber ein besonderes Augenmerk liegt auf der gemeinsamen Durchführung von Projekten, um das Potenzial datengetriebener Anwendungen in den Betrieben zu nutzen. Ein wichtiger Baustein hierbei wird das im Rahmen der Förderung aufzubauende Digital Lab sein. Das Digital Lab soll eine zentrale Transferplattform des DASU bilden, in der innovative KI-Anwendungen und -Prototypen, integriert in soziotechnische und cyber-physische Systeme, entwickelt und mittels entsprechender Technologie visualisier- und erfahrbar gemacht werden können. Die Kommunikation und Darstellung wird dabei auf KMU ausgerichtet, um die Chancen innovativer Data Science und KI-Lösungen zu verdeutlichen.

Konkret konnten bereits erste Projekte durchgeführt bzw. begonnen werden. Hierbei handelt es sich insbesondere um klassische, toolgetriebene Datenanalyse-Projekte für kleine, verarbeitende Unternehmen. Durch den Einsatz von Standardanwendungen in diesem Bereich ist es so beispielsweise gelungen, Maschinenauslastungen für eine Druckerei in Management-Dashboards darzustellen und Produktionskennzahlen an den Maschinen für Mitarbeiter zu visualisieren. Durch die Umsetzung dieser technologisch grundlegenden Digitalisierungsmaßnahmen, die bereits in dieser Form einen hohen Nutzen für die Anwender:innen bieten, ist die Grundlage für weiterführende analytische und datenwissenschaftliche Projekte innerhalb der KMU geschaffen.

4 Schlussfolgerungen und zukünftige Ausrichtung

Das DASU mit seiner innovativen Struktur aus Hochschuleinrichtungen und Partnern aus der Wirtschaft bietet eine unkomplizierte Möglichkeit für Unternehmen und Einrichtungen insbesondere, aber nicht nur, in der Region Ulm, um in Fragen der Digitalisierung voranzukommen. Die Mitarbeitenden des DASU helfen den Interessierten dabei, Möglichkeiten von künstlicher Intelligenz zu erkennen und operativ zu nutzen, aber auch sich deren Gefahren bewusst zu sein und diese zu umgehen. Zudem sind einfachere praktische Fragen aus dem Bereich Analytics, bei denen es oft um die Aufbereitung und anschauliche Darstellung von Informationen geht, als Ausgangspunkt für weitere Zusammenarbeit und als Quick-Wins von großer Bedeutung. Darüber hinaus dienen die mit Unternehmen und / oder öffentlichen Einrichtungen und Bürgern diskutierten Fragestellungen umgekehrt den beteiligten wissenschaftlichen Expert:innen, ihre Forschungen bedarfsgerecht auszugestalten.

Das DASU wird daher den eingeschlagenen Weg konsequent weiterverfolgen und die Vernetzung zu Fragen der Digitalisierung vorantreiben. Ein wichtiger Aspekt dabei wird der Aufbau des im Rahmen von RegioWIN 2030 geförderten Digital Labs sein. Mit dessen Hilfe wird gerade Künstliche Intelligenz begreifbar gemacht und die Übertragung auf konkrete Fragestellungen von Interessierten kann an konkreten Anwendungsbeispielen diskutiert werden.

Insbesondere wird das DASU auch daran arbeiten, weitere Stifter zu überzeugen, sich an der Stärkung der Wirtschaftsregion aktiv zu beteiligen. Sie werden dadurch Teil eines Netzwerks, innerhalb dessen sie sich austauschen und gegenseitig unterstützen können. Darüber hinaus leisten sie damit einen wichtigen Beitrag zu höherer Wettbewerbsfähigkeit der Akteure der Region im internationalen Umfeld durch intelligente Nutzung der neuen digitalen Möglichkeiten.

Literaturverzeichnis

- [DA21] DASU, Transferzentrum für Digitalisierung, Analytics & Data Science Ulm, www.dasu.digital, Stand: 07.07.2023.
- [EF21] EFRE 2021-2027, <https://2021-27.efre-bw.de/>, Stand: 07.07.2023.
- [MW22] Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg, <https://wm.baden-wuerttemberg.de/de/service/presse-und-oeffentlichkeitsarbeit/pressemitteilung/pid/wirtschaftsministerium-foerdert-dasu-mit-circa-49-mio-euro-aus-eu-und-landesmitteln>, Stand: 07.07.2023.

Die Hürden der Digitalisierung in KMUs erfolgreich nehmen

Eine strukturierte Vorgehensweise führt um Ziel


Klaus-Jürgen Meier¹ 

Abstract: KMUs liegen gegenüber Großunternehmen immer noch zurück, wenn es um die Digitalisierung von Geschäftsprozessen geht. Dies betrifft insbesondere die Bereiche Produktion und Logistik. Basierend auf Literaturrecherchen werden die maßgeblichen Hürden einer Digitalisierung bestimmt und Gegenmaßnahmen entwickelt. Anhand eines Fallbeispiels werden deren Wirksamkeit überprüft und weitere Erkenntnisse abgeleitet.

Keywords: KMU, Digitalisierung, Produktion, Logistik, Hürden, Lean Production, Standardisierung, Modularität, Komplexitätsreduktion, Mitarbeiterbindung.

1 Einleitung

Mit dem Begriff „Industrie 4.0“ (I4.0) verkünden Wahlster, Kagermann und Lukas [KLW11] im Jahr 2011 im Auftrag der Bundesregierung die vierte Industrielle Revolution in der Industrie. Mit Bezug auf laufende Veränderungen der IT-Technologien in der Wirtschaft sowie in der Gesellschaft wird die Notwendigkeit propagiert, zu einem nachhaltigen Wandel des Digitalisierungsgrades in u.a. deutschen Produktionsunternehmen [Ar13]. Erklärtes Ziel ist es, kundenspezifische Produkte mit den Methoden der automatisierten Massenfertigung bis hinab zur Stückzahl ‚1‘ wirtschaftlich herstellen zu können. Nur auf Basis dieses erfolgreichen Wandels erscheint aus Sicht der genannten Autoren eine langfristige Wettbewerbsfähigkeit am Hochlohnstandort realistisch. Obwohl die Resonanz groß ist und zahlreiche Anbieter mit Produkten und Dienstleistungen in Industrie, Handel und Logistik reagieren [Li20], scheint bis heute der Durchbruch nicht gelungen zu sein. So beschreiben VDE und DKE im Jahr 2023 Industrie 4.0 immer noch auf der Schwelle zwischen Konzeption und Umsetzung [DK23]. Während Großunternehmen die Herausforderung zumeist angenommen haben, ist bei kleinen und mittleren Unternehmen (KMUs) die Situation deutlich differenzierter zu bewerten. Junge Unternehmen (i.a. Start-ups) nutzen häufig die Digitalisierung gezielt im Sinne eines innovativen Geschäftsmodells. Bei langjährig existierenden KMUs hingegen, sind digitale Anwendungen noch stark unterrepräsentiert. Auch ist entsprechend der Unternehmensberei-

¹ Hochschule München, Institut für Produktionsmanagement und Logistik, Lothstr. 64, 80335 München, klaus-juergen.meier@hm.edu,  <https://orcid.org/0000-0003-0603-9435>

che zu differenzieren. Nachstehende Abbildung 1 zeigt eine Gegenüberstellung der Digitalisierungsrate aufgelöst nach Unternehmensgröße und Unternehmensbereiche bzw. Digitalisierungsschwerpunkte. Wie ersichtlich, werden zunächst Verwaltung und Marketingfunktionen digitalisiert. Insbesondere in den Bereichen ERP, CRM, SCM und RFID weisen KMUs einen Rückstand auf. Damit sind insbesondere Produktion und Logistik von den Defiziten betroffen [OE21].

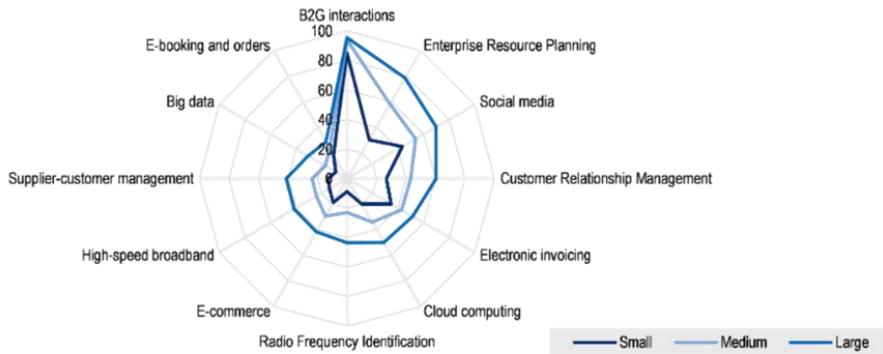


Abb. 1: Digitalisierungsrate in unterschiedlichen Unternehmensbereichen und Technologien in Abhängigkeit von der Unternehmensgröße [OE21]

Um eine Vorgehensweise zu entwickeln, wie zukünftig der Einstieg in die Digitalisierung in den Bereichen Produktion und Logistik für KMUs erleichtert werden kann, gilt es in Kapitel 2) zunächst, die Hürden für die diesbezüglich schleppende Auseinandersetzung mit der Digitalisierung in KMUs zu ermitteln. Basierend auf deren Kenntnis können in Kapitel 3) Maßnahmen auf Erfahrungsbasis und aus Literaturrecherchen abgeleitet werden. Ob diese in der Praxis von Bestand sind, soll im Rahmen eines Fallbeispiels überprüft werden. Kapitel 5 fasst die Ergebnisse nochmals zusammen und zieht Schlussfolgerungen.

2 Hürden der Digitalisierung für KMUs

Digitalisierung kann im engeren Sinn verstanden werden als „(...) das Überführen analoger Daten in ein diskretes System mit nur sehr wenigen Wertezuständen, im Extremfall sogar nur zwei (Binärsystem).“ [HEL18] Spricht man jedoch von Digitalisierung im Unternehmenskontext, so soll Digitalisierung als die „(...) Einführung digitaler Technologien in Unternehmen und als Treiber der digitalen Transformation“ bezeichnet werden [He19]. Das Angebot digitaler Technologien ist sehr groß und entwickelt sich auch laufend weiter [Sc19]. Die Beispiele digitaler Technologien reichen u.a. von Blockchain, Big Data, Anwendung Künstlicher Intelligenz bis hin zum Einsatz von Drohnen oder Kollaborativen Robotern [BDO20], [HKE22].

Die Literatur hat sich schon sehr intensiv mit der Fragestellung beschäftigt, welche Hürden insbesondere KMUs auf dem Weg zur Anwendung dieser digitalen Technologien überwinden müssen. Grundsätzlich können die Hürden sechs übergeordneten Kategorien zugeordnet werden. Diese sind [Me21], [Sh99]:

Orgaware: Hiermit wird die bestehende Aufbau- und Ablauforganisation eines Unternehmens beschrieben. Studien zeigen, dass Agilität und Anpassungsfähigkeit, Unternehmensgröße, finanzielle Ressourcen, betriebliche Fähigkeiten, Unternehmenskultur und Unternehmensalter einen erheblichen positiven oder auch negativen Einfluss auf die Bereitschaft zur Nutzung digitaler Technologien haben.

Manageware: Das Verhalten, die Einstellungen und das Profil des Unternehmenseigentümers und des Managements werden von dieser Kategorie erfasst. Sie wirken direkt auf die Unternehmenskultur und das Verhalten der Mitarbeiter. In diesem Zusammenhang werden vor allem die folgenden Faktoren genannt: Engagement und Offenheit des Eigentümers, strategische Ausrichtung, Managementwissen und -erfahrung (digitale Kompetenz und unternehmerische Fähigkeiten), Innovationsfähigkeit und Agilität, Alter und Geschlecht sowie die zeitlichen Ressourcen der Führungskraft.

Contextware: Kapazitäten und Ressourcen, die aus dem Unternehmensumfeld stammen, erleichtern die Nutzung digitaler Technologien. Die Kategorie untersucht in diesem Zusammenhang die Verfügbarkeit von Forschungsergebnissen oder Angebote von Dienstleistern. Als weitere relevante Einflüsse im Unternehmensumfeld gelten Kundenanforderungen, rechtliche Rahmenbedingungen, Wettbewerbsdruck, Zusammenarbeit mit anderen Unternehmen, Abhängigkeiten und Machtgleichgewicht innerhalb von Wertschöpfungsketten und Netzwerken, Marktvolumen, Branche, ökologischer, sozialer und geografischer Kontext sowie staatliche und politische Förderung.

Technoware: Diese Kategorie bezieht sich auf die physischen Kapazitäten und technischen Mittel, welche einem Unternehmen zur Verfügung stehen, um den Menschen im Rahmen seiner Tätigkeit zu unterstützen bzw. um den Menschen zu ersetzen. Als Beispiele gelten die vorliegende unternehmensinterne und -externe technische Infrastruktur, die Kompatibilität der in einem Unternehmen genutzten Technologien, die bestehenden Unsicherheiten bezüglich der technologischen Weiterentwicklung, die Komplexität einer Anwendung sowie die Angemessenheit eines Technologieeinsatzes. Große Bedeutung hat in diesem Zusammenhang auch, welches Investitionsbudget für den Technologieeinsatz notwendig bzw. vorhanden ist und wie sich später die laufenden Kosten verhalten.

Humanware: Diese Kategorie betrachtet den vorhandenen Mitarbeiterstamm. So stellt sich die Frage nach dem Durchschnittsalter der Belegschaft, den vorhandenen Personalressourcen bzw. dem Wissen und den Fähigkeiten der einzelnen Mitarbeiter sowie der Mitarbeiterintegration im Rahmen von Veränderungsprozessen.

Inforeware: Die Kategorie bezieht sich auf dokumentiertes und aufgezeichnetes Wissen und Fakten. Damit lässt sich der Kompetenzaufbau im Unternehmen beschleunigen und der Bedarf an Ressourcen verringern. Als wichtige Einflussgrößen werden vor allem die Datensicherheit und die wahrgenommenen Ergebnisse bzw. fehlende Leistungsmessungen angeführt, gefolgt von Datenqualität, offener Kommunikation und Datenmanagement.

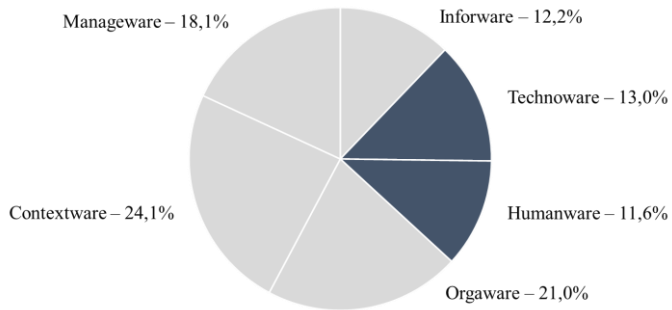


Abb. 2: Zuordnung von Hürden auf dem Weg zur Digitalisierung (in Anlehnung an Meier [Me21]).

Im Rahmen einer umfassenden Literaturrecherche ist es gelungen, in 77 einschlägigen Veröffentlichungen 38 unterschiedliche Hürden auf dem Weg zur Digitalisierung zu ermitteln. Diese wurden den vorstehenden Kategorien zugeordnet. Abbildung 2 zeigt deren numerische Verteilung. Dies erfolgte ohne eine Gewichtung oder Priorisierung der diskutierten Aspekte für KMUs [Me21].

Hürde	Anzahl	kum. Anteil	Kategorie
Wissen und Fähigkeiten	30	34,1%	Humanware
Bestehende technische Infrastruktur	16	52,3%	Technoware
Kompatibilität zu bestehenden Technologien	9	62,5%	Technoware
Unsichere Entwicklung	9	72,7%	Technoware
Komplexität	8	81,8%	Technoware
Personalressourcen	7	89,8%	Humanware
Mitarbeiterintegration	4	94,3%	Humanware
Investitionsbudget und Kosten	2	96,6%	Technoware
Angemessenheit	2	98,9%	Technoware
Alter der Mitarbeiter	1	100,0%	Humanware

Tab. 1: Hürden der Bereiche Produktion und Logistik und deren kumulierter prozentualer Anteil an den Nennungen der Kategorien Technoware und Humanware (in Anlehnung an Meier [Me21]).

Wie aus den vorstehenden Ausführungen abgeleitet werden kann (vgl. Abb. 1), stellen jedoch für KMUs insbesondere die den Unternehmensbereichen Produktion und Logistik zuzurechnenden Hürden die größte Bedeutung dar. Dies trifft auf die Kategorien Technoware und Humanware zu (in Abb. 2 dunkelblau hervorgehoben), weswegen diese nun im weiteren Verlauf herausgenommen und näher betrachtet werden sollen. Die beiden Kategorien umfassen ca. 25% Anteil an der Summe aller Nennungen. Welche Nennungen den beiden Kategorien im Detail zugeordnet sind und welchen kumulierten Anteil diese in den beiden Kategorien haben, zeigt Tabelle 1.

Für die erfolgreiche Digitalisierung im Bereich Produktion und Logistik ist es deshalb entscheidend, sich mit den in Tabelle 1 genannten Hürden auseinanderzusetzen. Gesucht sind methodische Ansätze, die zu einer erfolgreichen Bewältigung der Problemstellungen beitragen.

3 Methodische Ansätze zur Überwindung der Hürden

Zur Lösungsfindung soll auf bekannte und verbreitete Ansätze des Produktionsmanagements, der Produktionslogistik und der IT zurückgegriffen werden, um weitgehende Kompatibilität zu bestehenden Technologien, IT-Standards und auch vorhandenem Wissen der Mitarbeiter sicherzustellen. Als Randbedingung ist sicherzustellen, dass die Anforderungen einer Massenfertigung mit kundenspezifischer Stückzahl ‚1‘ erfüllt werden können, wie sie mit ‚Industrie 4.0‘ erreicht werden sollen – vgl. Kapitel 1) und Arbeitskreis Industrie 4.0 [Ar13]. Welche Maßnahmen zur Verfügung stehen, zeigt die Mindmap in Abb. 3. Sie ist das Ergebnis aus den langjährigen Erfahrungen des Autors bzw. einer Literaturrecherche.

Zu erkennen ist, dass es keine Maßnahme gibt, mit der es gelingt, alle erfassten Hürden gleichsam zu beeinflussen. Vielmehr wird ein Bündel an Maßnahmen benötigt, welches inhaltlich und in seiner Wirkung zeitlich aufeinander abgestimmt werden muss. Bei den identifizierten Maßnahmen aus Abb. 3 handelt es sich (im Uhrzeigersinn) um:

Nutzung von vorhandenem Equipment: Unternehmen versuchen zumeist, das vorhandene Equipment so gut wie möglich weiterhin zu nutzen. Je nach Alter und Ausstattung des Equipments ist diese Maßnahme im Rahmen der Digitalisierung vielfach nicht oder nur in Verbindung mit anderen Maßnahmen realistisch [DC19].

Nutzung einfacher Systeme: Die Installation hoch komplexer automatisierter Systeme verspricht zwar die maximale Entwicklungsgeschwindigkeit. In Abhängigkeit des Ausgangsszenarios besteht damit jedoch die Gefahr, dass das vorhandene Mitarbeiterpotenzial sowohl kompetenzmäßig als auch in seiner Akzeptanz überfordert ist. Zudem muss ein entsprechendes Investitionsbudget bereitgestellt werden, was gerade für KMUs eine Herausforderung darstellen kann. Liegen keine technologischen Erfahrungen vor, so besteht zusätzlich eine große Gefahr für Fehlinvestitionen mit unabsehbaren Auswirkungen auf die zukünftige Kostensituation und die langfristige

Nutzung des Investitionsobjekts. Der Kauf komplexer Systeme muss also wohl überlegt sein und dürfte insbesondere in der frühen Phase der Digitalisierung bei vielen KMUs auf Zurückhaltung stoßen. Besser ist es, zunächst in einfachere Systeme zu investieren und damit Erfahrungen aufzubauen [MRT20]. Auf diese Weise erfolgt eine langsame Erhöhung des Komplexitätsgrads in der Produktion, der es ermöglicht, die Mitarbeiterkompetenz in analoger Weise zu entwickeln.

Einführung einer Standardisierung: Die Schaffung von Standards in Geschäftsprozessen, IT (z.B. durch Nutzung des OPC-Standards zum Austausch von Daten [OP23]) und Anlagen bietet die Möglichkeit, Arbeitsergebnisse zu stabilisieren und gleichzeitig den Aufwand zu reduzieren. Unternehmen mit etablierten Standards verfügen zumeist über ein hohes Qualitätsniveau sowie große Flexibilität. Die eingebundenen Mitarbeiter verfügen über das benötigte Wissen, um auch in Ausnahmesituation schnell und sicher reagieren zu können. Damit ist die Einführung von Standards in jedem Fall ein attraktives Ziel.

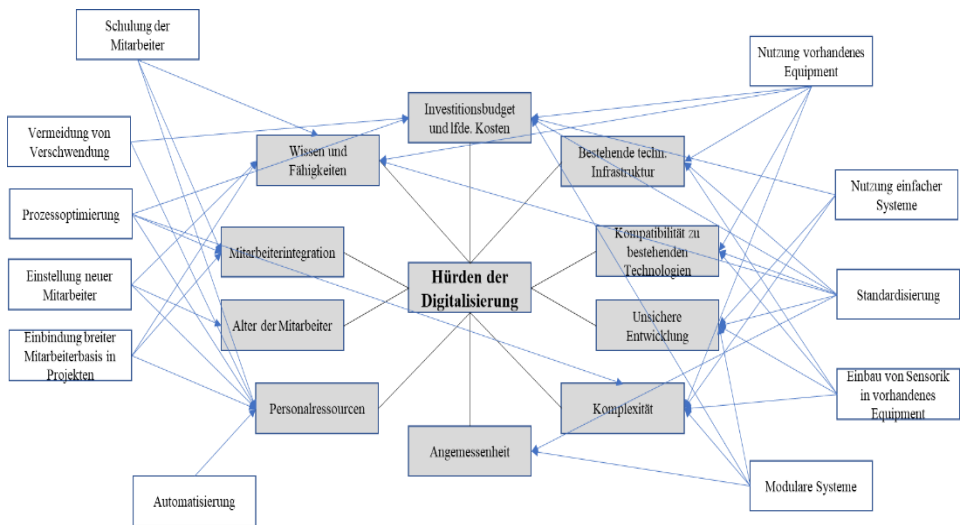


Abb. 3: Mindmap zur Ermittlung von Maßnahmen, welche dazu dienen, die Hürden der Digitalisierung im Bereich Produktion und Logistik (im Bild grau hinterlegt) zu überwinden.

1Einbau von Sensorik in vorhandenes Equipment: Zum Aufbau autonom ablaufender Prozesse muss dieser durch Sensoren überwacht werden. Entscheidungen müssen auf Basis der gewonnen Daten möglichst in Echtzeit getroffen und in Handlungen übersetzt werden. Je nach Maschinentyp und -alter ist die Nachrüstung von konventionellen Maschinen und Anlagen mit Sensoren möglich. In einigen Fällen kann dies sogar eine kostengünstige Alternative zur Neuinvestition darstellen. Ist die Nachrüstung möglich, so kann mit einem bekannten Maschinentyp weiter gearbeitet werden [ABN11], [KFS18].

Im Notfall ist der Schritt zurück auf den konventionellen Betrieb jederzeit möglich. Diese Option schafft Vertrauen und ermöglicht es, Know-How aufzubauen.

Nutzung modularer Systeme: Ist die Zusammenstellung von Systemen im Baukastenprinzip möglich, so kann die Digitalisierung in abgestuften Schritten erfolgen. Abgestimmt auf das vorhandene Investitionsbudget, das Mitarbeiterwissen sowie auf die Mitarbeiterakzeptanz lassen sich sukzessiv Entwicklungen planen und durchführen. Jeder Innovationsschritt kann mit geringem Investitionsrisiko ausgetestet werden und dient zum kontinuierlichen Aufbau des Erfahrungspotenzials. Verändern sich die Anforderungen in Bezug auf Spezifikation oder Volumen, so kann mit geringer Investition eine Anpassung vorgenommen werden. Das System kann schrittweise von einer einfachen Digitalisierungsebene kommend ausgebaut werden bis zu einer vollkommen autark laufenden Produktionslinie [APS20].

Einführung voll automatisierter Systeme: Mit der Einführung voll automatisierter Systeme machen Unternehmen in kürzester Zeit einen großen Sprung in Richtung Digitalisierung. Attraktiv erscheint die mögliche Reduktion der Mitarbeiterzahl. Die Maßnahme kann sinnvoll sein, wenn zum Beispiel beträchtliche Volumensteigerungen zu meistern sind. Liegen in dem KMU jedoch keine Erfahrungen hinsichtlich des Betriebs eines automatisierten Systems vor, so geht das Unternehmen ein großes Risiko ein. Um den stabilen und unterbrechungsfreien Betrieb zu gewährleisten, muss das Mitarbeiterwissen ebenfalls in kürzester Zeit auf den Stand der Technik gebracht werden. Geschieht dies nicht, so begibt sich das KMU in Abhängigkeit des Systemlieferanten oder anderer Dienstleister. Die Akzeptanz unter den Mitarbeitern aufgrund der disruptiven Veränderung kann extrem leiden. Es drohen Systemausfälle und hohe Betriebskosten.

Einbindung einer breiten Mitarbeiterbasis: Je größer die in das Projekt involvierte Mitarbeiterbasis ist, auf einen desto größeren Erfahrungsschatz kann zurückgegriffen werden. Gleichzeitig erhöht sich die Akzeptanz in der Belegschaft und die Flexibilität bei der Besetzung von Projektstellen. Schwierige und ggf. langwierige Diskussionen bei der Entscheidungsfindung stehen dieser Vorgehensweise oftmals entgegen. Es empfiehlt sich deswegen, eine Abwägung zu treffen und eine ausreichend große Basis an Mitarbeitern einzubinden, um den Projekterfolg abzusichern. Zeitgleich ist eine offene Kommunikation mit den nicht unmittelbar am Projekt beteiligten Mitarbeitern zu pflegen, um diese nicht zu Mitarbeitern ‚zweiter Klasse‘ zu degradieren [Bo13].

Einstellung neuer Mitarbeiter: Die Kompensation fehlenden Wissens oder fehlender Fähigkeiten im Unternehmen kann durch die Akquise neuer Mitarbeiter behoben werden – sofern auf dem Arbeitsmarkt für das KMU verfügbar. Erfolgt dies in großem Umfang und stehen dieser Maßnahme Entlassungen gegenüber, so ist mit innerbetrieblichen Spannungen zu rechnen.

Durchführung einer Prozessoptimierung: Die Digitalisierung schlecht geführter konventioneller Prozesse zieht zumeist hohe Prozesskosten nach sich und führt nicht zu dem erhofften Effizienzgewinn. Zudem sind die Kosten für die Implementierung unnötig hoch. Eine der Digitalisierung vorausgehende Optimierung der Prozesse ist deswegen

alternativlos [Hi20]. Je nach Prozessebene eignen sich u.a. Wertstromanalysen [RSW22] oder Flussdiagramme zur Visualisierung und zur nachfolgenden Optimierung [Fül14]. Je weitreichender die Prozessoptimierung gelingt, desto weiter können die für die Digitalisierung benötigten Investitionen gesenkt werden.

Vermeidung von Verschwendung: Einhergehend mit der Prozessanalyse und -optimierung empfiehlt es sich, eine Verschwendungsanalyse entsprechend der Lean-Methode durchzuführen. Sie ist aufwandarm zu erstellen, gibt aber wesentliche Anregungen bezüglich der Potenziale, welche in einem Prozess verborgen sind. Es werden folgende sieben Verschwendungsarten unterschieden: Überproduktion, Bestände, ineffiziente Prozesse, Bewegung, Transport, Wartezeit und Nacharbeit bzw. Ausschuss. Die Verschwendungsanalyse liefert damit nochmals einen Input für die Prozessoptimierung und sollte der Prozessoptimierung bzw. einer Digitalisierung vorausgehen [Bo13], [Sc18].

Schulung der Mitarbeiter: Gut ausgebildete Mitarbeiter sind die Voraussetzung für jede Form der Digitalisierung. Benötigt wird Wissen auf dem aktuellen Stand der Technik, um die Einführung der Digitalisierung zu begleiten und in der Folge einen unterbrechungsfreien Betrieb zu gewährleisten. Know-How wird jedoch bereits für die Entscheidungsfindung benötigt. Nur so lassen sich Unsicherheit im Rahmen der Investitionsentscheidung beheben und Fehlinvestitionen vermeiden. Kompetente Mitarbeiter haben keine Berührungängste mit neuen Technologien und müssen zudem nicht um ihren Arbeitsplatz fürchten. Damit wird der Prozess der Digitalisierung deutlich vereinfacht.

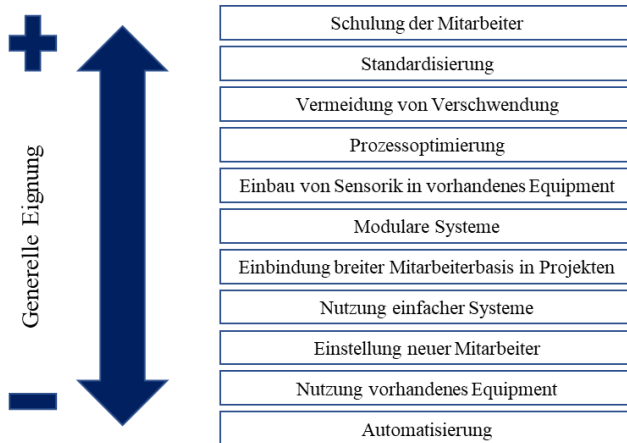


Abb. 4: Maßnahmen zur Überwindung der Hürden der Digitalisierung sortiert nach ihrer Eignung

Auf der Grundlage der vorangegangenen Diskussion der potenziell denkbaren Maßnahmen aus Abbildung 3 wird deutlich, dass nicht alle Ansätze in gleicher Weise geeignet sind, die Digitalisierung in KMUs zu unterstützen. Neben ihrer prinzipiellen

Eignung ist zudem der Zeitpunkt entscheidend, zu dem die jeweilige Maßnahme ergriffen wird. In der Abbildung 4 sind die Maßnahmen nochmals zusammengefasst und hinsichtlich ihrer generellen Eignung sortiert. Die Reihenfolge ist jedoch nicht allgemeingültig, sondern kann sich situationspezifisch ändern.

4 Nachweis der Wirksamkeit anhand eines Fallbeispiels

Die Maßnahmen aus Abbildung 4 wurden in Kapitel 3) aufgrund von Erfahrungswerten und einer Literaturrecherche abgeleitet. Ihre Wirksamkeit soll nun anhand von einem Fallbeispiel überprüft werden. Dazu wird bewusst auf einen Laborbetrieb zurückgegriffen, um die charakteristischen Kennzahlen in den einzelnen Entwicklungsständen der Digitalisierung und unter identischen Bedingungen aufzunehmen und vergleichen zu können.

In dem Fallbeispiel wird ein Fertigerzeugnis (FE) in mittlerer Stückzahl produziert, welches kundenspezifisch konfiguriert werden kann. So variieren nicht nur die zu fertigenden Spritzgussteile in ihrer Ausführung. Neben einer Anzahl an Standardmaterial werden in Abhängigkeit der Kundenbestellung auch unterschiedliche Rohmaterialteile zum Fertigerzeugnis verbaut. Das Rohmateriallager muss also die zum Kundenauftrag passenden Rohmaterialien an den Arbeitsplätzen zur Verfügung stellen. Die Endmontage kann parallel an drei separaten Montagearbeitsplätzen durchgeführt werden. Vor Auslieferung an den Kunden werden sämtliche FE in einer Hundertprozentprüfung von der Qualitätssicherung kontrolliert. In der Ausgangssituation befinden sich die Arbeitsplätze in der Anordnung gemäß Abb. 5.

Wie der Abbildung 5 zu entnehmen ist, handelt es sich um eine verrichtungsorientierte Aufstellung der Arbeitsplätze. Der Transport des Rohmaterials aus dem Rohmateriallager zu den Arbeitsstationen erfolgt in manueller Weise durch die Lagerlogistik. Der Weitertransport der Halbfertigerzeugnisse (HFE) zwischen den Arbeitsplätzen und der Transport der Fertigerzeugnisse (FE) zum Auslieferungslager muss von den Mitarbeitern in der Linie selbst übernommen werden. Aufgabe der Produktionssteuerung ist es, die eingehenden Kundenaufträge losgrößenoptimal einzusteuern, um einerseits die Rüstzeitanteile gering zu halten und andererseits die Einhaltung der Kundentermine zu gewährleisten. Insgesamt sind in der Ausgangssituation sieben Mitarbeiter in der Linie beschäftigt. Für den Betrieb können folgende charakteristische Kennzahlen erhoben werden:

Anzahl Fertigerzeugnisse in 90min:	18Stück
Termintreue:	15%
Nacharbeitsquote:	15%

Nach Abschluss der Zeitaufnahme wird unter Einbindung aller Mitarbeiter ein Workshop durchgeführt. In diesem werden in einem ersten Schritt die vorliegenden Verschwendungsarten erhoben. Zum Erstaunen der Mitarbeiter sind alle Verschwendungsarten mindestens einmal vertreten. Auf der Grundlage dieser Erkenntnis wird der Produktionsprozess und die Anordnung der Arbeitsplätze deutlich überarbeitet. Es resultiert das

flussorientierte Layout (Iteration 1) in Abb. 5. Als wesentliche Veränderung gegenüber der Ausgangssituation zeichnet sich die neue Linie durch eine Arbeitsteiligkeit in der Montage aus. Die Arbeitsplätze sind flussorientiert angeordnet. Dies führt zu einer deutlichen Verkürzung der Transportwege und Bewegungen, auch wenn diese weiterhin von den Mitarbeitern der Linie auszuführen sind. Die Transporte erfolgen in Losgröße, 1'. Die Aufträge werden nun gemäß Auftragsreihenfolge von der Produktionssteuerung vorgegeben. Auf eine Digitalisierung von Produktions- oder Logistikinhalten wird in dem ersten Entwicklungsschritt bewusst verzichtet. Es kommen lediglich Maßnahmen zur Anwendung, wie sie aus dem Lean-Umfeld seit Jahrzehnten bekannt sind. Die Anzahl der Mitarbeiter bleibt unverändert. Bedingt durch die Prozessoptimierung wird jedoch bereits deutlich, dass eine Digitalisierung zu einem späteren Zeitpunkt unmittelbar von den vereinfachten Abläufen profitiert. Die Struktur der Produktionslinie nach dem ersten Iterationsschritt ergibt sich gemäß Abb. 5.

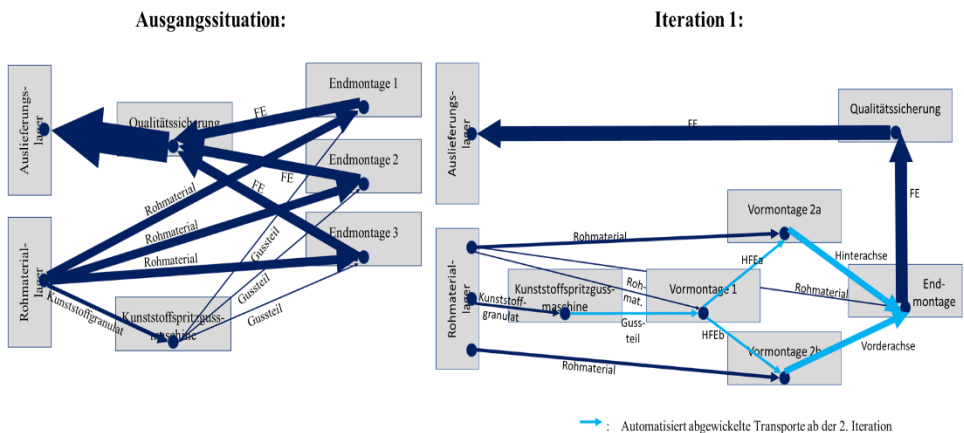


Abb. 5: Anordnung der Arbeitsplätze in der Ausgangssituation und ersten Iteration. Die Pfeile geben Auskunft über die Richtung und Menge an zu transportierendem Produktionsmaterial. Ladungsträger sind nicht dargestellt.

Die Kennzahlen weisen für die Produktionslinie eine signifikante Verbesserung aus:

Anzahl Fertigerzeugnisse in 90min:	45Stück
Termintreue:	96%
Nacharbeitsquote:	4%

Die Anordnung der Arbeitsplätze sowie der arbeitsteilige Produktionsprozess hat sich als richtig erwiesen. In dem Workshop mit den beteiligten Mitarbeitern werden dennoch weitere Verschwendungen erkannt. Unterstützt durch erste einfache Anwendungen der Digitalisierung sollen diese Verschwendungen behoben und zudem eine Effizienzsteigerung erreicht werden. Wesentliche Elemente sind fünf Förderbänder zur Verkettung der Vormontagen (s. farblich hervorgehobene Pfeile in Iteration 1 der Abb. 5 und ein kollaborativer Roboter für den Arbeitsplatz der Endmontage. Alle Elemente verfügen über

eigene Sensoren, eine unabhängige Steuerung und sind damit ‚stand-alone‘-fähig, sodass auf eine zentrale und damit komplexe Speicherprogrammierbare Steuerung (SPS) verzichtet werden kann. Es resultiert ein modularer Aufbau der Linie, welcher leicht an neue Gegebenheiten angepasst werden kann. Zudem ermöglicht dies den Einsatz der Elemente mit einem vergleichsweise geringen Einarbeitungsaufwand bzw. geringen IT-Systemwissen. Da Transport- und Montageaufgaben von den Bändern und dem Roboter übernommen werden, können zwei Mitarbeiter für andere Funktionen in dem Betrieb oder für eine Durchsatzsteigerung freigesetzt werden. Der Break-Even-Point für die erforderliche Investition liegt so bei unter einem Jahr.

Die Kennzahlen weisen für die zweite Iteration nachstehende Werte aus:

Anzahl Fertigerzeugnisse in 90min:	41 Stück
Termintreue:	97%
Nacharbeitsquote:	4%

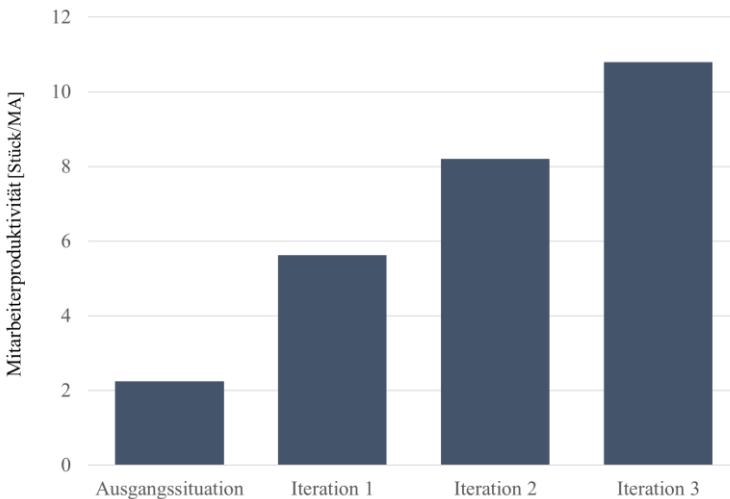


Abb. 6: Anstieg der Mitarbeiterproduktivität in den Iterationsrunden eines Fallbeispiels.

Damit resultiert gegenüber der ersten Iterationsrunde eine geringfügige Verschlechterung in den erfassten Leistungskennzahlen. Als Ursache ergibt sich in dem erneut abgehaltenen Workshop, dass der Umgang mit den neuen Arbeitsmitteln ‚Förderband‘ und ‚Roboter‘ für die Mitarbeiter noch ungewohnt ist. Auch die Programmierung weist noch Potenziale auf. Es entstehen im Produktionsprozess leichte Effizienzverluste. Berücksichtigt man jedoch die reduzierte Mitarbeiterzahl, so erzielt auch die Iterationsrunde 3 einen weiteren Produktivitätsfortschritt. Auf der Basis der gewonnenen Erkenntnisse wird eine vierte (und vorläufig letzte) Iterationsrunde eingeleitet. Sie greift die erkannten Potenziale auf. Mithilfe einer verbesserten Programmführung und einer erweiterten Sensorik gelingt es,

die Stückzeit nochmals zu senken. Dies belegen auch die erhobenen Kennzahlen für die dritte Iteration:

Anzahl Fertigerzeugnisse in 90min:	54Stück
Termintreue:	100%
Nacharbeitsquote:	0%

In Abb.6 ist die Entwicklung der Mitarbeiterproduktivität in der Ausgangssituation und den anschließenden Iterationsrunden zusammengefasst. Vergleicht man die Produktivität in Iterationsrunde 3 mit der Ausgangssituation, so ergibt sich eine Verfünffachung. Dieser beachtlichen Steigerung stehen überschaubare Investitionen gegenüber, die sich bereits innerhalb eines Jahres amortisieren.

5 Zusammenfassung und Fazit

Insbesondere KMUs weisen noch Defizite auf in ihrem Bestreben nach einer Digitalisierung der Prozesse. Von allen Unternehmensbereichen bilden die Produktion und Logistik dabei das Schlusslicht [OE21]. Die Gründe sind vielfältig, aber aus zahlreichen Studien weitgehend bekannt [Me21]. Auf der Basis von Erfahrungswerten des Autors sowie durch Auswertung von Literaturquellen können Ansätze abgeleitet werden, wie die Hürden bei der Digitalisierung überwunden werden können – vgl. Abb. 4. Deren Wirksamkeit wird in einem Fallbeispiel überprüft. In mehreren Iterationsrunden werden Maßnahmen innerhalb einer Produktionslinie ergriffen. Um welche Maßnahmen es sich handelt, ist in Abbildung 7 zusammengefasst.

Jede Iterationsrunde zeigt deutliche Steigerungen der Mitarbeiterproduktivität (s. Abb. 6). Deutlich wird jedoch auch, dass der Digitalisierung eine vorbereitende Phase vorausgehen muss, in der mithilfe konventioneller Ansätze insbesondere die Prozesse vereinfacht und standardisiert werden. Hilfreich sind Methoden, wie sie das klassische Lean Management seit Jahrzehnten zur Verfügung stellt. Die Anwendung von Lean-Methoden und -Werkzeugen wird damit nicht obsolet, sondern zur grundlegenden Voraussetzung für die Weiterentwicklung moderner Betriebe.

Ist die Basis geschaffen, kann durch einfache Formen der Digitalisierung der Einstieg erfolgreich geschafft werden. Dezentrale und modular aufgebaute digitale Arbeitsmittel erlauben es KMUs, erste Erfahrungen zu sammeln. Dies senkt die Hemmschwelle bei der Anwendung weiterer digitaler Hilfsmittel (z.B. Self Guided Vehicles). Gleichzeitig lassen sich modulare Systeme an zukünftige Entwicklungen (in den Bereichen Produkte, Volumen, Technologie) leicht anpassen und weiter ausbauen. Damit reduziert sich das Risiko einer Fehlinvestition deutlich. Unterstützend wirkt die Möglichkeit, Sensorik nachzurüsten. Die automatisierte Erfassung von Daten ist die Voraussetzung für die autonome Entscheidungsfindung bis hin in übergreifenden und komplexen Konstellationen.

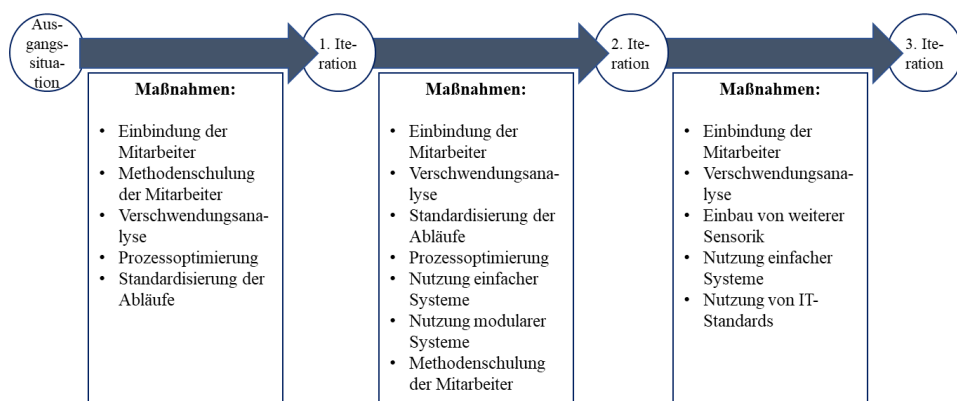


Abb. 7: Zusammenfassung, der im Fallbeispiel eingeleiteten Maßnahmen

Literaturverzeichnis

- [ABN11] Adam, W.; Busch, M.; Nickolay, B.: Sensoren für die Produktionstechnik. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [APS20] Aurich, J.C. et al. Hrsg.: Bedarfsgerechte Digitalisierung von Produktionsunternehmen. Ein modulares Transformationskonzept als praxisorientierter Ansatz. Synnovating, Kaiserslautern, 2020.
- [Ar13] Arbeitskreis Industrie 4.0: Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0 – Abschlussbericht, 2013.
- [BDO20] Bai, C. et al.: Industry 4.0 technologies assessment: A sustainability perspective. International Journal of Production Economics 229, S. 107776, 2020.
- [Bo13] Boppert, J.: Emotion ist Trumpf – Mitarbeiter für Veränderung begeistern. In (Günthner, W. A.; Boppert, J. Hrsg.): Lean Logistics. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 87–95, 2013.
- [DC19] Doyle, F.; Cosgrove, J.: Steps towards digitization of manufacturing in an SME environment. Procedia Manufacturing 38, S. 540–547, 2019.
- [DK23] DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE: Deutsche Normungsroadmap Industrie 4.0 – Version 5, 2023.
- [Fü14] Fűrermann, T.: Prozessmanagement. Kompaktes Wissen - Konkrete Umsetzung - Praktische Arbeitshilfen. Hanser Verlag, München, 2014.
- [He19] Hess, T.: Digitale Transformation strategisch steuern. Springer Fachmedien Wiesbaden, Wiesbaden, 2019.
- [HEL18] Heuermann, R.; Engel, A.; Lucke, J. von: Digitalisierung: Begriff, Ziele und Steuerung. In (Heuermann, R.; Tomenendal, M.; Bressemer, C.

- Hrsg.): Digitalisierung in Bund, Ländern und Gemeinden. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 9–50, 2018.
- [Hi20] Hierzer, R.: Prozessoptimierung 4.0. Den digitalen Wandel als Chance nutzen. Haufe Group, Freiburg, München, Stuttgart, 2020.
- [HKE22] Huth, M.; Knauer, C.; Ender, M.: BME-Logistikstudie 2022: Digitalisierung in Supply Chains, 2022.
- [KFS18] Kampker, A. et al.: Lernen von den Besten: Fünf Erfolgsfaktoren bei der Entwicklung von Smart Services. In (Meyer, K.; Klingner, S.; Zinke, C. Hrsg.): Service Engineering: Von Dienstleistungen zu digitalen Service-Systemen. Springer Fachmedien Wiesbaden, Wiesbaden, S. 151–165, 2018.
- [KLW11] Kagermann, H.; Lukas, W.-D.; Wahlster, W.: Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. VDI Nachrichten 13, S. 2–3, 2011.
- [Li20] Lichtenthaler, U.: Building Blocks of Successful Digital Transformation: Complementing Technology and Market Issues. International Journal of Innovation and Technology Management 01/17, 2020.
- [Me21] Meier, A.: Systematic Review of the Literature on SME Digitalization: Multi-sided Pressure on Existing SMEs. In (Schallmo, D. R. A.; Tidd, J. Hrsg.): Digitalization. Springer International Publishing, Cham, S. 257–276, 2021.
- [MRT20] McFarlane, D. et al.: Digital Manufacturing on a Shoestring: Low Cost Digital Solutions for SMEs. In (Borangiu, T. et al. Hrsg.): Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future. Springer International Publishing, Cham, S. 40–51, 2020.
- [OE21] OECD: The Digital Transformation of SMEs, 2021.
- [OP23] OPC Foundation: List of documents. <https://opcfoundation.org/developer-tools/documents/>, Stand: 17.2023.
- [RSW22] Rother, M. et al.: Sehen lernen. Mit Wertstromdesign die Wertschöpfung erhöhen und Verschwendung beseitigen. Lean Management Institut, Mühlheim an der Ruhr, 2022.
- [Sc18] Schmeisser, W.: Vom Lean Management zur Digitalisierung. UVK Verlagsgesellschaft mbH, Tübingen, 2018.
- [Sc19] Schallmo, D.: Die Digitale Transformation von Geschäftsmodellen als Erfolgsfaktor: Grundlagen, Beispiele und Roadmap. In (Meinhardt, S.; Pflaum, A. Hrsg.): Digitale Geschäftsmodelle – Band 1. Springer Fachmedien Wiesbaden, Wiesbaden, S. 47–66, 2019.
- [Sh99] Sharif, N.: Strategic Role of Technological Self-Reliance in Development Management. Technological Forecasting and Social Change 3/62, S. 219–238, 1999.

Extraction of Information from Invoices – Challenges in the Extraction Pipeline

Lukas-Walter Thiee , Felix Krieger and Burkhardt Funk ¹

Abstract: Data from invoices are key information for business processes. In order to use the data and create business value, the information must be captured in a digital and structured form. Leveraging digital tools and AI/ML is state-of-the-art in the extraction of information from invoices. However, the existing approaches are trained on specific languages and layouts, and while focusing on the performance of individual metrics, they neglect the demonstration of the pipeline from raw data to processable information. In this paper, we investigate the types of information on invoices and address the challenges in the extraction pipeline. We contribute by providing a morphological framework for the problematization and design of a pipeline as part of a design science study.

Keywords: Invoice recognition, Information extraction, Data pipeline.

1 Introduction, Motivation and Method

Extraction of information from business documents is an evolving area of research and practice, as structured, digital information support numerous business processes. While we focus on invoices in this paper, the research approaches can be applied to other document types, such as receipts or checks. Digitizing incoming invoices, i.e., capturing structured information, can not only save a considerable amount of time, but can also add value. E.g., supply chain management can utilize these data by automatically integrating delivery dates and quantities into ERP systems [KAD04]. In addition, structured invoice data enable business analytics, e.g., for purchasing patterns [Fa04], [Ra21]. Furthermore, auditing firms can leverage the data to simplify and enhance financial audits from sample testing to substantive test of details [KAM95]. Despite the possibilities of the electronic creation, transfer and standardized integration of documents [KAD04], it is still common practice today to send invoices on a paper or pdf basis, so that the information must be extracted from the document or file. This refers to both B2B and B2C invoicing processes. In contrast to digital invoice recognition human invoice reading (as well as annotation) is error-prone and costly with average “*processing costs of about 9 Euro*” [KAD04]. Nevertheless, humans are good at the cognitive task of information extraction, i.e., infer abbreviations, link tabular data, and form composite information.

As in many fields of digitalization and research, methods from artificial intelligence (AI) and machine learning (ML) are increasingly being examined and used. The ultimate goal

¹ Leuphana University, Institute of Information Systems, Universitätsallee 1, Lüneburg, 21335,

lukas-walter.thiee@stud.leuphana.de,  <https://orcid.org/0000-0002-1998-376X>

of this field of research is to digitally capture all (relevant) information from arbitrary business documents and make it available in a structured, processable format. In particular, the applications shall translate the raw data on a document into a machine-readable data type, so that the correct learning of the relationships in the data can be used to infer the original information. Existing industry solutions leverage these methods (see *Chapter 2*). However, the solutions are still far from comprehensive recognition of all relevant information, because invoices are designed in a plethora of layouts and languages. In addition, data protection concerns arise in the external processing of business documents, as sensitive information must not end up on external/foreign storages. Although the use of external services can provide access to (pretrained) models, cloud- and development environments, it creates a dependency that reduces both the ability to influence and the understanding of the model output. For these reasons – extraction quality, data privacy, and model dependency – there is a demand in research and practice to create and deploy proprietary models. However, two main difficulties become apparent when creating own models. On the one hand, the availability of a correspondingly large and qualitative/ labeled dataset, which is necessary for training, validation, and testing, and on the other hand, the model selection and the effort for the corresponding data preparation and pipeline implementation. An extraction process usually consists of several steps which depend on each other. We call this process a pipeline. It describes the data flow of an ML/AI model from raw data over algorithmically processable to complete and reliable information, also referred to as ‘*Information Extraction*’ [Cr18]. Since errors propagate within the pipeline, it is important to improve and optimally connect individual steps to ensure the best possible extraction of information, which strongly motivates this research.

However, in order to create such an ML/AI pipeline, a wide variety of decision points and development stages must be passed through. General process descriptions such as CRISP-DM [WH00] or DASC-PM [Sc22] are not sufficient in themselves for this case. While they provide a good structure of general project stages and stakeholders, they don’t offer practical guidelines that support the development and implementation of invoice recognition pipelines. Moreover, existing pipelines are not directly comparable, w.r.t. the deployed model, extracted classes and reported metrics, e.g., the F1-score in [Xu20b] refers to other classes than the F1-score in [Ma20]. Furthermore, the capability of the models to generalize to unknown layouts and languages is limited, far from human capabilities, so that we speak of weak, or artificial narrow intelligence (ANI) [Gi20]. In order to support standardization in the field of document analysis, and invoice recognition in particular, and to promote practical solutions towards artificial general intelligence (AGI), the specific challenges must first be captured systematically. Related literature has neglected the systematization of the plethora of information and the specific challenges of the extraction, in particular in the context of different languages and layout conventions. We identify this as an open research gap and postulate that a comprehensive description of challenges and systematic categorization of information types can close this gap. The proposed framework can ultimately support extraction performance through better understanding of the data pipeline and the model in various countries/languages. Therefore, we address the following research questions: *What challenges exist in ML/AI*

pipelines in the context of invoice recognition? What kind of information types are present on invoice documents that can influence the extraction pipeline?

Research Goal and Method. Our overall research goal is to improve invoice recognition pipelines, which covers the general phases of such projects, including data preparation, model selection, training, evaluation, and implementation. To achieve this goal we apply a design science research (DSR) method, that helps us to generate practice-oriented artifacts [He04], [Pe07]. Hevner defines three inherent research cycles for a DSR project – the *Relevance, Design and Rigor Cycle* [He07]. In this paper, we address the *Relevance Cycle*. At this early stage of our design study, we capture the details of the problem in order to design (software) artifacts in subsequent cycles. In our case, the derivation of the specific challenges and information types (*Chapter 3*), represents an intermediate result, that can be used both in our design study, and beyond, e.g., in the recognition of other business documents containing text, numbers, forms, or tables. The framework (*Chapter 4*) we present in this paper is the result of an initial design requirements cycle. It is a systematic catalog of challenges and information types encountered in invoice recognition, based on comprehensive review of literature and practice approaches (*Chapter 2*).

2 Review of Literature and Practice Approaches

Invoices are a characteristic type of business document, that represent proof of a business transaction, e.g., the purchase of goods or the provision of a service. *“Invoices contain always rather similar information, fostered by legal requirements for information items on invoices. However, the information items are distributed according to all different layout styles”* [KAD04]. Besides different languages, country-specific conventions are often reflected in their appearance and content. In contrast to pure, sequential text, invoices represent a form of visually rich documents [Li19a], on which various sources, especially semantics and 2-dimensional layout, contain the information. The plethora of information on invoices can range from dates and invoice numbers, over account and payment details, to product quantities and descriptions (line items), most of which is not arranged sequentially. Since the information on invoices do not follow a predefined sequence, conventional natural language processing (NLP) techniques are difficult to apply [Kr21]. Rather, meaning of words, their positions, and linguistic nuances, such as abbreviations and certain template conventions, form the superordinate signal basis of invoice information, e.g., a field for recipient address on the top left. To capture the information and to cope with the abundance of unstandardized data and layouts, a wide variety of approaches from the field of ML/AI have been presented in recent years. Therefore, we perform a literature review in relevant online (IS) libraries, such as *AIS (19 results)*, *IEEE (108 results)*, *Web of Science (47 results)*, *Scopus (194 results)*, *SpringerLink (4 results)* for literature from 2010 to 2022. We use combinations of the search terms ‘*invoice*’, ‘*recognition*’, ‘*information*’, and ‘*extraction*’, and enhance the results through forward and backward screening. Since we focus invoice document types, we do not include ‘*receipts*’ or similar search terms explicitly. We sort out duplicates and filter the results

according to whether concrete ML/AI models and pipelines are described, and obtain 26 results for in-depth analysis. We classify the literature results according to the described model types and find, that in digital invoice recognition, a distinction can basically be made between three successive approaches: rule-based, conventional machine learning, and deep learning approaches [Xu20b]. **Rule-based** approaches to invoice recognition [KAD04], [ASG09], [Ja12], [BB04] rely on consistent layouts or templates, defined patterns, and human input, which involves a lot of manual effort [Xu20b] that is difficult to scale, and therefore does not achieve the desired level of automation and cost reduction. ML-based approaches [Es12], [DK02], [Sc13] integrate training based inference and NLP-based models include language models [PWL13], [Si18].

Deep learning (DL) represents a subcategory of ML, characterized by the use of neural networks. For example, [PWL13] apply a Long-Short-Term-Memory network to match eight different key-value-pairs from labeled sequences. Other approaches include end-to-end deep learning models [Ya17], [PLW19]. Convolutional neural networks (CNN) are applied, for example, in [Da19], but also in **grid-based** approaches, such as Chargrid [Ka18], who represent invoice documents as a grid of characters and predict segmentation masks and bounding boxes for several key classes using a fully-convolutional encoder-decoder architecture. Other grid approaches include *Wordgrid* [De19], *BERTgrid* [DR19], *ViBERTgrid* [Li21], or *Tokengrid* [Ye22]. Another approach is presented by [Ma20], who use representation learning to identify specific key-value-pairs on invoices. For each value element, candidates are selected based on predefined data types and a priori knowledge (“*business rules*”), which are then compared to the corresponding key via a multi-stage embedding.

In **graph-based** approaches with Graph Convolutional Networks (GCN), entities (usually words or tokens) are transformed into a graph structure in which the nodes have relationships to each other via edges [LBB19], [Li19a], [Kr21]. Both nodes and edges can obtain features. [Xu20b] and [Ga21] present **transformer architectures**. LayoutLM, for example, an extension of the BERT model [De18], integrates 2D position (layout information) and image embeddings (visual information) in the document understanding task. Tab. 1 summarizes state-of-the-art DL approaches. The table illustrates, that while the approaches all address a similar extraction problem, they are difficult to compare against each other, w.r.t. the model, dataset, class labels and metrics used. For example, [Kr21] report a macro average F1-score of 0.875 for the classes ‘*invoice date*’, ‘*invoice number*’ and ‘*total amount*’, excluding, and 0.905 including the ‘*no label*’ class. [Ma20] report a macro-average F1-score of 0.878 for seven key items in their invoice dataset. [Li19] report F1-scores for six key items for two different datasets (Chinese, English), whereas [Xu20b] report an F1-score of 0.952 for the SROIE² dataset. Evidently, the approaches deploy different pipelines. While this leads to limited comparability, they share the assumption that context plays an important role in the recognition of key information on invoices [Kr21].

² Scanned Receipts OCR and Information Extraction Dataset: <https://rrc.cvc.uab.es/?ch=13>

Reference	Approach	Segmentation	Number of Classes: Keys
[Kr21]	GCN/GAT	Entity-level node classification	3: Invoice number, Invoice date, Total amount
[Li19a]	GCN/GAT BiLSTM-CRF	Entity-level node classification	6: Invoice number, Date, Price, Tax, Buyer, Seller
[LBB19]	GCN	Word-level node classification	27: Invoice number, Invoice date, Total without tax, Total tax amount, ...
[Xu20b]	LayoutLM	Token-level classification	FUNSD: 4 (question, answer, header, other) SROIE: 4 (company, date, address, total)
[Xu20a]	LayoutLMv2	Token-level classification	CORD: 4: menu, void menu, subtotal, total Kleister NDA: 4: date, jurisdiction, party, term
[Ma20]	Self-attention Network	Token-wise candidate similarity	7: Amount due, Due date, Invoice date, Invoice ID, Purchase order, Total amount, Total tax amount
[Ka18]	Chargrid	Character-wise segmentation	8: Invoice number, Amount, Date, Vendor name, Address, Line item description, Quantity, Amount
[DR19]	BERTgrid	Character-wise segmentation	6: Amount, Number, Date, Vendor name, LI mean, LI quantity
[Ye22]	Tokengrid	1D Anchor & Heatmap	12: Vendor/Importer address & name, Freight, Insurance, Invoice date & number, Product, Prices
[Ya17]	MFCN (fully conv. network)	Pixel-wise segmentation	6 format segments (figure, table, section heading, caption, list, paragraph)

Tab. 1: Selection of deep learning approaches in invoice recognition

Practice Review. The literature review shows state-of-the-art approaches from science, some of which are already being used in practice. [KAD04] study various suppliers for automated invoice processing. However, we need to account for technical progress since 2004. Therefore, we perform a Google search with the aforementioned search terms. We find a large number of providers of document processing software, API services, and receipt scanning apps (e.g., ‘*scan2bank*’ services from banks), which we cannot present here all. *Klippa*, *Nanonets*, *Kofax*, *Super.AI*, *Rossum*, *MS Azure Form Recognizer* or *Veryfi* represent only a selection that offer intelligent invoice document processing. In the latter various key-value-pairs can be extracted through a Graph Neural Network. *Google*³, offers a whole suite within their *Document AI*, e.g., *Docsumo*, or Human-in-the-loop training. Among other things, they provide services such as table recognition, optical character recognition (OCR) [Ne15], and (non-finite) key-to-value matching. The aforementioned services are not open-source and not free of charge. We sample test documents, for providers that offer to test one file. From this (non-standardized) test no conclusions can be drawn on transferable performance. Generally, there is hardly any (scientific) record of the deployed pipelines and their performance available. That is, the details and performance of these approaches are difficult to validate.

³ <https://cloud.google.com/document-ai>

3 Design Challenges in our Application Context

3.1 Information Types

Based on our literature review, we propose to distinguish seven different basic types of information on invoices, namely **segmental**, **syntactic**, **semantic**, **spatial**, **external**, **graphical**, and **logical** information, which are not equivalent to the type and number of classification labels, such as the class label *'total amount'*. They rather represent the underlying signal of the data. The categorization was performed through systematic grouping of themes within the literature results. **Segmentation** refers to the level of abstraction at which data from invoices are processed. These levels can also be thought of as a hierarchy. On the lowest level, images are represented by pixels using the RGB encoding [Gu14]. Depending on the language, individual symbols/characters/letters, words and sentences constitute further levels of abstraction. Tokens differ from words in the sense that they may represent only a part of a word, e.g., the word stem, or even only a single character. Furthermore, paragraphs, rows, columns and fields, if necessary, pages can be segmented. **Syntactic** information form an intermediate stage between segmentation and semantics. For example, they can be useful to label individual entities by their data type as string, digit, or alphanumeric. Upper and lower case or the number of letters or special characters can also be included as information. We refer to this category as syntax, because this information reflects rules for different entity types. For instance, a regular e-mail address requires three parts, a local name, an @ sign, and a server name with its top-level-domain. These three parts therefore represent the syntax. The syntactic category also includes so-called tagging, e.g., POS (Part-of-Speech) [FHM14], IOB- (Inside-Outside-Beginning) [TV99], or NER- (Named Entity Recognition) [NS07] tagging. While these tags can of course represent semantic information (logical semantics) [Pa99], e.g., whether a word is tagged as a noun or a verb, they mostly provide structure (grammar respectively). In terms of semiotics, syntax stands for the sign (and the rules for their concatenation) and semantics stands for the norm, i.e., the conceptual interpretation of a sign or a series of signs [St00]. Thus, we distinguish syntax from **semantics** by integrating so-called language or word models that transfer tokens, words, and sentences into a vector space (embedding), spanning a cognitive map [St93]. This part of semantics (lexical semantics) [Pa99] is concerned with the meaning of words and their relations. Semantically similar entities are located closer together in the vector space, indicated by their cosine similarity. Popular word models represent, for example, *Word2Vec* [Mi13] or *GloVe* [PSM14]. Another word model that has been increasingly used in recent years is *BERT*, which has been developed and open-sourced by *Google*⁴. It is available in a wide variety of variants, e.g., *RoBERTa* [Li19b] or *DistilBERT* [Sa19]. Its strength lies primarily in analyzing words in the context of sentences and thus extracting the semantic. An example of different meaning, that only arises in the context are homonyms, e.g., the word *drop* in the sentences *"I hope I don't drop my cup of coffee"* vs. *"I enjoyed every last drop"*

⁴ <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>

of my coffee”. In the invoicing context, subtle linguistic nuances and abbreviations emerge that are not necessarily used in everyday speech or prose, e.g., *Inv.#.* could mean *‘invoice number’*. These semantic characteristics have to be accounted for. Whereas syntax and semantics can be found in the NLP context as well, 2D spatial information add a new information. The **position** can be integrated explicitly (Cartesian coordinates) or implicitly through pixel position. Absolute and relative position of elements like the coordinates of corners or vertices of bounding boxes represent a significant signal. In addition, center points, areas and shapes can be used. Distance measures between elements, i.e., the Euclidean distance, can be applied to establish positional neighborhood relations [LBB19], e.g., the left, right, top and bottom neighbor of a word. Content neighborhoods are also conceivable through *‘semantic linking’* [Xu20b]. **Graphical** information are images such as logos, icons or product pictures, or visual elements such as table frames, boxes, arrows, dashed or dotted lines. Especially the latter can strongly contribute to the structure of the document. In addition, fonts, font styles and font sizes are also graphical information, e.g., an underlined, bold or highlighted element can have a special meaning in the document [Ne15]. Also, white-space can convey information. **External** information represent knowledge which helps to interpret the content of invoices in the form of data bases [LBB19] or knowledge graphs [Ma20]. External information can be distinguished in content and format-oriented elements. Format-oriented information are conventions for certain entities like tax or account numbers (IBAN), which are often country specific, e.g., five-digit ZIP codes. Content oriented external information can be drawn from data bases or self-generated dictionaries on certain topics [LBB19]. For example, order numbers from a company’s ERP system. The final category here is **logical** information [Ya17], [SAT03]. They are a result of the context and/or the interaction of several information. To illustrate that, we investigate the token *‘4.07’*. Syntactically it could be a decimal number or a date. Semantically it is ambiguous, it can express both a quantity or a currency amount. Only in combination of the currency sign (*‘\$4.07’* or *‘4,07€’*) the content is specified in this trivial example. Logic can also refer to calculations, e.g., the calculation of a gross amount from net amount and percentage tax amount. Another example of logic would be missing and multiple values, e.g., if an invoice states *“billing address same as shipping address”*, we could logically infer to fill both keys with the same value.

3.2 Pipeline Challenges

The information types form the substantive basis for the generation of input features from raw data for ML/AI pipelines. Several challenges arise in the development of such a pipeline. For example, the correct classification of single elements does not equate to the retrieval of the full information. This means, that the pre- and post-processing of labels and entities belongs to the extraction pipeline, as label data and raw data are usually not stored in the same location (Fig. 1). To elucidate this: the ground truth could be stored in a json file like this { *‘invoice date’*: *‘7/04/2022’* }. The corresponding OCR could yield four independent tokens: *‘July’*, *‘4’*, *‘th’*, *‘2022’*. Now, assuming correct positional data, we could assign the label *‘invoice date’* to all four tokens. If our model correctly classifies all

four tokens as ‘*invoice date*’, we achieve great performance, however, we would still need to perform post-processing steps to assemble the final value, here for example, connecting tokens of the same line. Since many scientific articles focus dataset performance and do not specify pre- and post-processing procedures [Kr21], [Ka18], [Xu20b], it is difficult to validate the deployed pipeline, in terms of full information retrieval.

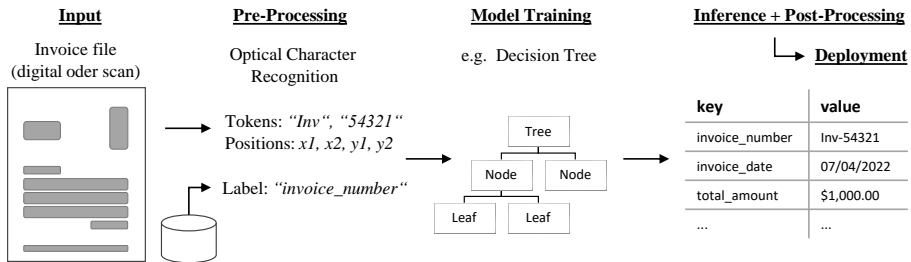


Fig. 1: Exemplary pipeline of invoice information extraction, author’s figure

In order to specify quality standards and comparison criteria in the future design and comparison of extraction pipelines, we distinguish here four categories of challenges: **Document-, Data-, Model- and Assembly Challenges** for ML pipelines in invoice recognition. While Document and Data Challenges refer to quality criteria, Model and Assembly Challenges are usually consequences of design decisions. Due to the fact that there are a variety of factors influencing the overall pipeline in these four categories, it is difficult to determine or to rank the influence of any single parameter. However, different quality criteria and design decisions influence and potentially reinforce each other. This interdependency may propagate errors through the pipeline, e.g., poor scan quality can lead to incorrectly recognized characters, which limits the tokenization quality. Although digitally created and sent documents have a great advantage over paper-based documents as they eliminate the factor of scan quality, and therefore generally provide better OCR results (or make them obsolete), they still face downstream challenges in the pipeline, such as adequate label quality. Document Challenges include the image/scan/file quality as well as the syntactic and semantic quality. The former is due to image noise, fading colors, or skewness [YCA20]. Syntactic and semantic quality refer to errors that arise when the invoice is created, e.g., spelling mistakes or incorrect abbreviations. Even if these errors are out of the scope of the actual invoice recognition task, they can still influence training and inference in ML pipelines. In the Data Challenges category we distinguish between text quality, visual information, tokenization, and label quality. Since we cannot assume that underlying meta data (i.e., words, labels, positions) of the documents are available and we cannot mine the content stream of the pdf file, an electronic conversion of the text must be performed by OCR engines, e.g., *Tesseract*, *Paddle*, *Abbyy*, *MS Azure Read*, *Google Cloud Vision* [Ta16]. If the pixels are not translated into the correct characters, words, sentences, or paragraphs, with corresponding bounding boxes, difficulties arise in the subsequent process. For pure computer vision approaches, the OCR step is omitted since only bounding boxes and class labels are needed. For graphical information, the

main challenge is to translate them into features, i.e., how to deal with table margins or white spaces, which give the document an essential structure, but are difficult to standardize in terms of data processing. Token quality refers primarily to pipelines that rely on language models. If the tokens are too small or big, information may be lost. Label quality refers to a major problem in data storage. Usually, image data and label data are stored separately, i.e., an assignment of labels to corresponding elements has to be programmed first. In addition, it may happen that several values are assigned to a single key or several keys, e.g., *'invoice date'* and *'shipping date'* receive the same value. That means, that the assignment could be ambiguous. The labels themselves can be inaccurate, as they usually stem from manual annotation. An essential quality criterion of the data is therefore the existence of an accurate ground truth, which also influences the number and kind of classification labels. Another challenge is the inherently strong class imbalance, i.e., the valuable information are usually contained in minority classes, which increases the chance of misclassification. Model Challenges. In ML features are used to represent the learning task. Feature engineering refers to the transformation steps and enrichment of the raw data, e.g., normalization or the procedure that calculates neighborhood relations [LBB19]. In general, converting raw data into model inputs can lead to information loss. The selected model of course influences the feature engineering, since the inputs must be prepared exactly as the model expects them to be, e.g., for pixel models with segments, segmentation masks must be created; or graph representations for graph models. The model itself then specifies how the corresponding outputs look like, e.g., probabilities per class. Furthermore, the integration of pre-trained models and the determination of hyperparameters also influence the pipeline. In addition, the characteristics of the dataset must be considered, i.e., the general sample size, the distribution of the classes, and the variance of layouts and tokens. Here, model and data challenges overlap. Lastly, attention must be drawn to challenges in pre- and post-processing [Gr19]. Depending on the outputs of the model, only fragments of the actual information are predicted. Assembly quality means how well abstract predictions at the entity level are converted into complete information. e.g., by defining pre- and post-processing heuristics. For example, it could be useful to assemble a numeric date from a string-based date.

4 Results and Conclusion

The result of our initial design cycle are the catalogs of information types and pipeline challenges, summarized in Tab. 2 and Tab. 3, which we apply in collaboration with our design process stakeholders (SME industry partner) in subsequent design cycles. The catalogs generally help to better understand the underlying data problem in each extraction case and to optimize the overall pipeline. As they address the *Relevance Cycle* of our design study, they can be used in two beneficial ways. First, they help to compare different approaches, as they offer structuring elements, e.g., whether and which OCR engine is used (Tab. 4). Second, they support the selection of models and data preparation strategies based on available datasets, and vice versa. The following questions are suitable for the application of the framework:

1. What type of document and what language are at hand?
2. Which classes and annotations are available and which outputs are required?
3. Which information types are involved?
4. Which models and which hardware/implementation structure can be used?
5. Which challenges are derived from the constellation of questions 2, 3 and 4?

Regarding the information types, we recommend using Tab. 2 to identify all the types that are relevant to the specific case. Based on model selection, the details and challenges in the implementation of the pipeline can then be identified (Tab. 3). We briefly showcase how we utilize the catalogs to design a pipeline, by means of the examples shown in Tab. 2 and 3. In our present use case, the goal is to develop a model that recognizes information on German invoices. We have a real-world dataset available, consisting of 977 pdf files with rule-based label annotations and OCR extracts for over 60 classes, which provides segmentation on word level. The labels to be classified in the model include among others the *invoice date*, *invoice number*, *total amount*, *payment information*, and also specifics such as the *IBAN* (international bank account number) and *commercial register number*, i.e., digit and character syntax. As for this design cycle we apply a neighborhood algorithm, however, we don't integrate external, graphical and logical information. We also exclude product line items and their details (quantities, amounts) in this cycle. We identify the general size of the dataset, class distributions (minority classes), and the label quality as potential challenges. Regarding the choice of model, we want to leverage the ability to integrate a word model, since we deal with German documents, instead of English documents (as predominant in science).

Inform. Type	Specification examples
Segmental	Pixel, Symbol/Letter/Character, Token/Word, Sentence, Paragraphs/Block/Columns, Pages/Slides
Syntactic	Data type (string, digit, decimal, date, alphanumeric), Case (upper, lower, mixed), Special characters (e.g., @, %), Count, Tagging (e.g., PoS, IOB, NER), character or byte pair encoding
Semantic	Numerical representation (e.g., vectorization) of an entity, based on context provided by a (word) model, such as Word2vec, GloVe, or BERT
Spatial Positional/	Explicit: Bounding box (vertices), coordinates: left, top, width, height, right, bottom (absolute/relative/scaled); Implicit: image pixel structure (Euclidean) Distance, Size, Area, Center points
Neighborhood	Spatial neighborhood, semantic connection, neighbor features, self-loop
External	Dictionaries, e.g., ZIP codes, City names Conventions/ Abbreviations, e.g. Tax ID, IBAN Spell check
Graphical	Font (type, size, style); Tables/Frames Lines, dots, arrows, connectors, boxes, contours Machine-readables (QR, barcode), Images
Logical	Calculations: sums, percentages, quantities, units, rebates, discounts; References

Tab. 2: Information Types with Examples

C1: Document	C3: Model	C4: Assembly
Image/Scan Quality <i>e.g. noise, skewness</i>	Feature Engineering <i>e.g. word model, normalization</i>	Pre/Post-Processing Heuristics <i>e.g. class aggregation, ambiguity</i>
Syntactic/Semantic Quality <i>e.g. false/missing information</i>	Model Design <i>e.g. outputs, hyperparameters</i>	
C2: Data		
Sample Quality <i>e.g. size, distribution</i>	Tokenization Quality <i>e.g. stemming</i>	Label Quality <i>e.g. incorrect annotation</i>
Text Quality <i>e.g. OCR conversion</i>	Graphics Quality <i>e.g. whitespace</i>	

Tab. 3: ML pipeline challenges in four categories

Therefore, the semantic aspect and neighborhood should be explicitly modifiable in the pipeline (Tab. 2). Based on these requirements, we prototypically decide to investigate graph-based models. With regards to our literature review, we consider the graph-based approaches from [Li19], [LBB19], and [Kr21]. We select [Kr21], because it is the most recent one and integrates semantic, syntactic and positional information types. This model was originally trained on a dataset of 1,129 manually labeled English invoices, which is comparable in terms of size. Differences emerge in the language and dataset variance, as only a single recipient was considered there, while we look at various recipients in our dataset. Tab. 4 showcases how the categorization helps to draw a comparison between approaches. Both GCN models include graph attention layers implemented in Pytorch and Deep Graph Library (DGL), which are accessible from open-source. For a detailed description and exact training parameters we refer to [Kr21]. We enhance their approach by integrating a German-BERT⁵ model for semantic features into the pipeline.

Quality category		Krieger et al. (2021) Pytorch, DGL	Adopted Prototype Pytorch, DGL
Doc.	Scan/born digital	n.a.	scanned
Doc.	Date range of invoice	n.a.	2010-2019
Data	# of invoices (vendors/ rec.)	1129 (277/1)	977 (494/531)
Model	Training method and split	supervised 80/20	supervised 80/10/10
Process.	Labeling method	manually	rule-based
Data	File Formats	n.a	.pdf, .json
Process.	OCR engine	Tesseract	Abbyy
Model	Model type	GCN	GCN
Model	Labels excl. 'no_label'	'date', 'number', 'total'	'date', 'number', 'total'
Metric	Test F1-score	0.905	0.823

Tab. 4: Exemplary comparison of approaches

⁵ <https://www.deepset.ai/german-bert>

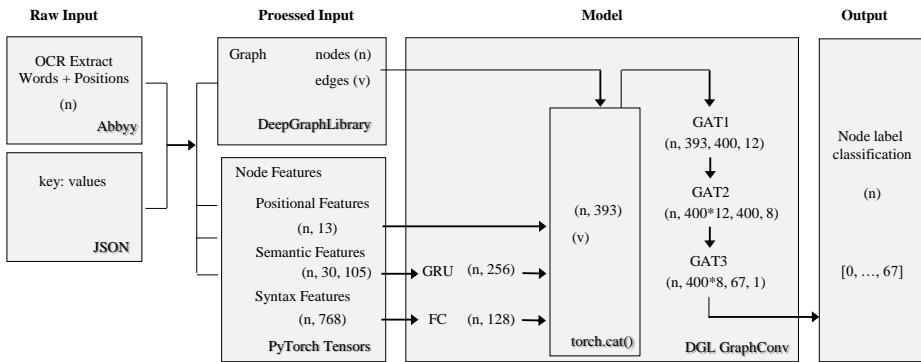


Fig. 2: Prototype pipeline section, author's figure.

Based on initial trainings on our dataset our prototype in this stage of the design study does not achieve the same performance (F1-score of 0.823) as the benchmark model (0.905). The difference could be due to the fact that we use a different word model and a different algorithm for neighborhood generation. Nevertheless, the application of the pipeline and the model results represent a proof of concept for our framework.

Contribution and Future Research. As the prior analysis and detailed (de-) composition of the data problem is an essential step in ML/AI projects, the catalogs are very beneficial tools. They form the morphological basis for the problem analysis and the derivation of design requirements in our specific as well as in general use cases beyond specific datasets and key items. With the help of the framework, model selection and data preparation can be better structured, and it improves the comparability between different approaches. Furthermore, as the classification problem is better understood, it helps to augment and synthesize further training data, which also highlights a path for future research. The novelty of our work is that the specific challenges and information types have not been catalogued before. The application of the framework supports both practice pipeline implementations as well as descriptive process research, into which these kinds of catalogs can be adopted, e.g., in DASC-PM. We validate our results here by demonstrating the selection, comparison, and enhancement of a model pipeline within our design study. We plan to use the framework in our future design path. The framework enables targeted work in the IS community and related disciplines towards more comprehensive document analysis and recognition, as it provides guidance to identify and implement improved pipelines. Especially, in terms of line item recognition many training opportunities remain [Si23], provided that appropriate (labeled) data can be accessed. Too little training data have already been discussed in other papers as a main source of suboptimal results [Li19b]. Future research could leverage our work to rank the categories' influence and develop a quantitative scoring model for extraction pipelines, which can be applied alongside standard performance metrics, such as F1-score. This would further enhance the comparability of different approaches from both research and practice.

Bibliography

- [ASG09] Adali, S.; Sonmez, A. C.; Gokturk, M.: An Integrated Architecture for Processing Business Documents in Turkish. In (Gelbukh, A. Ed.): Computational linguistics and intelligent text processing. 10th international conference. Springer, Berlin, Heidelberg, pp. 394–405, 2009.
- [BB04] Belaid, Y.; Belaid, A.: Morphological tagging approach in document analysis of invoices: Proceedings of the 17th International Conference on Pattern Recognition, 2004., 469–472 Vol.1, 2004.
- [Cr18] Cristani, M. et al.: Future paradigms of automated processing of business documents. *International Journal of Information Management* 40, pp. 67–75, 2018.
- [Da19] Davis, B. et al.: Deep Visual Template-Free Form Parsing. 15th International Conference on Document Analysis and Recognition, 2019.
- [De18] Devlin, J. et al.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2018.
- [De19] Denk, T. I.: Wordgrid: Extending Chargrid with Word-level Information, 2019.
- [DK02] Dengel, A. R.; Klein, B.: A Requirements-Driven System for Document Analysis and Understanding. *DAS 2002*, pp. 433–444, 2002.
- [DR19] Denk, T. I.; Reisswig, C.: BERTgrid: Contextualized Embedding for 2D Document Representation and Understanding. 33rd Conference on Neural Information Processing Systems, Vancouver, Canada., 2019.
- [Es12] Esser, D. et al.: Automatic indexing of scanned documents: a layout-based approach. *SPIE*, pp. 118–125, 2012.
- [Fa04] Fairchild, A.: Using Electronic Invoicing to Manage Cash Forecasting and Working Capital in the Financial Supply Chain. *ECIS 2004 Proceedings*, 2004.
- [FHM14] Farahmand, H.; Harounabadi, A.; Mirabedini, S. J.: Document features selection using background knowledge and word clustering technique. *Management Science Letters*, pp. 241–250, 2014.
- [Ga21] Garncarek, Ł. et al.: LAMBERT: Layout-Aware Language Modeling for Information Extraction. *Document Analysis and Recognition - ICDAR 2021* 12821, pp. 532–547, 2021.
- [Gi20] Girasa, R.: *Artificial Intelligence as a Disruptive Technology. Economic Transformation and Government Regulation*. Springer International Publishing; Imprint Palgrave Macmillan, Cham, 2020.
- [Gr19] Grüning, T. et al.: A two-stage method for text line detection in historical documents. *International Journal on Document Analysis and Recognition (IJDAR)* 3/22, pp. 285–302, 2019.
- [Gu14] Gupta, S. et al.: Learning Rich Features from RGB-D Images for Object Detection and Segmentation. *European Conference on Computer Vision (ECCV)*, 2014.
- [He04] Hevner, A. et al.: *Design Science in Information Systems Research*. *MIS Quarterly*

1/28, p. 75, 2004.

- [He07] Hevner, A.: A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems* 19, 2007.
- [Ja12] Janssen, B. et al.: Receipts2Go. In (Concolato, C. Ed.): *Proceedings of the 2012 ACM symposium on Document engineering*. ACM, New York, NY, 2012.
- [Ka18] Katti, A. R. et al.: Chargrid: Towards Understanding 2D Documents: *Proceedings of EMNLP*, 2018.
- [KAD04] Klein, B.; Agne, S.; Dengel, A.: Results of a Study on Invoice-Reading Systems in Germany. In (Hutchison, D. Ed.): *Document analysis systems VI. 6th international workshop, DAS 2004, Florence, Italy, September 8 - 10, 2004 ; proceedings*. Springer, Berlin, pp. 451–462, 2004.
- [KAM95] Koonce, L.; Anderson, U.; Marchant, G.: *Justification of Decisions in Auditing*. *Journal of Accounting Research* 2/33, p. 369, 1995.
- [Kr21] Krieger, F. et al.: Information Extraction from Invoices: A Graph Neural Network Approach for Datasets with High Layout Variety. *Wirtschaftsinformatik 2021 Proceedings*, 2021.
- [LBB19] Lohani, D.; Belaïd, A.; Belaïd, Y.: *An Invoice Reading System Using a Graph Convolutional Network*. Springer, Cham, pp. 144–158, 2019.
- [Li19a] Liu, X. et al.: *Graph Convolution for Multimodal Information Extraction from Visually Rich Documents*, 2019.
- [Li19b] Liu, Y. et al.: *RoBERTa: A Robustly Optimized BERT Pretraining Approach*, 2019.
- [Li21] Lin, W. et al.: ViBERTgrid: A Jointly Trained Multi-Modal 2D Document Representation for Key Information Extraction from Documents: *Proceedings of ICDAR 2021*, 2021.
- [Ma20] Majumder, B. P. et al.: *Representation Learning for Information Extraction from Form-like Documents*. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 6495–6504, 2020.
- [Mi13] Mikolov, T. et al.: *Efficient Estimation of Word Representations in Vector Space*, 2013.
- [Ne15] Nell, H.: *Quantifying the noise tolerance of the OCR engine Tesseract using a simulated environment*. Master Thesis, Karlskrona, Sweden, 2015.
- [NS07] Nadeau, D.; Sekine, S.: *Named Entities: Recognition, classification and use*. *Linguisticae Investigationes* 1/30, pp. 3–26, 2007.
- [Pa99] Partee, B. H.: *Semantics*. *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA, pp. 739–742, 1999.
- [Pe07] Peffers, K. et al.: *A Design Science Research Methodology for Information Systems Research*. *Journal of Management Information Systems* 3/24, pp. 45–77, 2007.
- [PLW19] Palm, R. B.; Laws, F.; Winther, O.: *Attend, Copy, Parse -- End-to-end information extraction from documents*. *ICDAR*, 2019.

- [PSM14] Pennington, J.; Socher, R.; Manning, C.: Glove: Global Vectors for Word Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1532–1543, 2014.
- [PWL13] Palm, R. B.; Winther, O.; Laws, F.: CloudScan - A Configuration-Free Invoice Analysis System Using Recurrent Neural Networks: 12th International Conference on Document Analysis and Recognition, pp. 406–413, 2013.
- [Ra21] Rahim, M. A. et al.: RFM-based repurchase behavior for customer classification and segmentation. *Journal of Retailing and Consumer Services* 61, p. 102566, 2021.
- [Sa19] Sanh, V. et al.: DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, 2019.
- [SAT03] Song Mao; Azriel Rosenfeld; Tapas Kanungo: Document structure analysis algorithms: a literature survey. *SPIE*, pp. 197–207, 2003.
- [Sc13] Schuster, D. et al.: Intellix -- End-User Trained Information Extraction for Document Archiving: 12th International Conference on Document Analysis and Recognition, pp. 101–105, 2013.
- [Sc22] Schulz, M. et al.: DASC-PM v1.1 A Process Model for Data Science Projects. *Universitäts- und Landesbibliothek Sachsen-Anhalt*, 2022.
- [Si18] Singh, S.: *Natural Language Processing for Information Extraction*, Australia, 2018.
- [Ši23] Šimsa, Š. et al.: DocILE Benchmark for Document Information Localization and Extraction, 2023.
- [St00] Stamper, R. et al.: Understanding the roles of signs and norms in organizations - a semiotic approach to information systems design. *Behaviour & Information Technology* 1/19, pp. 15–27, 2000.
- [St93] Stamper, R.: *A semiotic theory of information and information systems*, Enschede, 1993.
- [Ta16] Tafti, A. P. et al.: OCR as a Service: An Experimental Evaluation of Google Docs OCR, Tesseract, ABBYY FineReader, and Transym. In (Bebis, G. Ed.): *Advances in visual computing. 12th international symposium, ISVC 2016*. Springer, Cham, pp. 735–746, 2016.
- [TV99] Tjong Kim Sang, E.; Veenstra, J.: Representing Text Chunks. Ninth Conference of the European Chapter of the Association for Computational Linguistics, pp. 173–179, 1999.
- [WH00] Wirth, R.; Hipp, J.: CRISP-DM: Towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 2000.
- [Xu20a] Xu, Y. et al.: LayoutLMv2: Multi-modal Pre-training for Visually-Rich Document Understanding, 2020.
- [Xu20b] Xu, Y. et al.: LayoutLM: Pre-training of Text and Layout for Document Image Understanding, 2020.
- [Ya17] Yang, X. et al.: Learning to Extract Semantic Structure from Documents Using

Multimodal Fully Convolutional Neural Network, 2017.

- [YCA20] Yindumathi, K. M.; Chaudhari, S. S.; Aparna, R.: Analysis of Image Classification for Text Extraction from Bills and Invoices: 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, Piscataway, New Jersey, 2020.
- [Ye22] Yeghiazaryan, A. et al.: Tokengrid: Toward More Efficient Data Extraction From Unstructured Documents. IEEE Access 10, pp. 39261–39268, 2022.

Kriterien und Framework für intelligente Digitalisierung aus soziotechnischer Perspektive – zur Diskussion

Florian Allwein ¹, Maik H. Wagner²

Abstract: Die Digitalisierung in Deutschland kommt nur langsam voran; gleichzeitig bieten sich vielversprechende neue Möglichkeiten durch die Nutzung von Large Language Models. Das Konzept der intelligenten Digitalisierung ist vorgeschlagen worden, um den geeigneten Einsatz von Technologie zu erforschen und zu diskutieren, wie man ‚die richtigen Dinge‘ tun kann. Dieses Konzept ist jedoch noch nicht ausgearbeitet. Wir leiten eine Definition her, skizzieren ein Framework für intelligente Digitalisierung in soziotechnischen Informationssystemen und definieren Kriterien für eine intelligente Digitalisierung. Dies soll als Startpunkt für eine Diskussion des Konzepts dienen und Impulse geben, um es in der Fachöffentlichkeit zu etablieren.

Keywords: Intelligente Digitalisierung, Digitale Transformation, soziotechnische Informationssysteme

1 Einleitung

Neue Technologien wie Large Language Models (LLM), auf denen Anwendungen wie ChatGPT basieren, erhöhen das Potenzial für erhebliche Veränderungen und versprechen die Möglichkeit, eine neue Stufe der Digitalen Transformation einzuleiten [Dw23]. Gleichzeitig tun sich Wirtschaft und Verwaltung in Deutschland schwer, auch nur die bestehenden Potenziale zu nutzen. Um die Digitalisierung voranzutreiben, hat die EU für die 2020er Jahre die „digitale Dekade“ ausgerufen [Gi22]. Einstweilen findet sich Deutschland im EU-„Index für die digitale Wirtschaft und Gesellschaft“ auf Rang 13 von 27 Mitgliedsländern [Eu22].

Diese beschleunigte Digitalisierung erhöht gleichzeitig die Unsicherheit bei den Verantwortlichen. Wir folgen dem Aufruf von [Jo22], den „geeignete[n] Einsatz von KI [und Technologie allgemein] für alle Bereiche von Wirtschaft und Gesellschaft“ (S. 1696) zu erforschen und zu diskutieren, wie man „möglichst unverzüglich und transparent ‚die richtigen Dinge‘“ (ebd.) tun kann, um „Intelligente Digitalisierungen auf den verschiedenen Ebenen zu begünstigen“ (S. 1697). Dabei stellen wir die folgenden Forschungsfragen: „Was ist intelligente Digitalisierung?“ und: „Wie können intelligente Digitalisierungsprojekte geplant werden?“. Um eine Diskussion mit interessierten Wissen-

¹ IU Internationale Hochschule, Professur für Digitale Transformation, Juri-Gagarin-Ring 152, D-99084 Erfurt, florian.allwein@iu.org,  <https://orcid.org/0000-0002-2831-7259>

² MW Wissenskommunikation, Merianplatz 4, 60316 Frankfurt, mw@wissenskommunikation.eu

schaftler:innen und Praktiker:innen anzustoßen, schlagen wir eine Definition von intelligenter Digitalisierung vor und skizzieren ein Framework und Kriterien für intelligente Digitalisierung.

2 Hintergrund

2.1 Digitalisierung und Digitale Transformation

Wir verstehen Digitalisierung als „Übertragung von Aufgaben auf den Computer, die bisher vom Menschen übernommen wurden“ bzw. „Einführung digitaler Technologien in Unternehmen und als Treiber der digitalen Transformation“ [He19] und Digitale Transformation als „erhebliche Veränderungen des Alltagslebens, der Wirtschaft und der Gesellschaft durch die Verwendung digitaler Technologien und Techniken sowie deren Auswirkungen“ [Po19].

2.2 Aspekte einer intelligenten Digitalisierung

Die Psychologie definiert Intelligenz als „die Fähigkeit, aus Erfahrung zu lernen, Probleme zu lösen und das Wissen zur Anpassung an neue Situationen einzusetzen“ [My14, S.400]. Das beinhaltet insbesondere die Fähigkeit, Informationen wahrzunehmen und ggf. als Wissen zu speichern, auf die der englische Begriff *intelligence* abzielt.

Eine klare Definition des Begriffs ‚intelligente Digitalisierung‘ konnte in der Literatur nicht gefunden werden. Es finden sich jedoch zahlreiche Kriterien und Frameworks, wie bei der Digitalisierung vorgegangen werden kann [z.B. ABB22, Sc21, Vi20]. Die Gesellschaft für Informatik nennt etwa als Kriterien für eine gelungene Digitalisierung eine klare Zieldefinition und Zweckorientierung, ein ernst gemeintes Neudenken von Organisationen, Strukturen und Prozessen, die frühzeitige Prüfung und Berücksichtigung von Mindestanforderungen an digitale Systeme oder Technologien und geeignete Prozesse für Entwicklung und Change-Management [Me23].

Allerdings sind Digitalisierungsprojekte in der Praxis oft nicht erfolgreich. Die Beiträge eines Sammelbands zur Digitalisierung in Deutschland [Gi22] nennen etwa unzureichende technische Infrastruktur, Fachkräftemangel und einen nötigen Wandel der Führungskultur als Hindernisse für Digitalisierungsprojekte. Friesike und Sprondel gehen davon aus, dass die Verantwortlichen für solche Projekte tendenziell zu viel wollen – sei es der „Wunsch, den digitalen Wandel durch eine möglichst weitreichende Vision anzustoßen“ [FS21, S.17] oder eine „ausgeprägte Technikhörigkeit, die davon ausgeht, dass moderne Technologien so potent sind, dass ihre bloße Anwesenheit bereits zu Lösungen führt“ (ebd.). Unabhängig davon gibt es in der Literatur eine lange Geschichte gescheiterter IT-Projekte [Dw15].

2.3 Soziotechnische Ansätze

Wir konzipieren Informationssysteme schwerpunktmäßig als soziotechnische Systeme, wie es in der angelsächsischen Tradition des *Information Systems research* üblich ist. Auch in der deutschen Tradition der Wirtschaftsinformatik ist dieser Ansatz mindestens seit den 1990er Jahren verbreitet [Wk94].

Wir gehen also davon aus, dass beim Planen und Entwickeln von Informationssystemen neben der technischen Seite auch die menschliche und soziale Seite der Systeme diskutiert und gestaltet werden muss. Informationssysteme werden durch die Interaktion zwischen Technologie, Menschen und Organisationen als soziotechnische Systeme konstituiert [AC93]. Als eine praktische Anwendung schlägt etwa die *Soft Systems Methodology* vor, dass vor der Implementierung eines Informationssystems die gesamte Situation einschließlich persönlicher, organisatorischer oder politischer Konflikte reflektiert wird. Daraufhin kann ein System postuliert und mit den Beteiligten diskutiert werden. In jüngerer Zeit weisen [Ka19] auf die bedeutende Rolle von Menschen in Projekten zur digitalen Transformation hin.

3 Framework für intelligente Digitalisierung

3.1 Intelligente Digitalisierung

Ausgehend von den vorgestellten Konzepten definieren wir intelligente Digitalisierung als Digitalisierungsaktivitäten, die auf Basis eines angemessenen Maßes an Wissen über die Kontexte und Umwelten digitale Veränderungen herbeiführen, wodurch die Veränderungen mit den Kontexten und Umwelten adaptierbar sind. Intelligente Digitalisierung erkennt also an, dass Digitalisierungsinitiativen nicht im „luftleeren“ Raum passieren, sondern in soziotechnische Systeme eingebettet sind und diese beeinflussen. Das Informationssystem als soziotechnisches System bildet in diesem Sinne die unmittelbare Umwelt und den unmittelbaren Kontext der digitalen Technologie.

3.2 Framework

Um diese Zusammenhänge zu verdeutlichen, schlagen wir ein Modell vor, das helfen kann, Digitalisierungsinitiativen intelligenter zu steuern (Abb. 1). Digitale Technologien werden in unserer Herangehensweise als eine Komponente eines Informationssystems aufgefasst. Diese technologische Komponente steht im Zusammenspiel mit der Komponente Menschen, die von der Technologie betroffen sind, und der Komponente Organisation, in der die Technologie zum Einsatz kommt. In der Komponente Technologie vollzieht sich die digitalisierte Datenverarbeitung. Die technologische Komponente speist Daten in das soziotechnische System ein, die dort als Informationen weiterverarbeitet werden. Zur Steuerung der Adaptionsprozesse müssen Priorisierungen vorgenommen

werden. Diese Priorisierungen orientieren sich an den Unternehmenszielen bzw. an Digitalisierungszielen, die aus den Unternehmenszielen abgeleitet werden sollten.

Dieses Modell kann dabei helfen, die Zusammenhänge und Abhängigkeiten, die bei Digitalisierungsaktivitäten betroffen sind, besser in den Blick zu bekommen und zu steuern.

Modellhaft nehmen wir an, dass das System stabil funktioniert, bevor eine Veränderungsinitiative startet. Durch die Veränderung der Komponente Technologie gerät das Informationssystem in Instabilität. Die Interaktionen zwischen den Komponenten müssen neu adaptiert werden. Das System muss modifiziert werden, damit Effizienzsteigerungen, die durch eine neue Technologie möglich sind, auch tatsächlich realisiert werden. Die Orientierung an den Geschäftszielen (bzw. an daraus abgeleiteten Digitalisierungszielen) ermöglicht es Entscheider:innen, in einem Change-Prozess Prioritäten festzulegen. Eine neu zu implementierende Technologie hat zum Zweck, dass das Informationssystem effizienter läuft und einen (messbaren) Beitrag zur Erreichung der Geschäftsziele leistet. Der Change-Prozess sorgt dafür, dass die adaptierten Interaktionen zwischen den Komponenten im System stabil und effizient laufen. Die neue Technologie wurde intelligent in das Informationssystem implementiert, weil die Systemabhängigkeiten reflektiert und adaptiert worden sind.

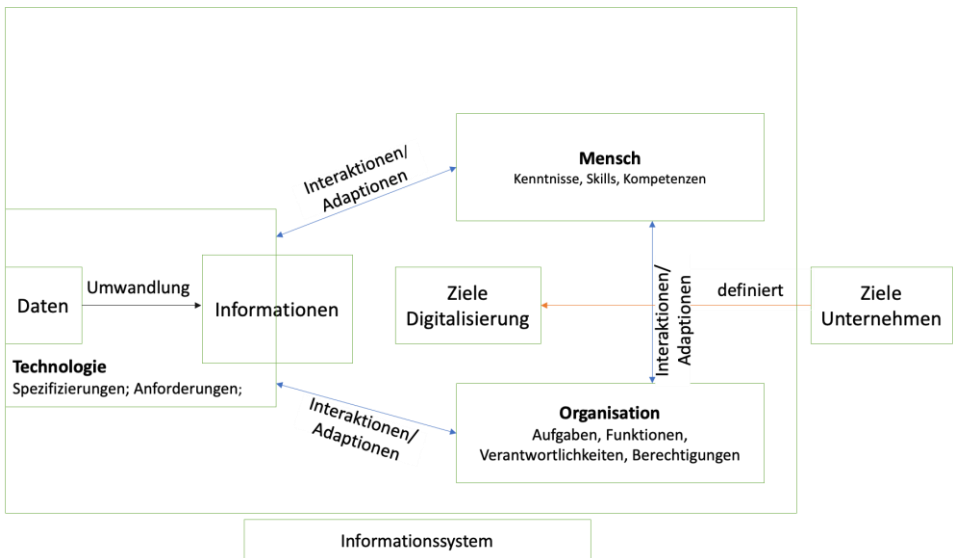


Abbildung 1 Framework: Intelligente Digitalisierung

3.3 Kriterien

Konkret schlagen wir vor, dass Digitalisierung als intelligent bezeichnet werden kann, wenn mindestens drei Kriterien erfüllt sind:

1. In die Planung von Digitalisierungsinitiativen ist ausreichend Wissen über die Komponenten des soziotechnischen Informationssystems eingegangen: Intelligenz als Einsatz von Wissen.
2. Bei der Implementierung und beim Betrieb der neuen Technologie wird darauf geachtet, dass die drei Komponenten zueinander neu ausgerichtet werden und dementsprechend weiterentwickelt werden (z.B. Usability; Ableitung der technischen Anforderungen aus den Bedarfen der Organisation; Upskilling der Mitarbeitenden; Definition neuer Verantwortlichkeiten und Zugriffsrechte etc.): Intelligenz als Adaption an Umwelten.
3. Das Ziel der Digitalisierungsinitiative wird aus den Zielen der Gesamtorganisation abgeleitet (z.B. welchen Betrag liefert die Digitalisierung zu den Unternehmenszielen?), damit bei der Adaption der drei Komponenten ein Maßstab für die Ressourcen-Allokation vorhanden ist (z.B. braucht das Unternehmen mehr Investitionen für das Upskilling der Leute oder muss die Ablauforganisation verändert werden?): Intelligenz als kluges Entscheiden.

4 Fazit

Ausgehend von den Forschungsfragen („Was ist intelligente Digitalisierung?“ und: „Wie können intelligente Digitalisierungsprojekte geplant werden?“ haben wir intelligente Digitalisierung definiert als „Digitalisierungsaktivitäten, die auf Basis eines angemessenen Maßes an Wissen über die Kontexte und Umwelten digitale Veränderungen herbeiführen, wodurch die Veränderungen mit den Kontexten und Umwelten adaptierbar sind“. Darüber hinaus haben wir ein Framework für intelligente Digitalisierung in Informationssystemen skizziert und drei Kriterien für intelligente Digitalisierung definiert.

Diese kurze Abhandlung kann allenfalls als Startpunkt für eine Diskussion mit interessierten Wissenschaftler:innen und Praktiker:innen dienen. Wir hoffen aber, damit wertvolle Impulse zu geben, um das Konzept der intelligenten Digitalisierung in der Fachöffentlichkeit zu etablieren. Aufgrund der Möglichkeiten, die Large Language Models bieten, glauben wir, dass diese Diskussion höchst zeitgemäß und relevant ist.

Literaturverzeichnis

- [ABB22] Aerssen, B.; Buchholz, C.; Burkhardt, N.: Das große Handbuch Digitale Transformation: 222 Methoden und Instrumente für mehr Wandlungsfähigkeit im Unternehmen. Vahlen, Franz, 2022.
- [AC93] Avgerou, C.; Cornford, T.: Developing Information Systems: Concepts, Issues and Practice. Macmillan, 1993.
- [Dw23] Dwivedi, Y.; Kshetri, N.; Hughes, L.; Slade, E.; Jeyaraj, A.; Kar, A.; Baabdullah, A.; Koohang, A. u. a.: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative

conversational AI for research, practice and policy. *International Journal of Information Management* 71 (2023), S. 102642, 2023.

- [Dw15] Dwivedi, Y.; Wastell, D.; Laumer, S.; Henriksen, H.; Myers, M.; Bunker, D.; Elbanna, A.; Ravishankar, M. u. a.: Research on information systems failures and successes: Status update and future directions. *Information Systems Frontiers* 17 (2015), S. 143–157, 2015.
- [Eu22] European Commission: The Digital Economy and Society Index — Countries' performance in digitisation, <https://digital-strategy.ec.europa.eu/en/policies/countries-digitisation-performance>, Stand: 11.5.2023.
- [FS21] Friesike, S.; Sprondel, J.: *Träge Transformation. Welche Denkfehler den digitalen Wandel blockieren.* Reclam Verlag, 2021.
- [Gi22] Gifford, A.: *Die digitale Dekade: Wie wir unsere Wirtschaft transformieren können.* Frankfurter Allgemeine Buch, 2022.
- [He19] Hess, T.: Digitalisierung, <https://wi-lex.de/index.php/lexikon/technologische-und-methodische-grundlagen/informatik-grundlagen/digitalisierung/>, Stand 25.5.2023.
- [Jo22] John, U.: Intelligentes Unternehmen, Intelligenter Staat, Intelligentes Europa – Was ist zu tun? – Anstoß einer kollegialen Experten-Diskussion. In (Demmler, D.; Krupka, D.; Federrath, H., Hrsg.): *Informatik 2022.* Gesellschaft für Informatik, S. 1695-1698, 2022.
- [Ka19] Kane, G.; Phillips, A.; Copulsky, J.; Andrus, G.: *The Technology Fallacy: How People Are the Real Key to Digital Transformation.* MIT Press, Cambridge, Mass, 2019.
- [Me23] Messer-Schmidt, K.; Kipker, D.; Allwein, F.; Winter, C.: *Gelungene Digitalisierung: Kriterien,* <https://gi.de/netzwerk/fachgremien/praesidiumsarbeitskreise/pak-digitalisierung/detailseite/digkriterien>, Stand 11.5.2023.
- [My14] Myers, D.: *Psychologie,* Springer, 2014.
- [Po19] Pousttchi, K.: *Digitale Transformation.* <https://wi-lex.de/index.php/lexikon/technologische-und-methodische-grundlagen/informatik-grundlagen/digitalisierung/digitale-transformation/>, Stand 25.5.2023.
- [Sc21] Schallmo, D.; Rusnjak, A.; Anzengruber, J.; Werani, T.; Lang, K. (Hrsg.): *Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente und Best Practices, Schwerpunkt Business Model Innovation.* Springer Fachmedien, Wiesbaden, 2021.

- [Vi20] Viaene, S.: Digital Transformation Know How: Connecting Digital Transformation, Agility and Leadership. Acco, Leuven, 2020.
- [Wk94] WKWI: Profil der Wirtschaftsinformatik. Ausführungen der Wissenschaftlichen Kommission der Wirtschaftsinformatik. Wirtschaftsinformatik 36/94, S. 80–81, 1994.

Wärmewende in einer Großstadt

Entwicklung und Anwendung eines zentralen Planungs- und Steuerungsinstruments


Maik Günther ^{1,2} und Guido Bayard³

Abstract: Der Wärmesektor hat in Deutschland einen signifikanten Anteil an den Treibhausgasemissionen. Bis spätestens 2045 sollen Heizwärme und Warmwasser ausschließlich aus erneuerbaren Energien und aus unvermeidbarer Abwärme erzeugt werden. Den Kommunen und ihren Stadtwerken kommt bei der Umsetzung und Gestaltung der Wärmewende eine entscheidende Rolle zu. Hierfür benötigen sie jedoch zentrale Entscheidungsunterstützungssysteme. Daher haben die Stadtwerke München im Rahmen ihrer Digitalisierungsstrategie mit dem Modell München ein gebäudescharfes Geoinformationssystem entwickelt, um die Wärmewende spartenübergreifend zu planen und umzusetzen. Zudem wurde ein Multiagentensystem für die Simulation von Zukunftsszenarien integriert. Inzwischen wird das Modell München nicht nur bei den Stadtwerken München für zahlreiche Anwendungsfälle eingesetzt, sondern auch von der Landeshauptstadt München für die kommunale Wärmeplanung.

Keywords: digitaler Zwilling, Entscheidungsunterstützungssystem, Geoinformationssystem, kommunale Wärmeplanung, Modell München, Multiagentensystem, Wärmewende

1 Einleitung

Dass die Energiewende nicht nur die Stromerzeugung und -verteilung betrifft, sondern auch den Wärmesektor mit Heizwärme- und Warmwassererzeugung, ist nicht zuletzt durch die Debatten rund um das Gebäudeenergiegesetz (GEG) und das Gesetz zur Wärmeplanung deutlich geworden. Die Transformation des Wärmesektors bis 2045 hin zum ausschließlichen Einsatz erneuerbarer Energien und unvermeidbarer Abwärme kann als Kraftakt bezeichnet werden. Denn laut Bundesverband der Energie- und Wasserwirtschaft (BDEW) werden in Deutschland noch überwiegend Erdgas und Heizöl in Heizungen eingesetzt [BD23]. Zudem sind die Investitionszyklen im Gebäudebereich relativ lang und es fehlt an Geld, an Material, an Planern sowie an Handwerkern.

¹ IU Internationale Hochschule, Juri-Gagarin-Ring 152, 99084 Erfurt, Deutschland, maik.guenther@iu.org 
<https://orcid.org/0000-0002-3926-4383>

² Stadtwerke München GmbH, Emmy-Noether-Str. 2, 80992 München, Deutschland, guenther.maik@swm.de

³ IU Internationale Hochschule, Juri-Gagarin-Ring 152, 99084 Erfurt, Deutschland, guido.bayard@iu.org

Den Kommunen und ihren Stadtwerken kommt bei der Umsetzung und Gestaltung der Wärmewende eine zentrale Rolle zu. So betreiben die Stadtwerke München (SWM) zahlreiche Netze für die Wärmeversorgung in und um München: Erdgas, Fernwärme und Strom. Außerdem verfügen sie über einen Anlagenpark zur Erzeugung von Fernwärme und Strom, wobei perspektivisch der Großteil der Fernwärme durch Tiefengeothermie bereitgestellt werden soll. Im Zuge der Wärmewende sind in Zukunft signifikante Veränderungen bei den jeweiligen Netzen und Anlagen zu erwarten, die vorausschauend und spartenübergreifend geplant werden müssen. Denn eine Änderung in einer Sparte, beispielsweise der Ausbau der Fernwärme in einem Gebiet, kann Auswirkungen auf andere Sparten haben. Im genannten Beispiel ist mit einem Rückgang des Gasabsatzes im betreffenden Gebiet zu rechnen. Zudem dürfte dort der zukünftige Anteil der Wärmepumpen eher gering ausfallen, was wiederum das Stromnetz in geringerem Ausmaß zusätzlich belastet und daher ggf. nur eine moderate Netzverstärkung erfordert.

Für die Umsetzung und Gestaltung der Wärmewende in München bedarf es wegen der zuvor genannten Wechselwirkungen zwischen den Sparten eines zentralen Entscheidungsunterstützungssystems. Zudem kann mit Karten in Papierform, Medienbrüchen, verteilten Datensilos, Datenlücken, Datenfehlern und Inkonsistenzen, fehlenden Schnittstellen und Berechtigungskonzepten, unterschiedlichen Arbeitsständen sowie unscharfen Zukunftsprojektionen mittels MS EXCEL (u.a. zur langfristigen Entwicklung der Sanierungsaktivitäten, der Treibhausgasemissionen, der Kosten und der Endenergie je Energieträger) die Komplexität bei der Planung und Steuerung der Wärmewende nicht beherrscht werden. Vielmehr sollten IT-Technologien unter dem Stichwort der intelligenten Digitalisierung angewendet werden, zu denen laut John [Jo22] u.a. die IT-basierte Prozessautomatisierung, die Künstliche Intelligenz, aber auch die Einbeziehung von Cloud-Computing zählen. Dabei unterstreicht John im Kontext der intelligenten Digitalisierung, dass Unternehmen, um erfolgreich zu sein, ihre Prozesse und Strukturen effizient gestalten und auf die gesteckten Zielen ausrichten müssen. So haben die SWM aufgrund der Bedeutung der Digitalisierung und der digitalen Transformation eine eigene Strategie im Jahr 2019 entwickelt. Als Teil der SWM-Digitalisierungsstrategie wurde beschlossen, das "Modell München" zu erstellen. Dieses zentrale Steuerungsinstrument soll u.a. eine einheitliche, aktuelle und konsistente Datenbasis inkl. Visualisierungskomponenten bereitstellen, so dass Änderungen spartenübergreifend und auf Basis eines Berechtigungskonzeptes (wg. Datenschutz, Unbundling⁴, Betriebsgeheimnis, kritische Infrastruktur) bekannt sind und zeitnah sowie ganzheitlich reagiert werden kann. Des Weiteren soll dieses Modell zur Planung der Wärmewende in München eingesetzt werden, indem Zukunftsprojektionen mittels Simulationen in einem Multiagentensystem (MAS) angestellt werden.

Der Aufbau dieses Beitrags gestaltet sich wie folgt: Das Modell München dient als Beispiel für ein gelungenes Projekt zur intelligenten Digitalisierung. Zunächst wird daher auf die Rahmenbedingungen für seine Entwicklung eingegangen. Daraufhin wird sein Einsatz

⁴ Gesetzlich vorgeschriebene Entflechtung des Netzbetreibers von anderen Geschäftsfeldern eines Energieversorgungsunternehmens, z.B. Trennung des Stromnetzbetreibers vom Stromlieferanten. Dies umfasst auch die Nutzung von Daten.

bei der kommunalen Wärmeplanung veranschaulicht, mit der die Umstellung des Wärme-sektors auf erneuerbare Energien und unvermeidbare Abwärme ganzheitlich geplant wird. Mit einer Zusammenfassung und einem Ausblick endet dieser Beitrag.

2 Rahmenbedingungen für die Entwicklung beim Modell München

Ende 2019 wurde begonnen, das Modell München zu erstellen. Hierfür wurden über einen Zeitraum von drei Jahren zunächst zwei Personen einer Abteilung eingesetzt (anfangs in Summe ein Vollzeitäquivalent). Dabei gab es eine klare Rollenverteilung nach dem Vorgehensmodell für Data-Science-Projekte DASC-PM [Sc22]: Eine Person hatte die Rollen Data Scientist und Data Engineer inne und die andere Person die Rollen Projektmanager sowie Domänenexperte. Inzwischen sind vier Vollzeitäquivalente im dem Thema aktiv. Der Zuwachs wurde möglich, da das Modell München rasch seinen Nutzen für die SWM unter Beweis gestellt hat. Die ergänzenden Rollen Compliance Support und Technical Support aus DASC-PM werden nicht innerhalb der Abteilung besetzt. Vielmehr werden für diese Rollen jeweils Experten bedarfsorientiert aus den entsprechenden Fachbereichen der SWM hinzugezogen.

Aufgrund der anfangs begrenzten personellen Ressourcen wurden zunächst relativ einfache Anwendungsfälle mit einem attraktiven Nutzenpotenzial (Quick Wins) akquiriert und umgesetzt, wie beispielsweise die Analyse von Wasserverlusten oder die Ermittlung des Absatzpotenzials für Wallboxen. Hierdurch ließ sich schnell die Bekanntheit des Modells im Konzern steigern. Zudem wurde das Modell in zahlreichen Jours fixes der Fachabteilungen und auf einer hauseigenen Digitalisierungsmesse vorgestellt. Inzwischen wird das Modell München innerhalb der SWM u.a. auch für die folgenden Anwendungsfelder eingesetzt:

- Transformationsplan für die Fernwärme.
- Potenzial für Wasserstoff in der dezentralen Wärmeversorgung.
- Auswirkungen von Elektromobilität auf das Stromnetz im Projekt E-Motion-to-Grid (E2G) [Kr23].
- Auswirkungen von Wärmepumpen auf das Stromnetz im Projekt Grid-for-Electrification (G4E).
- Kopplung des Projekts Geo.KW [Zo22] für das thermische Grundwassermanagement mit dem Modell München.
- Potenzial für Tunnelthermie bei U-Bahnen.

Für die Aufbereitung und Integration aller Daten in das Modell München werden R und Python genutzt, wobei die Daten in einer PostGIS-Datenbank abgelegt sind. Alle umfangreichen Berechnungen, wie beispielsweise das Schließen von Datenlücken beim Baujahr mittels Machine Learning [Ca20], werden in der Cloud ausgeführt. Für die Visualisierung

wird im Modell München zumeist Tableau genutzt. Zudem kommt ein Tableau Server zum Einsatz. Dieser hat den Vorteil, dass Nutzer über einen Browser auf relevante Tableau-Dashboards zugreifen können – es muss somit keine Applikation beim Nutzer installiert werden. Des Weiteren unterstützt der Tableau Server das umfangreiche Berechtigungskonzept für das Modell München. Je nach Anwendungsfall werden neben Tableau auch QGIS sowie browserbasierte Portale wie die Integrationsplattform 360 (SWM) oder das Geo Info Web der Landeshauptstadt München (LHM) eingesetzt. Gelegentlich sind Nutzer aber auch nur an bestimmten Daten zur Weiterverarbeitung oder an einzelnen Kennzahlen interessiert, sodass dann ggf. Exporte in den gewünschten Datenformaten vorgenommen werden. Die folgende Abbildung 1 zeigt beispielhaft ein Dashboard in Tableau (die beiden Hintergrundkarten stammen von © OpenStreetMap). Auf der linken Seite sind in der Karte Gebäude visualisiert und auf der rechten Seite Baublöcke (ein von Straßen eingefasstes Gebiet). Im unteren Bereich des Dashboards besteht die Möglichkeit, mit Radiobuttons verschiedene Inhalte in den beiden Karten für die Visualisierung auszuwählen.



Abb. 1: Beispielhaftes Tableau-Dashboard im Modell München, welches u.a. zur Bestandsanalyse sowie zur Definition von Quartieren eingesetzt werden kann

Eine Besonderheit am Modell München ist die Integration eines MAS der Energy Economic Group an der TU Wien zur gebäudescharfen Simulation von Sanierungsaktivitäten und des Austauschs von Heizungssystemen. Dabei wird jedes beheizte Gebäude (in München ca. 180.000) mit seinen individuellen Merkmalen wie Grundfläche, Höhe, Baujahr, Heizungsart, Nutzungsart etc. durch einen Agenten repräsentiert. Sobald nun Bauteile des Gebäudes wie das Dach, die Fassade oder die Heizung altern und ersetzt werden müssen, trifft der jeweilige Agent eine Entscheidung: Hierzu stehen ihm verschiedene Sanie-

rungsoptionen zur Verfügung – bis hin zum Abriss und Neubau – wie auch etwaige Optionen für das Heizungssystem. Da das Modell München alle Gebäude räumlich verortet, kann für das individuelle Heizungssystem jedes Gebäudes hinterlegt werden, ob beispielsweise die Fernwärme verfügbar ist, das Potenzial für eine Grundwasser-Wärmepumpe ausreicht, sich gegenseitige Beeinflussungen durch Kältefahnen im Grundwasser ergeben und die Luft-Wärmepumpe hinsichtlich der zulässigen Schallemissionen noch gewählt werden kann. Bei den Entscheidungen im MAS können auch Förderungen, Kostenänderungen, Verbote und Fristen berücksichtigt werden. So lassen sich beispielsweise die Effekte von Anpassungen am GEG quantifizieren.

Zur Plausibilitätsprüfung von Gebäudedaten, Parametern und Modellergebnissen gibt es zahlreiche Aktivitäten. Diese sind u.a.:

- Manuelle Überprüfung von Eingangsdaten in Stichproben mittels Begehungen und anhand von Satellitenaufnahmen (z.B. Baujahr, Anzahl Stockwerke, Nutzungsart, Strukturtypen).
- Abgleich von Parametern mit Studienergebnissen und mit Erkenntnissen aus umgesetzten Praxisprojekten (z.B. Sanierungskosten).
- Visualisierung von Gebäudemerkmalen in Tableau zur Identifikation von Ausreißern sowie unplausiblen Daten und Ergebnissen.
- Berechnung historischer Jahre und Abgleich mit tatsächlichen Energielieferungen für Fernwärme, Gas und Strom je Gebäude.
- Bildung von Kennzahlen zur besseren Prüfung auf Plausibilität (z.B. Nutzenergie je m² beheizte Fläche, Sanierungskosten je m², Sanierungsrate pro Jahr).
- Vergleich von Ergebnissen anhand von sogenannten “Zwillingsgebäuden“ (Gebäude auf der gleichen Straßenseite, benachbarte Grundstücke, gleiches Baujahr, gleiche Heizungsart und gleiche Grundfläche).
- Workshops mit Experten von Hochschulen und kommerziell agierenden Instituten sowie von den SWM und der LHM.
- Vergleich der Modellergebnisse für München mit den Ergebnissen anderer Städte.

3 Modell München in der kommunalen Wärmeplanung

Die kommunale Wärmeplanung ist eng mit dem GEG verbunden und zeigt auf, wie die Klimaneutralität in der Wärmeversorgung bis spätestens 2045 erreicht wird. Damit soll zum einen das Problem- und Lösungsbewusstsein der regionalen Akteure gestärkt werden. Zudem dient die Wärmeplanung dazu, Planungs- und Investitionssicherheit für Bürger, Unternehmen sowie Netzbetreiber zu schaffen. Zur Erstellung der Wärmeplanung sind mehrere Schritte vorgesehen, die nachfolgend kurz erläutert werden [PSB20, Bu23]:

1. Bestandsanalyse: Zunächst wird sich ein Überblick über den regional aufgelösten Wärmeverbrauch, die Gebäudestrukturen sowie die eingesetzten Energieträger verschafft. In der Bestandsanalyse geht es zudem um die Standorte bestehender Wärmeerzeugungsanlagen und um die Verläufe bestehender Netze.
2. Potenzialanalyse: In diesem Schritt werden die regional aufgelösten Potenziale für erneuerbare Energien und unvermeidbare Abwärme erfasst. Dies können Fernwärmeverdichtungs- sowie Fernwärmeausbaugebiete sein, aber auch regionale Potenziale für Grundwasser-Wärmepumpen, Erdkollektoren, Luft-Wärmepumpen etc.
3. Zielszenario: Ausgehend von der Bestands- und Potenzialanalyse sowie unter Einbeziehung weiterer Parameter wie beispielsweise der Entwicklung des Wärmebedarfs, der Energiepreise, der Förderbedingungen etc. werden Szenarien erstellt, aus denen ein Zielszenario für das gesamte zu planende Gebiet ausgewählt wird.
4. Meilensteine: Für die Zeitpunkte⁵ t+2, t+5, t+10, 2035 sowie 2045 wird je Teilgebiet angegeben, welche Wärmeversorgungsart erwartet wird und wie sich diese im jeweiligen Teilgebiet voraussichtlich entwickeln wird.
5. Umsetzungsmaßnahmen: Die zum Zeitpunkt der Beitragserstellung in Deutschland eingeführten Maßnahmen sind nicht ausreichend, um bis 2045 Heizwärme und Warmwasser ausschließlich aus erneuerbaren Energien und aus unvermeidbarer Abwärme zu erzeugen. Im letzten Schritt sind daher zusätzliche Maßnahmen zu definieren, mit denen die Ziele der Wärmeplanung erreicht werden.

Zusätzlich zu den Einsatzfeldern bei den SWM wird das Modell München seit 2022 in der kommunalen Wärmeplanung eingesetzt, für die das Referat für Klima- und Umweltschutz (RKU) der LHM federführend zuständig ist. Dabei findet das Modell München in allen fünf zuvor genannten Schritten Anwendung. Für die Bestandsanalyse stehen zahlreiche Visualisierungsmöglichkeiten zur Verfügung, um die Vielzahl an integrierten Daten zu Gebäuden und deren Wärmeversorgung zu untersuchen. Zudem enthält das Modell München umfassende Daten zu den regional verfügbaren Versorgungspotenzialen, sodass auch der Schritt der Potenzialanalyse vollständig abgedeckt ist. Zur Aufstellung des Zielszenarios wird das in das Modell München integrierte MAS der Energy Economic Group eingesetzt. Hierbei werden u.a. die Effekte des GEG, eine erhöhte Sanierungsrate und -quote, ein Verbrennungsverbot fossiler Energieträger etc. untersucht. Auch bei den Auswertungen für die Beschreibung der Meilensteine findet das Modell München Anwendung – ebenso bei der Definition konkreter Umsetzungsmaßnahmen, da die Effekte möglicher Maßnahmen häufig mit dem MAS quantifiziert werden können (für Details zum MAS und der Anwendung im Modell München siehe [Mü15], [GM20] sowie [GM22]).

Ohne das Modell München würde man in der kommunalen Wärmeplanung der LHM, wie bereits in der Wärmestudie 2035 [K121], zumeist MS EXCEL statt eines MAS für die Szenarioerstellung nutzen. Zudem würde voraussichtlich QGIS statt Tableau inkl. Tableau

⁵ Wenn beispielsweise t für das Jahr 2023 steht, dann meint t+2 das Jahr 2025.

Server für die Visualisierung und Analyse eingesetzt werden. Aber auch bei der Datenqualität und dem Aggregationsniveau gäbe es Abstriche. So standen in der Wärmestudie 2035 historische Lieferenergiebedarfe je PLZ zur Verfügung – beim Modell München liegen sie je Gebäude vor.

Das Modell München beinhaltet derzeit Daten für das Stadtgebiet der LHM und auch die kommunale Wärmeplanung findet aktuell nur für dieses Gebiet statt. Da die SWM aber auch außerhalb Münchens aktiv sind und dort Netze betreiben, werden bis voraussichtliche Ende 2023 die Eingangsdaten für das Modell München bei Gebäuden und Flurstücken teilweise auf umliegende Landkreise erweitert. Darüber hinaus sind Erweiterungen nicht ohne Weiteres möglich – denn Daten liegen häufig in anderen Formaten und in anderer Qualität vor, sodass jeweils individuelle Anpassungen an den Algorithmen im Modell München erforderlich wären, was äußerst zeitintensiv ist. Zudem bestehen Einschränkungen durch das Kommunalrecht.

4 Zusammenfassung und Ausblick

In diesem Beitrag wurde das Modell München als Teil der SWM-Digitalisierungsstrategie vorgestellt. Dabei wurden u.a. die Rahmenbedingungen für dessen Entwicklung erläutert. Hier ist deutlich geworden, dass es bereits mit zwei Personen nach dem Rollenmodell von DASC-PM möglich ist, ein durchaus komplexes Entscheidungsunterstützungssystem für die Planung und Steuerung der Wärmewende zu erstellen. Mit dem Modell München wurde das Ziel erreicht, einen zentralen, konsistenten und um Datenlücken bereinigten Datensatz zu schaffen, der in einem Entscheidungsunterstützungssystem mit integriertem MAS bei zahlreichen Anwendungsfällen zum Einsatz kommt – einer dieser Anwendungsfälle ist die kommunale Wärmeplanung der LHM. Aufgrund dieses Erfolgs und wegen des Einsatzes von Cloud Computing, der Erstellung eines digitalen Zwillings, der Anwendung von Machine Learning und von Multiagenten-Simulation sowie durch die IT-basierte Prozessautomatisierung (z.B. automatisierte Auskunft zu Versorgungsoptionen) kann das Modell München als Beispiel für ein gelungenes Projekt der intelligenten Digitalisierung bezeichnet werden.

Neben der kontinuierlichen Verbesserung der Algorithmen im Modell München (z.B. Zuweisung von Heizungsarten zu Gebäuden, Identifikation mitversorgter Gebäude, Zuweisung von Adressen zu Gebäuden, Bestimmung der Schallemissionen von Luft-Wärmepumpen, Ermittlung optimaler Trassenverläufe bei Wärmenetzen) sowie der Anbindung weiterer Datenquellen, steht derzeit das Thema Prozessexzellenz auf der Agenda. Denn bisher lag der Fokus eher auf der erfolgreichen Umsetzung von Anwendungsfällen, um das Modell München als leistungsfähiges Werkzeug zu positionieren. Mit dem eintretenden Erfolg sind Anfragen und Anwendungsfälle inzwischen so zahlreich, dass sich zur Verbesserung der Prozesse rund um das Modell München u.a. verstärkt dem Releasemanagement sowie dem Datenkatalog gewidmet werden muss – für Letzteres wird derzeit ein geeignetes Tool ausgewählt.

Literaturverzeichnis

- [BD23] BDEW: Beheizungsstruktur des Wohnungsbestandes in Deutschland. Anteile der genutzten Energieträger, 2023. <https://www.bdew.de/service/daten-und-grafiken/beheizungsstrukturwohnungbestand>, Abruf am: 02.07.2023.
- [Bu23] Bundesregierung: Referentenentwurf. Gesetz für die Wärmeplanung und zur Dekarbonisierung der Wärmenetze. Stand: Mai, 2023.
- [Ca20] Castell, C. et al.: Vergleich von Machine Learning Ansätzen zur Ermittlung des Alters der Gebäude in München. In: Simulation in den Umwelt- und Geowissenschaften. Shaker Verlag, S. 103-116, 2020.
- [GM22] Günther, M.; Müller, A.: Datenanalyse und Simulation zur gebäudescharfen Prognose von Wärmebedarfen und Heizungstechnologien in einer Großstadt. In: Simulation in den Umwelt- und Geowissenschaften. Shaker Verlag, S. 63-74, 2022.
- [GM20] Günther, M; Müller, A.: Ortsaufgelöste Transformation des Gebäudebestands in München zur Prognose der Fernwärmebedarfs. In: Simulation in den Umwelt- und Geowissenschaften. Shaker Verlag, S. 93-102, 2020.
- [Jo22] John, U.: Intelligentes Unternehmen, Intelligenter Staat, Intelligentes Europa – Was ist zu tun? In: INFORMATIK 2022, 14.Workshop KI-basiertes Management und Optimierung komplexer Systeme (MOC 2022), LNI P326, DOI = 10.18420/inf2022_145, S. 1695-1698, 2022.
- [Kl21] Kleinertz, B. et al.: Klimaneutrale Wärme München 2035. Mögliche Lösungspfade für eine klimaneutrale Wärmeversorgung in der Landeshauptstadt München. 2021.
- [Kr23] Kreutmayr, S. et al.: Auswirkungen von Elektrofahrzeugen auf städtische Niederspannungsnetze bei unterschiedlichen Ladestrategien. In: 13. Internationale Energiewirtschaftstagung (IEWT), 2023.
- [Mü15] Müller, A.: Energy Demand Assessment for Space Conditioning and Domestic Hot Water: A Case Study for the Austrian Building Stock. Dissertation, Technische Universität Wien, 2015.
- [PSB20] Peters, M.; Steidle, T.; Böhnisch, H.: Kommunale Wärmeplanung. Handlungsleitfaden, Stuttgart, 2020.
- [Sc22] Schulz, M. et al.: DASC-PM v1.1. A Process Model for Data Science Projects. Nordakademie, 2022.
- [Zo22] Zosseder, K.; et al.: Schlussbericht zum Verbundprojekt GEO-KW: Kopplung des geothermischen Speicherpotenzials mit den wechselnden Anforderungen des urbanen Energiebedarfs zur effizienten Nutzung der regenerativen Energiequelle Grundwasser für die dezentrale Kälte- und Wärmebereitstellung in der Stadt. DOI = 10.14459/2022md1692003, 2022.

A Proposal for Physics-Informed Quantum Graph Neural Networks for Simulating Laser Cutting Processes

Zurana Mehrin Ruhi,¹ Hannah Stein,¹ Wolfgang Maaß¹

Abstract: Simulations are crucial for production monitoring and planning in manufacturing. Still, the performance of simulations based on mathematical modeling and machine learning methods is limited and opaque to widespread application. Quantum computing offers the potential for exponential acceleration of these tools, while physically informed neural networks (PINN) improve learning and reduce ambiguity. Objective of this paper is to explore the concept of developing a tool for laser cutting simulation based on a quantum neural network that can be trained on thermal physics principles.

Keywords: Quantum Computing; PINN; Graph Neural Network; Predictive Simulation

1 Introduction

Laser cutting is a widely embraced manufacturing technique, e.g., in automotive. During the cutting process, thermal effects can lead to undesired outcomes such as distortion, and alterations in material properties; impacting the structural integrity and quality of end products. The process requires a balance of cutting speed, precision, and preserving edge quality. Analytical- or AI-based thermal simulations assist process optimization and design validation allowing manufacturers to make informed decisions. Still, these approaches often lack critical validation and present uncertainty due to oversimplification of the material's properties and boundary conditions [PH14]. The utilization of multiphysics modelling to simulate diverse conditional processes has a supportive effect but requires heavy computational power [Kn21]. Model-driven (i.e. ML) methods require less computational complexity but only reduce the maximum deformation by approx. 33% [Lu22]. Demonstrations showed that incorporating physics-based principles with neural networks (NN) enables them to capture the intricacies of the simulation, resulting in a faster-converging training process [RPK19]. These networks reduce dependence on input data by incorporating physical laws to help overcome learning plateaus. Thermodynamics-informed Graph NN (TIGNN) [He22] is one example of PINNs that successfully simulates fluid and metal changes exploiting Hamiltonian Physics. Using such methods can be a powerful tool for simulating the time-dependent and dynamic nature of the laser-cutting process. However, these solutions are often too resource-intensive whereas Quantum computing can offer great exponential speedup. Related work shows evidence of Quantum Machine Learning (QML) surpassing

¹ Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Smart Service Engineering, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Deutschland, {zurana_mehrin.ruhi;hannah.stein;wolfgang.maass}@dfki.de

its classical counterparts [ZJQ23]. Given the criticality of speed and precision in laser cutting simulations, quantum-enhanced simulations hold the potential to ensure accurate component reproduction. This paper proposes a model to optimize the accuracy, efficiency, and effectiveness of laser cutting simulations by utilizing physics-based solutions and leveraging the capabilities of QC.

2 Proposed method

We aim to train a NN model to solve the partial differential equations governing the behaviour of a continuous laser beam source on a 2D rectangular plate. The temperature $u(x,y,t)$ for this satisfies the following partial differential equation with initial and boundary conditions:

$$\frac{f-q}{\Delta z} = \rho c_p \cdot \frac{\partial u}{\partial t} - \nabla \cdot (k \cdot \nabla u)$$

$$q(x, y, t) = h \cdot (u(x, y, t) - u_\infty)$$

$$u(x, y, 0) = u_\infty, u(0, y, t) = u(a, y, t) = u(x, 0, t) = u(x, b, t) = u_\infty$$

To enforce the model’s learning process based on these rules, an additional loss function will be constructed that can be incorporated alongside the classical loss function. This model will be then transformed into a Quantum version of the PINN model. In typical QML approaches Parameterized Quantum Circuits [Be19] adopt the concept of setting predefined parameters and designing a gated parameterized model. The free parameters of the gates are adjusted using classical computing techniques which greatly simplifies the parameter-tuning process. The proposed structure to implement Physics-informed GNNs as Quantum based PINNs is inspired from a general implementation of quantum based PINNs [VB23].

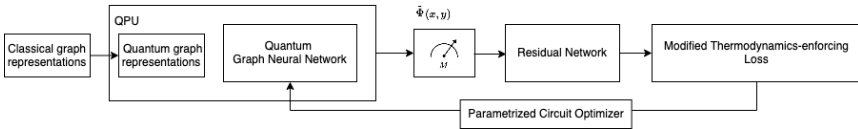


Abb. 1: Proposed circuit for Quantum-based PINN; illustration based on [VB23]

3 Conclusion

In this paper, we propose an idea to achieve improved performance in simulating the complex dynamics of laser-cutting processes, by harnessing the computational power and unique properties of QC. To implement the experiment, we will first encode the two-dimensional coordinates for the QNN unlike the one-dimensional cases introduced in previous works [Ma22]. After designing the tools to translate classical graph representations to quantum graph ones and vice-versa, the physics-based optimization will take place to generate the final model. The ultimate aim is to push the boundaries of QC further with PINNs to achieve trustworthy and effective simulation models for the manufacturing industry.

4 Acknowledgement

This work is part-funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) within the research project QUASIM, grant number: 01MQ22001A.

Literaturverzeichnis

- [Be19] Benedetti, Marcello; Lloyd, Erika; Sack, Stefan; Fiorentini, Mattia: Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4, 10 2019.
- [He22] Hernandez, Quercus; Badias, Alberto; Chinesta, Francisco; Cueto, Elias: Thermodynamics-informed Graph Neural Networks. *IEEE Transactions on Artificial Intelligence*, 2022.
- [Kn21] Knüttel, Daniel; Baraldo, Stefano; Valente, Anna; Wegener, Konrad; Carpanzano, Emanuele: Model based learning for efficient modelling of heat transfer dynamics. *Procedia CIRP*, 102:252–257, 2021. 18th CIRP Conference on Modeling of Machining Operations (CMMO), Ljubljana, Slovenia, June 15-17, 2021.
- [Lu22] Lu, Chang; Fei, Jiyu; Meng, Xiangzhong; Li, Yanshu; Liu, Zhibo: Thermal Error Prediction and Compensation of Digital Twin Laser Cutting Based on T-XGBoost. *Sensors*, 22(18), 2022.
- [Ma22] Markidis, Stefano: On physics-informed neural networks for quantum computers. *Frontiers in Applied Mathematics and Statistics*, 8, oct 2022.
- [PH14] Parandoush, Pedram; Hossain, Altab: A review of modeling and simulation of laser beam machining. *International Journal of Machine Tools and Manufacture*, 85:135–145, 10 2014.
- [RPK19] Raissi, M.; Perdikaris, P.; Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- [VB23] Vadyala, Shashank Reddy; Betgeri, Sai Nethra: General implementation of quantum physics-informed neural networks. *Array*, 18:100287, 2023.
- [ZJQ23] Zeguendry, Amine; Jarir, Zahi; Quafafou, Mohamed: Quantum Machine Learning: A Review and Case Studies. *Entropy*, 25(2), 2023.

Wirtschaft, Management Industrie -
Künstliche Intelligenz für kleine und
mittlere Unternehmen (KI-KMU
2023)

Von der Theorie zur Praxis: Erfahrungen bei der akademischen Begleitung von KI-Projekten in KMUs

Frederik Simon Bäumer,¹ Hans Brandt-Pook,² Falk Maoro,³ Sergej Schultenkämper,⁴
Benjamin Stecker⁵

Abstract: Die Integration von künstlicher Intelligenz (KI) in Unternehmen wird zunehmend wichtiger, um wettbewerbsfähig zu bleiben. In diesem Beitrag stellen wir unsere akademischen Erfahrungen bei der Begleitung von KI-Projekten in kleinen und mittelständischen Unternehmen (KMUs) vor. Wir präsentieren mehrere Fallstudien aus verschiedenen Branchen, in denen wir beratend tätig waren. Die Ergebnisse zeigen, dass eine strukturierte Umsetzung von KI-Projekten dazu beitragen kann, die Herausforderungen zu meistern. Der Theorie-Praxis-Transfer zielt darauf ab, Eintrittsbarrieren in das Themenfeld KI abzubauen und Ängste in den KMUs zu nehmen.

Keywords: Künstliche Intelligenz; Theorie-Praxis-Transfer

1 Einführung

Die fortschreitende Digitalisierung und der Einsatz von künstlicher Intelligenz (KI) spielen eine immer größere Rolle in der Wirtschaft und tragen maßgeblich dazu bei, die Prozesse in Unternehmen zu optimieren und die Wettbewerbsfähigkeit zu steigern [Du19]. Insbesondere kleine und mittelständische Unternehmen (KMUs) können von den Potenzialen von KI profitieren, indem sie ihre Geschäftsprozesse optimieren und ihre Produktivität erhöhen. Doch KMUs stehen oft vor der Herausforderung, nicht über die notwendigen Ressourcen, Expertise und Mittel zu verfügen, um eigene KI-Abteilungen aufzubauen [UF21]. Dieser Umstand wird oftmals als Eintrittsbarriere in die KI-Welt beschrieben. Aus diesem Grund suchen KMUs nach akademischer Unterstützung, um KI-Projekte erfolgreich umzusetzen.

In diesem Kontext spielt der Theorie-Praxis-Transfer zwischen Hochschulen und KMUs eine wichtige Rolle, indem die akademischen Partner ihre Technologieoffenheit und domänenübergreifenden Erfahrungen bei der Ausarbeitung und Umsetzung von KI-Roadmaps, bei der Anwendung von neuesten Technologien oder der Einführung in KMUs bereitstellen [LG21]. Speziell KI-Roadmaps sind ein Ansatz zur Planung und Umsetzung von KI-Projekten, der es Unternehmen ermöglicht, ihre KI-Aktivitäten strategisch zu planen

¹ Hochschule Bielefeld, Interaktion 1, 33619 Bielefeld, frederik.baeumer@hsbi.de

² Hochschule Bielefeld, Interaktion 1, 33619 Bielefeld, hans.brandt-pook@hsbi.de

³ Hochschule Bielefeld, Interaktion 1, 33619 Bielefeld, falk.maoro@hsbi.de

⁴ Hochschule Bielefeld, Interaktion 1, 33619 Bielefeld, sergej.schultenkaemper@hsbi.de

⁵ Hochschule Bielefeld, Interaktion 1, 33619 Bielefeld, benjamin.stecker@hsbi.de

und zu steuern. Dabei werden die Anforderungen, Ziele und Meilensteine der KI-Projekte festgelegt und in einen Zeitplan integriert. Aber auch konkreter in der Umsetzung können Standard-Vorgehensmodelle wie CRISP-DM und „vorbefüllte KI-Werkzeugkoffer“ dabei helfen, KMUs mit wenig KI-Erfahrungen den Zugang zu diesen Themen zu erleichtern.

In diesem Beitrag werden die Erfahrungen der Autoren beim Theorie-Praxis-Transfer in KMUs rund um das Thema KI vorgestellt und die Herausforderungen bei der Umsetzung von KI-Projekten erläutert. Insbesondere wird auf die spezifischen Herausforderungen eingegangen, denen sich KMUs gegenübersehen, wenn es darum geht, KI-Projekte zu planen und umzusetzen. Die Ergebnisse des Beitrags zeigen, dass eine fundierte Planung und Umsetzung von KI-Projekten dazu beitragen kann, die Herausforderungen zu meistern. Im folgenden Abschnitt stellen wir die Organisation und Vorgehensweise seitens der Arbeitsgruppe „Angewandte KI“ der Hochschule im Rahmen von Theorie-Praxis-Transferprojekten dar (s. Abschnitt 2). Folgend wird in Abschnitt 3 auf ausgewählte Kooperationen mit KMUs eingegangen und es werden Anwendungsfälle sowie Herausforderungen dargestellt, die in Abschnitt 4 diskutiert werden. Hier leiten wir *Lessons learned* ab, die helfen können, zukünftige KI-Projekte effizient durchzuführen.

2 Organisation und Vorgehensweise

Der Theorie-Praxis-Transfer im Bereich der KI wird in unserem Fall durch die Arbeitsgruppe „Angewandte KI“ im Bereich Wirtschaftsinformatik umgesetzt. Wie der Name bereits suggeriert, zielt die Arbeitsgruppe darauf ab, praxisnahe Lösungen zu entwickeln. Das Vorgehen der Arbeitsgruppe ist darauf ausgerichtet, KMUs bei der Einführung und Anwendung von KI-Technologien zu unterstützen. Dazu haben wir eine flexible und lösungsorientierte Herangehensweise entwickelt, die es uns ermöglicht, maßgeschneiderte Lösungen für die individuellen Bedürfnisse der Partnerunternehmen zu entwickeln und gleichzeitig den Herausforderungen des Mittelstandes gerecht zu werden.

Ein zentraler Aspekt der Arbeitsweise der Arbeitsgruppe ist der umfangreiche Werkzeugkasten an Methoden und Technologien rund um KI. Dieser wird stets auf dem neuesten Stand gehalten, indem sich die Arbeitsgruppe regelmäßig mit den neuesten Entwicklungen auseinandersetzt. Er umfasst:

- Datenanalyse- und -visualisierungswerkzeuge: Werkzeuge zur Extraktion, Bereinigung, Analyse und Visualisierung, um Muster und Erkenntnisse zu gewinnen.
- Machine Learning-Frameworks und -Bibliotheken: Beliebte Frameworks wie TensorFlow, PyTorch oder Scikit-learn, die es ermöglichen, Machine Learning-Modelle zu entwickeln, zu trainieren und zu evaluieren.
- NLP- und Textanalyse-Tools: Werkzeuge, die auf natürlicher Sprachverarbeitung (NLP) basieren, um Textdaten zu verstehen, zu klassifizieren oder zu generieren.

- Computer Vision-Tools: Bibliotheken und Frameworks zur Verarbeitung von Bildern und Videos, um Objekte zu erkennen, zu segmentieren oder zu klassifizieren.
- Big Data- und Cloud-Technologien: Werkzeuge und Plattformen zur Speicherung, Verwaltung und Analyse großer Datenmengen, einschließlich cloudbasierter Dienste wie Amazon Web Services (AWS) oder Google Cloud Platform (GCP).
- Modellinterpretations- und Erklärbarkeitswerkzeuge: Werkzeuge, die helfen, komplexe KI-Modelle zu interpretieren, um Vertrauen und Transparenz zu schaffen.
- Datenschutz- und Sicherheitswerkzeuge: Werkzeuge, die den Schutz der Daten und Modelle gewährleisten, wie beispielsweise Datenschutztechniken, Anonymisierungsmethoden und Modellsicherheitsprüfungen.
- Projektmanagement- und Kollaborationsplattformen: Werkzeuge zur Organisation und Verwaltung von KI-Projekten sowie zur Zusammenarbeit innerhalb des Teams und mit externen Partnern.
- Ressourcen für Best Practices und Fallstudien: Eine Sammlung von Methoden, Leitfäden, Fallstudien und anderen Ressourcen, um auf spezifische Probleme einzugehen und Ansätze für verschiedene Branchen und Anwendungsfälle zu demonstrieren.

Diese Vielfalt ermöglicht es, gemeinsam mit den Kooperationspartnern passgenaue Lösungen für deren spezifische Anforderungen zu finden. Eine enge Zusammenarbeit und offene Kommunikation mit den Partnern ist dabei entscheidend, um ein gemeinsames Verständnis der Problemstellung und der angestrebten Lösung zu erreichen. So werden die individuellen Stärken der Partner optimal genutzt und Synergien geschaffen, die zur erfolgreichen Umsetzung beitragen. Der KI-Werkzeugkoffer ist eine individuelle Auswahl an relevanten KI-Methoden, mit dem Ziel, den KMUs einen konkreten Überblick zu geben. Unsere Erfahrung zeigt, dass die bloße Masse an Verfahren, Tools, Frameworks und Cloud-Lösungen bereits eine abschreckende Wirkung haben kann. In den von der Arbeitsgruppe durchgeführten Projekten wurden ferner folgende Herausforderungen für KMUs deutlich: Unternehmen mangelt es häufig an ausreichenden Ressourcen [LG21], sowohl im Sinne von Hardware als auch an Fachpersonal mit entsprechendem Domänen- sowie KI-Wissen. Ferner besteht bei KMUs häufig zum einen ein Zeitdruck, KI-Lösungen zu entwickeln, um wettbewerbsfähig zu bleiben und auch aus Marketing-Sicht mit diesen Entwicklungen sichtbar zu werden. Zum anderen sehen sie sich oftmals der Tatsache gegenüber, dass Mitarbeiter nur schwer Arbeitsstunden in diese Themen investieren können, da es hierdurch potenziell bei anderen Themen zu Engpässen kommt. Ein Schwerpunkt unserer Arbeitsgruppe liegt daher auf dem Wissenstransfer, um Unternehmen langfristig dabei zu unterstützen, KI-Kompetenzen aufzubauen und die Technologien effizient einzusetzen.

Ein weiterer wichtiger Aspekt der Arbeitsgruppe ist die Orientierung an der agilen Methodik und dem CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Durch die Anwendung von CRISP-DM [WH00] gelingt es der Arbeitsgruppe, die Entwicklung und Implementierung von KI-Lösungen effizient und iterativ zu gestalten. Dabei ist CRISP-DM

weniger aufgrund der Passgenauigkeit auf das Thema KI gewählt worden, sondern vielmehr aufgrund der Bekanntheit innerhalb der KMUs.

3 Kooperationen: Unternehmensspezifische KI-Lösungen

KI-Projekte für KMUs können sehr unterschiedlich sein und reichen von der Automatisierung von Routineaufgaben bis hin zur Analyse großer Datenmengen zur Verbesserung des Kundenerlebnisses. Dabei spielt die Branche keine Rolle, denn KI kann in allen Branchen eingesetzt werden, sei es im Handel, in der Fertigung oder im Dienstleistungsbereich. Auch die Unternehmensgröße oder der Umsatz sind keine entscheidenden Faktoren, um von KI-Projekten zu profitieren. Schon kleine Unternehmen mit einem Mitarbeiter können von der Automatisierung einfacher Aufgaben wie der Beantwortung von Kundenanfragen oder der Verwaltung von Bestellungen profitieren [LG21]. Die Arbeitsgruppe hat in den letzten vier Jahren mehr als 20 KI-Kooperationsprojekte mit KMUs durchgeführt. Im Folgenden werden beispielhaft die Arbeiten zum Thema „*KI in der Kundenbetreuung*“ (Abschnitt 3.1) und „*KI-gestütztes multilinguales Content Management*“ (Abschnitt 3.2) vorgestellt.

3.1 KI in der Kundenbetreuung

In dieser Kooperation geht es um ein Unternehmen, das den Prozess zur Verarbeitung von eingehenden Tickets beim Kundenservice optimieren möchte. Das Unternehmen hat nur wenige Experten, die in der Lage sind, die eingehenden Tickets zu bearbeiten. Dies führte dazu, dass es zu langen Wartezeiten kommen konnte und Kunden nicht immer die angemessene Unterstützung erhielten. Um dieses Problem zu lösen, beschloss das Unternehmen, eine KI einzusetzen, um die Tickets zu sortieren und zu kategorisieren. Dies bietet den Vorteil, dass Fachexperten ausschließlich für sie relevante Tickets erhalten und sich auf die eigenen Fachgebiete konzentrieren, sodass sie mit weniger Arbeitslast konfrontiert werden. Ferner wird die Zeit, die Tickets bislang im Unternehmen kursierten, da diese nicht den richtigen Personen zugeordnet waren, minimiert. Die KI wurde darauf trainiert, die eingehenden E-Mails zu analysieren und automatisch die vorliegende Problemstellung zu erkennen, sodass die Tickets den richtigen Kategorien zugeordnet werden.

Der Einsatz von KI soll aber weitreichender sein: Im weiteren Projektverlauf ist es das Ziel, die KI zu befähigen, die Fragen der Kunden direkt zu beantworten. Hierzu wird eine bestehende Wissensbasis angedockt, um sicherzustellen, dass die KI auf eine umfassende Informationsquelle zurückgreifen kann. Während der Entwicklung dieses Projekts erkannte das Unternehmen, dass die aufgeworfenen Fragen so tiefgreifend sind, dass eine Überarbeitung der gesamten Wissensdatenbank notwendig sein könnte. Das Unternehmen setzte Experten ein, um die Wissensdatenbank zu überarbeiten und sicherzustellen, dass alle Informationen enthalten und auch von KI-Modellen verarbeitet werden können.

3.2 KI-gestütztes multilinguales Content Management

Dieses KI-Projekt beschreibt eine Kooperation mit einer Digitalberatungs- und Contentagentur. Diese führt komplexe Web- und eCommerce Projekte für Kunden durch und setzt zunehmend auf den Einsatz von KI zur schnellen und zuverlässigen Erstellung von hochwertigen Inhalten für den internationalen Rollout. In einem Rollout-Prozess kommt es gehäuft vor, dass Inhalte in einer Sprache vorliegen, jedoch in unterschiedlichen Formaten (Blogs, Tweets, Posts) und Sprachen (bis hin zu 30 Sprachen) ausgespielt werden müssen. Die Zusammenarbeit zwischen der Arbeitsgruppe und der Agentur zielte auf die gemeinsame Erprobung von KI-Lösungen im Bereich der Textübersetzung und Texttransformation ab. Der Fokus lag dabei auf Cloud-Lösungen sowie aktuellen Transformer-Modellen [Va17], deren effiziente Bereitstellung ebenfalls Beratungsgegenstand war.

Die erste Phase des Projektes umfasste die Akquise, Verarbeitung und Analyse von Daten zur Erstellung der Datenbasis für die KI-Modelle. In der zweiten Phase wurden State-of-the-Art Transformer-Modelle für die Textübersetzung und Textgenerierung ausgewählt, analysiert, trainiert und evaluiert. Das Ziel der Kooperation bestand darin, die Effizienz und Qualität der Inhaltsproduktion der Agentur zu steigern und damit ihre Wettbewerbsfähigkeit auf dem Markt zu erhöhen. Während der Projektlaufzeit ergaben sich immer mehr Anwendungsfälle bei der Agentur, die auf der Projektlösung aufbauen konnten (z. B. SEO-Optimierung), was zu einer breiten Akzeptanz der KI-Bestrebungen bei den Mitarbeitern führte.

Die erfolgreiche Zusammenarbeit mündete in einer Ausgründung der neu geschaffenen KI-Abteilung der Agentur, die sich heute auf die Entwicklung und den Vertrieb von Textübersetzungs- und Texttransformationslösungen spezialisiert hat.

4 Lessons Learned beim Theorie-Praxis-Transfer

Um unsere Erfahrungen bei der Durchführung der KI-Kooperationen zu strukturieren und zu verallgemeinern, orientieren wir uns im Folgenden an den CRISP-DM Schritten.

4.1 Business Understanding

Business Understanding ist ein wichtiger Schritt in jedem KI-Projekt, da er dazu beiträgt, sicherzustellen, dass das Projekt auf die Bedürfnisse und Anforderungen des Unternehmens zugeschnitten ist und das Problem, das gelöst werden soll, vollständig verstanden wird. Dabei stellen wir oft fest, dass die Erwartungen an KI-Lösungen überzogen sind oder dass das Problem, das mittels KI gelöst werden soll, sich hierfür gar nicht eignet. Dies sind beispielsweise einmalig auftretende Datentransformationen, die kein Training einer KI rechtfertigen, da Kosten und Nutzen nicht ausgeglichen sind. Ferner entstehen durch KI-Hypes in den Medien (zuletzt ChatGPT) Erwartungshaltungen, die KI-Lösungen in ihrer Flexibilität und Robustheit nicht liefern können.

Im Falle unseres KI-Projektes KI in der Kundenbetreuung (s. Abschnitt 3.1) stellte sich zu Beginn die Frage, wie der aktuelle Prozess der Bearbeitung von Tickets abläuft. Hierbei sind verschiedene Aspekte von Bedeutung, wie beispielsweise über welche Kanäle Tickets eröffnet werden und welche Systeme dabei zum Einsatz kommen. Zudem stellt sich die Frage, welche Personen die Tickets bearbeiten und wer für die Kategorisierung zuständig ist. Ein weiterer wichtiger Punkt betrifft die zeitaufwändigen Aufgaben innerhalb des Ticket-Bearbeitungsprozesses. Es gilt herauszufinden, welche Tätigkeiten besonders viel Zeit in Anspruch nehmen und ob es hierbei Möglichkeiten gibt, diese effizienter zu gestalten. In diesem Zusammenhang kann auch die Frage gestellt werden, welche Aufgaben möglicherweise durch den Einsatz von KI-Systemen unterstützt werden könnten. Durch gezieltes Fragen werden so die eigentlichen KI-Potenziale herausgearbeitet.

Learning: Klarheit über das Problem schaffen. Es ist wichtig, das Geschäftsproblem, das gelöst werden soll, vollständig zu verstehen und klar zu definieren, um sicherzustellen, dass die KI-Modelle tatsächlich dazu beitragen, das Problem zu lösen.

Diese Klarheit wird vor allem durch die Zusammenarbeit mit Fachleuten geschaffen. Es ist wichtig, Fachleute zu identifizieren und mit diesen im Unternehmen zusammenzuarbeiten, um ein tiefes Verständnis für das Geschäftsproblem und die spezifischen Anforderungen zu erhalten. Die Zusammenarbeit kann dazu beitragen, sicherzustellen, dass das KI-Modell auf die Bedürfnisse des Unternehmens zugeschnitten ist. Dies ist hervorzuheben, da KI-Projekte oftmals auf höheren Unternehmensebenen beschlossen oder initiiert werden, die Verantwortlichen jedoch nicht zwangsläufig diejenigen sind, die mit den KI-Lösungen arbeiten müssen. Dies zeigte sich beispielsweise bei unserer Kooperation zum KI-gestützten multilingualen Content Management (s. Abschnitt 3.2). Hier galt es zu verstehen, welche Übersetzungslösung das Unternehmen benötigte und welche sprachlichen Herausforderungen in der jeweiligen Fachdomäne bestanden. Behandelte Domänen umfassten beispielsweise Landwirtschaft und Maschinenbau. Dieses Verständnis ist von entscheidender Bedeutung für die Erstellung der Trainingsdaten und die spätere Evaluation der Übersetzungen. So stellte sich in der Diskussion mit den Fachleuten schnell heraus, dass bestehende Cloud-Übersetzungslösungen oftmals zu kurz greifen, da diese das Fachvokabular der jeweiligen Domäne der Endkunden nicht beherrschten. Fachleute in ihren jeweiligen Domänen sind allerdings ebenfalls nicht zwangsläufig mit KI vertraut, was oftmals zu einem Bedarf an Schulung und Aufklärung führt. Da KI-Technologie relativ neu ist, kann es wichtig sein, Schulungen und Aufklärungsmaßnahmen durchzuführen, um sicherzustellen, dass das Team und andere Beteiligte ein gutes Verständnis davon haben, was KI kann und was nicht. Dies kann dazu beitragen, Missverständnisse und unerfüllbare Erwartungen zu vermeiden.

Learning: Die Zusammenarbeit mit Fachleuten ist entscheidend, um KI-Modelle auf die Bedürfnisse des Unternehmens und Endnutzers anzupassen. Oftmals ist es notwendig, Schulungen und Aufklärungsmaßnahmen durchzuführen, um sicherzustellen, dass das Team und andere Beteiligte ein gutes Verständnis davon haben, was KI kann und was nicht. Dadurch können Missverständnisse und unrealistische Erwartungen vermieden werden.

In Schulungen kommen auch Ängste zum Vorschein. Diese betreffen vielfach Risiken, die mit dem Einsatz von KI in Verbindung stehen. Im Falle der Übersetzungslösungen wurde beispielsweise davor gewarnt, dass die Übersetzungsqualität unzureichend sein könnte und dies zum Abbruch bestehender Kundenbeziehungen führen würde. Hier ist es wichtig, potenzielle Risiken im Zusammenhang mit dem Einsatz von KI-Modellen zu identifizieren und Strategien zu entwickeln, um diese zu minimieren oder zu vermeiden. Dies kann helfen, Probleme zu lösen, bevor sie auftreten. Ferner ist es wichtig, in den Schulungen Ängsten einen Raum zu geben und begründete Risiken aufzunehmen. Risiken, die messbar sind, wie eine schlechte Übersetzungsqualität oder falsch zugeordnete Tickets im Falle der KI im Kundenservice (s. Abschnitt 3.1) lassen sich am besten durch ein *Monitoring* minimieren. Hier ist es wichtig, Metriken zu definieren, um den Erfolg des KI-Modells zu messen. Diese Metriken sollten auf das Geschäftsproblem abgestimmt sein und helfen, den Fortschritt des Projekts, die Passgenauigkeit der Lösung und die Ergebnisgüte zu verfolgen.

Learning: In KI-Schulungen werden oft Risiken besprochen. Gemeinsam sollten Strategien zur Risikominimierung bzw. -vermeidung entwickelt werden und messbare Risiken durch Monitoring reduziert werden. Es braucht Metriken, um den Erfolg der Lösung zu messen.

4.2 Data Understanding

Daten spielen in KI-Projekten immer eine große Rolle. Zum einen bilden sie als Trainingsdaten die Grundlage, zum anderen können sie auch Gegenstand der Modell-Anwendung sein. Zum Beispiel, wenn es um die Transformation bestehender Texte, Webseiten oder Bilder geht. Was wir erleben ist, dass viele Unternehmen ein falsches oder unvollständiges Bild von ihren Datenbeständen haben. Für viele stellt die Erfassung und Analyse von Daten eine Herausforderung dar. Häufig haben sie keine Daten oder nur wenige und unstrukturierte Daten, die dezentral und unzureichend qualitätsgeprüft gespeichert sind [Bä22]. Um jedoch erfolgreich von KI zu profitieren, ist es wichtig, die vorhandenen Datenquellen und die Art der Daten zu erfassen. Hierbei sollten sowohl interne als auch externe Datenquellen berücksichtigt werden. Auch die Qualität der vorhandenen Daten muss überprüft werden, um sicherzustellen, dass sie vollständig, konsistent, aktuell und korrekt sind.

Learning: Da die Datenlage oft nicht überschaubar, transparent und ideal ist, müssen akademische Partner mehr Aufwand einplanen, um daran zu arbeiten.

In unserer Kooperation zum KI-gestützten multilingualen Content Management (s. Abschnitt 3.2) wurde beispielsweise untersucht, welche Datenquellen verfügbar waren und wie die Daten beschaffen waren. Der Fokus lag hauptsächlich auf der Identifizierung geeigneter Datenquellen zur Extraktion von Texten für das Training der Modelle. Diverse Datenquellen wurden in Betracht gezogen, darunter Textdokumente, HTML-Seiten, Excel-Dokumente und Daten aus Content-Management-Systemen (CMS). Hinsichtlich der Datenbeschaffenheit wurde die Feinheit in Bezug auf domänenspezifische Begriffe betrachtet. Unter anderem bei mehrdeutigen Begriffen wie zum Beispiel „rape“, welcher sich auf eine Pflanzenart beziehen

kann, die auch als Raps bekannt ist, oder auch auf den gewaltsamen sexuellen Übergriff, der als Vergewaltigung bezeichnet wird. Die eigentliche Datenanalyse kann mithilfe verschiedener Analysemethoden durchgeführt werden, um Muster und Zusammenhänge in den Daten zu identifizieren. Allerdings sollten Datenschutz und Datensicherheit bei der Erfassung, Analyse und Verarbeitung der Daten berücksichtigt werden, um die Daten gemäß den geltenden Datenschutzbestimmungen zu speichern und zu verarbeiten [Bä22].

Learning: Die Verfügbarkeit von geeigneten Daten ist ein wichtiger Faktor bei der Umsetzung von KI-Projekten. Es ist wichtig, sicherzustellen, dass genügend Daten vorhanden sind, um ein Modell zu trainieren, das gut genug ist, um das Geschäftsproblem zu lösen.

Wie dargestellt, sind nicht immer Daten im Sinne von Trainingsdaten für KI-Modelle erforderlich. Allerdings ist Data Understanding hier auch umfassender zu verstehen. Eine effektive Datenvisualisierung ist beispielsweise ein weiterer wichtiger Aspekt, um komplexe Zusammenhänge und Muster in den Daten aufzudecken und verständlich zu machen. Hierfür können verschiedene Visualisierungstools eingesetzt werden, um die Ergebnisse der Analyse in Form von Diagrammen, Tabellen oder Grafiken darzustellen. So war es beispielsweise bei unserem Projekt KI im Kundenservice (s. Abschnitt 3.1) von erheblicher Bedeutung, die eingehenden E-Mails in gemeinsamen Gruppentreffen zu sichten und Herausforderungen auf Grund der teils niedrigen Datenqualität abzuleiten und zu diskutieren. In dem Kontext ist es wichtig, sich mit der Struktur und dem Inhalt der Daten zu beschäftigen. Konkret stellte sich die Frage, wie die Tickets aufgebaut sind und welche Informationen in welchem Format gespeichert werden. Es ist relevant zu klären, welche Informationen für die Kategorisierung und die Beantwortung einer Fragestellung relevant sind.

Learning: Eine effektive Datenvisualisierung kann helfen, komplexe Zusammenhänge und Muster in den Daten aufzudecken und für alle Projektbeteiligten verständlich zu machen.

4.3 Data Preparation

Unternehmen stehen oft vor der Herausforderung, einen Datenwildwuchs zu bewältigen. Dies bedeutet, dass Daten in verschiedenen Formaten und aus unterschiedlichen Systemen stammen, was zu Problemen bei der Verarbeitung und Analyse führen kann. Im Kontext des Projektes rund um KI-gestütztes multilinguales Content Management (s. Abschnitt 3.2) wurden spezifische Verarbeitungsmethoden entwickelt, die sich auf die Datenextraktion, -bereinigung und den Aufbau eines parallelen Korpus konzentrierten. Dabei erwies sich das automatische Extrahieren von Informationen aus PDF-Dateien als besonders anspruchsvoll im Vergleich zu anderen Formaten wie Excel-Dokumenten. Die Erstellung eines parallelen Korpus, also einer Sammlung von Texten in verschiedenen Sprachen, erforderte geeignete Verfahren zur Messung der semantischen Ähnlichkeit. Durch eine systematische Normalisierung der Datenformate und die Anwendung spezifischer Verarbeitungsmethoden können Unternehmen einen besseren Zugriff auf ihre Daten erlangen und so die Grundlage für erfolgreiche Datenanalyse und KI-Anwendungen schaffen.

Learning: Datenwildwuchs kann Unternehmen vor Herausforderungen stellen, wenn es darum geht, Daten in verschiedenen Formaten und Systemen zu verarbeiten und zu analysieren. Spezielle Verarbeitungsmethoden und eine systematische Normalisierung der Datenformate können helfen, den Zugriff auf die Daten zu erleichtern. In einigen Fällen ist eine Transformation der Daten notwendig, um sie nutzbar zu machen.

Neben Heterogenität in Herkunft und Format der Daten, ist die Datenqualität oftmals ebenfalls eine Herausforderung. Im Kontext des KI-Projekts zur Verbesserung des Ticket-Bearbeitungsprozesses (s. Abschnitt 3.1) war eine sorgfältige Vorbereitung der Daten von entscheidender Bedeutung. Eingehende E-Mails liegen zum Beispiel in sehr unterschiedlicher Qualität, Länge und Kodierung vor. Für die Klassifikation der Kategorie einer E-Mail ist lediglich der Inhalt bzw. die Intention der Nachrichten relevant. Andere Informationen, wie beispielsweise Kontaktdaten oder Signaturen, sind für diese Aufgabe überflüssig und können sogar zu einer Verzerrung der Ergebnisse führen. Um dies zu vermeiden, müssen Konversationsinhalte aus den Mailverläufen extrahiert werden, um ausschließlich relevante Informationen für die Klassifikation zu erhalten. Hierbei können auch neueste LLMs, wie beispielsweise GPT-4, eingesetzt werden, um die Genauigkeit der Klassifikation zu verbessern. Weiterhin müssen Telefonnummern und Adressen der extrahierten Konversationen durch reguläre Ausdrücke bereinigt werden, um eine einheitliche Datenbasis zu schaffen. Auch das extreme Ungleichgewicht der Kategorien in den verfügbaren Trainingsdaten muss berücksichtigt werden. Unterrepräsentierte Kategorien werden zu einer Restkategorie zusammengeführt, um eine effektive Verwendung der vorhandenen Daten zu gewährleisten.

Insgesamt ist eine effektive Vorbereitung der Daten von entscheidender Bedeutung, um eine erfolgreiche Implementierung von KI-Systemen im Ticket-Bearbeitungsprozess zu erreichen. Durch die Verwendung von modernen LLMs und die Bereinigung der Daten können die Klassifikationsgenauigkeit und die Effektivität des Systems verbessert werden.

4.4 Modeling

Die Auswahl der richtigen Tools ist ein wichtiger Aspekt bei der Implementierung von KI in KMUs. Ein KI-Werkzeugkoffer kann KMUs dabei helfen, die richtigen Tools für ihre Bedürfnisse auszuwählen. Die Wahl der richtigen Tools hängt von den individuellen Bedürfnissen und Zielen des KMUs ab. Basierend auf den Erfahrungen und Herausforderungen, die im Kontext des KI-gestützten multilingualen Content Management (s. Abschnitt 3.2) gesammelt wurden, ist es wichtig, bei der Entwicklung und Implementierung eines KI-Werkzeugkoffers für KMUs folgende Aspekte zu berücksichtigen: Wie im Abschnitt 2 beschrieben, ist die Auswahl des richtigen Frameworks und der passenden Tools entscheidend für den Erfolg von KI-Projekten. KMUs sollten sorgfältig evaluieren, welche Tools und Frameworks am besten zu ihren spezifischen Bedürfnissen passen und diese gegebenenfalls *fine-tunen*, um optimale Ergebnisse zu erzielen. Daraus ergibt sich ein Bedarf an spezieller Hardware. Die Implementierung von KI erfordert oft spezielle Hardware, um die notwendigen Berechnungen durchzuführen. KMUs sollten sicherstellen, dass sie Zugang zu dieser haben und diese

gegebenenfalls anschaffen und konfigurieren. Daneben ist der bereits angeführte Punkt der Datenmenge für das Training von KI-Modellen zu nennen. KI-Modelle benötigen große Datenmengen für das Training [De19]. KMUs sollten sicherstellen, dass sie Zugang zu ausreichend Daten haben, um Modelle zu trainieren. Hier kann die Generierung von synthetischen Daten helfen, den Trainingsprozess zu unterstützen. Oftmals nicht berücksichtigt wird Datenschutz und Sicherheit. Wie bei jeder Technologie, die personenbezogene Daten verarbeitet, ist es wichtig, dass KMUs sicherstellen, dass sie angemessene Datenschutz- und Sicherheitsmaßnahmen implementieren, um die Daten ihrer Kunden zu schützen.

Learning: Insgesamt kann ein KI-Werkzeugkoffer helfen, die Vorteile von KI zu nutzen und Geschäftsprozesse zu optimieren. Es ist jedoch wichtig, dass KMUs die Herausforderungen bei der Entwicklung und Implementierung von KI-Projekten berücksichtigen.

4.5 Evaluation

Die Evaluation von KI-Implementierungen in KMUs ist entscheidend, um sicherzustellen, dass die KI-Lösung die gewünschten Ergebnisse liefert. KMUs haben oft begrenzte Ressourcen, um solche Implementierungen durchzuführen und sind daher besonders anfällig für Fehler oder ineffektive Lösungen. Eine sorgfältige Evaluation kann helfen, diese Risiken zu minimieren und sicherzustellen, dass die KI-Lösung den tatsächlichen Bedarf des Unternehmens erfüllt. Es gibt jedoch auch Herausforderungen, die mit der Evaluation von KI-Implementierungen in KMUs verbunden sind. Eine Herausforderung besteht darin, dass viele KMUs oftmals nicht über das erforderliche Fachwissen verfügen, um die KI-Lösung zu verstehen oder zu bewerten. Eine weitere Herausforderung besteht darin, dass KMUs oft keine ausreichenden Daten haben, um eine effektive Evaluation durchzuführen. Daher kann es schwierig sein, die Qualität der KI-Lösung zu bewerten oder Fehler zu identifizieren.

Im Fall des multilingualen Content Managements (Abschnitt 3.2) war es wichtig, die Mitarbeiter des Unternehmens in den Evaluationsprozess einzubeziehen, um sicherzustellen, dass die Übersetzungsmodelle den Anforderungen des Unternehmens entsprechen. Da es keinen eindeutigen Standard für die Evaluierung von Übersetzungen gibt, waren die Mitarbeiter von entscheidender Bedeutung, um Fehler zu identifizieren und zu korrigieren.

In dem Beispiel der KI-gestützten Kundenbetreuung (Abschnitt 3.1) wurden bewährte Metriken wie *Accuracy*, *Precision*, *Recall* und *F1-Score* verwendet, um die Qualität der Klassifikationsergebnisse zu bewerten. Darüber hinaus wurden die exportierten Ticket-Daten in Trainings- und Testdaten aufgeteilt, um sicherzustellen, dass das Modell in der Lage ist, auf neuen Daten zu generalisieren. Die Anbindung an das Ticket-System ermöglichte es, das Modell in der realen Welt zu testen.

Learning: Insgesamt ist die Evaluation von KI-Implementierungen in KMUs von entscheidender Bedeutung, um sicherzustellen, dass die KI-Lösung den Bedürfnissen des Unternehmens entspricht und effektiv arbeitet. Durch die Verwendung bewährter Metriken,

die Einbindung von Mitarbeitern und die Bereitstellung ausreichender Daten können KMUs sicherstellen, dass ihre KI-Lösungen den gewünschten Nutzen bringen.

4.6 Deployment

Neben der Konzeption und Entwicklung in KI-Projekten ist das Integrieren in bestehende Prozesse und das Bereitstellen der KI-Lösungen immer wieder eine große Herausforderung.

Eine Option für KMUs, die KI-Modelle einsetzen möchten, besteht darin, Cloud-Lösungen zu nutzen. Hierbei wird die Verarbeitung der Daten und das Training der KI-Modelle von einem externen Anbieter übernommen, der die benötigten Ressourcen und Infrastruktur bereitstellt. Cloud-Lösungen bieten KMUs den Vorteil, dass sie schnell und unkompliziert auf die notwendigen Ressourcen zugreifen können, ohne selbst investieren zu müssen. Darüber hinaus sind Cloud-Lösungen in der Regel skalierbar und können an die Bedürfnisse des Unternehmens angepasst werden. Allerdings sollten KMUs sich der Risiken bewusst sein, die mit der Nutzung von Cloud-Lösungen einhergehen. Hierzu gehört zum Beispiel der Verlust der Kontrolle über die Daten und die Abhängigkeit von einem externen Anbieter. Um diese Risiken zu minimieren, sollten KMUs sorgfältig auswählen, welcher Anbieter ihre Daten verarbeitet und wie ihre Daten geschützt werden. Alternativ können KMUs auch entscheiden, ihre KI-Modelle selbst zu hosten, indem sie die notwendige Infrastruktur und Ressourcen bereitstellen. Hierbei sollten KMUs jedoch bedenken, dass sie über die erforderlichen Fähigkeiten und Ressourcen verfügen müssen, um die Modelle zu trainieren, zu implementieren und zu warten. Dies erfordert in der Regel ein hohes Maß an technischer Expertise und Investitionen in die IT-Infrastruktur. Auf der anderen Seite können KMUs durch das Hosting ihrer KI-Modelle die volle Kontrolle über ihre Daten behalten und haben die Möglichkeit, ihre Modelle vollständig an ihre spezifischen Bedürfnisse anzupassen.

Learning: Unabhängig davon, ob KMUs Cloud-Lösungen oder Selbsthosting wählen, müssen sie sicherstellen, dass sie die notwendigen Fähigkeiten und Ressourcen haben, um KI-Modelle erfolgreich zu betreiben. Hierzu gehört zum Beispiel ein fundiertes Verständnis der Technologien, die bei der Entwicklung von KI-Modellen zum Einsatz kommen, sowie die Fähigkeit, Daten effektiv zu sammeln, zu verarbeiten und zu analysieren.

5 Fazit und Ausblick

Die Integration von KI in Unternehmen wird immer wichtiger, um wettbewerbsfähig zu bleiben. Insbesondere KMUs stehen jedoch oft vor der Herausforderung, nicht über die notwendigen Ressourcen, Expertise und Mittel zu verfügen, um eigene KI-Abteilungen aufzubauen. Wie hier dargestellt wurde, werden diese projektspezifischen Herausforderungen oftmals als Eintrittshürde wahrgenommen. Anhand von Kooperationen aus verschiedenen Branchen wurde gezeigt, wie die Arbeitsgruppe Angewandte KI durch eine flexible und

lösungsorientierte Herangehensweise maßgeschneiderte Lösungen für die individuellen Bedürfnisse der Partnerunternehmen entwickelt. Dabei wurden spezifische Herausforderungen, wie die Begrenzung der Ressourcen, die fehlende Expertise und die unklare Vorstellung über die Vorteile von KI-Anwendungsfällen, aufgezeigt und Lösungen präsentiert.

Literatur

- [Bä22] Bäumer, F.; Denisov, S.; Sirvend, B.; Weber, J.: Tackling the We have no Data Challenge: Domain-Specific Machine Translation in SMEs. In (Kersting, J., Hrsg.): Proceedings of the Fourteenth International Conference on Pervasive Patterns and Applications (PATTERNS 2022): Special Track AI-DRSWA. Barcelona, Spanien, S. 1–4, 2022.
- [De19] Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In (Burstein, J.; Doran, C.; Solorio, T., Hrsg.): Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Association for Computational Linguistics, Minneapolis, MI, USA, S. 4171–4186, 2019.
- [Du19] Dukino, C.; Friedrich, M.; Ganz, W.; Hämmerle, M.; Kötter, F.; Meiren, T.; Neuhüttler, J.; Renner, T.; Schuler, S.; Zaiser, H.: Künstliche Intelligenz in der Unternehmenspraxis, hrsg. von Bauer, W.; Ganz, W.; Hämmerle, M.; Renner, T., 2019, URL: <https://publica.fraunhofer.de/handle/publica/300040>.
- [LG21] Lundborg, M.; Gull, I.: Künstliche Intelligenz im Mittelstand—So wird KI für kleine und mittlere Unternehmen zum Game Changer. Begleitforschung Mittelstand-Digital/WIK-Consult/, Abrufdatum: 04.07.2023, 2021, URL: <https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/ki-Studie-2021.pdf>.
- [UF21] Ulrich, P.; Frank, V.: Relevance and Adoption of AI technologies in German SMEs – Results from Survey-Based Research. *Procedia Computer Science* 192/1, Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 25th International Conference KES2021, S. 2152–2159, 2021.
- [Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I.: Attention is All You Need. In (von Luxburg, U.; Guyon, I.; Bengio, S.; Wallach, H.; Fergus, R., Hrsg.): Proceedings of the 31st International Conference on Neural Information Processing Systems. NIPS'17, Curran Associates Inc., Long Beach, California, USA, S. 6000–6010, 2017.
- [WH00] Wirth, R.; Hipp, J.: CRISP-DM: Towards a Standard Process Model for Data Mining. In (Mackin, N., Hrsg.): Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining. Blackpool, Lancashire, UK, 2000.

Synergien zwischen KI und menschlicher Expertise: Online-Sichtbarkeitsoptimierung durch Text-Transformation

Bastian Sirvend,¹ Alexander Stahlkopf,² Frederik Simon Bäumer³

Abstract: Dieser Praxisbericht stellt die Synergien zwischen KI und menschlicher Expertise bei der Online-Sichtbarkeitsoptimierung durch Text-Transformation dar. Es wird gezeigt, wie KI, insbesondere große Sprachmodelle, bei der Suchmaschinenoptimierung bestehender Texte unterstützt und gleichzeitig der „Human in the Loop“-Ansatz gewahrt wird. Eine effektive Integration in Prozesse und Content-Management-Systeme spielt hierbei eine wichtige Rolle. Wir sind davon überzeugt, dass es trotz aufstrebender KI-Lösungen im Enterprise-Umfeld den „Human in the Loop“-Ansatz bedarf.

Keywords: Künstliche Intelligenz; SEO; Online Marketing

1 Einführung

Die Optimierung der Online-Sichtbarkeit ist ein Erfolgsfaktor im heutigen digitalen Zeitalter. Unternehmen und Organisationen streben danach, ihre Webseiten und Inhalte so zu gestalten, dass sie von Suchmaschinen gut gefunden und in den Suchergebnissen möglichst hoch platziert werden. Suchmaschinenoptimierung (SEO) ist eine komplexe Aufgabe, die verschiedene Strategien und Techniken umfasst. Sie beinhaltet jedoch auch wiederkehrende, teils triviale Aufgaben – zum Beispiel das Umformen von Texten. Spätestens seit Veröffentlichung der leistungsstarken Transformer-Modelle [De19; Va17], hat künstliche Intelligenz (KI) einen bedeutenden Einfluss auf die SEO-Branche. Insbesondere große Sprachmodelle, wie GPT-4 [Op23], haben sich als äußerst nützlich erwiesen, um vorhandene Texte zu optimieren und die Sichtbarkeit zu verbessern. Diese KI-Modelle sind in der Lage, Texte zu analysieren, Schlüsselwörter zu identifizieren und Vorschläge zur Verbesserung zu machen.

Unserer Erfahrung nach, wird dieses Potenzial in vielen Unternehmen noch nicht in Gänze ausgeschöpft. Das hat aus unserer Sicht vor allem zwei Gründe: Zum einen sind bestehende KI-Tools häufig nur als externe Softwareprodukte zu erwerben und in der Regel schlecht in bestehende Content-Prozesse integriert. Dies führt zu ineffizienten „Copy-Paste-Prozessen“ und schnell zu Frustration. Zum anderen ist es eine Skepsis gegenüber Content-Automatisierung, insb. Transformation, mittels KI. Die zu pflegenden Webseiten sind das Aushängeschild der Unternehmen, auf dem falsche Informationen und fehlerhaftes Fachvokabular einen schlechten Gesamteindruck erzeugen. Aus diesem Grund leben wir in

¹ wonk.ai GmbH, Stapenhorststr. 50a, 33615 Bielefeld, Deutschland, bastian.sirvend@wonk.ai

² wonk.ai GmbH, Stapenhorststr. 50a, 33615 Bielefeld, Deutschland, alexander.stahlkopf@wonk.ai

³ Hochschule Bielefeld, Interaktion 1, 33619 Bielefeld, Deutschland, frederik.baeumer@hsbi.de

unseren KI-Integrationen den „Human in the Loop“-Ansatz. Menschliche Redakteure und SEO-Experten verfügen über ein tiefes Verständnis der Zielgruppe, der Markenbotschaft und der gewünschten Ziele. Die Kombination von KI und menschlicher Expertise ermöglicht eine effektive Seitenoptimierung. Diese Synergien zwischen KI und menschlicher Expertise werden in diesem Bericht genauer betrachtet. Es wird gezeigt, wie KI-Modelle wie GPT-4 bei der Seitenoptimierung von bestehenden Texten eingesetzt werden können, um die Online-Sichtbarkeit zu verbessern. Gleichzeitig wird der wichtige Beitrag der menschlichen Expertise hervorgehoben. Darüber hinaus wird auf die Bedeutung einer effektiven Integration von KI-Systemen in Prozesse und Content-Management-Systeme (CMS) eingegangen.

2 Gedanken zum Content-Erstellungsprozess

Die manuelle Pflege von Inhalten in CMS sowie die Anpassung an verschiedene Formate und Plattformen stellen Content-Teams vor immense Herausforderungen. Da ist zum einen die händische Pflege von Inhalten in CMS. Die manuelle Eingabe von Inhalten in ein CMS erfordert Zeit und Genauigkeit, um Formatierung, Stil und Konsistenz zu wahren. Dieser Prozess ist anfällig für Fehler und kann zeitaufwändig sein. Um diesen Herausforderungen zu begegnen, ist es entscheidend, automatisierte Workflows zu implementieren, die wiederkehrende Aufgaben vereinfachen und die Effizienz steigern. Dieser Gedanke ist nicht neu und wird von vielen CMS bereits ohne KI umgesetzt. Bei komplexen Aufgaben wie der Anpassung an verschiedene Formate und Plattformen kommen diese Lösungen aber oftmals an ihre Grenzen. Die Vielzahl von Formaten und Plattformen, auf denen Inhalte präsentiert werden müssen, erfordert manuelle Anpassungen, um eine konsistente und ansprechende Darstellung sicherzustellen. Dieser Schritt kann aufgrund der erforderlichen Änderungen an Bildgrößen, Textlängen, Schriftarten und anderen Elementen äußerst mühsam sein. Um diesen Aufwand zu minimieren, bieten sich die Verwendung von Templates und Style-Guides an, die die Formatierung automatisieren und die Inhalte auf verschiedenen Plattformen harmonisieren. Wenn allerdings ein Umschreiben, Kürzen oder Auffrischen eines Textes gefordert ist, greifen diese Ansätze zu kurz. Soll eine Meta-Description oder ein Tweet aus einem Beitrag erstellt werden, sind große Sprachmodelle heute das Werkzeug der Wahl.

Diese Situation verschärft sich im internationalen Kontext. Die Internationalisierung von Inhalten stellt Content-Teams vor zusätzliche Herausforderungen. Neben der reinen Übersetzung müssen auch kulturelle Aspekte und landesspezifische Anforderungen berücksichtigt werden. Dieser Prozess erfordert eine koordinierte Zusammenarbeit, um Inhalte in verschiedenen Sprachen zu erstellen, zu verwalten und zu lokalisieren. Die Integration von Content-Management-Tools mit Übersetzungs- und Lokalisierungsfunktionen können den internationalen Content-Prozess erleichtern und beschleunigen. In Fällen von großen internationalen Roll-Outs, die wir begleitet haben, waren dennoch sehr viele Aufgaben händisch zu erledigen (z. B. Zusammenfassungen). Unserer Erfahrung nach ist es daher notwendig, um den Content-Prozess zu optimieren, Content-Teams auf Lösungsansätze und Technologien zurückgreifen zu lassen, die die manuelle Arbeit reduzieren und die Effizienz

steigern. Automatisierte Workflows, die repetitive Aufgaben automatisieren, ermöglichen eine schnellere und präzisere Eingabe von Inhalten. Der nächste Schritt ist nun, KI-Lösungen zur Text-Transformation als weiteres Werkzeug in die CMS zu integrieren.

3 Text-Transformation in CMS mittels KI

Bei der Text-Transformation mittels KI können verschiedene Aufgaben durchgeführt werden. Nachgefragt sind (1) Automatische Textzusammenfassung, (2) Sprachübersetzung, (3) Stil- und Tonanpassung und (4) Keyword-Optimierung. Wir stellen im Folgenden drei KI-Services vor, wobei alle Services grundsätzlich beliebig kombiniert werden können.

Mittels der automatischen Textzusammenfassung (sog. SummarizR Service) können lange Texte automatisch zusammengefasst werden, um die wichtigsten Informationen hervorzuheben und Ziel-Kanälen gerecht zu werden hinsichtlich vorgegebener Zeichenrahmen. Hier knüpft die Sprachübersetzung an. Dieser KI-Service kann verwendet werden, um Texte automatisch von einer Sprache in eine andere zu übersetzen. Dies ist besonders nützlich für Unternehmen mit internationaler Präsenz, die ihre Inhalte in verschiedenen Sprachen anbieten möchten. Hervorzuheben ist hier der Bedarf im Enterprise-Bereich, auch in den Übersetzungen das Fachvokabular und die Tonalität zu wahren. Hier können KI-Services im Bereich der Stil- und Tonanpassung aushelfen (sog. TransformR-Service). Mit diesen können Texte entsprechend einem gewünschten Stil oder Tonfall angepasst werden. Der klassischen SEO-Aktivität eher zugehörig ist die KI-basierte Keyword-Optimierung. Durch den Einsatz von KI können Texte analysiert, geschrieben und mit relevanten Keywords optimiert werden, um die Sichtbarkeit in Suchmaschinen zu verbessern. Dies trägt dazu bei, dass die Inhalte von potenziellen Nutzern leichter gefunden werden können (sog. GhostwritR-Service).

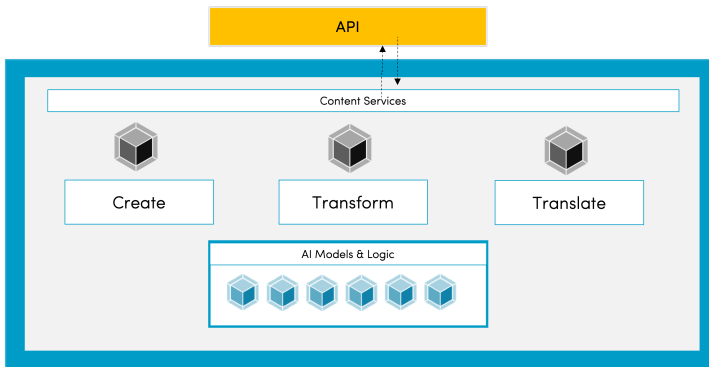


Abb. 1: Service-Bereitstellung trotz häufiger Umstellung des Modell-Fundaments durch Abstraktion

Bei der Integration sollten einige Punkte beachtet werden. Entwicklungen im Bereich der maschinellen Sprachverarbeitung unterliegen einer immensen Geschwindigkeit. Dies führt dazu, dass zum einen der Anspruch an generierte Resultate seitens der Anwender stetig steigt, zum anderen damit aber auch der Druck auf KI-Provider zunimmt, neuste

Modelle bereitzustellen. Jedoch bleibt ungeachtet von Veränderungen im technologischen Unterbau oder der Modellstruktur die konsistente Bereitstellung von Services das oberste Ziel. Selbst bei einem Wechsel im Modell-Unterbau ist alles daran zu setzen, Anwendern ein kontinuierlich effizientes und zuverlässiges Serviceerlebnis zu bieten. Hierfür werden umfassende Planung, gründliche Tests und kontinuierliche Optimierung eingesetzt, um sicherzustellen, dass der Übergang nahtlos verläuft und keine Kompromisse in Bezug auf die Qualität eingegangen werden. Am besten ist dies durch eine stark modulare Integration zu erreichen, in dem KI-APIs von den eigentlichen Modellen abstrahiert werden (s. Abb. 1).

Auch gehören gerade im Enterprise-Kontext die Punkte Sicherheit und Datenschutz in die Integrationsüberlegungen. Bei der Verwendung von KI-APIs sollten Maßnahmen ergriffen werden, um sicherzustellen, dass die Daten während der Transformation sicher übertragen und verarbeitet werden. So ist darauf zu achten, wo die Daten verarbeitet werden (z. B. USA im Falle von OpenAI). Daneben ist ein wichtiger Aspekt die Feinabstimmung. Beispielsweise ist es bei Übersetzungslösungen möglich, bestehende TMS-Systeme als Trainingsgrundlage heranzuziehen und die Übersetzungsqualität zu erhöhen [Bä22]. Schlussendlich ist Qualitätssicherung wichtig. Die generierten Texte sind regelmäßig auf Qualität zu prüfen. Dies kann manuell oder mithilfe von Feedback-Schleifen erfolgen, um sicherzustellen, dass die Texte den gewünschten Standards entspricht.

Es ist wichtig, dass der Mensch stets die Hoheit über den Text behält. Durch eine Zusammenarbeit zwischen Mensch und Maschine können Synergien entstehen, wobei jeder seine Stärken einbringt und die Integration in das CMS eine zentrale Rolle spielt. Im ersten Schritt erstellen und pflegen die Autoren die Inhalte im CMS. Sie haben die volle Kontrolle über den Inhalt, die Struktur und den Stil. Sobald ein Text erstellt oder aktualisiert wird, tritt die KI in Aktion und unterstützt bei der Text-Transformation. Die KI analysiert den Text und bietet Vorschläge zur Verbesserung an. Dies umfasst stilistische Anpassungen, Grammatik- und Rechtschreibkorrekturen sowie Optimierung für SEO (sog. TransformR-Service). Zusätzlich kann die KI relevante Informationen aus anderen Quellen oder bereits vorhandenen Inhalten im CMS ziehen und dem Autor als Vorschläge präsentieren (sog. GhostwritR-Service). Der Mensch steht dabei im Mittelpunkt des Prozesses. Die KI generiert Vorschläge und markiert potenzielle Änderungen im Text, die der Autor überprüfen kann. Der menschliche Autor behält die Kontrolle und kann die Vorschläge der KI akzeptieren, ablehnen oder bearbeiten. Darüber hinaus kann der Autor zusätzliche Informationen oder Kontext hinzufügen, die für die KI nicht zugänglich sind. Letztendlich trifft der Autor die Entscheidung über die Text-Transformation und übernimmt die Verantwortung für den Inhalt.

4 Fazit

Die Synergien zwischen Mensch und Maschine liegen auf der Hand. Die KI kann dem menschlichen Autor Zeit sparen, indem sie automatisch stilistische Verbesserungen oder Korrekturen vorschlägt. Zudem kann sie den Autor mit relevanten Informationen und Ergänzungen unterstützen, indem sie auf bereits vorhandene Inhalte im CMS oder externen

Quellen zugreift. Dadurch lassen sich konsistente und hochwertige Inhalte erstellen. Die KI kann dabei große Textmengen schnell analysieren, stilistische Muster erkennen und präzise Verbesserungsvorschläge liefern. Dennoch gibt es Bereiche, in denen der Mensch besser ist. Der menschliche Autor beherrscht die spezifische Markenstimme, die Stilrichtlinien und die Zielgruppenpräferenzen besser als die KI. Der Autor kann kontextbezogene Entscheidungen treffen und komplexe Emotionen oder Nuancen im Text berücksichtigen, die für eine KI möglicherweise schwierig sind. Der Autor kann kreatives Denken, Originalität und persönlichen Ausdruck in den Text einbringen, die für eine KI schwer zu replizieren sind. Schließlich hat der Autor das letzte Wort bei der Text-Transformation und kann persönliche Präferenzen und Branchenkenntnisse einbringen, um den Text optimal anzupassen.

Um einen reibungslosen Workflow zu gewährleisten, ist die KI-Lösung nahtlos in das CMS zu integrieren. Sobald ein Text eingegeben oder aktualisiert wird, erfolgt automatisch eine Analyse durch die KI. Die generierten Vorschläge werden dem Autor in einer intuitiven Benutzeroberfläche präsentiert, die es ihm ermöglicht, die Vorschläge zu betrachten, zu bearbeiten und zu übernehmen. Dieser Prozess wird nahtlos in den bestehenden CMS-Workflow integriert, um eine effiziente Zusammenarbeit zwischen Mensch und Maschine zu gewährleisten. Nur so ermöglicht dieser Prozess eine harmonische Zusammenarbeit zwischen Mensch und Maschine bei der Inhaltspflege im CMS. Die KI unterstützt den menschlichen Autor, indem sie Vorschläge zur Text-Transformation liefert, während der Autor die ultimative Kontrolle über den Text behält und die finale Entscheidung trifft. Letztendlich ermöglicht dieser Miteinander-Ansatz von Mensch und Maschine im CMS eine effiziente Text-Transformation und fördert die Erstellung von hochwertigen Inhalten. Indem die Stärken der KI und des Menschen kombiniert werden, können Synergien entstehen, die zu einer verbesserten Produktivität und Qualität der Inhalte führen.

Literatur

- [Bä22] Bäumer, F.; Denisov, S.; Sirvend, B.; Weber, J.: Tackling the “We have no Data” Challenge: Domain-Specific Machine Translation in SMEs. In: Proc. of the 14. Int. Conf. on Pervasive Patterns and Applications: AI-DRSWA: Maturing Artificial Intelligence - Data Science for Real-World Applications. Barcelona, ES, 2022.
- [De19] Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proc. of the 2019 Conf. of the North American Chapter of the ACL: Human Language Technologies, Volume 1. ACL, Minneapolis, MI, USA, S. 4171–4186, 2019.
- [Op23] OpenAI: GPT-4 Technical Report, 2023, arXiv: 2303.08774 [cs.CL].
- [Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I.: Attention is All you Need. In (Guyon, I.; Luxburg, U. V.; Bengio, S.; Wallach, H.; Fergus, R.; Vishwanathan, S.; Garnett, R., Hrsg.): Advances in Neural Information Proc. Systems. Bd. 30, Curran Associates, Minneapolis, MI, USA, 2017.

Classifying figures and illustrations in electronics datasheets: A comparative evaluation of recent computer vision models on a custom collection of 4000 technical documents

Lymperis Perakis,¹ Julian Balling,² Frank Binder,² Gerhard Heyer,² Franz Kreupl³

Abstract: We report findings from a comparative evaluation of several recent object detection models applied to a domain-specific use case in technical document analysis and graphics recognition. More specifically, we apply models from the EfficientDet and YOLO model families to detect and classify figures in electronics datasheets according to a custom classification scheme. We identify YOLOv7-D6 as the most accurate model in our study and show that it can successfully solve this task. We highlight an iterative approach to figure annotation in document page images for creating a comprehensive and balanced custom dataset for our use case. In our experiments, the object detection models show impressive performance levels on par with state-of-the-art results from the literature and related studies.

Keywords: Computer Vision; Object Detection; Document Analysis; Graphics Recognition; Electronic Design Automation; Machine Learning

1 Introduction

The design and development process of electronics components, for example, for automotive or IoT markets, is complex and time-consuming. The market need for embedded software and hardware engineers is continuously growing and cannot be covered by the number of professionals in the market. Using intelligent software algorithms and AI can automate some development tasks, enabling a broader spectrum of companies, engineers, and technicians to contribute to the development of embedded systems. In this context, many challenges must be addressed to better assist engineers and to automate certain process steps in hardware and software development.

The KIRESys project aims to contribute to the automation in hardware development towards an automatic generation of PCB layouts. A prerequisite for such automatic generation is the availability of machine-processable knowledge and information on electronic components. However, to date such knowledge is primarily represented texts, tables, and figures in PDF

¹ CELUS GmbH, Ridlerstraße 57, 80339 Munich, Germany lymperis.perakis@celus.io

² Institute for Applied Informatics (InfAI) e.V. at Leipzig University, Goerdelerring 9, 04109 Leipzig, Germany
{balling,binder,heyer}@infai.org

³ Department of Hybrid Electronic Systems, TUM School of Computation, Information and Technology,
Technical University of Munich, Arcisstraße 21, 80333 München, Germany franz.kreupl@tum.de

documents, i.e., datasheets and reference designs, intended to be consumed by humans. This paper targets one particular challenge in information extraction from these richly structured documents: Extracting and classifying illustrations and figures from electronic datasheets and reference designs.⁴

This paper reports the results of our study on the following questions: How well can current machine learning approaches recognize, extract, and classify illustrations and figures from electronic component datasheets. Which results can currently be achieved with different state-of-the-art and efficient object detection models? To answer those questions, we developed and curated a custom dataset from datasheets from the design automation domain and compared the performance of several state-of-the-art object detection models from computer vision when applied to detecting and classifying figures and illustrations in these documents.

2 Background and related work

2.1 Information extraction from technical documents

Information extraction from documents is a long-standing research area with a multitude of subfields, as documented, for instance, by the International Association for Pattern Recognition (IAPR)'s established series of conferences, workshops, and journals in the field.⁵ Information extraction from non-textual contents in document images had started to thrive as early as the 1980s, driven by industry needs for establishing computer-assisted design processes (CAD) and geographical information systems (GIS), among others [LR14]. An early example of aligning visual and textual information for product and component data, such as interpreting electronic diagrams and engineering drawings for conversion into operational CAD data, was presented in 1997 by Boeing [BBK98], as reported in [LO14]. It is also noted there, that as of 10 years ago (and still today) no generalized and comprehensive solution has been found, but that many approaches focus on information spotting and partial interpretation [LO14].

However, in recent years, significant developments in machine learning and computer vision, as well as the availability of large curated datasets for training deep learning models have brought new potential into the field of information extraction from technical documents. For example, state-of-the-art computer vision models, such as Mask R-CNN [He20], can be successfully leveraged for document layout analysis, which also entails the detection

⁴ *Reference designs* are manufacturer recommendations for the correct and specification-compliant implementation of electronic components in an overall system. They are usually available as PDF documents.

⁵ Consider, for example, the International Conference on Document Analysis and Recognition (ICDAR), the International Journal on Document Analysis and Recognition (IJ DAR), the International Workshop on Document Analysis Systems (DAS), the International Workshop on Graphics Recognition (GREC), some of which host serial competitions, such as the *Robust Reading Competition* linking the document analysis and computer vision communities [Ka13; Ya17], cf. <https://rrc.cvc.uab.es/>.

of figures and illustrations on document page images [Sh21]. Such improvements are fueled by the availability of large public training data sets as PubLayNet [ZTY19] and DocLayNet [Pf22]. In addition, recent systems for information extraction from documents “born digital”⁶ successfully integrate data from multiple modalities into coherent knowledge bases by incorporating textual, structural, tabular and visual cues [Wu18]. One common aspect of most comprehensive systems and related approaches, as seen from 1998 [BBK98] through 2018 [Wu18] and beyond [OSP21], has been the effective integration of user feedback [Bi22] and the design for efficient human-machine-cooperation [Op22]. To achieve that, the automatic information extraction steps have to be performed accurately and fast. For that reason, closely related studies, such as recent work on figure extraction from electronics datasheets by Chen et al. [Ch21], have relied and improved upon (earlier) state-of-the-art object detection models from computer vision.

2.2 Object detection with EfficientDet and the YOLO model families

Object detection has seen an increasing interest over the past years, mainly due to the general advance of technology in computer vision with the help of deep learning. While image classification refers to categorizing the whole image into classes, object detection involves classification and localization tasks where multiple objects may be present in one image. More precisely, object recognition discerns instances of objects of a specific class in an image and then highlights them, usually delimiting them with bounding boxes [SSD19]. Historically we can divide the progress of object detection into two periods: the “traditional object detection period (before 2014)” and the “deep learning-based detection period (after 2014)” [Zo23].

Modern deep learning object detection models can be divided into two main types: one-stage and two-stage detectors. Two-stage detectors have an extra step that is called region proposal. One-stage object detection models skip the region proposal and run detection directly over a dense sampling of locations. These models usually have faster inference (possibly at the cost of accuracy). They are more suitable for real-time applications, and sample models include YOLO [Re16], SSD [Li16], RetinaNet [Li17b], and EfficientDet [TPL20]. Two-stage methods prioritize detection accuracy, and sample models include FPN [Li17a], Faster R-CNN [Re17], and Mask R-CNN [He20].

Two families of object detectors were used for the present study: EfficientDet and YOLO. EfficientDet [TPL20] is a recent object detection family that is based on EfficientNet [TL19]

⁶ For a definition of “documents born digital” versus printed media and scanned documents, we refer to [HL14]. Documents born digital do not require the application of optical character recognition and can be processed with specific software libraries, such as *Pdfminer.six* or *pdfplumber*. However, extracting content in a meaningful and robustly automated way still poses challenges, as the PDF format by design is not intended to be machine-readable but rather encodes a visual document representation for human consumption.

as a backbone⁷, and adds a new neck layer to extend the functionality of the *Path Aggregation Network* (PAN). Similarly to EfficientNet, it applies compound scaling to optimize both accuracy and efficiency.

YOLO (“You only look once”) is a series of prevalent one-stage object detection model families that has seen several iterations with industry-leading performance when considering speed and accuracy in object detection. The initial release was based on its creator’s *Darknet* framework [Re13; Re16], followed by iterative improvements, YOLOv2[RF17] and YOLOv3 [RF18]. The YOLOv4 architecture was then proposed by the maintainer of the Darknet framework at that time [BWL20]. Its backbone utilizes a variant of the *Cross Stage Partial* (CSP) network [Wa20]. It also included techniques, such as mosaic data augmentation and hyper-parameter optimization using genetic algorithms. This made YOLOv4 the state-of-art detector based on its speed and accuracy compared to available alternative detectors.

The YOLO versions used in this study are YOLOv5 [Jo22], YOLOv6 [Li22], and YOLOv7 [WBL22]. An overview of the distinct EfficientDet and YOLO architectural choices can be found in Table 1.

Models	Backbone	Neck	Head	Loss Function	Anchor	Augm.
EfficientDet	EfficientNet	Bi-FPN	Coupled	Focal	Yes	No
YOLOv5	CSPDarknet53	PAN	YOLOv3	BCE with Logits Loss	Yes	Yes
YOLOv6	RepVGG or CSPRepStack	Rep-PAN	Efficient Decoupled	VFL + IoU Loss	No	Yes
YOLOv7	E-ELAN	PAN	Lead + Auxiliary	VFL + DFL	Yes	Yes

Tab. 1: Overview of the architectures of EfficientDet and different YOLO versions. The models are compared based on their backbone, neck, head, loss function, anchor method choices, and whether they use augmentation in their implementation. This table was developed based on the following sources: [Li22; TPL20; UI21; WBL22].

YOLOv5 is a model developed and published by Ultralytics shortly after YOLOv4 was introduced. The implementation of the network changed from Darknet to Pytorch. Similarly to YOLOv4, the model’s improvements include data augmentation and hyper-parameter optimization. Since its initial release, Ultralytics have continued to modify YOLOv5’s architecture to further improve exportability and speed [UI21]. Chronologically, YOLOv7 was published next [WBL22], which was developed in cooperation with the creator of YOLOv4. YOLOv7 achieved state-of-the-art for real-time object detectors. Compared to

⁷ The basic architectural pattern of one-stage object detectors comprises a *backbone* network for feature extraction from input images, a *neck* for multi-resolution feature aggregation, and a *head* for predicting the objects with their bounding boxes and classes.

YOLOv5 on a model with a similar *mean Average Precision* (mAP), YOLOv7 is 120% faster on inference time in a batch size of 1. Shortly after, YOLOv6 was released [Li22]. The authors propose two scaled re-parameterizable backbones and necks to serve efficient and hardware-friendly models of different sizes.

3 Methods

3.1 Data collection, preprocessing, and pre-labeling

Supervised learning techniques require labeled data to train the models. For our specific use case we collected and developed a custom dataset. We manually downloaded over 4000 electronic component datasheets from the website of a large international electronic components distributor. Only documents from the *Interface ICs* category and its subcategories were considered, since that already yielded a fairly large collection of born-digital PDF documents with a sufficient number and variability of figures and illustrations.

Since object detection models need a visual image as input, we converted the PDFs to document page images in PNG format. We then used LayoutParser [Sh21] with an existing object detection model to generate pre-labeled bounding boxes for figures on the document pages, and to remove all page images without figures from the dataset.⁸

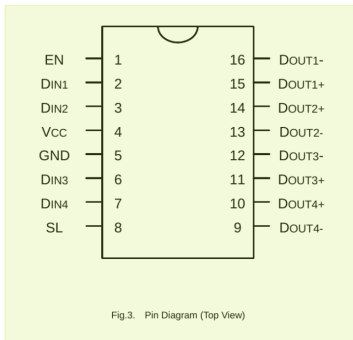
3.2 Data labeling and bounding box optimization

Given our use case, the following classes of figures in electronics datasheets were identified as relevant by CELUS' electronics engineering experts: *Package General*, *Schematic*, *Block Diagram*, *Timing*, *Pin Assignment*, *Memory*, *Plot*, *Package Drawing*, *Footprint*, *Other*. After manually labeling more than 1,000 figures on more than 600 pages, we outsourced the remaining labeling workload to a data annotation and labeling service provider, for which we composed and piloted a comprehensive labeling guide. The resulting bounding box labels were optimized using a technique inspired by [Ch21] to improve the precision and consistency of the ground truth annotations. We implemented their *Bounding Box Refinement* method to shrink and fit the bounding boxes to their respective visual objects. Figure 1 shows an example of the technique used on our dataset.

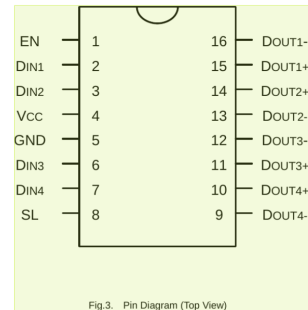
3.3 Model training with iterative pre-labeling of data piles

To prepare for model training and evaluation, we randomly split our labeled dataset into training, validation, and test subsets of 80%, 10%, and 10% respectively. The test set

⁸ From LayoutParser's model zoo we used an available Mask R-CNN model, which had been trained on the PubLayNet dataset [ZTY19] and which could identify (but not yet classify) figures in our dataset.



(a)



(b)

Fig. 1: Comparison of the bounding box before (a) and after (b) applying the *Bounding Box Refinement* on a *Pin Assignment* label. The excess white space around the object is removed. Note that we include figure captions inside our labels whenever possible.

underwent extended quality assurance, to ensure that the evaluation mirrors real-life results. Finally, we exported and transformed the respective data to the different data formats required by the various models.

We then trained our models using the following Python libraries: For *EfficientDet*, we used Ross Wightman’s PyTorch implementation [Wi], whose performance on the COCO dataset is on par with the official results published in [TPL20]. We trained 4 model variants for EfficientDet.⁹ For *YOLOv5*, we used version 6.0 from the official repository [UI21]. Here we trained 6 models. As an ablation study, we trained the YOLOv5m model a second time without augmentations to observe their effect on the model’s performance. For *YOLOv6*, we trained 5 models using the official repository [Me]. For *YOLOv7*, we trained 4 models using the official repository of the authors [Ki].

Table 2 lists the models and their characteristics from literature. For all model trainings, we used the default parameters and hyperparameters. All models were trained for 100 epochs with maximum batch sizes ranging from 4 for larger models to 64 for smaller ones. We utilized the default weights of the models, gained through pre-training on the COCO dataset, as provided by the authors’ libraries. Note that the implementation of EfficientDet does not include any augmentations, contrary to the YOLO families. We expect that the augmentations will improve the YOLO models’ performance without sacrificing speed. Performance of the models was evaluated on the COCO dataset.

In order to ensure a sufficient number of training items per class, we iterated over the steps from pre-labeling to model training as follows: As noted above, we first used Mask-RCNN, trained on PubLayNet, to filter out the pages without figures and to pre-annotate a first pile

⁹ The plan was to train the larger models as well, but our cluster could not provide the required amount of RAM.

of data. We then had the figures labeled for this first pile, and trained the EfficientDet model with those labels. Thus, we then had a model that could both detect and classify figures in our datasheets. This allowed us to estimate the distribution of figure classes, which we used for the subsequent iteration to *roughly balance* the dataset. So, we used the predictions of EfficientDet for pre-labeling and filtering pages when creating the second pile. We then had the second pile labeled, and trained YOLOv5 with it, which we used for pre-labeling and filtering of the third pile, and so on. After training the various models we evaluated their performance on the held-out test set. Results are reported in section 4.2.

Model	Image Size	mAP %	mAP ⁵⁰ %	Speed V100 b1 (ms)	Speed V100 b32 (ms)	Params (M)	FLOPS (G)
EfficientDet-D0	512	34.2	53.0	-	-	3.88	2.5
EfficientDet-D1	640	39.4	59.1	-	-	6.62	6.1
EfficientDet-D2	768	43.4	62.7	-	-	8.1	11
EfficientDet-D3	896	47.1	65.9	-	-	12.0	25
YOLOv5n	640	28.0	45.7	6.3	0.6	1.9	4.5
YOLOv5s	640	37.4	56.8	6.4	0.9	7.2	16.5
YOLOv5m	640	45.4	64.1	8.2	1.7	21.2	49.0
YOLOv5l	640	49.0	67.3	10.1	2.7	46.5	109.1
YOLOv5x	640	50.7	68.9	12.1	4.8	86.7	205.7
YOLOv5l6	1280	53.7	71.3	15.8	10.5	76.8	445.6
YOLOv7-tiny	640	30.8	47.3	-	-	6.2	5.8
YOLOv7	640	51.4	69.7	6.2	2.8	36.9	104.7
YOLOv7-X	640	53.1	71.2	8.8	4.3	71.3	189.9
YOLOv7-D6	1280	56.6	74	22.7	15.0	154.7	806.8
YOLOv6-N	640	36.3	51.2	-	-	4.3	11.1
YOLOv6-T	640	41.1	56.6	-	-	15.0	36.7
YOLOv6-S	640	43.8	60.4	-	-	17.2	44.2
YOLOv6-M	640	49.5	66.8	-	-	34.3	82.2
YOLOv6-L	640	52.5	70.0	-	-	58.5	144.0

Tab. 2: Characteristics of the models used in this study as reported in [Ki; Li22; Me; TPL20; U121; WBL22; Wi] prior to conducting our own experiments (cf. Table 4). The primary challenge metric on the COCO dataset is mAP calculated at IoU=.50:.05:.95. Speed metrics for YOLOv5 and YOLOv7 are reported for batch sizes of 1 and 32 in ms on NVIDIA Tesla V100 GPUs. As other models were tested on different setups, their speed metrics are omitted here, but findings from our own experiments are listed in Table 4. The number of model parameters is given in millions. FLOPS are largely determined by input resolution, i.e. 640x640 or 1280x1280.

4 Results and discussion

4.1 Data set

For our use case in supporting electronics design automation with information extraction from datasheets and reference designs, we built a custom dataset of 17019 page images with figures from over 4000 born-digital PDF documents of electronic component datasheets. This page image dataset contains 28093 high-quality labels for a custom set of 10 categories of figures. The diverse dataset contains labeled figures for electronic components from 26 subcategories of *Interface ICs*. We used datasheets from over 66 manufacturers with more than 2300 components to achieve that. Around 3% of the pages in our dataset include no figures. As reported in Table 3, the dataset is fairly balanced.¹⁰

Figure Type	Number of Figures				Number of Pages
	All	Train	Val	Test	All
Block Diagram	2941	2347	296	298	2581
Footprint	1306	1015	143	148	1220
Memory	2796	2243	282	271	1860
Other	2843	2357	273	213	2190
Package Drawing	3488	2818	340	330	3207
Package General	1749	1406	161	182	1458
Pin Assignment	2789	2286	252	251	2289
Plot	3436	2901	255	280	1038
Schematic	3017	2432	281	304	2093
Timing	3728	2945	394	389	2469
Total	28093	22750	2677	2666	-

Tab. 3: Overview of the number of figures in our self-created electronic components datasheet collection. We display the number of samples in each figure category and the number of pages containing figures from each category.

4.2 Model performance scores after fine-tuning

Using our custom dataset, we fine-tuned and comparatively evaluated object detection models of various sizes from four recent model families: *EfficientDet*, *YOLOv5*, *YOLOv7*, and *YOLOv6*. We ran all validation experiments on the same hardware setup using an NVIDIA GTX 1080 Ti GPU, with 11GB RAM. Table 4 lists the inference performance scores of the fine-tuned models regarding their *mAP*, *precision*, *recall*, and *speed*.

¹⁰ The categories *Footprint* and *Package General* occur only limited times in our datasheet collection. Fortunately, they appear very homogeneous between datasheets, which will still allow to successfully train the models.

Model	mAP %	mAP ⁵⁰ %	P ⁵⁰ %	R ⁵⁰ %	Speed b1 (ms)	Speed b16 (ms)	Speed b32 (ms)
EfficientDet-D0	82.0	88.4	-	-	23.0	10.9	10.9
EfficientDet-D1	83.5	88.7	-	-	35.0	21.6	21.3
EfficientDet-D2	84.5	89.6	-	-	49.0	33.8	33.3
EfficientDet-D3	85.6	90.4	-	-	84.0	64.18	-
YOLOv5n	83.37	91.12	85.69	86.46	6.6	2.3	2.2
YOLOv5s	85.62	90.73	85.5	88.64	7.7	4.0	4.2
YOLOv5m	87.22	91.51	85.58	88.65	13.0	8.3	8.0
YOLOv5m*	80.09	86.38	84.36	82.04	14.0	8.8	8.4
YOLOv5l	87.88	91.87	85.65	88.97	19.8	13.5	13.4
YOLOv5x	87.93	91.32	87.1	87.52	34.6	23.2	22.9
YOLOv5l6	88.37	91.75	86.46	88.26	62.9	52.8	-
YOLOv7-tiny	85.68	91.58	85.64	87.84	6.9	4.7	4.4
YOLOv7	87.75	91.98	87.32	88.38	19.8	14.9	15.7
YOLOv7-X	87.75	91.87	85.34	88.54	31.0	24.4	26.6
YOLOv7-D6	88.75	93.08	87.64	89.18	101.1	85.5	-
YOLOv6-N	85.6	90.4	-	-	7.1	2.4	2.3
YOLOv6-T	86.7	90.7	-	-	7.1	4.3	3.4
YOLOv6-S	86.3	90.3	-	-	8.1	5.4	4.6
YOLOv6-M	87.3	90.8	-	-	12.9	10.2	9.3
YOLOv6-L	87.3	90.6	-	-	19.8	14.5	14.4

Tab. 4: Performance comparison after finetuning: Overall mean Average Precision (mAP); mAP, precision and recall at an IoU=0.5; inference speed with batch sizes of 1, 16, and 32. EfficientDet and YOLOv6 libraries did not provide recall and precision metrics. *YOLOv5m** is the same model as the YOLOv5m without any augmentation during training. Best scores in each family are shown in bold.

4.3 Comparison of model performances

As expected from the literature, the *EfficientDet* models fall behind in performance and speed compared to the YOLO variants. Additionally, the library used for training EfficientDet posed problematic RAM requirements, it was impossible to complete the training for models larger than EfficientDet-D3. In general, we observe that the bigger the model the better its performance on all metrics, except for inference speed, which was slower for larger models.

The *YOLOv5* family' larger models showed the best trade-off between mAP and speed, as seen in Figure 2. Regarding the effect of augmentations, we find that the model finetuned without any augmentation (*YOLOv5m**) delivers the worst performance among all experiments. Specifically, its mAP is 7% worse than the equivalent YOLOv5m model. This underlines the importance of using augmentations during fine-tuning. Figure 2 also shows that the mAP stagnates for those large models of YOLOv5 and YOLOv7 that still use an

image size of 640. YOLOv516 and YOLOv7-D6 use a larger input image size of 1280 and show a steeper improvement compared to their smaller counterparts.

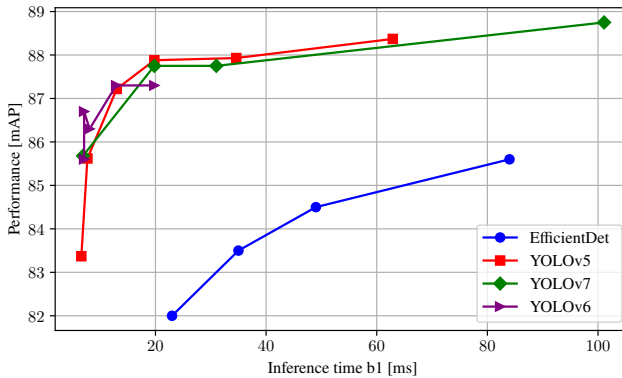


Fig. 2: Performance (mAP) of the trained models regarding their inference time in ms with a batch size of 1.

YOLOv7-D6 of the *YOLOv7* family has the best accuracy metrics of all models in our study, but it is also the slowest. We will analyze it further in Section 4.4. The smaller models of the newest *YOLOv6* family have the best trade-off between mAP and inference time. Specifically, the YOLOv6-T achieves a mAP of 0.87 with a speed of almost 300fps when using a batch size of 32. Its bigger siblings, YOLOv6-M and YOLOv6-L, are not improving much compared to equivalent ones from the other YOLO families. This may be because the larger models have different architectural choices, as listed in table 1. Another essential aspect is that the YOLOv6 models were faster to train and would converge sooner, thus potentially saving computing cost.

4.4 Best model results and related SOTA

The best model in terms of accuracy in our study is YOLOv7-D6. In Table 5, we can see its performance across the different classes. Notice that the recall is relatively high in all classes except the *Other*. Hence, the model can identify most figures in the dataset. The precision is also relatively high, showing that the model primarily identifies relevant objects.

The confusion matrix in figure 3 shows that some classes are interchanged more than others. Specifically, as expected, the model confuses the categories *Block Diagram* and *Schematics*. They have many similarities, and even electronic experts have difficulties categorizing some examples in our dataset. Another occasion that the confusion matrix shows improvement potential for our dataset is the *Timing vs. Plot*. Timing is a specific category of plots, and probably not perfectly consistent across our dataset.

Figure Type	mAP	mAP ⁵⁰	p ⁵⁰	R ⁵⁰
all	0.888	0.931	0.876	0.892
Block Diagram	0.883	0.904	0.822	0.873
Footprint	0.878	0.913	0.831	0.918
Memory	0.826	0.891	0.8	0.886
Other	0.802	0.836	0.794	0.715
Package Drawing	0.88	0.977	0.951	0.939
Package General	0.938	0.977	0.94	0.951
Pin Assignment	0.929	0.98	0.937	0.96
Plot	0.952	0.971	0.964	0.909
Schematic	0.872	0.916	0.858	0.851
Timing	0.915	0.944	0.867	0.916

Tab. 5: Metrics over classes for our best model YOLOv7-D6. The table gives the values of mAP as well as mAP, precision, and recall at an IoU=0.5 over each class.

Although comparing experimental results obtained on different datasets is debatable, we will try to analyze the similarities and differences to the study by Chen et al. [Ch21]. They use the following figure types: *Schematic Diagram*, *Package Diagram*, *Timing Diagram*, *Footprint Diagram* and *Characteristic Diagram* and also include the categories *Header*, *Footer*, and *Table* in their data set. When considering only the categories *Block Diagram*, *Schematic*, *Plot*, *Pin Assignment*, *Package Drawing*, and *Timing*, which appear to be very similar between our study and theirs, the performance scores of our model are: $mAP^{50} = 0.949$, $P^{50} = 0.9$, and $R^{50} = 0.908$ (cf. table 5). Their model has the following scores: $mAP^{50} = 0.907$, $P^{50} = 0.933$, and $R^{50} = 0.903$. Hence, it appears that the model in our experiments outperforms theirs regarding mAP and recall, but falls short in precision. We assume that, since we have defined more specific and mutually similar categories, e.g., *Block Diagram*, and *Schematic*, the precision of our model regresses as it confuses these categories, while Chen et al. [Ch21] group them into one figure type, which avoids confusion.

4.5 Results discussion and potential improvements

To gain further improvements, our approach could be improved in several ways. Additional rigorous data QA with domain experts and data scientists, possibly guided by confusion matrices such as in figure 3, would most certainly identify some remaining labeling inconsistencies, and could further improve data quality and model performance. During preprocessing, we used a pretrained model to filter out document pages without any figures. False negatives from that step are missing in our dataset. To determine the impact of this method, we should evaluate our models in real-world scenarios where all the pages of the datasheets are assessed [AF18]. Using layout analysis models trained on DocLayNet [Pf22] (instead of the earlier PubLayNet [ZTY19]) during prelabeling could also have a positive effect on the results. Furthermore, we could extend our dataset to contain more images from

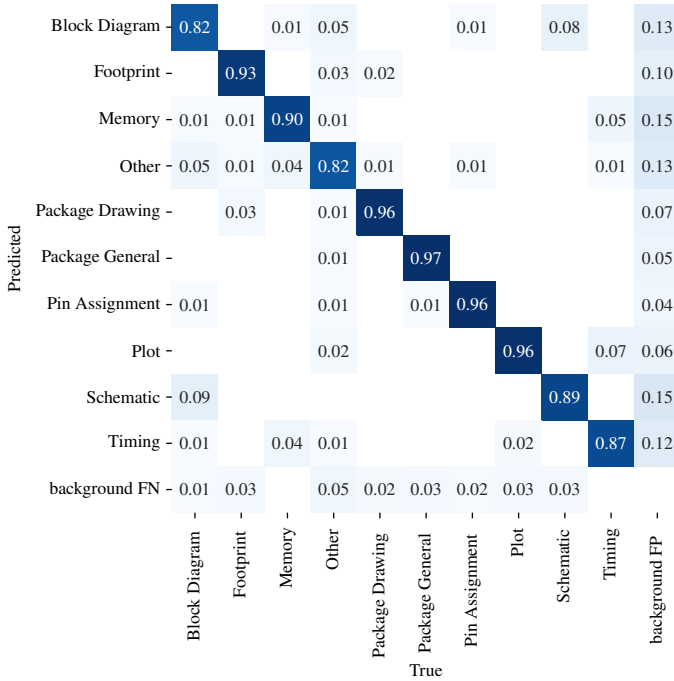


Fig. 3: Confusion matrix of the best model (YOLOv7-D6). The model’s predictions are on the vertical axis, while the ground truth labels are aligned horizontally. For instance, 97% of the *Package General* were correctly classified. The *background FP* are the false positives, and the *background FN* are figures that the model failed to recognize.

an even broader range of datasheets. Lastly, we can define further relevant categories, such as *Decision tree*, that our models would have to identify.

Albeit their impressive performance, there are areas where the models can be improved. The hyperparameters, model architectures, and augmentation techniques we utilized are mainly optimized for the COCO dataset. We could adapt those to our dataset that has very different characteristics compared to COCO. To adjust those parameters automatically, *hyperparameter evolution* could be used [UI]. However, that would require many more model training runs and considerably increase computing costs. We could also attempt to optimize the feature extraction process to more closely adapt it to our dataset of document page images (instead of the COCO data that was used for pretraining). Chen et al. [Ch21] achieve significant performance improvements through such adaptations. Finally, as object detection models continue to evolve, it would be interesting to update our research with the recent YOLOv8 family that was published after we finished our experiments [UI23].

5 Conclusion

With this study on extracting and classifying figures from electronics datasheets we show that recent object detection models from the computer vision domain can be successfully leveraged for specific tasks in document analysis and graphics recognition. We add to prior studies by considering newer generations of the YOLO model families, and compare against the also established EfficientDet. We highlight an iterative approach to outsourced data labeling, as well as the importance of using augmentations during model training. Our Experiments on our comprehensive custom dataset of page images from technical documents show impressive performance levels that are on par with SOTA results from the literature and related studies.

Acknowledgements

This research was funded by the German Federal Ministry of Education and Research (BMBF) under grants no. 01IS20091A, 01IS20091B (Project KIRESys). Computations for this work were done (in part) using resources of the Leipzig University Computing Center.

References

- [AF18] Auer, F.; Felderer, M.: Shifting Quality Assurance of Machine Learning Algorithms to Live Systems. In (Tichy, M.; Bodden, E.; Kuhrmann, M.; Wagner, S.; Steghöfer, J.-P., eds.): Software Engineering und Software Management 2018. Gesellschaft für Informatik, Bonn, pp. 211–212, 2018, URL: <https://dl.gi.de/handle/20.500.12116/21162>.
- [BBK98] Baum, L. S.; Boose, J. H.; Kelley, R. J.: Graphics recognition for a large-scale airplane information system. In: Graphics Recognition Algorithms and Systems. GREC 1997. Springer Berlin Heidelberg, pp. 291–301, 1998.
- [Bi22] Binder, F.; Diels, J.; Balling, J.; Albrecht, O.; Sachunsky, R.; Philipp, J. N.; Scheurer, Y.; Münsch, M.; Otto, M.; Niekler, A.; Heyer, G.; Thorun, C.: Putting Users in the Loop: How User Research Can Guide AI Development for a Consumer-Oriented Self-service Portal. In (Rauterberg, M., ed.): Culture and Computing. HCHI 2022. Vol. 13324. Lecture Notes in Computer Science, Springer, pp. 3–19, 2022.
- [BWL20] Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y. M.: Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934/, 2020.
- [Ch21] Chen, K.; Lee, C.; Lin, M. P.; Wang, Y.; Chen, Y.: Massive Figure Extraction and Classification in Electronic Component Datasheets for Accelerating PCB Design Preparation. In: 3rd ACM/IEEE Workshop on Machine Learning for CAD, MLCAD 2021. IEEE, pp. 1–6, 2021, URL: <https://doi.org/10.1109/MLCAD52597.2021.9531275>.

- [He20] He, K.; Gkioxari, G.; Dollár, P.; Girshick, R.: Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42/2, pp. 2961–2969, Feb. 2020.
- [HL14] Hu, J.; Liu, Y.: Analysis of Documents Born Digital. In: *Handbook of Document Image Processing and Recognition*. Springer London, pp. 775–804, 2014.
- [Jo22] Jocher, G.; Chaurasia, A.; Stoken, A.; Borovec, J.; NanoCode012; Kwon, Y.; TaoXie; Michael, K.; Fang, J.; imyhxy; Lorna; Wong, C.; Yifu, Z.; V, A.; Montes, D.; Wang, Z.; Fati, C.; Nadar, J.; Laughing; UnglvKitDe; tkianai; yxNONG; Skalski, P.; Hogan, A.; Strobel, M.; Jain, M.; Mammana, L.; xylieong: ultralytics/yolov5: v6.2 - YOLOv5 Classification Models, Apple M1, Reproducibility, ClearML and Deci.ai integrations, version v6.2, Aug. 2022, URL: <https://doi.org/10.5281/zenodo.7002879>, visited on: 11/05/2022.
- [Ka13] Karatzas, D.; Shafait, F.; Uchida, S.; Iwamura, M.; i Bigorda, L. G.; Mestre, S. R.; Mas, J.; Mota, D. F.; Almazan, J. A.; de las Heras, L. P.: ICDAR 2013 Robust Reading Competition. In: *2013 12th International Conference on Document Analysis and Recognition*. IEEE, Aug. 2013.
- [Ki] Kin-Yiu, W.: Official YOLOv7, URL: <https://github.com/WongKinYiu/yolov7>, visited on: 11/05/2022.
- [Li16] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A. C.: SSD: Single Shot MultiBox Detector. In: *Computer Vision – ECCV 2016*. Springer, pp. 21–37, 2016.
- [Li17a] Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S.: Feature pyramid networks for object detection. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*. IEEE, pp. 2117–2125, July 2017.
- [Li17b] Lin, T.; Goyal, P.; Girshick, R. B.; He, K.; Dollár, P.: Focal Loss for Dense Object Detection. In: *IEEE International Conference on Computer Vision, ICCV 2017*. IEEE Computer Society, pp. 2999–3007, 2017, URL: <https://doi.org/10.1109/ICCV.2017.324>.
- [Li22] Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W., et al.: YOLOv6: a single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*, 2022.
- [LO14] Lamiroy, B.; Ogier, J.-M.: Analysis and Interpretation of Graphical Documents. In: *Handbook of Document Image Processing and Recognition*. Springer London, pp. 553–590, 2014.
- [LR14] Lladós, J.; Rusiñol, M.: Graphics Recognition Techniques. In: *Handbook of Document Image Processing and Recognition*. Springer London, pp. 489–521, 2014.
- [Me] Meituan: YOLOv6, URL: <https://github.com/meituan/YOLOv6>, visited on: 11/05/2022.

- [Op22] Opasjumruskit, K.; Böning, S.; Schindler, S.; Peters, D.: OntoHuman: Ontology-Based Information Extraction Tools with Human-in-the-Loop Interaction. In (Luo, Y., ed.): Cooperative Design, Visualization, and Engineering - 19th International Conference, CDVE 2022. Vol. 13492. Lecture Notes in Computer Science, Springer, pp. 68–74, 2022.
- [OSP21] Opasjumruskit, K.; Schindler, S.; Peters, D.: Automatic Data Sheet Information Extraction for Supporting Model-Based Systems Engineering. In (Luo, Y., ed.): Cooperative Design, Visualization, and Engineering - 18th International Conference, CDVE 2021. Vol. 12983. Lecture Notes in Computer Science, Springer, pp. 97–102, 2021.
- [Pf22] Pfitzmann, B.; Auer, C.; Dolfi, M.; Nassar, A. S.; Staar, P.: DocLayNet: A Large Human-Annotated Dataset for Document-Layout Segmentation. In: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. KDD '22, Association for Computing Machinery, Washington DC, USA, pp. 3743–3751, Aug. 2022, ISBN: 9781450393850, URL: <https://doi.org/10.1145/3534678.3539043>.
- [Re13] Redmon, J.: Darknet: Open Source Neural Networks in C, 2013, URL: <http://pjreddie.com/darknet/>.
- [Re16] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A.: You only look once: Unified, real-time object detection. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016. IEEE Computer Society, pp. 779–788, 2016.
- [Re17] Ren, S.; He, K.; Girshick, R.; Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 39/6, pp. 1137–1149, June 2017.
- [RF17] Redmon, J.; Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Pp. 7263–7271, 2017.
- [RF18] Redmon, J.; Farhadi, A.: Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767/, 2018.
- [Sh21] Shen, Z.; Zhang, R.; Dell, M.; Lee, B. C. G.; Carlson, J.; Li, W.: LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis. In (Lladós, J.; Lopresti, D.; Uchida, S., eds.): Document Analysis and Recognition – ICDAR 2021. Springer International Publishing, Cham, pp. 131–146, 2021, ISBN: 978-3-030-86549-8.
- [SSD19] Sultana, F.; Sufian, A.; Dutta, P.: A Review of Object Detection Models based on Convolutional Neural Network. CoRR abs/1905.01614/, 2019, arXiv: 1905.01614.

- [TL19] Tan, M.; Le, Q.: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In (Chaudhuri, K.; Salakhutdinov, R., eds.): Proceedings of the 36th International Conference on Machine Learning. Vol. 97. Proceedings of Machine Learning Research, PMLR, pp. 6105–6114, June 2019.
- [TPL20] Tan, M.; Pang, R.; Le, Q. V.: EfficientDet: Scalable and efficient object detection. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020. IEEE, pp. 10778–10787, 2020.
- [U1] Ultralytics: Hyperparameter Evolution, URL: <https://docs.ultralytics.com/tutorials/hyperparameter-evolution/>, visited on: 11/25/2022.
- [UI21] Ultralytics: Release V6.0 - yolov5n 'nano' models, Roboflow integration, tensorflow export, opencv DNN support · ultralytics/yolov5, Oct. 2021, URL: <https://github.com/ultralytics/yolov5/releases/tag/v6.0>, visited on: 11/05/2022.
- [UI23] Ultralytics: ultralytics/ultralytics: NEW - YOLOv8 in PyTorch > ONNX > CoreML > TFLite, 2023, URL: <https://github.com/ultralytics/ultralytics>, visited on: 01/12/2023.
- [Wa20] Wang, C.-Y.; Liao, H.-Y. M.; Wu, Y.-H.; Chen, P.-Y.; Hsieh, J.-W.; Yeh, I.-H.: CSPNet: A new backbone that can enhance learning capability of CNN. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. Pp. 390–391, 2020.
- [WBL22] Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y. M.: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696/, 2022.
- [Wi] Wightman, R.: EfficientDet (PyTorch), URL: <https://github.com/rwightman/efficientdet-pytorch>, visited on: 11/05/2022.
- [Wu18] Wu, S.; Hsiao, L.; Cheng, X.; Hancock, B.; Rekatsinas, T.; Levis, P.; Ré, C.: Fondue: Knowledge Base Construction from Richly Formatted Data. In: Proceedings of the 2018 International Conference on Management of Data (SIGMOD 2018). SIGMOD '18, ACM, Houston, TX, USA, pp. 1301–1316, May 2018.
- [Ya17] Yang, C.; Yin, X.-C.; Yu, H.; Karatzas, D.; Cao, Y.: ICDAR2017 Robust Reading Challenge on Text Extraction from Biomedical Literature Figures (DeTEXT). In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE, Nov. 2017.
- [Zo23] Zou, Z.; Chen, K.; Shi, Z.; Guo, Y.; Ye, J.: Object Detection in 20 Years: A Survey. Proc. IEEE 111/3, pp. 257–276, Mar. 2023.
- [ZTY19] Zhong, X.; Tang, J.; Yepes, A. J.: Publaynet: largest dataset ever for document layout analysis. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE, pp. 1015–1022, 2019.

Wirtschaft, Management Industrie -
(Agiles) Enterprise Architecture
Management in Forschung und
Praxis

Visual analytics for knowledge management

Advantages for organizations and inter-organizational teams

Gergana Vladova¹, André Ullrich², Eldar Sultanow³, Marinho Tobolla⁴, Sebastian Sebrak⁵, Christian Czarnecki⁶, and Carsten Brockmann⁷

Abstract: The management of knowledge in organizations considers both established long-term processes and cooperation in agile project teams. Since knowledge can be both tacit and explicit, its transfer from the individual to the organizational knowledge base poses a challenge in organizations. This challenge increases when the fluctuation of knowledge carriers is exceptionally high. Especially in large projects in which external consultants are involved, there is a risk that critical, company-relevant knowledge generated in the project will leave the company with the external knowledge carrier and thus be lost. In this paper, we show the advantages of an early warning system for knowledge management to avoid this loss. In particular, the potential of visual analytics in the context of knowledge management systems is presented and discussed. We present a project for the development of a business-critical software system and discuss the first implementations and results.

Keywords: knowledge management, visual analytics, knowledge transfer, teamwork, knowledge management system, tacit knowledge, explicit knowledge

1 Introduction

In organizations, developing both individual and collective knowledge is one of the significant knowledge management goals. This is related to the challenge of establishing an appropriate framework to support knowledge transfer between single knowledge bearers and the organizational knowledge base. For this purpose, formal knowledge management rules are usually applied [He12] that help shape a framework for transferring explicit and explicable knowledge as data and information. In new or non-routine situations, however, rules and activity-specific guidelines have shortcomings and to some degree lack applicability. In order to avoid this, appropriate knowledge management

¹ University of Potsdam, Chair of Business Informatics, Processes and Systems, Karl-Marx-Str. 67, 14482 Potsdam, Germany, gergana.vladova@wi-uni-potsdam.de / Weizenbaum Institute for the Networked Society, Research Group Education for the Digital World, Hardenbergstr. 32, 10623 Berlin, Germany

² Weizenbaum Institute for the Networked Society, Research Group Digitalization, Sustainability and Participation, Hardenbergstr. 32, 10623 Berlin, Germany, andre.ullrich@weizenbaum-institut.de

³ Capgemini, Bahnhofstraße 30, 90402 Nürnberg, Germany, eldar.sultanow@capgemini.com

⁴ Bundesagentur für Arbeit, IT-Systemhaus, Südwestpark 26, 90449 Nürnberg, Germany, marinho.tobolla@arbeitsagentur.de

⁵ Capgemini, Bahnhofstraße 30, 90402 Nürnberg, Germany, sebastian.sebrak@capgemini.com

⁶ FH Aachen, Eupener Straße 70, 52066 Aachen, czarnecki@fh-aachen.de

⁷ SAP Deutschland SE & Co. KG, Rosenthaler Str. 30, 10178 Berlin, carsten.brockmann@sap.com

should address not only compliance with defined rules but also organizational culture and structure, awareness of processes and interdependences ([So94]; [MTK15]; [RKN14]), as well as leadership [LTK21] communication and interaction activities [Bh02].

Furthermore, the nature of tacit knowledge and its dependence on the knowledge bearer are essential criteria for the success of its transfer ([Po58]; [Ma00]). This is important since organizational knowledge creation also depends on personal knowledge and thus, in particular, on tacit knowledge ([NT96]; [RKN14]). Furthermore, organizational culture influences the knowledge-related behaviors of individuals, teams, organizational units, and organizations, setting values and norms about knowledge sharing and transfer practices [MTK15].

The literature has shown the importance of knowledge management for project work and project success [MS21]. In large organizational projects involving external actors, designing and maintaining formal and informal knowledge management and transfer rules, common culture, and interaction activities becomes a significant challenge. Over the long term, the work of external team members is of organizational importance, however, they are not fully integrated as organizational members. Furthermore, staff turnover more often affects external employees than internal employees. Implementing an appropriate framework for team knowledge management (e.g. [ES00]) always causes restrictions.

Regarding the management of explicit knowledge, the analysis of massive and complex data and the extraction of relevant information are becoming increasingly important. However, the capabilities and possibilities of extracting useful information from the collected data are relatively limited compared to developing technologies for data collection, storage, and management (e.g. [TC06]). Visual analytics (VA) can overcome information overload and turn it into an opportunity by offering tools and techniques to facilitate human judgment in the process of knowledge discovery in databases. So far, software visualization approaches have rarely been used in knowledge management systems (KMS) and only in isolation with a narrow focus on the source code. However, the combination of software visualization and KMS has great potential, such as providing context-specific relevant knowledge to employees with expert knowledge.

Against this background, this paper discusses on the example of one concrete project the challenges that organizational knowledge management meets within knowledge-intensive projects and shows how visual analytics can be used to overcome these challenges. The main question of the paper is therefore: Which benefits for the management and transfer of project-relevant knowledge can be delivered using visual analytic methods?

The paper extends existing methodological concepts for the use of VA in project and knowledge management and describes the advantages and challenges of a concrete application, which can generally support teamwork and the achievement of central knowledge management goals in different use cases.

2 Knowledge management and visual analytics

In the following, the nature of knowledge and the differences and interdependencies between tacit and explicit and collective and individual knowledge will be elaborated. Relevant in this context is the recent development of knowledge management systems (KMS). These systems enable the generation and appropriate analysis of various data and information in organizations. However, their analysis is often context- and individual-specific. Furthermore, it requires meta-information and meta-knowledge. Visual analytics, a field under dynamic development with high potential in analyzing and accurately representing data and information in organizations will be also shortly introduced.

2.1 Knowledge and knowledge management

Knowledge can be described as the dynamic human process of justifying personal beliefs about truth [NT96, p.58]. Organizational knowledge creation is a dialectical process in which various contradictions are synthesized through dynamic interactions among individuals, an organization, and the environment [NT15]. Although there is no clear definition of knowledge, considering different knowledge types [AL01] is crucial in describing and managing it. In this paper, in particular, tacit vs. explicit and individual vs. collective knowledge are discussed. Tacit knowledge [Po58], which is "personal, context-specific, and very difficult to communicate" [NT96] versus explicit knowledge [Po58], which can be distributed in formal and systematic language. Individual knowledge is "possessed" by a single entity whereby collective knowledge is transferred and shared between entities [No94] – in organizations, this can be described also as organizational knowledge [He12]. Knowledge creation can be seen as a social process in an organizational context. Personal knowledge and collective knowledge are discoverable, identifiable, and connected. [Bh02] defines interactions as the main components of organizational knowledge. If the levels of interaction and possibilities for exchange are reduced to a minimum, most knowledge remains under the control of individuals [RKN14].

2.2 Knowledge management systems

Given the complexity and variety of knowledge management tasks and challenges in organizations and the massive amounts of data and information created, transferred, and stored, information systems are applied to manage organizational knowledge. These are IT systems for enhancing the creation, exchange, retrieval, storage, usage, retainment, or application of knowledge, information, and data via platforms. Such platforms are usually referred to as knowledge management systems. KMS are "a class of information systems applied to managing organizational knowledge" [AL01], thus, enabling technologies for effective and efficient knowledge management [MH05]. The essential system functions include content management, information retrieval, visualization and aggregation of

knowledge, and collaboration. The success of KMS is related to the knowledge quality, the service quality perceptions, the intention to use, the user satisfaction, and the system quality influences the success (Halawi et al., 2008). An open organizational culture is crucial for KMS success [AL99]. In the context of digitalization, enterprise KMS are increasingly cloud-based. The benefits include information search and retrieval in a time and location-independent manner and knowledge sharing and reuse in distributed environments. Cloud-based KMS also enable the handling of big data and the application of analyses [Li15]. [TCS11] propose a new service layer called Knowledge as a Service (KaaS), which facilitates personal knowledge management, information retrieval, evaluation, and knowledge organization. The cloud offers infrastructure services (storage, communication, etc.), knowledge services (knowledge creation, sharing, reuse, etc.), and platform services (databases etc.) [Su10]. It enables knowledge workers to integrate external content and build individual knowledge-facilitating environments.

2.3 Visual analytics for knowledge management

Organizational project work often includes many knowledge tasks and process-related data and information. Their interpretation and application are crucial to the project's success. Therefore, accurate data representation is essential. One way to maintain an overview of team performance and knowledge development and meet this challenge is the application of visual analytics. Building on the human mind's ability to understand complex information received through visual channels and facilitate the analytical thinking process [CT05], visual analytics can explore vast amounts of data and find interesting patterns [GJ15].

Visual analytics is “the science of analytical reasoning facilitated by interactive visual interfaces” [CT05]. It can be further described as an “iterative process that involves information gathering, data preprocessing, knowledge representation, interaction, and decision-making” [Ke08]. The method combines the strengths of machines with those of humans. It fosters the constructive evaluation, correction, and rapid improvement of processes and models, thus improving the knowledge base and decision-making [Ke08]. This is achieved through an “effective delegation of perceptive skills, cognitive reasoning, and domain knowledge on the human side and computing and data storage capability on the machine side, and their effective coupling via visual representations” [Sa14].

Visual analytics has already been investigated in various studies with knowledge management focus and its potential has been demonstrated in different contexts. [Co16] examined the use of visual analytics in personnel management in connection with the development and maintenance process of software systems. In particular, the study focused on the benefits for project leaders and managers responsible for decision-making, including the area of human resources. The authors point out the harmful effects of staff rotation and distributed software development as barriers to building long-term relationships between project managers and software engineering projects and between managers and companies. As a solution, visualization and interaction techniques can enable project managers to determine the contribution of each programmer and identify

both the software elements modified jointly by a group of programmers and the people involved. This supports decisions for software project leaders and managers regarding assigning tasks to developers and substituting staff due to unexpected situations or staff turnover. [GJ15] explore and develop practical ways to integrate visual analytics with a computational model for medical decision-making. Based on three case studies, the paper shows ways to improve the conceptual phase of the design process by using visual analytics. [Em18]) show a concept and first implementation results for visual analytics in product development by transferring linked maintenance data and knowledge to visual features of a product design representation and making it easier for users to use understand product concepts. [Hr01] show the advantages of applying visual analytics for knowledge discovery in the medicine domain.

3 Case study description

The paper discusses the role of VA for knowledge management in the context of one project of the German Federal Employment Agency (FEA). The FEA employs more than 96,000 people and operates one of the country's most extensive information technology (IT) infrastructures, with more than 160,000 networked computer workstations, 9000 servers, two central data centers, and several regional data centers. E-mail, bank transfers, mail, and print products require processing a vast amount of data by the FEA.

The technical solution managing this complex IT infrastructure is the mission-critical software system ALLEGRO. The software was developed over the last seven years and remained one of the essential FEA projects. ALLEGRO is used to retrieve, administrate, and process data, calculate unemployment benefits and periods, process payment orders for benefits, report and pay into social insurance funds, prepare decisions, and central and decentralized document printing.

The system comprises more than 718,000 lines of code, and the development team consists of about 90 developers. As with all large private or public projects like this, external experts and consultants' power, skills, and knowledge are involved. Unfortunately, highly specialized knowledge developed by various external experts during the project is often not sufficiently shared within the agency. The teamwork depends on substantial tacit knowledge, which relies on experience and concrete individuals. Due to the high employee turnover, various knowledge management problems arose: peaks and valleys of knowledge flows and objects (artifacts, documents, source code, etc.), project-relevant knowledge loss due to changes in external consultants, and shortcomings in the identification of relevant knowledge and knowledge carriers, thus increasing knowledge gaps. In addition, the fluctuation makes it more challenging to establish a consistent knowledge management culture. This type of project work, however, at least makes it possible to use the software developer's digital footprints to identify work areas and concrete tasks. This is important to knowledge management for two reasons: identification

of knowledge bearers (and thus tacit knowledge) and identification of explicit knowledge already created.

The FEA knowledge management team recorded the current and existing problems and needs of the ALLEGRO development team based on interviews with project members, participating observations, analysis of information and knowledge flows, and development of concrete strategic goals. Investigative aspects included:

- relevant areas of knowledge of the team members,
- responsible persons and their representatives,
- breadth and depth of knowledge required in the context of different tasks,
- existing knowledge bottlenecks, gaps, and hotspots,
- possible future scenarios,
- overview of the current actual knowledge situation.

As a result, the main knowledge management challenges for the team of internal and external developers within the knowledge-intensive project ALLEGRO have been identified, especially an unintentional loss of knowledge and high significance and dependence on external (developer and consultant) knowledge.

An essential result of this internal project was the development and establishment of an Early Warning Knowledge Management System (EWKMS), which, among other things, demonstrates the potential of visual analytics in the context of knowledge-intensive projects. The aim of the implementation and further development of the EWKMS is to improve knowledge management, taking into account tacit and explicit and internal and external knowledge. Thus, the EWKMS highlights knowledge positions and critical knowledge resources by visualizing the development, atrophy/degeneration, and endangerment of knowledge in business-critical projects. It provides a well-structured approach to identifying knowledge management challenges and developing the best possible strategies and concepts.

As a result, the leading knowledge management challenges for the team of internal and external developers within the knowledge-intensive project ALLEGRO have been identified, especially an involuntary loss of knowledge and high significance and dependence on external (developer and consultant) knowledge. The organizational benefits include automatically using existing data from different sources (for example, Git, CVS, Confluence, and Innovator) and exploiting the hidden potential in existing team-relevant data. In addition, the knowledge management team becomes aware of relevant data protection guidelines. For example, anonymization and pseudonymization functions were implemented. One of the significant advantages is recognizing and visualizing knowledge-intensive areas (hotspots) that are particularly important to teamwork; the new knowledge management system algorithm makes this possible.

4 Visual analytics for the Early Warning Knowledge Management System

The EWKMS project combines human and machine skills to demonstrate the possibilities for resolving the volatility of critical external knowledge. The scope of visual analytics is shown in Fig. 1, with the associated applied functionalities and benefits for the ALLEGRO knowledge management (KM) team listed on the right.

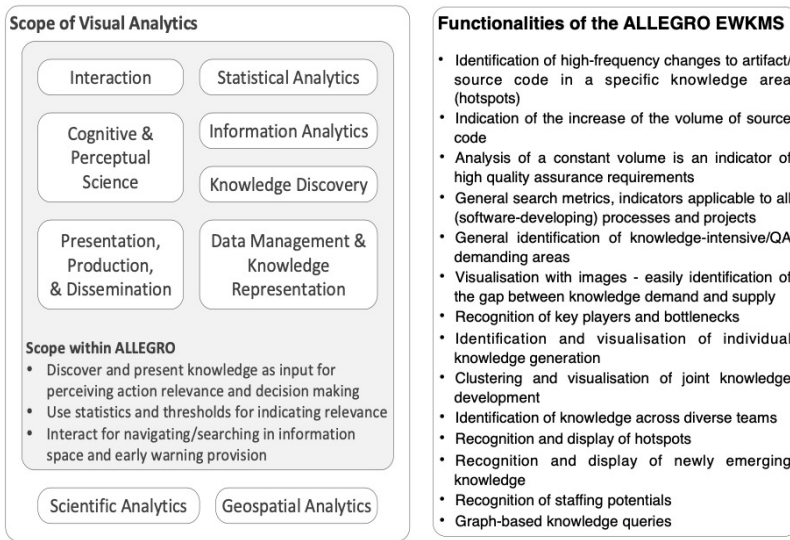


Fig. 1: Scope of visual analytics [Ke08] (on the left) applied to the ALLEGRO EWKMS (on the right)

Code statistics, commits, and reviews are the basis for the statistical analysis. Interaction, cognitive science, and perceptual science play a role in the use of the graphical interface of the KMS, which visualizes knowledge areas and hotspots based on large amounts of data. Information analysis is used to determine which artifacts (code fragments, documents, summary articles, etc.) belong to which field of knowledge. The pilot system issues early warnings as part of the knowledge-finding process, whereby presentation, production, dissemination, and data management are also part of the visual analytics.

The ALLEGRO EWKMS enables the analysis of project-related data to methodically determine the knowledge situation in the context of a specific project (including bottlenecks, undivided knowledge, etc.). It refers to knowledge management and transfer issues and challenges in the project - e.g. the permanent creation of countless data and codes from different developers in distributed teams; unpredictable staff turnover, making it difficult to draw necessary conclusions about personal (developer) knowledge and experience; jointly created knowledge as result of teamwork and joint tasks. As a result, explicit knowledge is linked with tacit knowledge during the entire project work.

The analysis aims at a general knowledge transfer and management level. It is generated by the tool depends only on the amount of data (lines of code, Git repository, etc.), which can be correspondingly large. Adverse side effects, for example, possible inference from the person-related service, are explicitly excluded.

The architecture of the pilot system corresponds to a cloud-based architecture that can process large amounts of data and includes four main layers: a web-based front end, a service, a data, and a back-end process (compare [Su17]).

4.1 Benefits of visual Analytics for the knowledge-intensive teamwork

Visual analytics plays a central role in the EWKMS. In the following, the most important cases for ALLEGRO covered by the KMS pilot are listed and described, including application examples for each benefit, and the advantages for organizational and team-oriented knowledge management are discussed.

Identification and visualization of jointly developed knowledge in clustered knowledge areas, as well as of undivided knowledge

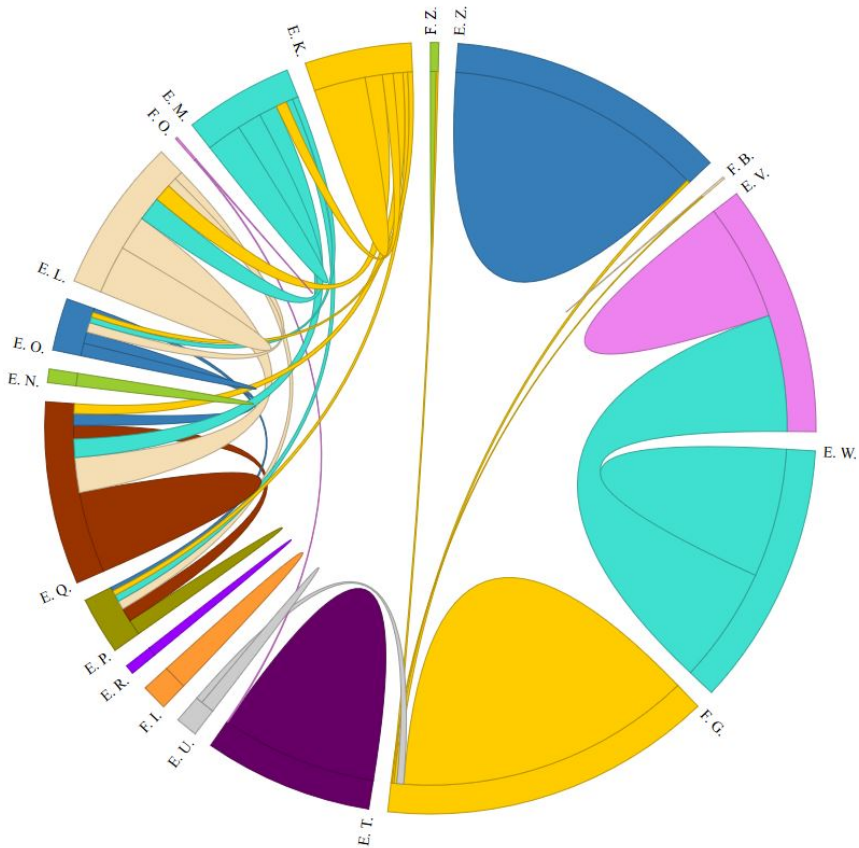


Fig. 2: Identification and visualization of jointly developed knowledge in clustered knowledge areas, as well as of undivided knowledge

The outer edge of the circle in Fig. 2 represents the orientation axis with the developers (responsible for checking in the source code). If a line runs from one end of the circle to the other, it is the set of source code (that two developers developed or integrated). In other words, there are commits to this source code from both developers. The developers (anonymized) are located on the edge/circle axis.

This refers to the visualization of jointly developed collective and relational knowledge (over three years). The D3 creates a chord diagram that illustrates which developers have

transferred source code to the ALLEGRO source repository and other developers/colleagues. The outer edge of the circle represents knowledge carriers; in this case, committers of the ALLEGRO source. A diagonal line between two or more committers means that they (effectively working together) have worked in the same place and have edited assembled code fragments (classes, methods, etc.).

As an additional benefit, this visualization can identify developers who are solely involved in specific areas as carriers of undivided tacit knowledge.

Identification of hotspot areas of knowledge

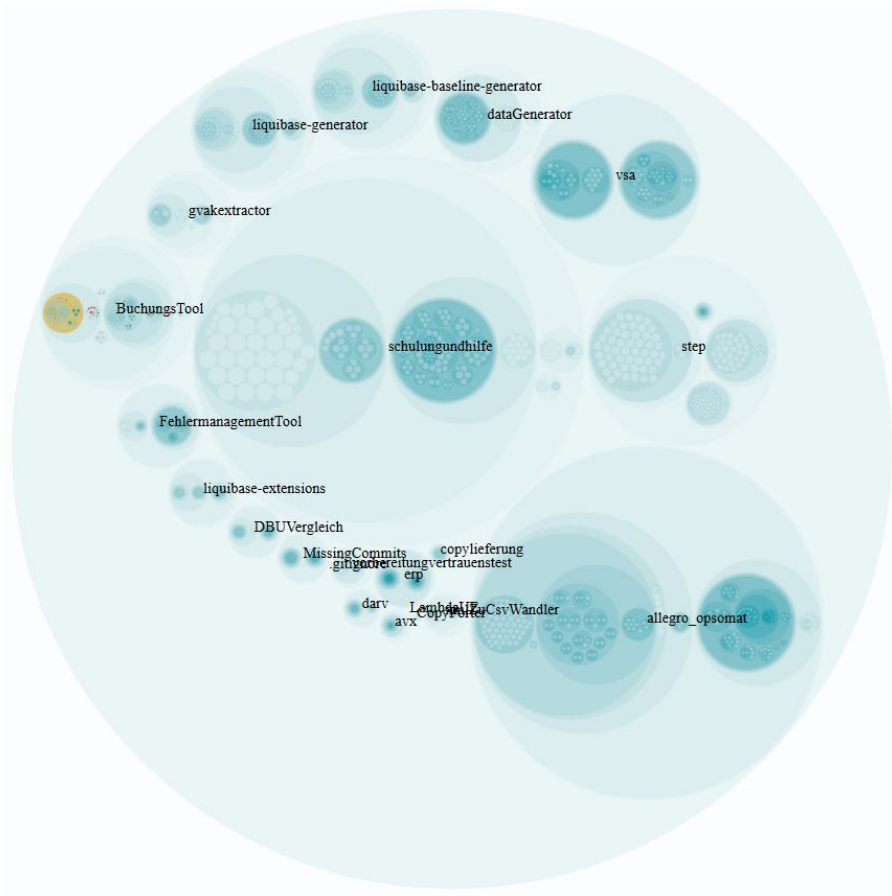


Fig. 3: Identification of hotspot areas of knowledge

The bubbles in Fig. 3 represent source code sets (the darker, the more commits; the bubble/circle size indicates the number of code lines).

The D3 was used to create a bubble chart that includes a circle packing algorithm to illustrate hotspots. The diameter represents the volume of the lines of code in a field of knowledge. The darker the color, the more commits there are in the area. It is possible to show how the focus shifts to areas by not displaying the timeline additively; each year can be analyzed separately by ignoring commits from the previous year. Another hotspot metric would be quoting the total number of commits and the number of commits in a range.

This view is a starting point for investigating and visualizing causal knowledge. ALLEGRO is complex in its business logic. In some areas, code is often changed to do not increase the volume of code. Such areas of knowledge indicate high quality requirements. This case is intended to ensure that this knowledge is passed on. It is more problematic if an area is identified as undivided knowledge (for example, in the second form of use described above).

Identification of experts, key players, and knowledge and information bottlenecks

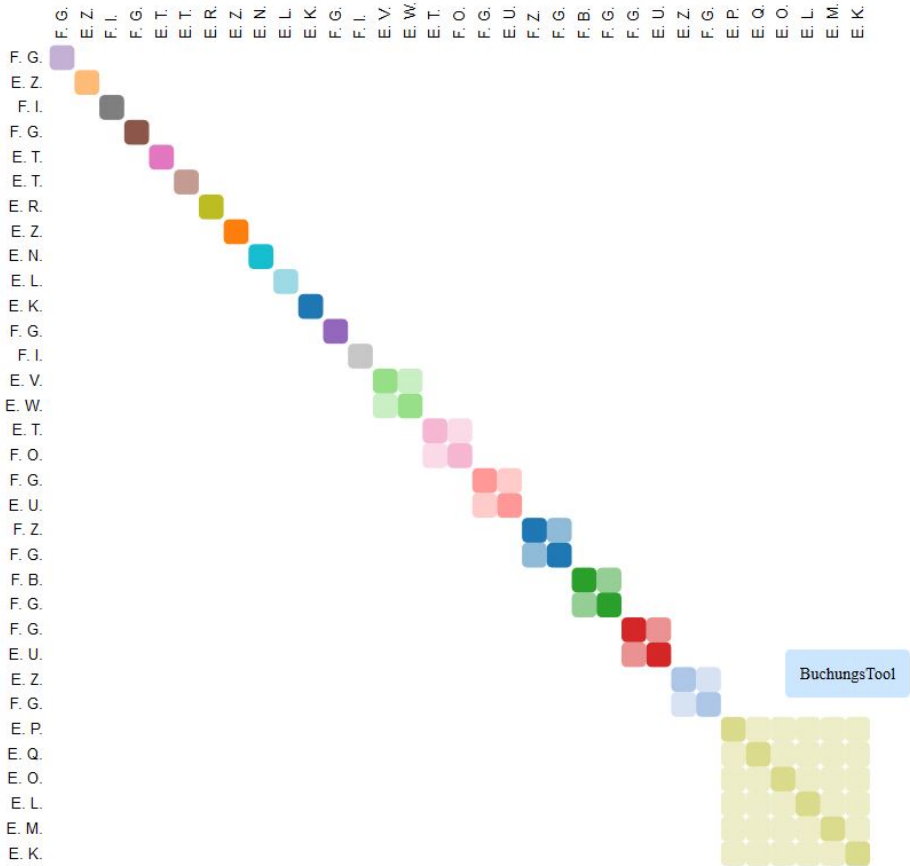


Fig. 4: Identification of experts, key players, and knowledge and information bottlenecks

The x- and the y-axes in Fig. 4 represent two different developers (square matrix). The matrix entries are the fields of knowledge in which the respective developers are active. Another view of this analysis shows the degrees to which individuals cover specific knowledge areas.

This visualization includes an analysis of crawled documents, the commit history, an overview of key individuals who are sources of specific, explicit knowledge in different

areas, and owners of tacit knowledge. The prerequisite is that knowledge owners have specified their relevant knowledge areas in databases or available documents.

Identification and visualization of the delta between knowledge need and supply (automatically)

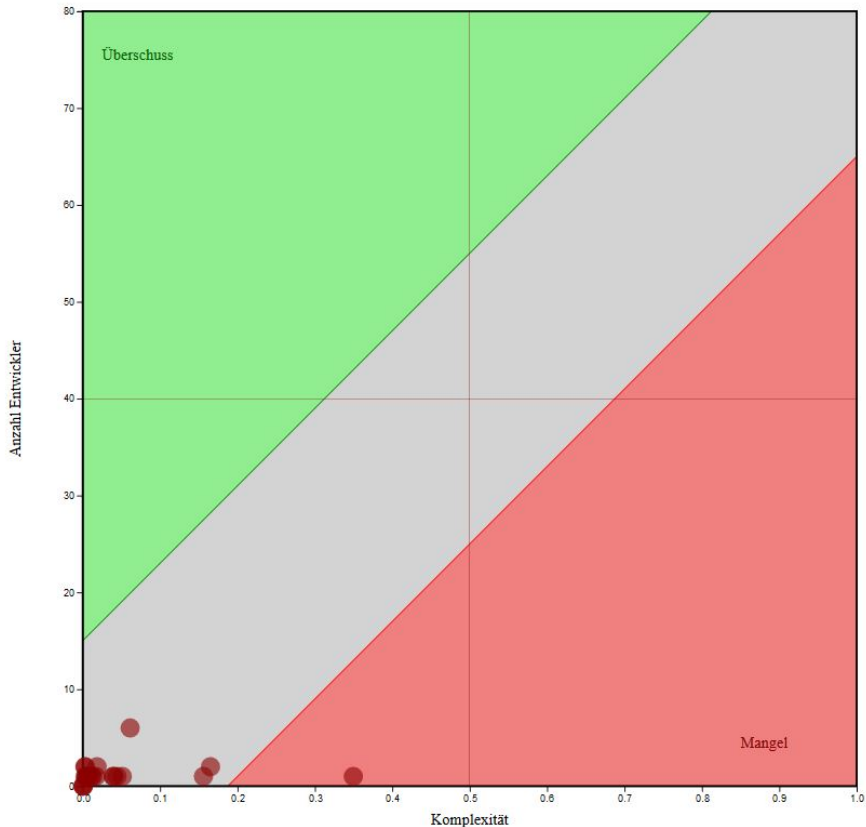


Fig. 5: Identification and visualization of the delta between knowledge need and supply (automatically)

The x-axis in Fig. 5 represents the complexity of a knowledge area (dependencies, frequency of changes, scope, etc.). The y-axis represents the number of developers working in this area. The basis for this visualization and analysis step is the comparison of search queries (sorted by frequency) in the company-wide Wiki, Confluence and Knowledge Portal with the delivered search results. This provides relevant indications of a delta

between knowledge demand and knowledge supply (between knowledge demanders and providers).

Identification of recently created knowledge

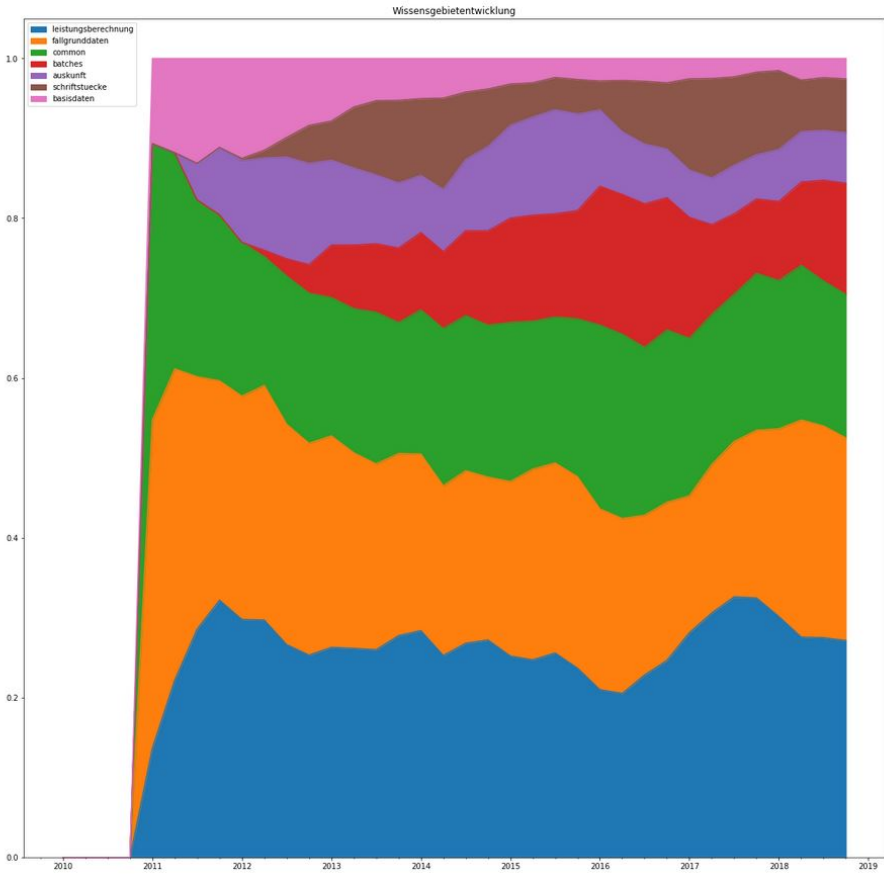


Fig. 6: Identification of recently created knowledge

The x-axis in Fig. 6 represents the development time (in yearly intervals). The y-axis represents the size of a knowledge area (described as lines of code). If, for example, a knowledge area grows over the years, then a horizontal color bar becomes wider from left to right.

Through the analysis of content creation and its history in the company-wide Wiki, Confluence and Knowledge Portal, the pilot project provides a comprehensive picture of knowledge development in existing areas and new knowledge areas relevant to the team. Such knowledge discovery is not limited to content analysis alone. It also includes the analysis of models created by the requirements engineering team with the Innovator modeling tool of the FEA. Therefore, this function deals with the dynamics of the knowledge-building process.

4.2 Initial application and lessons learned

The pilot system's functionality was extensively tested and evaluated by various members of the ALLEGRO team, with employees in different roles and positions serving as test participants. The tests began with a brief introduction to the use of the system and its functionalities. The testers were assigned various tasks covering the use cases described above. Their tasks included identifying knowledge islands, staffing a team for a potential new project, and identifying knowledge gaps within an ongoing project. The system's performance was assessed by comparing the results of the testers and interviewing the testers about their impressions. The most important results of the interviews are summarized as follows:

On the *strategic level*, the visualization and analyze method brings advantages for different decision-makers within the organization:

Knowledge Managers use visual analytics to obtain a holistic and atomic overview of the organizational knowledge situation. They can encourage both project management and developers to take appropriate measures, for example, to pass on their knowledge or to provide capacities for countermeasures against loss of knowledge. Furthermore, "this kind of identification of knowledge bottlenecks supports the short-, medium-, and long-term planning and development of domain experts and strategic knowledge areas" (quote from one respondent). Knowledge managers also found the efficient identification of newly created knowledge to be an advantage considering the significant effort previously required to collect and systematize new knowledge and its carriers. One respondent added that "the identification of similar knowledge across disciplines enables the implementation of interdisciplinary expert groups." This in turn can lead to synergistic effects between disciplines and departments. Finally, the test group stressed the pilot's potential to transform individual knowledge into collective knowledge. Applying the method, knowledge managers comprehend development and stagnation knowledge areas and monitor the coverage of knowledge areas (e.g. tacit and explicit knowledge, employees working in the area, missing activities or lack of manpower). Project managers use the analytical results as a decision-support tool for creating knowledge acquisition and transfer budgets and investing in knowledge outsourcing, exchange, or renewal. The pilot project also enables them to efficiently identify key knowledge owners without contacting the Domain Knowledge Manager. Project leaders regarded the pilot project as "a useful way of identifying knowledge gaps in teams and staffing, as technical experts are easy to

identify” (quote from one respondent). Overall, the pilot project helps to counteract a lack of causal and relational knowledge in project teams by integrating relevant roles into project teams. Through visualization, competency management receives additional support – responsible managers can easier identify critical areas with competencies shortage and address them on time. A benefit from the point of view of human resources responsible manager is the establishment of training and qualification measures for relevant employees and the acquisition of necessary knowledge. In case this knowledge can be found within the organization or the project team (including external developers), it can be organized across the organization. Team leaders get an overview of how and which knowledge areas are affected by the loss of personnel from the own team by monitoring well and insufficiently covered by personnel. Chief architects can automatically anticipate the need for new knowledge (about technologies, etc.) and act immediately and appropriately in case of upcoming technological challenges.

On an *operative level*, the software developers of the ALLEGRO team can easily recognize what kind of knowledge is required of them and prepare through training or qualification in areas of knowledge that are or will become relevant to them. Most developers appreciated the architecture of the pilot, which allows quick and relatively easy integration of new functionalities into the system. It was often commented that “now, we know which technology know-how (for example, rich client or server) they need from us for the development of the next release.” Perceived benefits of members the developer's group are e.g.: quick overview of the knowledge areas and their sizes, avoidance of additional, operative, time-consuming tasks, e.g., manual data entry and maintenance required, and the possibility of searching for and identifying organizational members who can provide valuable assistance in specific areas.

However, in the context of the first evaluation, there have been some critical statements. The most significant doubts about the project are stated by the human resources managers and high-level managers. The main issues are employee data privacy as well as employee performance. The contradiction is that the system must provide an anonymous mode, but the knowledge manager can only use the tool in a non-anonymous mode when required. Furthermore, the analysis results can be used to monitor employee performance.

5 Conclusion

Within the framework of this paper, we discussed the established and structured knowledge management framework of the FEA as the decisive factor in the success of the ALLEGRO project. Although the development of an early warning system with visual analytics was initiated from the bottom up to respond to concrete challenges within the knowledge-intensive ALLEGRO project, the basis for the immediate support and establishment of appropriate activities to address the challenges are provided by the existing organizational knowledge management strategy. This is where substantial knowledge management difficulties for organizations are encountered: On the one hand,

relevant relational and causal knowledge leaves the organization along with its external bearers; on the other hand, explicit knowledge (in the form of data, artifacts, documents, source code, etc.) created by external knowledge bearers during their employment may remain abstract and unknown, the meta-information and knowledge not transferred to or internalized by other employees. As a result, knowledge developed could be forgotten if left unused for long periods. On an organizational level, that could lead to substantial future effort and investment to achieve the same outcome.

The application of the pilot system addresses all relevant taxonomic aspects of knowledge – tacit, explicit, individual, and collective. The system supports handling knowledge objects (such as documents and data) and knowledge topics (collaboration and communication, learning processes, and competence development). Given the peculiarities of process knowledge implicitly anchored in individuals, internal and external experts involved in the project were the essential sources for identifying critical success factors of processes.

In addition to the benefits, which can be directly correlated with concrete visual analytics, further knowledge management benefits have been identified, including the following:

- Improvement in the staffing process: By visualizing and highlighting insufficiently covered areas of knowledge, new employees can be recruited, trained in these areas, and assigned to specific tasks.
- Graph-based knowledge queries: As described, a diagram database stores the current knowledge available in ALLEGRO. This enables the formulation of knowledge queries with specific knowledge objects linked to each other. For example, it is possible to query knowledge carriers and the connected persons with whom they share knowledge.
- Identification of relationships between knowledge across an area with scattered disciplines: This is achieved by searching the metadata of a shared file frequently used by the requirements engineering team, test team, and design and implementation team, including matching that metadata with commit metadata. The result makes it possible to predict whether the design and implementation team will encounter a required knowledge bottleneck when developing the next ALLEGRO release. The pilot project extracts various knowledge areas from requirements engineering artifacts, thus several taxonomic aspects of knowledge, and concentrates on relational knowledge. The requirements engineering team is currently working on the next release and compares these knowledge areas with the knowledge available in the design and implementation team.

Despite the pilot project's success, acceptance, and positive application in analytical knowledge management, there are other requirements to meet. The pilot in the current phase was developed and deployed exclusively by the design and development team. The expansion of the entire ALLEGRO team is pending. Although the pilot searches various sources (Git, Gerrit, Innovator, and Confluence), relationship matching and identification can be optimized. Moreover, the pilot cannot detect whether knowledge remains tacit in a

person. It also does not capture knowledge passed on face to face. However, we point out that developers may have much undivided tacit knowledge if their code ownership is high in a particular field of knowledge.

The pilot project provides valuable insights into the organizational situation regarding existing, emerging, fluctuating, and required knowledge. However, it has yet to reach its full potential. The system will be available as a mobile app for use on a device in the future. A Knowledge-as-a-Service platform will be provided that integrates machine learning and in-memory technologies to gain better insight into knowledge management at the FEA.

On the example of this factual knowledge-intensive project, we demonstrate the potential of visual analytics for knowledge management in combination with implementing and using knowledge management systems and extend the existing research on this topic.

Bibliography

- [AL99] Alavi, M.; Leidner, D.: Knowledge management systems: Emerging views and practices from the field. Proc. of the 32nd Annual Hawaii International Conference on Systems Sciences, HICSS-32, 1999.
- [AL01] Alavi, M.; Leidner, D.: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, pp. 107-136, 2001.
- [Bh02] Bhatt, G. D.: Management strategies for individual knowledge and organizational knowledge. *Journal of Knowledge Management* 6/1, pp. 31–39, 2002.
- [CT05] Cook, K. A.; Thomas, J. J.: Illuminating the path: The research and development agenda for visual analytics (No. PNNL-SA-45230). Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2005.
- [Em18] Emmanouilidis, C. et.al.: Internet of Things-Enabled Visual Analytics for Linked Maintenance and Product Lifecycle Management, *IFAC-PapersOnLine*, 51/11, pp. 435-440, 2018.
- [ES00] Eppler, M. J.; Sukowski, O.: Managing team knowledge: core processes, tools and enabling factors. *European Management Journal*, 18/3, p. 334-341, 2000.
- [GJ15] Giabbanelli, P. J.; Jackson, P. J.: Using visual analytics to support the integration of expert knowledge in the design of medical models and simulations. *Procedia Computer Science*, 51, p. 755-764, 2015.
- [Co16] González-Torres, A. et.al.: Knowledge discovery in software teams by means of evolutionary visual software analytics. *Science of Computer Programming*, 121, pp.55-74, 2016.
- [He12] Hecker, A.: Knowledge beyond the individual? Making sense of a notion of collective knowledge in organization theory. *Organization Studies*, 33/3, pp. 423-445, 2012.

- [Hr01] Hristovski, D. et.al.: Supporting discovery in medicine by association rule mining in Medline and UMLS. *Studies in health technology and informatics*, 2, pp. 1344-1348, 2001.
- [Ke08] Keim, D. et.al.: Visual analytics: Definition, process, and challenges. *Information visualization*. Springer, Berlin, Heidelberg, pp. 154-175, 2008.
- [Li15] Li, J. et.al.: Big data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, 81/1-4, pp. 667-684, 2015.
- [LTK21] Liu, G.; Tsui, E.; Kianto, A.: Revealing deeper relationships between knowledge management leadership and organisational performance: a meta-analytic study. *Knowledge Management Research & Practice*, 1-15, 2021.
- [MH05] Maier, R.; Hadrich, T.: Knowledge Management Systems. In (Schwartz, D. Ed.): *Encyclopedia of knowledge management*. IGI Global, 2015.
- [Ma00] Martin, J.: Knowledge management. The basis of corporate and institutional knowledge management., www.spottedcowpress.ca/KnowledgeManagement/pdfs/06MartinJ.pdf, 2000.
- [MTK15] Mojibi, T.; Hosseinzadeh, S.; Khojasteh, Y.: Organizational culture and its relationship with knowledge management strategy: a case study. *Knowledge Management Research & Practice*, 13/3, pp. 281-288, 2015.
- [MS21] Moutinho, J. D. A.; Silva, L. F. D.: Knowledge management in project management: mapping bibliographic convergence. *Knowledge Management Research & Practice*, 1-13, 2021.
- [No94] Nonaka, I.: A dynamic theory of organizational knowledge creation. *Organization Science*, 5/1, pp. 14-37, 1994.
- [NT96] Nonaka, I.; Takeuchi, H.: The knowledge-creating company: How Japanese companies create the dynamics of innovation. *Long-range planning*, 4/29, pp. 592, 1996.
- [NT15] Nonaka, I.; Toyama, R.: The knowledge-creating theory revisited: knowledge creation as a synthesizing process. In *The essentials of knowledge management*. Palgrave Macmillan, London, pp. 95-110, 2015.
- [Po58] Polanyi, M.: *Personal Knowledge. Towards a post-critical philosophy*. NY: Harper Torchbooks, 1958.
- [RKN14] Razmerita, L.; Kirchner, K.; Nabeth, T.: Social media in organizations: leveraging personal and collective knowledge processes. *Journal of Organizational Computing and Electronic Commerce*, 24/1, pp. 74-93, 2014.
- [Sa14] Sacha, D. et.al.: Knowledge generation model for visual analytics. *IEEE transactions on visualization and computer graphics*, 20/12, pp. 1604-1613, 2014.
- [So94] Spender, J. C.: Organizational knowledge, collective practice and Penrose rents. *International business review*, 3/4, pp. 353-367, 1994.
- [Su10] Sultan, N.: Cloud computing for education: A new dawn?. *International Journal of Information Management*, 30/2, pp. 109-116, 2010.

- [Su17] Sultanow, E. et.al.: Visual Analytics Supporting Knowledge Management: A Case Study of Germany's Federal Employment Agency. In Proceedings of i-KNOW- Data-Driven Future Conference, 11-12 October 2017, Graz, Austria. http://ceur-ws.org/Vol-2025/paper_hci_1.pdf, 2017
- [TC06] Thomas, J. J.; Cook, K. A.: A visual analytics agenda. IEEE computer graphics and applications, 1, pp. 10-13, 2006.
- [TCS11] Tsui, E.; Cheong, R. K.; Sabetzadeh, F.: Cloud-based personal knowledge management as a service (PKMaaS). In 2011 International Conference on Computer Science and Service System (CSSS), IEEE, pp. 2152-2155, 2011.

Process view on interface development of ERP and legacy systems

Dr. Carsten Brockmann¹

Abstract: Interfaces allow the exchange of data with systems outside of the ERP. When creating the ERP architecture, the interfaces shall be considered as vital elements of the company's processes rather than being single points of exchange. In this article, the holistic consideration of interfaces using the ACTIVATE methodology is being described

Keywords: SAP Activate, interfaces, Enterprise Resource Planning (ERP)

1 Introduction

When introducing new ERP systems, the ACTIVATE methodology has a proven record of success. The methodology provides ready to use business processes, a methodology as well as accelerators, which allow a timely go-live of ERP systems.

An important part of the methodology is related towards the interfaces. The connection to systems is usually seen as a single point-to-point connection. However, when it comes into play with the ERP system a deeper analysis if the connection is required, since the decision which data is exchanged is the last part of the chain.

Die Schriftenreihe LNI umfasst Tagungsbände, die Seminarreihe der Informatiktag und die im Rahmen des Dissertationspreises der GI prämierten Dissertationen. Soweit hier Unterschiede in der Formatierung bestehen, wird im Folgenden explizit darauf hingewiesen.

2 Procedure to determine the interface

This section describes the procedure to determine the interface design.

2.1 Interfaces in Activate

The activities for interfaces are subsumed under the term integration in Activate [SAP23].

¹SAP Deutschland SE & Co. KG, George-Stephenson-Str. 7-13, 10557 Berlin,
carsten.brockmann@sap.com

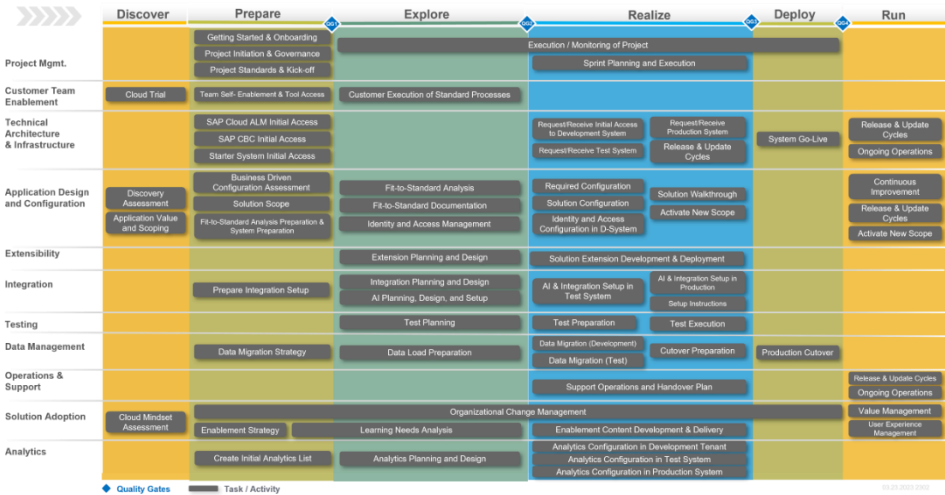


Fig. 1: Activate for 3SL Public Cloud ERP

Briefly described (Fig. 1), the Prepare phase consists of initiating the activities, the requirements are gathered in the Explore phase and implemented in the Realize phase.

2.2 Determine process flow between ERP and legacy system

Based on the overall enterprise architecture, the process flow between the ERP and legacy system shall be determined.

As an example, a client calls in to repair a device. If a ticket for that service request is entered within the legacy system, it should be defined how spare parts, that are required to conduct, are ordered and their arrival tracked so that the service technician arrives on-site with the required equipment.

The modelling of processes flows concludes with a discussion on which process step is done in which system. In addition, it is should be determined which system steers the process steps and how that integrates into the flow overall.

2.3 Derive required data objects and processes

Once the process flow between the systems is determined, the next steps consist of identifying data objects which shall be exchanged between the two systems. A simple way

of identifying these objects are process steps that are performed in the other system. In order to perform the steps, the data needs to be present.

2.4 Build interfaces

Once the process flow and data objects to be exchanged are known, the next step consists of developing the interface.

When deploying a standard ERP, it is recommended to use standard interfaces whenever possible since standard interfaces are maintained within the regular development cycles of the ERP. Hence they are expected to work after a release upgrade.

Upon the successful development of the interface, the flow as a whole should be tested between the two systems.

3 Conclusion

The development of an interface shall start with a clear process definition, followed by the definition of exchanged data. After both are defined, the interface itself should be developed.

By following this approach, it is possible create the interface in a structured manner.

References

[SAP23] SAP, Activate Roadmap Viewer, <https://go.support.sap.com/roadmapviewer/>.

Wirtschaft, Management Industrie -
Workshop zum Stand, den
Herausforderungen und Impulsen
des Geschäftsprozessmanagements
(ZuGPM 2023)

ChatGPT – Modelliere mir meinen Geschäftsprozess!

Ralf Laue¹

Abstract: In den letzten Monaten wurde vielerorts experimentiert und spekuliert, inwiefern große Sprachmodelle wie ChatGPT Arbeiten abnehmen oder unterstützen können. Da Geschäftsprozessmodelle auch sprachliche Modelle sind, liegt es nahe, die Eignung von ChatGPT für die Geschäftsprozessmodellierung zu prüfen.

In einer Live-Demo soll sowohl gezeigt werden, was möglich ist, als auch, wo ChatGPT derzeit noch grandios scheitert.

Keywords: generative KI; Sprachmodelle; ChatGPT; Geschäftsprozessmodellierung

1 Geschäftsprozessmodellierung mit ChatGPT

Dass ChatGPT grundsätzlich in der Lage ist, Geschäftsprozessmodelle zu erstellen, wurde kürzlich in [FFK23] gezeigt. Allerdings war das dort verwendete Beispiel so klein und einfach, dass es keine Schlüsse über mögliche Grenzen des Sprachmodells zulässt.

Wir wollen in einer Live-Demo versuchen, etwas komplexere (jedoch immer noch kleine) Modelle zu erzeugen. Wir nutzen das Prompt-Muster *Meta-Spracherzeugung (Meta Language Creation Pattern)* [Wh23], um eine Notation für eine Graphstruktur einzuführen. In dieser lassen sich dann einzelne Modelle darstellen, zu denen von ChatGPT sinnvolle Vorschläge etwa für Erweiterungen macht. Als hilfreich erweist sich dabei die Diskussion möglicher Ausnahmefälle.

Wie wir weiter sehen werden, zeigt ChatGPT dann aber Schwächen bei der Anwendung von Alltagswissen. Zum Beispiel bleiben offensichtlich falsche Reihenfolgen (wie: etwas wird erst verschickt, danach verpackt) unbemerkt.

In der Demonstration wird auch sichtbar, dass Notationselemente zur parallelen Ausführung von Aktivitäten von ChatGPT nicht genutzt werden, was angesichts der Tatsache, dass das Sprachmodell auf (sequentieller) natürlicher Sprache trainiert wurde, nicht verwunderlich ist.

Als eine interessante Anwendung zur Unterstützung des Geschäftsprozessmodellierers wird das Prompt-Muster *Umgekehrte Interaktion (Flipped Interaction Pattern)* [Wh23] gezeigt,

¹ Westsächsische Hochschule Zwickau, Frachgruppe Informatik, Dr.-Friedrichs-Ring 2a, 06056 Zwickau,
Ralf.Laue@fh-zwickau.de

bei dem ChatGPT so lange Fragen stellt, bis genügend Informationen vorliegen, um ein Geschäftsprozessmodell zu erzeugen.

Der Vortrag soll beim 8. Workshop zum Stand und den Herausforderungen des Geschäftsprozessmanagements eine Anregung für Diskussionen um Möglichkeiten und Grenzen der Nutzung künstlicher Intelligenz im Bereich des Geschäftsprozessmanagements geben.

Literaturverzeichnis

- [FFK23] Fill, H-G; Fettke, P.; Köpke, J.: Conceptual Modeling and Large Language Models: Impressions From First Experiments With ChatGPT. Enterprise Modelling and Information Systems Architectures, 18:3, 2023.
- [Wh23] White, J.; Fu, Q.; Hays, S.; Sandborn, M.; Olea, C.; Gilbert, H.; Elnashar, A.; Spencer-Smith, J.; Schmidt, D. C.: A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. CoRR, abs/2302.11382, 2023.

Specification and Application of a Domain Specific Modeling Language for Social Services

Felix Holz, Dennis Vogel and Michael Fellmann ¹

Abstract: For developing process-oriented capabilities and henceforth advancing digitization, there is a need to structure and formalize the processes in the domain of social services to an extent where it can support social workers. Processes in social services are challenging to model due to their unstructuredness and knowledge-intensity; they require high flexibility from the workers, primarily caused by intensive work with human beings – the clients. Declarative languages like CMMN and ConDec promise to be a solution. Still, they are too complex and overloaded for persons without any experience in modeling to use them in a practical context. In this paper, a specification for a domain-specific modeling language (DSML) is defined to enable representing the process knowledge in the domain by simplifying the language to the domain's core concepts: (client) *goals* to be achieved, occurring *situations*, and *actions* to be carried out. We applied the modeling language in a German real-world social care organization and elicited models, facilitating a workshop-based approach. The domain experts confirmed the resulting model's usefulness, especially for guiding, training, and planning purposes.

Keywords: Social Services, Personal Services, Domain Specific Modeling Language, Notation, Flexibility, Knowledge-intensive Processes, Declarative Modeling

1 Introduction

Digitization and IT-Support are important factors for organizations to improve the work in a more effective and efficient manner. As a prerequisite for advancing these factors in an organization, existing enterprise aspects are to be formalized and structured. Especially for processes in social service companies, this task proves to be difficult; hence the work is weakly-structured, highly flexible and knowledge-intensive [HL17, HFL22, LBL19]. Compared to routine processes modeled with imperative languages, declarative modeling languages like ConDec [Pv06] or the Case Management Model and Notation (CMMN) [OM14] are more suited to model knowledge-intensive processes [GVC15, HSW22]. In the past, however, weaknesses have emerged in the practical application of these modeling languages [HL17]. On the one hand, CMMN is indeed usable for depicting the casework on a general level but seems to only partially correspond to specific needs of the domain, especially regarding the core work like client interaction [HL17]. On the other hand, ConDec is difficult to read and access for non-method-experts [MGP15]. Consequently, an approach to design a domain specific modeling language (DSML), more targeted to the focus group,

¹ University of Rostock, Institute of Computer Science, Albert-Einstein-Str. 22, 18059 Rostock, Germany
{felix.holz,dennis.vogel,michael.fellmann}@uni-rostock.de

promises to be more suitable for the domain's needs [Fr13]. To enable modeling the social core work with high client interaction, we stated three central concepts in our previous work - *goals* to be achieved, *situations* to be encountered, and *actions* to be taken [HFL22]. We expect easier manageability by adapting the domain vocabulary to a simplified version of existing languages. In order to collect the implicit process knowledge of the social workers for analyzing, reusing, and explicating process patterns, a modeling language for the domain of social services is proposed in this paper, formally specified and applied in practice in a German social service enterprise.

The remainder of this work is structured as follows: Chapter 2 covers the theoretical background of the domain. Chapter 3, gives the DSML's formal specification while loosely following the guidelines for designing a DSML according to Frank et al. [Fr13], alongside a comparison with CMMN. Chapter 4 illustrates a workshop-based approach for gathering process knowledge and representing it with the given language. The application and evaluation of this approach are found in Chapter 5. Finally, it concludes with limitations and future work in Chapter 6.

2 Background: Processes in the Domain of Social Services

Social services are classified as personal services aiming to improve the client's mind or body [Ha67]. Therefore, service providers work in close cooperation with clients while considering their situation and context [LBL19]. Consequently, social workers must bring their experience and knowledge to the work process to achieve given, usually long-term goals under uncertain conditions. The operations comply with knowledge-intensive processes and require high flexibility of process support [MGP15]. Knowledge-intensive processes are characterized by both unpredictability and knowledge intensity, i.e., the process course and outcome depending on the actor's tacit knowledge [SL20]. According to Boissier et al. [BRL19], and Işık et al. [IMv13], the process success is not apparent before and after execution. This is caused by highly individual processes for different customer requests. Each request represents a separate process instance, namely cases. These are complex activities that require creativity, collaboration and the simultaneous inclusion of a large amount of information as unplanned situations occur frequently. An additional characteristic of German social service organizations lies in the regulations they must comply with (e.g., social security codes). Moreover, collaborating with governmental carriers, such as social and youth welfare offices, is essential. These factors are accompanied by a reporting obligation for which the social workers must document the activities with their clients. Ultimately, case management and client interaction form the core of social work that needs to be supported.

Mertens et al. state that such dynamic, non-standard, and knowledge-intensive processes require the run-time flexibility of declarative process modeling [MGP15]. According to Sadiq et al. [SSO01], flexibility is the ability of a workflow process to be executed from a "loose" or only partially specified model, while the full process specification only occurs at runtime, where each process instance can be unique. Schonenberg et al. [Sc07] define a

taxonomy with four ways in which flexibility can be achieved in process representations, whereas Reichert and Weber [RW12] state four requirements for the performance of flexible process specifications subjected to process-aware information systems (PAIS).

As stated in [HFL22], the most important criterion for social service providers is “flexibility by underspecification” or “looseness”, i.e., a free process specification that also includes unplanned activities at run-time. [RW12]. The former defines the ability to execute processes at run time that are incomplete at design time [Sc07]. The latter concerns a vague process definition at design time to make process decisions and deviations possible at run time, whereby a superordinate process goal is pursued [RW12].

Since it is necessary for social workers to have a full scope of action during the processing of their tasks and to be able to make situation-specific decisions, no stoic process should be modeled. Still, a selection of meaningful activities should be highlighted which are helpful for the current situation.

3 Language for Modeling Processes in Social Services

To create a DSML, Frank et al. [Fr13] provide guidance in several phases. Initially, the language’s scope and purpose are to be clarified. Then, generic and specific requirements for a modeling language within the domain are to be given. In the domain, essential generic requirements (GR) for a DSML are (GR1) *reconstruction and representation of existing domain concepts*; (GR2) *different levels of detail and abstraction* for better *readability* and *transferability* to several cases; (GR3) *simplicity* by restricting to the most important concepts for the purpose of DSML; (GR4) *comprehensibility* of the representations; (GR5) *usability* – the language should be usable for the purpose, particularly for social workers with little modeling experience regarding processes.

In our previous work [HFL22], the following benefits that (process) models can provide to the domain were identified: (M1) Possible agenda to support case and action planning, as well as estimation of client interaction expectations; (M2) illustrativeness of the current status of the case; (M3) overview and the possibility of comparison of actions taken by colleagues as a recommendation for action and induction of new staff; (M4) basis for discussion to support teamwork; and (M5) assistance for self-reflection after case completion.

Alongside these insights combined with the given literature, the following specific requirements (SR) emerge: Regarding (SR1) support of knowledge-intensive work [SL20], first (SR1.1) be able to list and accommodate frequent (unforeseen) circumstances. And second, (SR1.2) be able to deal with the individuality of cases. Therefore, providing enough abstraction supports most cases and avoids patronizing the workers. They should be free to do what is right, according to their experience. The model should have a support function. (SR2) Support of highly flexible work means to enable specifying processes and workflows loosely and possibly deviating from the process definition at run-time [RW12, Sc07]. (SR1.2) applies here as well. (SR3) Recognizing work patterns to enable repeatability and reuse of positive patterns and identifying negative patterns; optimally, deriving reasons for certain patterns.

3.1 Specification and Meta Model

In this section, a modeling language is specified which is adapted to the domain of social services and, in particular, should enable the modeling of processes with high client interaction, i.e., the social worker's core work. Thereby, the strong inspiration of CMMN cannot be denied. Still, simplifications are needed for use in the domain because CMMN offers too many possibilities, which can overwhelm a social worker [HL17, HFL22].

For the specification of a DSML, Frank et al. [Fr13] recommend creating glossaries identifying and defining the most important domain-specific concepts and developing a concept dictionary and a meta model. The domain's three core concepts are derived from [HFL22]. Bork et al. [BKP20] highlight a wide range of approaches and techniques for meta modeling and therefore modeling language specifications. In every case, a language specification is composed of three central parts: defining the *syntax*, the *semantics*, and the *graphical notation*.

Figure 1 shows a simple meta model of our proposed Social Service Notation (SSN), graphically representing the syntax and how the core concepts are related. Here, we differentiate between Model Elements (depicting the core concepts) and Relations. The Model Elements can be associated to each other with two different Relation types. The Connector type relates any two model elements with each other. A Model Element can be connected with multiple Model Elements. The Affiliation inherits from the Composition relation for building a hierarchy. Only Container Model Elements can be a Composition of further Model Elements.² Henceforth, a Goal can contain further Goals, Actions, and Situations. On the other hand, these composed Model Elements belong to exactly this Goal.

In Table 1, the descriptions of the Model Elements can be found alongside their semantics and symbols. Since the SSN is a simplification of CMMN [OM14, Ro21], similarities can be found. The Goal concept has a similar purpose as the CMMN *Case* and *Stage Plan Item*, i.e. can be opened and closed to direct the user focus (fulfilling GR2 and the principle of complexity management [Mo09]); Actions are the domain specific form of tasks with less options to customize them with additional symbols (like human task or repeatability); and the Situation resembles most to the *Event Listeners*. Since Situations have higher importance in the domain, the symbol reflects it with larger size than the *Event Listeners*.

The possible Relations are defined in Table 2. The Affiliation relation is intended to give the corresponding Model Elements the context, in which they placed in. Like in the declarative paradigm [Pv06], the Affiliation represents that an Action can be carried out at every time, if it corresponds to the affiliated Goal, unless there is a Prerequisite to be met. The Prerequisite connector does not mean that one process step must necessarily follow to another (but certainly can). Regarding Syntax, both the Prerequisite and the Recommend connectors can be used the same way – connecting each Model Element with each other. Semantically, the former implies a strong, obligatory binding between model elements (i.e., it *has to* be done before), whereas the latter softens this relation to show a

² Note, that Composition is an abstract meta element class to specify the syntax. Affiliation is the actual class to be used in the model.

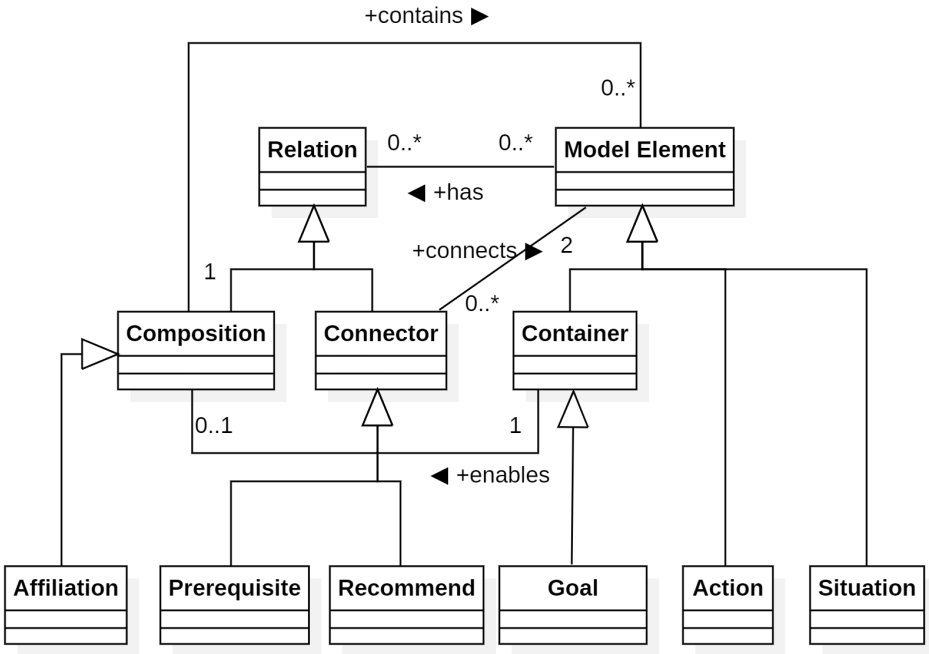





Fig. 1: Meta model of the Social Service Notation (SSN)

noteworthy association. (e.g., after doing homework, sleeping is recommended. But it is possible to sleep without doing homework before.)




Figure 2 shows a model sample in the given modeling language. There is to be seen that regarding the Goal of *Supporting Financial Handling, Networking Debt Counseling* is only a valid Action, if the Situation [Client has] *Debts* occurs. *Reflection of Financial Behavior* is reasonable to be carried out at every time, if the Goal is relevant to the client. Same goes for the Action *Supporting w/ Financial Authorities*, however, this Action is especially recommended, if the client is *Unable to Pay Bills*.

3.2 Comparison with CMMN

Since CMMN was created to model flexible and knowledge-intensive processes [OM14, Ro21] and the SSN is a simplified form of it, both languages will be briefly compared in this section. For the empirical comparison on model level, several authors enlisted constructs for evaluating the model quality (e.g., [MLv15, Ar07, OBS12]); in [Di14, MHL13] metrics for comparing meta models are given. However, since the SSN is designed to express domain concepts to domain experts, we compare the languages regarding cognitive capabilities. Therefore, we are using the concepts from the framework to evaluate model comprehensibility

Model Element	Description
Goal 	Describes client goals to be achieved while coping with the case. A goal can contain sub-goals, situations, and actions. Depending on the client, the importance of individual goals differs, yet reflects possible “areas of work” or “work phases” (context) in the entire case.
Situation 	Describes the state in which clients may be. A social worker must be able to respond to such situations, so they are often a trigger.
Action 	Describes a client-related task to be carried out by the social worker.

Tab. 1: Central Model Elements of the SSN

Relation	Description
Affiliation 	Represents that the corresponding elements are relevant while working on the associated Goal, i.e., they belong to the Goal. Only Goals can contain Model Elements. The Affiliation builds a hierarchy and is transitive.
Prerequisite 	Represents a directed connection between two Model Elements. The element from which the connector originates acts as a precondition for the element into which the connector points. The latter can only be executed or occur if the preconditions are fulfilled (e.g. completion of an action, the occurrence of a situation, resolving a goal).
Recommend 	Represents a directed noteworthy connection between two Model Elements, which are recommended to execute together. This connection forms a meaningful relationship, but not an obligatory prerequisite for each other – the connected Model Elements are allowed to occur independently of each other.

Tab. 2: Central Relations of the SSN

according to Aranda et al. [Ar07] empirically on a language level.

First, we refer to “External Cognition” as a criterion for evaluating models. The representation should be adapted to the human way of thinking to find practical solutions. As part of it, “Offloading” reduces a person’s cognitive effort by minimizing knowledge that must be kept in mind while interacting with the model and the language.

A model supporting “Re-representation” means making problem-solving easier due to their cognitive fit; “Graphical Constraining” is defined as allowing a limited number of conclusions for using cognitive power more effectively. The more direct and simple the representations are kept, the faster the actual solution process can be started [Ar07].

The advantage of the SSN consists of being tailored for the use case and domain; CMMN applies to a broader range of knowledge workers. The main differences are the number of usable model elements and relations. The SSN has three model elements, CMMN has eight with further features (like task types). This gives the modeler significantly more possibilities to design a model. According to Routis et al. [Ro21], the modelers distinguish the importance

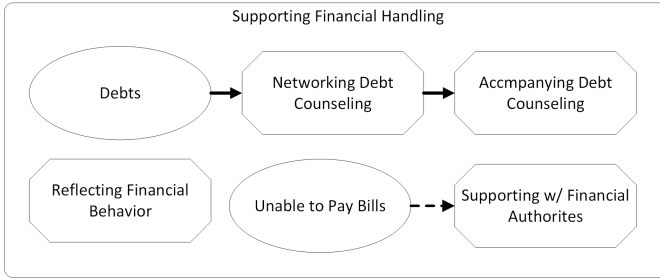


Fig. 2: Model sample for the SSN

of the model elements; stages and discrete items were perceived as important, whereas milestones and event listeners “were not considered that important”. Additionally, “the ‘Ease of Use’ usability factor got the lowest score regarding its prospect of the participants’ intention to adopt CMMN [...]” [Ro21]. Too many possibilities to represent a real-world aspect can be overwhelming for persons without experience in process modeling. The same is true for interpreting the models [MLv15].

We argue from past practical interactions that cutting unnecessary parts out and adapting the important aspects to the domain might increase the intention to use [HL17, HFL22]. As part of the “Cognitive Dimension Framework”, the term “Hidden dependencies” refers to the aspect of whether essential links between entities are visible and if users have to go through a complicated process to uncover these dependencies [Ar07]. In CMMN, sentries define dependencies between model constructs, but no visual distinction in the type of dependencies can be recognized. The SSN distinguishes dependencies with different types of relations. A disadvantage of the SSN is to be stated here: SSN lacks differentiation between planning and modeling, while CMMN can facilitate this through discretionary items. Both parts are merged in the SSN, resulting in concept overload [Mo09], because goals have several meanings.

4 Workshop-Based Procedure for Gathering Process Knowledge

Process modeling endeavors in the past, utilizing a semi-automated approach on the internal client case documentation, were lacking in accuracy [HLF21]. The need to consult domain experts for higher accuracy and, henceforth, intent to use became apparent. According to Sandkuhl et al. [Sa14] and Dumas et al. [Du21], there are different ways to collect necessary information for modeling a (business) model, such as (automatic and manual) document analysis, observations/shadowing, form-based surveys, or interviews. This section presents a workshop-based procedure to collect and externalize the implicit process knowledge of domain experts so that precise close-to-reality models can be derived from it in the given modeling language.

The approach envisions that the domain experts will interact and discuss with the method

experts and each other. The domain experts should be workers in the domain's operational business who are involved in the processes; thus, social workers with high client interaction are optimal for modeling with the SSN. The workshop-based approach suggests grouping multiple domain experts. Opposing to an individual interview, the advantage of a group of domain experts is the discussion of processes through different perspectives on similar activities, leading more naturally and quickly to a more generalized form of the processes [ØL17]. Additionally, group members can correct and enrich the facts described. Expert group members should remain the same throughout the workshops. To evaluate the models in the last workshop, additional people can be brought in to review the results for more unbiased opinions. Interviewing and modeling experience and familiarity with the used modeling language are necessary for method experts.

Figure 3 shows an overview of our proposed procedure, divided into three workshops. The first introductory workshop is used to acquire information on facts to be modeled, especially collecting the relevant Model Elements (Goals, Situations, and Actions). In the second workshop, the results of the first are refined, corrected, and generalized. Here, the focus lies on consolidating existing and acquiring additional Relations between the model elements. Thus, the Prerequisite and Recommended relationships are to be emphasized. For the third workshop, the method experts transform the validated information into a graphical representation beforehand. The workshop itself is to be used for presenting and evaluating the final models. Our proposed evaluation criteria for assessing the models are comprehensibility (derived from [Fr13, Hi20]), correctness [OBS12], complexity [MLv15] and usefulness for the domain. Further evaluation criteria can be found in [Ar07, Mo09, Ro21].

With the SSN at hand, the information to be gathered and presented follows a certain pattern. Starting with the Container concepts Goals and their sub goals to place further Model Elements in an appropriate context. For each of these Goals, the Situations and Actions are discussed alongside their interrelationships with each other. In the following Chapter 5, the application of the procedure with corresponding results can be found.

5 Practical Application of the Notation

With the objective of structuring the knowledge-intensive work of the domain, the procedure mentioned in chapter 4 was applied in practice as an experimental evaluation. This chapter presents the application's results and their evaluation in excerpts. We worked with a German social institution in the field of child, youth, and social welfare. This organization is active in both ambulatory and stationary areas, with about 150 employees. The different client's thematic areas (such as educational assistance and addiction disorders) [HLF21] are divided into so-called settings. The study was conducted in the setting of the "Clean Living" project, a semi-stationary setting for people with addiction issues who are also supported with matters concerning their living situation. About 36 clients are accommodated in this setting. Four social workers and pedagogues (4 f, 0 m), working mainly in this setting, could be

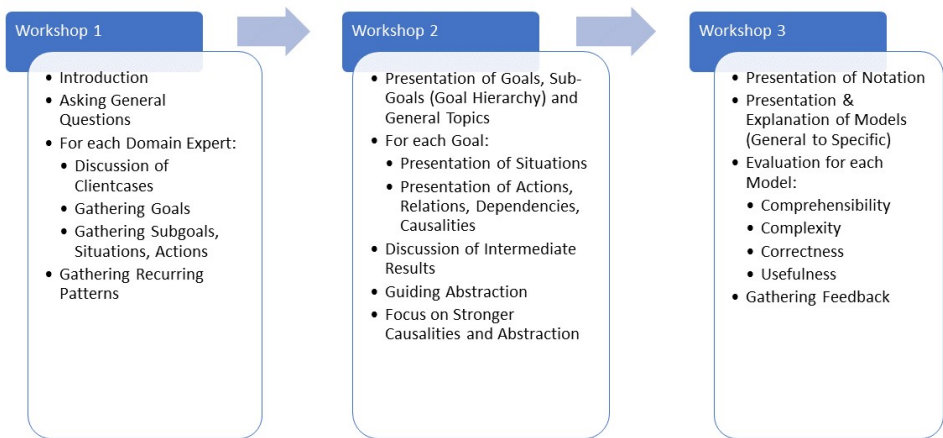


Fig. 3: Overview of the workshop-based approach for each group to model client interaction in social services alongside the workshop's contents

acquired as participants. The participants' ages differ from 28-36 years, and they work in this setting for about two to three years. The workshops were carried out online in March 2022.

Following the mentioned approach, three participants were present in the first workshop, and one had to be interviewed separately. It was noticeable that the contents of the individual interview were more instantiated and exemplified. In the second workshop also, only three of the four participants attended. In the third workshop, the fourth participant took part in examining the results. Due to time restrictions, the fourth participant was absent from the first two sessions, so she had a more unbiased point of view on the results. Schematically, we followed the same procedure for gathering or presenting the information in the workshops: Description of goals and, for each goal, the associated sub goals, situations, actions, and their relationships.

5.1 Results: Elicited Models

The final result of the workshops is an overall model, which represents the knowledge of domain experts about processes with high client interaction. The overall model can be divided into the following goal areas: (G1) *Abstinence*, (G2) *Household and residential situation*, (G3) *Parenting assistance and family stabilization*, (G4) *Vocational orientation*, (G5) *Healthcare*, and (G6) *General Processes*.

For the purpose of structuring the work, a catalog of possible actions is an additional outcome that emerged from the workshops. During the workshops, we noticed that certain actions are often repeated, so we formed a set of finite actions. However, since the work is characterized by high individuality [HLF21, LBL19], the same actions may address

different topics/subjects depending on the client and the objective. This is facilitated by the action description in the form *Action [subject]*, whereas [subject] is a replaceable, customizable non-terminal from an infinite set of possible client topics. The advantage of such a catalog is partial standardization and formalization by reduction of actions. At the same time, it still contains flexibility through a certain degree of abstraction since it describes what has to be done but leaves the action's arrangement to the domain expert (fulfilling SR1). This form of activity description was used during the creation of the models. Such a catalog could not be designed for situations since individuality is too high. The challenge lies in the generalization of such situations. The catalog contains following actions:

- Observation [of what?]
- Support/Guidance [subject]
- Networking [contact]
- Conversation/talk [subject]
 - Advice [subject]
 - Motivation [subject]
 - Reflection [subject]
 - Clarification [subject]
 - Relationship work
- Assessment [subject, situation]
- Accompanying [subject, client activity]
- Organisation [subject]
- Control [subject]
- Strengthening [subject]
- Strategy development [subject]
- Training [subject, competence]
- Competence development [subject, competence]
- Prevention [subject]

Since the overall model is too large and partly confidential, Figure 4 shows the sub model with the goal area (G1) *abstinence*. For demonstrating the notation in practice, some of the model's statements are highlighted and explained in the following:

1. The Goal of this model is (G1) *abstinence*, and it has the sub goal for *preventing addiction pressure*. This goal has further sub goals to divide the means for prevention into short-, mid-, and long-term actions. They may not appear as classical goals but more like work phases. A **goal hierarchy** is built with the nested *Affiliation* relations to provide better structure and an overview of which measures can be taken depending on time criticality.
2. Acting as a start-situation, if a new client has arrived with the goal of *abstinence*, it has to be clarified how the client deals with the addiction and the reasons for it. Note that *clarification*, compared with *conversation*, has a more binding and committing means to it. This fact is implicitly known but not explicitly modeled. Before the social worker can begin with the *networking* action, the aforementioned actions have to be carried out, i.e., they are a *Prerequisite* for this action.
3. A Situation like *acute addiction pressure* can also be a *Prerequisite* for an Action. In this instance, the action for *assessing the addiction pressure* is only necessary (or can only be carried out) when the social worker is aware of the situation of *acute addiction pressure*. The awareness of how the client is coping with it (either successful or with recidivism) is only possible after the assessment. Depending on success, further actions are to be carried out.
4. Actions like *Control addiction pressure* or *on-call service* do not have any *Prerequisites*, meaning they can be carried out whenever the corresponding Goal is active for the client.
5. The *Recommended* relation in this model is mainly depicted around *acute addiction pressure*, meaning first, if the pressure is perceived, it is recommended to initiate the Goal for *prevention*. On the other hand, it is recommended to have worked on means for prevention before the addiction pressure arises. Using *Prerequisite* instead of *Recommend* connectors would result in a deadlock. *Listlessness* can result from the addiction pressure, which is an issue for *long-term prevention* actions.

Also noticeable while modeling were additional model elements (drawn from BPMN or CMMN) that also seem appropriate here. A model element that references existing documentation (such as checklists for new admissions, help plans, or the emergency plan for child endangerment) proved to be useful in other sub models. Regarding domain concepts, an additional model element, “Third Party or Organization” was found during the workshops. In particular, the *networking* activities occur frequently and have a high value in the domain. It is reasonable to highlight which network partners are appropriate for which goals visually.

5.2 Preliminary Evaluation of the Models

Figure 4 shows one of five sub models; each sub model was presented separately and rated by the domain experts based on comprehensibility, complexity, correctness, and usefulness on

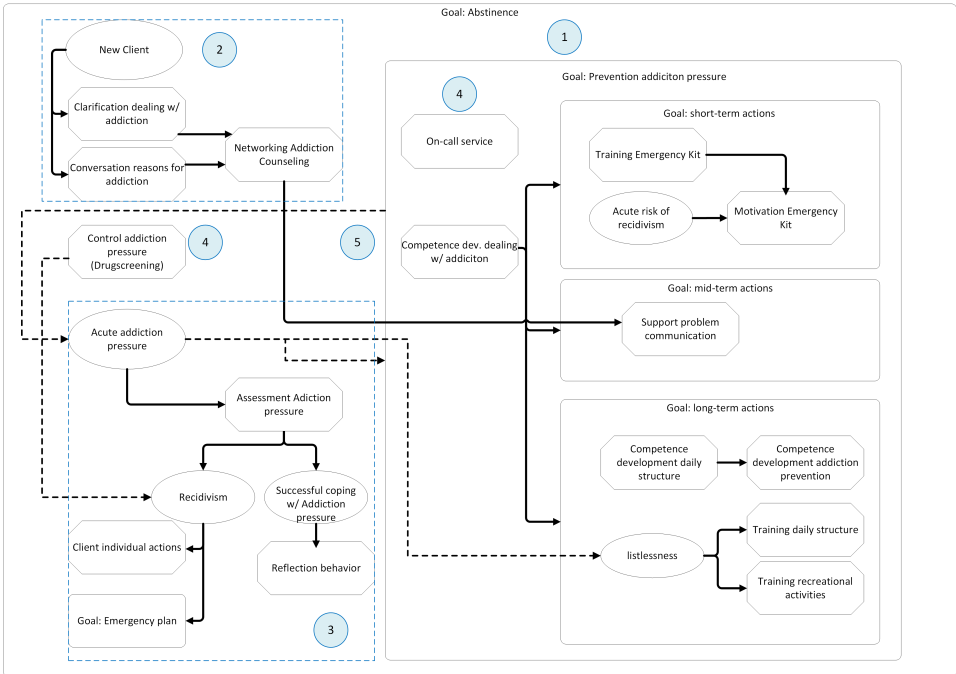


Fig. 4: Process (knowledge) model in the SSN-Modeling language for dealing with client’s goal of maintaining (G1) abstinence

a 5-point Likert Scale. The results can be seen in Table 3. In terms of comprehensibility, the participants agreed that, at first glance, the models were “overwhelming” and complicated (which would also correspond to reality). Still, after the models were worked through step by step and explained, the participants considered them to be comprehensible. The sub model (G1) *Abstinence* shown in Figure 4 earned the highest ratings for comprehensibility and complexity (4.75 each). Correctness and usefulness have also been rated as high (4 each). This also shows that this is the primary client goal in the chosen setting, and accordingly, it was talked about the most and worked on in the most detail. Regarding complexity, the participants stated that too much detail would be too limiting, and a more abstract approach would help to facilitate flexibility (fulfilling SR1.2, SR2). In terms of correctness, the naming of some model elements was partly perceived as misleading. Furthermore, the level of detail varied between the sub models, and the participants wished for more completeness in some of them. Therefore, due to the individual client cases, some modeled areas are sometimes more, sometimes less important for certain clients.

The model rated as least comprehensible (2.5/5) belongs to the (G4) *Vocational Orientation*. Here the correctness is also the lowest (2.75). This was explained with too specific element descriptions; the model seemed in parts more like an example, and many aspects that should belong to it did not occur. (G5) *Health care* had the least complexity. The sub model was

also rated as incomplete, but unlike (G4) *Vocational Orientation*, it was not really wrong. Nevertheless, the usefulness (with 3/5) is the lowest for health care, mainly due to the low information content.

The participants perceive a correlation between correctness and usefulness (the more correct, the more useful). With the little data at hand, no such conclusions can be made. The most useful are the (G6) *General processes*, i.e., processes that cannot be assigned to any direct goal and tend to represent the majority of clients. Here, the connections between individual goals are also summarized, especially in the case of emergency plans ³ or help plans. So we assume standardization of processes that can be standardized is desired (matches SR3, M3). Here also, the comprehensibility and the complexity (with 4,5/5 each) are quite high. (G3) *Stabilization of the family* also has a high score (4.5/5) in usefulness, but correctness (3.5) is rated lower than in G1. This sub model has by far the most model elements listed. This indicates a need for structured information, especially for large and complex topics (where it is easy to lose track of all the possibilities).

The domain experts considered the usefulness of models to be very high, especially for newcomers to the profession, to improve the induction of new employees (matches M3). Mainly, the participants saw an added value of the models in having an overview and orientation for the entire scope of action as well as being useful as a reminder, especially compared to already existing concepts like the text-based help plan (matches M3). The models were further seen as an action-guiding tool to counteract “gut feeling decisions”. It was positively received that it depicts not a strict directive but more a guideline, possibly assisting in a rather abstract but comprehensible manner (matches SR1.2). To embed the models as support in the work, the participants see the highest potential in planning support or help plan preparation, less support in the everyday work.

Concept	Comprehensibility	Complexity	Correctness	Usability	No. of ME
(G1) Abstinence	4,75	4,75	4	4	28
(G2) Household	3,25	4	3,5	3,5	27
(G3) Parenting Assistance	3,75	4,5	3,5	4,5	58
(G4) Vocational Education	2,5	4	2,75	3,25	37
(G5) Healthcare	3,5	3	3,25	3	16
(G6) General Processes	4,5	4,5	3,75	4,75	32
Average	3,71	4,125	3,46	3,83	33,17

Tab. 3: Evaluation criteria scores and number of model elements (ME) for the elicited sub models

³ Emergency plans summarize, what kind of emergencies can arise, and what is to do in such situations.

6 Conclusion, Limitations, and Future Work

In this paper, we formally specified a domain specific modeling language for the domain of social services – the Social Service Notation (SSN) – and applied it in practice to gather and model the social worker’s implicit process knowledge. Whereas CMMN can also be appropriate for modeling flexible, knowledge-intensive processes, the proposed DSML could be seen as a simplified version of it, aiming at reducing the complexity to the necessary parts and facilitating domain specific concepts for increasing the intention to use. The use case for the language is narrowed down to the core work area of the domain – processes with high client interaction.

The application of the SSN was carried out in practice, and the preliminary results are promising. The approach to gathering the implicit process knowledge works fine, and the resulting models are perceived to be useful in practice, especially for inducting new employees. Goals showed to be the concepts for structuring the other model elements: situations and actions. Additional model elements (like third-party organization or documentation reference) were found and are reasonable to include in future versions of the SSN with their own symbols. Regarding symbols, future considerations about visually highlighting more important or frequent model elements should be made. Regardless of the modeling language used, the given procedure helps structure the work of social service providers. However, there are many limitations in the results and the workshop-based study carried out: The amount of participants is relatively low (four persons), and the order of presented models influences their evaluation, e.g., the first presented models are rated higher in complexity since it is the first shown model. Furthermore, we focused more on some modeled sub goals over others since time is a heavy restriction while working with practitioners. Henceforth, the perceived model quality varies considerably for each sub model.

For future work, we see potential in tool support to embed the models in an information system with role-specific views. Therefore, exchange formats (i.e., XML standards) and execution semantics (i.e., token-based like in BPMN) for the DSML have to be created and implemented. Addressing a full-fledged evaluation, the next step is to carry out an empirical investigation, where participants model and interpret models written in both languages, CMMN and SSN. This enables comparing the languages regarding model quality to further prove the proposed language’s usefulness and validity.

Bibliography

- [Ar07] Aranda, J.; Ernst, N.; Horkoff, J.; Easterbrook, S.: A Framework for Empirical Evaluation of Model Comprehensibility. In: 2007 International Workshop on Modeling in Software Engineering. IEEE, Piscataway, NJ, p. 7, 2007.
- [BKP20] Bork, D.; Karagiannis, D.; Pittl, B.: A survey of modeling language specification techniques. Information Systems, 87:101425, 2020.
- [BRL19] Boissier, F.; Rychkova, I.; Le Grand, B.: Challenges in knowledge intensive process management. Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOCW, 2019-October, 2019.

- [Di14] Di Rocco, J.; Di Ruscio, D.; Iovino, L.; Pierantonio, A.: Mining metrics for understanding metamodel characteristics. In (Atlee, J. M.; Kulkarni, V.; Clark, T.; France, R. B.; Rumpe, B., eds): Proceedings of the 6th International Workshop on Modeling in Software Engineering. ACM, New York, NY, USA, pp. 55–60, 2014.
- [Du21] Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H. A., eds. Grundlagen des Geschäftsprozessmanagements. Springer Berlin Heidelberg, Berlin, Heidelberg, 2021.
- [Fr13] Frank, U.: Domain-Specific Modeling Languages: Requirements Analysis and Design Guidelines. In (Reinhartz-Berger, I.; Sturm, A.; Clark, T.; Cohen, S.; Bettin, J., eds): Domain engineering, pp. 133–157. Springer, Berlin and Heidelberg, 2013.
- [GVC15] Goedertier, S.; Vanthienen, J.; Caron, F.: Declarative business process modelling: principles and modelling languages. *Enterprise Information Systems*, 9(2):161–185, 2015.
- [Ha67] Halmos, P.: The Personal Service Society. *The British Journal of Sociology*, 18:13, 1967.
- [HFL22] Holz, F.; Fellmann, M.; Lantow, B.: Ein Modellierungskonzept zur Prozessstrukturierung für Soziale Dienstleister. In (Riebisch, M.; Tropmann-Frick, M., eds): Modellierung 2022. Gesellschaft für Informatik e.V, Bonn, pp. 171–180, 2022.
- [Hi20] Hildebrandt, T. T.; Andaloussi, A. A.; Christensen, L. R.; Debois, SØ.; Healy, N. P.; López, H. A.; Marquard, M.; Møller, N.L.H.; Petersen, A.C.M.; Slaats, T.; Weber, B.: EcoKnow: Engineering effective, co-created and compliant adaptive case management systems for knowledge workers. *Proceedings - 2020 IEEE/ACM International Conference on Software and System Processes, ICSSP 2020*, 2020.
- [HL17] Herzog, P.; Lantow, B.: Adaptive case management in social institutions [Adaptive Case Management in sozialen Einrichtungen]. *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft für Informatik (GI)*, 275, 2017.
- [HLF21] Holz, F.; Lantow, B.; Fellmann, M.: Towards a Content-Based Process Mining Approach in Personal Services. In (Augusto, A.; Gill, A.; Nurcan, S.; Reinhartz-Berger, I.; Schmidt, R.; Zdravkovic, J., eds): ENTERPRISE, BUSINESS-PROCESS AND INFORMATION SYSTEMS MODELING, volume 421 of *Lecture Notes in Business Information Processing*, pp. 62–77. SPRINGER NATURE, [S.l.], 2021.
- [HSW22] Haarmann, S.; Seidel, A.; Weske, M.: Modeling Objectives of Knowledge Workers. In (Marrella, A.; Weber, B.; Marrella, eds): *Business Process Management Workshops*, volume 436 of *Lecture Notes in Business Information Processing*, pp. 337–348. [s.n.], [S.l.], 2022.
- [IMv13] Işık, Ö.; Mertens, W.; van den Bergh, J.: Practices of knowledge intensive process management: quantitative insights. *Business Process Management Journal*, 19(3):515–534, 2013.
- [LBL19] Lantow, B.; Baudis, T.; Lambusch, F.: Mining Personal Service Processes. In (Abramowicz, Witold; Corchuelo, Rafael, eds): *Business Information Systems Workshops*, volume 373 of *Springer eBook Collection*, pp. 61–72. Springer International Publishing and Imprint Springer, Cham, 2019.
- [MGP15] Mertens, S.; Gailly, F.; Poels, G.: Enhancing Declarative Process Models with DMN Decision Logic. In (Nurcan, S.; Guerreiro, S.; Ma, Q.; Schmidt, R., eds): *Enterprise, Business-Process and Information Systems Modeling*, volume 214 of *Springer eBook Collection Computer Science*, pp. 151–165. Springer, Cham, 2015.

- [MHL13] Ma, Z.; He, X.; Liu, C.: Assessing the quality of metamodels. *Frontiers of Computer Science*, 7(4):558–570, 2013.
- [MLv15] Marin, M. A.; Lotriet, H.; van der Poll, J. A.: Metrics for the Case Management Modeling and Notation (CMMN) Specification. In (Barnett, R. J., ed.): *Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists*. ACM Digital Library, ACM, New York, NY, pp. 1–10, 2015.
- [Mo09] Moody, D.: The “Physics” of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. *IEEE Transactions on Software Engineering*, 35(6):756–779, 2009.
- [OBS12] Overhage, S.; Birkmeier, D. Q.; Schlauderer, S.: Qualitätsmerkmale,-metriken und-messverfahren für Geschäftsprozessmodelle. *Wirtschaftsinformatik*, 54(5):217–235, 2012.
- [ØL17] Ørngreen, R.; Levinsen, K.: Workshops as a Research Methodology. *Electronic Journal of E-learning*, 15(1):70–81, 2017.
- [OM14] OMG: , Case Management Model and Notation, version 1.0. Technical Report May, OMG, May 2014, 2014.
- [Pv06] Pesic, M.; van der Aalst, W. M. P.: A Declarative Approach for Flexible Business Processes Management. In (Eder, J.; Dustdar, S., eds): *Business process management workshops*, volume 4103 of *Lecture Notes in Computer Science*, pp. 169–180. Springer, Berlin, 2006.
- [Ro21] Routis, I.; Bardaki, C.; Dede, G.; Nikolaidou, M.; Kamalakis, T.; Anagnostopoulos, D.: CMMN evaluation: the modelers’ perceptions of the main notation elements. *Software and Systems Modeling*, 20(6):2089–2109, 2021.
- [RW12] Reichert, M.; Weber, B.: *Enabling flexibility in process-aware information systems: challenges, methods, technologies*. Springer Science & Business Media, 2012.
- [Sa14] Sandkuhl, K.; Stirna, J.; Persson, A.; Wißotzki, M.: *Enterprise modeling*. Springer, 2014.
- [Sc07] Schonenberg, M.H.; Mans, R.S.; Russell, N.C. ; Mulyar, N.A.; Aalst, van der, W.M.P.: *Towards a taxonomy of process flexibility (extended version)*. BPM reports. BPMcenter.org, 2007.
- [SL20] Szelagowski, M.; Lupeikiene, A.: *Business Process Management Systems: Evolution and Development Trends*. Informatica (Netherlands), 31(3):579–595, 2020.
- [SSO01] Sadiq, S.; Sadiq, W.; Orłowska, M.: Pockets of Flexibility in Workflow Specification. In (Kunii, H. S., ed.): *Conceptual modeling*, volume 2224 of *Lecture Notes in Computer Science*, pp. 513–526. Springer, Berlin, 2001.

Towards Identifying GDPR-Critical Tasks in Textual Business Process Descriptions

Leonard Nake ¹, Stephan Kuehnel ¹, Laura Bauer ¹, and Stefan Sackmann ¹

Abstract: Complying with data protection regulations is an essential duty for organizations since violating them would lead to monetary penalties from authorities. In Europe, the General Data Protection Regulation (GDPR) defines personal data and requirements for dealing with this type of data. Hence, organizations must identify business activities that deal with personal data to establish measures to fulfill these requirements. Especially for large organizations, a manual identification can be labor-intensive and error-prone. However, textual business process descriptions, such as work instructions, provide valuable insights into the data used in organizations. Therefore, we propose a first approach to automatically identify GDPR-critical tasks in textual business process descriptions. More specifically, we use a supervised machine learning algorithm to automatically identify whether a task deals with personal data or not. A first evaluation of our approach with a dataset of 37 process descriptions containing 509 activities demonstrates that our approach generates satisfactory results.

Keywords: Legal Compliance, General Data Protection Regulation (GDPR), Business Process, Task Identification

1 Introduction

Protecting the IT infrastructure is an essential task for companies since information security incidents become more frequent and the costs of managing and mitigating breaches have been growing over the years [Ac21]. Due to this phenomenon, legislators and companies define overarching requirements for IT security that companies must comply with. For instance, Article 32 (1) of the EU General Data Protection Regulation (GDPR) requires an organization to implement appropriate technical and organizational measures to ensure compliance with the protection goals of confidentiality, integrity, availability, and resilience when processing personal data. To fulfill this requirement, technical precautions (e.g., encryption and pseudonymization of personal data) and procedural configurations (e.g., activities and controls to ensure compliance in business processes) become necessary [Kü21]. Therefore, compliance with IT security requirements is a cost-intensive task [La15], which makes it essential to determine where potential IT security measures have to be applied since economic efficiency in this context

¹ Martin Luther University Halle-Wittenberg, Chair for Information Management, Universitaetsring 3, 06108 Halle (Saale), {leonard.nake, stephan.kuehnel, laura.bauer, stefan.sackmann}@wiwi.uni-halle.de,

 <https://orcid.org/0000-0001-8324-5641>,  <https://orcid.org/0000-0002-6959-9555>,

 <https://orcid.org/0000-0001-8911-0879>,  <https://orcid.org/0000-0002-3370-6785>

is important [CCR04].

One of the requirements established by the GDPR is defined in Article 32 (2): “In assessing the appropriate level of security account shall be taken in particular of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to personal data transmitted, stored or otherwise processed.” This requirement makes it necessary for organizations to determine where personal data is dealt with in their business processes. These business processes consist of sets of activities performed in the organization [We19]. Only when the activities dealing with personal data are identified, adequate measures can be established. For small organizations with only a few business processes, analyzing each activity manually might be a viable solution. However, larger organizations agglomerate vast numbers of business processes since even single projects can result in the creation of hundreds or more new business processes [BRU00]. Each of these business processes can contain hundreds of activities of different departments and legal entities [RMv09]. Hence, a manual analysis of all of these business processes regarding the use of personal data becomes a difficult and labor-intensive problem.

One widely used possibility to store information about business processes is in textual business process descriptions. They contain valuable information that could be utilized in such an analysis since organizations usually maintain hundreds of textual process descriptions [LvR18]. For this reason, some approaches automatically analyze textual process descriptions to gain insights into the business processes of organizations [LvR18], [FMP11]. However, to the best of our knowledge, there is no approach that analyzes such textual business process descriptions to gain insights into the use of personal data.

To address this research gap, we propose an approach that analyzes textual business process descriptions to gain insights into the transmission, storage, or processing of personal data in business processes. More precisely, the proposed approach analyzes the natural language inside the textual business process descriptions to automatically classify each task as either GDPR-critical or GDPR-uncritical. A task is seen as GDPR-critical if it deals with personal data defined by Article 4 (1) of the GDPR. Hence, we raise the following research question:

How to automatically identify GDPR-critical tasks in textual business process descriptions?

To address this research question, we base our research design on the method proposed by Leopold et al. [LvR18]. Our research contributions are threefold: we introduce three features, train a model, and conduct an evaluation. Our study shows that the proposed approach with the current dataset achieves a F1-score of 0.81. This paper is structured in the following way: In section 2, we discuss the theoretical background and related work. In section 3, we present our research design. Section 4 deals with the conceptual approach as it describes the proposed concept in detail. Section 5 contains an evaluation of the conceptual approach. Section 6 concludes the paper.

2 Theoretical Background and Related Research

Our paper is thematically located in the area of business process compliance. Compliance refers to the adherence to rules, i.e., acting in accordance with applicable regulations, which can originate from various sources, such as laws, directives, standards, etc. [Ra12]. More specifically, business process compliance deals with identifying, formalizing, implementing, checking, analyzing, and optimizing compliance requirements before, during, or after the execution of business processes [SKS18]. An important source of regulations on the security of information and data in Europe, which we focus on in this study, is the GDPR. Data protection laws in Europe have long been inconsistent and the development of the GDPR has addressed this issue [Se20]. The regulation provides individuals with more protection and control over their personal data in light of new technological developments [To20], [Se20]. To this end, Article 4 (1) EU GDPR writes the following about personal data: “personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;” Since its entry into force, violations of, e.g., Article 32 EU GDPR (“security of processing”) in conjunction with Article 83 (“general conditions for imposing administrative fines”) have been punishable by fines of up to EUR 20 million or 4% of the total worldwide annual turnover of the previous fiscal year. Thus, the level of potential sanctions increased dramatically compared to previously applicable rules, and so did companies’ needs for security to protect them against compliance violations resulting from the GDPR. Since it is not an easy task to keep track of all information and data processing activities of a company, our approach starts exactly at this point and makes use of task analysis with Natural Language Processing (NLP) technologies and supervised machine learning (SML). While NLP uses computer-based methods to understand, produce and learn human language content [HM15], SML uses labeled datasets to train algorithms for classifying data or accurately predicting outcomes [Li11].

This paper relates to two major streams of research. Firstly, technologies for identifying requirements and checking completeness in the context of privacy policies. Secondly, the application of NLP technologies in the context of business process analysis. The first stream of research can be divided into two types of approaches: identifying requirements of privacy policies and completeness checking of privacy policies. Regarding the first type, Caramujo et al. [Ca19] propose a domain-specific language for specifying privacy policies in the context of mobile and web applications. This language allows policy authors to define a privacy policy as a set of declarative statements. They apply this language to support the analysis and comparison of policies from different companies. Pullonen et al. [Pu19] present a multi-level model as an extension of BPMN (Business Process Model and Notation). This enables the user to visualize, analyze, and communicate the characteristics of privacy policies in business processes. Kumar and

Shyamasundar [KS14] use information flow controls as a means to specify and enforce privacy policy requirements. Although some of these papers also use business processes to identify privacy policy requirements, the overall goal is different since we try to identify tasks that deal with personal data. Additionally, they are not strictly based on the GDPR. Regarding completeness checking, Torre et al. [To20] and Rahat et al. [RLT22] propose approaches that check the completeness of privacy policies against the GDPR. In summary, the approaches from the first research stream deal with similar problems. However, there is no approach that addresses our research question. The second stream of research is task analysis utilizing natural language texts in business process management. In this field, process-related textual documents are analyzed to gain insights about the tasks of organizations [RRM21]. Relevant work in this area extracts process models from textual business process descriptions [FMP11], compares textual descriptions with process models [vLR17], identifies candidates for task automation [LvR18], or extracts relevant task content aspects from textual task descriptions [RRM21]. Our research can be seen as a bridge between the two main research streams. It uses methods from the task analysis in business process management domain to solve a problem regarding privacy policies, more specifically, the GDPR. To the best of our knowledge, there is no existing approach for automatically identifying GDPR-critical tasks using textual business process descriptions.

3 Research Design

This research is part of a larger Design Science Research (DSR) project [Kü21] following the procedure proposed by Vaishnavi and Kuechler [VK15] with the goal to create a method that automatically identifies GDPR-critical tasks in textual business process descriptions. In this paper, we present the results of the first iteration, where we took an already existing method, applied it to our identified problem, and evaluated the results. In this study, we base our research design on an existing approach by Leopold et al. [LvR18]. In their research, tasks are identified by analyzing the natural language in textual business process descriptions following the approach by Friedrich et al. [FMP11]. Subsequently, the analyzed textual business process descriptions are used to compute a set of features. Using these features, the identified tasks are then classified into manual, user, or automated tasks using SML. This is done to identify candidate tasks for robotic process automation. We follow the overall structure of this three-step approach by Leopold et al. [LvR18] but adapt the second and the third step for our approach to answer our research question. Figure 1 shows a visualization of the proposed approach. Firstly, a textual description of a business process is parsed to identify the linguistic entities in the text, such as the verbs and objects that describe a task. This step results in an annotated textual process description with identified tasks and their respective linguistic entities. Secondly, the linguistic entities of the annotated textual process description are analyzed to compute features for our model. More specifically, the object and data item of the identified task are extracted as features and the customer relation of the task is calculated. This step produces a table of the tasks as well as their features. Thirdly, the classification is performed using the features. We use a support vector machine (SVM), a SML approach,

to classify tasks as GDPR-critical or GDPR-uncritical based on manually classified training data. The output of this step is a list of classified tasks.

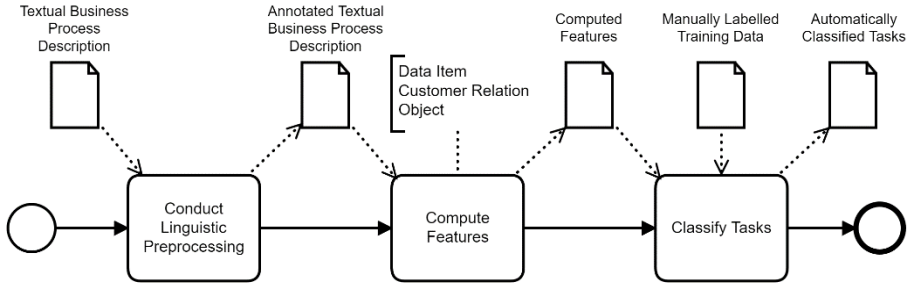


Fig. 1: Research Design

4 Concept of the Approach

4.1 Idea for the Approach

The idea behind our approach is to automatically identify an organization's tasks that deal with personal information as defined by the GDPR. This is an essential step in organizations' effort to comply with data protection regulations since possible sources of sensitive information must be identified to be able to apply adequate security measures. There are tasks where this is a simple problem that could be solved by only analyzing the task name, e.g. "Receive Customer Data" (Technical University of Berlin). However, many tasks in business processes have generic names, e.g. "Process Data" (Private Companies). To be able to accurately classify these tasks, it is necessary to include more context of the business process. Ideas for including additional information could be to analyze the data object of a task. For instance, if the data object is a "medical file" (Eindhoven University of Technology) then the task should be classified as GDPR-critical. If the data object is a "job description" (BPMN Handbook) on the other hand, the task should be classified as GDPR-uncritical. Another idea is to analyze the resource that executes the task. If the task "Process Data" is executed by the "Engineering Department" (Humboldt University of Berlin), it is unlikely that it deals with personal data. However, if the same task is executed by the "Customer Service" (Technical University of Berlin), the data referred to in the task name likely contains personal information.

4.2 Linguistic Preprocessing of Textual Business Process Descriptions

The linguistic preprocessing step is based on the work of Friedrich et al. [FMP11] excluding the actual creation of business process models. This step takes a textual business process description as an input and extracts verbs, objects, roles, and data items for each identified task. To achieve this, it uses NLP techniques such as the Stanford Parser [KM03] and VerbNet [Sc05]. The first step in the linguistic preprocessing is the application of the Stanford Parser that detects the part of speech of each word and the grammatical relations between words, also called part-of-speech tagging. It takes a sentence in natural language as input and identifies verbs as well as subjects and objects to which these verbs relate. The approach of Friedrich et al. [FMP11] uses the output of the tagging to automatically extract tasks consisting of a verb, an object, the executing role, and data items used.

4.3 Computation of Features

Given the ideas formulated in a prior section, we propose the following features for our model.

- **Data item (categorical).** The *data item* feature is categorical and relates to the name of the data item used in a task in the business process. We use data item as a synonym of the term data object defined in BPMN. Choosing the data items as a feature in our classification is a logical choice since we are interested in the data processed in each task. The rationale behind this feature is that there might be data items that almost always contain GDPR-critical data. Examples of such data items are “customer data” or “patient file”. But there also might be data items like “job description” that are likely to contain uncritical data since such documents tend to be publicized by the organization.
- **Customer relation (related/unrelated).** The *customer relation* feature is a binary feature that reveals whether the respective task has any relation to a customer. This is important because it allows the classification of tasks with ambiguous names. For instance, it is hardly possible to correctly classify a task called “enter data”, when only considering the name. However, if it can be determined that the data is from customers, it is possible to classify the task as GDPR-critical. Such a determination could be made by examining the context of each task. For instance, when the resource executing the task is the customer service, the processed data is likely from customers.
- **Object (categorical).** The *object* feature is categorical and relates to the grammatical object used in the name of a task. The idea behind this feature is that there are objects likely associated with the type of data used in a task. For instance, tasks with the objects “order” or “invoice” have an increased probability to deal with personal data. On the other hand, it is unlikely that tasks with certain objects, such as physical products (e.g., bicycles), deal with personal data.

The *data item* and *object* features are obtained during the linguistic preprocessing step while the *customer relation* feature is obtained through further analysis. The computation is done by analyzing the natural language used in the identified tasks and the relevant sentence in the process description. Should they contain the word *customer* or a synonym (e.g. *patient*), the *customer relation* feature has the value: *related*.

4.4 Classification of Tasks

Finally, we use a model to classify tasks from unseen business process descriptions as either GDPR-critical or GDPR-uncritical. Such a classification is not a trivial task as the context of each task varies. For instance, the task “send data” can hardly be correctly classified using a single feature. It is therefore necessary to combine the features from the previous chapter and apply a SML algorithm. A SML algorithm analyzes the training data and infers a function that is subsequently used to map new observations [Ru10]. As we are dealing with categorical variables where no ordinal relationship exists, we apply one-hot encoding. We then use a linear SVM [CV95] since it can deal with a small dataset, has a low risk of overfitting, and scales well. Because of this, SVMs are often used in text categorization problems to classify documents into predefined categories because they perform well in this particular set of problems [GDS19]. Other authors used SVMs for very similar text classification problems [Jo05], [LvR18], [TK01]. As input for our SVM, we provide manually labeled tasks as well as their computed features to train the model.

5 Dataset

We use the dataset of textual process descriptions introduced by Friedrich et al. [FMP11] with two differences. Firstly, we do not use all 47 process descriptions from the original dataset since 14 process descriptions from one source have a low language quality. Secondly, we expanded the dataset by one source of four textual business process descriptions from two companies. All of the 509 tasks were classified as either GDPR-critical or GDPR-uncritical by two researchers. If the classifications differed, a third researcher resolved conflicting labels. Some of these classifications were only possible by taking the whole business process description into account. For example, a task called “store information” was classified as GDPR-critical because the term “information” referred to the health information of a hospital patient. While most tasks were classified with confidence, there were some tasks where the classification of the first and second researcher differed. For instance, in a business process about receiving customer data, there was disagreement about a task called “initiate measures” that is executed in the case of unusual occurrences. One researcher argued that the mere initiation of measures does not involve personal data. The other researcher thought that since these measures are linked to customer data and cannot be initiated without directly referring to this customer data, the task should be classified as GDPR-critical. The third researcher agreed with the second researcher and the task was classified as GDPR-critical.

Table 1 shows the characteristics of the dataset. It should be noted that the identified tasks in Table 1 are not the direct result of the software tool by Friedrich et al. [FMP11]. Some of the identified tasks only consisted of a single verb or did not contain work to be performed and therefore violated the definition of an activity. Such tasks were excluded from the dataset. However, due to this exclusion, there is some subjectivity regarding the number of identified tasks. The textual business process descriptions of each source differ in several ways. Especially the length of the descriptions and sentences distinguishes the sources from each other. Another important difference is that the amount of GDPR-critical tasks varies greatly in the sources. This is due to the variety of business processes in the dataset. For instance, creating a job description (BPMN handbook) does not deal with personal data. The opposite is the reception of customer data (private companies) where almost all tasks explicitly deal with personal information about customers.

Type	Sources	D	S	SL	CT	UT
Academic	Humboldt University of Berlin	4	10.0	18.1	6	51
	Technical University of Berlin	2	34.0	21.2	37	37
	Queensland University of Technology	8	6.1	18.3	38	30
	Eindhoven University of Technology	1	40.0	18.5	23	21
Textbook	BPMN Handbook	3	4.7	17.0	3	14
	BPMN Guide	6	7.0	20.8	14	43
Industry	Vendor Tutorials	4	9.0	18.2	19	24
	inubit AG	4	11.5	18.4	15	40
	BPM Practitioners	1	7.0	9.7	6	1
	Private Companies	4	26.8	25.7	63	24
Total		37	15.6	18.6	224	285

Tab. 1: Details about the Dataset.

Legend: D = Number of textual business process descriptions, S = Average number of sentences, SL = Average sentence length in words, CT = Number of GDPR-critical tasks, UT = Number of GDPR-uncritical tasks, BPMN = Business Process Model and Notation, BPM = Business Process Management.

6 Evaluation

6.1 Description of the Evaluation

In a former section, we described the proposed approach on a conceptual level. In this section, we will evaluate this concept using our dataset by testing whether it can reliably predict the GDPR-criticality of tasks in unseen textual business process descriptions. The approach was implemented in Java using the prototype of Friedrich et al. [FMP11] and

the machine learning library Weka [Ha09]. We conducted a 5-fold cross-validation [HTF09] with a 60/20/20 split in training, validation, and test data with our dataset. Although we present our results by showing correctly and incorrectly classified tasks, the cross-validation is done on the business process descriptions to ensure that tasks from the same business process cannot be in the training data and the test data at the same time since this would lead to information leakage. To evaluate our approach as well as the proposed features, we used three configurations. Firstly, we trained only on the *data item* feature. Secondly, we trained on the *data item* and *customer relation* features. Thirdly, we trained on the *data item*, *customer relation* and *object* features. The quality of the configurations is measured using of precision, recall, and F1-score.

6.2 Results

Class	Metric	Data Item (DI)	DI & Customer Relation (CR)	DI & CR & Object
Tasks	Correct	63	86	86
	Incorrect	43	20	20
GDPR-critical	Precision	0.90	0.83	0.84
	Recall	0.29	0.83	0.81
	F1-Measure	0.44	0.83	0.83
GDPR-uncritical	Precision	0.53	0.79	0.78
	Recall	0.96	0.79	0.81
	F1-Measure	0.68	0.79	0.80
Total	Precision	0.73	0.81	0.81
	Recall	0.59	0.81	0.81
	F1-Measure	0.55	0.81	0.81

Tab. 2: Results of the 5-fold cross validation.

The detailed results of our 5-fold cross-validation with a 60/20/20 split in training, validation, and test data are demonstrated in Table 2. It shows the precision, recall, and F1-measure for both classes as well as the total amount of correctly and incorrectly classified tasks. The columns differ in the features used for each model. The left column shows the results when only using the *data item* feature, while the right column demonstrates the results of using all three features in a model. The results show that the discriminating power of the *data item* feature is not very high with an overall F1-measure of 0.55. Still, the precision regarding GDPR-critical tasks is high with a value of 0.90. When only using this *data item* feature, many factors are not taken into account. Some of these factors are considered when using a combination of the *data item* and the *customer relation* features. This addition improves the performance drastically, leading to an F1-measure of 0.81. The addition of the *object* feature to the two prior features leads only to a slight improvement regarding the F1-score of the GDPR-uncritical class. A model using

the *data item*, *customer relation*, and *object* features results in a F1-measure of 0.81. Figure 2 shows the receiver operating characteristic (ROC) curves for the full configuration (*data item*, *customer relation*, *object*). The curves are based on only three data points due to the calculation in Weka but they still provide valuable information about the classifier. More precisely, the curve illustrates the quality of a binary classifier by representing the true positive rate and false positive rate. The curve of a random classifier would result in the point (0.5, 0.5) creating a diagonal from (0, 0) to (1, 1) [Fa06]. The points of the two classes are (0.18, 0.81) for GDPR-critical and (0.19, 0.81) for GDPR-uncritical. This means that our classifier performs satisfactorily.

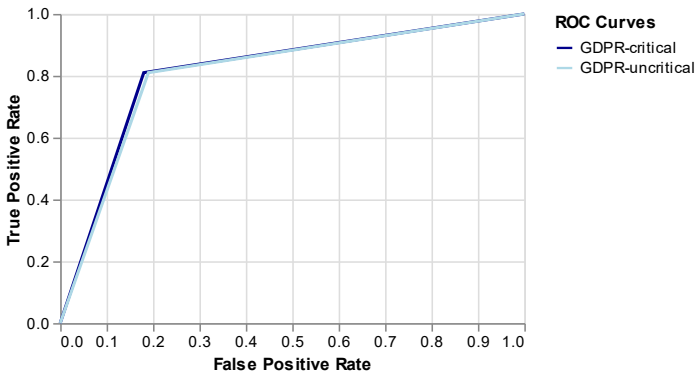


Fig. 2: ROC curves for both classes

When analyzing the errors of the classification, we observe that most errors were caused by the same feature values being used differently. For instance, tasks with the *object* “information” tend to be misclassified because the term information is vague and is used with different meanings. In one process, the task “enter information” refers to parts and quantities necessary to create a product. In another process, the task “record information” refers to the health status of a patient. The proposed *customer relation* feature can remedy some of these errors as it gives additional insight about the *object* feature “information” since the information that is mentioned in direct contact with a customer or patient is more likely to contain data about the respective customer or patient. Still, there are cases without direct relation to the customer where personal data is dealt with and where ambiguous terms are used. For example, this could be the case when customer data is processed by internal departments after it was obtained from the customer service. Another similar type of error is caused by polysemes, for example in tasks with the *object* value “order”. In the task “receive order”, the word order refers to data sent by a customer including information about the customer, such as the name, address, and other. In the task “give order” in the context of a hotel, the word order refers to an order to a sommelier to fetch wine. These polysemes seem to be a problem for our approach since they lead to the same feature value but a different GDPR-criticality. Another cause of errors is the lack of training data for

some feature values. Tasks with the *data item* feature value “customer data” were never misclassified since there are several tasks with this feature value. A task with the *data item* feature value “treatment plan” was misclassified although a medical treatment plan will always contain personal data. The problem is that the task “formulate treatment plan” only occurred once in the dataset, which means that it can be misclassified when in the test set. In summary, the first main cause of errors is ambiguity in features leading to feature values with different usage. The second main cause of errors is a lack of training data for some values. It should be noted that Leopold et al. [LvR18] faced similar types of misclassifications in their research. Although these misclassifications exist, we consider the results of the evaluation to be positive. The proposed approach generated satisfactory results in identifying GDPR-critical tasks in textual business process descriptions.

7 Conclusion and Next Steps of Research

The goal of our ongoing research is to create an approach that automatically identifies GDPR-critical tasks in textual business process descriptions. In this paper, we proposed a conceptual approach and conducted a first evaluation. The classification of unseen textual business process descriptions using the proposed features achieved an overall F1-measure of 0.81 meaning that the approach generated satisfactory results.

Although the results of our evaluation were positive, there are limitations to our approach that need to be considered. Firstly, our dataset is not representative. Textual business process descriptions in practice might differ from the ones in our dataset. Depending on the degree of such differences, this could make a correct classification difficult. However, by choosing a dataset that includes different types of descriptions from various sources, we attempted to maximize the validity of our evaluation. Secondly, dealing with tasks from business processes as a source for categorical features poses challenges. Unseen tasks from textual business process descriptions often contain previously unobserved objects and data items since activities and wording in business processes are heterogeneous. As discussed in the previous chapter, this can lead to misclassifications. This can be improved by using other NLP methods, such as word embeddings or language models. Nevertheless, our evaluation has shown that despite these limitations our model can predict the GDPR-criticality of tasks from unseen textual business process descriptions of different sources with satisfactory results. Lastly, in practice, the highest recall of 0.83 for GDPR-critical tasks is not high enough for a fully automatic identification since fines for not complying with the GDPR can be steep. Therefore, our approach cannot be seen as a substitution for a compliance officer when identifying GDPR-critical tasks. Nevertheless, it provides valuable insights and can propose a classification of tasks to support a compliance officer when making the final decision. This means that it can play a role similar to machine learning applications in medicine as described by Forsting [Fo17] for example, where physicians are supported by machine learning models but make the final decision since they are the ones responsible. At the very least, it can give an overview of which business processes often deal with personal data.

During the next steps of our ongoing research, we plan to conduct a second iteration of our DSR project. In this iteration, we will improve the approach by applying a large language model and test it on different datasets. This would improve our model regarding the limitations of the unrepresentative dataset as well as of the heterogeneity of the feature values extracted from business processes. Also, we plan to improve our model by testing other features. After these steps, a summative evaluation of the approach can be made [VPB16]. In future work, we plan to apply similar approaches to different types of sensitive data to generate more insights from textual business process descriptions in this context.

Acknowledgements

The project on which this study is based was funded by the German Federal Ministry of Education and Research under grant number 16KIS1331. The responsibility for the content of this publication lies with the authors.

References

- [Ac21] Accenture: State of Cybersecurity Resilience 2021. <https://www.accenture.com/content/dam/accenture/final/a-com-migration/custom/us-en/invest-cyber-resilience/pdf/Accenture-State-Of-Cybersecurity-2021.pdf>, accessed 8 Jun 2023.
- [BRU00] Becker, J.; Rosemann, M.; Uthmann, C. von: Guidelines of Business Process Modeling: Business Process Management. Springer, Berlin, Heidelberg, pp. 30–49, 2000.
- [Ca19] Caramujo, J. et al.: RSL-IL4Privacy: a domain-specific language for the rigorous specification of privacy policies. *Requirements Engineering* 1/24, pp. 1–26, 2019.
- [CCR04] Cavusoglu, H.; Cavusoglu, H.; Raghunathan, S.: Economics of IT Security Management: Four Improvements to Current Security Practices. *Communications of the Association for Information Systems* 14, 2004.
- [CV95] Cortes, C.; Vapnik, V.: Support-vector networks. *Machine Learning* 3/20, pp. 273–297, 1995.
- [Fa06] Fawcett, T.: An introduction to ROC analysis. *Pattern recognition letters* 8/27, pp. 861–874, 2006.
- [FMP11] Friedrich, F.; Mendling, J.; Puhmann, F.: Process model generation from natural language text: Advanced Information Systems Engineering: 23rd International Conference, CAiSE 2011, London, UK, June 20–24, 2011. *Proceedings* 23, pp. 482–496, 2011.

- [Fo17] Forsting, M.: Machine learning will change medicine. *Journal of Nuclear Medicine* 3/58, pp. 357–358, 2017.
- [GDS19] Ghosh, S.; Dasgupta, A.; Swetapadma, A.: A study on support vector machine based linear and non-linear pattern classification: 2019 International Conference on Intelligent Sustainable Systems (ICISS), pp. 24–28, 2019.
- [Ha09] Hall, M. et al.: The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter* 1/11, pp. 10–18, 2009.
- [HM15] Hirschberg, J.; Manning, C. D.: Advances in natural language processing. *Science* 6245/349, pp. 261–266, 2015.
- [HTF09] Hastie, T.; Tibshirani, R.; Friedman, J. H.: *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009.
- [Jo05] Joachims, T.: Text categorization with support vector machines: Learning with many relevant features: *Machine Learning: ECML-98: 10th European Conference on Machine Learning Chemnitz, Germany, April 21-23, 1998 Proceedings*, pp. 137–142, 2005.
- [KM03] Klein, D.; Manning, C. D.: Accurate unlexicalized parsing: *Proceedings of the 41st annual meeting of the association for computational linguistics*, pp. 423–430, 2003.
- [KS14] Kumar, N. N.; Shyamasundar, R. K.: Realizing purpose-based privacy policies succinctly via information-flow labels: *2014 IEEE Fourth International Conference on Big Data and Cloud Computing*, pp. 753–760, 2014.
- [Kü21] Kühnel, S. et al.: *Towards a Business Process-based Economic Evaluation and Selection of IT Security Measures*, 2021.
- [La15] La Rosa, M.: Strategic business process management: *Proceedings of the 2015 International Conference on Software and System Process*, pp. 177–178, 2015.
- [Li11] Liu, B.: *Supervised learning*. Springer, 2011.
- [LvR18] Leopold, H.; van der Aa, H.; Reijers, H. A.: Identifying candidate tasks for robotic process automation in textual process descriptions: *Enterprise, Business-Process and Information Systems Modeling: 19th International Conference, BPMDS 2018, 23rd International Conference, EMMSAD 2018, Held at CAiSE 2018, Tallinn, Estonia, June 11-12, 2018, Proceedings* 19, pp. 67–81, 2018.
- [Pu19] Pullonen, P. et al.: Privacy-enhanced BPMN: enabling data privacy analysis in business processes models. *Software and Systems Modeling* 18, pp. 3235–3264, 2019.
- [Ra12] Ramezani, E. et al.: Separating compliance management and business process management: *Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part II* 9, pp. 459–464, 2012.
- [RLT22] Rahat, T. A.; Long, M.; Tian, Y.: Is Your Policy Compliant? A Deep Learning-based Empirical Study of Privacy Policies' Compliance with GDPR: *Proceedings of the 21st Workshop on Privacy in the Electronic Society*, pp. 89–102, 2022.

- [RMv09] Reijers, H. A.; Mans, R. S.; van der Toorn, R. A.: Improved model management with aggregated business process models. *Data & Knowledge Engineering* 2/68, pp. 221–243, 2009.
- [RRM21] Rizun, N.; Revina, A.; Meister, V. G.: Analyzing content of tasks in Business Process Management. Blending task execution and organization perspectives. *Computers in Industry* 130, p. 103463, 2021.
- [Ru10] Russell, S. J.: *Artificial intelligence a modern approach*. Pearson Education, Inc, 2010.
- [Sc05] Schuler, K. K.: *VerbNet: A broad-coverage, comprehensive verb lexicon*. University of Pennsylvania, 2005.
- [Se20] Serrado, J. et al.: Information security frameworks for assisting GDPR compliance in banking industry. *Digital Policy, Regulation and Governance* 3/22, pp. 227–244, 2020.
- [SKS18] Sackmann, S.; Kuehnle, S.; Seyffarth, T.: Using business process compliance approaches for compliance management with regard to digitization: evidence from a systematic literature review: *Business Process Management: 16th International Conference, BPM 2018, Sydney, NSW, Australia, September 9-14, 2018, Proceedings* 16, pp. 409–425, 2018.
- [TK01] Tong, S.; Koller, D.: Support vector machine active learning with applications to text classification. *Journal of machine learning research* Nov/2, pp. 45–66, 2001.
- [To20] Torre, D. et al.: An ai-assisted approach for checking the completeness of privacy policies against gdpr: 2020 IEEE 28th International Requirements Engineering Conference (RE), pp. 136–146, 2020.
- [VK15] Vaishnavi, V. K.; Kuechler, W.: *Design science research methods and patterns: innovating information and communication technology*. Crc Press, 2015.
- [vLR17] van der Aa, H.; Leopold, H.; Reijers, H. A.: Comparing textual descriptions to process models-the automatic detection of inconsistencies. *Information Systems* 64, pp. 447–460, 2017.
- [VPB16] Venable, J.; Pries-Heje, J.; Baskerville, R.: FEDS: a framework for evaluation in design science research. *European journal of information systems* 25, pp. 77–89, 2016.
- [We19] Weske, M.: *Business process management. Concepts, languages, architectures*. Springer, Berlin, Heidelberg, 2019.

Verhalten und Ausführungssemantik optionaler Kanten in Geschäftsprozessen

Thomas Bauer¹


Abstract: Für viele Geschäftsprozesse (GP) ist erforderlich, dass das zugrundeliegende Prozess-Management-System (PMS) eine große Flexibilität bei der GP-Ausführung ermöglicht. Ein in der Literatur vorgeschlagener Ansatz, um die Flexibilität zu erhöhen, sind optionale Kanten. Diese geben eine gewünschte Ausführungsreihenfolge von Aktivitäten vor, die von den Benutzern in Ausnahmefällen aber auch ignoriert werden kann. Allerdings ist bisher ungeklärt, welches Verhalten eines PMS und welche Ausführungsreihenfolgen zur Runtime der Prozesse in solchen Fällen exakt gewünscht sind. Dies wird in diesem Beitrag anhand von Prozessbeispielen aus der Praxis geklärt. Außerdem wird eine formale Ausführungssemantik für derartige GP entwickelt. Sie ist erforderlich, damit eine GP-Engine solche Prozesse zur Runtime tatsächlich ausführen kann.

Keywords: Flexibilität, Kontrollfluss, Sequenz, Runtime, Prozess-Engine, Workflow-Engine.

1 Einleitung

Das generelle Ziel des Projektes CoPMoF (Controlable Pre-Modeled Flexibility) ist, die Flexibilität bei der Ausführung von Geschäftsprozessen (GP) zu erhöhen (d.h. zur Runtime), indem vorhersehbar benötigte Funktionalität bereits zu Buildtime „vor-modelliert“ wird. Hierfür wurden diverse Ansatzpunkte für den Kontrollfluss [Ba20, Ba21] und weitere Perspektiven [Ba19] von GP skizziert, wie z.B. die Verwendung optionaler Kontrollflusskanten. Solche Kanten geben die gewünschte Ausführungsreihenfolge von Aktivitäten vor; diese muss von den Benutzern aber nicht unbedingt eingehalten werden. Bestimmte, eigentlich noch nicht zur Ausführung anstehende, Aktivitäten werden ihnen deshalb bereits in ihren Arbeitslisten angeboten, allerdings mit einem entsprechenden Warnhinweis. Dadurch sind sie in Ausnahmefällen vorzeitig bearbeitbar. Allerdings wird ein entsprechendes Verhalten in [Ba20, Ba21, Ba23] lediglich angedeutet. Es existieren jedoch noch keine Arbeiten, die das gewünschte exakte Verhalten eines GP mit optionalen Kanten zur Runtime klären. Deshalb wird im Folgenden die offene Forschungsfrage betrachtet, welche genaue Bedeutung optionale Kanten haben sollen und wie die Ausführungsregeln einer GP-Engine hierfür gestaltet sein können.

Optionale Kanten ermöglichen die frühere Bearbeitung von Aktivitäten. Es werden aber weiterhin alle Aktivitäten ausgeführt (d.h. es handelt sich nicht um optionale Aktivitäten [Ba21]). Die parallele Modellierung (d.h. nach einem AND-Split) der mit optionalen Kanten verbundenen Aktivitäten würde zwar alle gewünschten Ausführungsreihenfolgen

¹ Hochschule Neu-Ulm, Fakultät Informationsmanagement, Wileystr. 1, 89231 Neu-Ulm,
thomas.bauer@hnu.de,  <https://orcid.org/0000-0001-8360-8430>

ermöglichen, jedoch führt dies zu einem unübersichtlichen GP-Modell und die „bevorzugte“ Ausführungsreihenfolge wäre für die Endbenutzer nicht mehr erkennbar.

Ziel dieses Beitrags ist, ein sinnvolles Verhalten und eine formale Ausführungssemantik für solche GP zu entwickeln. Nach einer Literaturdiskussion (Abschnitt 2) werden hierzu in Abschnitt 3 relevante Beispielprozesse und Szenarien vorgestellt. Mittels dieser wird untersucht, welche Art des Verhaltens zur Runtime sinnvoll ist. In Abschnitt 4 wird hierfür eine formale (operationale) Ausführungssemantik definiert. Eine solche ist unbedingt erforderlich, damit Prozess-Engines entsprechende GP zur Runtime steuern können. Die resultierende Ausführungssemantik ist ähnlich zu der von BPMN 2.0 [OMG11], jedoch werden zusätzliche Ausführungszustände für Aktivitäteninstanzen und zusätzliche bzw. veränderte Ausführungsregeln benötigt. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick.

2 Stand der Forschung

Beim Case Handling [AWG05], dem CMMN-Standard und Constraint-basierten Ansätzen [APS09, Pe07, RW12] wird kein Kontrollflussgraph vorgegeben. So wird sogar eine noch größere Flexibilität als mit dynamischen Änderungen [RD98] ermöglicht, weil die Benutzer auszuführende Aktivitäten recht frei wählen können. Unser Ziel ist hingegen eine höhere Prozesssicherheit (und damit weniger Flexibilität), indem erlaubte Abweichungen vor-modelliert werden (und damit geprüft und genehmigt werden können).

Außer den bereits angesprochenen eigenen Vorarbeiten des Projekts CoPMoF enthält die wissenschaftliche Literatur kaum Erwähnungen optionaler Sequenzkanten. Lediglich im Projekt CrossFlow [Gr00, Kl00] wird ein entsprechender Kantentyp vorgeschlagen und als Optional Execution Order bezeichnet: Um in Ausnahmesituationen z.B. eine vorgegebene Ausführungsdauer erreichen zu können, wird bei diesem Kantentyp von der eigentlich vorgegebenen Reihenfolge abgewichen. Durch eine parallele Ausführung von Aktivitäten kann dadurch die gewünschte maximale Prozessdurchlaufzeit eingehalten werden. In den Kontrollfluss-Pattern [RH06], dem BPMN-Standard [OMG11], sowie in sonstigen auf GP-Graphen basierenden Ansätzen, werden optionale Sequenzkanten nicht betrachtet. Deswegen wird im Folgenden auf optionale Reihenfolgebeziehungen in Ansätzen eingegangen, die keine Kanten und GP-Graphen im klassischen Sinn verwenden.

Bei Constraint-basierte Ansätzen (vgl. [APS09, Pe07, RW12]) werden die erlaubten Ausführungsreihenfolgen mittels Constraints (Regeln) festgelegt. Durch die Verwendung eines optionalen Constraints (Recommendation) kann das Verhalten einer optionalen Sequenzkante erreicht werden. Jedoch erfolgt die Steuerung einer Prozessinstanz, d.h. die Berechnung der aktuell ausführbaren Aktivitäten, unmittelbar durch die Constraints. Deshalb ist bei solchen Ansätzen keine separate Ausführungssemantik erforderlich. Dementsprechend wird in Arbeiten zu optionalen Constraints nichts mit der nachfolgend

vorgestellten Ausführungssemantik Vergleichbares entwickelt. Ebenso wird nicht diskutiert, welche Arten des Verhaltens in komplexeren Szenarien erforderlich sein können (vgl. Abschnitt 3.1 sowie den in Abschnitt 3.2 eingeführten Typ „Soft“).

Beim Case Handling [AWG05] ergibt sich der Status einer Prozessinstanz aus den Inhalten von Datenobjekten, d.h. diese definieren, welche Aktivitäten aktuell ausführbar sind. Bei DOPA [Li06] handelt es sich um solch einen datengetriebenen Ansatz, bei dem durch zusätzliche Regeln sichergestellt werden kann, dass nur tatsächlich erwünschte Ausführungsreihenfolgen möglich sind. Die Regeln definieren also einen Teil des Kontrollflusses. Da eine solche Regel bei DOPA optional sein kann, entspricht dies einer optionalen Sequenzkante. Weil sich die tatsächliche Ausführungsreihenfolge ausschließlich aus Dateninhalten und diesen Regeln ergibt, ist auch bei solchen Ansätzen eine zusätzliche formale Ausführungssemantik nicht erforderlich

Zusammenfassend lässt sich feststellen, dass keine Arbeiten existieren, welche die gewünschten Arten des Verhaltens optionaler Kanten diskutieren oder eine formale Ausführungssemantik hierfür definieren. Damit GP-Modelle mit solchen Kantentypen von einer GP-Engine tatsächlich gesteuert werden können, ist es erforderlich, diese Forschungslücke zu schließen. Dies schafft zudem eine Grundlage, um GP-Modellierungssprachen wie BPMN [OMG11] um optionale Kanten erweitern zu können.

3 Nutzen und Verhalten optionaler Kanten

Im Folgenden wird der Verwendungszweck von optionalen Sequenzkanten erläutert. Außerdem wird geklärt, welches Verhalten für solche Kanten zur Ausführungszeit gewünscht ist. Diese Klärung bildet die Basis für die Entwicklung einer Ausführungssemantik in Abschnitt 4.

3.1 Sequenzen optionaler Kanten

Folgendes Klinikbeispiel dient dazu, den Nutzen optionaler Kanten zu veranschaulichen: Bei bestimmten Diagnosen (Akt. a in Abb. 1a) wird bei einem Patienten üblicherweise (entsprechend dem Klinikstandard) zuerst ein EKG erstellt (Akt. b), danach ein Röntgenbild (Akt. c) und dann wird die Akt. d (MRT-Aufnahme) ausgeführt. Nach Abschluss dieser Untersuchungen wird dem Patienten eine Behandlung vorgeschlagen (Akt. e). Falls aber eine der Untersuchungseinrichtungen aktuell nicht verfügbar oder überlastet ist, kann von dieser Standardreihenfolge abgewichen werden: Nach Beendigung der Akt. a befinden sich die Untersuchungsaktivitäten b bis d alle in den Arbeitslisten der Benutzer. Jedoch ist ausschließlich die Akt. b „regulär startbar“ (d.h. ohne auszunutzen, dass eine Kante optional ist). Die auf Akt. b mit optionalen Kanten folgenden Aktivitäten c und d sind z.B. für den Patienten Fritz Müller in den Arbeitslisten von Peter Meier und von Anne Schmidt (Abb. 1b und c) aber bereits enthalten, jedoch als „eigentlich noch nicht auszuführen“ gekennzeichnet. Sind z.B. aktuell alle EKG-Geräte belegt und der Patient wird

deshalb direkt zum MRT geschickt, so kann die dortige Mitarbeiterin Anne Schmidt die Akt. d bereits problemlos ausführen.² Die Akt. b und c finden dann später statt. Wird zuerst die Akt. b ausgeführt, so verschwindet der Warnhinweis aus der Arbeitsliste von Peter Meier (Abb. 1b), weil die Akt. c nun regulär bearbeitbar ist.³

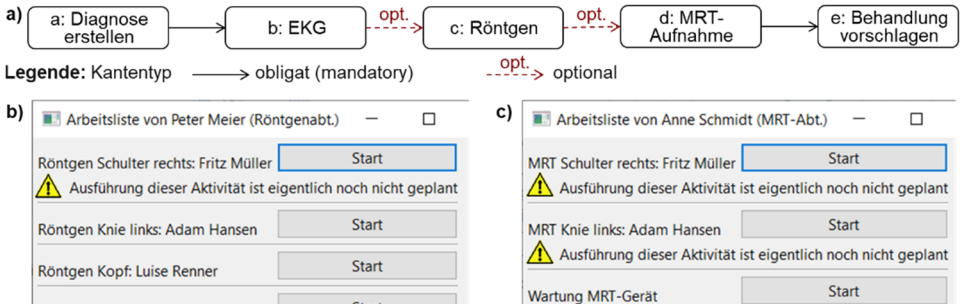


Abb. 1: a) Untersuchungsprozess sowie b) Arbeitsliste für die Akt. c und c) für Akt. d

Die zur Akt. e „Behandlung vorschlagen“ führende Kante ist nicht optional (sondern obligat, engl.: mandatory). Deswegen darf die Akt. e wirklich erst nach Beendigung der Akt. d ausgeführt werden. Allerdings bleibt folgende Frage zu klären: Falls die Akt. d vorzeitig gestartet und auch bereits beendet wurde, muss Akt. e dann dennoch auf die Beendigung der Aktivitäten b und c warten? Dies ist in diesem Beispiel sinnvoll, weil eine konkrete Behandlung erst dann vorgeschlagen werden sollte, wenn alle Untersuchungen durchgeführt wurden. Das bedeutet, die Kante von Akt. d zu e mit dem „Strictness-Typ“ mandatory „synchronisiert“ die Ausführung aller vorherigen Aktivitäten, so dass ab Akt. e nur noch eine einzige Aktivität dieser „Sequenz“ bearbeitbar ist, sich also in den Arbeitslisten befindet oder aktuell ausgeführt wird (so wie bei klassischen Sequenzen). Wie nachfolgend dargestellt, gibt es aber auch Szenarien, in denen ein anderes Verhalten wünschenswert ist.

3.2 Strictness-Typ „Soft“

Bei dem in Abb. 2 dargestellten Ausschnitt eines GP zum Testen eines Bauteils ist solch ein abweichendes Verhalten gewünscht. Nach dem (nicht dargestellten) Bestellen eines Bauteils bei einem Zulieferer und dem Wareneingang, soll das eingetroffene Bauteil zuerst im Testing-IT-System verbucht werden (Akt. a). Danach wird es in Akt. b für den Test

² Beim vorliegenden Prozessmodell können die Aktivitäten b bis d prinzipiell auch überlappend ausgeführt werden. Dies wird aber wohl nicht passieren, weil der Patient gleichzeitig nur an einer Untersuchungseinrichtung anwesend sein kann. Deshalb werden für ihn nicht mehrere Aktivitäten (gleichzeitig) gestartet werden, bevor die anderen beendet sind. In anderen Szenarien kann eine solche parallele Bearbeitung jedoch erwünscht sein, um in Ausnahmesituationen Zeit zu sparen (z.B. die Aktivitäten a und b in Abb. 2).

³ Damit Arbeitslisten mit solcher Funktionalität realisiert werden können, muss über die Schnittstelle (API) der GP-Engine abfragbar sein, ob eine bestimmte Aktivität regulär oder vorzeitig bearbeitbar ist. Dies ist einfach realisierbar, weil dies direkt am Aktivitätenzustand (vgl. Abschnitt 4.1) Active (regulär) bzw. PreActive (vorzeitig) erkennbar ist.

vorbereitet (es werden z.B. Mess-Sensoren angebracht). Der eigentliche Test (Belastbarkeit, Korrosionsbeständigkeit, o.ä.) findet in Akt. c statt, bevor das Testergebnis im bereits erwähnten Testing-IT-System dokumentiert wird.

Um Entwicklungszeit einzusparen, kann mit der Bauteilvorbereitung begonnen werden, bevor die Verbuchung abgeschlossen wurde oder sogar bevor sie gestartet wurde. Deshalb befindet sich zwischen den Akt. a und b eine optionale Kante. Auch mit dem Test (Akt. c) kann vor Abschluss der Akt. a begonnen werden. Eine optionale Kante zwischen den Akt. b und c wäre jedoch nicht sinnvoll, weil mit dem Testen erst begonnen werden kann, wenn die Bauteil-Vorbereitung abgeschlossen ist. Eine „echte“ Mandatory-Kante hätte hier jedoch den Nachteil, dass die Akt. c dann auf die Beendigung der Akt. a warten müsste (so wie Akt. e auf die Untersuchungen in Abb. 1a). Dies ist jedoch nicht erforderlich und hierdurch würde unnötig Zeit verschwendet werden. Deshalb wird zwischen den Akt. b und c eine Kante mit dem Strictness-Typ soft verwendet.⁴ Dieser stellt einen „Mittelweg“ zwischen optional und mandatory mit folgendem Verhalten dar: Die Akt. c muss auf die Beendigung der unmittelbaren Vorgängerakt. b warten (der Startaktivität dieser Soft-Kante). Akt. c muss aber nicht auf die Beendigung von deren mit optionalen Kanten verbundenen Vorgängeraktivitäten (hier nur Akt. a) warten.

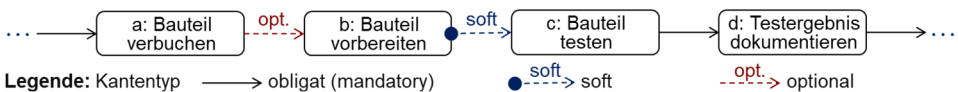


Abb. 2: GP zum Testen eines Bauteil mit dem zusätzlichen Strictness-Typ „Soft“

Die Dokumentation des Testergebnisses in Akt. d kann erst erfolgen, wenn dieses Bauteil durch Akt. a im Testing-IT-System erfasst wurde. Deshalb wird zwischen den Akt. c und d eine Kante mit Strictness-Typ mandatory verwendet. Dies führt (wie bereits bei Abb. 1a erläutert) dazu, dass vor dem Start der Akt. d alle Vorgängeraktivitäten beendet sein müssen.

Die Verwendung einer Kante mit Strictness-Typ soft ist nur nach optionalen Kanten und anderen Soft-Kanten relevant. Bei einer Kante mit Strictness-Typ mandatory müssen ohnehin alle Vorgängeraktivitäten beendet sein, so dass das Verhalten einer danach folgenden Soft-Kante identisch zu einer (klassischen) Mandatory-Kante wäre.⁵

⁴ Der ausgefüllte Kreis am Kantenbeginn deutet an, dass auf die Beendigung dieser Aktivität gewartet werden muss. Dass die Kante gestrichelt ist, symbolisiert ihre „Optionalität“, d.h. auf die Vorgängeraktivitäten von Akt. b muss nicht gewartet werden. Allerdings sind alle verwendeten Symbole nur ein erster Vorschlag und kein Forschungsergebnis dieses Beitrags.

⁵ Dies kann von einem GP-Modellierungswerkzeug ausgenutzt werden, indem dieser Strictness-Typ nur nach Aktivitäten mit solchen eingehenden Kanten angeboten wird. Dadurch muss der GP-Designer diesen Typ bei anderen Aktivitäten gar nicht berücksichtigen. Bei Aktivitäten, bei denen Soft-Kanten relevant sind, kann das Tool zudem erklären, welche Auswirkungen die Verwendung eines solchen Typs hat. Dies kann am konkret vorliegenden GP erfolgen. Im Beispiel aus Abb. 2 wird dann die Information angezeigt, dass Akt. c zwar auf die Beendigung von Akt. b warten muss, aber nicht auf Akt. a. Ein solches Verhalten vereinfacht die Verwendung der neuen Kantentypen.

Im Beispiel aus Abb. 2 ist der Zweck der Soft-Kante, dass Akt. c auf die Beendigung der Akt. b warten muss, aber nicht auf deren Vorgängerakt. a. Es ist auch modellierbar, dass auf die Beendigung mehrerer (aber wieder nicht aller) Vorgängeraktivitäten gewartet werden muss: Befindet sich vor einer Akt. c eine Sequenz von mit optionalen Kanten verbundenen Aktivitäten $a_1 \dots a_m$, $b_1 \dots b_n$ (in beliebiger Reihenfolge) und es soll nur auf die Beendigung der Aktivitäten $b_1 \dots b_n$ gewartet werden, dann kann dies wie in Abb. 3 dargestellt modelliert werden: Ausgehend von diesen Aktivitäten $b_1 \dots b_n$ werden Soft-Kanten modelliert, die alle in die Akt. c münden. Da alle ihre Vorgängeraktivitäten (irgendwann) tatsächlich ausgeführt werden, muss diese Akt. c eine „AND-Join-Semantik“ haben. Bei diese Modellierungsweise wird zwar auf die Aktivitäten $b_1 \dots b_n$ gewartet, nicht aber auf die Aktivitäten $a_1 \dots a_m$, von denen ausgehend keine Soft-Kante zur Akt. c führt.

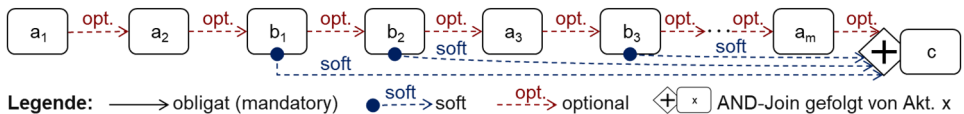


Abb. 3: Sequenz von Aktivitäten, bei denen Akt. c nur auf die Aktivitäten b_i warten muss

3.3 Kombination optionaler Kanten mit Gateways

Gateways (vgl. BPMN [OMG11]) dienen dazu, einen Pfad in mehrere aufzuspalten (Split) oder diese wieder zusammenzuführen (Join). Hierfür existieren einige grundlegende Gateway-Typen, die sich dadurch unterscheiden, dass genau ein Pfad (XOR), mehrere Pfade (OR) oder alle Pfade (AND) ausgeführt werden. Bei den ersten beiden Typen wird mittels Bedingungen festgelegt, welche Pfade ausgeführt werden sollen. Das ist für den Pfad i der Fall, wenn die zur ersten Kante dieses Pfades gehörende Bedingung (condition) $cond_i$ zum Wert $true$ evaluiert.

Das in Abb. 4 dargestellte Beispiel enthält einen OR-Split, so dass abhängig von den Bedingungen $cond_1$ bis $cond_5$ manche der Pfade mit den Aktivitäten b_i bis d_i ausgeführt werden. Dies erfolgt parallel, d.h. zeitlich überlappend. Betrachten wir nun das Verhalten am OR-Join: Hier wird in klassischen GP-Metamodellen auf die Beendigung aller ausgeführter Pfade gewartet. Da der Prozess aus Abb. 4 optionale Kanten enthält, soll jedoch nicht auf die Beendigung aller Pfade gewartet werden, bevor der OR-Join und damit auch Akt. e ausgeführt werden können. Stattdessen ist folgendes Verhalten gewünscht:

- Die Aktivitäten d_1 und d_2 müssen beendet sein, bevor der OR-Join und Akt. e ausgeführt werden, weil die Kanten zwischen diesen Aktivitäten und dem Join den Strictness-Typ **mandatory** haben. Da in die Akt. d_2 eine optionale Kante mündet, ist anzumerken, dass auch ihre Vorgängeraktivitäten (z.B. c_2) beendet sein müssen. Dies entspricht dem bereits in Abschnitt 3.1 definierten Verhalten.

- Die Kanten zwischen den Aktivitäten d3 bzw. d4 und dem Join sind optional. Deshalb kann die Akt. e (vorzeitig) gestartet werden, bevor Aktivitäten d3 und d4 beendet wurden.
Da in Akt. d4 eine optionale Kante mündet, muss auch ihre Vorgängerakt. c4 hierfür noch nicht beendet sein.
- Zwischen Akt. d5 und dem Join befindet sich eine Soft-Kante. Deshalb muss zwar die Akt. d5 selbst, aber nicht ihre (mittels einer optionalen Kante verbundene) Vorgängerakt. c5 beendet sein, um die Akt. e vorzeitig starten zu können (vgl. Abschnitt 3.2).⁶
- Der Strictness Typ der Kante zwischen dem Join und Akt. e wird ignoriert: Das Join-Gateway selbst führt zu keiner Verzögerung, weil es sofort und automatisch „durchschaltet“, sobald dies die Ausführungszustände seine Vorgängeraktivitäten d1 bis d5 erlauben. Deshalb ist irrelevant, ob Akt. e auf diesen Schaltvorgang wartet. Das gewünschte Verhalten bzgl. einer vorzeitigen Ausführbarkeit der Akt. e wird durch die in den Join eingehenden Kanten detailliert definiert (für jeden Pfad einzeln). Dieses soll nicht durch den Strictness-Typ der einzelnen aus dem Join ausgehenden Kante „überschrieben“ werden. Dies gilt generell für die Kante nach einem Join-Gateway und ebenso für in einen Split-Gateway mündende Kanten (in Abb. 4 die Kante von Akt. a zum OR-Split).⁷

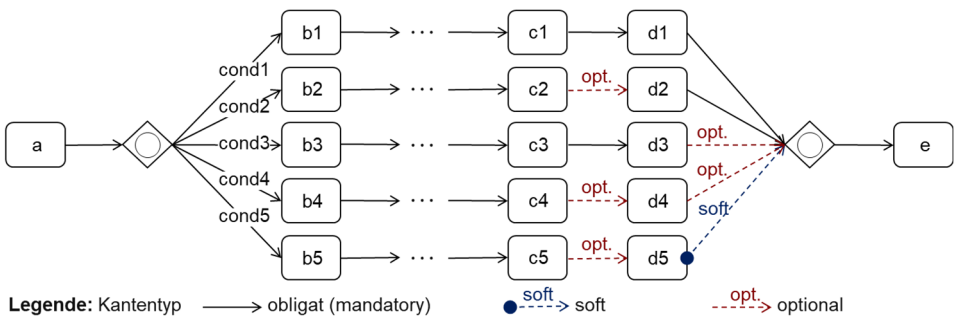


Abb. 4: OR-(Inklusives) Gateways mit Kanten verschiedenen Strictness-Typen

Der Gateway-Typ (XOR, OR, AND) legt ausschließlich fest, wie viele der Pfade ausgeführt werden. Er hat keinen Einfluss darauf, ob auf die Bearbeitung von Aktivitäten

⁶ Abb. 4 enthält kein analoges Beispiel mit einer in Akt. d5 eingehenden Mandatory-Kante, weil (wie in Abschnitt 3.2 erwähnt) auf diesem Strictness-Typ folgende Soft-Kanten dasselbe Verhalten haben, wie Mandatory-Kanten. Das Verhalten wäre also identisch zum oberen Pfad mit den Aktivitäten c1 und d1.

⁷ Da der Strictness-Typ dieser Kanten irrelevant ist, werden in den Abbildungen stets klassische Mandatory-Kanten verwendet. Dies hat keine Auswirkungen auf die in Abschnitt 4.2 entwickelte Semantik, weil diese Kanten in den Ausführungsregeln nicht verwendet werden (sondern nur die in den Join mündenden Kanten). Dass diese Kanten ignoriert werden, mag irritierend wirken. Alternativ könnte, so wie für Akt. c in Abb. 3 dargestellt, auch ein GP-Metamodell gewählt werden, bei der sich zwischen einem Join und der nachfolgenden Aktivität gar keine Kante befindet, sondern der Gateway-Typ (XOR, OR, AND) eine Markierung dieser Aktivität darstellt (vgl. ADEPT [RD98]). Da sich der BPMN-Standard aber gegen solch eine Notation entschieden hat, wird diese Darstellungsart im Folgenden nicht verwendet.

gewartet werden soll, weil in klassischen GP-Metamodellen auf die Beendigung aller ausgeführten Pfade gewartet wird. Deshalb ist dieses Verhalten bei anderen Gateway-Typen identisch zum eben erläuterten OR-Join. Ein analoges Verhalten ergibt sich bei verschachtelten Strukturen, z.B. wenn in einem Pfad nach einem OR-Split ein AND-Split folgt. Auch bei solch komplexen Strukturen entstehen durch den vorgestellten Ansatz keine prinzipiell neuen Problemstellungen, weil auch heutige Prozess-Management-Systeme (PMS) bereits analysieren können, welche Pfade bei dieser Prozessinstanz tatsächlich ausgeführt werden, so dass auf diese gewartet werden muss. Diese Fähigkeit von PMS nutzen die Def. 3 und 4 in Abschnitt 4.1.

Es ist nicht möglich, Soft-Kanten (vgl. Abb. 5a) durch Gateways zu ersetzen. Im GP aus Abb. 5b suggeriert die von Akt. a zum Join-Gateway verlaufende optionale Kante, dass nicht auf die Beendigung der Akt. a gewartet werden muss. Allerdings muss bei dem OR-Join, wegen der von Akt. b aus eingehenden Mandatory-Kante, auf die Beendigung aller Vorgängeraktivitäten gewartet werden (vgl. BPMN: lediglich überhaupt nicht ausgeführte Pfade werden von einem Join-Gateway ignoriert). In dem Beispiel aus Abb. 5b ist die Akt. a jedoch eine tatsächlich ausgeführte Vorgängeraktivität von Akt. b, so dass das Gateway auf diese wartet (trotz der Optionalität der Kante von Akt. a zum Gateway). Dies ist im Original-GP (Abb. 5a) jedoch nicht erwünscht. Auch die Verwendung einer optionalen Kante zwischen der Akt. b und dem OR-Join würde keine Lösung darstellen, weil dann nicht mehr auf die Beendigung der Akt. b gewartet wird (was jedoch erforderlich sein soll, vgl. Soft-Kante in Abb. 5a). Deshalb ist der Strictness-Typ *soft* tatsächlich erforderlich, um alle gewünschten Arten des Verhaltens realisieren zu können.

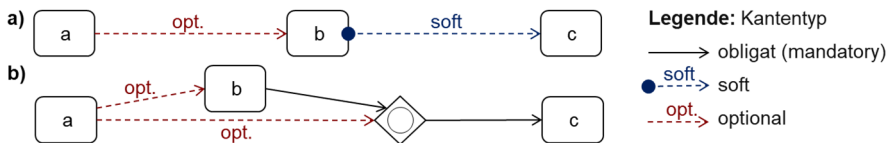


Abb. 5: Beispiel zur Erläuterung der Notwendigkeit von Soft-Kanten

4 Ausführungssemantik

Der nun folgende Unterabschnitt enthält einige grundlegenden Definitionen und es werden die bei optionalen Kanten zusätzlich erforderlichen Ausführungszustände für Aktivitäten eingeführt. Beides bildet die Basis, um im Abschnitt 4.2 eine formale Ausführungssemantik für optionale Kanten zu definieren. Hierzu werden Ausführungsregeln definiert, die auf Benutzeraktionen und Zuständen von Aktivitäten basieren.

Die alternativ mögliche Vorgehensweise, die Ausführungssemantik mittels eines bereits existierenden Formalismus (z.B. Petri-Netze) zu definieren, wurde nicht gewählt, weil der resultierende Ansatz unmittelbar zur Realisierung einer GP-Engine nutzbar sein soll. Existierende (z.B. kommerzielle) GP-Engines basieren häufig auf BPMN und den zugehörigen Ausführungsregeln, welche um die nachfolgend vorgestellten Regeln erweitert

werden können. Sie basieren jedoch nicht zwangsläufig auf Petri-Netzen o.ä., weshalb eine darauf aufbauende Ausführungssemantik nicht einfach zu integrieren wäre.

4.1 Definitionen und Ausführungszustände von Aktivitäten

Im Folgenden werden die für GP üblicherweise verwendeten (z.B. in [RD98, We19, OMG11]) Definitionen und die Menge der Ausführungszustände von Aktivitäteninstanzen so erweitert, dass sie als Grundlage für die Erstellung von Ausführungsregeln geeignet sind, welche die neu eingeführten Strictness-Typen berücksichtigen.

Def. 1: Ein Prozesstemplate $PT=(N, E)$ besteht aus einer Menge von Knoten N (engl.: Node) und einer Menge von Kontrollflusskanten E (Edge).

Eine Prozessinstanz $PI=(PT, State)$ besteht aus einem Prozesstemplate PT und einer Menge von Aktivitätszuständen (State). Jede Aktivität $a \in N$ kann einen unterschiedlichen Ausführungszustand $State(a)$ haben. Aktivitäten können die in Abb. 6 dargestellten Zustände einnehmen, d.h.:

$\forall a \in N$ gilt: $State(a) \in \{Inactive, Active, Running, Completed, PreActive, PreRunning, PreCompleted\}$.

Das PMS kennt den aktuellen Zustand jeder Aktivität und verändert diese Zustände während der Ausführung einer Prozessinstanz mittels vorgegebener Ausführungsregeln (siehe Abschnitt 4.2). Dadurch kann das PMS die Ausführungsreihenfolge der Aktivitäten steuern, Einträge in Arbeitslisten einfügen, Aktivitätenprogramme (z.B. Formulare) starten, automatisch ausgeführte Services aufrufen, etc.

Def. 2: Eine Kontrollflusskante $e \in E$ ist definiert als

$e = (SourceAct, TargetAct, Strictness)$ mit:

SourceAct: die Startaktivität der Kante, mit $SourceAct \in N$

TargetAct: die Zielaktivität der Kante, mit $TargetAct \in N$

Strictness: der Strictness-Typ der Kante, mit $Strictness \in \{optional, soft, mandatory\}$

Um diese Informationen in den Ausführungsregeln nutzen zu können, kann mit gleichnamigen Funktionen auf sie zugegriffen werden. So liefert z.B. $SourceAct(e)$ die Startaktivität der Kante e und $State(a)$ den aktuellen Ausführungszustand der Akt. a . Die nachfolgenden beiden Definitionen dienen dazu, die in Abschnitt 4.2 vorgestellten Regeln leichter verständlich formulieren zu können. Sie spezifizieren die relevanten Vorgängeraktivitäten einer Akt. a , wobei bei Def. 4 nur bestimmte Kantentypen berücksichtigt werden. Die Akt. a kann mehrere direkte Vorgängeraktivitäten V_{all} haben, wenn sich zwischen diesen und Akt. a ein Join-Gateway (AND, OR oder XOR) befindet. Wie in Abschnitt 3.3 bereits erläutert (vgl. Abb. 4), sind bei einem Join nur solche Vorgängeraktivitäten $v \in V$ mit $V \subseteq V_{all}$ relevant, die in Pfaden liegen, die aufgrund der Bedingung $cond$; überhaupt ausgeführt werden. Eine Analyse, welche Pfade tatsächlich ausgeführt werden, muss ein PMS auch bei klassischen GP-Metamodellen durchführen (vgl. BPMN 2.0 [OMG11]). Dieser Aspekt ist also unabhängig von den in diesem

Beitrag betrachteten Neuerungen, so dass dies in den nun folgenden Definitionen nur informell spezifiziert wird (Bedingung ii). Um die Lesbarkeit weiter zu erhöhen, werden in Definitionen 3 und 4 die Gateway-Knoten selbst ebenso ignoriert, wie die aus einem Join-Gateway ausgehende Kante⁸, weil deren Strictness-Typ ohnehin irrelevant ist (vgl. Abschnitt 3.3). So wird für die nachfolgenden Definitionen beispielsweise das Prozessmodell aus Abb. 4 so interpretiert, als wenn es direkte Kanten von den Aktivitäten d1 bis d5 zur Akt. e gäbe.

Def. 3: $\text{RelevantPredecessors}(a)$ beschreibt die Menge der direkten Vorgängeraktivitäten der Akt. a , die aufgrund der Gateway-Struktur des Prozesses und der Kantenbedingungen bei dieser Prozessinstanz tatsächlich ausführbar sind. Dabei beschreibt die Menge V_{all} alle Vorgängeraktivitäten, noch unabhängig von ihrer tatsächlichen Ausführbarkeit:

- i) $V_{\text{all}} = \{ V' \subseteq N \mid \forall v' \in V' \text{ gilt: } \exists e \in E \text{ mit } v' = \text{SourceAct}(e) \wedge a = \text{TargetAct}(e) \}$
- ii) $\text{RelevantPredecessors}(a) = \{ V \subseteq V_{\text{all}} \mid \forall v \in V \text{ gilt, dass } v \text{ bei dieser Prozessinstanz tatsächlich bereits ausgeführt wurde oder noch ausgeführt werden wird} \}$

Während mit der durch Def. 3 definierten Funktion alle relevanten (d.h. tatsächlich ausgeführten) Vorgängeraktivitäten ermittelt werden, berücksichtigt Def. 4 zusätzlich die Strictness-Typen. Hierzu wird ein bestimmter Typ s vorgegeben und es werden nur über entsprechende Kanten verbundene Aktivitäten berücksichtigt. So ergeben sich für das Beispiel aus Abb. 4 mit $\text{RelevantPredecessors}(e)$ alle ausführbaren Vorgängeraktivitäten d1 bis d5 der Akt. e, während z.B. mit $\text{RelevPredWithEdge}_{\text{optional}}(e)$ ausschließlich die Aktivitäten d3 und d4 ermittelt werden.

Def. 4: $\text{RelevPredWithEdge}_s(a)$ beschreibt die Menge V der direkten Vorgängeraktivitäten der Akt. a , wobei nur Kanten mit der vorgegebenen Strictness s berücksichtigt werden:

- i) $V_{\text{all}} = \{ V' \subseteq N \mid \forall v' \in V' \text{ gilt: } \exists e \in E \text{ mit } v' = \text{SourceAct}(e) \wedge a = \text{TargetAct}(e) \wedge \text{Strictness}(e) = s \}$
- ii) $\text{RelevPredWithEdge}_s(a)$ ist identisch zu $\text{RelevantPredecessors}(a)$ definiert (vgl. ii) in Def. 3)

Bei klassischen GP-Metamodellen wie BPMN 2.0 [OMG11] befinden sich beim Start einer Prozessinstanz alle Aktivitäteninstanzen im Zustand Inactive. Eine Ausnahme bilden die am GP-Anfang liegenden Aktivitäten: diese starten mit dem Zustand Active. Bei den anderen Aktivitäten wechselt der Zustand zu Active (vgl. 1 im oberen Zweig von Abb. 6), wenn die Bearbeitung ihrer Vorgängeraktivitäten abgeschlossen wurde. Dann wird eine solche Aktivität in die Arbeitslisten der Benutzer eingetragen und wechselt in den Zustand Running, sobald einer von diesen mit der Bearbeitung beginnt. Wird diese

⁸ Die zur Umformung in einen Graphen ohne Gateways erforderliche Transformation ist recht offensichtlich, so dass diese hier nicht detailliert beschrieben wird. Auf die Verwendung eines GP-Metamodells, das ganz auf Gateway-Knoten und daraus ausgehende Kanten verzichtet (vgl. Akt. c in Abb. 3 und ADEPT [RD98]), wird – wie bereits erwähnt – verzichtet, weil diese auch beim Standard BPMN existieren.

beendet, so findet ein Zustandswechsel zu Completed statt, so dass die Nachfolgeraktivitäten bearbeitbar werden, also selbst zu Active wechseln.

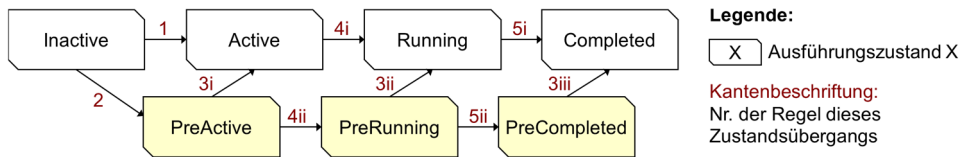


Abb. 6: Aktivitätszustände und Zustandsübergänge

Die Menge der Zustände einer Aktivitäteninstanz wurde gegenüber solchen klassischen Ansätzen erweitert, um die neuen Strictness-Typen realisieren zu können. Die zusätzlich erforderlichen Zustände sind in Abb. 6 farbig hinterlegt und ihre Bezeichnungen beginnen „Pre“. Diese Zustände haben die Bedeutung, dass die entsprechende Akt. a „vorzeitig gestartet“ wurde, also nur ausführbar ist, weil eine eingehende Kante den Strictness-Typ optional oder soft hat. Sobald ihre Vorgängeraktivitäten regulär beendet wurden, wäre die Akt. a auch regulär startbar, so dass sie in den gleichnamigen Zustand ohne „Pre“ wechselt (3i bis 3iii in Abb. 6, die Zahlen an den Zustandsübergängen verweisen auf die Nummer der zugehörigen Regel aus Abschnitt 4.2).

In diesem Beitrag steht das Verhalten für die neuen Strictness-Typen im Fokus. Deshalb werden keine Zustände oder Regeln betrachtet, die hierfür nicht relevant sind. Solche sind z.B. für einen Abbruch einer Aktivität, deren Kompensation, erfolglos beendete Aktivitäten (Zustand Failed in BPMN 2.0 [OMG11]) oder bei Sprüngen zu anderen Aktivitäten einer Prozessinstanz [Ba22] erforderlich.

4.2 Ausführungsregeln

Der Startzustand Inactive einer Aktivität(eninstanz) wechselt zu Active (1 in Abb. 6), sobald alle relevanten Vorgängeraktivitäten (vgl. Def. 3) beendet wurden. Da es sich um den Zustandswechsel zu Active – d.h. nicht zu PreActive – handelt, müssen alle Vorgängeraktivitäten tatsächlich den Endzustand Completed erreicht haben und der Strictness-Typ der Kanten spielt keine Rolle (dieser wird in Regel 2 berücksichtigt). Konkret bedeutet das, dass jede Vorgängerakt. v der Akt. a regulär beendet sein muss, also den Zustand Completed erreicht hat. Es genügt nicht, dass sie den Zustand PreCompleted erreicht hat, d.h. bei dem GP aus Abb. 4 ist Akt. e erst dann regulär startbar, wenn alle ausgeführten Vorgängeraktivitäten $d1$ bis $d5$ regulär beendet wurden.

Regel 1: Wenn für eine Akt. $a \in N$ mit $\text{State}(a) = \text{Active}$ gilt:

$\forall v \in \text{RelevantPredecessors}(a)$ ist $\text{State}(v) = \text{Completed}$,
dann wird $\text{State}(a)$ auf Active gesetzt.

Ist der eben beschriebene Zustandsübergang nicht möglich, so kann eine Akt. a dennoch startbar werden (und zwar vorzeitig), falls Kanten mit den Strictness-Typen optional oder soft in sie münden. Damit die „Vorzeitigkeit“ erkennbar ist, wechselt die Akt. a

dann in den Zustand PreActive (anstatt Active). Bei optionalen Kanten genügt es, dass deren Quellakt. v startbar (Zustand Active) oder selbst vorzeitig startbar (PreActive) ist, oder einen Nachfolgezustand hiervon erreicht hat, d.h. der Zustand von Akt. v darf lediglich nicht Inactive sein. Bei Soft-Kanten muss die Bearbeitung der Quellakt. v beendet sein, wobei jedoch die Vorgänger von Akt. v noch nicht beendet sein müssen. Das ist daran erkennbar, dass sich die Akt. v im Zustand PreCompleted befindet (d.h. nach einer vorzeitigen Bearbeitung beendet wurde). Natürlich ist ein solcher Zustandswechsel der Akt. a auch möglich, wenn eine bestimmte Vorgängerakt. v regulär beendet wurde (Zustand Completed). Alle über Mandatory-Kanten verbundenen Vorgängeraktivitäten müssen, so bei Regel 1, beendet sein, also den Zustand Completed erreicht haben.

Regel 2: Wenn für eine Akt. $a \in N$ mit $\text{State}(a) = \text{Inactive}$ gilt:

Die Bedingung von Regel 1 ist nicht erfüllt \wedge

$\forall v \in \text{RelevPredWithEdge}_{\text{mandatory}}(a)$ ist $\text{State}(v) = \text{Completed} \wedge$

$\forall v \in \text{RelevPredWithEdge}_{\text{soft}}(a)$ ist $\text{State}(v) \in \{\text{Completed}, \text{PreCompleted}\} \wedge$

$\forall v \in \text{RelevPredWithEdge}_{\text{optional}}(a)$ ist $\text{State}(v) \neq \text{Inactive}$,

dann wird $\text{State}(a)$ auf PreActive gesetzt.

Eine Akt. a befindet sich in einem Pre-Zustand, wenn sie startbar ist, obwohl nicht alle direkten Vorgängeraktivitäten beendet wurden. Erreichen all diese irgendwann später den Zustand Completed, dann wurden sowohl diese, als auch ihre Vorgängeraktivitäten (d.h. die indirekten Vorgängeraktivitäten der Akt. a) regulär beendet. Deshalb kann die Akt. a in den entsprechenden Zustand ohne „Pre“ wechseln. Befand sich Akt. a im Zustand PreActive (i in Regel 3), dann ist sie ab jetzt regulär startbar, d.h. die Warnung an den Benutzer entfällt (wie bei Abb. 1b erläutert entfällt die Warnung für die Akt. Röntgen, sobald die Akt. EKG beendet wurde). Die nachfolgende Regel definiert alle Zustandsübergänge vom unteren zum oberen Pfad von Abb. 6.

Regel 3: Wenn für eine Akt. $a \in N$ mit $\text{State}(a) \in \{\text{PreActive}, \text{PreRunning}, \text{PreCompleted}\}$ gilt:

$\forall v \in \text{RelevantPredecessors}(a)$ ist $\text{State}(v) = \text{Completed}$,

dann ergibt sich der neue Zustand dieser Akt. a folgendermaßen:

$$\text{State}(a) = \begin{cases} \text{Active}, & \text{falls } \text{State}(a) = \text{PreActive} \text{ war} & \text{(i)} \\ \text{Running}, & \text{falls } \text{State}(a) = \text{PreRunning} \text{ war} & \text{(ii)} \\ \text{Completed}, & \text{falls } \text{State}(a) = \text{PreCompleted} \text{ war} & \text{(iii)} \end{cases}$$

Wenn eine Akt. a vom Benutzer gestartet wird, dann wechselt sie normalerweise in den Zustand Running (vgl. BPMN 2.0 [OMG11]). Im Kontext optionaler Kanten wird zusätzlich der Zustand PreRunning erforderlich. Durch diesen wird erkannt, dass die Akt. a noch nicht regulär startbar war, sondern vorzeitig gestartet wurde. Analog wechselt sie später (Regel 5) nicht in den Zustand Completed, sondern PreCompleted.

Regel 4: Wenn die Akt $a \in N$ einen Zustand $\text{State}(a) \in \{\text{Active}, \text{PreActive}\}$ hat, dann kann diese von einem Benutzer gestartet werden. Ihr Zustand ändert sich dann in:

$$\text{State}(a) = \begin{cases} \text{Running}, & \text{falls } \text{State}(a) = \text{Active} \text{ war} & \text{(i)} \\ \text{PreRunning}, & \text{falls } \text{State}(a) = \text{PreActive} \text{ war} & \text{(ii)} \end{cases}$$

Eine aktuell ausgeführte Akt. a kann vom Bearbeiter abgeschlossen werden. Wurde sie vorzeitig gestartet (d.h. PreRunning), dann geht ihr Zustand nicht direkt in Completed über, damit erkennbar ist, dass die Bearbeitung vorzeitig erfolgt ist. Dadurch bleibt z.B. eine Nachfolgerakt. n , die mit einer nicht-optionalen Kante verbunden ist (d.h. Strictness-Typ mandatory oder soft), nicht startbar, weil die Regel 1 zwar den Zustand Completed berücksichtigt, nicht aber PreCompleted.

Regel 5: Wenn die Akt $a \in N$ einen Zustand $\text{State}(a) \in \{\text{Running}, \text{PreRunning}\}$ hat, dann kann diese von einem Benutzer beendet werden. Ihr Zustand ändert sich dann in:

$$\text{State}(a) = \begin{cases} \text{Completed}, & \text{falls } \text{State}(a) = \text{Running} \text{ war} & \text{(i)} \\ \text{PreCompleted}, & \text{falls } \text{State}(a) = \text{PreRunning} \text{ war} & \text{(ii)} \end{cases}$$

5 Zusammenfassung und Ausblick

Wie in Abschnitt 2 dargestellt, wird die Relevanz optionaler Kanten in der wissenschaftlichen Literatur zwar erwähnt, aber ihr gewünschtes Verhalten wurde noch nicht exakt definiert. Dies betrifft sowohl das Verhalten des PMS gegenüber den Benutzern (z.B. das Anbieten von Einträgen in Arbeitslisten) als auch das der GP-Engine zur Runtime. Bei der Analyse dieser Forschungslücke wurde erkannt, dass das erforderliche Verhalten keineswegs eindeutig ist, wenn nicht nur eine einzelne optionale Kante betrachtet wird, sondern eine ganze Sequenz von derart verbundenen Aktivitäten. Deshalb wurde der zusätzliche Strictness-Typ „Soft“ eingeführt und in den Ausführungsregeln geeignet behandelt. Diese berücksichtigen zudem, dass sich vor einem Join-Gateway mit unterschiedlichen Strictness-Typen verbundene Aktivitäten befinden können.

Die Strictness-Typen verursachen kaum Aufwand bei der GP-Modellierung, weil sicherlich die meisten Kanten weiterhin mandatory sein werden, was die Standardeinstellung des Modellierungswerkzeugs sein sollte. Da zur Wahl der anderen Typen z.B. lediglich deren Typ ausgewählt werden muss, sind sie für GP-Designer sehr einfach verwendbar. Endanwender müssen lediglich einen Hinweis in ihrer Arbeitsliste beachten und können solche Aktivitäten ansonsten wie gewohnt bearbeiten. Deshalb stellen die neuen Kantentypen für beide Personengruppen keine Herausforderung dar, welche eine Evaluation der „Usability“ erforderlich machen würde. Eine technische Evaluation der neu vorgestellten Ausführungsregeln mittels eines Prototyps steht aktuell noch aus. Da diese den klassischen Ausführungsregeln einer GP-Engine stark ähneln, ist hierbei jedoch nicht mit besonderen Erkenntnissen zu rechnen.

Literaturverzeichnis

- [APS09] Aalst, W. van der; Pesic, M.; Schonenberg, H.: Declarative Workflows: Balancing between Flexibility and Support. *Computer Science-Research and Development* 23, S. 99–113, 2009.
- [AWG05] Aalst, W. van der; Weske, M.; Grünbauer, D.: Case Handling: A New Paradigm for Business Process Support. *Data & Knowledge Engineering* 2/53, S. 129–162, 2005.
- [Ba19] Bauer, T.: Pre-modelled Flexibility for Business Processes. *Proc. 21th Int. Conf. on Enterprise Information Systems*, S. 547–555, 2019.
- [Ba20] Bauer, T.: Business Processes with Pre-designed Flexibility for the Control-Flow. *Proc. 22nd Int. Conf. on Enterprise Information Systems*, 631–642, 2020.
- [Ba21] Bauer, T.: Pre-modelled Flexibility for the Control-Flow of Business Processes: Requirements and Interaction with Users. *Enterprise Information Systems*, 833–857, Springer, 2021.
- [Ba22] Bauer, T.: Requirements for Dynamic Jumps at the Execution of Business Processes. *Proc. 12th Int. Symposium on Business Modeling and Software Design*, S. 35–53, 2022.
- [Ba23] Bauer, T.: Modelling of Advanced Dependencies Between the Start and the End of Activities in Business Processes. *Proc. 25th Int. Conf. on Enterprise Information Systems*, S. 457–465, 2023.
- [Gr00] Grefen, P. et al.: CrossFlow: Cross-organizational Workflow Management in Dynamic Virtual Enterprises. *Computer Systems Science & Engineering* 5, S. 277–290, 2000.
- [Kl00] Klingemann, J.: Controlled Flexibility in Workflow Management. *Proc. Int. Conf. on Advanced Information Systems Engineering Stockholm*, S. 126–141, 2000.
- [Li06] Li, M. et al.: DOPA: A Data-Driven and Ontology-Based Method for Ad Hoc Process Awareness in Web Information Systems. *Proc. Int. Conf. on Web Information Systems Engineering*, S. 114–125, 2006.
- [OMG11] Object Management Group: Business Process Model and Notation (BPMN) 2.0, 2011.
- [Pe07] Pesic, M. et al.: Constraint-based Workflow Models: Change Made Easy. *Proc. 15th Int. Conf. on Cooperative Information Systems Vilamoura*, S. 77–94, 2007.
- [RD98] Reichert, M.; Dadam, P.: ADEPTflex - Supporting Dynamic Changes of Workflows Without Losing Control. *Journal of Intelligent Information Systems, Special Issue on Workflow Management Systems* 2/10, S. 93–129, 1998.
- [RH06] Russell, N.; Hofstede, A. ter: Workflow Control-Flow Patterns. A Revised View, *BPM Center Report BPM-06-22*, 2006.
- [RW12] Reichert, M.; Weber, B.: Enabling Flexibility in Process-Aware Information Systems. *Challenges, Methods, Technologies*. Springer, 2012.
- [We19] Weske, M.: Business Process Management: Concepts, Languages, Architectures. Springer, 2019.

The Reuse of Business Process Automation Artefacts

Peter A. François¹ and Ralf Plattfaut²

Abstract: Reusing business process models and process-related artefacts is deemed an essential factor in BPM initiatives by researchers and practitioners. The reuse of corresponding artefacts has been researched and practiced for almost 40 years. Yet, with the growing adoption of component- and platform-based information systems, a current review of how we view and conduct reuse in Business Process Automation becomes necessary. The advent of citizen developers further increases the need for a consolidated overview of the knowledge on reuse. As citizen developers are not trained to conduct reuse, guidance on how to harness the potential of reuse is needed. Based on a literature review, we identify four main layers on which process automation artefacts are reused and list exemplary technologies and methodologies. In addition, we systematize nine characteristics of reused automation artefacts and nine characteristics that describe a concrete instance of reuse. Finally, we provide a research agenda.

Keywords: Automation Reuse, Business Process Automation, Literature Review, Process Reuse, Reuse Layers, Citizen Developer

1 Introduction

Both supporting digital innovation and enabling quick adaptation of processes and automation solutions are current challenges for business process management (BPM) [Ke21]. There is also wide agreement that – where sensible – business processes and associated artefacts should be reused [Du18, Da15, RV15, RMR15].

While there is broad knowledge about the reuse of business process models (e.g., reference models, reuse of sub-processes, see e.g., [Du18, Ho10, BDK07]), scholarly research has identified the need for more insights on both process reuse [RMR15] and reuse of process automation solutions [As19, Sy20, LW18]. Especially since process automation solutions are considered digital innovations and are said to grow in importance [By15].

Digital innovations have disrupted several markets [Vi19]. The ability of Information Technology (IT) to automate current processes can be a deciding factor in the success of companies [WH20]. Recent disruptions in supply chains and operations have forced entire industries to rapidly adapt their processes or even business models. This pressure also led to the need to adapt corresponding Information Systems (IS) [CCT22, Rö22]. However, companies are partially hindered in building adequate automation solutions themselves

¹ South Westphalia University of Applied Sciences, Lübecker Ring 2, 59494 Soest, francois.peter@fh-swf.de, <https://orcid.org/0000-0003-1537-1026>

² University of Duisburg-Essen, Chair of Information Systems and Transformation Management, Universitätsstraße 2 45141 Essen, ralf.plattfaut@icb.uni-due.de, <https://orcid.org/0000-0002-1442-4758>

due to a shortage of competent IT workforce. This shortage has persisted almost consistently since the 1980s, leading to a decreased ability to create new IT artefacts and maintain existing ones [Ko84, KS98, NR06, Di15, BNK22].

The reuse of IT artefacts for process automation has been named one measure to cope with both the shortage of competent IT workforce and the need for innovation through quick creation and adaptation of IS (see, e.g., [Ha99, SRK12, Ry19, RJS17]). Next to resolving this dilemma, reuse also brings several benefits, such as improved maintainability, quality, flexibility, productivity and the ability to scale (see, e.g., [LW18, AP10, ZT05, Wi06, Ap90, RJS17, As19]). Especially in process automation, the need for reuse has been highlighted in the literature. For example Lacity and Willcocks [LW18, LW21] describe the reuse of automation artefacts as necessary to be able to reach a significant scale under consideration of cost and time. Dumas et al. [Du18] describe “standardization and reuse across projects” as criteria for a high BPM Maturity.

Since the reuse of artefacts brings such a wide range of benefits, we argue that there is merit in further exploring the reuse of automation artefacts. The ongoing relevance and extensive history of reuse within research (c.f. [Mc68, HHL22]) lend themselves to a literature review. Without clear knowledge and guidelines on automation reuse in the entire artefact lifecycle, several issues arise. Such issues include the inability to design and create IT artefacts in a way that makes them easy to reuse, yet effective when reused (e.g., [ZT05, Ap90, 422, DUB11]), realizing the highest possible value using reusable artefacts using the least resources (see, e.g., [LW18, Wi06, LMS07]) and systematically applying reuse across different hierarchical layers of automation reuse (see, e.g., [KS98]). In this article we therefore aim to uncover the following research questions:

RQ 1. How is reuse conducted in Business Process Automation?

RQ 2. Which Characteristics of artifacts and reuse instances (concrete cases of reuse) are mentioned in the literature?

To further our understanding of the reuse of automation components and to enable BPM and process automation research to make use of related concepts, we first identify layers reuse can be conducted on and then develop characteristics of both the Reuse Artefacts (RAs) and the Reuse Instances (RIs).

2 Background

Following existing literature, we define the reuse of automation solutions as using the same IT-artefact in new circumstances or taking (parts of) an existing artefact [KLJ18, Br01] in order to save on effort when creating or extending an artefact [KS98, Br01, FK05]. The initial artefact may have been previously developed in-house [KS98] or supplied externally [KLJ18].

As such, the reuse of process automation solutions is a cognitive process that involves analysis of the existing artefact in light of new requirements, selecting suitable parts of the

existing artefact, and (if necessary) adapting these parts or adding additional elements (other existing artefacts or specifically developed ones) (see, e.g., [PB13, EJP21, PSH03, As19]) to enable additional automation. Each complete execution of such a process is referred to as a reuse instance (see e.g., [JRS06, RJS17, BB91]).

There has been a long history of reusing automation artefacts. Montecinos [Mo96] argues that humans have always utilized reuse in their reasoning processes and that experience is nothing other than the reuse of existing knowledge and methods. The reuse of information artefacts originated in the 1960s (see, e.g., [Ko07, Mo96, Mc68]). Reuse research first focused on in-house code reuse [RJS17, AP10]. Newer concepts tend to be component-based and work by integrating smaller subsystems (e.g., accessed through such APIs [HHL22]), focusing on interoperability between different systems and languages as well as swift re-configurability and reusability (see, e.g., [Bl00, SJ03, SR05, ZS02]). The ability to quickly adapt systems is especially important when processes change [RL08]. At the same time, some of the “old” concepts of reuse, like functions [KS98] and objects [SJ03] have become an essential part of artefact development. However, a structured overview regarding the concepts of reuse in automation is still missing, which we will provide in the following.

3 Method

As the reuse of artefacts has a long history in IS research and practice, we conducted a structured literature review to answer our research questions. Literature reviews allow for gathering and systemizing rigorously published knowledge within a domain [Pa15]. We followed a three-step sequential process [WW02, Vo15]:

We first applied a keyword-based approach to identify relevant literature and give a comprehensive overview of the concepts used. We used automation reuse (as two words without operators) as keywords across all search fields to first identify relevant articles. As research outlets, we chose the databases of key journals and IS conferences [Vo15]. To uncover BPM-specific concepts, we included the BPM conference. In order to achieve a broad view of our field, we included the AIS e-library as well as the senior scholar basket of eight (EJIS, ISJ, ISR, JIT, JMIS, JSIS, JAIS, MISQ). These outlets were chosen since they represent a broad picture of our field, combining both high scientific relevance (ad established journals) and state-of-the-art publications (ad conferences) [GMT14]. We did not limit the publications reviewed to a specific timespan since older reuse concepts may still be relevant or transferrable to current contexts.

This initial search led to 2075 identified articles. After removing 76 duplicates, the first author manually screened the remaining articles by title for their potential relevance to the topic and research questions. Articles where the title implied that the articles were out of the scope of this review (e.g. “Driver Trust in Automated Driving Systems”). 1.745 articles were excluded based on this. For the remaining 254 articles, the first author manually

screened the abstract and excluded 118 articles mentioning the reuse of software and automation only in passing. This resulted in 136 articles for our thorough analysis. During this analysis, we added another 34 articles through a selective backward search, yielding a total of 170 articles. A foreword search was not deemed necessary due to the broad search term and the inclusion of conference proceedings (see Figure 1).

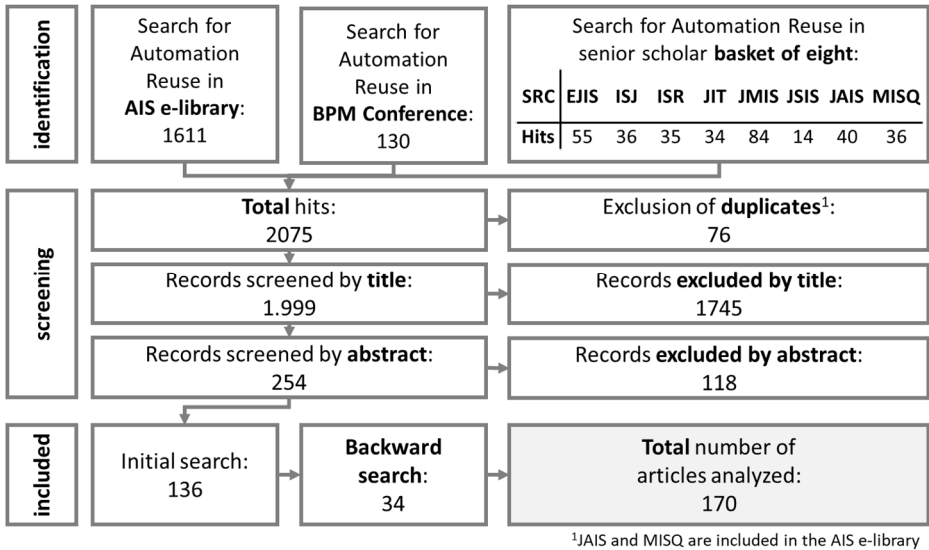


Fig. 1. Literature search process

In the second step, we followed the principles of grounded theory to analyze the literature [WFW13, Ba15]. All relevant articles were initially coded using open coding by both involved researchers to uncover the fundamentals of automation reuse and further relevant themes and theories. Next, we reviewed the resulting codes and aligned similar codes in an iterative process [Pe19] (e.g. “SoftwareComponents_ChoosingGranularity” or “F_Reuse_Layer”). We yielded 89 of these aggregated codes, found in 826 coded segments in the sources.

As a third step, we applied another final coding iteration, in which we applied axial coding within the codes formed in Step 2 to uncover relevant features of automation reuse in a similar iterative process [WFW13]. This second round of axial coding was equivalent to the categories mentioned in our main findings, as depicted in Figure 2 and Table 1. Finally, we compare the results of our grounded-theory-based literature review with the results of prior overview articles (i.e., [KS98, PCH93]).

4 Findings

4.1 Layers of Automation Reuse

Based on our analysis of 170 unique research articles, we extend the work of Poulin et al. [PCH93] and identify four main layers on which automation reuse can occur: Knowledge, Specifications, Software and Architecture (see Figure 2). Here, reuse on a higher layer also includes some degree of reuse on lower layers (e.g., reuse of specifications includes reuse of the corresponding knowledge).

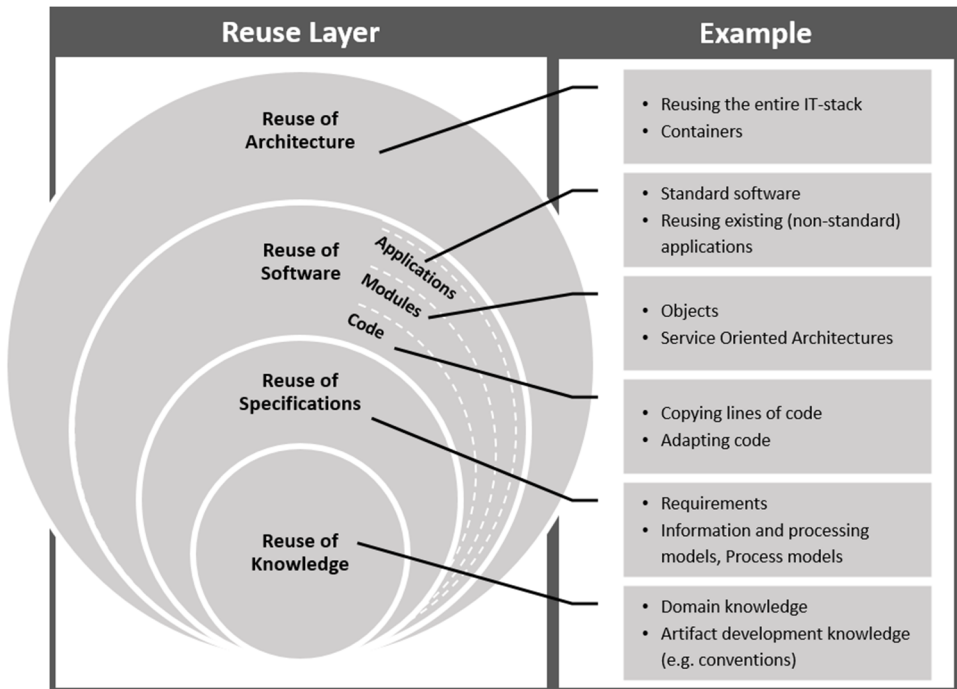


Fig. 2. Layers of Automation Reuse and corresponding examples

The innermost layer identified is the layer of knowledge: Actors can reuse knowledge gained in previous projects for new development [JT08, Ko07, EJP21, Mo96, SM04]. This knowledge includes, for example; domain knowledge [JT08, PSH03, Mo96], the results of data analysis [Ap90], and knowledge on the artefact development processes itself [JT08, FK05, Ap90, EJP21, SM04]. The knowledge reused may be obvious to the actors or hidden in higher layers [SR05, JT08]. One example of structured knowledge reuse are patterns. Patterns offer possible paths to a solution when a specified (often occurring) problem is faced. The knowledge of how to solve such a problem can therefore be reused [Ha99, ACL09, EJP21]. Patterns can also be used together artefacts on other layers such as software [EJP21, Ha99].

In the second layer, knowledge that has been explicated in the form of specifications can be reused [PSH03, ACL09, Ha99, KS98, Da15]. This includes requirements [ACL09, BB91, ZS02], conceptual information/data models [KS98, HAG08], process models or process descriptions (see e.g., [Ap90, MPR20, KS98, Bē15, VS17, ACL09, BDK07]), graphical design [BB91, Ko07], and software documentations [BB91, KS98, Ra99, ACL09]. These specifications may be rigid and (partially) reused as is (see e.g., Fr14, JW08, SD16) or created with the possibility of flexibility / adaptability / configurability, allowing easy reuse in various environments [Fr14, MPR20, BDK07, va13]. Reference processes are another significant concept allowing the reuse of generalized processes across several organizations (see e.g., [K112, BDK07, Zh08a, HSV09]).

In the third layer, software created based on the specifications can be reused. Here, the literature differentiates between the reuse of source code, modules, or whole applications. With regards to source code [KS98, Ry19, SM04, Ap90], scholars highlight that single lines of codes [KS98, SM04] or code fragments [KS98] can be reused.

A comparably large research stream focuses on using modules or components (see, e.g., [AP09, BI00, GRB17]). We understand a module as “a physical packaging of executable software with a well-defined and published interface” [Ho00]. The literature describes modules for use within one technology or system (e.g., one programming language run on one machine) and for use across several. During artefact creation in modern languages, classes or objects [Ba05, Br01, KS98, SJ03, SR05, Wi06, FH00], methods, functions and subroutines [Ry19, KS98, Wi06], as well as standard libraries or packages [KS98] are routinely used (see, e.g., [Ba05]). These artefacts are commonly reused within the same artefact generation process but can also be reused later on (see, e.g., [FH00, SR05]). Such practices have been verified in practice [Sw91] with companies developing repositories for module-based automation reuse [LW18].

Technologies for module (re)use across multiple technologies and systems have gained traction more recently (see, e.g., [SJ03, SR05]). Related technological and organizational questions are still under research (see, e.g., [SWK22, RJS17, Li21]) and several technologies and methodologies in this vein have been proposed [Ha99, BI00, JRS06, JW07]. Stemming from object-oriented development, concepts such as Component-Based Development [SR05, ZS02], Microservices [Li21, KLJ18] and Services based development (see, e.g., [Li21, RL08, Zh08b]) allow independent components to be freely arranged and thus follow business processes closely. Such services are often designed for reuse or have reuse in mind at the development stage [RL08, KS98]. It may therefore be hard or even impossible to distinguish between the “use” and “reuse” of such components. Kleinert et al. [K112] combine the concept of reference processes with reference automations for these processes under the name of “service blueprints”. Increasingly there are platforms that allow creating and sharing of modules, enabling users to connect different components created by themselves, the platform owner or third-party suppliers to create information systems from loosely coupled modules [VK19, HHZ18].

Moreover, the era of the reuse of whole applications came with the advent of packaged software applications in the 1960s. Standard software created by software vendors can be

bought and reused in the organization, taking advantage of the proven “standard” functionalities built into them (see, e.g., [RK00, MPR20]). Software-as-a-Service approaches are another way of reusing entire or partial applications (see, e.g., [KLJ18, WOK18]). Next to this, organizations can also reuse their existing artefacts. Schreieck et al., for example, describe using an existing application as part of a new information artefact [SWK22]. In this RI, the legacy system provides the core functionality for a platform ecosystem, allowing (from the platform component creators’ viewpoint) the reuse of this core functionality of the system.

In the fourth and final layer, whole IT architectures can be reused. One example of this is containerization, where every system needed for a specific task is bundled into one executable environment [Li21]. In addition, Schreieck et al. describe reusing parts of an IT infrastructure across multiple associated organizations [SWK22]. Karhu et al. describe competitors reusing open-source IT architectures [KGL18].

4.2 Characteristics of Reuse Artefacts and Reuse Instances

The literature describes several characteristics as relevant for automation reuse. The characteristics of RAs (artefact characteristic A1-9) and the characteristics of the specific RI concerning a specific artefact (instance characteristics I1-9) are systemized in Table 1. While some of these characteristics may be relevant when using any (predefined or ad hoc) reuse process on any reuse layer (like described in 4.1) or any combination of reuse layers, some may only be meaningful (or possible) in specific manifestation compositions.

Characteristic		Manifestation						
Reuse artefact characteristics	Artefact creation	(A1) Artefact Layer	Knowledge	Specifications	Software			Architecture
		(A2) Size			Code	Module	Application	
		(A3) Reuse Orientation	Designed for reuse		Not designed for reuse			
		(A4) Specificity	Customized	Function specific		Generic		
		(A5) Dependency	Independent	Interconnected		Dependent		
	Preparation for Reuse	(A6) Intended approach	Not intended	Bottom-up		Top-down		
		(A7) Visibility	Black box			White box		
		(A8) Documentation	No documentation		Interface documentation		Full documentation	
		(A9) Reuse methodology	Model-based method	Mixed method	Component-based method	Configur. based method	Not method based	

Characteristic (contd.)		Manifestation (contd.)						
Reuse <i>instance</i> characteristics	This instance	(11) Reuse Layer	Knowledge	Specifications	Software			Architecture
		(12) Reuse volume	Artefact partially reused		Artefact mostly reused		Artefact fully reused	
		(13) Adaptation	Significant		Some		None	
	Environment	(14) Process Logic	Identical		Similar		Different	
		(15) Project affiliation	Different project			Same project		
	Reuse methodology	(16) Reuse approach	Bottom-up			Top-down		
		(17) Reuse methodology	Model-based method	Mixed method	Component-based method	Configur. based method	Not method based	
		(18) Class of Actor	Original developers	Other developers	Other organizations	Private users	Malicious entities	
		(19) Software Support	Software supported			Not software supported		

Tab. 1: Characteristics of Reuse Artefacts and Reuse Instances

First, we already established that artefacts can be reused on several layers (A1). The reused artefact is associated with a specific layer, yet when the reuse process itself is conducted, it can be reused on the same layer or a different layer (I1). A finished application (e.g., a chatbot automating a customer service process) may be reused as a module in another system (e.g., in a self-service portal), or knowledge explicit in specifications may be reused in another project (e.g., a new chatbot for an internal human resource process).

Another characteristic of RAs is their size (A2). Some artefacts require a certain size in order to be reused. Size is especially important on the component layer. The reuse of large components brings higher value, as more is reused, yet more modification may be required [NR06, SRK12, KS98, PCH93], but small components can be reused on more occasions and can thus produce value more often since they cover less functionality and requirements, yet are harder to find [NR06, RJS17, Ap90, KS98, PCH93]. This trade-off implies the existence of an ideal component size, yet this ideal size is hard to find (see, e.g., [KS98, AP09, Jo09]) and also context-specific [NR06, SRK12].

In the reuse process, the artefact size can also affect the reuse volume (I2). Artefacts can be fully reused, mostly reused (with minor alterations), or only partially reused. Artefacts in the right context and the right size can be fully reused. In other cases, adaptation (I3) is needed (see, e.g., [KS98]). This adaptation may be necessary from a technological perspective since the process logic (I4) differs in the new usage environment.

The amount of adaptation required for reuse depends on the reuse orientation (A3). Here the literature describes artefacts designed for reuse or not designed for reuse (see e. g. [KS98, Ap90, SM04]). When designing an artefact for reuse, the developers had reuse in

mind during development. Ideally, this would include reusability on each layer presented above. This can include artefacts that have been designed for the sole purpose of existing as reusable automation components (see, e.g., [SM04]) or artefacts that have been developed for a specific purpose but with reuse in mind (see, e.g., [ACL09, KS98]).

If the artefact has not been developed with reuse in mind, it can still be adapted for future reuse through some adaptation process (see, e.g., [Br01, AP10, Bē15]). This can happen as part of the specific RI (in I3; see, e.g., [SWK22]) or with the goal of making the artefact reusable in general to potentially be reused at a later stage (e. g, the discovery of relevant legacy artefacts like described in [ACL09]). Also, components not designed for reuse and not adapted for reuse can still be reused using the software (of even single lines of codes) as it is.

According to Kim and Stohr, artefacts have a specificity (A4); They can be customized, functional or generic [KS98]. Customized artefacts are intended to satisfy specific use cases in specific processes according to specific requirements. Functional artefacts are more general in that they are reusable multiple processes within a specific application area or function, such as finance. Generic artefacts can be applied in any field and process where their functionality is required (e.g., a printing operation) [KS98].

Montecinos describes several dependency levels (A5) between artefacts [Mo96]. Independent artifacts can be reused without considering other artifacts. Interconnected artifacts rely on other artifacts and their characteristics and functionality may need to be considered during the reuse process. Dependent artifacts are closely intertwined with the RA and may require major rework for reuse.

Another factor in the reuse process is project affiliation (I5). Here we differentiate between reusing artefacts in the same or a different project [KS98, FH00, SR05]. Reuse in the same project may occur when creating different functionalities or across different lifecycle stages (e.g., reusing knowledge, documentation or existing code during maintenance) [Ap90, KS98, LMS07]. Similarly, artefacts can be reused in a different project.

The reuse approach (I6) can either be top-down or bottom-up. In a top-down approach, artefact creators try to predict the artefacts necessary and create those with the intent of reuse at a later (but not predetermined) time (see, e.g., [SM04]). In a bottom-up approach, existing artefacts are automatically or manually screened to identify potentially reusable components or to extract specific functionalities for specific use cases or later reuse (see, e.g., [DMC21]).

Artefacts can be reused with or without visibility (A7) regarding the inside workings. In a white box reuse case, the artefact is completely accessible to the reusing agent [KS98, SRK12]. In black box reuse, the reusing party does not have insights into the workings of the artefact but can “only” reuse the artefact by interacting with some form of interface and only knows the information given by the artefact developer [Mc68, JT08, JRS06], especially the documentation (see below). Of course, black box reuse may prevent adaptability and partial reuse of the artefact (see, e.g., [JRS06, SRK12]).

Documentation (A8) can be an essential factor in deciding if and how to reuse an automation artefact [Ra99, JT08, Ap90] since documentation increases “understandability” [KS98], trust [Ry19, RL08] and discoverability [DUB11]. In some instances, no documentation may be available for a specific component. In other instances, interface documentation [FSW00] or full documentation [KS98] may be available.

According to Han [Ha99], reuse can be based on different reused methodologies. We have divided this into the intended reuse methodology (A9) and the applied reuse methodology (I7). The literature names five different reuse methodologies [Bē15, Ha99]. First, the component-based methodology relies on assembling components to create larger systems [Ha99]. The model-based methodology uses (and adapts) meta-models to create new artefacts [Ha99]. Model-based and component-based methods can be mixed [Ha99]. Bērziša et al. describe that it is also possible to create RAs under a configuration-based method, where artefacts have built-in decision points to be usable in several contexts [Bē15, BDK07]. In addition, we argue that it is possible to reuse automation not method based (for example, ad hoc copying of single lines of code). The intended reuse methodology can be adhered to in the reuse process or omitted – hence the distinguished applied reuse methodology (I7).

Reuse can be carried out by several classes of actors (I8). These can be the original developers, other developers in the same company, or developers in other organizations [KS98]. Moreover, private persons and malicious entities may reuse artefacts [So14, Ry19, KGL18].

Finally, we found software support (I9) to be one of the important aspects in designing the reuse process (see, e.g., [Ra99, As19]), since the cost of finding out if there is a suitable artefact and where it is located can offset the benefit of reuse [Ra99, KS98, VS17]. Software-supported reuse processes can, for example, offer such a digital repository of RAs (see, e.g., [As19, Ko07, VS17, KS98]), software to extract RAs from legacy code (see, e.g., [ACL09]), or software recommending RAs while developing (see, e.g., [Wa14]).

5 Discussion

Our paper has three main contributions to existing theory. First, we showed the different layers on which automation artefacts can be reused. Here, we showed relevant technologies and concepts used on each layer. Secondly, we collected and systematized nine characteristics regarding RAs and which manifestations of these are mentioned in the literature. Finally, we found and systematized nine characteristics regarding RIs. We discussed the characteristics of RAs and the manifestations described in the literature.

The knowledge of the layers, the characteristics of RAs and RIs can help researchers to understand reusable artefacts better (what are the characteristics of a specific artefact), RIs (how was reuse conducted in this specific act of reuse) and existing reuse process guidelines (which combination of characteristics do the guidelines focus on / support / prohibit).

By defining the layers and characteristics, we enable researchers and practitioners to further advance the topic of automation reuse under a unified, common vocabulary. In addition, new research on reuse process directives can consider the individual layers and characteristics. The characteristics in Table 1 could be combined in various ways to develop previously unknown procedures for reuse support.

By systematizing the layers and characteristics as well as showing examples of relevant technologies, we have shown that there are many different types of automation reuse. Research on reuse in information systems, especially in BPM, should, instead of emphasizing the usefulness and need for reuse of automation artefacts, find out how reuse can best be realized in business process management and automation. We, therefore, (further down in this chapter) formulate a corresponding research agenda.

In addition, our paper holds several implications for practice. According to Rothenberger and Kulkarni, knowledge about the artefact is required to retrieve and find the right component [RK00]. The characteristics of RAs and instances discovered in this paper can be used to describe RAs (e.g., when used in a repository of RAs) by researchers and practitioners. They can additionally help discover and evaluate suitable RAs (e.g., in artefact databases). The overview can also help practitioners decide which layer(s) to perform reuse on and plan for reuse on appropriate layers when generating automation artefacts. When developers design and implement automation artefacts, they can – using the systematized layers and characteristics – be aware of and plan for the reusability of different components of their artefacts (and the artefacts on lower layers). Furthermore, they can use the characteristics mentioned in Table 1 (especially those in the section “Preparation for Reuse”) to pave the way for easy reusability.

Still, this paper has three main limitations. Firstly, since it is a pure literature review, we do not know which concepts are (still) relevant in practice. Secondly, adjacent fields have only been covered through the backward search. Finally – since the literature search and paper selection has been conducted by one author – there may be selection bias.

6 Research agenda

Based on our literature review, we derived a research agenda on the reuse of process automation technology. These questions originate from gaps identified in the literature.

First we would recommend research on which combination of characteristics (of the RA, RI, new artefact or additional elements) leads to successful reuse. In this paper we have defined several characteristics of RAs and RIs. Further research should advance these findings and validate them in practice. In addition, real-world data on the influence of characteristics combinations in reuse projects could be used to extend, evaluate and advance reuse process guidelines. Here, it is possible that multiple equifinal combinations of characteristics exist.

Furthermore, we need to find out how the knowledge regime used in developing the initial artefact or in the reuse process influences the success of reuse. Bygstad differentiates between heavyweight and lightweight IT knowledge regimes [By17]. In heavyweight IT, development is “driven by IT professionals, enabled by systematic specification and proven digital technology, and realized through software engineering” [By17]. In lightweight IT, users are enabled to fulfil their own IT needs and generate IT innovation through “cheap and easy-to-use” [By17] systems. Since development under such lightweight knowledge regimes is end-user-driven, users will have varying IT skill levels and knowledge about reuse. Well-designed support for reuse could help harness its potential in this area as well.

Item	Question(s)
Overarching	Which combination of characteristics leads to successful reuse?
	When should organizations rely on the reuse of process automation artefacts? Which environmental factors or characteristics of the components make reuse viable?
	How does the knowledge regime used in developing the initial artefact or in the reuse process influence the success of reuse?
	What can automation reuse research learn from adjacent streams such as knowledge management or requirements engineering?
	Since automation reuse stems from reuse in manufacturing in the 1960s (see above) are there newer concepts in manufacturing that could again be transferred for reuse?
	Which types of reuse are mutually beneficial, and which are obstructive / incompatible?
Reuse Layer	How to best combine / stack the different layers, technologies, and concepts of automation reuse in Business Process Automation?
Artefact	How to methodically decide which parts of an artefact to reuse?
Reuse Process	Which of the existing reuse process guidelines support reuse with which configuration of characteristics well?
	Which new reuse process guidelines could be developed to improve additional configurations?
Additional development	Which degree of additional development (also in the sense of requirements engineering and testing) should be performed during reuse?
New Artefact	How do properties of the reuse process or properties of the reused artefact affect the new artefact (short or long term)?
	How can artefacts developed with reuse be reused and what does re-reuse entail?

Table 2. Research agenda

Moreover, it would be valuable to evaluate further when organizations should rely on the reuse of process automation artefacts. While several proposals for the structure of reuse processes itself exist (see, e.g., [HAG08, ACL09, Sw91, PSH03, HHL22, BI00]), there seem to be few up-to-date articles focusing on concrete instructions on when to reuse (e.g.,

[CB91]). The general consensus seems to be, that reuse should be conducted when it is financially beneficial [KS98, RK00], leads to better quality or increases development speed [CB91]. However, we argue that these characteristics are complex for programmers to evaluate when “out in the field”. This issue increases when many possible reuse sources are available (e.g., large codebase, many components in a platform ecosystem) [RK00, NR06]. Therefore, we argue that research should be done on how these characteristics can be determined algorithmically to guide developers.

We also call for more research on how to best combine or stack the different layers of automation reuse and combine the technologies and concepts developed for automation reuse. While some concepts (e.g., model-driven development [LMS07]) rely explicitly on utilizing different layers to enable reuse, we have not seen any guidelines on how to systematically reuse artefacts across all layers in different contexts.

Another interesting problem is the degree of additional development that should be accepted during reuse (in regular or lightweight programming), like functional changes, such as adapting or expanding functionality, (additional) requirements engineering and testing for the new environment [KS98, SRK12, Ra99, PB13].

In addition, very little work seems to be done on the workings of the artefacts created through reuse. Baskerville et al. describe a case study where adaptations to the new artefact did not negatively affect the new artefact [Ba05]. However, it is unclear what the long-term effects of reuse are, which configurations of artefact characteristics, reuse process and additional elements lead to successful new artefact creation. It is also unclear how these new artefacts can be re-reused and its implications.

7 Conclusion

One of the basic rules in BPM is that the reuse of process models should take precedence before the new development [Du18, RMR15, RV15]. This precedence for reuse also extends to automation artefacts (Davenport, for example, states this for the knowledge layer [Da15]). The need for research on reusing automation solutions has been well described [As19, Sy20, LW18].

In this paper, we have demonstrated the state of the reuse of process automation artefacts in the IS literature. We have shown and systematized layers of automation reuse, related technologies, the characteristics of RAs, and the reuse process.

There is still much work for researchers in the area of automation reuse, as seen in the proposed research agenda. In 1968, McIlroy already called for adapting principles and technologies from the field of mechanical engineering for use in software engineering [Mc68]. He called for “manufacturers of standard parts”, which we now see in the existence of standard software and external service vendors, and for the ability to “order parts to individual specifications of size, ruggedness, speed, capacity, precision or character set” which we now see in requirements engineering, and “catalogues of standard parts”, which

we now see in the form of reuse repositories and platform ecosystems stores. While we now have access to these technologies (and many others as described above) and could utilize them for further reuse, the reuse of automation artefacts still seems to pose hurdles for researchers and practitioners. Thus, we hope our findings can help to advance the understanding of reuse and to tap into the potential that reuse holds for BPM and Business Process Automation.

Acknowledgements

This work has been partially developed in the project KEBAP at South Westphalia University of Applied Sciences. The project (reference number: 13FH034KX0) is partly funded by the German Ministry of Education and Research (BMBF).

References

- [ACL09] Aldin, L.; Cesare, S. de; Lycett, M.: Semantic Discovery And Reuse Of Business Process Patterns. MCIS 2009 Proceedings, 2009.
- [AP09] Antikainen, J.; Pekkola, S.: Factors influencing the alignment of SOA development with business objectives. ECIS 2008 Proceedings, 2009.
- [AP10] Allen, G.; Parsons, J.: Is Query Reuse Potentially Harmful? Anchoring and Adjustment in Adapting Existing Database Queries. *Information Systems Research* 1/21, p. 56–77, 2010.
- [Ap90] Apte, U. et al.: Reusability-Based Strategy for Development of Information Systems: Implementation Experience of a Bank. *MIS Quarterly* 4/14, p. 421–433, 1990.
- [As19] Asatiani, A. et al.: Organizational Implementation of Intelligent Automation as Distributed Cognition: Six Recommendations for Managers. ICIS 2019 Proceedings, 2019.
- [Ba05] Baskerville, R. et al.: Extensible Architectures: The Strategic Value of Service Oriented Architecture in Banking. ECIS 2005 Proceedings, 2005.
- [Ba15] Bandara, W. et al.: Achieving Rigor in Literature Reviews: Insights from Qualitative Data Analysis and Tool-Support. *Communications of the Association for Information Systems* 37, 2015.
- [BB91] Barnes, B. H.; Bollinger, T. B.: Making reuse cost-effective. *IEEE Software* 1/8, p. 13–24, 1991.
- [BDK07] Becker, J.; Delfmann, P.; Knackstedt, R.: Adaptive Reference Modeling: Integrating Configurative and Generic Adaptation Techniques for Information Models. In (Becker, J.; Delfmann, P. Hrsg.): *Reference Modeling. Efficient Information Systems Design Through Reuse of Information Models*. Physica, Heidelberg, p. 27–58, 2007.
- [Bē15] Bērziša, S. et al.: Capability Driven Development: An Approach to Designing Digital Enterprises. *Business & Information Systems Engineering* 1/57, p. 15–25, 2015.

- [BI00] Blake, M. B.: A Development Approach for Workflow-Based E-Commerce using Reusable Distributed Components. AMCIS 2000 Proceedings, 2000.
- [BNK22] Bossler, L. F.; Noeltner, M.; Kroenung, J. S.: Attitude Discrepancy and Its Influence on Turnover Intention Among IS Professionals. ICIS 2022 Proceedings 15, 2022.
- [Br01] Brash, D.: Reuse in Information Systems Development: Classification and Comparison. ACIS 2001 Proceedings., 2001.
- [By15] Bygstad, B.: The Coming of Lightweight IT. ECIS 2015 Completed Research Papers, 2015.
- [By17] Bygstad, B.: Generative Innovation: A Comparison of Lightweight and Heavyweight IT. *JIT (Journal of Information Technology)* 2/32, p. 180–193, 2017.
- [CB91] Caldiera, G.; Basili, V. R.: Identifying and qualifying reusable software components. *Computer* 2/24, p. 61–70, 1991.
- [CCT22] Compeau, D.; Correia, J.; Thatcher, J.: When Constructs Become Obsolete: A Systematic Approach to Evaluating and Updating Constructs for Information Systems Research. *MIS Quarterly* 2/46, p. 679–712, 2022.
- [Da15] Davenport, T. H.: Process Management for Knowledge Work. In (Vom Brocke, J.; Rosemann, M. Hrsg.): *Handbook on Business Process Management 1. Introduction, Methods, and Information Systems*. Springer, Berlin, Heidelberg, p. 17–35, 2015.
- [Di15] Dinger, M. et al.: Does Professionalism Matter in the IT Workforce? An Empirical Examination of IT Professionals. *Journal of the Association for Information Systems* 4/16, 2015.
- [DMC21] Diaz, O.; Medina, H.; Contell Jeremias P.: Promoting Design Knowledge Accumulation Through Systematic Reuse. HICCS 2021 Proceedings, 2021.
- [Du18] Dumas, M. et al.: *Fundamentals of Business Process Management*. Springer, Berlin Heidelberg, 2018.
- [DUB11] Dudek, S.; Uebernickel, F.; Brenner, W.: Towards Computer Aided It Service Engineering. MCIS 2011 Proceedings, 2011.
- [EJP21] Ebel, M.; Jaspert, D.; Poeppelbuss, J.: Pattern-Based Smart Service Innovation. PACIS 2021 Proceedings, 2021.
- [FH00] Ferguson, R. I.; Hunter, A.: On the Use of Object Oriented Techniques to Support the Construction of CASE Tools. AMCIS 2000 Proceedings, 2000.
- [FK05] Frakes, W. B.; Kang, K.: Software Reuse Research: Status and Future. *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING* 7/31, p. 529–536, 2005.
- [Fr14] Frank, U. et al.: The Research Field “Modeling Business Information Systems”. *Business & Information Systems Engineering* 1/6, p. 39–43, 2014.
- [FSW00] Fan, M.; Stallaert, J.; Whinston, A. B.: The adoption and design methodologies of component-based enterprise systems. *EJIS (European Journal of Information Systems)* 1/9, p. 25–35, 2000.
- [GMT14] Gogan, J. L.; McLaughlin, M.-D.; Thomas, D.: Critical Incident Technique in the Basket. ICIS 2014 Proceedings, 2014.

- [GRB17] Gruettner, A.; Richter, J.; Basten, D.: Explaining The Role Of Service-Oriented Architecture For Cyber-Physical Systems By Establishing Logical Links. ECIS 2017 Proceedings, 2017.
- [Ha99] Han, T.-D.: Automating Reuse for Systems Design. AMCIS 1999 Proceedings, 1999.
- [HAG08] Hammoudi, S.; Alouini, W.; Gammoudi, M. M.: On The Challenge Of A Semi-Automatic Transformation Process In Model Driven Enterprise Information Systems. MCIS 2008 Proceedings, 2008.
- [HHL22] Huang, J. C.; Henfridsson, O.; Liu, M. J.: Extending digital ventures through templating. *Information Systems Research* 1/33, p. 285–310, 2022.
- [HHZ18] Hahn, C.; Huntgeburth, J.; Zarnekow, R.: Leverage Once, Earn Repeatedly—Capabilities for Creating and Appropriating Value in Cloud Platform Ecosystems. In (Lamboglia, R. et al. Hrsg.): *Network, Smart and Open*. Springer International Publishing, Cham, p. 143–164, 2018.
- [Ho00] Hopkins, J.: Component primer. *Communications of the ACM* 10/43, p. 27–30, 2000.
- [Ho10] Hofstede, A. H. M. et al.: *Modern Business Process Automation*. Springer, Berlin, Heidelberg, 2010.
- [HSV09] Hoyer, V.; Stanoevska-Slabeva, K.; Vom Brocke, J.: On the Contribution of Reference Modeling for Organizing Enterprise Mashup Environments. In (Rinderle-Ma, S.; Sadiq, S.; Leymann, F. Hrsg.): *LNBIP 43. Business Process Management Workshops*, 2009.
- [Jo09] Joachim, N. et al.: Examining the Organizational Decision to Adopt Service-Oriented Architecture (SOA) - Development of a Research Model. *DIGIT 2009 Proceedings*, 2009.
- [JRS06] Jain, H.; Rothenberger, M.; Sugumaran, V.: Flexible Software Component Design Using A Product Platform Approach. *ICIS 2006 Proceedings*, 2006.
- [JT08] Juhrisch, M.; Thies, G.: Service Management Beyond UDDI - A Design Science Approach. *PACIS 2008 Proceedings*, 2008.
- [JW07] Juhrisch, M.; Weller, J.: On the Reuse of SOA Components on Business Process Analysis Stages. *PACIS 2007 Proceedings*, 2007.
- [JW08] Juhrisch, M.; Weller, J.: Connecting Business and IT - A Model-driven Web Service based Approach. *PACIS 2008 Proceedings*, 2008.
- [Ke21] Kerpedzhiev, G. D. et al.: An Exploration into Future Business Process Management Capabilities in View of Digitalization. *Business & Information Systems Engineering* 2/63, p. 83–96, 2021.
- [KGL18] Karhu, K.; Gustafsson, R.; Lyytinen, K.: Exploiting and Defending Open Digital Platforms with Boundary Resources: Android’s Five Platform Forks. *Information Systems Research* 2/29, p. 479–497, 2018.
- [K112] Kleinert, T. et al.: Systematic Identification of Service-Blueprints for Service-Processes - A Method and Exemplary Application. *BPM 2012 Workshops, LNBIP 132*, p. 598–610, 2012.

- [KLJ18] Krancher, O.; Luther, P.; Jost, M.: Key Affordances of Platform-as-a-Service: Self-Organization and Continuous Feedback. *Journal of Management Information Systems* 3/35, p. 776–812, 2018.
- [Ko07] Kotlarsky, J. et al.: Globally Distributed Component-Based Software Development: an Exploratory Study of Knowledge Management and Work Division. *JIT (Journal of Information Technology)* 2/22, p. 161–173, 2007.
- [Ko84] Konsynski, B. R.: Advances in Information System Design. *Journal of Management Information Systems* 3/1, p. 5–32, 1984.
- [KS98] Kim, Y.; Stohr, E. A.: Software Reuse: Survey and Research Directions. *Journal of Management Information Systems* 4/14, p. 113–147, 1998.
- [Li21] Li, S. et al.: Understanding and Addressing Quality Attributes of Microservices Architecture: A Systematic Literature Review. *Inf. Softw. Technol. (Information and Software Technology)* 106449/131, 2021.
- [LMS07] Lycett, M.; Marcos, E.; Storey, V.: Model-driven systems development: an introduction. *European Journal of Information Systems* 16, p. 346–348, 2007.
- [LW18] Lacity, M.; Willcocks, L. P.: Innovating in Service: The Role and Management of Automation. In (Willcocks, L.; Oshri, I.; Kotlarsky, J. Hrsg.): *Dynamic innovation in outsourcing. Theories, cases and practices*. Palgrave Macmillan, Cham, 2018.
- [LW21] Lacity, M.; Willcocks, L.: Becoming Strategic with Intelligent Automation. *MIS Quarterly Executive*, p. 169–182, 2021.
- [Mc68] McIlroy, M. D.: Mass Produced Software Components. NATO Software Engineering Conference. Garmisch, Germany, 7th to 11th October 1968, 1968.
- [Mo96] Montecinos, F.: An experience in the establishment of a software reuse culture in a real environment. *AMCIS 1996 Proceedings.*, 1996.
- [MPR20] Mending, J.; Pentland, B. T.; Recker, J.: Building a complementary agenda for business process management and digital innovation. *European Journal of Information Systems* 3/29, p. 208–219, 2020.
- [NR06] Nazareth, D. L.; Rothenberger, M.: Does the ‘Goldilocks Conjecture’ Apply to Software Reuse? *Journal Of Information Technology Theory And Application* 2/8, p. 57–67, 2006.
- [Pa15] Paré, G. et al.: Synthesizing information systems knowledge: A typology of literature reviews, *Information & Management. Information & Management* 2/52, p. 183–199, 2015.
- [PB13] Pahlke, I.; Beck, R.: *End User Empowerment through Platforms for Situational Applications*. All Sprouts Content, 2013.
- [PCH93] Poulin, J. S.; Caruso, J. M.; Hancock, D. R.: The business case for software reuse. *IBM Systems Journal* 4/32, 1993.
- [Pe19] Pekkola, S. et al.: The Magical “We”: Enhancing Collaboration Transparency in Grounded Theory Method in Information Systems Research. *Communications of the Association for Information Systems* 45, p. 251–269, 2019.

- [PSH03] Purao, S.; Storey, V. C.; Han, T.: Improving Analysis Pattern Reuse in Conceptual Design: Augmenting Automated Processes with Supervised Learning. *Information Systems Research* 3/14, p. 269–290, 2003.
- [Ra99] Ravichandran, T.: Software reusability as synchronous innovation: a test of four theoretical models. *European Journal of Information Systems* 3/8, p. 183–199, 1999.
- [RJS17] Rothenberger, M. A.; Jain, H.; Sugumaran, V.: A Platform-based Design Approach for Flexible Software Components. *Journal of Information Technology Theory and Application* 2/18, p. 29–56, 2017.
- [RK00] Rothenberger, M. A.; Kulkarni, U. R.: Software Reuse in Information Systems Development. *AMCIS 2000 Proceedings*, 2000.
- [RL08] Ren, M.; Lyytinen, K. J.: Building Enterprise Architecture Agility and Sustenance with SOA. *Communications of the Association for Information Systems* 22, 2008.
- [RMR15] Reijers, H. A.; Mendling, J.; Recker, J.: Business Process Quality Management. In (Vom Brocke, J.; Rosemann, M. Hrsg.): *Handbook on Business Process Management 1. Introduction, Methods, and Information Systems*. Springer, Berlin, Heidelberg, p. 167–185, 2015.
- [Rö22] Röglinger, M. et al.: Exogenous Shocks and Business Process Management. A Scholars' Perspective on Challenges and Opportunities. *Business & Information Systems Engineering* 5/64, p. 669–687, 2022.
- [RV15] Rosemann, M.; Vom Brocke, J.: The Six Core Elements of Business Process Management. In (Vom Brocke, J.; Rosemann, M. Hrsg.): *Handbook on Business Process Management 1. Introduction, Methods, and Information Systems*. Springer, Berlin, Heidelberg, 2015.
- [Ry19] Ryan, T. J. et al.: The Influence of Personality on Code Reuse. *HICSS 2019 Proceedings*, p. 5805–5814, 2019.
- [SD16] Saarsen, T.; Dumas, M.: Factors Affecting the Sustained Use of Process Models. In (La Rosa, M.; Loos, P.; Pastor, O. Hrsg.): *Business Process Management Forum. BPM Forum 2016, Rio de Janeiro, Brazil, September 18-22, 2016, Proceedings*. Springer, Cham, 2016.
- [SJ03] Satzinger, J. W.; Jackson, R. B.: Making the Transition from OO Analysis to OO Design with the Unified Process. *Communications of the Association for Information Systems* 12, 2003.
- [SM04] Sherif, K.; Menon, N. M.: Managing Technology and Administration Innovations: Four Case Studies on Software Reuse. *Journal of the Association for Information Systems* 7/5, 2004.
- [So14] Sojer, M. et al.: Understanding the Drivers of Unethical Programming Behavior: The Inappropriate Reuse of Internet-Accessible Code. *Journal of Management Information Systems* 3/31, p. 287–325, 2014.
- [SR05] Sharp, J.; Ryan, S.: A Review of Component-Based Software Development *ICIS 2005 Proceedings*, 2005.

- [SRK12] Subramanyam, R.; Ramasubbu, N.; Krishnan, M. S.: In Search of Efficient Flexibility: Effects of Software Component Granularity on Development Effort, Defects, and Customization Effort. *Information Systems Research* 3-part-1/23, p. 787–803, 2012.
- [Sw91] Swanson, K. et al.: The Application Software Factory: Applying Total Quality Techniques to Systems Development. *MIS Quarterly* 4/15, p. 567–579, 1991.
- [SWK22] Schreieck, M.; Wiesche, M.; Krcmar, H.: Governing innovation platforms in multi-business organisations. *European Journal of Information Systems*, p. 1–22, 2022.
- [Sy20] Syed, R. et al.: Robotic Process Automation: Contemporary themes and challenges. *Computers in Industry* 115, p. 103162, 2020.
- [va13] van der Aalst, W. M. P.: Business Process Management: A Comprehensive Survey. *ISRN Software Engineering Article ID 507984*,/Volume 2013, 2013.
- [Vi19] Vial, G.: Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems* 2/28, p. 118–144, 2019.
- [VK19] Vestues, K.; Knut, R.: Making Digital Infrastructures More Generative Through Platformization and Platform- driven Software Development: An Explorative Case Study. *10th Scandinavian Conference on Information Systems*, 2019.
- [Vo15] Vom Brocke, J. et al.: Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems* 37, 2015.
- [VS17] Veitch, R.; Seymour, L. F.: Business Process Model Reuse In A Multi-Channel / Multi-Product Environment–Problem Identification And Tentative Design. *MCIS 2017 Proceedings*, 2017.
- [Wa14] Wang, D. et al.: The Design of a Workflow Recommendation System for Workflow as a Service in the Cloud. In (Lohmann, N.; Song, M.; Wohed, P. Hrsg.): *Business Process Management Workshops*. Springer International Publishing, Cham, p. 251–263, 2014.
- [WFW13] Wolfswinkel, J. F.; Furtmueller, E.; Wilderom, C. P. M.: Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems* 1/22, p. 45–55, 2013.
- [WH20] Wiesböck, F.; Hess, T.: Digital innovations. *Electron. Markets (Electronic Markets)* 1/30, p. 75–86, 2020.
- [Wi06] Witman, P. D.: Tracing Patterns of Large-Scale Software Reuse. *AMCIS 2006 Proceedings*, 2006.
- [WOK18] Willcocks, L.; Oshri, I.; Kotlarsky, J. Hrsg.: *Dynamic innovation in outsourcing. Theories, cases and practices*. Palgrave Macmillan, Cham, 2018.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly* 2/26, p. xiii–xxiii, 2002.
- [Zh08a] Zhu, L. et al.: Challenges Observed in the Definition of Reference Business Processes. In (Hutchison, D. et al. Hrsg.): *Business Process Management Workshops*. Springer Berlin Heidelberg, Berlin, Heidelberg, p. 95–107, 2008.

- [Zh08b] Zhao, J. L. et al.: Impact of Service-Centric Computing on Business and Education. Communications of the Association for Information Systems 22, 2008.
- [ZS02] Zhao, L.; Siau, K.: Component-Based Development Using UML. Communications of the Association for Information Systems 9, 2002.
- [ZT05] Zhang, Y.; Tanniru, M.: Business Flexibility and Operational Efficiency - Making Trade-Offs in Service Oriented Architecture. AMCIS 2005 Proceedings, 2005.

Wirtschaft, Management Industrie -
Skills and Qualifications for a
Digitalized Future (SKILLS'23)

Digitalization and Sustainability in German Continuing Education

Felix Derksen,¹ Jens Dörpinghaus²

Abstract: Skills and qualifications are at the heart of designing digitalized and sustainable futures, as they link society and the labour market. While the design of digitalized futures raises several issues for society, society in turn, and in particular skills and qualifications, can be a bottleneck for maintaining productivity and the workforce. The labor market relies heavily on both vocational and academic education and training, retraining, and continuing professional development to meet these future challenges. Thus, at the individual level, the question arises as to what qualifications and skills are needed in a digitalized and sustainable future. In this paper we present a novel approach to analyze advertisements for continuing vocational education and training (CVET) in order to identify if skills and qualifications needed for digitalization and sustainability are currently considered, and if the so-called green economy is also important for CVETs.

Keywords: CVET advertisements, Labor market data, Text Mining, Digitalization, Sustainability, Computational Social Sciences

1 Introduction

Skills needs for digitalization and sustainability can be analyzed among different stakeholders, namely educational institutions, employers, employees and financiers. In this context, continuing education advertisements complement data from curricula, job advertisements, and labor market surveys. The German labor market is dynamic: technical innovations and social changes lead to new skills needed by employees. Vocational education and training (VET), retraining and continuing vocational education and training (CVET) are key to responding to these new demands [Do19; He16]. The central importance of continuing vocational training and lifelong learning has been increasingly addressed by policymakers as part of the National Continuing Education Strategy, see [Ge19; Sc22]. Here, continuing education is described as a central prerequisite for securing a skilled workforce, for ensuring the employability of all employees and thus also for national competitiveness and innovation.

However, the German education system offers different pathways for professionals [GL21] and we need to distinguish between initial (training, “Ausbildung” or retraining, “Umschulung”) and continuing vocational education, which includes advanced training (“Weiterbildung”, unregulated, e.g. continuing professional development) and upgrading training

¹ TH Köln, Cologne, Germany felix.derksen@smail.th-koeln.de and Federal Institute for Vocational Education and Training (BIBB), Bonn, Germany

² Federal Institute for Vocational Education and Training (BIBB), Bonn, Germany and University of Koblenz, Germany jens.doerpinghaus@bibb.de

(“Fortbildung”). Upgrading trainings are mostly regulated (on federal level [BBiG/HwO] or by federal states [DD20]), while advanced trainings are comparatively less regulated. They are usually summarized as continuing vocational education and training (CVET). Important actors in continuing vocational education and training in Germany include: (a) educational institutions, (b) companies and enterprises, (c) employees and (d) financiers.

All stakeholders have to deal with changing conditions and requirements as a result of transformation processes such as increasing digitalization [He16] and the move towards a more sustainable way of doing business [St22]. They all need to respond to emerging situations, such as the advancing digitalization and the move toward a sustainable economic system, by developing the educational content offered accordingly. From a research point of view, an important task is to determine which continuing education contents are increasingly being offered and demanded in order to draw conclusions about the development needs of the vocational training system and, in particular, of the continuing education system. The research-based further development of the vocational training system should not only ensure the competitiveness of the economy on a systemic level, but also contribute to counteracting unemployment and stabilizing the social security system [Ge19; Zi13].

In this paper, our research question is if we can determine whether the skills and qualifications needed for digitalization, sustainability, and the green economy are currently being considered in vocational training advertisements. To this end, we present a novel workflow and a detailed analysis, and discuss further research in this area.

2 Related Work

Green Economy is an important topic nowadays. Sustainability is becoming increasingly relevant, even in vocational education [Be22]. Moreover, in the realm of digitization, the question arises: “Does industrial structure optimization and green innovation matter?”, see [Ha23]. Even data centers as well as green data centers are being reassessed in terms of their environmental sustainability [Wa22].

To analyze vocational education we find an increasing interest in mining data from educational databases, advertisements and information systems, see [DIH17; MT13; RV07]. However, there are still several challenges on data and data integration, see [KKM20]. Skill concepts have been heavily used for the analysis of job advertisements (job ads) [De18] their visualization [Kr18] and offer a good starting point for matching open positions to corresponding employees. The proposed technologies range from automated mapping of search terms to classification of skills, see [Zi12]. In education research and pedagogy, several approaches towards competencies and skills exist which provide divergent resources: Not only with different meaning but also with different synonyms. While in English Competences, Skills and Knowledge are often used as synonyms, this does not hold for German language: Some words refer to multiple concepts, while some concepts are labeled with multiple words, see [KM22].

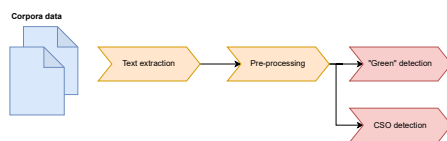


Fig. 1: An illustration of the proposed workflow

There is very little research on the identification and analysis of green economy and environmental sustainability in German texts. Binnewitt and Schnepf presented a solution to the question of to what extent training enterprises in the green economy use the attractiveness potential of ecological sustainability for their training advertisements and whether a change can be observed in recent years, see [BS22].

Since we can only rely on very limited previous work, we will continue with a detailed discussion of the methods to encourage further research in this field.

3 Method

3.1 Data

Our first evaluation data was obtained from the data portal KURSNET of BA³. We retrieved data for initial and continuing vocational education and training, which leads to 23,460 (training), and 66.549 (re-training) results in February 2023. Within the advertisement text, we only considered the section access information (“Zugangsinformationen”) with the subsection access (“Zugang”) which usually holds natural language describing the access limitations.

In Germany, the BA offers with the “Klassifikation der Berufe” (KldB), part C, a hierarchy of continuing vocational education, which is used to construct more granular subject areas, e.g., IT, ecology, technology, and so on. We consider a nearly equally distributed random subset of 120-160 ads per field.

3.2 Workflow

The proposed workflow consists of four different steps, see Figure 1. First, we retrieve the data for different job advertisements from different sectors and extract the relevant texts, i.e. the descriptive texts. Second, we use text preprocessing and NLP methods from the Spacy library. Then, we use pattern matching to extract entities to analyze and detect “green” ads and those related to digitalization. We will describe the methods for the last two steps in more detail in the next subsections.

³ See <https://www.arbeitsagentur.de/kursnet>.

3.3 Detecting ecological sustainability in textual data

We rely on the work of Binnewitt and Schnepf, who presented a solution to the question of to what extent training enterprises in the green economy use the attractiveness potential of ecological sustainability for their training advertisements and whether a change can be observed in recent years, see [BS22]. They were able to identify OJAs that advertise environmental sustainability. They used a German Federal Environment Agency (Umweltbundesamt) word list on the green economy and enriched this word list with additional terms using a lexical-semantic word network (GermaNet) and the method of word embedding (Word2Vec). All terms in the enriched word list are categorised according to their weak, medium and strong association with environmental sustainability.

3.4 Detecting digitalization in textual data

To support the detection of computing tools and software, we also consider word lists based on the Computer Science Ontology (CSO 3.2), which was introduced to cover CS-related publications. It is a comprehensive ontology of research areas, topics and concepts in computer science. It provides a formal description of the research areas, topics and concepts within the discipline of Computer Science that can be used to facilitate the understanding and discovery of research content. The CSO is extensive in nature and has been generated using 16 million publications related to the field of computer science [Sa20]. The CSO was created to improve the accuracy of scientific bibliographic databases and the ability of researchers to identify relevant research publications [Os16]. Gaps in an ontology refer to areas or concepts within a domain that are not described or covered by the ontology, in this case we are missing several alternative terms and in particular German terms. However, as CSO links to DBpedia and Wikidata, we have obtained German labels and alt-labels for all terms in CSO.

In addition, we can use CSO to classify the entities found, as it provides several top-level topics such as artificial intelligence, computer security or hardware.

4 Results

4.1 Ecological sustainability in CVET ads

The main question for Green CVET ads is if schools and institutes use green and sustainability related terms to improve their offered courses and trainings and are green skills part of their curriculum? We are also interested in whether these terms are mostly related to framing the offer by green topics or if they are due to novel methods, technologies and skills.

In Figure 2 we present a detailed analysis of CVET ads by categories and the amount of ads which have a small, middle or high association level with sustainability and green economy.

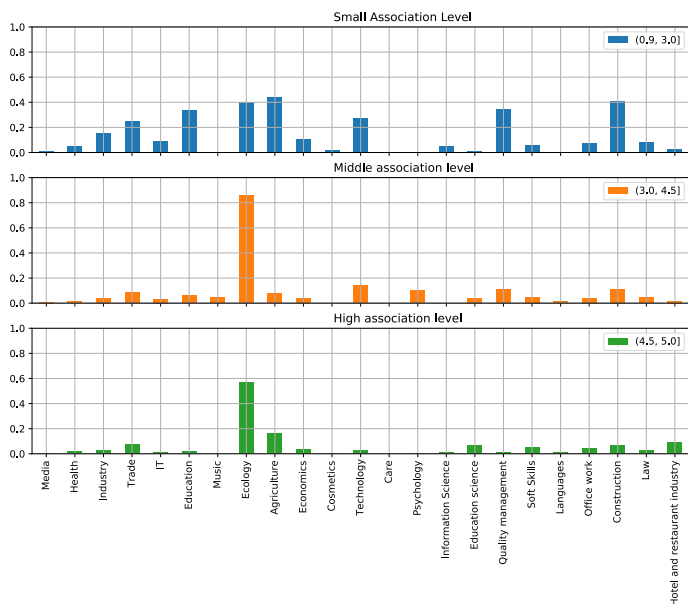


Fig. 2: Detailed analysis of CVET ads by categories and the amount of ads which have a small, middle or high association level with sustainability and green economy according to [BS22]

The small association level shows those advertisements having rather a relation to green technologies. Here, we find 44.5% of all ads in agriculture, 41% in Construction, 39.0% Ecology, 34.3% Quality Management, 33.7% in Education. While in IT 9% of all ads have a small association level, in Information Science it is only 4.8%.

Example entries for IT are Cloud Computing, Virtualization, Cloud-Computing, Power Supply, Public Transport. For Information Science we only find “fire protection”. Ecology, however, includes topics like “eco-management and audit scheme”, “ISO 14001” or “ISO 50001”. As we can see, for CVET advertisements, sustainability is not a widely considered for developing new skills for digitalization-related areas of industry.

In general, the number of ads for middle and high associated levels decreases, the only exception is the area Ecology, where 85.6% of all ads have a middle association level and 57.2% a high association level. Terms are related to “environment protection”, “environmental policy” etc. For IT, these values are 3,4% and 0,9%, for Information Science we find no middle association levels, but 0,9% have a high association level. Here, terms are mostly related to concepts like Smart Cities, Ecology, Resource management and Sustainability.

We find several education courses which are explicitly related to green economy: “Projektmanager:in Erneuerbare Energien”, “Elektromobilität”, “Nachhaltigkeit und gesellschaftliche Entwicklung”, or “CSR- und Nachhaltigkeitsmanager:in”.

Summarizing, while topics like “Green IT” and sustainability are important topics, we cannot find them in CVET ads. However, other sectors like Ecology, Agriculture, Technology and even Construction refer to green tools and concepts and use sustainability and green ecology widely to advertise their offers.

4.2 Digitalization in CVET ads

We will continue our analysis with details on those continuing education ads that have at least one CSO entity. Obviously, we find a lot of hits for IT and Information Science (90% and 97%). However, they are split into several smaller subject areas, see Figure 3 for the top 8 subject areas. We see that AI, Internet and Software Engineering are most important for IT, while Information Technology, Social Psychology and Data Mining are more important for the Information Science sector. We do not find any specifically green topics here.

If we compare this with sectors previously identified as green, in particular ecology (71% coverage) and agriculture (52%), see Figure 4, we see that the only two overlaps can be found in Internet technologies and human-computer interaction. We could consider AI as another topic relevant to both areas. Interestingly, this is also true for other sectors such as technology (71%), construction (67%) and education (79%).

As we have been able to show, both the coverage and the topics for digitalization and computer science in continuing vocational training differ from sector to sector, but we can see overlaps and differences between greener and less green sectors. We find several examples trainings which underline this connection, for example “Elektromobilität” and “Smart Factory Manager Produktion 4.0” which are clearly linked to AI and human-computer interaction.

5 Conclusion and Outlook

Our initial research on sustainability and digitalization in German continuing education advertisements could show that while most continuing education is concerned with computer science and digitalization (the lowest coverage is in the sector “Health” with 57%), the green economy is not a big topic in most advertisements. It is particularly interesting that while sustainability is a topic for specific tools and skills (low association), it is rarely used to “frame” and advertise courses and training (high association). We were able to show that training advertisements related to sustainability are mostly related to Internet technologies, AI, and human-computer interaction. This may be due to the fact that concepts first have to gain enough attention to be used in contexts other than their original ones.

However, our analysis has some serious limitations. First, we only considered very broad sectors to conduct our analysis. A more detailed analysis based on specific ads will allow a better analysis, and in particular a better analysis of the correlation between

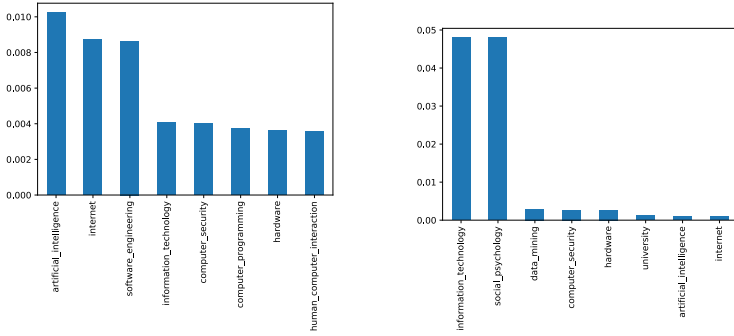


Fig. 3: Detailed analysis of training advertisements with at least one entity from a specific subject area (top 8) in CSO. Left: advertisements from the IT sector, right: Ads from the Information Science sector.

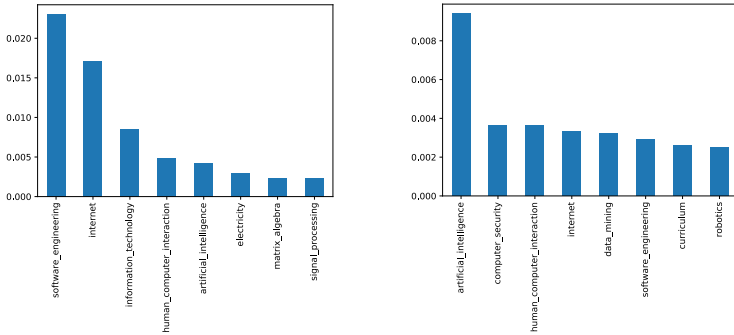


Fig. 4: Detailed analysis of CVET advertisements with at least one entity from a specific subject area (top 8) in CSO. Left: advertisements from the ecology sector, right: Ads from the agriculture sector.

specific terms. Second, we considered topics to analyze CSO terms. Combining topics with entities will provide a better understanding of the underlying topics for digitization. Third, a careful comparison of the two approaches based on word lists with other AI-driven methods could improve the quality of our results. However, we must also consider questions about the representativeness of the associations and results presented here, or a further contextualization of why the associations are surprisingly low.

In conclusion, our research is only a beginning, although it provides a unique perspective on a data resource that is currently rarely considered in research and shows that sustainability is rarely considered in continuing education in IT-related fields.

References

- [Be22] Beer, M.: Arbeits- und Lernaufgaben zur Entwicklung von Nachhaltigkeitskompetenzen in der Berufsbildung: Didaktische Überlegungen zur Berufsbildung für nachhaltige Entwicklung. *berufsbildung–Zeitschrift für Theorie-Praxis-Dialog* 76/1, pp. 56–58, 2022.
- [BS22] Binnewitt, J.; Schnepf, T.: Join us to turn the wor(l)d greener! – Investigating online apprenticeship advertisements’ reference to environmental sustainability. In: *Zum Konzept der Nachhaltigkeit in Arbeit, Beruf und Bildung - Stand in Forschung und Praxis*. 2022.
- [DD20] Dobischat, R.; Düsseldorf, K.: Organisation, Recht und Finanzierung der beruflichen Weiterbildung. *Handbuch Berufsbildung/*, pp. 579–595, 2020.
- [De18] Degenhardt, S.: Kompetenzen für eine digitalisierte Arbeitswelt–Anforderungen an Aus- und Weiterbildung. In: *Digitaler Wandel in der Sozialwirtschaft*. Nomos Verlagsgesellschaft mbH & Co. KG, pp. 259–272, 2018.
- [DIH17] Dutt, A.; Ismail, M. A.; Herawan, T.: A systematic review on educational data mining. *Ieee Access* 5/, pp. 15991–16005, 2017.
- [Do19] Dobischat, R.; Käßlinger, B.; Molzberger, G.; Münk, D.: Digitalisierung und die Folgen: Hype oder Revolution? *Bildung 2.1 für Arbeit 4.0?/*, pp. 9–24, 2019.
- [Ge19] Germany. Federal Ministry of Labour and Social Affairs (BMAS) Germany. Federal Ministry of Education and Research (BMBF): *National Skills Strategy: continuing education and training as a response to digital transformation./*, 2019.
- [GL21] Graf, L.; Lohse, A. P.: Advanced skill formation between vocationalization and academization: The governance of professional schools and dual study programmes in Germany. *Governance Revisited. Challenges and Opportunities for Vocational Education and Training*; Gonon, P., Bürgi, R., Eds/, 2021.
- [Ha23] Hao, X.; Li, Y.; Ren, S.; Wu, H.; Hao, Y.: The role of digitalization on green economic growth: Does industrial structure optimization and green innovation matter? *Journal of Environmental Management* 325/, p. 116504, 2023.
- [He16] Helmrich, R.; Tiemann, M.; Trotsch, K.; Lukowski, F.; Neuber-Pohl, C.; Lewalder, A. C.; Gunturk-Kuhl, B.: *Digitalisierung der Arbeitslandschaften: keine Polarisierung der Arbeitswelt, aber beschleunigter Strukturwandel und Arbeitsplatzwechsel*. Wissenschaftliche Diskussionspapiere, 2016.
- [KKM20] Kovalev, S.; Kolodenkova, A.; Muntyan, E.: Educational data mining: current problems and solutions. In: *2020 V International Conference on Information Technologies in Engineering Education (Inforino)*. IEEE, pp. 1–5, 2020.

- [KM22] Krebs, B.; Maier, T.: Die QuBe-Kompetenzklassifikation als verdichtende Perspektive auf berufliche Anforderungen, tech. rep., Wissenschaftliche Diskussionspapiere, 2022.
- [Kr18] Kreuzer, C.: Visualisierung der Opportunity Recognition-Kompetenz von Industriekaufleuten. *Zeitschrift für Berufs- und Wirtschaftspädagogik* 114/2, pp. 247–271, 2018.
- [MT13] Mohamad, S. K.; Tasir, Z.: Educational data mining: A review. *Procedia-Social and Behavioral Sciences* 97/, pp. 320–324, 2013.
- [Os16] Osborne, F.; Salatino, A.; Birukou, A.; Motta, E.: Automatic classification of springer nature proceedings with smart topic miner. In: *The Semantic Web–ISWC 2016: 15th International Semantic Web Conference, Kobe, Japan, October 17–21, 2016, Proceedings, Part II* 15. Springer, pp. 383–399, 2016.
- [RV07] Romero, C.; Ventura, S.: Educational data mining: A survey from 1995 to 2005. *Expert systems with applications* 33/1, pp. 135–146, 2007.
- [Sa20] Salatino, A. A.; Thanapalasingam, T.; Mannocci, A.; Birukou, A.; Osborne, F.; Motta, E.: The computer science ontology: A comprehensive automatically-generated taxonomy of research areas. *Data Intelligence* 2/3, pp. 379–416, 2020.
- [Sc22] Schiersmann, C.: Weiterbildungsberatung im Kontext der Nationalen Weiterbildungsstrategie: Finanzielle und strukturelle Aspekte. *Hessische Blätter für Volksbildung* 72/1, pp. 43–53, 2022.
- [St22] Steeg, S.: Die Wasserstoffwirtschaft in Deutschland: Folgen für Arbeitsmarkt und Bildungssystem; eine erste Bestandsaufnahme./, 2022.
- [Wa22] Wang, F.; Nian, V.; Campana, P. E.; Jurasz, J.; Li, H.; Chen, L.; Tao, W.-Q.; Yan, J.: Do ‘green’ data centres really have zero CO2 emissions? *Sustainable Energy Technologies and Assessments* 53/, p. 102769, 2022.
- [Zi12] Ziegler, P.: Zur Verwendung von Berufsinformation im Hinblick auf Matching in Deutschland und Österreich, tech. rep., AMS info, 2012.
- [Zi13] Zimmermann, K. F.; Biavaschi, C.; Eichhorst, W.; Giulietti, C.; Kendzia, M. J.; Muravyev, A.; Pieters, J.; Rodriguez-Planas, N.; Schmidl, R., et al.: Youth unemployment and vocational training. *Foundations and Trends® in Microeconomics* 9/1–2, pp. 1–157, 2013.

Ethical Challenges of Digitalization in Education

Reflections from a Theological Perspective

Matthias Meister¹

Abstract: The dominance of technology in Western societies is becoming increasingly visible in everyday life. It influences the way members of these societies perceive the world, information, and themselves. This influence has various consequences, both positive and negative. In this paper, some of the consequences are explained in order to reflect on how to counteract the negative tendencies resulting from the dominance of technology in academic education. The conclusion argues for the importance of skills and qualifications of ethical reflection and interdisciplinary discourse in academia to foster competent reflection on the dangers of the dominance of technology.

Keywords: Digitalization, Technological Society, Shift of Values, Ethics, Theology, Academic Education, Ethical Reflection, Interdisciplinary Discourse

1 Introduction

When I enter the classroom in the morning, the students' laptops are already open. Although the students are talking to each other, they all have their laptops out and their smartphones either on the table, tucked in their pockets, or even just in their hands, looking at the latest news from friends, family, or pictures on Instagram or movies on YouTube.

The room is flooded with stimuli. Students' minds are not only on the conversation they just had, but also on people who are not even physically there, and on information that is unnecessary and needless. They are inundated with information whose importance is more than doubtful; they can't help it.

As I teach theology, I always start the morning class with a prayer. Most of the time the students close their laptops and don't look at their smartphones, but often I must remind them. And during class? I don't even want to know how often people aren't taking notes, but are looking at memes, pictures, or news. This happens although I work at a private college, so the students pay for their studies.

Although there are numerous benefits from digital gadgets, its constant presence is one scary side of digitalization: Technology is not only everywhere, but it also determines and influences our behavior quite decisively.

¹ Interkulturelle Theologische Akademie Bad Liebenzell, Heinrich-Coerper-Weg 11, 73578 Bad Liebenzell, matthias.meister@ita-info.de.

Another scary side is how digital technologies alter the way we perceive contexts, knowledge, truth and ourselves.

This paper is not intended to highlight the influence of digitalization and mechanization on education, or rather teaching in a classroom on the pragmatic level. Every lecturer or teacher nowadays should be technically competent enough to cope with PowerPoint, various digital sources of the Internet, and knowledge transfer through technology. If this is not the case, there exist various sources that explain these technical applications better than I could ever do. However, I want to question the focus on the technical level. I'd like to argue that we might need to consciously set counter-accents against this overwhelming influence of digitalization in education, training, and adult education. Therefore, I want to focus on the moral and ethical level in this paper.

2 The Process of Digitalization of Life

The word 'digitalization' is used inflationary in the news. When talking to laypersons everybody thinks to know what digitalization is. However, I get the impression that the word 'digitalization' is just a symbol for a vague understanding that it is the distribution of computers and smartphones. However, digitalization is not equipping and operating areas of everyday life or work with computers. Digitalization is the measuring of the world. The fact that we measure all areas of life and express them in numbers is digitalization. It is the prerequisite for algorithms to function and for computers to be used profitably. That's why wearables are being marketed so strongly because human behavior can be recorded in numbers by means of sensors and the corresponding measurements, so that statistical values can be collected and optimizations for people can function. [Fr18]

This development is worrying because it is unconsciously influencing every area of life, as I have tried to illustrate in my brief introduction. The digital surveying of the world has several effects on human behaviors, society, worldviews, ideologies, and thus also on the practice of teaching. They are subliminal effects that influence how we perceive the world; how we deal with ourselves and other people; and what we consider to be true, beautiful, and good. [E11]

With the entry of the digital measurement of our lives, society, information, and the world come along the determination of how these data are to be categorized and evaluated. While in human history this was done by institutions like the family, the royalty and state's administration, stories, or by religion, this process was taken over by mathematics, science, technology, and the respective experts. The categorization and evaluation of human knowledge have therefore changed accordingly. During this still ongoing process, certain actions and procedures are re-evaluated. This has a significant influence on what is considered rational and generally valid in human lives. [E116] As this ongoing process runs mostly unnoticed, this must be reflected and questioned. Again and again, and increasingly so in the last 2 or 3 years, my theology students question values and ethical Christian standards, not on a theological and principled level, but on a pragmatic level.

They ask: ‘How does this work and is there a positive outcome for me?’ While these questions are not false per se, they are technical in a non-technical area of human life. On those occasions, I observe the shift from a philosophical, or rather in my specific case theological approach to life and society, towards a technical perception of life. This change has its origin in the increase of the dominance of technology in society which is a parallel development to the digitalization of life. [Va20]

3 The Influence of Technology on Society

Technology is created to make the world available, measurable, and controllable. [Ro18] Technology is much more than machines and computers. There are techniques of behavior, communication, decision-making, design, cleaning, cooking, and so on (and I don't just mean the equipment and tools that might be associated with those concepts). These concepts describe standardized procedures in enterprises, institutions, and private life, which exist to give us measurable and controllable successes and assistance so that life succeeds better. Humans have the tendency to automatically reach for and try to learn and master the technology that is most productive and efficient as this increases the chances of beneficial outcomes. Consequently, the highest and most desirable values are efficiency and productivity. [Ri11] However, those values are not fitting to evaluate long-term goals or determine the human meaning that action implies, but rather the means to achieve those goals. This overemphasizes “means over ends”. [Mo21]

The dominance of “means over ends” evokes a type of utilitarianism, [El16] ‘What works is good.’ It is a particular challenge, therefore, for general education to teach the students not only in purely technical terms how to do things but to focus on the ends. What are we doing these things for? Why learn professional techniques? And which areas of life are non-technological and will suffer if technology and digitalization are the criteria for our decision-making? A well-known example of this problem is the development of industrialization in the 20th century: The technological influences and possibilities had many great advantages for societies and states which led to a general increase in the standard of living. However, one of the ends of this process was the destruction of ecosystems and the exploitation of the Earth to an extent that potentially leads to the collapse of our planet. The dominance of means and short-term ends, which in turn became again “means” because the consumption of products was necessary to fuel the process of developing expensive technologies, was so strong that this long-term result, although it was anticipated by many scientists and intellectuals, had little influence on the decisions on how to shape further development in the past.

4 Human Freedom in a Technological Society

Another challenge I have observed in recent years is the view that freedom is a status, not a mandate. In my opinion, this also relates to the dominance of technology and

digitalization. Digitization in connection with the corresponding techniques serves, as described earlier, to make the world available and to use it as productively and efficiently as possible. This leads to the fact that when techniques work, we are accustomed to no longer having to worry about the processes. Water just comes out of the tap. Electricity flows from the socket. The car drives without our effort. The movie from Netflix or YouTube plays with a click from the couch. It just works. If it doesn't just work, then the technology is broken. There's something wrong with it and then there are usually bigger problems that cannot be solved individually. Today this technological reality is transferred to human life in general. Humans want their bodies to function like a machine. If I feed it the right food, get enough sleep, and give it enough breaks, then it should run like clockwork. [Ro18] Unfortunately, it doesn't. We have good days and bad days. This fact alone drives some of my students to despair. Especially the most motivated and strong students often despair of themselves in their studies. They want to be as efficient and productive as possible. Is this really the best thing to do? After all, we humans are not machines. However, what is even more important is the understanding that human freedom is not a status, but a call to action. [Va17] According to Ernst Cassirer (1874-1945), a German philosopher, who developed a philosophy of culture, technology is the symbolic form to translate human will into reality. [Ca85] As for Cassirer, freedom is not a form of human existence, but a mandate; technology is the means by which humans experience themselves as free. [Ca16] Through the dominance of "the means over the ends" the expression of the human will in reality and therefore human freedom that needs to be created and acted out constantly is lost. Technology itself which is the means to serve the human will and create freedom through servitude becomes the end. [Or87] This leads to an illusion of freedom as a status that pushes humans into passivity which is simply more convenient than actively shaping one's own life. [Ca16]

The common notions of freedom that have been shaped by the media and our culture are not inspired by a concept of positive freedom. The current definitions of freedom are usually concepts of negative freedom [Be69]: Humans must free themselves from power structures, cultural symbols, definitions of terms, understanding of roles, and so on. But liberated for which end? The concepts of positive freedom are missing. Technology also is a means of negative freedom. The goal of technology is negative freedom: To set people free from negative, hindering, or exhausting circumstances, so that other things can be achieved. [El90] But what is to be achieved? As soon as technology is no longer used to achieve other things, but only for its own sake, i.e. for entertainment, to fulfill work even more productively and efficiently, which then, in turn, enables us to work even more and so on, then the question has to be posed: What is this negative freedom for?

The absence of good definitions of positive freedom goes hand in hand with the process, of making the world available and thus controlling it as I have described above. If this process is applied to humans, human nature is diminished to something that can simply be controlled and changed so that it corresponds to human will, ideas, and desires. [Wa06] Therefore, human nature, especially when it causes weaknesses and problems – for example to fit in a community or a specific gender role – is ultimately considered in a way that it must be overcome. Because the technical world we build for ourselves takes place

primarily in our heads – virtual worlds, entertainment worlds, virtual identities, social media, and so forth – this technical world has to do with our bodies themselves only insofar as we have to use our fingers to operate the technology so that we can move in this world. Thus, there is a disembodiment that has become the reality of human life. For young people, the friend who is texting on WhatsApp is now just as present as the girlfriend sitting right next to him/her. Therefore, it would be impolite not to respond to the WhatsApp message. Perceptions of space are thus shifting into virtual realms and mixing with material realities. [Zo14]

5 Problems of the Dominance of Technology in Society

Ultimately, it has to do with a mixing of categories. The experience of disembodiment is nothing new. Everyone who listens to a good story at a campfire in the evening, or reads a thrilling book forgets himself/herself about it. All humans know this experience. Also, two people in love, who talk intensively with each other and forget about space and time, create a similar experience. The problem is on another level. By means of mathematical language in technology, it has become possible to make arbitrary wishful thinking at least visible, if not even experienceable. Thus, humans can produce whatever they want. [Mo21] We are not yet medically and technically ready to completely materialize these realities, but the trend is heading in this direction. Not only does transhumanism play a crucial role here, but also stories of Hollywood, Marvel and co. These stories are the mythologies of the current time. In one way or another, these stories tell that reality must be overcome in favor of wishful thinking. [Mo21] Reconciliation with individual or even human limitations and adverse circumstances is no longer a desirable option. Personal circumstance, even physical circumstance, is there to be controlled and made available to change. [Ro18] Fitness, wellness, and mindfulness industries thrive because this mythology has become commonly accepted. Overall, the trend is toward disembodiment, disembodiment that creates reality and is not mere phantasy. Digitalization, therefore, gives human beings the power to create a reality that has the possibility to bring ideas of the category of phantasy into reality. When reality is in crisis and this disembodiment is no longer possible or helpful the danger is that humans can lose their ability to cope with these crises. [Jo03]

In addition, due to the dominance of technology in every detail, responsibilities are perceived no longer. Everybody understands only a small part of reality that taking responsibility for their actions is almost impossible. If mistakes happen, it is not the mistake of an individual, but the failure of a system, or it is a mistake that happens because individuals or contexts do not fit into the system. [E118] Contexts of meaning which do not fit into the mathematical description of the world become suspect. Related to this, it can be observed that ambiguities in human language shifts in the interpretation of life that disagree with the images of man accepted within technological society are no longer considered seriously. As much as the dominance of technology allows plurality because it makes it possible for certain decisions or desires to become reality, even if only in virtual

space, digitalization and this dominance of technology do not allow any transcendence. Only immanence can be measured and controlled. Something that is greater and different from what digitalization can describe is by definition non-technological. This makes certain aspects of religion unavailable, i.e., when religion talks about a transcendent God. [Lo91] Ultimately, this shift from the responsibility of individuals to the responsibility of a system, results in the fact that commonly accepted morality no longer exists. Religion or other transcendent concepts are no longer commonly acceptable in a technological society. Therefore, a common moral foundation is missing. Additionally, there are two other reasons for missing responsibility as well:

First, the individual comes into a situation of powerlessness: his individual deeds no longer change anything in the course of the great whole. Secondly, there is no longer any need to take responsibility for one's own deeds. The reason for failure is the system that has not worked. [E190]

Thus, there is no longer any generally accepted moral right or wrong in a technologized society. Only the means that seem to promise the greatest effect are applied. This could be witnessed quite nicely in recent years: In the corona policy, the measures that promised the best effect according to the corresponding computer models were implemented. The fact that these computer models were only models and could only depict the effects according to the programmers' specifications is problematic, because there is no longer a uniform goal or a viable image of man. These statistical models applied in computer simulations are therefore not neutral. The way these specifications are chosen reflects certain value decisions. However, there was no public moral discussion in mainstream media about problems this way of conduct may cause. This emphasis on technology as the criteria for decision-making results in the perspective that human beings become a statistical mass and computer models and statistical analyses are believed more than the so-called "common" sense. When individuals warned that it cannot do children any good to isolate them for weeks on end and that this would create greater problems than the benefits of containing the pandemic, it was not heard because this cannot be mapped in a computer program. The inaccuracies are too great and therefore this effect cannot be statistically calculated. This example is not a critique of the actual policy. The political level is not in focus. It shows simply that the necessary moral discussions have become so difficult in a technological society that they are avoided and impossible. This is one of the reasons why nobody wants to take the moral consequences and responsibility for such negligence and belief in technology. Because who should be able to do that? The programmers, the physicians, the politicians, or those who implement the measures? We are all just a wheel in the big system and cannot bear any responsibility for the whole. The technical system is too interconnected and too intertwined. [Cé18]

At the same time, the technical system is always pulling forward into the future. [Jo03] Because every detail, every optimization is made more efficient by the latest technology, the newer gadgets, and the latest scientific findings, within the value system of a technological society the application of the latest technologies become more meaningful. Efficiency and productivity as the highest values in the system prioritize the

implementation of better technology. This changes the reality of human lives. The newer and the less known is actually the better. [El90] This was different in the past. When technology was more handicraft than mathematical optimization, the elderly were the ones who managed life better. Accordingly, traditional wisdom today is no longer what to strive for, but what to question and test. Therefore, elderly people and their wisdom are to be questioned and examined. A fortiori, ethical standards and assertions about humans and life that are thousands of years old are difficult to accept and certainly not credible.

6 Conclusion

I will stop here. There are many other effects of digitalization and the dominance of technology that could be listed but that would only take us into unnecessary detail. I think my point has been made. The need to question thinking on a purely pragmatic, technological level and the need to question our technological developments on an ontological, metaphysical, and ethical level is more urgent than ever before in human history.

Of course, the listed effects are only trends. They are not irreversible, nor do they affect everyone in the same way. But as technology penetrates more and more areas of our daily lives, certain developments are intensifying and leading to the value shifts I have mentioned above. These are trends and generalizations, but they need to be countered by a greater emphasis on ethical reflection on the purposes and ends of technology. What does the use of digital tools mean for human societies and individuals in the long run? Are the results really worth the consequences and costs of achieving them? Critical ethical reflection and interdisciplinary discussions are therefore imperative in academic education.

Bibliography

- [Be69] Berlin, I.: *Four Essays on Liberty*. Oxford University Press, Oxford, London, New York, 1969.
- [Ca85] Cassirer, E.: *Form und Technik*. In (Krois, J. M.; Orth, E. W. Hrsg.): *Symbol, Technik, Sprache. Aufsätze aus den Jahren 1927 - 1933*. Meiner, Hamburg, S. 39–92, 1985.
- [Ca16] Cassirer, E.: *Vom Mythos des Staates*. Meiner, Hamburg, 2016.
- [Cé18] Cérézuelle, D.: Jacques Ellul: From Technique to the Technological System. *the Ellul Forum* 62, S. 59–66, 2018.
- [El90] Ellul, J.: *The technological bluff*. William B. Eerdmans, Grand Rapids, Mich., 1990.

- [El11] Ellul, J.: *The Technological Society. A Penetrating Analysis of our Technical Civilization and of the Effect of an Increasingly Standardized Culture on the Future of Man. With an Introduction by Robert K. Merton.* Vintage Books, New York, NY, 2011.
- [El16] Ellul, J.: *Presence in the Modern World. A New Translation.* Wipf and Stock Publishers, Eugene, OR, 2016.
- [Fr18] Friedrich, J. P.: *Die Herkunft der digitalen Welt. Zeitschrift für Kulturphilosophie* 1/12, S. 11-19, 2018.
- [Jo03] Jochum, U.: *Kritik der Neuen Medien. Ein eschatologischer Essay.* Wilhelm Fink Verlag; Fink, München, 2003.
- [Lo91] Lovekin, D.: *Technique, discourse, and consciousness. An introduction to the philosophy of Jacques Ellul.* Associated University Presses, London, Toronto, 1991.
- [Mo21] Moyses, A. J.: *The Art of Living For a Technological Age. Toward a Humanizing Performance.* Fortress Press, Minneapolis, MN, 2021.
- [Or87] Orth, E. W.: *Zum Begriff der Technik bei Ernst Cassirer und Martin Heidegger. Phänomenologische Forschungen* 20/20, S. 91–122, 1987.
- [Ri11] Riis, S.: *Zur Neubestimmung der Technik. Eine Auseinandersetzung mit Martin Heidegger.* Francke, Tübingen, 2011.
- [Ro18] Rosa, H.: *Unverfügbarkeit.* Residenz Verlag, Wien, Salzburg, 2018.
- [Va17] Van den Heuvel, S. C.: *Bonhoeffer's Christocentric Theology and Fundamental Debates in Environmental Ethics.* Pickwick Publications, Eugene, OR, 2017.
- [Va20] Van Vleet, J.E. ; Rollison, J. M.: *Jacques Ellul. A Companion to His Major Works.* Cascade Books, Eugene, OR, 2020.
- [Wa06] Waters, B.: *From Human to Posthuman. Christian Theology and Technology in a Postmodern World.* Ashgate, Aldershot, 2006.
- [Zo14] Zorn, I.: *Selbst, Welt und Technologieverhältnisse im Umgang mit Digitalen Medien.* In (Marotzki, W.; Meder, N. Hrsg.): *Perspektiven der Medienbildung.* Springer VS, Wiesbaden, S. 91–120, 2014.

Enhancing educational insights: A real-time data analytics stack for project-based learning

Gian-Luca Gücük¹, Dejan Simic¹, Stephan Leible², Tom Lewandowski², and Emir Kučević²

Abstract:

This paper presents a real-time data analytics (DA) stack designed for a project-based course utilizing Jira for project management at a university. The DA stack follows an Extract, Transform, and Load process to visualize students' usage data within dashboards. The DA stack supports course management by providing insights into students' activities and progress. We demonstrate the DA stack's effectiveness through an evaluative case study, which was found to support course objectives and foster improved behavioral adaptations from lecturers to students. Furthermore, we propose a generic DA stack for generalizing and adopting it for similar applications, considering the extensibility and maintainability inherent in the open-source tools used. Moreover, we provide the GitHub repository to view our source code. This study contributes to the relatively underexplored field of real-time learning analytics and offers a starting point for the customization and adoption of the proposed DA stack in different educational contexts.

Keywords: Learning analytics, educational data mining, data analytics, project-based learning, business intelligence

1 Introduction

Recognized as a potent catalyst for student learning in higher education, project-based learning formats have garnered considerable attention [Gu20]. The relentless march of digitalization has further intensified the demand for digital learning modalities, particularly in the realm of higher education [ZR21]. Notably, project-based courses, e.g., in software engineering, that leverage project management frameworks such as Scrum have demonstrated a profound impact on the quality of software produced and the overall learning trajectory of students [Pe11]. The utilization of project management tools like Jira in such courses has proven to significantly enhance student productivity [RF21].

Given the voluminous data generated by these software tools, they present an excellent opportunity to gauge student activity through the application of data analytics (DA). This

¹ Universität Hamburg, Department of Informatics, Vogt-Kölln-Straße 30, 22527 Hamburg, Germany, {gian-luca.guecuk,dejan.simic}@studium.uni-hamburg.de

² Universität Hamburg, Department of Informatics, Vogt-Kölln-Straße 30, 22527 Hamburg, Germany, {stephan.leible,tom.lewandowski,emir.kucevic}@uni-hamburg.de

paper endeavors to devise a generic DA stack capable of analyzing software related to project-based courses, including project management tools (e.g., Jira), version control systems (e.g., Gitlab), and documentation systems (e.g., Confluence). Such an approach would offer insights to lecturers and students, enabling them to monitor the activity within these courses. Consequently, lecturers could offer tailored assistance based on the extracted and analyzed data to student groups exhibiting lower activity or facing challenges with specific tasks. The proposed DA stack also aims to deliver real-time insights, a marked improvement over most strategies that typically analyze data at periodic intervals [SKO18], often delaying evaluation until the completion of courses and impeding immediate interventions.

The overarching terms for analyses of this nature are learning analytics (LA) and educational data mining (EDM), both burgeoning areas of research [Va15]. Despite the potential benefits of LA for student learning progression, its adoption has been somewhat sluggish [Pr19]. The slow uptake of LA and EDM can be attributed to the complexity of implementing requisite technologies such as databases and business intelligence software, coupled with the limited technical expertise among lecturers, especially outside of the computer science discipline [If20]. Furthermore, interpreting the generated data effectively necessitates adequate data literacy, which is an important future skill but also not universally presented among key stakeholders like teaching staff [GD12].

This study's context originates from a project-based software engineering course conducted at a university. Here, students tracked their project activities via Jira and documented their time spent on tasks. The initial conception of developing a DA stack was limited to using Jira as a data source, but the design is intended to be generic, thereby accommodating other tools. This research proposes and assesses an approach to conducting analytics during an ongoing project-based course based on data generated by students using their respective digital tools - Jira, in this case. A DA stack is planned, implemented, and evaluated to fulfill this objective. This facilitates the interpretation of ongoing courses by lecturers and students, enabling data-informed decisions to achieve their distinct (learning) goals.

The collected data is displayed on business intelligence dashboards, offering a user-friendly interface for lecturers. The stack and dashboard's evaluation encompasses a proof-of-concept to verify the stack's architecture and a qualitative assessment with the lecturers. To affirm the generalizability of our work, each component of the DA stack was scrutinized for its transferability. Consequently, we explore the following research question (RQ): *How can an adaptable real-time data analytics stack be constructed based on student usage data?*

2 Related research

The application of data mining to enhance education is encapsulated within the research domains of LA and EDM [Sá21]. A comprehensive study by Papamitsiou and Economides [PE14] elucidated the objectives of researchers and their contributions to LA and EDM. Our paper aligns with their category of "Increase (self)-reflection and (self)-awareness",

which emphasizes enhancing lecturers' awareness, identifying underperforming students, and evaluating visualizations to inform students about their progress and performance. It also encourages proposing modifications to student behavior to improve their work results.

Beyond the scope of Papamitsiou and Economides' work [PE14], our endeavor aims to make visualizations accessible to students for their own data analysis. This approach to self-assessment echoes the work of Ifenthaler et al. [ISK22] and has demonstrated positive implications for both students and course outcomes. An aspect that is further of interest is the frequency of analysis refresh. Shimada and Konomi [SKO18] delineated three analysis frequencies: annual, weekly, and real-time feedback. Annual feedback involves the analysis of learning logs for the past year or term. Weekly feedback centers on predicting students' academic performance based on attendance and learning logs. Real-time feedback is employed in virtual learning environments. Studies exemplifying the use of real-time LA include Minovic and Milovanovic [MM13], who tracked students' learning progress while utilizing an education game, evaluated via a visualization tool. Raj and Renumol [RR22] proposed a learning path recommendation system based on artificial intelligence powered by real-time LA. However, most existing proofs-of-concept in real-time LA are designed for a specific purpose and not prepared for adaptation to new technical infrastructures.

Several studies have proposed the use of Extract, Transform, and Load (ETL) processes to gather learning data. Barriocanal et al. [Ba18] utilized data from the Jira and GitHub Rest API in problem-based learning conducted with agile methods. They extracted the data into a MySQL database via Jupyter notebooks and visualized them within the notebooks. Thereby, the work by Drachsler and Greller [DG16] emphasizes the critical role of privacy and data security in the design, implementation, and utilization of EDM and LA systems, reminding researchers and practitioners of the ethical considerations inherent in handling student data. However, the frequency of these analyses remains unspecified. Given the software utilized, a real-time approach appears improbable as the processing of notebooks must be manually triggered. Sclater et al. [SBW15] proposed a reference model for an open learning analytics (OLA) approach, allowing organizations to develop their own architecture. The concept of constructing a LA approach with open-source software, facilitating easy model extension, is termed OLA [MCG20]. It is a research field founded by the Society for Learning Analytics Research (SoLAR) to support the integration of diverse LA techniques and assist researchers, developers, and technologically proficient lecturers in enhancing these platforms with their own tools and methods [Si11]. Overall, our work can be categorized as creating a generic, extensible, and maintainable solution towards an OLA platform.

3 Method

3.1 Case description

The DA stack was developed and evaluated in two iterations within the context of a project-based course with the course lecturers, teaching staff, and 22 students at a university. The

course is conducted in a hybrid mode with a four-week presence phase at the university, allowing lecturers to supervise students' activities and provide immediate assistance or suggest behavioral modifications as required. Subsequently, the course transitions to an online phase facilitated through the open-source web conference system, BigBlueButton. The course's primary objective is to cultivate students' ability to engage in collaborative, decentralized software development utilizing digital tools. To facilitate this objective, students employ Jira for project management, which helps in coordinating software development, task decomposition among project teams, and task assignment to individual students during sprint planning.

Furthermore, the time spent on each task is logged in Jira to ensure a requisite amount of work is dedicated to the project for successful completion. However, during the online phase, despite scheduled plenary sessions, supervising students' activities becomes challenging for the lecturers. Moreover, reduced communication among students make productivity comparison difficult. An existing solution, employing data extraction from the Jira database, transformation through the low-code ETL tool Pentaho Spoon, and data warehousing for subsequent business intelligence with Microsoft PowerBI, is currently in use. However, this solution is complex, requires manual refreshes, and is challenging to augment with new tools, making it unsustainable and limiting timely analysis and extension efforts.

3.2 Research approach

To address these challenges and provide a universally applicable solution for similar project-based courses, we propose the construction of a DA stack. This stack allows data extraction from various digital tool sources, data processing, and representation in business intelligence dashboards. The stack is designed with an emphasis on scalability and extensibility to accommodate a diverse range of digital tools and data formats. The implementation was evaluated through demonstrations to the stakeholders (lecturers and students), and further improvements were suggested. For this, we adopted a prototyping approach [WH07], which began with the definition of requirements and goals, followed by iterative development, demonstration, and evaluation of the DA stack. The evaluation was conducted through continuous feedback from the stakeholders gathered during their utilization and observation of the DA stack to ensure its usability and effectiveness in a real-world educational setting.

4 Results

4.1 Data analytics stack

The DA stack, in the context of this study, fundamentally relies on data collected via Jira Webhooks. The detailed architecture of the DA stack is depicted in Figure 1. The process adheres to the ETL approach, responsible for data extraction from Jira (a), (b), and (c),

transformation using Apache Hop (d), and loading into the data warehouse (e), culminating in the visual representation of data in Metabase (f). Notably, this process operates in real-time, continuously updating the data. The functions and roles of the individual tools within this DA stack are elaborated in Table 1.

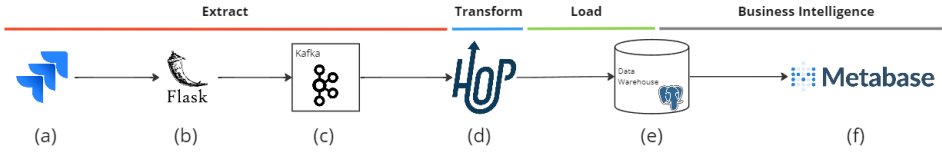


Fig. 1: Data analytics stack with Jira Webhooks

Tool	Description
(a) Jira Webhook	Selected events trigger the webhook to send a corresponding payload to the Flask webservice (b).
(b) Flask Webservice	Accepts the payload from the Jira Webhook (a) and transmits it to the Kafka message queue (c) using the Kafka producer API.
(c) Apache Kafka	Functions as a message queue that stores events produced by Jira in JSON format.
(d) Apache Hop	Extracts events from Kafka (c) via the Kafka consumer API, transforms the data into a suitable format, and loads it into the data warehouse (e), leveraging a low-code visual programming interface.
(e) PostgreSQL	Serves as a data warehouse, housing the data in a star schema format, which forms the basis for subsequent business intelligence operations.
(f) Metabase	Presents the data visually, refreshing every 60 seconds, and facilitates the creation of data visualizations through an intuitive query builder.

Tab. 1: Tools of the data analytics stack

The choice of individual tools was primarily driven by their compatibility with each other. Additionally, we prioritized tools with a robust community presence and comprehensive documentation to ensure long-term maintainability and availability of support (e.g., material and contact points). Furthermore, we focused on user-friendliness, particularly for non-technical lecturers who interact with the system regularly, e.g., by using low-code approaches.

4.2 Demonstration and evaluation

This section is dedicated to the demonstration and critical evaluation of the constructed DA stack, achieved by presenting the architecture to the stakeholders, implementing it in a practical context, and soliciting feedback. The stakeholders encompass three distinct groups: the course lecturers, the teaching staff (e.g., student tutors), and the students. The demonstration primarily involved showcasing the resulting dashboards during regular sprint planning sessions conducted live. The central objective was to evaluate the DA stack's utility in facilitating a more nuanced understanding of student activities, generating insightful suggestions for behavioral modifications, and implicitly testing the underlying technologies' robustness. One crucial aspect of the demonstration was discerning the boundaries of data

extraction and visualization. Certain analyses, such as tracking the "time spent" by students during live sessions, were perceived as overly intrusive by stakeholders, thereby underscoring the need for balancing informative insights with respectful privacy considerations.

The DA stack proved instrumental in identifying groups not adhering to established conventions, such as labeling Jira Issues in a prescribed manner, assigning tasks appropriately, or conducting sprint planning within Jira. It was also possible to identify groups that were having problems with certain tasks because they were working on them for a long time based on the visualized data in the dashboards, which allowed individuals from the teaching staff to proactively reach out to them. Consequently, the lecturers and teaching staff could rectify these deviations by providing immediate feedback, thereby facilitating the students' progression and proficient use of the digital tools deployed in the project-based course. However, with the student course data now available, it was also determined that the communication structure should be adjusted in the future, e.g., by having a regular, standardized feedback meeting at the end of the sprints based on the group data. Overall, the demonstration and evaluation emphasized our approach's utility and the DA stack, as the real-time analysis functioned as envisaged, providing live insights into students' activities.

4.3 Generalizing the data analytics stack

This DA stack was primarily designed with a broad application in mind. This design philosophy allows the integration of an unlimited number of sources with minimal integration effort when the technical infrastructure is initially established. As such, every stage delineated in Figure 1 can be supplanted or augmented with other tools or technologies. The following details considerations for adopting our DA stack to similar use cases. The initial step (a) can be extended by any desired webhook. In this specific instance, the Flask webserver (b) is deployed as a docker container and is not accessible via the web. Nonetheless, since Jira (a) is also deployed as a docker container within the same network, it can directly address the webserver (b) within the docker network. Consequently, if services that are not self-hosted are being used, the webserver (b) needs to be web-accessible to receive webhooks. The Flask webserver (b) is fully generic and can be replaced by any other webserver. It executes a plain Python script, which forwards every incoming webhook in JSON format to a Kafka message queue (c) using the Kafka Producer API. Kafka (c) stores the webhooks and facilitates rapid forwarding when invoked. Any other message queue or similar tool or technology can be selected in place of Kafka.

The extraction is executed by Apache Hop (d) via the Kafka consumer API. Hop boasts a visual low-code programming interface, facilitating the intuitive construction of ETL pipelines. Although Hop was primarily chosen for its user-friendliness, numerous other (open-source) ETL tools are also viable options. For accommodating new sources, new pipelines in Hop must be created, as they are specific to a particular structure. The data warehouse (e) is built on PostgreSQL, a widely utilized database management system. Its performance is adequate for use as a data warehouse in this specific use case. The only

requirement is that the data warehouse structure, such as a star schema, must be established in advance, for instance, by an ETL tool like Hop. However, it could potentially become a bottleneck for more substantial data volumes. Metabase (f) was selected for business intelligence due to its user-friendly query builder and web-based operation. Any other business intelligence tool capable of connecting to PostgreSQL can be selected based on user preferences. Our generic model is depicted in Figure 2. To foster insights and provide an adaptable and evaluated DA stack as starting point, we have made our source code available on GitHub to facilitate the generalizability of our approach².

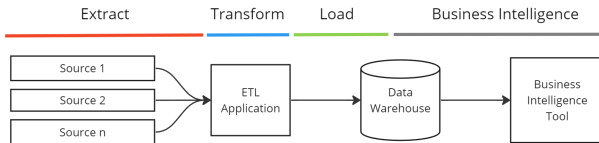


Fig. 2: Generic DA stack

It is noteworthy that the webserver and message queue forwarding are deliberately not incorporated in the ETL application in Figure 2. They are optional tools explicitly used in our process, representing one of many possible implementation strategies. However, other ways exist to model this subprocess, but they are beyond the scope of this work.

5 Discussion and conclusion

In conclusion, this study presents a DA stack utilized in a project-based course at a university where Jira is employed for project management. The stack operates in real-time, adheres to an ETL process, and is built on webhooks from Jira. These webhooks are received by a webserver and dispatched to a Kafka message queue. The data is then extracted and loaded into Apache Hop, which handles its transformation and loading into a PostgreSQL data warehouse. Ultimately, this data is visualized on business intelligence dashboards in Metabase, refreshing every 60 seconds. Moreover, this study delineates the evaluation of the DA stack and provides guidance on how to generalize it for similar application scenarios.

Our RQ is addressed by the DA stack depicted in Figure 1. Furthermore, we demonstrate and evaluate our DA stack approach as a data-driven provider of insights into students' activity and progress for stakeholders of a project-based course. Based on enhanced suggestions for behavioral adaptations and qualitative feedback, particularly from the teaching staff, the DA stack is recognized as valuable support to foster the achievement of course objectives. The generalizability of our solution is addressed in subsection 4.3 and Figure 2. As previously discussed, any tool in the DA stack can be replaced with other appropriate tools, considering compatibility aspects. We also provide a GitHub repository containing our approach as a starting point for adopting the DA stack and tailoring it to similar use cases.

² <https://github.com/GLucaDEV/ola-realtime-stack>

The field of real-time LA is relatively unexplored [SKO18]; hence, this study contributes to this research area with the proposed DA stack. This work can be classified as OLA [MCG20], presenting a strategy that offers extensibility—as the selected tools are open-source and inherently support various sources—and maintainability, given the minimal need for actual coding. Stakeholders are provided with a DA stack that leverages the potential of LA in education, as evidenced by its implementation and evaluation in our project-based course. Lecturers and teaching staff found gaining a comprehensive overview of their ongoing course beneficial, especially during student group meetings, whose outcomes are vital for future work. We posit that this stack is suitable for adaptation to other project-based courses that require timely analysis or wish to have real-time access to the most current data.

However, this study is not without its limitations. Although we offer a DA stack approach that is extensible and adaptable to diverse use cases, it is not a solution that can be employed entirely without prerequisite technical knowledge. The construction of the DA stack’s generic model is predicated on the induction of the utilized stack, implying that new challenges may arise when implemented in new technical infrastructures not previously identified. We propose that expanding the DA stack to include version control systems like GitHub or documentation software like Confluence could offer more detailed oversight of course progression and student activities. Furthermore, process mining alongside business intelligence could enhance course processes, endorse specific actions, identify bottlenecks, and facilitate the derivation of measures to mitigate them [AP22].

Bibliography

- [AP22] AlQaheri, H.; Panda, M.: An Education Process Mining Framework: Unveiling Meaningful Information for Understanding Students’ Learning Behavior and Improving Teaching Quality. *Information* (Switzerland), 13(1), 2022.
- [Ba18] Barriocanal, E.; Sicilia, M.; Sánchez-Alonso, S.; Cuadrado, J.: Agile methods as problem-based learning designs: setting and assessment. In: *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*. New York, NY, USA, pp. 339–346, 2018.
- [DG16] Drachsler, H.; Greller, W.: Privacy and Analytics: It’s a DELICATE Issue a Checklist for Trusted Learning Analytics. In: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM Press, New York, NY, USA, pp. 89–98, 2016.
- [GD12] Greller, W.; Drachsler, H.: Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, 15(3):42–57, 2012.
- [Gu20] Guo, P.; Saab, N.; Post, L. S.; Admiraal, W.: A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research*, 102, 2020.
- [If20] Ifenthaler, D.: Learning Analytics im Hochschulkontext – Potenziale aus Sicht von Stakeholdern, Datenschutz und Handlungsempfehlungen. In (Fürst, R. A., ed.): *Digitale Bildung und Künstliche Intelligenz in Deutschland: Nachhaltige Wettbewerbsfähigkeit und Zukunftsaagenda*, pp. 519–535. Springer, Wiesbaden, 2020.

- [ISK22] Ifenthaler, D.; Schumacher, C.; Kuzilek, J.: Investigating students' use of self-assessments in higher education using learning analytics. *Journal of Computer Assisted Learning*, 39(1):255–268, 2022.
- [MCG20] Muslim, A.; Chatti, M.; Guesmi, M.: Open Learning Analytics: A Systematic Literature Review and Future Perspectives. In (Liu, S.; Pinkwart, N., eds): *Artificial Intelligence Supported Educational Technologies*, pp. 3–29. Springer, Cham, 2020.
- [MM13] Minović, M.; Milovanović, M.: Real-time Learning Analytics in Educational games. In: *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*. New York, NY, United States, pp. 245–251, 2013.
- [Pe11] Persson, M.; Kruzela, I.; Alder, K.; Johansson, O.; Johansson, P.: On the use of scrum in project driven higher education. In: *Proceedings of the International Conference on Frontiers in Education: Computer Science and Computer Engineering*. 2011.
- [PE14] Papamitsiou, Z.; Economides, A.: Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Journal of Educational Technology & Society*, 17(4):49–64, 2014.
- [Pr19] Prieto, L.; Rodríguez-Triana, M.; Martínez-Maldonado, R.; Dimitriadis, Y.; Gasevic, D.: Orchestrating learning analytics (OrLA): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australasian Journal of Educational Technology*, 35(4), 2019.
- [RF21] Reha, M.; Fai, V.: Project Management Tools in the Classroom: Using the Atlassian Tool Suite in the Classroom. *Journal of Instructional Research*, 10:85–92, 2021.
- [RR22] Raj, N. S.; Renumol, V. G.: An improved adaptive learning path recommendation model driven by real-time learning analytics. *Journal of Computers in Education*, 2022.
- [Sá21] Sáiz-Manzanares, M.; Rodríguez-Díez, J. J.; Díez-Pastor, J. F.; Rodríguez-Arribas, S.; Marticorena-Sánchez, R.; Ji, Y. P.: Monitoring of Student Learning in Learning Management Systems: An Application of Educational Data Mining Techniques. *Applied Sciences*, 11(6), 2021.
- [SBW15] Sclater, N.; Berg, A.; Webb, M.: Developing an open architecture for learning analytics. Technical report, *EUNIS Journal of Higher Education*, 2015.
- [Si11] Siemens, G.; Gasevic, D.; Haythornthwaite, C.; Dawson, S.; Buckingham Shum, S.; Ferguson, R.; Duval, E.; Verbert, K.; J. Baker, R. S.: Open Learning Analytics: an integrated & modularized platform Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques Project Overview, 2011.
- [SKO18] Shimada, A.; Konomi, S.; Ogata, H.: Real-time learning analytics system for improvement of on-site lectures. *Interactive Technology and Smart Education*, 15(4):314–331, 2018.
- [Va15] Vahdat, M.; Ghio, A.; Oneto, L.; Anguita, D.; Funk, M.; Rauterberg, M.: Advances in learning analytics and educational data mining. In: *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. pp. 297–306, 2015.
- [WH07] Wilde, T.; Hess, T.: *Forschungsmethoden der Wirtschaftsinformatik Eine empirische Untersuchung*. *Wirtschaftsinformatik*, 49(4):280–287, 2007.
- [ZR21] Zawacki-Richter, O.: The current state and impact of Covid-19 on digital higher education in Germany. *Human Behavior and Emerging Technologies*, 3(1):218–226, 2021.

Monitoring of Digitization and Sustainability on Twitter

Informatik 2023 Festival - Skills'23

Ali Vahdatnia, Danoosh Peachkah

Abstract: In today's era of technological advancements, particularly in the realm of cyber technologies like the internet and social media platforms, conducting research on the reception of Digitization and Sustainability has become more accessible, affordable, and convenient. Among these platforms, Twitter stands out as a pioneering and appealing medium for such research endeavors. Therefore, we selected Twitter social media platform to analyze public opinions expressed through tweets and how these concepts are connected to skills and tools and the labor market. By gathering data related to specific keywords and utilizing SpaCy for sentiment analysis in our customized pipeline, we observed that a majority of tweets originating from Germany exhibited a positive sentiment towards sustainability strategies but expressed more negative views, especially regarding methods of leftover disposal. Additionally, we concluded that even a proficient analyzer may generate irrelevant results if the keywords being searched inherently possess negative sentiments.

Keywords: Digitization, Sustainability, Sentiments, Germany, Twitter

1 Introduction

It is known that Twitter as a micro-blogging platform is used for journalism and is a platform to share public opinion, which means that the users can state their thoughts in the form of "Tweets". [Cp20] Twitter has had a significant impact on research and is widely used in research and business, because of both being more text-based (or less visual content) compared to other networks and providing an API allowing researchers to access the public Twitter data based on their access level. On the other hand, jobs are no different from other subjects discussed on the Internet, meaning that people have different opinions about sustainability and digitization whether very favorable or rather disliked. In fact, as the author states that users expressed their sadness and happiness regarding these in the context of a text, as an effect of rejection or a specific decree. Whereas other Tweets are only a combination of different meaningless text like Memes. [Rn19] This context extends towards different industries like finance and the economy. Throughout the journey of Tweet analysis, different approaches for research were conducted, like using ML algorithms such as SVM, KNN, etc.

New technologies like web crawling and content scraping have made collecting people's opinions faster. Whereas, Machine learning and AI have improved knowledge extraction from internet data, making it more efficient and accurate than traditional methods.

Based on the information above, we got motivated to conduct a study leading to formulate the following research questions in this paper:

RQ1. The distribution of sentiments around states of Germany regarding sustainability and digitization and if this data leads to results related to the labor market.

RQ2. What are the most discussed relevant topics and how biased the related sentiments were to positive or negative poles? Can we infer tools and skills?

From previous experience of mining tweets using the 5-digit labels for KldB (Klassifikation der Berufe) [Bf11], it was discovered that not enough data regarding sustainability could be gathered within the scraped data based on the relevant topics as the data was mainly generated using occupation terms which are not relevant to the topic, hence the need to find another path to obtain the needed relevant data emerged. Therefore, in order to increase our accuracy, we retried scraping but this time using a set of keywords that suggest digitization and sustainability mainly from the related lists of BIBB¹. [BS22] Lastly, figured out the opinion of the audiences, specifically from Germany, on Twitter. This would give us some potentially insightful and useful data regarding the subject.

As for the starting point, some relevant words were first extracted and cleaned for scraping. Then a set of operations was conducted to analyze the stats extracted from the scraping based on tweet text sentimental analysis. Using the extracted raw data shows a hefty number of discoveries. In fact, we discovered that the majority of tweets originating in Germany had good perspectives toward sustainability efforts but voiced more negative opinions, particularly about residual disposal techniques. Furthermore, we determined that even a skilled analyzer might produce irrelevant results if the terms being searched already had negative connotations.

2 Background

Web scraping is a technique used to collect data from the internet, often automated using scripts. Web scraping extracts data from webpages, social media, and other online sources and converts it into useful information, whether it is stored data or real-time streaming data. [Mn13]

Twitter, established on March 21, 2006, is a popular social platform that allows users to post limited-size messages called Tweets. Tweets can contain not only visible text but also additional information like Metadata. [Cp20] Twitter is not just a platform for individual

¹ <https://www.bibb.de/de/153062.php>

status sharing but also provides additional functionalities through its API. An API, or Application Programming Interface, is a way to publicly share data from an application. The Twitter API is a scalable and flexible solution that allows users to authenticate and access various Tweet data within defined limitations. [MK09]

Sentiment analysis is important for analyzing products in various industries. Microblogging platforms like Twitter are commonly used for opinion analysis. Sentiment analysis techniques employ Natural Language Processing (NLP) to assess unstructured data, such as Tweets, and determine whether the sentiment is positive, negative, or neutral. [Ms17]

3 Literature review

Twitter, a popular social platform with vast databanks of unstructured information, is utilized for sentiment analysis. The API from Twitter is used with two approaches: a Dictionary-based method categorizes user information into positive, negative, or neutral classes using lexicon analysis, while a Machine Learning approach employs AI models like K-Nearest Neighbors (KNN) and Naïve Bayes to extract information from unstructured data. Challenges include multilingual divisions, limited character count for opinions, and the use of discrete hashtags. Preprocessing and text cleaning are essential for accurate sentiment calculation. [Cp20]

Sentiment analysis on Twitter is a crucial topic for understanding public opinion about a company. It provides valuable insights to enhance product quality by analyzing criticisms on social media. Sentiment analysis is a widely used technique to extract people's opinions from online platforms. In this study, seven levels of sentiment were extracted, ranging from strongly positive to weakly negative. Tweepy and TextBlob were utilized as Twitter API and text processing tools, respectively, for tasks such as tokenization, stemming, lemmatization, and natural language processing. The findings revealed that the majority of sentiments expressed were neutral. [Bg19]

In an interesting article, the public reaction to a political incident in India and its neighboring countries was studied using Twitter data. The overall response was found to be predominantly positive, although some individuals expressed concerns and worries. The article outlined a 4-step research procedure, starting with data collection using the Tweepy library and a fixed set of keywords. Preprocessing steps involved lowercasing the data and removing hashtags to enhance accuracy. Sentiment analysis was performed using the Textblob library, which provided polarity values ranging from -1 to 1, representing negative and positive sentiments respectively, with 0 indicating neutrality. Finally, data were categorized and visualized using Matplotlib to gain different perspectives on the subject. [Rn19]

4 Methodology

In summary, the majority of research in this field used Tweepy for data collection, and seemingly its querying procedure was done one time where in our case we have a series of complex memory-optimized procedures to collect the data. In the case of sentiment analysis, mostly the older NLTK and Textblob libraries were used whereas we used the newer faster library of SpaCy for data preprocessing and sentiment analysis. In fact, our system works in this way: the three-part pipeline functions by first reading the scraped data from the database, then transferring the data using a custom Type System to the second part which analyzes the sentiment of words within a Tweet, and ultimately passes the result to the writer for storing the output. In the end, the output result can be used in our analysis that we ourselves undertook one type of analysis and depicted it at the end of this paper. Therefore, in this part, the description of the whole design and development of the Python scripts is provided. Additionally, along with this document, the source code of scraper, pipeline, and analyzer are provided as a repository hosted on GitHub².

As per how the data is processed, we used a custom SQLite database consisting of different aspects of the scraped data with the overall 682,399 recent tweets covering the two focused topics of sustainability and digitization. Moreover, the included information in the database is the tweet topic field, tweet id, hashtags, and tweet texts. As well as other relevant metadata like scraped and tweeted times, authors' id, and such.

5 Analysis

5.1 General Analysis

Before diving deep into the details, a general overview of the scraped tweets indicates in Table 1 that the majority of most tweeted speeches were in English and with relatively moderate negative sentiments, with the exception of the words that are related to the historic events of the earth like prehistoric phenomena. To be precise, “Megafauna” is a hot topic on Twitter relating to the environment and partially regards sustainability to some extent. On the other hand, not only the number of German tweets were fewer, but also those tweets experienced a significantly low amount of keyword reuse. In fact, the most used keyword “fachgerechte Entsorgung” was used ten times less than the most used keywords regarding digitization and sustainability, showing the mere importance of such topics in the said country. Seemingly, out of these keywords, it is apparent that the 2nd and 3rd most frequently used keywords are half of the 1st ones in Germany. While on the other hand, the most used keywords internationally, exceeded our table limits and seem to be a very popular subject outside of the country. Though not all of those international tweets had location metadata and it is not possible to pinpoint which country cares the most

2 <https://github.com/AliVn85/Professions-Sentiment-Analysis-on-Twitter>

regarding digitization and sustainability, pursuing the topic also is outside of our research context.

	Keyword	Avg Sent	Count
0	cloud computing	-0.289921	500
1	atommüll	-0.908605	500
2	hybridbus	-0.303546	500
3	tandem	-0.016890	500
4	big data	-0.642017	500
5	corporate citizenship	-0.372641	500
6	anteil erneuerbaren	-0.041272	500
7	led	-0.330205	500
8	green building	-0.347264	500
9	velo	-0.715579	500
10	e-bikes	-0.502839	500
11	megafauna	0.179740	500
12	iso 14051	-0.905288	500
13	micro grid	-0.410693	500
14	abgasnorm	-0.185469	500
15	biodiesel	-0.250407	500
16	flotation	-0.337020	500
17	artenschutz	-0.198053	500
18	schalldämpfer	-0.380035	500
19	composting	-0.385795	500

Table 1 - Most Related Frequent Keywords

	Keyword	Avg Sent	Count
0	fachgerechte entsorgung	0.155056	50
1	leergut	-0.105909	24
2	leihrad	-0.136431	23
3	fensterbau	0.480846	22
4	straßenreinigung	-0.331140	20
5	prozessmanagement	-0.000299	19
6	straßenbegleitgrün	-0.247830	18
7	gemeindepark	-0.010263	18
8	radverkehrskonzept	-0.189757	18
9	müllumladestation	-0.196089	18
10	liegerad	0.003656	17
11	ressourcenmangel	-0.483404	16
12	müllplatz	-0.848637	15
13	sperrmüll	-0.417641	15
14	grünflächenpflege	-0.147253	14
15	vertragsnaturschutz	-0.050061	14
16	baumschutz	-0.176297	14
17	radwegkonzept	-0.067871	14
18	hobelspäne	-0.196387	14
19	fahrbahnbelag	-0.585849	14

Table 2 - Most Frequent Keywords in Germany

5.2 German-Specific Analysis

Speaking of the context, in the second part. We are focusing in-depth on German-originated tweets with location metadata. The aforementioned data expresses that the strategies of sustaining energy have a priority over any other type of nature preservation activities, and the levels of attention to this subject intertwine with digitization. Though the ranks of reused words indicate the first subject would be a tidbit more considerable to the German Twitter users rather than the latter “Digitization” matter.

Whilst practically it became clear that the south-west areas of Germany seem to be more involved in the two topics as the image shows how dense the sources are closely put together and how sparse northern and eastern Germany are compared to the said sites. Though in metropolises like Berlin and Leipzig, there seem to be more Twitter users with worries regarding digitalization and sustainability issues (figure 1).

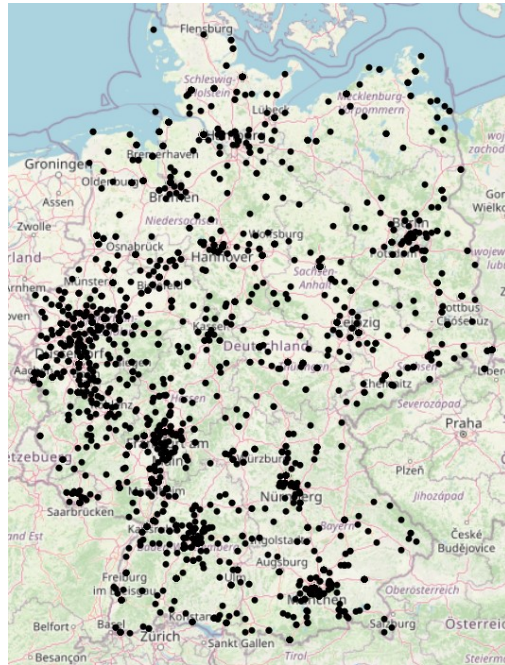


Figure 1 – German Digitization and Sustainability distribution map

But how these worries and thoughts are expressed? To which degree the users are talking positively or negatively regarding the topics? To find out, we simply thoroughly ran a sentiment analysis and therefore the result is written below.

To obtain answers, we sorted the 20 most frequently talked about keywords from German Tweets and that created the table 2.

The most frequently tweeted keyword among Germans is "Leihrad," which also has a positive sentiment bias. The following keyword, "Fensterbau," ranked second but with

Unnamed: 0	Keyword	Avg Sent	Count
11	megafauna	0.179740	500
283	range-extender	0.090998	491
732	sandfilter	0.068844	475
497	radwegekonzept	0.039996	485
1166	dichtstoffe	0.036961	381
682	naturchutzpark	0.028252	477
1142	landschaftsgarten	0.026398	405
494	naturkosmetik	0.024372	485
279	naturpark	0.021165	491
905	naturtourismus	0.018050	465
481	regionalpark	0.014241	485
1182	rohstoffeffizienz	0.011034	368
784	naturschutzbund	0.003298	473
513	energieausbeute	0.002656	484
75	baubiologisch	0.000921	498
208	prozessleittechnik	-0.002365	493
440	waldpädagogik	-0.002504	487
154	fernüberwachung	-0.003348	495
1049	umweltmanagementsystem	-0.003719	446
421	siedlungswasservirtschaft	-0.005203	487

Table 2- German Keywords Average Positive Sent Descending

Unnamed: 0	Keyword	Avg Sent	Count
1	atommüll	-0.908605	500
12	iso 14051	-0.905288	500
523	energieverbrauch-einsparung	-0.889776	484
1079	schrott	-0.871486	436
68	zoologischer garten	-0.820056	498
261	trinkwasserschutz	-0.803569	492
595	schornsteinfeger	-0.775359	481
805	sondermüll	-0.774697	471
1133	strommix	-0.754836	412
1069	restmüll	-0.753084	439
876	müllkippe	-0.749691	467
779	müllsäcke	-0.742953	473
39	greenit	-0.738231	499
142	design to cost	-0.734348	495
417	müll	-0.731030	487
542	thin clients	-0.723182	483
613	müllsack	-0.723018	480
9	velo	-0.715579	500
336	abwasseraufbereitung	-0.709001	490
1001	deinking	-0.707984	455

Table 1- German Keywords Ordered By Average Negative Sentiment

only half the frequency count of the first keyword. Similarly, the keywords "Straßenreinigung," "Prozessmanagement," and "Straßenbegleitgrün" were ranked subsequently, with minor differences in their frequencies. This observation suggests that Digitization and Sustainability were not prioritized over each other. However, it should be noted that the first keyword can also refer to the enjoyment of cycling, making it difficult to draw a definitive conclusion from the obtained tables. Two other tables display keywords sorted by negative and positive sentiments, with the important distinction that location metadata limitations were disregarded to obtain a larger dataset for analysis.

Table 3, shows the majority of people express positive sentiments about nature and sustainability in Germany, as opposed to digitization. The theme of the words also indicates associations with pleasure and vacation. However, it is noteworthy that discussions about nature tourism often focus on recreational activities rather than addressing environmental issues. Another observation is that German audiences tend to have a more critical tone, with a scarcity of significant positive sentiment numbers.

Noteworthy based on Table 4, German Twitter users are inclined to talk with a negative view over topics regarding trash and recycling, and leftovers. This however could very well be inaccurate as such words have negative sentiments in them, though it is worth reminding you that the sentiment analysis is done not only per keywords of a tweet, but the whole tweet text itself. Nearly, the 20 most negative keywords have the meaning Trash in them.

Additionally, one more exercise was done with the achieved tweets data. That is, we found even more accurate results by merging the popularity and the average sentiment values of keywords. This enabled us to take all entries into account and was possible by multiplying the two factors together without losing any information. Ordering it descending, showed the most positively popular keywords at the top. And on the other hand, the most negative ones are at the bottom.

	Unnamed: 0	Keyword	Avg Sent	Count	value
11	11	megafauna	0.179740	500	0.051918
283	283	range-extender	0.090998	491	0.025812
732	732	sandfilter	0.068844	475	0.018891
497	497	radwegekonzept	0.039996	485	0.011206
1166	1166	dichtstoffe	0.036961	381	0.008135
...
261	261	trinkwasserschutz	-0.803569	492	-0.228397
68	68	zoologischer garten	-0.820056	498	-0.235926
523	523	energieverbrauch-einsparung	-0.889776	484	-0.248788
12	12	iso 14051	-0.905288	500	-0.261493
1	1	atommüll	-0.908605	500	-0.262451

[1731 rows x 5 columns]

Table 3 - Alternative way of sentimental analysis

In this case, Table 5 showed a result with a high correlation to the last two ones, providing proof for the level of accuracy is satisfying. As per our previous experiment, sometimes a simple ordering of top frequent keywords per sentiment level would not show meaningful knowledge, however, in this case, the experiment expressed otherwise.

6 Conclusion

To conclude, it was that almost every one of the tweets originating in Germany possessed good emotion toward initiatives to promote sustainability – which covers mainly the engineering part of the labor market (“Sandfilter”, “Dichtstoffe”) – but communicated more negative opinions, particularly about residual disposal techniques. Besides, we noticed that even a proficient analyzer could return unrelated results if the terms being searched already had negative connotations. To be precise, the origins of the published texts are mainly metropolitan locations which also leads to further questions on the difference in the labor market between rural and metropolitan areas. Moreover, Germans tend to talk with negative sentiments concerning the said themes, especially in the case of strategies to sustain energy or digitize. But when they tweet about a related matter with positive sentiments, they tend to talk about pleasing recreational activities like cycling and we receive little to no information on the labor market. It is worth mentioning that including AI technologies in the pipeline to predict the origin of the tweets based on the German samples is a future work suggestion relating to this matter. The other suggestion for future work is to study the tweet numbers above the limitations that the Twitter API offers.

Our suggestion regarding future work is to add an NLP feature in the scraper script so that according to the language of the import, it can tokenize and clean up the characters. This leads to avoiding manually micromanaging the input like removing the special or illegal characters.

On the other hand, the other suggestion is to add the already scraped Twitter user data in the Analysis part in order to further get the user profile details and then extract the knowledge related to the Twitter account that sends a specific Twitter post, so that more data and analytics can be inferred.

One of the notable limitations of our research was that for scraping data per each keyword, a maximum of 500 tweets could have been retrieved at most. That is because of how the Tweeter API at the time of research was working. To get over the limit one has to split scraping about a specific keyword in different time ranges. For example, a specific keyword must be scraped from 2006 to 2007, then such an endeavor would be repeated for the time ranging from 2007 to 2008 and so on. Although this would eliminate the maximum 500 tweets cap and thus would yield a more accurate result for calculating the tables regarding frequencies of terms and keywords in this research, due to the short timeframe it was not possible for us to do so, and thus it is a suggestion for future work.

References

- [Bg19] Baweja, A; Garg, P.: Sentimental Analysis of Twitter Data for Job Opportunities. *International Research Journal of Engineering and Technology (IRJET)*, 6/11, 2019.
- [Bf11] Bundesagentur für Arbeit. *Klassifikation der Berufe 2010–Band 1: systematischer und alphabetischer Teil mit Erläuterungen*, 2011.
- [BS22] Binnewitt, J.; Schnepf, T.: Join us to turn the wor(l)d greener! – Investigating online apprenticeship advertisements’ reference to environmental sustainability. In: *Zum Konzept der Nachhaltigkeit in Arbeit, Beruf und Bildung - Stand in Forschung und Praxis*, 2022.
- [Cp20] Kariya, C.; Khodke, P. Twitter Sentiment Analysis, 2020 *International Conference for Emerging Technology (INCET)*, pp. 1-3, 2020.
- [Ms17] M. Trupthi et al.: Sentiment Analysis on Twitter Using Streaming API, 2017 *IEEE 7th International Advance Computing Conference (IACC)*, Hyderabad, India, 2017, pp. 915-919, doi: 10.1109/IACC.2017.0186.
- [MK09] Makice, K.: *Twitter API - Up and Running: Learn How to Build Twitter Applications*, 2009.
- [Mn13] Marres, N.; Weltevrede, E.: SCRAPING THE SOCIAL? *Journal of Cultural Economy* 6, 313-335, 2013.
- [Rn19] R. Patil et. al: Twitter Data Visualization and Sentiment Analysis of Article 370, 2019 *International Conference on Advances in Computing, Communication and Control (ICAC3)*, pp. 1-4, 2019.

What social media can tell us about essential occupations

Contextualising twitter data to understand shifts in occupational valuations


Dr. Michael Tiemann ¹, Stefan Udelhofen ² and Lisa Fournier ³


Abstract: Societal debates about essential occupations in the context of the pandemic have raised questions about the valuation of occupational tasks. In a first step we compare two lists of essential occupations, one from the start, the other from the end of the pandemic, to describe differences in their valuation based on characteristics such as wages, prestige and workload. Between these lists it becomes apparent that there has been a broadening shift, with essential occupations at the end of the pandemic being different to what they were at the beginning. This is based on data from the BIBB/BAuA Employment Survey 2018. We then investigate the use of twitter data for generating insights on how the valuation of occupations were discussed and changed during the pandemic and thus helped leverage said shift in essential occupations.


Keywords: essential occupations, COVID-19, Twitter

1 Essential occupations under discussion⁴

Occupations relating to the provision of key supplies and services – otherwise known as “essential jobs” or “essential work” – are still new categories in the field of labour market and occupational research [HSZ23]. Different occupational groups are subsumed under this heading depending on the context of a specific crisis situation, although no clear definition for essential occupations exists. They gained some attention in the light of the first coronavirus lockdown in March 2020. In Germany, the Federal Office for Civil Protection and Disaster Assistance categorised them as occupations which “help ensure provision to the population of important and in some cases vital goods and services” (own translation, cited from [SSS21]). Workers in these occupations enjoyed special rights (e.g. emergency childcare), but were also subject to restrictions (such as bans on taking a leave

¹ Federal Institute for Vocational Education and Training, AB 12, Friedrich-Ebert-Allee 114-116, 53113 Bonn, tiemann@bibb.de,  <https://orcid.org/0000-0001-7136-2744>

² Federal Institute for Vocational Education and Training, AB 12, Friedrich-Ebert-Allee 114-116, 53113 Bonn, stefan.udelhofen@bibb.de,  <https://orcid.org/0000-0002-1749-7231>

³ Federal Institute for Vocational Education and Training, AB 12, Friedrich-Ebert-Allee 114-116, 53113 Bonn, lisa.fournier@bibb.de,  <https://orcid.org/0009-0008-9544-5577>

⁴ The parts of this paper dealing with the analysis of BIBB/BAuA-Employment Survey data (section 1) are in large parts based on our article [TUF23].

and the lack of opportunities for working from home). These “initial” [Ko20]⁵ lists of essential occupations were adjusted and extended as the pandemic continued. Teaching staff and financial services workers, for example, were added. Regional differences between individual federal states also emerged, such as in how agricultural occupations were considered.

Our article seeks to take this as a starting point in order to link a description of the valuation of essential occupations with the question of the distribution of requirements and qualification levels. A consideration of the categorical dynamics produced via the adjustment of lists of essential occupations appears to be of particular interest, also extending the focus of previous studies. For this reason, we use data from the 2018 BIBB/BAuA Employment Survey⁶. The Classification of Occupations (KldB 2010) is used to operationalise and differentiate between two lists: a narrower list of occupations (referred to below as the “Berlin list”), which were defined on an ad hoc basis as being essential in the wake of the coronavirus pandemic (20 occupations), and an extended list (62 occupations), which takes account of subsequent additions made by policy makers and of the findings from academic research studies.

In our research, we looked at characteristics associated with valuation. For this, we have looked at pay, job prestige, workload and demands, and skill levels of essential jobs. We also investigated whether there is a mismatch between the requirement levels of the tasks and the qualification levels of the employees.

As indicators of occupational esteem, we examined pay and prestige. Initially, our comparative investigation of remuneration and occupational prestige as indicators of the valuation and esteem of individual occupations is essentially in line with findings from previous studies. In the essential occupations on the “Berlin list”, the average hourly wage is significantly lower (€16.32) than that paid in non-essential occupations (€19.22). This discrepancy is not reproduced in the “extended list”. This is similar with regard to occupational prestige. The average ISEI value⁷ for occupations on the “Berlin list” (40.15) is significantly below the value of the other occupations (47.88). Although the average value for occupations on the “extended list” is slightly above that for all other occupations (47.05 as compared to 44.96), this difference is not significant.

⁵ Actually, there have been several lists from the beginning due to most federal states issuing their own list.

⁶ The Employment Survey 2018 is a representative survey of 20,012 employees in Germany on changes in work and occupations and on the acquisition and utilisation of professional and vocational qualifications. The data [HHR18a; HHR18b] were collected by Kantar Public of Munich via computer-assisted telephone interviews (CATI) carried out between October 2017 and April 2018. Selection was based on a random procedure (Gabler-Häder sampling process) in order to ensure representativity. The statistical population comprises members of the labour force aged 15 and above (not including trainees). Employment activity is considered to be regular work activity of at least ten hours per week for which payment is received (“core workers”). Data was adapted to the structures of the population via weighting in accordance with central characteristics on the basis of the 2017 Microcensus.

⁷ The “International Socio-Economic Index of Occupational Status” (ISEI) maps the social status of an occupation by providing a value between 16 and 90 and thus ranks occupations according to status. Further information is available at: <https://metadaten.bibb.de/de/classification/detail/11>.

There are other measures to take into account when looking at valuations. In the context of essential occupations the subjective valuation of ones own work should be taken into account as well. Here it is worthwhile to note that employees in occupations on both lists are similarly likely to rate their work as important. Differences between occupations on the two lists are shown with regard to consideration of the employees' situation. Workers in occupations on the "Berlin list" tend to evaluate their overall work conditions as being worse and also state that workloads are higher than average. They are also more likely to be less well informed about tasks they have to fulfil and report that they enjoy a lesser degree of occupational autonomy (for more details cf. electronic supplement of [TUF23], Table 3).

Linear models further demonstrate that, although they assess their work as subjectively relevant, workers in essential jobs also perceive the value of their tasks to be less positive and show themselves to be less satisfied the longer they work in an occupation. In the case of employees in occupations on the "extended list", workloads are only perceived as being significantly worse in relation to physically demanding tasks and with regard to noise in the workplace (cf. [TUF23] electronic supplement, Tables 6 and 7). In contrast, work-life balance is perceived to be worse.

We also took a closer look at qualifications in essential jobs. Deviations in respect to the shares of requirement level of "no vocational qualification" and "academic qualification" exhibit differences to the totality of occupations. Employees in occupations on the "Berlin list" show a comparatively large share (21.4%) of workers stating that no vocational qualification is required to work in their jobs (all occupations: 17.5%). On the other hand, the proportion of employees with an academic requirement level (17%) lies below that for all occupations (24.4%). These ratios are reversed if we consider employees in occupations on the "extended list": The requirement level of "no vocational qualification" is significantly lower there (14.5%), whereas the proportion of "academic qualifications" is higher (27.7%).

A consideration of the qualification level also produces a similar structure. Initially, a significant preponderance of 65.23 percent of employees with vocational qualifications can be observed in occupations on the "Berlin list". In the "extended list", this share stands at 55.35 percent (all occupations: 55.67%). In occupations on the "Berlin list", 10.8% of employees have no vocational qualification. The corresponding share for the "extended list" is 7.3%. These figures are respectively above and below the value for all occupations (8.9%).

A comparison of the matching of requirements and qualification level (cf. Tables 4 and 5 in electronic supplement to [TUF23]) also shows differences between the two lists. In objective⁸ terms, employees on the "Berlin list" do not tend to be under-qualified for the

⁸ "Objective" relates to the type of measurement. A comparison is drawn between highest vocational qualification and the requirement levels according to the Classification of Occupations (KldB 2010).

work they do. Subjectively⁹, they believe they may not be deployed appropriately. By way of contrast, employees on the “extended list” are objectively more likely to be appropriately deployed rather than being under-qualified. Subjectively, however, they are more likely to work over-qualified and their deployment does not match their qualification. We were, however, unable to ascertain any statistically significant mismatch between requirements and qualification level.

Overall, our evaluations indicate a more differentiated picture of essential occupations than the one hitherto addressed in societal discourses or in academic research. Over the course of time, and in a way which diverges from the public perception of essential occupations at the beginning of the pandemic, we can assume an increase in occupations with an academic requirement level and in employees with academic training. Essential occupations, thus, have become more “average”. How this happened cannot be inferred from the Employment Survey data, which is why we turned to a different data source.

2 Reviews of essential work on twitter

Twitter is a short message service that has become established primarily in the media industry, but is also used by institutions, companies, organizations, political parties and educational institutions, as well as by private individuals. Compared to other social media there is the particularity of a limitation of text messages to a maximum of 280 characters (until November 2017 it was 140 characters). Twitter is a medium for information, communication and opinion in one place, and more or less in real time. In 2017, around 320 million people actively used the service, i.e., logged into the network, and a further 500 million followed tweets without logging in [Sa17]. Current statistics published by Twitter owner Elon Musk show 253.8 million users in 2020 [Tw20]. When the company celebrated its tenth anniversary in 2016, user figures for Germany were announced for the first time. According to these, a total of 12 million active and passive German users were recorded at the time. Although registration and use of the service is free of charge, the company behind it, Twitter Inc., is financed by advertising, evaluating user data and selling the flow of information in its network to third parties [Sa17]. Twitter uses an algorithm that is difficult for users to understand. Since the takeover of the company in October 2022 by Elon Musk, the new owner announced in March 2023 a "new era of transparency for Twitter" [Tw23], in which this should be made more comprehensible.

There is only limited information on the exact user figures of Twitter in Germany and the algorithm to which Twitter is subject. This must be considered when interpreting the data presented below and it is especially noteworthy since data in social sciences are usually either statistically representative (like the Employment Surveys analysed here) or reveal types of people, behaviours or structures (with more ethnographic, anthropological or so

⁹ “Subjective” also refers to the type of measurement. In this case, a comparison is made between highest vocational qualification and the response to the question as to which qualification is normally necessary for the exercising of one’s own occupation.

called “qualitative” data like observations or in-depth interviews). Twitter data falls somewhere between these kinds of data. On the one hand there are masses of individual data points, on the other hand they are in no way representative. Their meaning is also only revealed when looking into detailed individual, or typical, contents [Pf16].

But social media data is also a source depicting trends in opinions and discussions. We would expect topics like the pandemic to be discussed widely enough to be visible in twitter tweets of 2020 and 2021. In a sense, it is a real-time seismograph of relevant topics, where “relevant” simply stands for “important enough to many persons to write/tweet about them”. Twitter has been described very enthusiastically by young scholars as a medium which is simultaneously a mass medium as well as constructing our reality while covering all (social and individual) topics [We21]. It has as well been assessed as polarising opinions and even influencing elections [GKD19] and examined as a political sphere [BR20]. Overall, Twitter data are seen as a valuable source [Pf16].

Following our initial findings on the changes in essential jobs we want to answer one question with analysing twitter data: Is there a shift in discussions of essential jobs? The idea behind this is as follows: A change in the composition of essential jobs must have been justified. This should have happened in political discussions, as the lists are results of governmental processes. As inclusion of a job on the list of essential jobs meant rights and restrictions for their incumbents, these decisions will have been at least commented on publicly. Given the possible explanation of occupational interest groups shifting discussions and opinions towards the inclusion of specific occupations, these comments will not just have occurred after decisions had been made. A mass medium which is found to be able to foster polarisation and form opinions, even influence elections, will at least show traces of a discussion affecting people in a crisis situation like a pandemic. Whether this was the case and if we can trace such a shift in debates of essential jobs will be discussed in the following.

3 Preparation and analysis of twitter data

Since discussions as we have depicted them will include communication of sentiments, we first conducted a sentiment analysis. We focussed on tweets¹⁰ mentioning essential jobs and therefore only included tweets that had some mentioning of jobs or occupations. When such mentions were being made, they were coded with a systematic position of the German Classification of Occupations (KldB2010).¹¹ Out of all the tweets being connected to and coded with an occupation in this way we filtered out those that were on the initial “Berlin” list of essential occupations and those on the “expanded” list. It is here

¹⁰ They were scraped by students of Koblenz University, utilising BIBB’s TM4VETR (<https://github.com/TM4VETR>) and twitter’s API, checking for tweets containing mentions of official occupation names or synonyms.

¹¹ This was done automatically utilising lists of occupational names and their aliases.

where we would expect statements and comments fitting the shift in debates explained above.

After removing doubles, 1.2 million tweets remained that related to occupations on at least the expanded list from 2007 until 2023. Besides information on which list of essential jobs tweets were related to we also have information about tweet sentiment and the full text of tweets. Tweet sentiment is interesting since a change in sentiment could occur in correspondence to a shift in opinions and the debate. Tweet contents would also possibly reveal topical shifts. Due to the nature of data processing though we have tweets that contain a lot of very similar features (words, numbers, emojis and hashtagged topics). This makes topic modelling challenging, which is why we will look at predefined topics for the time being.¹²

Sentiments were assessed within the data preparation process with adapted standard procedures.¹³ Negative or positive sentiments were scarce, compared to neutral ones (four fifths neutral, around 16 percent negative, four percent positive). For our analyses we work with mean sentiments over groups of tweets (e.g. occupations or the group of essential occupations) instead of classified dichotomous measures.

While data scraping and building the dataset of tweets with sentiments, occupations and tweet texts were done with a pipeline devised by researchers at BIBB, further analysis was conducted in R (R Core Team 2023). Topics were searched for within tweets using regular expressions.

4 Assessment of essential occupations on twitter

Our data shows an increase in tweets related to occupations over time, from only 17 in 2007 up to more than 240.000 in 2023 (where scraping stopped in March) with the most tweets scraped from 2022. The share of posts relating to essential occupations from the shorter and more restrictive Berlin list show some variation, but interestingly are stable (at around 35 percent) during the pandemic, indicating that they did not raise more attention or possibly even that not more persons in these occupations tweeted about them.

Figure 1 shows the increase in absolute numbers, but only for occupations on the Berlin list. We do not see movements that could be traced back to events occurring during the pandemic, e.g. there is no peak in tweets around the start of the pandemic, only at the beginning of the first lockdown, but not at the start of the second. There does seem to be an increase around winter months, though, but we could not find seasonal shifts over time.

¹² To give an example: In topic modelling one can trim data and leave out words (features) occurring only in very few or in almost all tweets. Trimming in our data would regularly cut of more than 95 percent of features.

¹³ The routines used can be found here: <https://github.com/AliVn85/Professions-Sentiment-Analysis-on-Twitter>

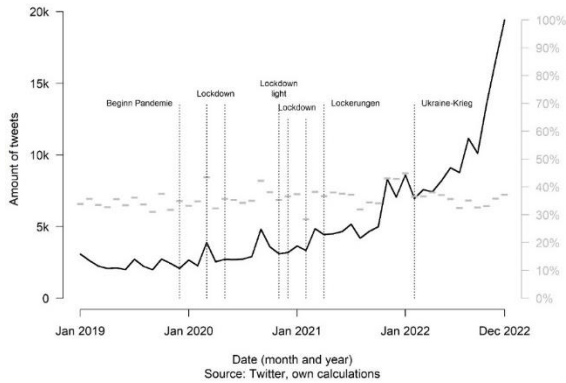


Figure 1: Tweets in essential jobs of the Berlin list

We observe negative sentiments in essential jobs to increase over time: In March 2023 we have a mean negative sentiment of .3862 in occupations on the Berlin list, in January 2019 it was .1295.

There is an interesting finding in mean sentiments regarding the differences of essential jobs of the Berlin list and the expanded list. Figure 2 shows that generally occupations of the Berlin list are connected to higher negative sentiments than those from the extended list. Taking into account that the extended list is indicative of a mean over all occupations we see an interesting pattern: Before and after the pandemic, and even within it, when there was a time of relative easement, we see that essential jobs from the Berlin list are connected to more negative sentiments than other jobs – while during the pandemic this es levelled out and at times even turned (in January, April and August 2020). This could be due to essential jobs of the Berlin list generally being discussed based on their discriminating features like long working hours, relatively low pay, relatively bad work situations. But during the pandemic there was a wave of support and appraisal towards these occupations (especially care and health jobs), which could have lead to the rapprochement of both curves. This also corresponds to workers’ self assessment (see above) of the value of their tasks as less positive in essential occupations of the Berlin list than in other occupations, as well as their work loads being higher and their work conditions being worse.

Looking at which topics were discussed figure 3 shows some interesting differences between an occupation¹⁴ from the Berlin list (sales occupations selling foodstuffs and one of the extended list (insurance and financial services). In sales occupations the negative sentiments pronounce the differences of pandemic and non-pandemic times. In the first

¹⁴ In this case we look at three-digit occupations from the German classification. These are relatively detailed regarding occupational content, as there is only one more level with the fourth digit, but they include every level of skill requirement (from no qualification needed up to academic qualifications). Even though there are many data points in our data set and we could in theory look at occupations on the five-digit level, we would run into problems of low numbers in some of these cases.

year of the pandemic, valuation was a topic with this occupation, in the second lockdown we see a peak in topics relating to acknowledging what incumbents do. In insurance and finance services we do not observe higher than average negative sentiments. Also, when they peak, they do so later than occupations from the Berlin list. Topics discussed are generally less prominent, except for a discussion of pandemic measures during the second lockdown (January 2021) and of acknowledgement right in the following month (March 2021). Valuation topics tend to be relevant during phases of easement, but never as much as with sales occupations.

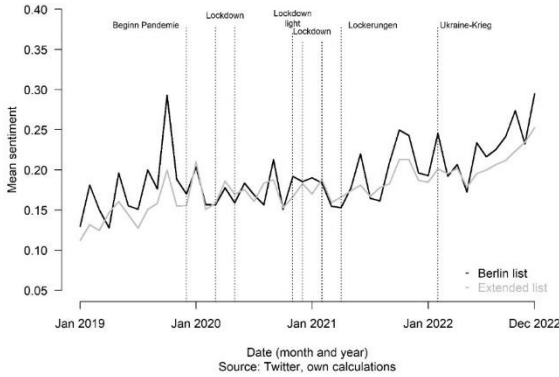


Figure 2: Negative sentiments in essential jobs

Overall the analysis of topics shows the data to reflect discussions and debates at the times they were also discussed in the media. This might be seen as a kind of verification. Combined with distinctive occupations they shed some light on how discussions on twitter were structured thematically.

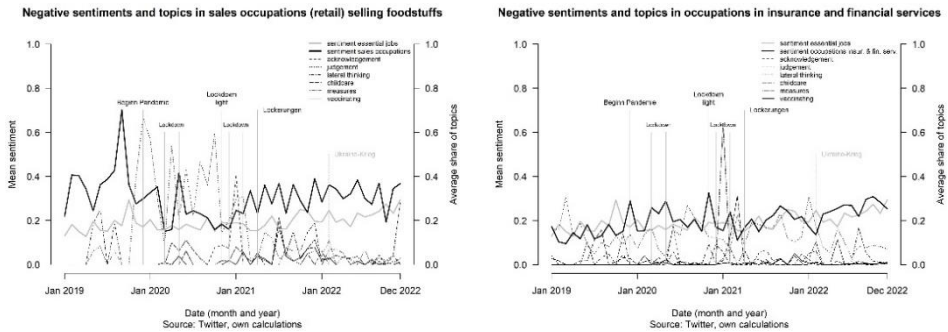


Figure 3: Topics in two essential occupations

5 Conclusion

This paper shows that valuation of essential jobs can be traced in very different data types. We see, among others, lower wages, higher mismatch with respect to their qualifications but also higher negative sentiments in occupations on the Berlin list. These findings might be seen as hints for discussions in which essential occupations were restructured in a sociological meaning. Considering developments in sentiments and topics discussed, Twitter seems to be a valuable data source. However, not least due to the ambiguous user numbers and the algorithm Twitter is subject to, the data is not easy to interpret. To further examine these initial indications, on the Twitter data in more detail and also how it can supplement other data sources, future research should establish a pipeline and examine and compare other cases, as well as other research questions in more detail.

References

- [BR20] Bouvier, G., Rosenbaum, J.E. (eds.): *Twitter, the Public Sphere, and the Chaos of Online Deliberation*. Cham, 2020.
- [GKD19] Grovera P., Kumar Kara A., Dwivedib, Y. K., Janssen, M.: Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences? *Technological Forecasting and Social Change* 145, pp. 438-460, 2019.
- [HHR18a] Hall, A.; Hünefeld, L.; Rohrbach-Schmidt, D.: BIBB/BAuA-Erwerbstätigenbefragung 2018 – Arbeit und Beruf im Wandel. Erwerb und Verwertung beruflicher Qualifikationen. SUF_1.0. Bonn 2020a. doi.org/10.7803/501.18.1.1.10
- [HHR18b] Hall, A.; Hünefeld, L.; Rohrbach-Schmidt, D.: BIBB/BAuA-Erwerbstätigenbefragung 2018 – Arbeit und Beruf im Wandel. Erwerb und Verwertung beruflicher Qualifikationen. SV_1.0. Bonn 2020b. doi.org/10.7803/501.18.1.5.10
- [HSZ23] Herzog, L.; Sold, K.; Zimmermann, B.: Essential Work: A Category in the Making? Unsettling the Ways We Think about Jobs, Labor, and Activities. In: (Herzog, L.; Zimmermann, B., eds.): *Shifting Categories of Work. Unsettling the Ways We Think about Jobs, Labor, and Activities*. New York und Abingdon, pp. 252–264, 2023.
- [Ko20] Koebe, J. et.al.: *Systemrelevant, aber dennoch kaum anerkannt: Entlohnung unverzichtbarer Berufe in der Corona-Krise unterdurchschnittlich*. Berlin, 2020.
- [Pf16] Pfaffenberger, F.: *Twitter als Basis wissenschaftlicher Studien. Eine Bewertung gängiger Erhebungs- und Analysemethoden der Twitter-Forschung*. Wiesbaden, 2016.
- [R23] R Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, 2023, <https://www.R-project.org>, retrieved: 13.07.2023.
- [Sa17] Samuelis, T.: *Twitter in der Bildung*. Bundeszentrale für politische Bildung. Bonn, 2017.

- [SSS21] Schrenker, A.; Samtleben, C.; Schrenker, M.: Applaus ist nicht genug. Gesellschaftliche Anerkennung systemrelevanter Berufe. Aus Politik und Zeitgeschichte 71, pp. 13–15, 2021.
- [TUF23] Tiemann, M.; Udelhofen, S.; Fournier, L.: Esteem, equivalence and valuation of essential occupations. *Berufsbildung in Wissenschaft und Praxis (bwp)* 52, pp. 22–26, 2023.
- [Tw20] Twitter. Elon Musk. Slides from my Twitter company talk. 2020, <https://twitter.com/elonmusk/status/1596718851097755648?s=20&t=kKOJe8nXDwbmD3i2LgXj-Q>, retrieved: 13.07.2023.
- [Tw23] Twitter. Twitter's Recommendation Algorithm. 2023, ULR: https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm, retrieved: 13.07.2023.
- [We21] Wende, F.: *Twitter: Das Verbreitungsmedium für alle. 140 Zeichen zur Selbstbeschreibung der Gesellschaft.* Grin Verlag, 2021.

Wirtschaft, Management Industrie -
8th Industrial Automation and
Control Systems Standardization
Workshop (IACS 2023)

Optimizing OPC UA Deployments on Node.js through Advanced Logging Techniques

Erkin Kirdan ¹, Josef Schindler¹ and Karl Waedt¹

Abstract: This paper explores the value and benefits of implementing advanced logging techniques within OPC UA deployments in Node.js. OPC UA is a leading protocol for interoperable and secure data exchange in industrial automation and IoT, among other complex data communication systems. Adopting sophisticated logging strategies can optimize its deployments on Node.js. The paper uses a case study to demonstrate the real-world impact of integrating robust logging solutions into OPC UA deployments. It underscores how such practices can improve system reliability, increase debugging efficiency, enhance security, and understand system performance. This valuable insight aids developers and system administrators in managing and maintaining complex OPC UA deployments, reinforcing the critical role of a well-implemented logging strategy. By analyzing a specific instance of an OPC UA server-client pair implemented in Node.js, the paper invites a broader discussion around the optimization strategies that could further strengthen the robustness and security of OPC UA systems. It aims to open avenues for more research, encouraging a continuous drive towards more efficient and secure industrial automation and data communication systems.

Keywords: opc ua; logging; debugging

1 Introduction

Optimized data communication and industrial automation are at the heart of contemporary technology operations. OPC UA is an industry-leading protocol to facilitate secure and interoperable data exchange. Particularly on Node.js, JavaScript-based OPC UA deployments are prevalent, forming the backbone of diverse data communication systems. However, these systems often exhibit a need for enhancement in efficiency and reliability, an issue that this paper aims to address through the implementation of advanced logging techniques.

In the broader context of technological advancements, [Pa21] illustrates the shift towards decentralized and cross-sector energy systems, which involve complex data exchange and management. While the authors' primary focus is on energy management, their discussion about critical components of IoT, including wireless communication and interoperability middleware, resonates with our paper's core theme of OPC UA deployments in Node.js. As our research explores how advanced logging techniques can enhance system reliability and security, the study above's analysis of emerging technology trends underlines the potential for further optimization in complex data

¹Framatome GmbH, ICSP, Paul-Gossen Straße, 91052 Erlangen, Germany {name.surname}@framatome.com

communication systems, which includes industrial automation and IoT applications.

Logging is a means to record application behaviors and subsequently store these details in a lasting medium. This indispensable process is integral to system monitoring and troubleshooting and even serves as a resource for non-technical teams tracking user behavior or measuring marketing campaign impacts. Logging's potential to provide multifaceted value within an organization largely depends on factors, including the specificity of logged details, the format, and how these logs are analyzed and utilized.

In a Node.js environment, the logging process begins at a rudimentary level with the built-in console module. However, when addressing the needs of a complex system like OPC UA, the basic console module approach needs to be revised. Effective logging in production applications demands more advanced features, including log levels, structured JSON logging, timestamps, and the ability to log to multiple destinations simultaneously. This is where logging frameworks come into play, bringing a sense of structure and detail to the logging process that the console module cannot deliver.

Frameworks like Winston, Pino, Bunyan, and Roarr, to name a few, have shown themselves to be invaluable tools in creating robust logging strategies. Winston, as the most popular and comprehensive logging framework for Node.js, will serve as our primary reference for demonstrating the implementation of an advanced logging strategy.

Our primary focus is enhancing OPC UA deployments by integrating advanced logging techniques. By considering OPC UA as a use case, we aim to increase system reliability, debugging efficiency, and security and foster a more profound comprehension of system performance. This paper ventures to illustrate these enhancements through a case study, sparking a broader conversation around optimization strategies that could solidify the robustness and security of OPC UA systems. Ultimately, this dialogue aspires to propel further research and motivate a continuous endeavor toward more efficient and secure industrial automation and data communication systems.

As our study delves deeper into the implications of advanced logging techniques for OPC UA deployments, addressing the broader context of cybersecurity is vital, particularly given the significant consequences of security breaches in industrial and data communication systems. Here, we draw upon the insights presented in [Wa22] by Framatome GmbH. Since 2015, this organization has engaged numerous Cybersecurity Ph.D. candidates, Bachelor's and Master's thesis students in solving complex technical challenges from the nuclear cybersecurity domain. This extensive engagement in cybersecurity R&D, nationally and internationally, offers a rich backdrop against which we can better understand the complexities of maintaining robust and secure OPC UA deployments. Furthermore, their future outlook on challenges provides context for our ongoing exploration of how advanced logging techniques can contribute to mitigating such future risks in OPC UA systems.

2 Identifying Log-worthy Events in OPC UA

A key aspect of effective logging is determining what to log. While troubleshooting and debugging are conventional uses of logs, their functionality extends to studying application systems, improving business logic, and making informed decisions. They are instrumental in understanding customer behavior, facilitating data mining, and offering insights into application workings.

Several events should be logged in the context of OPC UA deployments in Node.js. Firstly, requests that record the execution of services, such as authentication, authorizations, system, and data access, are critical to understanding how the application is being interacted with. This becomes even more crucial in a scenario where the application is involved in complex data communication systems, as is the case with OPC UA.

Logs should include resource-related issues such as exhausted resources, exceeded capacities, and connectivity problems. These logs can provide invaluable insights into the areas of the system that are under strain or improperly configured.

Thirdly, the availability of the application should be tracked by logging the runtime at the start and end of an application session. Availability logs capture faults and exceptions, providing a clear view of the system's stability and availability over time. This is particularly beneficial in maintaining a high level of service in industrial automation and IoT applications, where uptime is often critical.

Fourthly, potential threats to the system should be logged. These could include invalid inputs, security issues such as invalid API keys, failed security verifications, failed authentications, and other warnings triggered by the application's security features. Logging such information can significantly enhance the security of OPC UA deployments, alerting administrators to potential breaches and vulnerabilities.

Lastly, significant system or application changes and data changes, such as creating or deleting data, should be logged. This also extends to user interactions like button clicks and context changes. Logging these events can lead to a better understanding of how users interact with the system and how changes might impact system performance.

In essence, capturing a wide range of events and changes within logs can lead to a deeper understanding of OPC UA deployments. This can ultimately contribute to developing more robust, efficient, and secure systems, reaffirming the critical role of advanced logging strategies.

3 Related Work

The study [Mü20] provides a pertinent point of reference in our research. This paper

dives into the security models of four widely-used open-source OPC UA implementations - open62541, node-opcua, UA-.NETStandard, and python-opcua. Given the centrality of OPC UA in remote industrial control systems, security is a critical consideration, and this work provides a valuable assessment of how these popular open-source implementations manage security concerns. Furthermore, it explores the scalability of these implementations in terms of the number of clients and OPC UA nodes. These aspects align closely with our research focus, as our advanced logging techniques aim to enhance security and the overall scalability of OPC UA deployments on Node.js, particularly by improving system reliability and debugging efficiency.

The paper [Mü21] contributes significantly to our understanding of the varying functionalities across different OPC UA implementations. The authors carefully compare the most common open-source OPC UA implementations, focusing on their support for essential features and overall interoperability. Their findings indicate that most of these implementations have good coverage of features, particularly open62541 and UA-.NETStandard, and they found no significant interoperability issues in their tests. This research provides a critical background for our work, as it helps establish the diverse range of features and functionalities offered by various OPC UA implementations, providing a broader context for applying our advanced logging strategies.

The authors in [Sc22] expand upon a crucial facet of our work, particularly ensuring the secure execution of OPC UA deployments in JavaScript environments. In this paper, the authors analyze the potential of Deno as a runtime system for the widely used JavaScript implementation of OPC UA, known as node-opcua. Deno, developed by the original author of Node.js, promises enhanced security features like a restricted file system and network access by default. The paper's authors present and evaluate measures to overcome errors when running node-opcua on Deno, considering the current lack of a dedicated OPC UA implementation and the limited research literature on this topic. This research aligns with our exploration of advanced logging strategies in the context of Node.js, offering insights into the evolving landscape of secure runtime environments for OPC UA deployments.

4 Implementation

The advanced logging techniques for optimizing OPC UA deployments can be implemented by leveraging Winston, a versatile logging library for Node.js, which provides several advanced features.

The heart of this approach is about replacing the regular console logs with more detailed and structured log messages. This enhancement can lead to improved debugging, a better understanding of system behavior, and as a result, more efficient OPC UA deployments.

In your server code, initialize Winston and replace `console.log` with `logger.info`:

```
const winston = require('winston');

const logger = winston.createLogger({
  // configuration
});

// Replace console.log with logger.info
logger.info("initialized");
```

You can adjust the configuration to change the log level, format, and destination. Winston also allows you to log not just info messages, but also warnings and errors.

In case of exceptions, it's crucial to log them for investigation:

```
try {
  await server.initialize();
  logger.info("Server initialized successfully");
} catch (error) {
  logger.error("Error initializing server", error);
}
```

The client-side implementation is similar. Initialize Winston and replace `console.log` with `logger.info`. In cases of exceptions, you should use the `logger.error` function to record the error details:

```
try {
  await client.connect(endpointUrl);
  logger.info("Successfully connected to the server");
} catch (error) {
  logger.error("Error connecting to the server",
    error);
}
```

When the client reads data from the server, logging the response can help with debugging:

```
const dataValue2 = await session.read({
  nodeId: "ns=1;s=free_memory",
  attributeId: AttributeIds.Value
});
logger.info(`Read value: ${dataValue2.toString()}`);
```

This setup, both on the server and client-side, provides an advanced logging mechanism using Winston, which is a powerful way to optimize the OPC UA deployments by

making the process of troubleshooting, monitoring, and analyzing the server-client interactions significantly more effective.

After applying the advanced logging technique using Winston, the structured logging output in the console appears as shown in Figure 1. Each log entry contains a level indicating the severity of the message, the actual log message, and a service identifier. This standard format provides easy traceability and efficient debugging for OPC UA deployments.

```
{ "level": "info", "message": "Server initialized", "service": "user-service" }
{ "level": "info", "message": "Variable1 read with value 1", "service": "user-service" }
{ "level": "info", "message": "Variable2 read with value 10", "service": "user-service" }
{ "level": "info", "message": "Server is now listening", "service": "user-service" }
{ "level": "info", "message": "The primary server endpoint url is opc.tcp://erkin-VirtualBox:4334/UA/MyLittleServer", "service": "user-service" }
{ "level": "info", "message": "Variable1 updated to 2", "service": "user-service" }
```

Fig. 1: Screenshot of the console output of Winston

These logs show how the system state changes over time, from the server initialization to the reading and updating of variables. They also demonstrate how a structured and clear logging approach improves system understanding and management. The power of this logging approach is that it provides valuable real-time insights and forms a precise system activity history, which is crucial for diagnosing past incidents and predicting future system behavior.

5 Conclusion

The implementation and application of advanced logging techniques in OPC UA deployments on Node.js, as discussed in this workshop paper, have shown significant potential for optimizing such systems' overall performance and security. With the adoption of structured logging using the Winston library, system administrators and developers gain a more detailed, organized, and traceable record of events. This enables them to better understand the system's behavior, quickly identify and resolve issues, and effectively prevent potential vulnerabilities.

Moreover, using a case study demonstrated the tangible benefits of these advanced logging techniques in a real-world scenario, thereby establishing their practical significance. These findings invite future research into additional optimization strategies, including but not limited to advanced log analysis, predictive maintenance based on logged data, and integration of logging with automated incident response mechanisms.

As industrial automation and IoT continue to evolve, it's clear that the need for more efficient, secure, and robust data communication systems will continue to increase. By advancing logging techniques and other optimization strategies, we can effectively support this progress and open new avenues for the growth and success of these critical sectors. Through sharing and building upon the insights presented in this workshop paper, the technology community can continue driving towards improved OPC UA deployments, resulting in more efficient and secure systems for the future.

Drawing insights from [SKW21], we are inspired to further optimize the use of OPC UA in diverse contexts. This study investigates Machine-To-Machine (M2M) communication, akin to Direct-To-Consumer (D2C) marketing, for battery-booster charging stations through a setup that involves an OPC UA client and server. The server in their study represents a smart charging station where the client reserves time slots for charging sessions. This configuration prompts us to consider the vast array of potential applications of OPC UA, further underscoring the need for robust logging strategies to manage, debug, and secure these systems effectively. Additionally, the author's use of an OPC UA Metasploit module on two differently configured OPC UA servers is a salient reminder of the security considerations our work must address.

Bibliography

- [Mü20] Mühlbauer, N., Kirdan, E., Pahl, M. O., & Carle, G. (2020, September). Open-source OPC UA security and scalability. In 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) (Vol. 1, pp. 262-269). IEEE.
- [Mü21] Mühlbauer, N., Kirdan, E., Pahl, M. O., & Waedt, K. (2021). Feature-based comparison of open source OPC-UA implementations.
- [Pa21] Paiho, S., Kiljander, J., Sarala, R., Siikavirta, H., Kilkki, O., Bajpai, A., ... & Weisshaupt, T. (2021). Towards cross-commodity energy-sharing communities—A review of the market, regulatory, and technical situation. *Renewable and Sustainable Energy Reviews*, 151, 111568.
- [Sc22] Schindler, J., Belaidi, S., Kirdan, E., & Waedt, K. (2022). Securing javascript runtime of OPC UA deployments. *INFORMATIK 2022*.
- [SKW21] Schindler, J., Kirdan, E., & Waedt, K. (2021). Secure OPC UA Server configuration for smart charging stations.
- [Wa22] Waedt, K., Ben Zid, I., Schindler, J., & Kirdan, E. (2022). Cybersecurity Education Programmes & Laboratories Brainstorming. In *Cybersecurity for Critical Infrastructure Protection via Reflection of Industrial Control Systems* (pp. 100-107). IOS Press.

Controlled Run-Time Adaptivity in Industrial Agent Systems - Challenges and Research Prospects

Jan Sudeikat¹ and Michael Köhler-Bußmeier¹

Abstract: Developing Cyber-physical Systems (CPS) inherently requires enabling run-time adaptivity. These systems integrate physical components, which operate in changeable contexts. In addition, objectives may change, due to socio-technical aspects. Industrial agents have been proposed for enhancing future industrial automation and control systems and integrating agents in industrial systems is an active field of research. Besides enabling technical compatibility of agent concepts and frameworks, the design and orchestration of agent activities have to be fine-tuned and a goal-directed *adaption@run.time* requires that the system can analyze its own structure at run-time; therefore, the system's structure has to be reflected inside the system. Here, we outline current work and research challenges on how explicit organizational modeling can facilitate developing industrial agent systems. We discuss architectural aspects and outline how adaptations of organizations can be enabled, modeled and automated following the MAPE-K approach.

Keywords: Industrial agents, Cyber-physical System, Multi-agent System, Organizational modeling, Self-adaptation, Self-organization, MAPE-K

1 Introduction

Cyber-physical Systems (CPS) are built by designing interrelated *cyber* and *physical* elements and integrating these into coherent structures. It has been argued that augmenting these systems with software agents is attractive as it allows embedding necessitative system properties [Le16]. Among these, self-adaptation, with regard to environmental changes and performance enhancements, is a key property for industrial agent systems and is particularly addressed within embedded, cyber-physical and IoT systems [We23]. Therefore, the integration of (software) agents and multi-agent systems (MAS) in industrial systems is an active field of research. It requires bridging the gap between industrial control systems and distributed multi-agent platforms (e.g. see [Le21; Ló23; Sa19]).

The resulting challenges are discussed in [MGS20], among others these concern design support for cooperative agent groups and purposeful global, i.e. macroscopic system behaviors. Next-level adaptive features are enabled by *autonomous* decision making, but the autonomy of system elements also challenges design practices. In this respect, we advocate the use of organizational models [ASA15; DP13] for orchestrating the co-action of system elements within CPS. Defining a flexible execution overlay for organizations (cf. Sect. 3) allows to adapt the logical relation of agents. Adapting these

¹ Hamburg University of Applied Sciences, Department of Computer Science, Berliner Tor 7, 20099 Hamburg, Germany, jan.sudeikat@haw-hamburg.de

requires a corresponding vehicle (see Sect. 4) based on a unifying meta model and adjusting relations within this overlay.

In this respect, the organization of system elements is a *crosscutting* design aspect. Augmenting systems with a coherent organization requires means to inspect, modify and enact organizations at run-time. In prior works [SK22b], we related modeling organizations to the development of CPS [SK22b] and proposed the use of industrial M2M protocols, i.e. OPC UA (cf. Sect. 2), for this purpose. Here, we build on these efforts and discuss integrating SONAR-based organizations in agent-based CPS. Architectural aspects and implications of enabling industrial adaptive organizations. We particularly outline challenges and work in progress on adapting the organizations themselves, thus adapting adaptive systems.

In the next section, we outline related work. In Section 3, we propose a meta model for integrating organizations in CPS. Here, we argue that there is a mutual benefit for both sides: On the one hand, CPS benefit from Org-MAS as the latter provides a conceptual grounding for the *autonomy vs. coherence* design; on the other hand, OPC UA provides a powerful information model to express all the aspects of an Org-MAS. However, the overlay provides the possibility to modify the organisational overlay, but does not imply any structured process how to do so. Therefore, we need a complementary run-time perspective, which is given by the well-known MAPE-K approach [We18]. This combination is presented in Section 4. From this combination we identify research challenges to specify run-time adaptivity in industrial agent systems in Section 5, before we finally conclude.

2 Related Work

The benefits of integrating agent technology in industrial settings has been identified in numerous works. In [Le16], the use of software agents in cyber-physical systems is motivated and fundamental design principles, based on [FR15], are outlined. In addition, a trend towards applying agents for thorough system control has been identified. Initially, uses of agent technology were focusing on high-level objectives, e.g. planning and scheduling [Le16]. In [Sa19], projects and demonstrators, which apply agents in industrial systems, have been analyzed and pattern for industrial agents have been identified. A predominant concern is accessing physical devices. It has also been shown how the identified agents can be logically located within associated reference architectures, i.e. the *Reference Architecture Model Industrie 4.0* (RAMI 4.0).

However, these efforts are focusing on static settings. The logical architecture, including the logical location and functioning of agents, reflect physical settings and predefined coordination scenarios. This is also reflected by engineering approaches. E.g. in [Le16], design principles for agent-based CPS are discussed. Starting point are high-level (e.g. holonic, bio-inspired, service-oriented) design principles, which are then gradually concretized using reference architectures and leading to comprehensive system

integration. Detailed, structured design practices are an open challenge [MGS20]. We aim at supporting these principles by explicit modeling. In addition, run-time change and adaptation is discussed. Thus, not only the initial design an implementation on of a design principle but also their adjustment at run-time can be addressed.

The design and implementation of organizations is an engineering discipline within the field of MAS and agent-oriented software engineering [ASA15; DP13]. In [Co19], a conceptual model for integrating agents in CPS is proposed. Our work takes inspiration from this approach, since we also integrate agents, physical and software systems in coherent organizations. However, we discuss integrating SONAR-based organizational models [KW09; KWM09]. This approach provides basic formal foundation [Kö07; Kö12] and tool-set for formal analysis [DK10] and thus allows to reason about the effects of organizational adjustments [KS23]. This is particularly beneficial for run-time adjustments of systems where a system abstraction is examined at run-time in order to decided adjustment of the system structure, i.e. the organization.

The *Open Platform Communications Unified Architecture (OPC UA)*² is a comprehensive framework for Machine to Machine (M2M) Communication. Developed as an open international standard it is gaining widespread and tool support in different application domains. It particularly addresses the interoperability by providing an extensible, object- oriented information modeling approach. Nodes provide information and model elements can reference each other, thus leading to a distributed information model [Iv21; Pa16] which particularly facilitates interoperability of system elements.

3 The Need for Organizations as a Logical Overlay

We argued in [SK22b], that organizations, as used in MAS, can benefit the development of agent-based CPS, since the organization principles of an application can be expressed explicitly. This particularly allows to express and reason about *High Level Designs* (cf. Section 2). Supporting this approach leads to the definition of an *organizational overlay* that describes the logical structure of the system (design aspect) and allows to assess the dynamical behavior the system is capable to show (run-time aspect). Enacting this model means to adjust and constrain agents in order to comply.

3.1 Architecture

A common approach to integrating MAS in CPS is conceive agents as an overlay in which agents controls underlying physical assets (see Fig.1, inspired by e.g. [Bo13]). Organizations are commonly understood as an additional layer, e.g. adopting the principles of multi-agent oriented programming [Bo13]. The *Asset Layer* (bottom)

² <https://opcfoundation.org/about/what-is-opc/>

contains the physical assets that are to be managed by software agents. Asset management can be realized by resource access pattern [Sa19]. Here, we distinguish only between $n:1$ (a), $1:1$ (b) and $1:n$ (d) relations between agents and assets. Access is mediated via a *Communication Layer*, which is realized via field bus and industrial communication protocols. Not all agents in the adjacent *Agent Layer* are responsible for controlling assets, (c) these may also provide additional services, like Agent Management Systems (AMS), Directory Facilitators (DF) as well as mediation of third-party components and services [Sa19]. Passive artifacts are to be integrated as well (e), signified by the absence of explicitly managing agents. The topmost *Organization Layer* contains the representation of the organizational structure (see Section 4). A connection link (*Organization Enactment*) allows to execute organizations in heterogeneous MAS. Semantic information about assets, agents and the inter-agent organization are contained in the *Information Model* (left hand side).

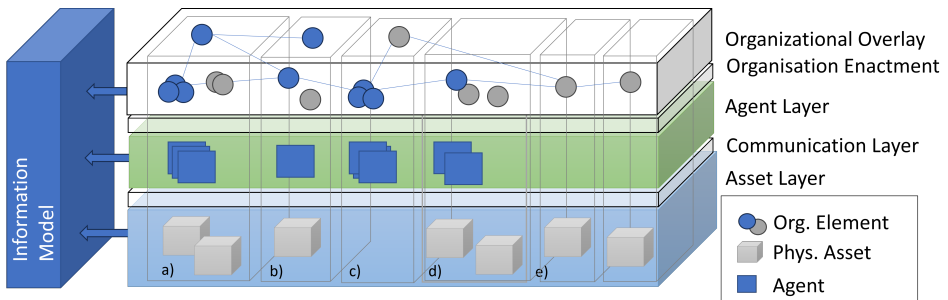


Fig. 1: Typical, layered structure of a multi-agent overlay, supplemented by a spanning information model and connection layers

Since the organization describes the interrelation of system elements, it supplements traditional development approaches which use reference architectures, e.g. the *Smart Grid Architecture Model (SGAM)* and the *Use Case Methodology* [GUD17]. These approaches can be used to build the Asset Layer elements and supporting software structures. Agents can be integrated to provide autonomy for selected system elements [Sa19] and organizations structure the co-action [SK22b]. Providing a comprehensive system reflection, according to [SK22b], the organizational model is a crosscutting concern, used at various implementation levels. Next, we outline the semantic structure of this extended *Information Model*.

3.2 A Unifying Meta Model

A unifying meta model for annotating a CPS with an agent/organization overlay, is shown in Figure 2. Comparable to other models for agent-based cyber-physical models, e.g. [Co19], we aim at integrating agents in multi-paradigm CPS. Thus, our model reflects different types of system elements. Besides physical devices and agents the system may also comprise sets of artifacts, i.e. legacy software systems, which can range

from (micro)-services to actors or third party software systems.

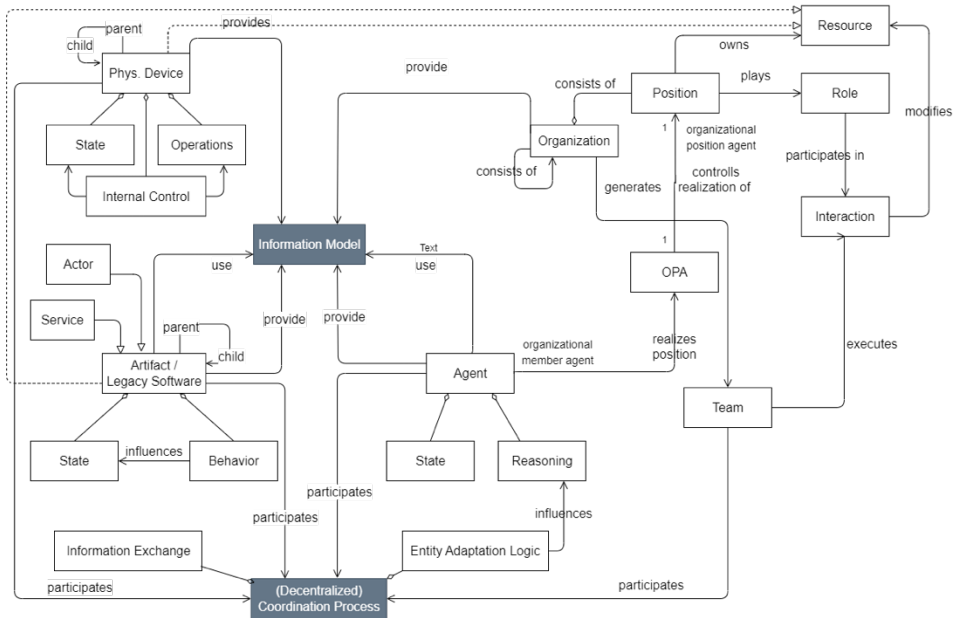


Fig. 2: Metamodel (combining CPS, MAS and SONAR-Organizations via Information Modeling)

Basic building blocks are *Physical Devices* (i.e. assets or units) that are to be controlled and orchestrated. These are typically reactive in nature and have at least a *State*, offer *Operations* and contain differing forms of *Internal Control* e.g. simple control loops for protecting the equipment and maintaining devices operational. In this respect, these operate with differing degrees of autonomy and thus it is a design effort to integrate these system elements into coherent structures, e.g. as argued in [SK22b].

Additional functions of a CPS can be integrated by legacy artifacts, i.e. micro-services, functions, as conceived by using SGAM-based architecture models and Use Case Methodology [GUD17]. Also actor-based system components, as integrated in [Co19], fall into this category.

Sophisticated control functions are realized as software *Agents*, i.e. industrial agents. While we aim at integrating different agent models, these elements will be extended for specific implementation architectures. Basically, *Agents* also contain *State* and *Reasoning*.

Finally, we aim to represent *Organizations* as an explicit modeling construct. The representation of *Organizations* is based on the SONAR formalism [Kö07; KWM09]. Basic building blocks of organizations are so-called *Positions* and *Teams*. *Teams* prescribe and constrain interactions among agents, i.e. *Team* members. *Positions* are

filled by (member) Agents and imply that specific Roles are adopted by agents. A Position may also require access to specific *Resources*. In this context, these are mainly the *Phys. Devices* but also contain computational Artifacts. Associating agents to Positions is done via intermediary agents, so-called *Organization Position Agents* (OPA). These Agents control and constrain the execution of a specific position (1:1 relation). The Positions are actually realized by individual agents, which are called *Organizational Member Agent* (OMA) whenever linked to their OPA. Comparable to a plug-in mechanism these agents join the organization and provide an implementation for the activities and responsibilities of the Position.

In [SK22a], we argued that OPC UA is a practical Framework that allows information modeling at design time as well as model provisioning and model updates at run-time. System analysis at run-time requires expressing the static structure and run-time configuration of the whole system. Thus each system element is providing an *Information Model* about its state. Here, an object-oriented model is used to describes the structure, variables, properties and methods/procedures of system elements.

Agent interplay and team-based interactions are used to adjust the system to external influences. Thus, the system is exhibiting an adaptive *Dynamic Process* [SR08a; SR08b; SR09; Su10]. Building blocks are the *Information Exchange* between system elements and the localized *Entity Adaptation Logic*. Modeling these allows to assess the application dynamic, i.e. the dynamic behavior of the system as is.

4 Run-Time Perspective: Adapting Organizations using MAPE-K

The organizational overlay presented so far provides the possibility to modify the organization. As argued before, we need the MAPE-K loop [We18] as a complementary run-time perspective (see Figure 3). The principal phases for adjusting systems are *monitor*, *analyse*, *plan*, and *execute* [KC03]. Typically, a knowledge-base (K) is associated as well. This is provided by the information model, which is augmented with the configuration of the managed system (cf. Figure 1). The principal phases signify the foundational steps for adapting systems, here organizational structures expressed in the information model. Conceptually, the control of cyber-physical assets and systems (monitoring and execution) are combined with techniques from distributed adaptive systems (analysis and planning).

Initially, we aim at supporting the manual inspection and adjustment of CPS organizations by inspecting and modifying the organization via OPC client applications. The loop delineates automated adjustments at run-time. The so-called *SONAR** overlay, e.g. based on [KW09], allows to enact SONAR-based organizational models. Typically, OPAs observe the system execution and trigger adjustments, i.e., reconfigurations of agents or the restructuring of the organization [KR23b]. The run-time information about the organization under consideration is reflected within an *OPC UA*-based model information model [SK22a] of the organizational structure (*info@run.time*). Thus,

execute means in this context to enact the organization and update the OPC-based model. Monitoring activities allow assessing the compliance of the system at run-time with the model and run-time properties are updated by mining/refining the organizational structure (*Org-Miner*).

Based on the current organization and the perceived activities of agents, the *analyse*-phase identifies problematic configurations and/or the need for adjustments to enhance system performance. This implies a cost/benefit analysis [KR23a; KS23]. When required, the *planning* aims at identifying beneficial organizations. We plan to base this activity on a *models@run.time* approach [BGS19], thus mining activities generate a model-based system abstraction that augments the static structure (*Inf. Model, OPC UA*) with dynamic aspects that allow to infer beneficial organizations [KS23]. Planning activities apply SONAR-based operations [KW09; KWM09] on the model to infer an executable sequence of reconfiguration steps, which allow to establish the target organization. The actual reconfiguration, incl. failure recovery is executed within the overlay, completing the loop.

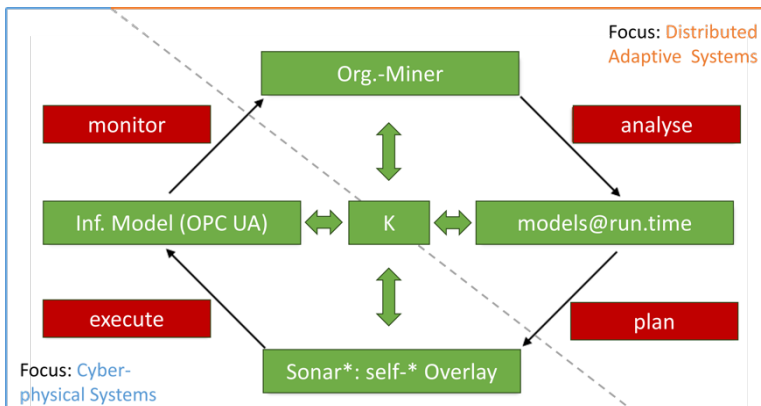


Fig. 3: MAPE-Loop based Self-Adjustment of Organizations

While industrial systems are inherently characterized by a static structure of physical assets and units, it is beneficial to enable flexible overlays which allow to adjust how coordination e.g. cooperation, task allocation, etc. are handled at run-time. Enabling this flexibility is a major argument for agent-based system designs in the first place.³ In [CMM23] the coordination in *local energy markets* and communities are discussed and a systematic taxonomy of categories of coordination approaches is given. Levels of this taxonomy are the *Agency* (direct/indirect control), the *Information* (mediated /

³ 3 Foundational inter-agent organizational models are reviewed in [HL04]. Prime examples are Hierarchies, Holarchies, Coalitions, Teams, Congregations, Societies, Federations, Markets, Matrix and Compound organizations. These modes relate to the high-level design principles, as denoted in [Le16]. For example, markets-based coordination can be understood as an organization and modeled in SONAR [LMV09, Ch. 3].

bilateral / implicit) used and the type of inter-agent *Game* (competition / cooperation). While mainly as a design-time aspect, we aim at adjusting the coordination types at run-time. Another application domain are manufacturing systems, e.g. in [GNF19] auctions have been applied for coordination. Both application domains signify the need to balance individual *autonomy* and organization *coherence*. Individual agents, i.e. market participants, typically do not want to reveal their individual preferences and reasoning.

5 Research Challenges

From our approach to combine run-time views on CPS with MAPE-based self-adaptivity (cf. Fig. 3) – both grounded in organizational MAS – we can identify the following research challenges.

Organization Mining and Evaluation (RC1) In many systems we are faced with the problem that we could not rely on an explicit organizational model as CPS are usually open systems and the integrated components are not developed with an organizational agent model in mind. Therefore, we have a chicken-and-egg situation here as the goal-directed planning of adaption relies on an explicit model of the system. So, our future work will address the mining and/or refining of organizations [Aa16; Xa22]. As a special challenge this mining also takes place at run-time, i.e., we consider on-line learning of organizational models here. Closely related to this is the appropriate evaluation of models as a goal-directed planning relies on the study of the cost/benefit ratio of possible transformations and the benefit is usually measured as the change of the evaluation. Based on initial works [KS23] the estimation of the effects of organizational changes at different time scales are to be examined and revised to guide automated adjustments.

Application Specific Analysis Support for Model Transformation (RC2) We argue to it will be necessary to develop application specific support for the adaption of systems since run-time transformation of systems is a complicated task, especially when properties of the system have to be preserved. Many approaches that cover the transformation of UML-like models are grounded in the theory of graph transformations [Eh06]. However, these approaches do not provide a general methodology how to preserve desired properties. On the other hand, there are approaches for special application domains, e.g. re-organizational aspects of workflow management systems are studied in [KI01; RR07]. However, these approaches are not directly transferable to the domain of CPS. We argue in [KR23b] that analytical as well as simulative approaches are needed to support the analysis and the planning phase of the MAPE-loop. Therefore, we have to study how to introduce self-reflective systems, i.e., we have to represent the organizational models in such a way that we could easily generate new model candidates via self-adaptation and evaluate their value in order to perform a cost-benefit analysis of adaptation.

Organization Embedding and Execution Environment (RC3) In parallel a tool set for the distributed *embedding and enactment* of SONAR-based organizations in industrial

MAS will be revised. We aim at enabling the integration and (semi-)automated deployment in heterogeneous MAS. Analogous to the deployment of agents in industrial settings, as guided in [Le21], deployment options are to be structured and evaluated in order enable informed decisions about embedding realizations of Organizations in agent-augmented cyber-physical Systems. A typical approach to enact organizations (Organization Enactment) is to provide centrally maintained models of the intended organization, e.g. using the artifact approach [Hü09]. In [SR08b; Su12], an architecture for the decentralized provisioning of inter-agent coordination is given. Thus centralized, decentralized, hybrid execution options are to be examined (cf. Section 3.1).

Development Support (RC4) The mapping of development principles and patterns to Organizations is to be examined. Organizational models allow to describe pattern of inter-agent structures in order to reflect best practices in industrial MAS development. Practical application development will be supported by providing a library of coordination media and artifacts for organization enactment (auctions, voting, social choice, etc.). This would supplement works on integrating agent-based reasoning an interaction patterns in industrial systems. In addition, systematic development support is needed. An initial effort is the design of a *testbed* infrastructure to evaluate the use of organizational overlays in-vitro. We aim at using this tool set and modeling approach to derive a systematic process (methodology) for deciding between alternative organizations.

6 Conclusions

In this paper we outlined work in progress to adjust, enact and adapt inter-agent organizations in industrial systems. We outline an approach that combines the widespread and flexible OPC UA framework with software agents and the SONAR organizational model in order to express, model and integrate inter-agent organizations. Consequently, we aim at adapting organizations at run-time, based on the *models@run.time* approach. This work is based on the argumentation that CPS development benefits from the concepts provided by Org-MAS [SK22b] and, that OPC UA provides a powerful information model to express all the aspects of an Org-MAS [SK22a], as a logical supplement to industrial information models.

Here, we discussed consequences and implications. We argued that realizing this supplement requires a complementary run-time perspective, i.e. an organizational overlay. The well-known MAPE-K approach is proposed to adapt organizations provided within a logical overlay, but it does not imply any structured process how to do so. From this combination we identified research challenges to specify run-time adaptivity in industrial agent systems, which we will address in future work.



Bibliography

- [Aa16] van der Aalst, W. M. P.: *Process Mining: Data Science in Action*. Springer, 2016.
- [ASA15] Abbas, H.A.; Shaheen, S.I.; Amin, M.H.: *Organization of Multi-Agent Systems: An Overview*. *Int. Journal of Intelligent Inf. Systems* 4/3, S. 46, 2015.
- [BGS19] Bencomo, N.; Götz, S.; Song, H.: *Models@run.time: a guided tour of the state of the art and research challenges*. *Software & Systems Modeling* 18/5, S. 3049–3082, 2019.
- [Bo13] Boissier, O.; Bordini, R. H.; Hübner, J. F.; Ricci, A.; Santi, A.: *Multi-agent oriented programming with JaCaMo*. *Science of Computer Programming* 78/6, S. 747–761, 2013.
- [CMM23] Charbonnier, F.; Morstyn, T.; McCulloch, M.: *Active Players in Local Energy Markets*. In: *Lecture Notes in Energy*. Springer, S. 71–111, 2023.
- [Co19] Cossentino, M.; Lopes, S.; Renda, G.; Sabatucci, L.; Zaffora, F.: *A Metamodel of a Multi-paradigm Approach to Smart Cyber-Physical Systems Development*. In: *20th Workshop From Objects to Agents*. Parma, Juni 2019.
- [DK10] Deliu, E.; Köhler-Bußmeier, M.: *SONAR/OREDI: A Tool for Creation and Deployment of Organisation Models*. In: *Multiagent System Technologies*. Springer, S. 76–87, 2010.
- [DP13] Dignum, V.; Padget, J.: *Multiagent organizations*. In (Weiß, G., Hrsg.): *Multiagent Systems*, 2nd ed. MIT Press, S. 51–98, 2013.
- [Eh06] Ehrig, H.; Ehrig, K.; Prange, U.; Taentzer, G.: *Fundamentals of algebraic graph transformation*. Springer, 2006.
- [FR15] Farid, A. M.; Ribeiro, L.: *An Axiomatic Design of a Multiagent Reconfigurable Mechatronic System Architecture*. *IEEE Transactions on Industrial Informatics* 11/5, S. 1142–1155, 2015.
- [GNF19] Gehlhoff, F.; Nabizada, H.; Fay, A.: *Optimization of multi-agent auctioning processes in flexible production networks*. In: *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*. IEEE, 2019.
- [GUD17] Gottschalk, M.; Uslar, M.; Delfs, C.: *The Use Case and Smart Grid Architecture Model Approach*. Springer, 2017.
- [HL04] Horling, B.; Lesser, V.: *A survey of multi-agent organizational paradigms*. *The Knowledge Engineering Review* 19/4, S. 281–316, 2004.
- [Hü09] Hübner, J. F.; Boissier, O.; Kitio, R.; Ricci, A.: *Instrumenting multi-agent organisations with organisational artifacts and agents*. *Autonomous Agents and Multi-Agent Systems* 20/3, S. 369–400, 2009.
- [Iv21] Ivanova, T.; Batchkova, I.; Gocheva, D.; Belev, Y.: *Information Modeling of Cyber-Physical Systems using OPC-UA*. In: *2021 International Conference Automatics and Informatics (ICAI)*. IEEE, 2021.
- [KC03] Kephart, J.; Chess, D.: *The vision of autonomic computing*. *Computer* 36/1, S. 41–50, 2003.

- [Kl01] Klarmann, J.: A Comprehensive Support for Changes in Organizational Models of Workflow Management Systems. In: Proceedings of the 4th International Conference on Information Systems Modeling (ISM'01). 2001.
- [Kö07] Köhler, M.: A Formal Model of Multi-Agent Organisations. *Fundamenta Informaticae* 79/3-4, S. 415–430, 2007.
- [Kö12] Köhler-Bußmeier, M.: Analysing SONAR Model Transformations. In (Accorsi, R.; Murata, T.; Ranise, S., Hrsg.): Workshop on Petri Net-based Security (WOOPS 2012). Bd. 853, CEUR Workshop Proceedings, S. 55–70, 2012.
- [KR23a] Köhler-Bußmeier, M.; Rölke, H.: Analysing Adaption Processes of Hornets. *Transactions on Petri Nets and Other Models of Concurrency XVII* 14150/, hrsg. von Koutny, M.; Ciardo, G.; Bergenthum, R., 2023.
- [KR23b] Köhler-Bußmeier, M.; Rölke, H.: Petri-Nets@Run.Time: Handling Uncertainty during Run-Time Adaptation using Digital Twins. In: *Petri Nets and Software Engineering 2023*, PNSE'23. Bd. 3430, CEUR, S. 34–52, 2023.
- [KS23] Köhler-Bußmeier, M.; Sudeikat, J.: Balance vs. Contingency: Adaption Measures for Organizational Multi-agent Systems. In: *Intelligent Distributed Computing XV*. Springer, S. 224–233, 2023.
- [KW09] Köhler-Bußmeier, M.; Wester-Ebbinghaus, M.: SONAR*: A Multi-Agent Infrastructure for Active Application Architectures and Inter-Organisational Information Systems. In: *Conference on Multi-Agent System Technologies, MATES 2009*. Bd. 5774, S. 248–257, 2009.
- [KWM09] Köhler-Bußmeier, M.; Wester-Ebbinghaus, M.; Moldt, D.: A Formal Model for Organisational Structures behind Process-Aware Information Systems. *Transactions on Petri Nets and Other Models of Concurrency*. 5460/, S. 98–114, 2009.
- [Le16] Leitao, P.; Karnouskos, S.; Ribeiro, L.; Lee, J.; Strasser, T.; Colombo, A. W.: Smart Agents in Industrial Cyber-Physical Systems. *Proceedings of the IEEE* 104/5, S. 1086–1101, 2016.
- [Le21] Leitão, P.; Strasser, T.; Karnouskos, S.; Ribeiro, L.; Barbosa, J.; Huang, V.: Recommendation of Best Practices for Industrial Agent Systems based on the IEEE 2660. Standard. In: *22nd IEEE Int. Conf. on Industrial Technology (ICIT)*. IEEE, S. 1157–1162, 2021.
- [LMV09] v. Lüde, R.; Moldt, D.; Valk, R., Hrsg.: *Selbstorganisation und Governance in künstlichen und sozialen Systemen*. Lit Verlag, Münster, 2009.
- [Ló23] López, A.; Casquero, O.; Estévez, E.; Armentia, A.; Orive, D.; Marcos, M.: An industrial agent-based customizable platform for I4.0 manufacturing systems. *Computers in Industry* 146/, S. 103859, 2023.
- [MGS20] Marik, V.; Gorodetsky, V.; Skobelev, P.: Multi-Agent Technology for Industrial Applications: Barriers and Trends. In: *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020.
- [Pa16] Pauker, F.; Frühwirth, T.; Kittl, B.; Kastner, W.: A Systematic Approach to OPC UA Information Model Design. *Procedia CIRP* 57/, S. 321–326, 2016.

- [RR07] Rinderle-Ma, S.; Reichert, M.: A Formal Framework for Adaptive Access Control Models. In: *Journal on Data Semantics IX*. LNCS, Springer, S. 82–112, 2007.
- [Sa19] Salazar, L. A. C.; Ryashentseva, D.; Lüder, A.; Vogel-Heuser, B.: Cyber-physical production systems architecture based on multi-agent’s design pattern—comparison of selected approaches mapping four agent patterns. *The Int. Journal of Advanced Manufacturing Technology* 105/9, S. 4005–4034, 2019.
- [SK22a] Sudeikat, J.; Köhler-Bußmeier, M.: Towards Integrating Multi-Agent Organizations in OPC UA for Developing Adaptive Cyber-Physical Systems. In: *INFORMATIK 2022*. Gesellschaft für Informatik, Bonn, 2022.
- [SK22b] Sudeikat, J.; Köhler-Bußmeier, M.: On Combining Domain Modeling and Organizational Modeling for Developing Adaptive Cyber-Physical Systems. In: *Proceedings of the 14th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART, INSTICC, SciTePress*, S. 330–336, 2022.
- [SR08a] Sudeikat, J.; Renz, W.: Building Complex Adaptive Systems. In: *Applications of Complex Adaptive Systems*. IGI Global, S. 229–256, 2008.
- [SR08b] Sudeikat, J.; Renz, W.: On the Encapsulation and Reuse of Decentralized Coordination Mechanisms: A Layered Architecture and Design Implications. *Communications of SIWN* 4/, S. 140–146, Juli 2008.
- [SR09] Sudeikat, J.; Renz, W.: MASDynamics: Toward Systemic Modeling of Decentralized Agent Coordination. In: *Kommunikation in Verteilten Systemen (KiVS)*. Springer Berlin Heidelberg, S. 79–90, 2009.
- [Su10] Sudeikat, J.: Systemic Modeling of Agent Coaction: A Catalog of Decentralized Coordinating Processes. *Electronic Comm. of the EASST* 27/, 2010.
- [Su12] Sudeikat, J.; Steghöfer, J.-P.; Seebach, H.; Reif, W.; Renz, W.; Preisler, T.; Salchow, P.: On the combination of top-down and bottom-up methodologies for the design of coordination mechanisms in self-organising systems. *Information and Software Tech.* 54/6, S. 593–607, 2012.
- [We18] Weyns, D.: Engineering Self-Adaptive Software Systems - An Organized Tour. In: *3rd International Workshops on Foundations and Applications of Self*Systems (FAS*W)*. IEEE, 2018
- [We23] Weyns, D.; Gerostathopoulos, I.; Abbas, N.; Andersson, J.; Biffi, S.; Brada, P.; Bures, T.; Salle, A. D.; Galster, M.; Lago, P.; Lewis, G.; Litoiu, M.; Musil, A.; Musil, J.; Patros, P.; Pelliccione, P.: Self-Adaptation in Industry: A Survey. *ACM Trans. on Autonomous and Adaptive Systems* 18/2, S. 1–44, 2023.
- [Xa22] Xavier, M.; Dubinin, V.; Patil, S.; Vyatkin, V.: Process mining in industrial control systems. In: *2022 IEEE 20th Int. Conf. on Industrial Informatics (INDIN)*. IEEE, 2022.

Reasoning about Causal Effects of Regulation and Legislation on Interconnected Markets

Christoph Legat¹ , Uwe Seebacher², Dominik Brunner³ 

Abstract: Regulatory interventions have far-reaching implications beyond the directly affected market, creating dependencies and interconnections among various markets. Understanding the causal effects and dependencies between different markets arising from regulatory actions is crucial for policymakers, industry stakeholders, and researchers.


This paper presents a novel causal effect model specifically designed to reason about the impact of regulation and legislation, with a specific focus on the dependencies between different markets beyond the one directly affected by regulation. The proposed model provides a comprehensive framework for analyzing the consequences of regulatory interventions across interconnected markets. By employing advanced causal inference techniques, the model enables the identification and quantification of the causal effects that emerge between markets due to regulatory actions.

Keywords: Regulation, Interconnected Markets, Causal Effects, Effect Analysis, Prediction


1 Introduction

Regulation plays a pivotal role in shaping markets, ensuring fair competition, protecting consumers, and promoting overall economic stability. However, the impact of regulation extends beyond the directly affected market, influencing interconnected markets, and creating dependencies among them. Understanding the intricate relationships and causal effects that arise between different markets due to regulatory interventions is crucial for policymakers, industry stakeholders, and researchers. This scientific paper presents a causal effect model specifically designed to reason about the impact of regulation, with a specific focus on the dependencies between different markets beyond the ones directly affected by regulation.

While existing research has explored the effects of regulation within individual markets, there is a gap in understanding how regulatory measures in one market can cascade and affect related markets. For instance, changes in regulations governing the financial sector can have ripple effects on industries such as real estate, insurance, and even technology. These interdependencies create complex dynamics that necessitate a comprehensive

¹ Augsburg Technical University of Applied Sciences, Augsburg, Germany, christoph.legat@hs-augsburg.de,  <https://orcid.org/0000-0002-5843-1845>

² HM Hochschule München University of Applied Sciences, Munich, Germany, uwe.seebacher@hm.edu

³ predictors.ai GmbH, Graz, Austria, dominik@predictores.ai,  <https://orcid.org/0009-0007-0814-1958>

causal effect model to analyze the consequences of regulatory actions across multiple markets.

The proposed model aims to fill this gap by providing a systematic and formal framework to reason about the impacts of regulation beyond the direct market. By employing advanced causal inference techniques, the model allows for the identification and quantification of the causal effects that emerge between interconnected markets as a result of regulatory interventions. This enables policymakers and industry stakeholders to gain deeper insights into the wider ramifications of regulatory actions and make informed decisions that consider the broader market ecosystem.

The objectives of this paper are twofold. Firstly, we aim to introduce the causal effect model specifically tailored to analyze the impact of regulation across interconnected markets. We will outline the key components of the model, including data collection methodologies, variables considered, and the causal inference techniques employed. Secondly, through the application of the model in concrete example, we aim to demonstrate its effectiveness in capturing the dependencies and causal effects between different markets arising from regulatory interventions.

By presenting this causal effect model, we seek to contribute to the field of regulatory analysis by offering a framework to reason about the dependencies between different markets beyond the one directly affected by regulation. The findings derived from this model can guide policymakers in understanding the potential spillover effects of regulatory interventions and help them design more holistic and effective regulatory strategies. Additionally, this research may inspire further studies and discussions on the interconnections between markets and the role of regulation in shaping broader market ecosystems.

The remainder of the paper is structured as follows: Sect. 2 introduces the general relationships between regulation, legislation, and standardization. The novel framework for modelling and analyzing causal relationships between interconnected markets and regulatory effects is presented in Sect. 3. Using a concrete application example, Sect. 4 explains the functioning of the model and, by that, evaluates its applicability. Summary and outlook are given in Sect. 5.

2 Perspectives on Standardization and Regulation

The central political intuition in Europe is the European Union (EU), which is an independent legal entity. The EU is an association of 27 European states. The political system of the EU contains both supranational and intergovernmental elements. This is reflected in the organization of the organs of the European Union: The European Council and the Council of the European Union represent individual states through their governments; the European Parliament, as the legislative body of the EU, directly represents the citizens of the European Union. The European Commission represents the

executive body, and the EU Court of Justice represents the judicial authority in the form of supranational institutions.

In European law, a distinction is made between directives and regulations. Regulations are a legal act of the European Union with general validity and direct effectiveness in the member states. An example is EU Regulation 2017/745 [Eur17] on the clinical investigation and sale of medical devices for human use. Directives are also legal acts of the European Union but are not directly applicable and must be transposed into national law by the member states. Directives are usually adopted jointly by the Council of the European Union and the European Parliament in accordance with the ordinary legislative procedure on a proposal from the European Commission. One example is the so-called Machinery Directive, Directive (EU) 2006/42/EC [Eur06], which regulates a uniform protection level for accident prevention for machines and partly completed machines when they are placed on the market. Similar examples are the AI Act proposal [Eur21], NIS 2 [Eur22d], Digital Markets Act [Eur22c] or the ESPR [Eur22b].

In a nutshell, regulators and public authorities create the policy framework and set legal requirements, e.g. product safety, environment protection. Accordingly, no further distinction between directives and regulations is made in the remainder of this paper. Rather, regulation is understood in the further course to mean all legal regulations. In this context, standards serve as a base defining the state-of-the-art solution to achieve defined policies and legal requirements.

In other countries and markets beyond Europe, the mechanisms are similar but may be slightly different but have major core relationships in common: Regulatory authorities and/or governments set legal requirements for the market they are responsible for. Some kind of specifications, typically preferably international or national standards, are used to refine these regulatory and/or legal requirements and may either be mandatory or recommended.

Ultimately, however, in the context of the causality of legislation and regulation considered here, the result is identical: the (direct or indirect) influence of an authority on the mandatory characteristics of products or services traded or marketed in the market.

3 Framework to Reason about Market Effects by Regulation

3.1 Market Upper Ontology

A general but simplified economic model is used which is optimized and aligned to the objective to reason about causal dependencies and the impact of regulation. The core conceptual elements and their relations are informally visualized in Fig. 1.

In hypothetical markets, things of all kinds are traded, e.g. objects, services, or other intangible goods. The concrete type of what is traded on a market is referred to in the

following as an asset. A market is characterized by the sales achieved and the demand for the assets under consideration. Consequently, concrete sales and concrete demands are explicitly related to a specific type of an asset. To characterize an asset and facilitate to compare or distinguish it from other assets, a uniform meta-model for describing assets is required. In literature, especially in standardization [DI20, PIZ23] and symbolic artificial intelligence [BK12, Ho05], so-called property value statements are proposed and applied which facilitates to describe elements, i.e. in context of this paper products, by means of their properties and their property-values.

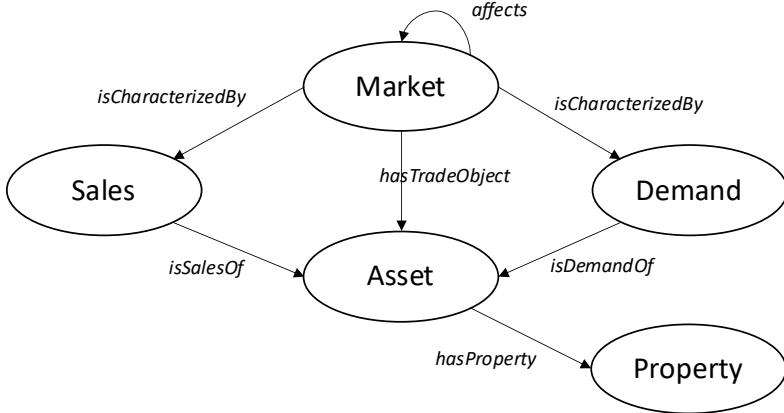


Fig. 1: Market Upper Ontology

The market upper ontology defined in this paper combines Knowledge Graphs [Ho21] with the strict Description Logics [BHS08] semantics of ontologies. An excerpt of the core axioms the market upper ontology's Description Logics formulation is given subsequently.

$$\begin{aligned} \mathbf{Market} \sqsubseteq & \exists \text{ isCharacterizedBy. Demand } \sqcap \\ & \exists \text{ isCharacterizedBy. Sales } \sqcap \\ & \exists \text{ hasTradeObject. Asset} \end{aligned} \quad (1)$$

$$\mathbf{Asset} \sqsubseteq \exists \text{ hasProperty. Property} \quad (2)$$

$$\mathbf{Sales} \sqsubseteq \exists \text{ isSalesOf. Asset} \quad (3)$$

$$\mathbf{Demand} \sqsubseteq \exists \text{ isDemandOf. Asset} \quad (4)$$

According to the background analysis (cf. Sect. 2Fehler! Verweisquelle konnte nicht gefunden werden.), regulations define by law prerequisites on assets' mandatory requirements, i.e. with respect to the Market Upper Ontology above, prerequisites on its properties.

3.2 Causal Effect Model

Let P and C disjunct sets of vertices, i.e. $P \cap C = \emptyset$. The set P of places $p \in P \subseteq \mathbb{P}$ represent any kind of data within a certain value range $\mathbb{W}(p)$. Respective values of places $p \in P$ are given by the set $V \subseteq \mathbb{P} \times \mathbb{W}$ so that for a place $p \in P$ the value $v \in \mathbb{W}(p)$ is given by $(p, v) \in V$. For simplifying the notation, the function $\mathbb{V}: P \rightarrow \mathbb{W}$ represents the value $v = \mathbb{V}(p)$ for a given place P based on the tuples $(p, v) \in V$. The set of causalities C represent causal dependencies within the formal model between one or more places $p \in P$. Causal dependent places by means of causalities are given by the set R of relations with $R \subseteq (P \times C) \cup (C \times P)$. For any causality $c \in C$, a transformation function $f_c: V \rightarrow V \in F$ exists which defines the causal mapping between values of places based on the defined causal dependency so that $\forall p_1, p_2 \in P: \forall c \in C: (p_1, c), (c, p_2) \in R \Rightarrow f_c(\mathbb{V}(p_1)) = \mathbb{V}(p_2)$ as well as $\forall (p_1, v_1), (p_2, v_2) \in V: \exists (p_1, c), (c, p_2) \in C \Rightarrow v_2 = f_c(v_1)$.

A Causal Effect Net (CEN) is, based on the concepts introduced above, defined as $z = (P, C, R, F) \in Z$ with a set of places P , a set of causalities C , a set of relations R , and a set of transformation functions F . Any CEN is element of the superset Z of all possible CENs.

CEN focuses on the causal dependencies between places and, by that, can represent interrelationships which are generally valid without further (external) influences. Extending the notion of CEN, a valued CEN (vCEN) contains additional data values of places $p \in P \subseteq \mathbb{P}$ which are given by the set $V \subseteq \mathbb{P} \times \mathbb{W}$ so that for a place $p \in P$ the value $v \in \mathbb{W}(p)$ might be given by $(p, v) \in V$ and $\mathbb{W}(p)$ the set of possible values, i.e. the value range. In a nutshell, a vCEN is a 4-tuple (P, C, R, F, V) with (P, C, R, F) forming a CEN and an additional set of associated values for places $V \subseteq \mathbb{P} \times \mathbb{W} \subseteq \mathbb{P} \times \mathbb{W}$.

Regulation is, by definition, the exertion of governmental influence to shape (internal) markets, i.e. manipulate values of interest. Accordingly, in context of the causality model described here, regulation is considered as a set of expected values $\emptyset \neq V_{\text{NEW}} \subseteq \mathbb{P} \times \mathbb{W}$ where $(p, v) \in V_{\text{NEW}}$ describe the expected outcome of the regulation for one or more aspects of interested, i.e. places with recourse to a corresponding CEN. By that, the vCEN model can be used to apply the regulatory effect by forming a vCEN $(P, C, R, F, V_{\text{NEW}})$. Based on the transformation functions F of the vCEN, the impact of the regulation to further aspects beyond the regulatory target can be inferred. Concretely, all values can be derived for places that can be computed by causal dependencies.

The reachability graph $\Psi(P, C, R, V) = (\pi, \varepsilon)$ of a vCEN (P, C, R, F, V) is defined as a set of vertices $\pi \subseteq P$ and a set of directed edges $\varepsilon \subseteq \pi \times \pi$. Based on the set of valued places of the vCEN given by $PV = \{p \in P \mid \exists v \in \mathbb{W} : (p, v) \in V\}$ and a set of directed edges E derived from vCEN as $E = \{(p_1, p_2) \in P \times P \mid \exists (p_1, c), (c, p_2) \in R\}$, the set of vertices π is defined as set of reachable places and the set edges ε is defined as the set of edges between reachable vertices π . The reachability graph defines the directly affected, interconnected markets.

4 Application Scenario

For evaluating and highlighting the applicability of the proposed approach to reason on causal effects of interdependent markets, a simplified version of an intensively discussed, recently announced regulatory intervention in the European market is used. As part of the European Commission's efforts to reduce greenhouse gas emissions, the "Fit for 55" initiative [Eur23a] has been launched. The aim is to ban the production of vehicles with combustion engines that run on fossil fuels from 2030. It is expected that this regulation will lead to a dramatic decrease in sales of new cars of combustion engines for fossil fuels. Assuming that the overall demand for (new) cars remains almost constant, the demand for cars with alternative drives (e.g. electrical or hydrogen) will arise. Beyond these obvious influences on the market, which are basically the intended objective of the regulation, other industries and markets will be affected by this regulatory intervention. Which markets are ultimately affected by this market intervention and how they are influenced (positively or negatively) can be examined by the mechanism presented in this paper. This is described in more detail in the following. To illustrate how the proposed concept is applied, the scope of the analysis has been limited to Germany, although this does not detract from its general validity.

In context of the application example considered here, the causal effect model is used for interconnected markets, i.e. causal dependencies are considered with respect to demand and sales of specific assets described in a domain ontology, here the Vehicle Domain Ontology, according to the Market Upper Ontology. This is described in detail in Sect. 4.1. Afterwards the derivation and definition of the causal effect model with focus on vehicles with fossil combustion engines described in Sect. 4.2; its application to reason about causal effects is presented in Sect. 4.3.

4.1 Vehicle Domain Ontology

An excerpt of the personal car domain ontology is given Fig. 2. Therein, an exemplary vehicle *asset_123* is given which is given which is *ComposedOf* an *CombustionEngine* using *Diesel*, i.e. such vehicles would be affected with the regulatory task considered here.

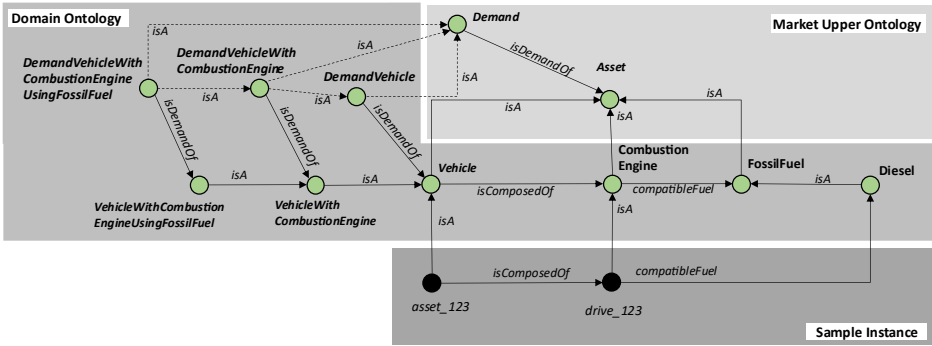


Fig. 2: Excerpt of the Vehicle Domain Ontology in relationship to the Market Upper Ontology including a sample instance.

For illustration in this example, additional concept definitions have been added for vehicles with combustion engine in general and fossil fuel powered vehicles in particular. The latter represents the vehicles affected by the regulation under consideration.

$$\begin{aligned} \mathbf{VehiclesWithCombustionEngine} &\equiv \mathbf{Vehicle} \sqcap \\ &\quad \exists \text{ isComposedOf. } \mathbf{CombustionEngine} \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{VehiclesWithCombustionEngineUsingFossilFuel} &\equiv \\ &\quad \mathbf{VehiclesWithCombustionEngine} \sqcap \\ &\quad \exists \text{ isComposedOf.} \\ &\quad \quad (\mathbf{CombustionEngine} \sqcap \exists \text{ compatibleFuel. } \mathbf{FossilFuel}) \end{aligned} \quad (6)$$

According to the equivalence definition, any *Vehicle* which *isComposedOf* a **CombustionEngine** which is compatible with **FossilFuel** will be classified by the classical reasoning task “classification”. Furthermore, in Fehler! Verweisquelle konnte nicht gefunden werden., the relationship between the Vehicle Domain Ontology and the Market Upper Ontology by means of the *isDemandOf* relation is depicted. Due to the formal nature of the Market Upper Ontology, it is automatically inferred by the reasoner that (a) the demands of vehicles and the specialized concepts introduced above are *Demands* according to the upper ontology and are related in a sub-concept, i.e. subset, relation similar to the conceptual definitions of the vehicles (indicated by the dashed lines). Analogously, sales are handled in the similar way.

4.2 Causal Effect Model for Vehicles with Fossil Combustion Engines

For building up the causal model, and, in particular, for the derivation of the causal dependencies between markets, German law was utilized, which requires an official registration of vehicles. Thereby, a large amount of data is collected, which is publicly available online free of charge. These statics about registrations of new cars are used to derive the sales market (as time series) of all passenger cars in Germany. For identifying

interconnected markets, spare part lists are used to exemplarily derive dependencies based on the supply chain. In accordance with the Market Upper Ontology, a market is defined by means of the asset, which is traded on the market and, by that, any spare part (i.e. asset) of a vehicle is defining a market considering the specific characteristics of the spare part. The transformation function between the markets is derived by means of the quantity of the spare part of a respective type required within the vehicle. In other words, if for example three windscreen wiper are required for assembling a respective car, it results in the transformation function $f_1(x) = 3 \cdot x$ between the sales of the markets and $f_2(x) = 3 \cdot x$ between the demands of the markets.

4.3 Reasoning about causal effects of the regulation of Fossil Combustion Engines in Europe

The ban of vehicles with combustion engines using fossils fuels as considered here will result in a sales market in Europe towards zero. In accordance with the property-based description of assets as included within the Market Upper Ontology, sales for all assets which are **VehiclesWithCombustionEngineUsingFossilFuel** will be reduced accordingly. Applying the causal effect model as exemplified in Sect. 4.2 using the causal effect model CEN will facilitate the reasoning among the impact of interconnected markets for windscreen wiper by applying the transformation function f_1 . The respective reasoner based upon the CEN formalism will determine that the dependent market share of windscreen wipers related to vehicles with fossil fuel combustion engines will be reduced also to zero in the European market, i.e. $f_1(0) = 3 \cdot 0 = 0$. Nevertheless, as discussed previously, this is not the overall sales market of windscreen wipers, for sure. The overall sales market, which is formed, among others, by the supply chain demand from all vehicles (in our simplified example) will be not directly affected. Nevertheless, assuming that the demand for vehicles remain constant, it is possible to derive by CEN that the overall demand and the overall sales for windscreen wiper (in context of vehicles) remain stable. Accordingly, a market shift is expected from vehicles with combustion engines for fossil fuels to vehicles with alternative drives. This is conformant with the expectations of the regulatory intention.

5 Conclusion & Outlook

This paper presented the challenge of how to evaluate the impact of regulatory and legislative influences on markets. Typically, regulatory influences address specific markets. The model proposed here can be used to examine the influence of regulation in interlinked economic areas or markets. By means of an exemplary application for the ban on the production of fossil fuel vehicles, it was shown that possible influences of regulation can be successfully examined by means of structured modelling using the approach proposed here of combining ontologies and a specific causal effect model.

References

- [BHS08] Baader, F.; Horrocks, I.; Sattler, U.: Description logics. In: Foundations of Artificial Intelligence, Vol. 3, pp. 135–179, 2008.
- [BK12] Bauer, F.; Kaltenböck, M.: Linked Open Data: The Essentials: A Quick Start Guide for Decision Makers, 2012.
- [DI20] DIN e.V., German Institute for Standardization: DIN SPEC 92000 - Data Exchange on the Base of Property Value Statements, August 2020.
- [Eur06] European Parliament; European Council: Directive 2006/42/EC of the European Parliament and the Council of 17 May 2006 on machinery, and amending directive 95/16/EC, In: Official Journal of the European Union, L.157, pp. 24-86, June 2006.
- [Eur17] European Parliament; European Council: Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending directive 2001/83/EC, regulation (EC) No 178/2002 and regulation (EC) No 1223/2009 and repealing Council directives 90/385/EEC and 93/42/EEC, In: Official Journal of the European Union, L. 117, pp. 1-175, May 2017.
- [Eur21] European Parliament; European Council: Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts, COM(2021) 206 final, 2021/0106 (COD), April 2021.
- [Eur22b] European Parliament; European Council: Proposal for a regulation of the European Parliament and Council establishing a framework for setting ecodesign requirements for sustainable products and repealing Directive 2009/125/EC, COM(2022) 142 final, 2022/0095 (COD), March 2022.
- [Eur22c] European Parliament; European Council: Regulation (EU) 2022/1925 of the European Parliament and Council of 14 September 2022 on contestable and fair markets in the digital sector and amending directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), In: Official Journal of the European Union, L. 265, pp. 1-66, October 2022.
- [Eur22d] European Parliament; European Council: Directive (EU) 2022/2555 of the European Parliament and Council of 14 December 2022 on measures for a high common level of cybersecurity across the union, amending regulation (EU) No 910/2014 and directive (EU) 2018/1972, and repealing directive (EU) 2016/1148 (NIS 2 directive), In: Official Journal of the European Union, L. 333, pp.80-152, December 2022.
- [Eur23] European Council: Fit for 55 - The EU's plan for a green transition - Consilium, www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition, visited: 25.08.2023.
- [Ho05] Horrocks, I.; Parsia, B.; Patel-Schneider, P.; Hendler, J.: Semantic web architecture: Stack or two towers? In: International Workshop on Principles and Practice of Semantic Web Reasoning. pp. 37–41, Springer, 2005.
- [Ho21] Hogan, A.; Blomqvist, E.; Cochez, M.; d'Amato, C.; de Melo, G.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S.; Polleres, A.; Ngomo, A.-C. N.; Polleres,

A.; Rshid, S.M.; Rula, A.; Schmelzeisen, L.; Sequeda, J.; Staab, S.; Zimmermann, A.: Knowledge graphs. In: ACM Computing Surveys Vol. 54, pp. 1–37, 2021.

[PIZ23] Plattform Industrie 4.0; Industrial Digital Twin Association; ZVEI: Details of the Asset Administration Shell - Part 1: The exchange of information between partners in the value chain of Industrie 4.0, Specification, Federal Ministry for Economic Affairs and Climate Action (BMWK), Version 3.0RC02, 2023.

Towards a Standardization Ecosystem for Industry 4.0

Selected Characteristics for Categorizing Standardization, Publication and Participation


Christoph Legat¹ , Marvin Böll²³

Abstract: The proliferation of standards, norms and specifications across different industries poses a significant challenge to organizations trying to navigate this complex landscape. This paper addresses this challenge by identifying and discussing selected characteristics that facilitate the distinction of different publication formats such as specifications, norms, and standards. It also examines the inherent characteristics of their development process, including consensus mechanisms, levels of influence and accessibility. In addition, inherent characteristics of the development process for standards, norms and specifications is discussed. The level of consensus is identified as a critical factor, as it reflects the degree of agreement among stakeholders during the development phase. The degree of influence exerted by different organizations or bodies is examined, shedding light on the authority and influence of different publications and organizations. Accessibility is identified as another important characteristic. The study of these characteristics is a first step towards a holistic navigation mechanism for users, experts, and organizations to gain a better understanding of the diverse landscape of standards, norms, and specifications, thus enabling informed decision-making on which publications to follow, which to prioritize, and where to bring in ideas and challenges to be addressed in the context of open innovation. In addition, the results can provide guidance to standards bodies on recent shortcomings in their offerings and limitations in current models of collaboration with other standardization and research institutions.

Keywords: Standardization, Norms, Industry 4.0, Ecosystem, Navigation

1 Introduction

In today's interconnected world, the seamless integration and compatibility of diverse systems, products, and services are vital for fostering innovation, driving economic growth, and ensuring the well-being of societies. To achieve this harmonization, standard setting organizations have played a crucial role by developing and publishing standards, norms, and specifications that provide a common framework for various industries and sectors.

¹ Augsburg Technical University of Applied Sciences, Augsburg, Germany, christoph.legat@hs-augsburg.de,  <https://orcid.org/0000-0002-5843-1845>

² DKE – German Commission for Electrotechnical, Electronic & Information Technologies, Standardization Council Industrie 4.0, Offenbach, Germany, marvin.boell@vde.com

³ DKE – German Commission for Electrotechnical, Electronic & Information Technologies, Standardization Council Industrie 4.0, Offenbach, Germany, marvin.boell@vde.com

Traditionally, global standards have been established by internationally recognized bodies such as the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). Their efforts have enabled the development of universal guidelines that facilitate interoperability, safety, and quality assurance across a wide range of domains. However, in recent years, a new trend has emerged, with associations assuming the role of standardization organizations and contributing to the growing jungle of standards. Nevertheless, standardization plays a crucial role nowadays for innovation [EC21, Le22, BRS12].

The proliferation of these emerging associations as standardization organizations has introduced both opportunities and challenges. On one hand, this diversification has allowed for the development of standards tailored to specific industries, addressing their unique needs, and accelerating technological advancements. On the other hand, navigating through this ever-expanding jungle of standards has become increasingly complex, creating significant challenges for stakeholders seeking to comply with multiple sets of requirements. Making the choice where to participate to innovate and develop standards and/or other types of specifications and publications is challenging [Wi22]. Nevertheless, standardization as well as consortia standardization – especially on international level – are important activities from industrial perspective [ZV21].

This paper aims to shed light on the challenge of navigating through the jungle of standards, norms, and specifications by analyzing major characteristics to be considered for diversifying different types of standards, norms, and specifications with special focus on Industrie 4.0.

The paper is structured as follows: In the subsequent Sect. 2, the term standardization is first analyzed and examined with reference to German framework as a kind of blueprint, before the basic mechanisms of international standardization are introduced. Afterwards, standardization in the context of Industry 4.0 is summarized. Building on this, the following Sect. **Fehler! Verweisquelle konnte nicht gefunden werden.**³ discuss key characteristics of standardization and derive resulting challenges. Finally, the paper is summarized and concluded in Sect. 4.

2 Terminology & Related Works

2.1 Terminology

The principles of standardization work in the Federal Republic of Germany are regulated within the DIN 820 norm series. Therein, in Germany, is a precise, conceptual distinction of technical specifications with respect to the consensus level and the involvement of all interested parties: A norm is the result of a process, in which tangible and intangible objects are normed by consensus for the benefit of the public. This means that all parties interested in the topic participate in the process in a planned and collaborative manner, the

public is involved, and the public is involved in the process [DI21] for achieving highest possible acceptance of the results. In contrast, standardization means the setting of technical rules without a mandatory involvement of all interested parties, without the obligation to involve the public and without the obligation to reach a consensus.

In Germany, three different terms in context of the technical rules for the design and execution of structural installations or technical objects exist: ‘recognized rules of technology’, ‘state of the art in technology’, and ‘state of the art in science and technology’. In a judgement (the so-called Kalkar decision) of the German Federal Constitutional Court in context of residual risk despite all technical, procedural, and economic risk reductions undertaken, a three-stage relationship is defined: Firstly, the lowest level are ‘recognized rules of technology’ which must be generally recognized and, because of this broad professional consensus, take up innovations and technical advances relatively late. Secondly, the ‘state of the art in technology’ is at the second level, which dispenses with such recognition and therefore helps technical innovations to gain acceptance more quickly. Thirdly, the ‘state of science and technology’ comprises the latest technical and scientific knowledge and is not limited by what is currently realized and feasible. It is therefore the highest level, but also the most difficult to determine, since in individual cases the technical and scientific state of knowledge must be specified with the decision of disputes. These considerations are relevant from a legal perspective in Germany but also provides a solid framework to distinguish the broad scope of technical and scientific findings.

Nevertheless, this describes the situation in Germany and is not generally valid for other countries or regions. Furthermore, the terminological separation between norms and standards is not spread and consistently used within the international discussions. In contrast, the term standardization is used mostly for any kind of specification [AB19, BPR17, Ka16].

2.2 Mechanisms of International Standardization

To provide necessary common understanding of the remainder of the paper, a brief overview of the mechanism is presented for both types of standardization activities as identified in the previous section.

- **International Standardization and Institutions:**

The so-called World Standards Cooperation (WSC) was established in 2001 by the International Electrotechnical Commission (IEC), the International Organization for Standardization (ISO), and the International Telecommunication Union (ITU) to strengthen voluntary consensus-based

international standards and working together to implement the strategic role of international ICT standards.

ISO and IEC are independent, non-governmental organizations made up of members from the national standardization bodies. Beyond various governmental and organizational structures, technical works is done within technical committees⁴. If a member, i.e. national standardization body, is involved in a technical committee, a national mirror committee within this standardization body exists. Members of a national mirror committee, i.e. natural persons, are involved in the technical works. Decisions are brought about by voting of national committees.

ITU is organized slightly different and made up of member countries, their industry and academia. The state membership is administered typically a governmental institution, e.g. the Federal Ministry for Digital and Transportation of Germany or the International Information and Communication Policy for the US [IT23].

In context of legislation and regulation, standardization organizations listed above play a dedicated role because (in most countries and markets) they are contracted, asked and/or prompted to develop standards defining the state of technology to comply with legal and regulatory frameworks. Even there are slight differences across the world, this basic rule is largely valid in the abstract, even if it is implemented differently in different countries, cultures, and markets.

- **Development of Standards in Consortia**

Consortia standards are developed by companies who agree to work together to solve a specific market need. In Europe, for example, interest groups representing individual sectors, such as mechanical engineering or the electrical industry, were established at an early stage. In this context, representation of interests is to be understood universally and includes loose exchange of information, social activities, trade fairs and the organized discussion of specific problems and topics in working groups, as well as the representation of interests in politics. For example, German industry representatives are organized in the German association of the machine and plant automation industry (VDMA) [VD23], the German association of the Electrical and Digital industry (ZVEI) [ZV23], the German Association for Information Technology, Telecommunications and New Media (BITKOM) [BI23] or the Association of the Internet Industry (eco) [EC23]. Topics that unite increased political and industrial interest and require close cooperation between industry, politics and research are often addressed in

⁴ Some other bodies developing standards or guides exist in ISO [IS23].

Germany but also in Europe via joint exchange platforms such as the Plattform Industrie 4.0 in Germany [Pla23] or Austria [Ass23].

Analogously, this also applies at the international level to institutions from specific countries that also represent the interests of their members internationally, as well as to international associations, societies, or other forms of organization, such as e.g.⁵ GAIA-X [GA23], Catena-X [Ca23], OPC Foundation [OP23] in the context of Industry 4.0. The institutions mentioned above develop and publish documents and specifications with detailed description of use cases and requirements but also technical challenges to be solved and their solutions. Therefore, these organizations are, in the sense of the above description, also to be understood as standardization organizations.

2.3 Standardization in Industrie 4.0

Standardization plays a crucial role for building up an international Industrie 4.0 ecosystem [KA16]. Whereas in the early years of Industrie 4.0 after its proclamation in 2011, the focus was on identifying, structuring and concretizing the topic area and the related challenges by means of application scenarios and defining the reference architecture model RAMI4.0, the need for (national and international) coordinated solutions increased continuously. The challenge of developing a solid and well-defined framework of standards for Industrie 4.0 is talked in Germany by the development of a Standardization Roadmap which is frequently updated with respect to the current state of standardization and knowledge regarding future needs since 2015 and currently available in Version 5 [Ge23]. For selected technologies and applications, specialized roadmaps have been additionally developed, as e.g. the Standardization Roadmap Circular Economy [St23]. Furthermore, surveys developed on international level are available, e.g. [La23]. A comprehensive discussion of relevant standards and norms for Industry 4.0 is beyond the scope of this article and, therefore, the interested reader is referred to these publications.

Standards and norms are developed based on the currently valid state of technology and science. Accordingly, there is a strong correlation between innovation, technological developments and scientific works. Due to this reason, the interplay and knowledge exchange plays a crucial role: considering regulatory aspects and existing standards in scientific works will facilitate the dissemination and foster applicability of the results in industry [Le22]. Various scientific works deal with surveying recent status of standards, as e.g. [LMF16, Lu19]. The integration of strong interplay will ease and speed up standardization [EC21]. This is expected to result also in novel opportunities to improve

⁵ The list of institutions mentioned here is not to be understood as complete and is only intended as an example.

education, foster understanding of systemic mechanisms and, by that, finally mutually leveraging education, science, and standardization.

3 Selected Characteristics for Categorizing Standardization, Publication and Participation

The necessity to provide a deeper analysis on the structural and systemic mechanisms of standardization as well as the (general) differences between institutions was given in the previous sections. Accordingly, selected properties and characteristics to gather insights behind standardization, publication and participation within activities is provided as a proposal towards a novel framework for classification of the standardization landscape for facilitating informed decision making in this context.

3.1 Process Model of Standards Development and Consensus

The main purpose of consensus processes is to establish the level of agreement on controversial issues [Fi84]. Different types of consensus strategies are a long-standing area of research both in terms of consensus building between humans [FS12, FS88] as well as in computer science in the case of distributed systems, such as blockchain [BMB20]. Any organization developing and publishing any kind of document, specification, standard or norm uses its own process which may contain well-defined intermediate versions as a kind of maturity level of the publications, considers (possibly limited) public feedback, rules to consider feedback in a structured way, etc. Some institutions have a well-documented, publicly visible and therefore transparent procedure or process model, which in this case is often accompanied by clearly defined rules of conduct for drafting (e.g. compliance) and, among other things, for reaching consensus or dealing with disagreements. The accessibility and transparency of these processes, guidelines and rules differ from institution to institution. For example, ISO and IEC developed exhaustive common procedures for technical work which are defined and publicly available [II23]; additionally, both institutions give simplified insights on their webpages [ISO23, IEC23]. In some cases, information is often only partially accessible to the public, not precisely defined and, by that, open to interpretations, and/or available only for a restricted user group (e.g. members). The reason why the process models for developing specifications and standards differ is not only because each institution wants to be individual. Rather, these processes reflect the self-image and objectives of the respective forms of publication and types of cooperation within the respective organization. Finally, the framework conditions for the definition of processes are predefined by other organizational circumstances, such as the type of membership and the resulting levels of participation (cf. Sect. 3.3). For some institutions, no public information is available about any aspect of the internal structure for working towards and developing of standards, specifications, or any other kind of results. Only when complete data is available can an informed, fact-based decision be made by outsiders as to whether a specific issue, aspect or challenge

should be worked on within an institution (collaboratively and in the spirit of open innovation). Likewise, without sufficient data, no solid assessment can be made as to whether and how publications and their content should be assessed, especially in terms of quality, soundness, and any supporter base.

The challenge, and by that, underlying structural patterns are similar as in science. In generalized but simplified terms: A scientific publication, which is reviewed and, by that, approved by reviewers is more trustful, and technically sound than a publication without any review. Depending on the way, how reviews are performed (peer, blind, etc.) and the acceptance rate of publications additionally scale the value. Accordingly, the challenge described thus exists both for scientific publications and for publications in the context of standardization.

Challenge: Untransparent processes for developing documents, especially standards, hinder comparability, understanding of underlying mechanisms, and finally comprehensive insights about the trustworthiness and necessary interpretation of the document.

Implication: There is a need to provide insightful implications resulting from the differences in the process model of standards development. Additionally, a uniform easy understandable approach to identify relevant action points for giving technical input, intervention, etc. would facilitate the handling of parallel, partially concurrent, and maybe contradictory activities in different institutions.

3.2 Classification of Publication Types

When reading a publication and rely own work on it, it is essential to get a general understanding how to process informational content of the publication correctly, i.e., to gain a deep understanding about the trustfulness of its content and how it should be used or interpreted by the reader. For example, if technical descriptions within a publication where never implemented and, accordingly, never evaluated in any way before, this should be kept in mind when following descriptions of a publication; in contrast, when reading a law, there is no option to follow the rules or not – but also laws leave room for interpretation. Accordingly, it is inevitable for a reader of a document to get deep and precise understanding of the conditions, interpretation rules and trustfulness of a publication.

Any organization or community can define its own publication types and procedures – typically reflecting the need of the document’s target group and the missions and objectives of the organization. Accordingly, the document type is relevant to be considered for determining the trustfulness, conditions, and interpretation rules.

With regard to the fundamental challenge, similarities with scientific works are again apparent: in scientific research, different values are implicitly assigned to the different types/classes of publication formats. For example, the publication of a research report is

valued differently compared to a scientific publication at conferences or even in journals⁶. Furthermore, in the daily practice of academic research, rules of thumb with recourse to such a categorization are not uncommon, like "for a dissertation, you need a certain amount of accepted scientific journal articles". The background to this rule of thumb, which varies depending on the subject area and branch of science, stems from the fact that a successful publication in scientific journals implies that peer review of the content by internationally recognized scientists from the respective scientific field and is a solid indicator of the quality of the scientific findings.

In general, two major types of publications can be distinguished regarding the impact and interpretation consequences for the reader: specifications within a **normative publication** can be seen as mandatory if one wants to be compliant with it whereas **informative publications** provide some explanatory or descriptive information, guidelines, or implementation processes and, by that, can be seen as recommendations.

Challenge: Overviewing all relevant institutions is time- and resources-intensive; investigating different publication types and their approval level is very complex and, in some cases, not possible without costs or active participation.

Implication: A uniform measure for indicating the level of approval and, by that, indicate the level of trust and technical soundness would facilitate the comparison of different publications among different publishers, institutions, and sources.

3.3 Classification of Participation Level

As discussed previously, a huge number of different institutions provides the opportunity to work collaboratively on (technical) challenges resulting in the publication of some kind of specification or standard. These organizations differ significantly in terms of the possibility of participation: while some offer membership exclusively to natural persons, others offer institutional membership only (e.g. companies, research institutions). The latter typically send representatives to work on specific topics or groups. A (non-exhaustive) collection of some relevant organization is given in Tab. 1.

In general, the variety of participation models is a great achievement. Nevertheless, it is influencing, among others, the process of voting and consensus building (cp. Sect. 3.1). Some organizations limit participation in voting regarding the sending institution, others do not. Organizations whose members are natural persons do not to a large extent do not have such restrictions.

While, for example, in the IEEE every natural person has the same right to vote, the DIN and DKE bylaws limit this to one vote per legal entity and, by that, follows the mechanisms

⁶ This is a highly simplified presentation and is in no way intended to diminish any kind of scientific work or publication. Rather, this should only be considered as a simplified example for illustration. Furthermore, various additional criteria are relevant to value scientific contributions, e.g. the importance and acceptance rate of a conference, the impact factor of journal, etc.

of ISO and IEC on national level. This means that legal entities that send employees to IEEE committees have proportionally more influence.

The relevance of the participation level is also illustrated at European level: in Europe there are three registered standards organizations - CEN, CENELEC, and ETSI. The national standards organizations of European countries are the members of CEN and CENELEC which systematically follow the basic decision-making procedure mentioned above for the IEC. In ETSI, institutions, in particular companies, participate directly. Accordingly, the influence of individual companies at ETSI is therefore be rated as higher than in the case of CEN and CENELEC. This is reflected, among other things, in the European Commission's standardization requests which, in the case of the AI Act for example, are addressed exclusively to CEN and CENELEC; ETSI is given solely an advisory role [EC22].

Institution	Regional scope	Membership type	Fee parameter
CEN [Eur23a] / CENELC [Eur23b]	Europe	National SSOs	N/A
DIN [DI23]	Germany	Institutions	No. of employees
DKE [DK23]	Germany	Persons	Free of charge
ETSI [ET23]	International	Institutions	Company annual turnover in specific areas
IDTA [ID23]	International	Institutions	Company annual turnover
IEC [IEC23]	International	National SSOs	N/A
IEEE [IE23]	International	Persons	Per person
ISO [ISO23]	International	National SSOs	N/A
OPC Foundation [OP23]	International	Institutions	Company annual turnover
VDE [VE23]	International	Institutions / Persons	No. of employees / per person
VDI [VI23]	Germany	Persons	Per Person

Tab. 1: Excerpt of selected institutions developing and/or publishing standards in context of Industrie 4.0 (status Q2/2023) in alphabetical order.

Challenge: The lack of transparency about influencing possibilities or the consequences resulting from organisational boundary conditions with regard to possible dominant influences by individual or several interest groups leads to misinterpretations with regard to the interpreted meaning and significance of possible publications of an institution. A possible lack of coordination between individual institutions can lead to contradictions between standards, which may not be immediately visible or recognisable. This results in a not inconsiderable risk of contradictory demands between the contractor and the client (in the case of a demand for compliance with possibly contradictory standards) and

ultimately in not inconsiderable additional work in the development/revision of standards and in the development.

Implication: A uniform cooperation model for knowledge and information exchange between institutions, simplified – in best case – automatic analysis of contradictions would ease the combination of standards from different institutions. A uniform model to reason about implications of an institutional organizations would ease the systemic understanding and role of an organization and guide contribution and finally, give transparency to activities of those institutions not all yet registered in transparency registers. Furthermore, a uniform process model for collaboration and knowledge sharing would guide members of institutions when to share information about a recent status of a document or development for avoiding contradictions and facilitate its application in a broader scope.

3.4 Accessibility of Standards / Business Models of Institutions

Organizations are faced with various activities around their services which take time and costs, like providing infrastructure for developing publications (e.g. committee meetings), infrastructure for publication, promotion, organizing and hosting meetings, etc. For financing these activities, business models are set up for financing these activities. Even if some institutions are operating non-profit, at least operating costs must somehow compensated. Independently of a respective business model, two major services of institutions can be identified: (i) offering to participate in (an organized) development of one or more publications and (ii) accessing a developed and published document. Both aspects may be charged for financing the operation of the institution in sense of (a part of) the business model or may be free of charge.

Whereas often a whole institution is considered, an initially conducted study indicated that business models of institutions may vary depending in the way how to participate, depending on the publication type originated to be developed, or different combinations of them including additional service or product offering. Accordingly, it is neither sufficient to differentiate on institutional level nor regarding the publication types. Therefore, offered products and services and their bundles must be considered for classification.

Due to the large number of institutions and their different types of publications, detailed research is very time- and cost-intensive. Some of this information is not publicly available. Not every organization that develops specifications is non-profit or operates non-profit. Some of them represent a kind of organized lobby organization which, in addition to the development of technical specifications, also try to assert their interests on the market and in politics, not least with an increasing number of supporters. Initial investigations indicate that lots of institutions that call themselves standardization organization or interest are not listed in transparency or lobby registers, even if they develop technical publications, are thus (at least implicitly) are a kind of rule-makers influencing industry and politics.

Challenge: Missing transparency of the business models of some institutions hinders comparability. Relevant information for neutral assessing and decision making about participating and contributing is limited. By that, participating to existing activities or initiating new activities for introducing own challenges is difficult. This results in significant barriers to get involved, especially individuals and SMEs for whom a corresponding investment of time and money must be targeted and may be critical.

4 Summary & Outlook

In the context of this paper, the need for a more detailed consideration of the mechanisms in the development of technical specifications was identified based on different approaches, which are highly heterogeneous between different institutions. Different levels of participation and process models in the development and adoption of publications have a direct impact on the level of consensus achieved. Different models of participation and business models also affect the possibilities of exerting influence and the availability of standards. In summary, it could be shown that there is a great need to make existing structures more transparent, but also to make the mechanisms of standardization more understandable. A distinction between norms and standards for the terminological separation of different mechanisms of standardization and significance, as is the case in Germany (at least formally), makes sense, but cannot be achieved internationally and is not practised consistently in Germany, too. In this respect, a new classification scheme is needed to create transparency for users, in particular SMEs, about the carrying capacity and significance of documents, but also suitable participation opportunities for (small and medium-sized) companies. The work presented here lays a first foundation to ultimately act as a navigator in the jungle of standardisation - also in the context of the increasing trend towards collaborative development and open innovation.

However, the question of well-defined and suitable process models to enable the transfer of knowledge and content between different institutions - research and standardisation as well as different standardisation organisations - in order to take regulatory necessities into account remains open for future works.

References

- [AB19] Abdelkafi, N.; Bolla, R.: Understanding ICT Standardization: principles and practice, TREDITION GMBH, 2019.
- [Ass23] Association Industrie 4.0, Plattform Industrie 4.0, plattformindustrie40.at/?lang=en, visited: 25.08.2023.
- [BI23] BITKOM e.V., www.bitkom.org, visited: 25.08.2023.

- [BMB20] Bamakan, S.M.H.; Motavali, A.; Babaei Bondarti, A.: A survey of blockchain consensus algorithms performance evaluation criteria. In: *Expert Systems with Applications*, Vol. 154, pp. 113385, 2020.
- [BPR17] Blind, K.; Petersen, S.S.; Riillo, C.: The impact of standards and regulation on innovation in uncertain markets. *Research Policy*, Vol. 46, pp. 249–264, 2017.
- [BRS12] Brunsson, N.; Rasche, A.; Seidl, D.: The Dynamics of Standardization: Three Perspectives on Standards in Organization Studies. In: *Organization Studies*, Vol. 33, pp. 613-632, 2012.
- [Ca23] Catena-X, catena-x.net/en/, visited: 25.08.2023.
- [DI21] DIN e.V., German Institute for Standardization: DIN 820-3: Standardisation Work – Part 3: Terminology, February 2021.
- [DK23] DKE - German Commission for Electrotechnical, Electronic & Information Technologies, VDE Association for Electrical, Electronic & Information Technologies, <https://www.dke.de/en>, visited: 25.08.2023.
- [DI23] DIN e.V., German Institute for Standardization, www.din.de, visited: 25.08.2023.
- [EC21] EC JRC, European Commission. Joint Research Centre: Linking research and innovation with standardisation: operational activities in the collaboration between JRC and CEN CENELEC. Publications Office, 2021.
- [EC22] EC, European Commission, Draft Commission Implementing Decision on a standardisation request to the European Committee for Standardisation (CEN) and the European Committee for Electrotechnical Standardisation (CENELEC) in support of safe and trustworthy artificial intelligence, 2022.
- [EC23] eco, Association of the Internet Industry, international.eco.de, visited: 25.08.2023.
- [ET23] ETSI, European Telecommunications Standards Institute, www.etsi.org, visited: 25.08.2023.
- [Eur23a] European Committee for Standardization (CEN), www.cencenelec.eu, visited: 25.08.2023.
- [Eur23b] European Committee for Electrotechnical Standardization (CENELEC) www.cencenelec.eu, visited: 25.08.2023.
- [Fi84] Fink, A.; Kosecoff, J.; Chassin, M.; Brook, R.H.: Consensus methods: Characteristics and guidelines for use. In: *American Journal of Public Health*, Vol. 74, pp. 979–983, 1984.
- [FS12] Farrell, J.; Simcoe, T.: Choosing the rules for consensus standardization. In: *The RAND Journal of Economics*, Vol. 43, pp. 235–252, 2012.
- [FS88] Farrell, J.; Saloner, G.: Coordination Through Committees and Markets. In: *The RAND Journal of Economics*, Vol. 19, 1988.
- [GA23] GAIA-X, GAIA-X: A Federated Secure Data Infrastructure, gaia-x.eu, visited: 25.08.2023.

- [Ge23] German Standardization Roadmap Industrie 4.0 - Version 5. DIN e.V. & DKE German Commission for Electrical, Electronic and Information Technologies, January 2023.
- [ID23] IDTA, International Digital Twin Association e.V., industrialdigitaltwin.org/en, visited: 25.08.2023.
- [IE23] IEEE SA, IEEE Standards Association, standards.ieee.org, visited: 25.08.2023.
- [IEC23] IEC, International Electronic Commission, iec.ch, visited: 25.08.2023.
- [II23] ISO/IEC, ISO/IEC Directives, Part 1 + IEC: Supplement: Procedures for the technical work — Procedures specific to IEC, May 2023.
- [IS23] ISO, International Organization for Standardization: Other bodies developing standards or guides in ISO, www.iso.org/other-bodiesdeveloping-standards.html, visited: 25.08.2023.
- [ISO23] ISO, International Organization for Standardization, www.iso.org, visited: 25.08.2023.
- [IT23] ITU, International Telecommunication Union: Members Directory - ITU Hub, <https://www.itu.int/hub/membership/ourmembers/directory/?myitu-members-states=true&request=countries>, visited: 25.08.2023.
- [Ka16] Kagermann, H.; Anderl, R.; Gausemeier, J.; Schuh, G.; Wahlster, W.: *Industrie 4.0 in a Global Context: Strategies for Cooperating with International Partners*. Utzverlag, 2016.
- [La23] Landscape Analysis Reports | StandICT.eu 2026, www.standict.eu/landscape-analysisreports, visited: 25.08.2023.
- [Le22] Leitao, P.; Cristalli, C.; Paone, N.; Chiariotti, P.; Utz, W.; Stojanovic, N.; Barata, J.; Woitsch, R.: *Integrating Standardization in Research and Innovation Projects: the GOODMAN Experience*. In: 2022 IEEE 31st International Symposium on Industrial Electronics (ISIE). pp. 1147–1152. IEEE, Anchorage, AK, USA, 2022.
- [LMF16] Lu, Y.; Morris, K.; Frechette, S.: *Current Standards Landscape for Smart Manufacturing Systems*. National Institute of Standards and Technology, 2016.
- [Lu19] Lu, Y.; Huang, H.; Liu, C.; Xu, X.: *Standards for Smart Manufacturing: A review*. In: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE). pp. 73–78. IEEE, Vancouver, BC, Canada, 2019.
- [OP23] OPC Foundation, opcfoundation.org, visited: 25.08.2023.
- [Pla23] Plattform Industrie 4.0, www.plattform-i40.de/IP/Navigation/EN/Home/home.html, visited: 25.08.2023.
- [St23] Standardization Roadmap Circular Economy. DIN e.V. & DKE German Commission for Electrical, Electronic and Information Technologies & VDI - The Association of German Engineers, January 2023.
- [VE23] VDE Verband der Elektrotechnik Elektronik Informationstechnik e.V., Association for Electrical, Electronic & Information Technologies, www.vde.com/en, visited: 25.08.2023.
- [VD23] VDMA e.V., German Association of the Machine and Plant Automation Industry Registered Association, www.vdma.org, visited: 25.08.2023.

- [VI23] VDI Verein Deutscher Ingenieure e.V., Association of German Engineers, www.vdi.de/en/home, visited: 25.08.2023.
- [Wi22] Wiegmann, P.M.; Eggers, F.; De Vries, H.J.; Blind, K.: Competing Standard-Setting Organizations: A Choice Experiment. *Research Policy*, Vol. 51, No. 104427, 2022.
- [ZV21] ZVEI, Electro and Digital Industry Association: Policy paper on standardisation: Key statements and recommendations, 2021.
- [ZV23] ZVEI e.V., Electro and Digital Industry Association, www.zvei.org/en/, visited: 25.08.2023.
- [Fe19] Ferreira, A. et al.: Data generation for intrusion detection systems in industrial control systems, 2019.
- [Gi21] Giraldo-Forero, A. F. et al.: Adversarial generative networks for data augmentation in industrial control system intrusion detection, 2021..
- [Hj19] Herrera, J.: *Industrial Internet of Things: Cybersecurity and ICS*, 2019.
- [IA22] IAEA NSS 17-T:2022: IAEA Nuclear Security Series No. 17-T; *Computer Security Techniques for Nuclear Facilities*, 2022.
- [IE20] International Electrotechnical Commission: IEC 63096: 2020-Nuclear power plants-Instrumentation, control and electrical power systems-Security controls, 2020.
- [IE20a] International Electrotechnical Commission: IEC 62541-1: OPC unified architecture—part 1: Overview and concepts.
- [IE20b] IEC TR 62451-2: OPC Unified Architecture - Part 2: Security Model, 2020.
- [Me21] Miuyin Y. W. et al.: *An Inside Look into the Practice of Malware Analysis.*, 2021.
- [OP12] OPC Foundation: *OPC UA Information Modeling*.
- [OP17] OPC Foundation: *OPC UA Security Analysis*.
- [SEK21] Schindler, J.; Kirdan, E.; Waedt, K.: *Secure OPC UA Server configuration for smart charging stations.*, 2021.
- [SSE16] Shin, I., Song, B., Eom, D.: Auto-Mapping and Configuration Method of IEC 61850 Information Model Based on OPC UA. In: *Energies*, MDPI, Open Access Journal, vol. 9(11), pp. 1-16, 2016.
- [St15] Stouffer, K.; Pillitteri, V.; Lightman, S.; Abrams, M.; Hahn, A.: *Guide to industrial control systems (ICS) security*. NIST Special Publication 800-82r2, 2015.
- [To20] Thomas, R.: *Cybersecurity for Industrial Control Systems: A New Breed of Cyber Attacks and Security Framework for ICS Protection.*, 2020.
- [Wa22] Waedt, K.; Ben Zid, I.; Schindler, J.; Kirdan, E.: *Cybersecurity Education Programmes & Laboratories Brainstorming.*, 2022.
- [We17] Wei, H.; Huang, Q.; Li, S.; Wang, P.; Zhang, S.: *The consideration of OPC UA security in constrained environments draft-wei-ace-opc-ua-security-00.*, 2017.

- [Zh20] Zhang, Y. et al.: Deep adversarial learning for industrial control systems intrusion detection., 2020.

DETECTION FOR INTER-ZONE IACS PROTOCOLS

Natasha Edeh¹, Robert Altschaffel² and Karl Waedt³

Abstract: ISO/IEC 27002:2022 distinguishes between the following types of security controls corrective: preventive, detective and corrective. The focus of this paper is on the support for testing of detective security controls for Industrial Automation and Control Systems. More specifically we will only address the generation of synthetic data that can be used for the detection of selected, advanced detective security controls. The proposed approach will be justified, while a comprehensive validation of the effectiveness of the synthetic data is beyond the scope of this paper. This work aims to contribute to the comprehension and improvement of security measures in Industrial Automation and Control Systems by focusing on the development of synthetic data and its consequences for the identification of specific detective security controls.

Keywords: Stego-Malware, Inter-Zone, Synthetic data, IACS, IACS protocols

1 Introduction

Industrial Automation and Control Systems (IACS), which ensure the effective and secure operation of numerous industrial processes, are the foundation of critical infrastructure sectors. Strong security controls are now more important than ever because of the IACS networks' expanding interconnectedness and the increasingly sophisticated cyber threats. In order to safeguard IACS environments against unauthorized access, malicious activity, and data breaches, detective security controls—controls made to recognize and respond to security incidents and anomalies—are essential.

To ensure the dependability and durability of detective security controls inside IACS, testing and evaluating their efficacy is crucial. However, operational limitations, potential infrastructure hazards, and the dearth of adequate test scenarios make it difficult to undertake thorough testing in real-world production systems. In order to overcome these difficulties, the creation of synthetic data has shown to be a useful method for mimicking actual attack scenarios and assessing the effectiveness of detective security systems in controlled test environments.

This research paper's main objective is to investigate and suggest approaches for creating synthetic data that is especially suited for the detection of particular, advanced detective security controls in IACS. We can help the thorough testing and assessment of detective

¹ Framatome GmbH, ICSP, Paul-Gossen Straße, 91052 Erlangen, Germany, natashaedeh@gmail.com, <https://orcid.org/0000-0001-9627-7001>

² Otto-von-Guericke-University Magdeburg, Germany, robert.altschaffel@iti.cs.uni-magdeburg.de

behaviour.

The detective security controls that are regarded as cutting-edge and essential for IACS security will be the focus of our study. These controls may use behavior-based detection, intrusion detection systems, network traffic analysis, anomaly detection algorithms, or any other pertinent methods. The performance and strength of these sophisticated investigative security mechanisms will be evaluated using a realistic test environment that we will create using synthetic data creation techniques.

The findings of this study report will aid in improving testing procedures for detective security controls in IACS. We can offer a controlled and realistic environment to test their performance, identify weaknesses, and improve their efficiency against up-and-coming cyber threats by producing synthetic data suited to the detection of specific advanced security mechanisms.

In the sections that follow, we will review pertinent literature, talk about stego-malware, talk about the difficulties in producing synthetic data for testing detective security controls, suggest methods for producing synthetic data, and show tentative findings to illustrate the efficiency of our approach in assessing the chosen advanced detective security controls.

2 Need for Synthetic Dummy Data

It is necessary to have access to representative attack data that appropriately reflects the changing threat landscape in order to detect and assess the performance of detective security mechanisms within Industrial Automation and Control Systems (IACS). However, because there are so few identified and published representative attacks, acquiring such data is extremely difficult. The delicate nature of IACS environments, organizations' reluctance to reveal security issues, and the ongoing advancement of attack methods are the main causes of this scarcity. To get over these restrictions and support thorough testing of detective security mechanisms, the creation of synthetic dummy data becomes crucial [Bu20].

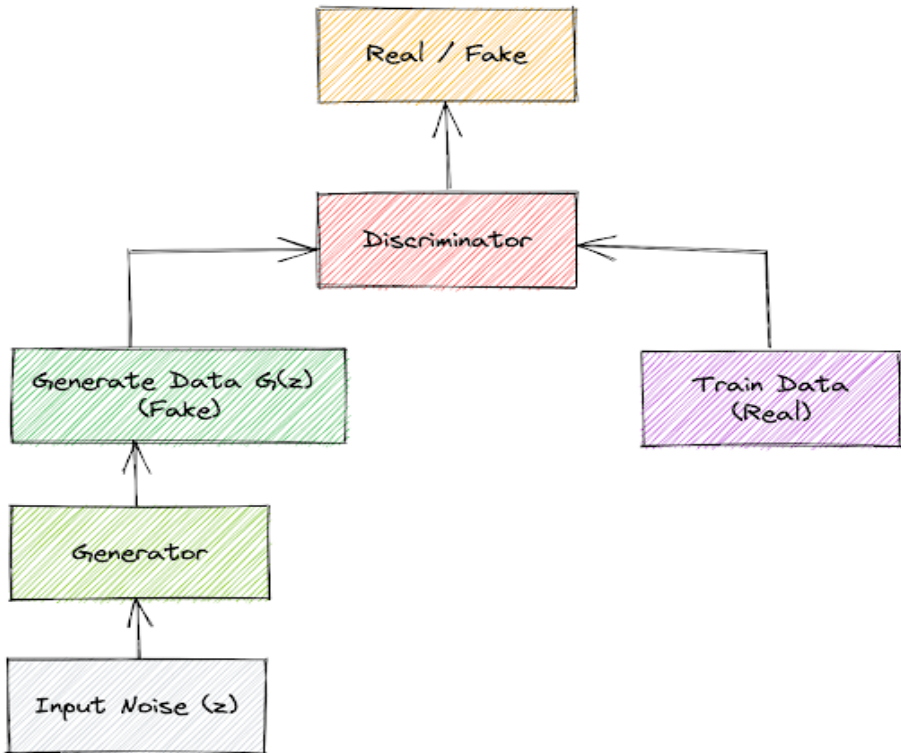


Figure 2: Synthetic Data Generation Process

2.1 Scarcity of Known Representative Attacks

The evaluation of detective security controls for IACS is hampered by the dearth of known representative attacks. Due to the possible harm to their reputation, potential legal repercussions, and potential disruption of essential operations, the majority of IACS operators are hesitant to reveal security events. As a result, there aren't many available datasets with actual attack data relevant to IACS setups, and when there are, they sometimes lack the depth and breadth needed for thorough testing. Studies like the one by [De17] emphasized the necessity for alternate strategies like synthetic data generation due to the dearth of publicly available representative attack data for IACS intrusion detection systems. Due to this dearth, detective security procedures that rely on actual attack patterns and behaviour cannot be developed or evaluated.

The threat landscape is dynamic, which makes it more difficult to find representative attack data. It is difficult to gather and share up-to-date attack data because attackers constantly modify their tactics, methods, and procedures (TTPs) to avoid detection. The reliance simply on known assaults becomes inadequate for assessing the efficacy of detective security procedures as new attack vectors arise and existing ones' change. The limitations of relying only on known assaults and the importance of tackling the dynamic nature of IACS threats were both emphasized by [Zh20]. To supplement current datasets, they suggested using techniques for creating synthetic data. This would enable the simulation of novel and developing attack patterns and allow for a thorough assessment of detective security controls.

2.3 Comprehensive Testing of Detective Security Controls

The detective security controls in IACS must be thoroughly tested in order to guarantee their dependability and efficacy. The evaluation of detective security systems against a variety of attack scenarios, including those that are not well-documented or publicly available, is made possible by the controlled and repeatable environment provided by synthetic dummy data. [Gi21] emphasis on the value of thorough testing with synthetic data generation methods for assessing the performance of anomaly detection algorithms for IACS. Researchers and practitioners are able to recreate real-world attack situations, test the effectiveness of detective security mechanisms, and spot potential vulnerabilities that might not have been found otherwise by creating synthetic dummy data.

3 Steganography

The art and science of steganography involves concealing information in seemingly innocent carrier data, such as pictures, audio files, or written texts, without raising any red flags. Steganography seeks to hide the existence of the information itself, as contrast to encryption, which concentrates on safeguarding the privacy of information by converting it into an unintelligible form. Steganography permits clandestine communication between adversaries by embedding the secret data within the carrier and getting over conventional security measures.

Steganographic approaches use the properties of the carrier data, such as the silent areas in an audio file or the least important bits of an image, to insert the hidden information. Authorized recipients who have the skills and resources to decrypt the steganographic technique can then

extract the embedded data.

3.1 Malware

Any software program or code that is intended to interfere with, harm, or gain unauthorized access to computer systems, networks, or data is referred to as malware, short for malicious software. Malware is a broad category of malicious software that includes spyware, Trojan horses, worms, viruses, and ransomware. Malware can have a variety of goals, from stealing confidential data to seizing control of computers and using it for nefarious ends [Me21].

Malware can be delivered into IACS settings in a number of ways, including as through corrupted external devices, malicious files, or network-based attacks. Once installed, malware has the potential to threaten the reliability, security, and confidentiality of IACS components and data by spreading throughout the network, lying inactive, or engaging in harmful actions.

3.2 Stego-Malware

A hybrid of steganography and malware, stego-malware uses steganographic techniques to hide dangerous code or payloads within carrier data. Stego-malware's main goal is to avoid typical security measures by concealing dangerous components inside of seemingly innocent files, making detection and investigation more difficult. Bypassing intrusion detection systems, antivirus programs, and other security measures that primarily concentrate on recognizing known malware signatures or distrustful network traffic is made possible by the use of steganography in the context of malware. Stego-malware can evade network defenses by hiding itself inside carrier files that appear to be legal and then carry out its malicious operations once inside the target system.

3.3 Challenges in Stego-Malware Detection

Due to its clandestine nature and the attackers' dynamic strategies, stego-malware detection is extremely difficult. Stego-malware is frequently missed by traditional signature-based detection techniques since the embedded payloads are specifically intended to avoid conventional checks. Specialized detection techniques that can locate and extract concealed payloads from carrier data are needed for the study of stego-malware. These algorithms may use statistical analysis, machine learning, or heuristic-based methods to identify anomalies and departures from typical data behaviour. Further complicating the procedure is the identification of stego-malware in inter-zone Industrial Automation and Control Systems (IACS) protocols. IACS protocols, which are intended to facilitate communication between various zones or subnetworks inside an IACS, exhibit

particular traits and might need specialized detection techniques to properly detect stego-malware.

4 Why Inter-zone?

Detective security controls for Industrial Automation and Control Systems (IACS) must handle potential risks resulting from communication across various systems or zones within an IACS environment in addition to internal dangers present in a single system. In the context of stego-malware detection, this section examines the necessity of focusing on inter-zone network protocols and the explanations why they necessitate special attention.

4.1 Communication Between Systems in IACS

To ensure the smooth operation of industrial processes, a number of systems and components come together in an IACS environment. Based on the functionality, criticality, or security requirements of these systems, various zones can be defined. Through network protocols created expressly for communication between various zones within the IACS, communications and data exchange among these systems take place.

The inter-zone communication network protocols are essential to the total security of the IACS system. Attackers could enter or disrupt the system using any weaknesses or security holes in these protocols, which could have serious repercussions including safety mishaps, operational disruptions, or unauthorized access to vital infrastructure.

4.2 Unique Characteristics of Inter-Zone Protocols

In contrast to protocols used within a single system, inter-zone network protocols in IACS have special qualities. These traits result from the demand to maintain safe and dependable communication between various zones while upholding the operational specifications of the industrial operations. Inter-zone procedures have several important components, including:

Heterogeneity and Complexity: Inter-zone protocols frequently make use of a variety of communication technologies, from antiquated systems to contemporary industrial networking standards. Compatibility, security settings, and interoperability may be further complicated by this complexity and heterogeneity.

Zone isolation: Inter-zone protocols are created to make it easier for systems to communicate with each other while ensuring separation and access control to stop unauthorized system interactions. Potential security issues are segmented and contained

to some extent by this isolation, which also lessens the impact they have on the entire IACS.

Traffic Patterns and Data Structures: Inter-zone protocols' traffic patterns and data structures may be very different from those of a single IACS system. These variants can take the shape of various message formats, temporal traits, and data representations, necessitating specialized detection techniques to spot abnormalities and possible stego-malware activities.

4.3 Significance for Stego-Malware Detection

The following factors make concentrating on inter-zone network protocols for stego-malware identification essential:

Cross-Zone Protection: Stego-malware detection can successfully safeguard the entire IACS environment by studying the connection between various zones. Stego-malware can be stopped from spreading and possible lateral movement within the IACS by being detected at the inter-zone level.

Targeting Communication Flaw: Attackers frequently use inter-zone network protocol flaws or incorrect setups to access or steal data from an IACS. Security measures can particularly target the communication flaws that attackers may exploit by addressing stego-malware detection at this level.

Detection Challenges and Solutions: The distinctive features of inter-zone protocols present particular difficulties for the detection of stego-malware. Researchers and practitioners might create custom detection algorithms by concentrating on certain protocols, taking into account the uniqueness of message structures, inter-zone communication patterns, and protocol-specific vulnerabilities.

5 Generating Synthetic Dummy Data using OPC UA or MQTT

Two widely used communication protocols in the area of industrial automation and control are OPC UA and MQTT. Within an IACS context, these protocols offer effective and secure communication between equipment, systems, and programs. Synthetic dummy data can be created to imitate a variety of realistic circumstances for stego-malware detection by utilizing the flexibility and extensibility of OPC UA or MQTT. To create synthetic dummy data using OPC UA or MQTT, the first step is to create data models and message formats that precisely reflect the features of IACS communication. It entails locating the pertinent data points, variables, and qualities present in an IACS and mapping them to the protocol's chosen data model.

5.1 Incorporating Normal and Malicious Data

Data depicting stego-malware activity as well as typical (legal) data should both be included in the creation of synthetic dummy data. The harmful data should contain stego-malware payloads concealed inside the normal data, whilst the normal data can be generated by replicating operating scenarios and common behaviour found in an IACS context. The altered synthetic data can be used to implement stego-malware operations by hiding malicious payloads inside the legitimate data utilizing steganographic methods. In order to imitate actual stego-malware instances, the hidden payloads can be deftly positioned inside the data streams, offering a realistic testing ground for the assessment of detective security systems.

5.2 Scalability and Realism

Scalability and realism are key factors for creating synthetic dummy data. The method should enable the production of substantial amounts of data that accurately reflect the complexity and diversity of IACS communication. To enable thorough testing of detective security controls, the synthetic data should be indicative of diverse scenarios, system configurations, and operational settings. Additionally, the synthetic dummy data ought accurately to depict the subtleties of timing, sequencing, and data dependencies included in IACS protocols. As a result, the generated data is guaranteed to be realistic, and the test situations for stego-malware identification are made more faithful.

5.3 Validation and Evaluation

It is critical to validate and assess the synthetic dummy data's performance in testing detective security mechanisms for stego-malware detection once it has been developed. This entails evaluating the efficacy of the selected security controls against the synthetic data in terms of their performance in properly identifying stego-malware activities and minimizing false positives. Controlled trials, simulations, or a comparison of the findings with information on known stego-malware attack patterns can all be used for validation and evaluation. The iterative process of creating fictitious data, evaluating the effectiveness of detection, and optimizing the strategy ensures that detective security controls in IACS are improved continuously.

6 Conclusion

The necessity of evaluating detective security controls in Industrial Automation and Control Systems, as specified by ISO/IEC 27002:2022, has been addressed in this study. We have investigated a useful method for boosting the general security posture of these systems by concentrating primarily on the generation of synthetic data for the detection of advanced detective security mechanisms. Although the effectiveness of the synthetic data


is not thoroughly validated in this report, it does establish the groundwork for further investigation and experimentation. It is advised that more research be done in the future to assess the viability and practical application of the synthetic data approach. To validate the detection capacities of the chosen detective security controls using the supplied synthetic data, this could entail conducting case studies or controlled experiments.

References

- [Bu20] Buczak, A. L. et al.: Synthetic data generation for anomaly detection in industrial control systems, 2021.
- [DC19] Dinculeană, D.; Cheng, X.: Vulnerabilities and limitations of MQTT protocol used between IoT devices, 2019.
- [De17] Delgado-Mohatar, C. et al.: Generation of synthetic data sets for network intrusion detection in industrial control systems, 2017.
- [Fe19] Ferreira, A. et al.: Data generation for intrusion detection systems in industrial control systems, 2019.
- [Gi21] Giraldo-Forero, A. F. et al.: Adversarial generative networks for data augmentation in industrial control system intrusion detection, 2021..
- [Hj19] Herrera, J.: Industrial Internet of Things: Cybersecurity and ICS, 2019.
- [IA22] IAEA NSS 17-T:2022: IAEA Nuclear Security Series No. 17-T; Computer Security Techniques for Nuclear Facilities, 2022.
- [IE20] International Electrotechnical Commission: IEC 63096: 2020-Nuclear power plants-Instrumentation, control and electrical power systems-Security controls, 2020.
- [IE20a] International Electrotechnical Commission: IEC 62541-1: OPC unified architecture—part 1: Overview and concepts.
- [IE20b] IEC TR 62451-2: OPC Unified Architecture - Part 2: Security Model, 2020.
- [Me21] Miuyin Y. W. et al.: An Inside Look into the Practice of Malware Analysis., 2021.
- [OP12] OPC Foundation: OPC UA Information Modeling.
- [OP17] OPC Foundation: OPC UA Security Analysis.
- [SEK21] Schindler, J.; Kirdan, E.; Waedt, K.: Secure OPC UA Server configuration for smart charging stations., 2021.
- [SSE16] Shin, I., Song, B., Eom, D.: Auto-Mapping and Configuration Method of IEC 61850 Information Model Based on OPC UA. In: Energies, MDPI, Open Access Journal, vol. 9(11), pp. 1-16, 2016.
- [St15] Stouffer, K.; Pillitteri, V.; Lightman, S.; Abrams, M.; Hahn, A.: Guide to industrial control systems (ICS) security. NIST Special Publication 800-82r2, 2015.

- [To20] Thomas, R.: Cybersecurity for Industrial Control Systems: A New Breed of Cyber Attacks and Security Framework for ICS Protection., 2020.
- [Wa22] Waedt, K.; Ben Zid, I.; Schindler, J.; Kirdan, E.: Cybersecurity Education Programmes & Laboratories Brainstorming., 2022.
- [We17] Wei, H.; Huang, Q.; Li, S.; Wang, P.; Zhang, S.: The consideration of OPC UA security in constrained environments draft-wei-ace-opc-ua-security-00., 2017.
- [Zh20] Zhang, Y. et al.: Deep adversarial learning for industrial control systems intrusion detection., 2020.

Security challenges and best practices for resilient IIoT Networks: Network Segmentation

Romarick Yatagha ¹, Karl Waedt¹, Josef Schindler¹ and Erkin Kirdan¹

Abstract: The surging prominence of the Industrial Internet of Things (IIoT) introduces both unique prospects and complex issues for industrial control systems, notably within the cybersecurity sphere. Cybersecurity concerns are particularly acute for smart factories, entities that leverage IIoT capabilities like networked sensors and machine learning to streamline production. The heterogeneous devices from diverse manufacturers and vast interconnected networks heighten their susceptibility to cyber threats. This paper examines the contemporary cybersecurity landscape within smart factories, pinpointing current vulnerabilities and imminent threats. Drawing on this analysis, we put forth a suite of best practices and strategic measures to fortify IIoT networks, including but not limited to network segmentation and stringent access controls. We pay specific attention to network segmentation, a technique used to break down a computer network into manageable subnetworks, thus mitigating the risk of attacks. We propose an innovative network segmentation policy that leverages clustering, an unsupervised learning algorithm. This algorithm classifies network traffic into distinct categories based on, but not limited to, source and destination IP addresses, employed protocol, and packet size. This data-driven classification simplifies network segmentation and configuration, minimizing their complexity. The paper also underlines the critical role of employee training and awareness in establishing robust security practices, particularly for the design, integration, and deployment of IIoT devices and edge computing. Our findings offer actionable insights for industrial control systems operators and cybersecurity professionals, empowering them to fortify their IIoT networks against cyber threats effectively.

Keywords: Network Segmentation, Machine Learning, Cybersecurity, IIoT, Edge Computing.

1 Introduction

As with any other revolution, industry 4.0 has advantages and drawbacks. Massive productivity, added comfort, and ergonomics are just a few of the benefits incurred when working in a smart and connected environment, and several industries are now integrating this technology, from smart factories to smart cities [Pa21]. A smart factory is a manufacturing establishment that uses advanced technologies such as the Internet of Things (IoT), Machine Learning (ML), robotics, and cloud computing to automate the production process and increase efficiency. A major challenge currently faced by smart factories is securing and protecting the network environment against cyber threats. A study by the multinational corporation specialized in cybersecurity Fortinet [SaF23] on security experts in several countries worldwide and leaders of the Industry 4.0 indicates that almost

¹ Framatome GmbH, ICSP, Paul-Gossen Straße 100, Erlangen, 91052, {romarick.yatagha, karl.waedt, josef.schindler, erkin.kirdan}@framatome.com

all (94%) of the surveyed organizations had to deal with at least one cyberattack in the year 2022. It is an alarming situation, as the estimated cost of these attacks is up to \$8 trillion worldwide [Ja23], with the most common form of attacks being malware intrusion and phishing, often leading to a ransomware attack [SaF23]. It is therefore important for industries and researchers to always be informed on the latest attack vectors, trends, and developments on cyber-attacks and countermeasures to effectively secure their organizations as new hacking strategies are devised every day [Wa22].

Several reviews on challenges and best practices in computer networks have been published in recent years to raise awareness, but very few are specifically directed towards smart factories, and the countermeasures proposed do not take into account the particularities of smart factories. In [Re23], for example, Rehman et al. analyse how smart factories secure communication between their equipment and address privacy using encryption mechanisms and authentication. They focus only on global security and privacy issues such as authentication and data encryption and do not propose any countermeasure against the outlined security threads in fintech companies. In [SS21], Soofun et al. propose a review of security challenges alongside their countermeasures. They proposed a layered security assessment, considering issues at the application, network, and transport layer, with a comprehensive list of open security issues in IIoT. In [DKP22], the authors examine how the cryptographic algorithms of TLS, an established Internet security standard, affect resource consumption. They measure the consumption of CPU cycles, memory, and network bandwidth with the IANA-recommended TLS 1.2 and TLS 1.3 cipher suites. In [Co21], the authors analyse threats in the manufacturing system through the CIA Triad, pinpointing the sensitive digital assets that need to be protected in the manufacturing system. Some vulnerabilities mentioned in the article, like the use of unencrypted protocols, are patched already, and the main recommendation for improving security in industries is developing cybersecurity awareness by training employees to limit and minimize the impact of attacks. In [Ki18], the authors propose a high-performance and scalable network traffic shaper and evaluate it in a DDoS mitigation scenario. Alireza et al. in [Za19] highlight cybersecurity issues that can affect maintenance in smart manufacturing systems. As maintenance aims to have fewer or no defects on the manufacturing devices, making the production environment safe and continually productive, they emphasize how cybersecurity threats can impact this smooth functioning. By so doing, the paper focuses most on cyber threats that affect the availability, performance, and quality of devices in the IIoT networks, leaving aside other threats and attack vectors like phishing and social engineering. In [Za18], the authors use game theory to analyse and predict what an attacker will do during an attack on industrial systems and propose a responsive strategy to defend the system, should the analysed type of attack eventually happen. Frederico et al. in [Ma21] consider threats on a smart factory test bed comprising seven workstations with PLCs, HMI, and sensors, highlighting four attack scenarios on different entry points on the overall system. They illustrate lateral attacks in smart manufacturing networks as the main drawback, given that a single security flaw on one device can be enough for an attacker to access other machines in the factory successfully. The authors fail to analyse and propose efficient countermeasures like

network segmentation, which is studied in this paper.

We propose to analyse some recent attacks on smart factories critically. We simulate these attacks in a sandbox environment to better understand the attack scenario and propose efficient countermeasures. Based on the attack vector used, contextual international standards, and our observations, we propose some countermeasures that can be adopted in a smart factory environment to minimize the probability of attack success. Accordingly, we propose a Machine Learning architecture for network segmentation.

2 Security challenges in Smart Factory

Securing smart factories is an ever-growing challenge due to many interconnected, heterogeneous devices. Numerous devices configured with different protocols and naming conventions, such as IP/mac addresses or Radio Frequency Identification (RFID), must send, receive and analyse data packets in different formats. The high growth rate at which smart factories are being attacked daily sufficiently demonstrates the many challenges smart factories have to deal with. Therefore, listing all the recent attacks and challenges smart factories face is difficult. For this reason, we identify recent attacks belonging to some threat catalogues specified in standard [II10] and replay these attacks to the best of our knowledge and equipment to identify the breaches and propose corresponding countermeasures.

2.1 Malware intrusion

Malware intrusion is an attack scenario in which hackers take over control of a production system via malware like a Trojan or a virus. The malware is often downloaded and installed along with other programs and software packages called wrappers, which now appear as benign programs undetected by anti-malware definition tools. This is becoming a common attack vector, as freely cracked software is publicized online. Unadvised employees want to save money by using this cracked software, not paying for licenses of the original copyrighted software. But in reality, they install malware that enables attackers to control their devices. This is the case of a modified version of the software “WinX DVD ripper Platinum” used to copy information from copy-protected and encrypted Digital Versatile Disks (DVD) free of charge, as opposed to the official software subject to purchased licenses.

We install this software in our analysis sandbox and run it against the intelligence platform and anti-malware tools VirusTotal [To23]. This platform compares the program’s signature with databases of several anti-malware vendors and industries. The result is presented in tab 1, where some security vendors tag the program as containing malware and Trojan, and some others did not detect any malware for the following reasons:

- The malware can modify its script and hence have a new unknown signature

- There is a wrapper software that effectively copies encrypted files from Digital Versatile Disks
- The malware definition engine of these security vendors is not current.

The latter two reasons are less probable, as the software was installed on a Windows 10 with an updated Defender in our analysis environment without raising any alarm.

The intelligence platform also captures and analyses outgoing network traffic to determine the contacted IP addresses

Security vendor	Malware	Trojan	Undetected	Unable to analyse
McAfee GW-Edition	x	x		
Avast			x	
Kaspersky		x		
Acronis			x	
Avira			x	
AVG			x	
Avast mobile				x
Microsoft			x	
Node 32			x	
Google			x	
Bit defender				x
Trust look				x

Tab. 1: Detection comparison between anti-malware vendors

Using these IP addresses with the whois lookup table [Lo23], we have an address of the contacted location. But this can be a “zombified” computer or IP address used in a distributed denial of service attack. Table 2 presents these results.

IP	Port	Country
192.229.211.108	80	California, US
20.99.184.37	443	Washington, US
20.99.185.48	443	Washington, US
23.216.147.76	443	Massachusetts, US

Tab. 2: Outbound network analysis

The malicious program also modified, replaced, and added new files, as shown in table 3. This leaves a backdoor on the compromised system for the attacker to carry out several other attacks, such as ransomware or Distributed Denial of Service.

Files modified	Files deleted	Files Added
.../Desktop/setup.exe	.../Temp/WER216D.tmp. csv	.../Temp/WER216D.tmp.c sv
.../AppPatch/sysmain.sdb	.../Temp/WER21CD.tmp .txt	.../Temp/WER21CD.tmp.t xt
.../SYSTEM32/SSPICLI. DLL	.../Temp/WER21CE.tmp. txt	.../Temp/WER21CE.tmp.t xt
.../SYSTEM32/apphelp.dl l	.../Temp/WER217E.tmp. csv	.../Temp/WER217E.tmp.c sv
.../SYSTEM32/SSPICLI. DLL	.../store/cache/cache.dat	.../Temp/WER217E.tmp

Tab. 3: Files modified, deleted and added

Countermeasures

- Employees must be trained to be vigilant and cautious when opening email attachments, clicking on links, or downloading files from the internet. This will help a long way in avoiding such attacks. Employees should be encouraged to report suspicious activities promptly by educating them on the severity and cost of such attacks.
- Segmenting a network into several blocks, each with its security controls and access restrictions, will contain the spread of self-modifying viruses and limit their impacts on critical systems. This will also prevent lateral movement in networks.
- The use of smart IDS that monitors and analyzes the behavior of applications and processes rather than relying solely on static signatures. This approach can help detect and block unknown or self-modifying malware based on their actions. These IDS can also provide a regular audit of system files that were modified, added, or deleted.
- Permission to make changes to the computer's defense system, such as disabling windows defender or antivirus, should not be granted to regular employees unless the security professionals are aware and can monitor the activity.
- Factories and organizations should have customized notebooks for work, where employees cannot install applications on the notebooks.
- In a graded security approach, corporate notebooks with access to sensitive assets should be customized to only recognize and read data from specific flash drives, on which IT security specialists have control.
- Keep all software, including operating systems, applications, and antivirus definition programs, up to date. Software updates often include security patches that address vulnerabilities that malware can exploit.

- Available Intelligence platforms such as VirusTotal [To23] and McAfee should be used as browser extensions to automatically scan and analyze downloaded files on personal computers and compare them with anti-malware vendors' databases for any eventual intrusion on personal notebooks.

2.2 Phishing and social engineering

Phishing is the most used attack vector in smart factories, and the number of organizations who reported successful phishing attacks is increasing [SaF23]. This is particularly true as most companies and factories implement mechanisms to identify and combat phishing emails only on corporate computers, leaving employees' personal computers and emails vulnerable. The concept of separation between corporate and personal notebooks is not always respected. Payloads installed on personal computers via phishing might be transferred to the corporate notebook whenever the employee copies a document from the personal to the corporate notebook using a flash drive, for example. This is how the Melissa virus, the fastest-spreading infection, was spread through Microsoft Word Macros in 1990. It infected many users by attaching itself to a Word document and spreading along with this file whenever it was copied or emailed.

Phishing attacks are becoming more and more successful as hackers now use spear phishing, sending a tailored message to a specifically targeted user with some private information obtained during the reconnaissance and information-gathering phase of the attack. For example, a vast campaign of phishing emails supposedly sent by UNICEF was sent to users who registered or applied for scholarships on the UNICEF online learning portal after a data leak of personal data in September 2019 [Ch19]. Publicly available and purchased via some underground sources, the leaked data and others, e.g., from the recent T-Mobile [Co23], from NVIDIA and Airbus [Aa23], are used to send specific information to targeted users to obtain sensitive information or install payloads on computers on a single click. Another example is spear phishing, in which North Korean hackers target 150 well identified South Korean experts in a phishing campaign [Sh23], aiming at stealing sensitive information such as industry know-how and security information. Information gathering on social media enables hackers to obtain the personal information of employees who seem to be environmentally friendly. The hackers then impersonate an academic or journalist seeking information or opinions by sending targeted emails to these employees. If any target replies to the first benign email, they follow up and ask them to anonymously participate in a survey that could enable the government to sign and adopt some environmentally friendly policy, such as closing nuclear power plant sites. By downloading and filling out the form, the target computer is infected with a payload attached to the form that helps carry out future attacks.

Countermeasures

- Security and Monitoring should be extended to employees' devices: Cybersecurity teams should ensure that employees' notebooks have strong protection against phishing. This can be done with the Netcraft anti-phishing toolbar extension for browsers [Gr10] or the VirusTotal intelligence platform [To23].
- The cybersecurity teams should properly identify employees' notebooks, and any communication between the two devices should be properly monitored.
- Phishing campaign simulation with scenarios explained above must be regularly reviewed with employees to raise their awareness, as recommended in standard [IA21].
- Scenarios, as those explained above, should be generated with the help of machine learning and used to train employees. Innovative tools must be used to prepare and raise awareness. This is the only way to stand against present days attacks that already leverage these tools.

2.3 Insecure implementation of security measures

In our previous research [SWK21], we used two different configuration manuals to illustrate the necessity of using cyber-secured hardware configuration in a smart factory environment. On the default/insecure configuration proposed by one configuration manual, we used a Metasploit module for OPC UA to show the possibility of login to controllers with empty credentials and revealed the nodes reserved for privileged users only and unencrypted communication.

This highlights the importance of training engineers to configure such hardware, along with several other security policies prescribed in standards [IE16a] and [IE16b]

Ransomware is another type of computer attack that deals with the confidentiality and availability features of the CIA Triad, disrupting the functions of an organization by encrypting sensitive information. With the advent of Ransomware as a Service (RaaS), ransomware is the attack most organizations encounter and are concerned about. This is also fueled by the availability of AI-powered tools which assist in creating and using ransomware. The same study [FO23] shows that of the 94% of organizations attacked, half of the surveyed organizations were victims of ransomware attacks at least once, of which 44% of the organizations paid the total amount demanded by the hackers. This is the case of the Swiss technology leader in electrification and automation ABB, recently the victim of a ransomware attack in May 2023.

Insights [He23] say the ABB attack was conducted using malicious web link techniques, such as Search Engine Optimization (SEO) poisoning and fake browser updates, to install the Qakbot Trojan malware, which now “deployed a type of ransomware that is not self-propagating, and ex-filtrated specific data. We understand that the malware does

not spread through emails or attachments and does not automatically self-propagate to other systems across a network”, said the cybersecurity experts [Te23]. This further demonstrates the need to have resilient defense mechanisms to combat evolving and self-modifying malware, as this is not the initial description of the Qakbot Trojan malware [In21].

SEO poisoning is an attack vector consisting of malicious users making links to their malicious websites appear as the first result of a search engine-like advertisement. Inattentive users click on these links, which direct them to malicious websites or download malware. As illustrated in the previous sections, SEO is an attack vector like malware intrusion or phishing, giving the hacker access to the target system. The Qakbot Trojan was initially developed to steal banking and financial sensitive information but has evolved to give a malicious hacker access and control of the complete target system. It is not ransomware but serves as a dropper that, once installed, gives the hacker a backdoor to launch a ransomware attack. Extensively analyzed by Microsoft Security Intelligence [In21], Qakbot attempts to identify a virtual environment by checking for installed software and identifying running processes to compare it with a predefined blacklist. The malware stops all execution upon identification of a sandbox environment.

Countermeasures

- Ensure continuous technological watch on the latest developments in the cyber threats and proposed solutions
- Implement innovative access control based on the principle of least privileged
- Always ensure that anti-malware solutions are up-to-date
- Employee training and awareness is the game-changing factor in the fight against cybercrime, especially in Industry 4.0.

3 Securing IIoT networks with network segmentation

A strict application of contextual guidelines provided by standardization bodies is a good step toward securing IIoT networks. Most of them include network segmentation. Network segmentation is a policy used to optimize network security and enhance performance while isolating critical components of the network, to contain an eventual attack and prevent lateral movement in the network. It is implemented by isolating sections of the network from each other with devices like routers and is a continuous process as networks, especially in smart manufacturing environments, never cease to evolve and grow. It is a highly recommended protection policy by international standards, but its implementation is vaguely described. For instance, network segmentation is recommended and described in different standards:

- “To cover the different information security requirements for the countries an international organization is doing business in, segmentation of a network in effect in line with country borders can be an effective broad solution. [II09]”
- In [II10] as “a tool used to augment system control by segregating critical applications from other applications or databases from most users.”

Neal Wagner et al. in [Wa16] approached the question of network segmentation as a decision problem, questioning how an appropriate segmentation architecture for a given network can be selected. They proposed a decision support system based on statistical analysis to select a segmentation architecture.

In traditional networks, segregation is either based on the following:

- IP addresses, where devices with addresses in a specific range are regrouped together
- on the physical location of the devices, where devices on the same floor or building are regrouped in the same network, for example
- the intensity of traffic flow between the devices in the network
- the protocol and port number used.

In an IIoT environment, we have all the features to implement any of the above segmentation architectures for network segmentation, but it is not a trivial task to identify which segmentation architecture is best. Moreover, after a segmentation architecture is implemented, the manual configuration and maintenance of the network is tedious, especially when several devices are being added to the network, like in smart factories.

In this research, we use Machine Learning to propose a network segmentation policy that considers the source and destination IP addresses, the used ports and protocols, and the intensity of traffic flow between network devices. The proposed model can also maintain the network segments when new devices are added efficiently.

3.1 Data collection

We use a data set generated by Al-Hawawreh et al. [ME21] for evaluating and training machine learning-based intrusion detection systems for IoT and IIoT systems, as it suits the heterogeneity and interoperability of IIoT systems, including several IIoT devices and protocols such as MQTT, CoAP, Modbus, and WebSocket. Generated by the test bed illustrated in Fig 1, which is based on the Industrial Internet Reference Architecture (IIRA) model, the dataset follows industrial standards [ME20].

This test bed consists of a variety of Machine-to-Machine, Machine-to-Human, and Human-to-Machine connectivity protocols, IIoT devices that are deployed on the three tiers of an IIoT system, including its

- edge tier (i.e., physical devices and edge computing),
- platform tier (i.e., cloud storage and analytics) and
- enterprise tier (i.e., service and application devices)

The data is captured with the appropriate tools [ME21] and is carefully correlated with well-defined algorithms to ensure data consistency and limited data redundancy. This dataset consists of 421,417 observations/instances for normal and 399,417 for attacks with 67 features. For network segmentation, some features like *attack_label* are less relevant and are dropped from the dataset, as illustrated in the next section.

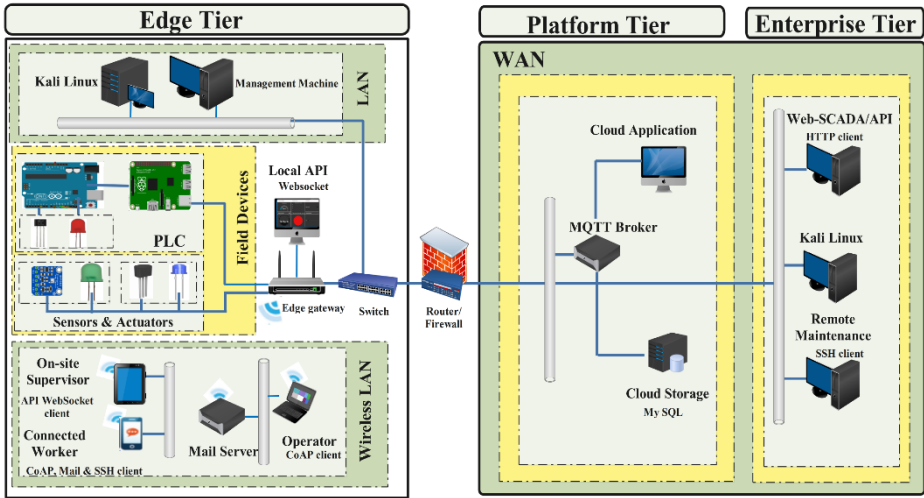


Fig. 1: Brown-IIoT-bed Architecture for data capture [ME21]

3.2 Data preprocessing

Data preprocessing involves filtering out irrelevant information, scaling the data, and converting it into a suitable format for machine learning algorithms. Tab. 4 details the features used in our model.

Feature Name	Type	Description
Timestamp	Discrete	Timestamp to synchronize connections and packets
Scr_IP	Discrete	Source endpoint's IP address
Des_IP	Discrete	Destination endpoint's IP address
Scr_Port	Discrete	Source endpoint's TCP/UDP port,
Des_Port	Discrete	Destination endpoint's TCP/UDP
Protocol	Discrete	Link layer protocol (eg. TCP, UDP)
Scr_bytes	continuous	Number of bytes from source to destination
Des_bytes	continuous	Number of bytes from destination to source
missed_bytes	continuous	Number of bytes missing
Service	Discrete	Application layer protocol (CoAP, MQTT, modbus)
Duration	continuous	The time difference between the last packet and the first packet seen
Scr_pkts	continuous	Number of packets sent
Des_pkts	continuous	Number of packets received
total_bytes	continuous	total number of bytes exchanged between source and destination
byte_rate	continuous	Total number of bytes per second
total_pkts	continuous	Total number of packets exchanged between source and destination
pkts_rate	continuous	Total number of packets per second
Conn_state	Discrete	Connection status (1 complete, 2 Rest, 3 partial)
SYN	Discrete	If connection has packet with SYN flag (1 or 0)
SYN_ACK	Discrete	If connection has packet with SYN-ACK flag (1 or 0)
is_with_payload	Discrete	If connection has packet with payload (1 or 0)
Scr_packt_ratio	continuous	Ratio of packets sent to total number of packets
Des_packt_ratio	continuous	ratio of packets received to total number of packets
Scr_bytes_ratio	continuous	ratio of source bytes to total number of bytes
Des_bytes_ratio	continuous	ratio of bytes received to total number of bytes
Login_attempt	Discrete	1 if attempt to login, 0 otherwise

Tab. 4: Features used

The Timestamp feature is a good indicator for network segmentation, as it can tell if there is a pattern in the data exchange between devices at some specific moments to regroup them. For its missing values, we convert the date value in the initial dataset into a timestamp to get the corresponding packets captured.

The IP addresses and ports are also relevant for network segmentation, as they will indicate the specific devices that exchange information regularly and the frequency at which the ports are used. Protocol and Service features indicate the link layer protocol (TCP, UDP...) and the application layer protocol (CoAP, MQTT...), respectively, used in the data packets. The source, destination, and *missed_bytes* will bring information on the intensity of traffic

flow between devices of the network. Synchronization, acknowledgment, and *with_payload* will bring information about the type of packet being considered. Some highly correlated features are dropped, as they will not provide extra information. Examples are *successful_login* and *is_activity_privileged* which are highly correlated (97%) with the *login_attempt* feature. Looking at the very high number of unique values for some categorical features in figure 2, one hot encoding will not yield good results for the problem at hand but will lead us to the curse of dimensionality. It will also make us lose the relation between the Source and Destination of the IP addresses or ports. For encoding these features, we use the pseudocode in Algorithm 1, which regroups all unique feature elements and encodes them according to their frequency in the dataset. The more frequent an element, the higher its encoding.

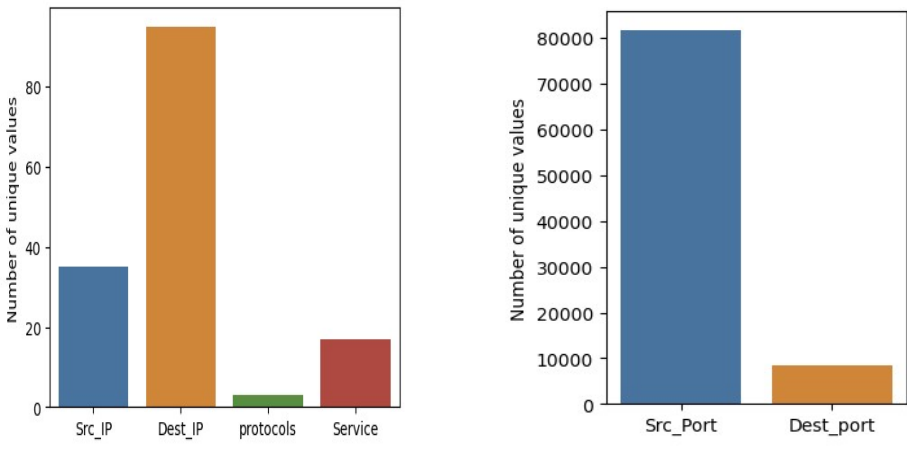


Fig. 2: Number of Unique values of selected features

```

Input Data: dataframe
Initialization: D={};
for each cat_feat in dataframe do {
    N = count(uniq_elts);
    count(uniq_elts_freq);
    D = order(uniq_elts) by freq;
    map(D) in dataframe by (1-N);
}

```

Alg. 1: Encoding algorithm

3.3 Model implementation, results and discussion

K-Means clustering is a machine learning model used to regroup the data into a specified

number of segments or clusters. It iterates through the dataset, regrouping all objects with similar characteristics. To identify the ideal number of segments or clusters, we use the Elbow method [Na18]. This technique evaluates the sum of squared error for several clusters. Considering that the higher the number of clusters, the lower the sum of squared error, it assumes the best number of clusters is obtained when the sum of square error does not decrease beyond a given threshold value. In practice, we plot the values of the Within Sum of Square (WSS) on a line graph, and the threshold value is considered to be where we have an elbow. The Elbow technique is summarized by the pseudocode in Algorithm 2. The Within Sum of Square is calculated as follows:

$$WWS = \sum_{i=1}^m (x_i - c_i)^2$$

where the x_i represents the data points and the c_i the center point of the considered cluster.

```

Input Data: k_range, threshold
Result: best_k
WSS(init) = 0;
for each k in cluster_range do {
    fit data with k;
    evaluate WSS(k);
    if (WSS(k) - WSS(k-1) > threshold) {
        k = k+1;
    }
    else {
        best_k = k;
        stop
    }
}
Best_k = last_k;
}

```

Alg. 2: Elbow Algorithm

Implementing this technique, Fig. 3 depicts the sum of squared errors with the corresponding number of clusters. From this diagram, 7 is chosen as the appropriate number of segments the traffic data will be divided into. This is because 7 is the value at which we have an “elbow” on the line graph.

Two different metrics are used to evaluate the performance of our model. These are the Silhouette Coefficient and the Calinski-Harabasz index.

The Silhouette coefficient measures cohesion among the data points in a cluster. It calculates the quality of clusters formed by clustering algorithms [BH21]. Two parameters evaluate this coefficient,

- the mean distance between a sample and all other samples in the same cluster (a) and
- The mean distance between a sample and all other samples in the nearest cluster

(b)

Its value ranges from -1 to 1, where higher values indicate better-defined clusters and is evaluated by the formulae

$$S = \frac{b-a}{\max(a,b)}$$

The Calinski-Harabasz index calculates the ratio of “between-cluster” dispersion to “withincluster” dispersion for a given clustering solution. Also known as the CH index or the variance ratio criterion, the higher its value, the better the clustering result is considered to be. It aims at finding clusters that are well-separated and compact internally. The formula for the Calinski-Harabasz index is as follows:

$$CH = \left[\frac{\sum_{k=1}^K n_k \|c_k - c\|^2}{K-1} \right] / \left[\frac{\sum_{k=1}^K \sum_{i=1}^{n_k} \|d_i - c_k\|^2}{N-K} \right]$$

Where K is the number of clusters, d_i the data points, n_k and c_k the number of points and centroid of the k^{th} cluster respectively and N the total number of data points. The values of the evaluation metrics are consigned in Tab. 5.

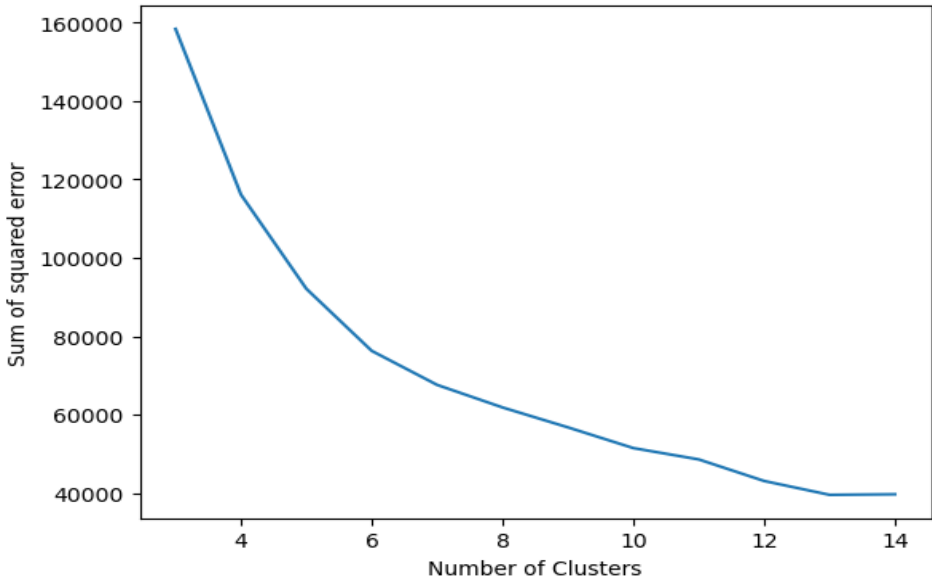


Fig. 3: Identifying appropriate number of clusters with elbow method

	SC	CH
Results	0.52	417887.1

Tab. 5: Evaluation metrics

With the assumption that the IP addresses uniquely identify the devices in the considered dataset, we regroup the addresses belonging to each segment by their frequencies to analyze if the traffic clusters correlate with device clusters. Network segmentation is the regrouping of devices and not traffic data into segments. This result is extracted in Tab. 6. The complete result and code are hosted on GitHub [Ya23]. The IP addresses are represented in the table by their encoded values. This will show us how often the considered device transmits data in the network.

Cluster 0		Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6	
IP	Freq	IP	Freq	IP	Freq	IP	Freq	IP	Freq	IP	Freq	IP	Freq
1.00	2990	0.98	9074	1.00	3805	1.00	6834	0.98	4232	1.00	1007	0.96	1754
	3		1		8		6		5		9		3
0.97	5052	0.97	1394	0.93	9457	0.90	3201	0.95	1752	0.93	459	1.00	6920
			0						8				
0.93	2746	0.96	1016	0.82	242	0.93	2399	0.94	1727	0.96	345	0.97	1211
									3				
0.98	1425	1.00	132	0.83	238	0.97	2286	0.91	3648	0.98	323	0.98	417
0.96	68	0.93	37	0.81	189	0.98	2048	0.88	1429	0.89	205	0.89	389
0.90	3	0.50	11	0.96	186	0.89	1932	0.87	1211	0.97	118	0.85	20
		0.90	7	0.89	147	0.96	1684	0.86	618	0.90	2	0.93	2
				0.90	139			1.00	566			0.62	2
				0.79	107			0.85	240			0.81	2
				0.97	105			0.93	191			0.79	2
				0.98	54			0.80	130			0.83	2
				0.72	32			0.78	88			0.34	1
				0.65	28			0.77	82			0.82	1

Tab. 6: Segmentation results

- A group of IP addresses is strongly represented in several clusters: Address 1.00 in clusters 0, 2, 3, 5, and 6; address 0.93 in clusters 0, 2, 3, and 5. These are probably the addresses of routers and gateways of the network, as they are the most frequent addresses in the dataset.
- A second group of addresses is represented in one cluster and less represented in others: 0.98 in cluster 1; 0.89 and 0.90 in cluster 3.
- Others are strongly Others in just one of the clusters, e.g., 27 IP addresses are present only in cluster 4, and 32 IP addresses are present only in cluster 2.

4 Ongoing challenges and future works

The ideal scenario for machine learning-based segmentation of the IIoT network is to

capture and analyse data from an actual IIoT system operating in real time and not on a simulated test bed. Having such data traffic is challenging as Industry 4.0 is a relatively new technology, and Industries are unwilling to give their data for scientific research. Data collected from real Industries will include features like the physical location of each device, which can significantly impact the segmentation process, as the mobility of devices is fundamental in the IIoT environment.

One of the research results presented in this paper is the successful labelling of segmentation data. Supervised learning algorithms can be applied to this data to predict which segment will belong to new devices to be connected to the network. this will help optimize the maintenance of the segmented network as it grows.

Using passwords, badges, and PINs to access resources in Industry 4.0 is not a good practice.

Badges can be duplicated, and passwords and PINs can be guessed or brute-forced with AI-powered tools such as ChatGPT, leaving the target systems vulnerable. Fig. 4 and 5 show how easily they have used innovative techniques not available in the past decades to crack passwords and PINs. This calls for an advanced, next-generation access control scheme like ABAC, which provides dynamic, fine-granular, and context-aware access control to resource [Aea22].

TIME IT TAKES A HACKER TO BRUTE FORCE YOUR PASSWORD IN 2022

Number of Characters	Numbers Only	Lowercase Letters	Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters, Symbols
4	Instantly	Instantly	Instantly	Instantly	Instantly
5	Instantly	Instantly	Instantly	Instantly	Instantly
6	Instantly	Instantly	Instantly	Instantly	Instantly
7	Instantly	Instantly	2 secs	7 secs	31 secs
8	Instantly	Instantly	2 mins	7 mins	39 mins
9	Instantly	10 secs	1 hour	7 hours	2 days
10	Instantly	4 mins	3 days	3 weeks	5 months
11	Instantly	2 hours	5 months	3 years	34 years
12	2 secs	2 days	24 years	200 years	3k years
13	19 secs	2 months	1k years	12k years	202k years
14	3 mins	4 years	64k years	750k years	16m years
15	32 mins	100 years	3m years	46m years	16b years
16	5 hours	3k years	172m years	38m years	92bn years
17	2 days	69k years	96b years	179bn years	7m years
18	3 weeks	7m years	467bn years	11m years	438bn years

Learn our methodology at hivesystems.io/password

(a) Hacking password in 2022

TIME IT TAKES A HACKER TO BRUTE FORCE YOUR PASSWORD IN 2023

Number of Characters	Numbers Only	Lowercase Letters	Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters, Symbols
4	Instantly	Instantly	Instantly	Instantly	Instantly
5	Instantly	Instantly	Instantly	Instantly	Instantly
6	Instantly	Instantly	Instantly	Instantly	Instantly
7	Instantly	Instantly	1 sec	2 secs	4 secs
8	Instantly	Instantly	28 secs	2 mins	5 mins
9	Instantly	3 secs	24 mins	2 hours	6 hours
10	Instantly	1 min	21 mins	5 days	2 weeks
11	Instantly	32 mins	1 month	10 months	3 years
12	1 sec	14 hours	6 years	59 years	226 years
13	5 secs	2 weeks	132 years	3k years	15k years
14	52 secs	1 year	17k years	202k years	1m years
15	9 mins	27 years	898k years	12m years	77m years
16	1 hour	713 years	46m years	276m years	60m years
17	14 hours	18k years	35m years	48bn years	380bn years
18	6 days	481k years	128bn years	2m years	26bn years

Learn how we made this table at hivesystems.io/password

(b) Hacking passwords in 2023

Fig. 4: Hacking password without AI

USING CHATGPT HARDWARE TO BRUTE FORCE YOUR PASSWORD IN 2023

Number of Characters	Numbers Only	Lowercase Letters	Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters, Symbols
4	Instantly	Instantly	Instantly	Instantly	Instantly
5	Instantly	Instantly	Instantly	Instantly	Instantly
6	Instantly	Instantly	Instantly	Instantly	Instantly
7	Instantly	Instantly	Instantly	Instantly	Instantly
8	Instantly	Instantly	Instantly	Instantly	1 secs
9	Instantly	Instantly	4 secs	21 secs	1 mins
10	Instantly	Instantly	4 mins	22 mins	1 hours
11	Instantly	6 secs	3 hours	22 hours	4 days
12	Instantly	2 mins	7 days	2 months	8 months
13	Instantly	1 hours	12 months	10 years	47 years
14	Instantly	1 days	52 years	608 years	3k years
15	2 secs	4 weeks	2k years	37k years	232k years
16	15 secs	2 years	140k years	2m years	16m years
17	3 mins	56 years	7m years	144m years	1bn years
18	26 mins	1k years	378m years	8bn years	79bn years



> Learn how we made this table at hivesystems.io/password

Fig. 5: Hacking password in 2023 with ChatGPT

A smart IDS system capable of combating self-modifying viruses must be proposed, as malware evolves by encrypting itself with different keys, hence having a signature different from what was recorded in threat intelligence systems using the YARA rules, for example. These are future research directions that need to be considered.

Bibliography

- [Aa23] Aaron, D.: Data Breaches That Have Happened in 2022 and 2023 So Far. <https://tech.co/news/data-breaches-updated-list>, 2023.
- [Aea22] Altschaffel, R. et al.: Sino-German white paper on IT Security Tests for Industrie 4.0 and Intelligent Manufacturing. Technical report, Deutsche Gesellschaft für Internationale Zusammenarbeit, 2022.
- [BH21] Belyadi, H.; Haghighat, A.: Chapter 4 - Unsupervised machine learning: clustering algorithms. In (Belyadi, Hoss; Haghighat, Alireza, eds): Machine Learning Guide for Oil and Gas Using Python, pp. 125–168. Gulf Professional Publishing, 2021.
- [Ch19] Chadwick, V.: UNICEF data leak reveals personal info of 8,000 online learners. <https://www.devex.com/news/unicef-data-leak-reveals-personal-info-of-8000-online-learners-95558>, 2019.

- [Co21] Corallo, A.; Lazoi, M.; Lezzi, M.; Pontrandolfo, P.: Cybersecurity Challenges for Manufacturing Systems 4.0: Assessment of the Business Impact Level. *IEEE Transactions on Engineering Management*, pp. 1–21, 2021.
- [Co23] Cox, J.: T-Mobile Confirms It Was Hacked. <https://www.vice.com/en/article/y3d4dw/t-mobile-confirms-it-was-hacked>, 2023.
- [IE16a] IEC/TR 62541-1, CEI/TR 62541-1. OPC Unified Architecture, Part1: Overview and Concepts. IEC, 2016.
- [IE16b] IEC/TR 62541-2, CEI/TR 62541-2. OPC Unified Architecture, Part2: Security Model. IEC, 2016.
- [II09] ISO/IEC 27033-1 Information technology — Security techniques — Network security. Part 1: Overview and Concepts. ISO, 2009.
- [II10] ISO/IEC 27033-2 Information technology — Security techniques — Network security. Part 3: Reference networking scenarios — Threats, design techniques and control issues. ISO and IEC, 2010.
- [In21] Intelligence, Microsoft Security: , A closer look at Qakbot’s latest building blocks (and how to knock them down). <https://www.microsoft.com/en-us/security/blog/2021/12/09/a-closer-look-at-qakbots-latest-building-blocks-and-how-to-knockthem-down/>, 2021
- [Ja23] James, N.: 90 Cybercrimes statistics in 2023: Cost, Industries and Trends. <http://www.getastra.com/blog/security-audit/cyber-crime-statistics/>, 2023.
- [JJT19] Jill, West; Jean, Andrews; Tamara, Dean: *Network+ Guide to Networks*. CENGAGE, 2019.
- [Ki18] Kirdan, Erkin; Raumer, Daniel; Emmerich, Paul; Carle, Georg: Building a traffic policer for ddos mitigation on top of commodity hardware. In: 2018 International Symposium on Networks, Computers and Communications (ISNCC). IEEE, pp. 1–5, 2018.
- [Lo23] Lookup, W.: Whois Lookup. <https://research.domaintools.com/>, 2023.
- [Ma21] Maggi, F.; Balduzzi, M.; Vosseler, R.; Rösler, M.; Quadrini, W.; Tavola, G.; Pogliani, M.; Quarta, D.; Zanero, S.: Smart Factory Security: A Case Study on a Modular Smart Manufacturing System. *Procedia Computer Science*, 180:666–675, 2021. Proceedings of the 2nd International Conference on Industry 4.0 and Smart Manufacturing (ISM 2020).
- [ME20] Muna, A.; Elena, S.: Developing a Security Testbed for Industrial Internet of Things. *IEEE Internet of Things Journal*, 8:5558 – 5573, 4 2020.
- [ME21] Muna, A.; Elena, S.: X-IIoTID: A connectivity-and device-agnostic intrusion dataset for industrial internet of things. *IEEE Internet of Things Journal*, 9:3962–3977, 8 2021.

- [Na18] Nainggola, R.; Perangin, A.; Simarmata, E.; Astuti, F.: Improve the performance of K-Means cluster using the sum of square error, optimized by the elbow method. *Journal of Physics: Conference Series*, 1361:12015, nov 2018.
- [Pa21] Paiho, S.; Kiljander, J.; Sarala, R.; Siikavirta, H.; Kilkki, O.; Bajpai, A.; Duchon, M.; Pahl, M.; Wüstrich, L.; Lübben, C. et al.: Towards cross-commodity energy-sharing communities—A review of the market, regulatory, and technical situation. *Renewable and Sustainable Energy Reviews*, 151:111568, 2021.
- [Re23] Rehman, F.; Attaullah, H.; Ahmed, F.; Ali, S.: Data Defense: Examining Fintech’s Security and Privacy Strategies. 04 2023.
- [SaF23] SUNNYVALE, C.; Maddison, J.: , State of Operational Technology and Cybersecurity Report. <https://www.fortinet.com/corporate/about-us/newsroom/press-releases/2023/fortinet-global-report-finds-75-percent-of-organizations-experienced-intrusion-last-year>, 2023.
- [Sh23] Shreyas, R.: North Korean hackers target some 150 experts in phishing campaign: ROK police. <https://www.nknews.org/2023/06/north-korean-hackers-target-some-150-experts-in-phishing-campaign-rok-police/>, 2023.
- [SS21] Soofun, T.; Samsudin, A.: Recent Technologies, Security Countermeasure and Ongoing Challenges of Industrial Internet of Things (IIoT): A Survey. *Sensors (Basel, Switzerland)*, 21, 10 2021.
- [SST23] Semenchenko, A.; Salvio, J.; Tay, R.: , YouTube Pirated Software Videos Deliver Triple Threat: Vidar Stealer, Laplas Clipper, XMRig Miner. <http://www.fortinet.com/blog/threat-research/youtube-pirated-software-videos-deliver-triple-threat-vidar-stealer-laplas-clipper-xmrig-miner>, 2023.
- [SWK21] Schindler, J.; Waedt, K.; Kirdan, E.: Secure OPC UA Server configuration for smart charging stations. *Gesellschaft für Informatik e.V. GI*, pp. 1753–1757, 2021.
- [Te23] Team, ABB Security: , IT Security Incident Customer and Supplier FAQ. https://library.e.abb.com/public/f1fbab4a291041ae84769e52fa42b0bd/IT%20security%20incident_customer%20FAQ.pdf, 2023.
- [To23] Total, V.: Virus Total Intelligence platform. <https://www.virustotal.com/gui/home/upload>, 2023.
- [Wa16] Wagner, N.; Sahin, C.; Winterrose, M.; Riordan, J.; Peña, J.; Hanson, D.; Streilein, W.: Towards automated cyber decision support: A case study on network segmentation for security. pp. 1–10, 12 2016.
- [Wa22] Waedt, K.; Ben Z.; Schindler, J.; Kirdan, E.: Cybersecurity Education Programmes & Laboratories Brainstorming. In: *Cybersecurity for Critical Infrastructure Protection via Reflection of Industrial Control Systems*, pp. 100–107. IOS Press, 2022.
- [Ya23] Yatagha, R.: , Implementation of Network segmentation in IIoT. https://github.com/Mentor-Roma/Network_Segmentation_with_ML.git, 2023.

- [Za18] Zarreh, A.; Saygin, C.; Wan, H.; Lee, Y.; Bracho, A.; li N.: Cybersecurity Analysis of Smart Manufacturing System Using Game Theory Approach and Quantal Response Equilibrium. *Procedia Manufacturing*, 17:1001–1008, 2018. 28th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2018), June 11-14, 2018, Columbus, OH, USA Global Integration of Intelligent Manufacturing and Smart Industry for Good of Humanity.
- [Za19] Zarreh, A.; Wan, H.; Lee, Y.; Saygin, C.; Janahi, R. Al: Cybersecurity Concerns for Total Productive Maintenance in Smart Manufacturing Systems. *Procedia Manufacturing*, 38:532–539, 2019. 29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM 2019), June 24-28, 2019, Limerick, Ireland, Beyond Industry 4.0: Industrial Advances, Engineering Education and Intelligent Manufacturing.

Data Spaces as the Distributed Communication Means for Industrial Automation and Control Systems

Semantic data spaces governed by digital twins

Jan deMeer¹

Abstract: A data space is more than just a repository for data. Even data is more than an item of data. However, in research a data space is not a new philosophy of communication. In this paper the concept of a data space shall be developed for its application in industrial automation and control systems (IACS). For this purpose, the existing reference architecture models, e.g. for I4.0 manufacturing or, electricity transportation and distribution but also for altruistic data dissemination in realms of smart infrastructures like cities, buildings, agriculture etc. Almost all examples of infrastructures shall be extended with something that is called in this paper a ‘fourth dimension’ in addition to the three regular dimensions comprising life cycle value streams, communication protocols and system component hierarchies. The fourth dimension of various reference architecture models can be represented by a combination of the two axes of the life cycle value stream of the system assets i.e., data, products, energy etc. with the axis of layered interoperability dealing with the representation of semantics in the given reference model. Thus, semantics means the state changes performed over time by the considered assets which requires the semantic interoperability between locations of a site or between device of a production chain. A state change issued by actors or processes during the life cycle value stream is an event that is represented in a data model and shall be accessible to other actors via the data space. Thus, the communicating actors or processes interconnected by the data space do not need the traditional layered communication protocols of the architectural models since they are interconnected through a distributed data space which plays the role of a distributed data repository to all actors and recipients of an application.

Keywords: IEC 62443 IACS, Data Provenance, Reference Model, Digital Twin, Smart Data Space, Category of Objects, Interoperation Semantics, Knowledge Pyramid, Life Cycle, Asset Value Streams, AAS, Data-Mesh.

1 Introduction to the concept of a semantic data space

1.1 What is data?

In [BK23] the answer to the question ‘What is data?’ is given that ‘*Data is a product!*’ Since a product is an asset, data is an asset as well. Data however cannot be used for the data analysis because data as such is not clean. Thus [BK23] suggest to manage data as an asset like a product. A product as an asset obeys a blueprint that guides the assembly from

¹ Smartspacelab.eu GmbH c/o HTW Berlin, 12205 Berlin, Berner Str. 21B, demeer@smartspace.de

components and pieces that fit together into the anticipated product. For data as an asset the same procedures shall become applicable i.e. to have a plan by which data can be assembled into meaningful interpretable pieces of a domain knowledge. These pieces are called semantic artifacts that can be compiled into a larger picture of a dynamic domain configurations or system states.

A semantic piece of data comprises not only data-of-interest but also explaining metadata and policies as access rights (R) to data, identities (I) of interested parties and stakeholders and, last not least ‘narratives of provenance’ (P) which tell the history of data manipulations during data’s life cycle value streams. This explaining metadata 3-tuple is defined as $\langle R, I, P \rangle$ and will be annotated to a piece of data-of-interest that is operated by processes represented by variables $(\Delta v/\Delta t)$ according to [VDE23].

In [BK23] it is suggested that the so-called *DevOps* principle shall be applied to the digital and physical asset analysis by tightly combining aspects of the asset development with asset operations into a *short story of data provenance*.

1.2 What is a data space?

At the end of the last century the paradigm of distributed communication became mature enough to provide communication means at first called a ‘data lake’ among loosely coupled processes. When the data lake becomes more structured the lake was renamed to ‘data space’ since space-based programming technologies became apparent such as *Jini*TM *Technology* leveraging distributed communication means of higher-order operations like ‘leasing’, ‘distributed events’, ‘transactions’ etc. that are available throughout a common data space. For more information see [FHA99].

The new concept of a data space is semantics based and thus can be seen as an additional dimension added to the various reference architectures respectively models (RAM). The fourth dimension represents semantics which is added to the three-dimensional RAM without semantics.

A data space is so-to-say a closed shop that serves for specific purposes of e.g. I4.0 production, or energy transportation and distribution, smart cities, smart buildings, smart manufacturing etc. In theory a *class* of (data, phenomenological, ontological) objects usually have a purpose that is defined by a *category* based on a family of morphisms that define any purpose semantically making consistency of data objects much easier to maintain.

During life cycle the physical objects of a value stream e.g. production, passes different (physical) maturity stages. An example of a kind of asset twinning is a blueprint i.e. a compound set of data and meta-data that may specify the assembly of the twinned engine.

However, all operational stages of the machinery shall always be *iso-morphic* to the digital (blueprint) data stages. Assuming this is the case the asset of a blueprint can be trusted and

thus be provable against required phenomena and state changes of the physical asset. The prove of the isomorphy comprises a list of digital rights, actor identities and data provenance narrative i.e. RIP meta-data.

By having a purpose i.e., a set of ‘real’ physical or chemical phenomenological narratives implemented by rules and constraints that must permanently be checked against both value streams of data and of involved IoT devices and machinery. Thus, both data and machinery can be handled as an integrated digital and physical semantic asset which means the twinning of physical and digital representations of a thing.

The data maintained by the data space represents digital the physical phenomena of things of interest. Currently the architectural reference models are three-dimensional models that do not consider semantics as a dimension of a compound cyber-physical system. The purpose anticipated by tightly coupled life cycle value streams (LC VS) of data and physical (production) maturity stages of e.g. energy distribution, industrial production, altruistic data provisioning in smart cities etc. defines the semantics of the LC VS since the production process shall yield a product of good quality.

1.3 What is semantics?

Semantics has to do with expressing (and achieving) the purpose of an operational plan incorporating in parallel value streams of data assets and IoT things or machine assets in terms of generating, using, or manipulating these assets. Asset can be everything that has worth to the applying stakeholders i.e., assets are data, machineries, systems, IoT devices that are applicable by I4.0 production, construction, or even by volatile energy distribution processes.

The purpose-oriented semantics implemented by a data space can be applied to maintain data consistently and to generate trustworthiness to the stakeholders, actors or recipients connected to a data space they use in common. For purposes of synchronization between the production maturity state in terms of measured and analyzed data and the state of the (physical, IoT) machinery of production there must be a kind of cooperation between both states respectively value streams.

In the webinar [UN23] the question ‘What is semantics?’ has been raised. In 9 presentations at least 9 answers have been given. Already in the 1970ies technical committees of several SDO, UN and many others have developed semantic interoperability standards such as the UN/EDIFACT standard, which is based on *syntactic rules for message building*, data elements which are the atoms of EDIFACT messages, segments which are groups of data elements and, messages that represent sequences of segments.

To-day TC184/SC4 provides ISO/IEC standards on industrial data comprising standards on *system life cycle processing* (ISO15288), *interoperability barriers* (ISO11354), common data definitions (ISO19650), ontological specifications (ISO15926), Common

Data Environment backbone (IEC81346 CDE) or, Resource Description Framework Schema of the W3C (ISO15926 RDFS).

The smart city data model of ISO30182 that combines data for high-level decision-making or ISO5087 series [ISO22a] of Common Data Modelling (CDM) is described by a *horizontal three levels model* intended to overcome with the approaches of vertical data silos. On the lowest CDM level there is the set of foundational definitions like time location, activity, event, resource etc. Above there is the city level that represents actors, objects and infrastructures of the city, e.g. households, residents, organizations, transportation services, buildings, sensors, persons etc. and the top-most level CDM provides the service level that comprises the shared services of the city.

As these examples from [UN23] tell us, there's no unique style to address semantics. While EDIFACT still stresses 'syntactic rules for message building' the standards for the Smart City Model CDM of ISO5087-2:2014 [ISO22a] address already some basics of a data space like 'semantic web' of the W3C obeying rules for handling data. Furthermore, semantics and knowledge about city services is expressed in terms of formal terminologies of Description Logic (DL), Ontology Web Languages (OWL) or, Resource Description Framework Schemas (RDFS).

A CDE/CDM considered from data space point-of-view comprises data sets with overlapping semantics which must be transformed into data sets useful to applications and trusted by users.

Rights on handling and access data, Identities of actors allowed to operate on data contained in the data spaces, narratives of provenance compiled during the LC VS of data, build the so-called RIP metadata constraints (cp. with figure 2).

The LC VS of data twins the LC VS of physical phenomena of interest respectively a useful system configuration of a machinery needed for an *anticipated purpose* of I4.0 production, energy distribution or, smart city services etc. The anticipated purpose that is achieved stepwise during the twinned functional and data LC VS processing can be set equal to the operational semantics of the purpose of the machinery.

1.4 What is metadata?

Metadata is defined by the RIP model of figure 2. In the RIP model there are three classes of metadata defined, i.e. digital rights, digital identities and, provenance narratives. The digital rights control access to data, digital identities identify data publisher and subscriber to the data space and, provenance narratives are intended to provide provenance information on data such that anticipated data use becomes meaningful to subscribers. In legal contexts data use must be trustful thus, verifiable. For certain data there could be more than one provenance narration which means that there are alternatives in the development of data. A narrative is a trajectory of events reported for certain data and is represented in terms of graph theoretical artifacts.

2 Common Data Space Reference Architecture Model

2.1 Data space Knowledge

A data space is more than a collection of containers of various data. A data space is application oriented hence, the data space requires behavioural means that enable the application's stakeholders to communicate semantically.

Pure layered connections as it is currently documented in existing RAMs do not support semantic communication since the messages transmitted over the connections do not obey information being derivable from meta data annotated to the data sets of interest which are managed by the data space.

Hence a data space implements up to some extent the semantics of the data space application. The data space hosts data and meta-data where data models the behaviour of physical processing during the life cycle of an SM asset in case of I4.0 production. It produces a value stream with a certain purpose (e.g. production of an engine, energy transmission, city data usage etc.). The meta-data that annotates data sets of interest with RIP meta-data lists that can be derived from the data space.

Consequently, a data space represents the knowledge about the value stream semantics of the physical asset processing that is described by a set of physical differential variables communicating with each other. In order to exploit the knowledge captured by a data space, data, meta data and procedures e.g. to perform Collaborative Condition Monitoring CCM of an application based on one or more data spaces, the application semantics is partly implemented in terms of behavioural artifacts contained in the data space.

2.2 Data space Behaviour

The application of the URMSM - *Universal Reference Model for Smart Manufacturing* of the [IEC23a] to SMRM – *Smart Manufacturing Reference Models* is not an obvious case. The gap between these two RAMs² is identified as the provability of coherence and compatibility between standards developed.

From practice we know processing takes time and yields behavioural maturity states which shall be modelled in an anticipated three-dimensional RAM comprising both the set of data spaces comprising various kinds of knowledge of an application and the application itself.

However, a fourth dimension associated with the *Life Cycle Value Stream* (LC VS) of an asset should be suitable because all technical processes shall make sense and thus obey a

² In this document the various reference models for smart manufacturing such as URMSM, SMRM, SMIM, SM system etc. are referred to by RAM since the architecture with three axes are focused.

meaningful intension or a certain purpose which is semantics. The behavioural state of development or operation of an asset e.g. the maturity states of a product during its design phases or while production processing is going-on or, during its use in an application, is a set of sequences of events. Sequences of events are semantically modelled as trajectories through a *directed graph* as it has been invented by [VDE23]. Events are state changes of a value stream over time that are semantically represented by a set of *discrete time-differential variables* $\{v^{\text{dot}}|\Delta v|\Delta t\}$, for more information see [VDE23]. Technical processes are modelled by such variables e.g. $\{s:\text{const}, s^{\text{dot}}:\text{linear}, s^{2\text{dot}}:\text{exponential}\}$ obeying the behaviours of being constant, linear, exponential or even of higher order.

2.3 Data space State Changes

The state changes as part of the behaviour of an asset are formally modelled by graph types and graph trajectories, where the latter are runs through given graph types. Graph types are undirected graphs and graph trajectories are directed graphs. Directed graphs model sequences of events and undirected graphs model the capability of processes respectively of differential variables that may happen in a certain application environment.

Meta-data must be derived from any event in the data space resulting from domain-specific (e.g. energy bidding in the given example) activities. In the following, it is exemplified how to derive meta-data from value stream operations and how meta-data is compiled into a list of meta-data items represented as n-tuples and annotated to observed semantic domain operations marked by *OPERation* or *CONDition*:

<OPER operation! (list of RIP metadata)> and
<COND operation? (list of RIP metadata)>.

The above two syntactic templates of operations and conditions are applied to construct the **graph type** representing the options of communication among a publisher stakeholder ‘SGOperator’ and a set of subscriber stakeholders that are loosely coupled through the data space ‘energy bidding’. The graph type is the representation of the intended semantics of ‘energy bidding’.

For example, in a smart grid (SG) data space the semantics of bidding energy from green energy providers is defined in terms of certain published and subscribed capabilities.

The two specifications from above *OPER* and *COND* operations actualized with actual system state parameters demonstrate the capabilities of the data space of the domain ‘energy bidding’ which are declarative statements, and which declare the options of one publisher and a non-empty set of subscribers to the ‘energy bidding’ domain.

The publisher shall be able to define a time phase of a demand on certain bids between the operations ‘startofbidding’ and ‘endofbidding’ for which the publisher ‘SGOperator’ needs to be licensed by an entry in the ‘digital rights’. The data provenance annotations record the initialization and the closing of the compound event ‘energy bidding’. During

the defined phase the ‘SGOperator’ may receive bids issues from subscribed ‘SGProviders’ and may analyze them for use in the SGOperator’s domain. The result of analysis could be a sorted list of bids with respect to the priorities of the SGOperator.

Similarly, ‘SGProviders’ may subscribe independently to the domain ‘energy bidding’ and wait asynchronously for the notification of a ‘biddemand’ on requested energy demands from a publishing stakeholder. Once received and accepted the notified ‘SGProvider’ will prepare and publish their bids on their capabilities of energy loads. If the bids are notified in-time to the ‘SGOperator’ the ‘SGProviders’ may wait for a ‘bidaccept’ event which contains the contractual conditions of the energy delivery load.

2.4 RIP meta-data Annotations

In the data space basic data is represented by $\langle n\text{-tuples} \rangle^3$ and managed by the ‘space’ which is organized as a publish-subscribe data space without channels or protocols interconnecting actors and recipients. Hence a data space is over-arching the historic communication layers of existing reference architecture models. All entries of an n-tuple are data atoms while lists comprise other lists and atoms. Data n-tuples of the data space are annotated with so-called meta-data RIP collection that provide annotated information to the application which is generating, acquiring and using data with respect to the intended semantics related to **digital rights (R)**, with respect to incorporated stakeholder categories by **digital identities (I)** or, with respect to trustworthiness of data manipulations by stakeholders and inally by **data provenance narratives (P)** by which the historic data maintenance can be taken into consideration when evaluating the value of a data asset.

Digital rights prescribe to associated data assets the allowances for stakeholders who want to apply the data. Digital identities identify stakeholders requesting the application of data assets. And finally the data provenance narrative exploit the type graph semantics for the reasoning on data provenance and maintenance.

A possible trajectory through the provenance narrative which could have more than one path that is gained from the exploitation of a **graph type** from the domain ‘energy bidding’ (see clause 2.1). To represent trajectories the operators ‘||’ and ‘→’ are applied. The operator ‘||’ demonstrates **independency** which is also called ‘**parallel**’ operator and the ‘→’ operator demonstrates **sequency**. If two operations are independent from each other they could occur in any sequence. If two operations are sequential, they must occur in the prescribed sequence.

Hence a possible trajectory representing a certain behaviour would be observable:

```
(startofbidding! || biddemand?)
```

³ The term $\langle n\text{-tuple} \rangle$ represents the the tuple itself with n entries. In the following text the brackets are omitted since the meaning of ‘n-tuple’ or simply ‘tuple’ always refer to the mathematical theory with induction.

→ (bidprepare! → bidanalysis!)
→ (bidsorting! → bidaccept!)
→ (endofbidding) .

In the trajectory above four states of the asset value stream are marked by brackets ‘(.)’.

2.5 Value Stream Domain Semantics

The LC asset value stream uses the concept of a ‘semantic flow of objects’ (e.g. data, energy, products, goods etc.) from stage to stage of a value stream of interest, from generation to usage, from creation to obsolescence, from distribution to consumption etc.

A semantic flow means there exists ‘knowledge’ about the purpose of an object flowing through the data space from the originator to its recipients. Originators and recipients are interconnected through a shared data space which obey the identified artifacts to serve for some purposes of interest. A purpose is the semantics of a certain application domain of interest, for example, a bidding phase in a smart grid data space for producing green electricity to be injected into the smart grid (SG). Another example is a customer that want to subscribe for a car rental service at a published selling point etc.

Hence domain semantics in the data space is defined using the principles of offers and demands. Thus, there is at least one originator for demands and there are many recipients of the demands which are interested in offers respectively bidding.

In an electrical smart grid ecosystem, the smart grid (SG) operator is allowed to define the domain semantics of a bidding phase because the SG operator is responsible for the stable operating of the whole SG. After checking the current situation, the SG operator will publish the bidding demand to the data space. For the definition of the semantics the SG operator uses a library of semantic artifacts to assemble the provided service with the necessary semantics.

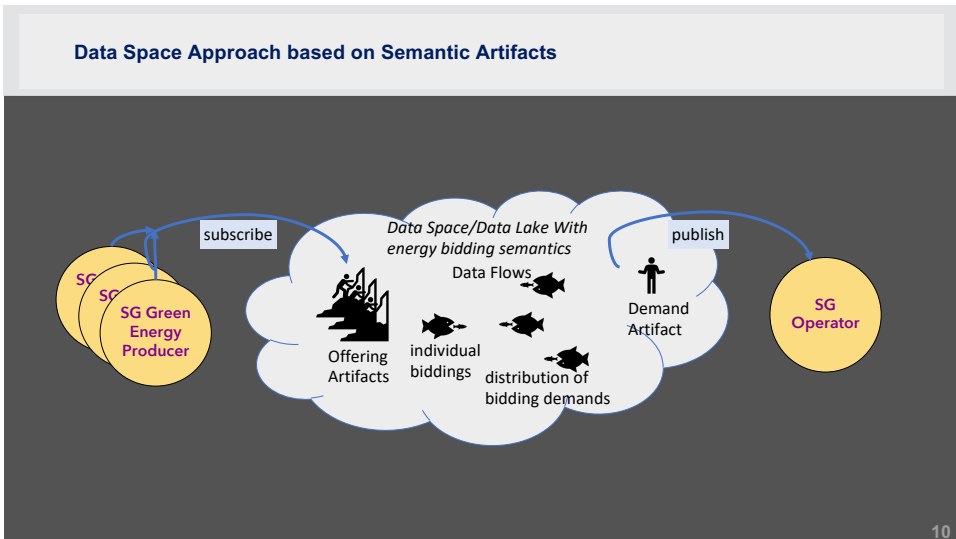


Figure 1: Artifact-based design of a data space

In the example from above and in figure 1 one kind of artifact is the **demand artifact** that may support to describe the object needed, the quality of the object and time period of demanding etc. Another kind of artifact is the **offering artifact** that allows recipients registered e.g., at a semantic domain of energy bidding to subscribe for notifications of published demands, in order to prepare a bid and to publish it back into the data space.

If the bidding phase has not yet been closed the energy bids published via the offering artifacts are returned to the demanding artifact i.e. the originator of the bidding phase. There, analysis and sorting *OPERations* can happen before the bids are notified to the demanding SG operator.

2.6 The data space RAM

The operational program outlined in section 2.3 and the paradigm of the data lake of figure 1 both sketch the semantic interoperations of the applied artifacts of ‘bid demanding’ proxy and ‘bid offering’ proxy both, derived from the application domain ‘energy bidding’.

Thus, figure 1 and the program sketch both illustrate how the semantic artifact proxies of a publishing stakeholder ‘SG Operator’ and a set of subscribing stakeholders ‘SG Providers’ can be implemented in the data space. While stakeholders operate outside the data space proxies operate inside the data space and represent the outside world. The inside world performs a hierarchy-less communication based on the sketched unconditioned operations (OPER) and conditioned operations (COND).

In figure 2 you will find the sketch of a data space reference architecture model (RAM) that comprises the three dimensions of the data space stakeholders (DSS), the architecture of the data application governed by the rules of the underlying common data space (CDS) either via a data platform (DPF) or via the data marketplace (DMP).

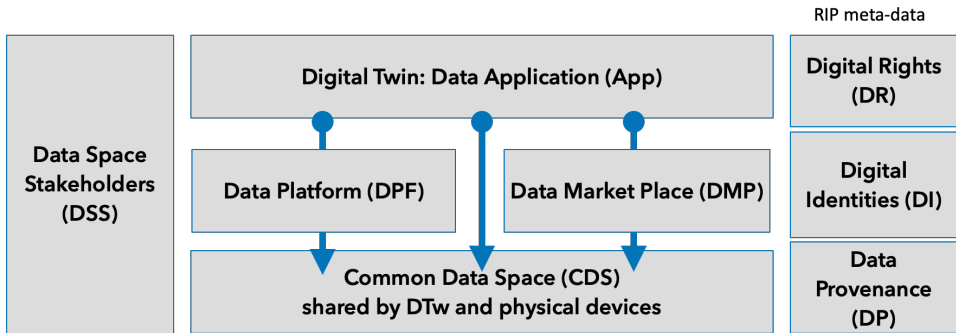


Figure 2: The Reference Architecture Model of a Data Space [Source: SC27/WG4 data provenance project]

The application in figure 2 applies data from the common data space (CDS) for certain purposes. If the application would be a digital twin then the purpose could be keeping control of the physical phenomena of the machinery and tools in an industrial environment. According to figure 2 there are three tiers of an application operating on the common data space.

- Application respectively digital twin using data from the common data space,
- Access provisioning by data platforms or marketplaces and stakeholders,
- Common data space ruled by digital rights (DR), digital identities (DI) and data provenance narratives (DP) of trust.

2.7 Common Data Space Industrial Application

Industrial applications are defined by a network of highly specialized stakeholders required to collaborate in dependency from the real-time condition monitoring in the network which is called CCM – Collaborative Condition Monitoring. The CCM approach presented in [Jo23a, Jo23b, Jo23c] is AAS - Asset Administration Shell-based which means that all components respectively digital data sub-models of an asset (e.g. production machinery) are accessible through the AAS.

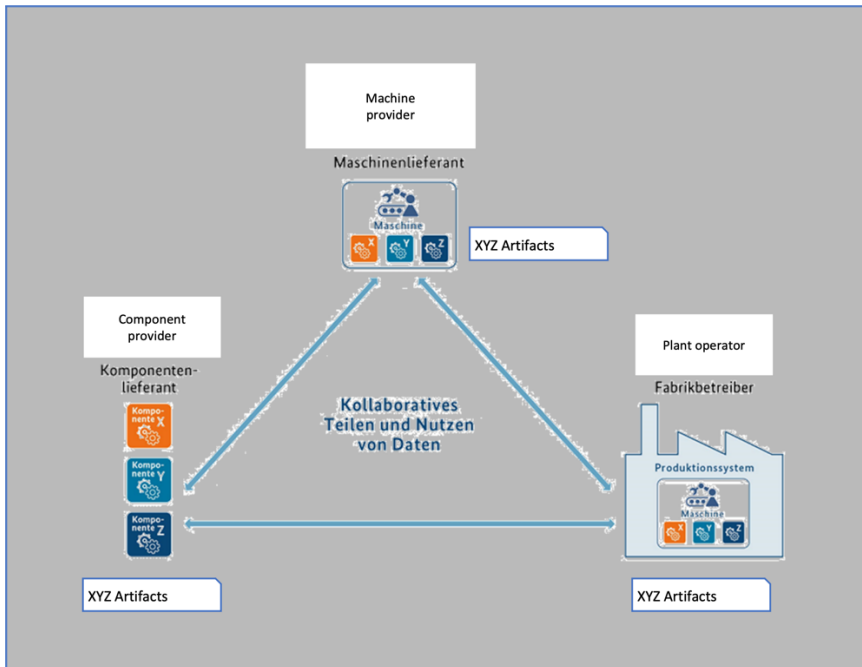


Figure 3: Artifacts of Collaborative Condition Monitoring [Source: Michael Jochem, VDE PI4.0]

In figure 3 an industrial production network is sketched. It comprises the three basic stakeholders of production comprising machine provider, component provider and plant operator. However, a plant may accelerate a production by more than one system X_s, Y_s, Z_s and each system with more than one machine from different machine providers X_m, Y_m, Z_m , and each machine comprises more than one component from different component providers X_c, Y_c, Z_c .

Thus, for $n=3*3$ stakeholders $\langle X, Y, Z \rangle$ from the three categories $\langle s, m, c \rangle$ a network of $1/2*n(n-1)=36$ connections would be required to make the stakeholders aware of conditions monitored. To approach the interconnection issue more straight forwardly, would be to invent a data space common to all 9 stakeholders of the CCM example from above. Then, all stakeholders are all enabled to publish and subscribe their offers and demands into the data space asynchronously. The data space plays the role of a data broker among the three categories of data space participating stakeholders $\langle s, m, c \rangle$.

In figure 3 the application of the so-called CCM fractals (for short: fractals) is sketched. It looks like an extension of publish-subscribe collaboration paradigm of figure 1. In figure 1 there is only one pair of publish-subscribe artifacts. Hence to provide with a relationship between figure 1 and figure 3 only three pairs of publish/demand-subscribe/offer artifacts must be implemented into the data space such that the three stakeholder kinds (machine provider – plant/system operator – component provider) get enabled to collaborate with

all other stakeholders by referring to proxy artifacts of the pairs publish/demand and subscribe/offer.

2.8 Security and Privacy of the RIP meta-data

In critical infrastructures (CRITIS) RIP meta-data can be secured by the triple <C, I, A> of Confidentiality, Integrity, and Authenticity or, by encoding RIP metadata and data using Distributed Ledger Technologies (DLT). In that case, City users can access trustworthy data obeying identified sources and provenance narratives that fit with the user's own narratives.

3 Conclusions

The 4th dimension of semantics added to Reference Architecture Models (RAM) is applied to phenomenological (physical) assets twinned with data set assets aiming at various purposes and goals such as I4.0 production achieving high quality of products, stable energy distribution comprising renewable energy generators and bulky energy industrial or home consumers, smart city services relayed on altruistic data provisioning, smart building administrations relayed on real-time data sensing and IoT device actuating.

In practice the data asset value stream twinning the phenomenological (physical) asset value stream shall be maintained by the concept of a digital twin. Hence the digital twin may analyze in real-time the aggregated data sets including RIP meta-data information and may give feed-back to adjust the physical processes of the phenomenological twin in the realms of I4.0 production, electricity, smart infrastructures of cities, buildings, fabrics or others.

Operational semantics is defined by the stepwise achievement along the RAM axis of LC VS. The LC VS are asset oriented. One asset obeys the hierarchical configuration of the physical machinery with the specific purpose of production or, energy distribution or, the altruistic data provisioning of all stakeholders of a city or a building etc. The physical asset is twinned by data that reflects the state transformation of the product towards defined quality constraints to be achieved. Hence a physical production machinery state has its equivalence by a digital production model state. This dual state representation is the intrinsic semantic state of a system of systems.

Semantics is often understood from an interoperation point-of-view. In the various RAMs interoperation is achieved by the axis of modelling a set of hierarchical communication protocols. The interoperation hierarchy comprises the asset, the integration of LC VS phases of asset design and asset operation, the communication means comprising all stages of the LC VS, the information models respectively *data spaces* of the *asset type design and asset instance operation*, the functional and business interoperability.

Since a data space based on the publish-subscribe paradigm that provides interoperability to all entities that need interoperation capabilities, communication protocols are not required any more for establishing semantic interoperability.

Bibliography

- [FHA99] Freeman, Hupfer, Arnold JavaSpaces™ Principles, Patterns, and Practice, 1999 Sun Microsystems.
- [UN23] UNECE Webinar on Semantics and **Semantic Interoperability** – eB MoU – Program and presentations, logged-in June 27, 2023
- [ISO22a] ISO/IEC JTC1/WG11 WD/DIS 5087-2 ‘**City Data Model**: Part 2 City Level Concepts.
- [IEC23a] IEC TC184 **JWG21** N332 TR Draft Application of IEC63339 to SMRM r7: Application of IEC63339 to Smart Manufacturing Reference Models.
- [Jo23a] Jochem M.: PI4.0 Workshop Presentation on Security Approaches 24.2.2023.
- [Jo23b] Jochem M.: PI4.0 Presentation on The **AAS Dataspace** for Everybody, proposed to IDTA.
- [Jo23c] Jochem M.: PI4.0 Workshop presentation on Security Approaches related to PI4.0/IDTA, IDSA, Catena-X(EDC), 2.2.2023
- [VDE23] VDE SPEC 90010 v1.0 (DE, in German language) Guidelines for establishing **executable semantic standards** (SemNorm), 2023 Project Leader Jan deMeer, HTW Berlin, smartspacelab.eu GmbH et al.
- [BK23] Digitalverband Bitkom **Data Mesh** – On finding and using Data Potentials (DE, in German Language), 2023 David Schönwerth et al. Bitkom e.V. 10117 Berlin.

No Attacks Are Available: Securing the OpenPLC and Related Systems

Wael Alsabbagh¹, Chaerin Kim¹ and Peter Langendörfer¹

Abstract: The use of Programmable Logic Controllers (PLCs) expands in industrial domains, which makes ensuring the security of Industrial Control Systems (ICSs) become paramount. The OpenPLC project, the first open-source initiative, provides flexible and cost-effective PLC solutions to build up affordable test-beds, as well as conduct experiments and academic researches. This project has wildly grown in the last few years, thus it is essential to address the most emerging security challenges it encounters. This paper introduces a new OpenPLC architecture, called OpenPLC Aqua, provided with a set of security solutions designed specifically to overcome the vulnerabilities that the current OpenPLC versions are prone to. The new OpenPLC architecture includes four security features: 1) user credentials encryption, securing the *Webserver*, Whitelisting and secure SSL/TLS communication channel. The OpenPLC Aqua software was tested against several attack scenarios that were feasible against the old OpenPLC versions. Our experimental results showed our enhanced OpenPLC software is secure and resistant against several attack scenarios e.g., authentication, injection, Man-in-the-Middle and replay attacks. The OpenPLC Aqua is publicly available and a proof of concept demo is also published with this paper.

Keywords: OpenPLC, Security Solutions, Mitigation Solutions, Industrial Control Systems.

1 Introduction

Programmable Logic Controllers (PLCs) play a crucial role in monitoring and controlling various industrial processes and applications e.g., manufacturing, energy management, transportation, etc [ALS2]. Real hardware PLCs are pricey and sometimes are unaffordable for scientists to design small experimental labs, and conduct their experiments or academic researches. The OpenPLC [ALV1] project provides an appropriate alternative solution for implementing low cost PLCs and Industrial Control System (ICS) environment. Its open-source and collaborative nature offers a unique platform for developers, researchers, and even industrial experts to contribute, innovate, and test new PLC based solutions and technologies. With a user-friendly interface, extensive documentation, and online community forums, the OpenPLC allows users to learn, exchange ideas, distribute PLC codes and compatibilize it with a wide range of hardware platforms, including popular Single Board Computers (SBCs) such as Raspberry Pi and Arduino [ALV2]. However, since the use of the OpenPLC has significantly grown

¹IHP – Leibniz-Institut für innovative Mikroelektronik, Frankfurt (Oder), Germany, { alsabbagh, chaerin, Langendoerfer }@ihp-microelectronics.com.

in the last few years few years, it is essential to ensure the integrity, availability, and confidentiality of the project. Like other ICS components, the OpenPLC is susceptible to a range of security issues that can compromise the operation of OpenPLC based systems and disrupt the physical processes they control. These security issues arise from various factors e.g., software design flaws, insecure communication protocols, weak authentication protocols, and others.

One of the primary security concerns in OpenPLC is the potential for unauthorized access to the OpenPLC configuration [ALS1]. If an attacker gains access, he can maliciously modify the logic controlling industrial processes, potentially leading to physical damages, production disruptions, and safety hazards. Another security challenge in the OpenPLC is the presence of software vulnerabilities. In our former paper [ALS1], we showed that the OpenPLC still has some design flaws that allow adversaries to steal ready-to-execute programs from the *Websvr*, and conduct successful control logic injection attacks. Furthermore, the communication channels used by the OpenPLC systems are also prone to security risks. The absence of encryption exposes sensitive data e.g., user-credentials [ALS1], control commands [ALS1, ALV2, ALS3], configurations [ALS1, ALS3], etc. to attackers who are capable of intercepting, manipulating, and compromising the integrity of the data transmitted over the network. Therefore, we introduce in this paper a more secure OpenPLC software that is integrated with robust security solutions to enhance the overall security posture and trustworthiness of the OpenPLC project. First, we implemented an Advanced Encryption Standard (AES) algorithm to encrypt the user credentials i.e., the *username* and *password*. The outputs of the AES algorithm are then encoded from binary to ASCII, using the Base64¹ algorithm before they are finally stored in the *openplc.db* database. This managed successfully to close the vulnerability we found in our former paper [ALS1] and attackers are no more able to neither sniff the user credentials from the network nor to steal them from the *openplc.db*. Furthermore, the *Websvr* in the OpenPLC Aqua is accessible only by users with root permissions. This prevents external adversaries from accessing ST file copies of the programs that the user uploaded into the OpenPLC in the past [ALS1]. We also implemented a whitelisting approach that allows only pre-approved or trusted users to upload a new ST file into the OpenPLC. This approach aims at detecting any malicious attempt to manipulate the running program by an unauthorized user. Finally, the OpenPLC Aqua secures the data transmitted between the client (user) and server (OpenPLC) over the internet by using SSL/TLS (Secure Sockets Layer/Transport Layer Security) protocols. This ensures that the information remains confidential and cannot be intercepted or read by malicious attackers i.e., it helps to prevent eavesdropping and data tampering.

The rest of the paper is structured as follows: Section 2 provides related works, while Section 3 introduces the OpenPLC Aqua architecture. In Section 4 we show our experimental results and conclude this paper in Section 5. Please note that our experiments and the open-source code of OpenPLC Aqua are publicly available as Section 6 shows.

¹ <https://docs.python.org/3/library/base64.htm>

2 Related Work

The OpenPLC project was first released by Alves [ALV1] in 2014, and consists of Editor, *RunTime*, and Human Machine Interface (HMI) Builder see figure 1.

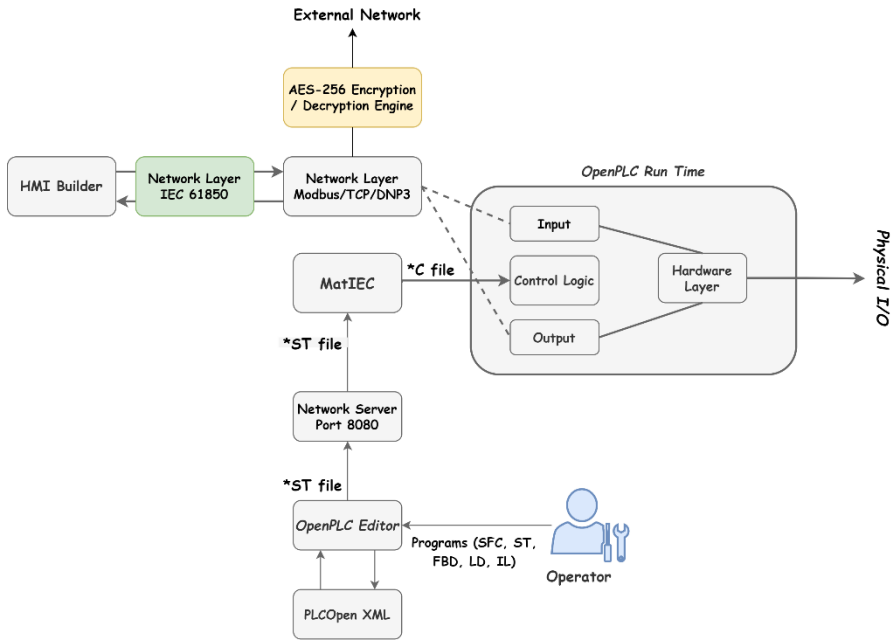


Fig. 1: OpenPLC Internal Architectures (boxes highlighted in gray color indicates the original OpenPLC; in yellow color the features added in OpenPLC NEO, in green color the features added in OpenPLC61850)

The development environment that is used to create programs is called Editor. This Editor supports program-development using several programming languages identified by IEC 61131-3 e.g., Function Block Diagram (FBD), Ladder Diagram (LD), Structured Text (ST), Instruction List (IL), and Sequential Function Chart (SFC). Based on IEC 61131-3, the programs are saved as XML files. Once the program is created, the built-in module in Editor compiles all programs into an ST file. This ST file is then utilized in the OpenPLC *RunTime* for the execution of control logics. The communication protocols supported by OpenPLC are Modbus and Distributed Network Protocol 3 (DNP3) using the default ports 502 and 20000 respectively.

In 2017, Alves [ALV3] introduced an enhanced version of project, called OpenPLC Neo, by adding an AES Encryption Layer placed between the Network Layer and the external network (see figure 1, the yellow box). This AES Encryption Layer encrypts all the

messages sent from the OpenPLC to an external client e.g., user, HMI, etc. by using a symmetric key and then forwards the appropriate ciphertext to the external network. In opposite, it decrypts the messages it receives by using a symmetric key provided by the user, and sends it then to the network layer for further processing. The OpenPLC Neo enables a secure end-to-end encrypted channel between the PLC and the user or HMI, without requiring any external hardware to encrypt the data. However, the project relies totally on the secrecy and integrity of the encryption keys. If the keys are not properly managed, such as being weak, compromised, or improperly stored, it can weaken the overall security of the AES encryption and expose OpenPLC Neo to cyber-attacks. Roomi [ROO1], in 2021, introduced the OpenPLC61850 that was also an enhancement of the OpenPLC project (see figure 1, the green box). The new version supports the International Electrotechnical Commission (IEC) 61850 protocols², including IEC 61131-3 standard [TIEG] for programming the PLC logics and uses MatIEC compiler [DESO] for the compilation of the programmed logics. In a follow up work Roomi [ROO2] evaluated the design of the OpenPLC61850 under network-based attacks e.g., false data and command injections. Their attack scenarios managed successfully to manipulate different Circuit Breaker (CB) that are communicated to the OpenPLC. However, Roomi did not introduce particular security solutions to protect the OpenPLC61850 against cyber-attacks and recommended only to include the security measures proposed in [HUSS] as a part of forthcoming research.

3 OpenPLC Aqua Security Features

Figure 2 depicts the internal architecture of our OpenPLC Aqua software. As can be seen from the figure, four additional security features are integrated in this version: 1) AES-128 Encryption Algorithm, 2) locked *Webserver*, 3) Whitelisting function and 4) SSL/TLS communication channels. In the following we elaborate our security features in more detail.

3.1 AES-128 Encryption Algorithm

The encryption process that we implemented in the OpenPLC Aqua aims at encrypting the user credentials i.e., username and password, by applying a customized AES-128 encryption algorithm as depicted in Figure 3. The plaintext messages, here the *username* and *password*, are divided in chunks of 128 bits each. If the original message size is not multiple of 128, the last block is padded with random bits to form a full 128-bit block. The Cipher Block Chaining (*CBC*) is used then to process the 128-bit blocks. At this step a random Initialization Vector (*IV*) is provided as the starting of pseudo-block.

² libIEC61850. <https://webstore.iec.ch/searchform&q=61850>.

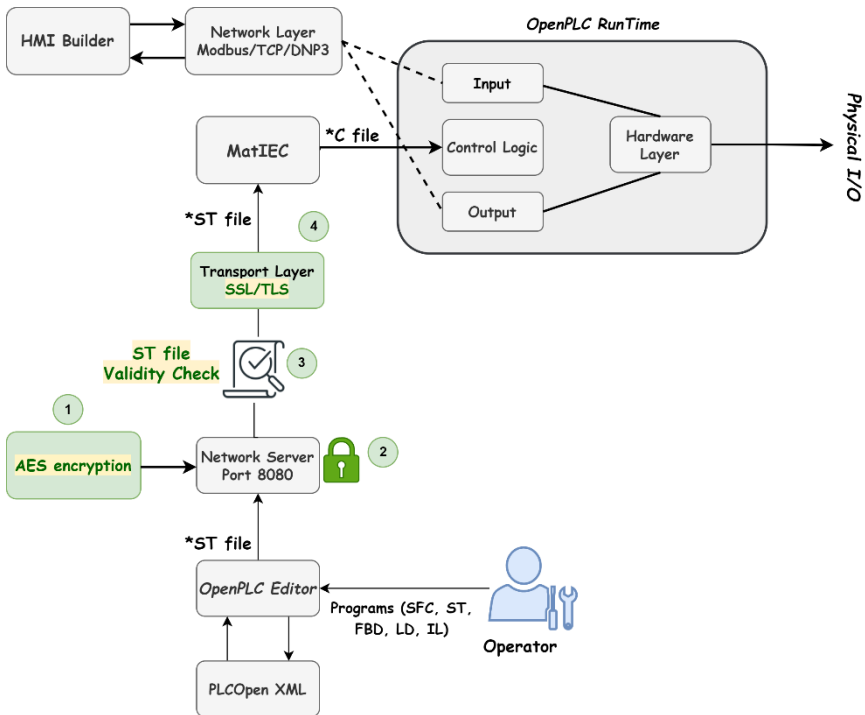


Fig. 2: OpenPLC Aqua Internal Architecture Diagram

In this work, we used a function called *os.urandom*³ that is provided by python library to generate two random 16 bytes. One is used as a secret key (K) and the other as an IV . Both K and IV are stored in the *Webserver*. For the *password*, we used the IV to encrypt the initial block while the K is used to encrypt the entire blocks. Whereas in the *username*, we used the K and IV in a reverse way. The K here is utilized as an initialization vector to encrypt the initial block, while the IV is used as a key to encrypt the entire blocks. Once the ciphertext of both *username* and *password* are calculated, the ciphertexts are then encoded with the help of the Base64 encoding algorithm. The results of the encoding process are finally presented in ASCII format and stored in the *openplc.db* database see figure 4.

3.2 Securing the Webserver

Since the *Webserver*, which contains sensitive data and information about the entire project e.g., user credentials, user programs, etc., was the hugest security gap in the

³ <https://docs.python.org/3/library/random.html>

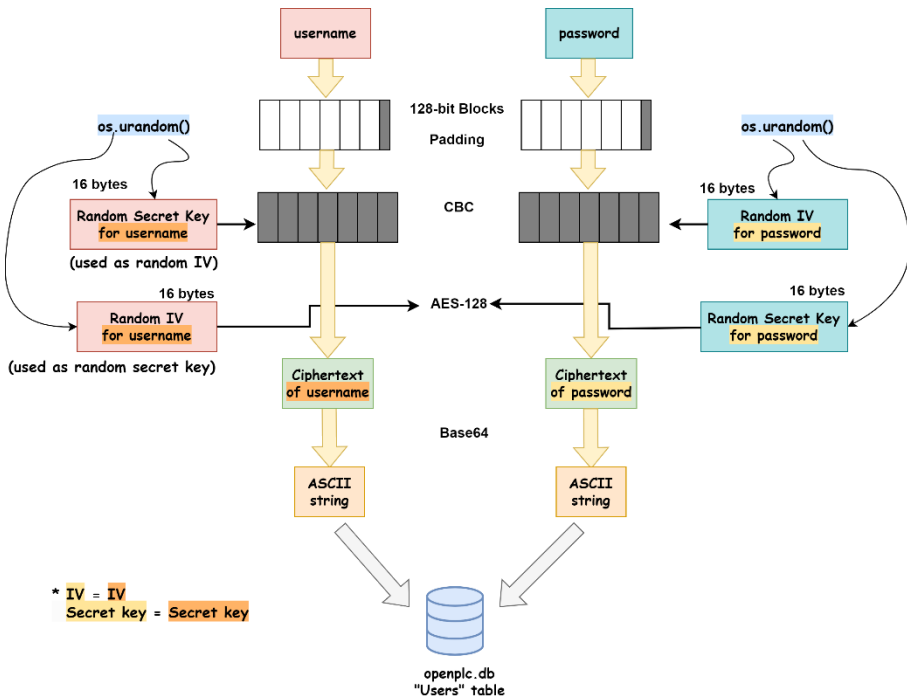


Fig. 3: The AES-128 Encryption Algorithm used in OpenPLC Aqua

OpenPLC as our investigations showed in [ALS1], the OpenPLC Aqua limits the access to the *Webserver* to only legitimate users with root permission. In addition, we also noticed that the OpenPLC keeps copies of all uploaded programs in the *Webserver*. These copies are ready-to-execute ST files and attackers who are still capable of accessing the *Webserver* can retrieve the programs and use them to inject the OpenPLC e.g., replace the currently running program with an old one. The problem in the OpenPLC software is that those copies are irremovable. Meaning that even if the user deletes the programs from the user-dashboard, their corresponding copies in the *Webserver* are not removed [ALS1]. To avoid abusing those copies, we improved the OpenPLC Aqua security by adding a script that removes any copy from the *Webserver* once the user deletes its corresponding program from the user-dashboard.

```

-----Users-----
UID: 10 | Name: OpenPLC User | ID: openplc | Email: openplc@openplc.com | Pwd: o
penplc | pict: None

```

Fig. 4a: The *username* and *password* before the encryption as they are stored in the *openplc.db*

```

-----Users-----
UID: 10 | Name: OpenPLC User | ID: wHky27E3NF3fs1E6TpELtg== | Email: openplc@ope
nplc.com | Pwd: lTcuJRKU9apvtJcW495/gw== | pict: NULL

```

Fig. 4b: The *username* and *password* after the encryption as they are stored in the *openplc.db*

Fig. 4: The representation of the *username* and *password* in both OpenPLC software versions. Fig. 4a: in the OpenPLC, Fig 4b: in the OpenPLC Aqua

However, removing the copies from the *Webserver* has limitations. The OpenPLC starts launching any program (already uploaded), by first reading the name of the program from the *Active_Program* index, and recalling the corresponding ST copy file in the *Webserver* the *Active_Program* indicates. Since the OpenPLC Aqua deletes a copy together when the user removes its original program from the user-dashboard list, the name of the copy still remains in the '*Active_Program*' index. Meaning that, the *Active_Program* indicates to a missing copy (removed) in the *Webserver*. This causes an error in launching the software itself.

To overcome this challenge, we configure the OpenPLC Aqua with a blank program (fake program) that assists the software with launching itself. The objective of this program is not to run on hardware controllers, so the software keeps giving an alert to the user to launch a new proper program see figure 5. It is worth mentioning that we designed this blank program for a very specific use i.e., to launch the software itself. Therefore, if any user attempts to start the blank program in the OpenPLC, the session will be immediately closed. This adds more security to the software as it prevents any attempt to abuse the blank program.

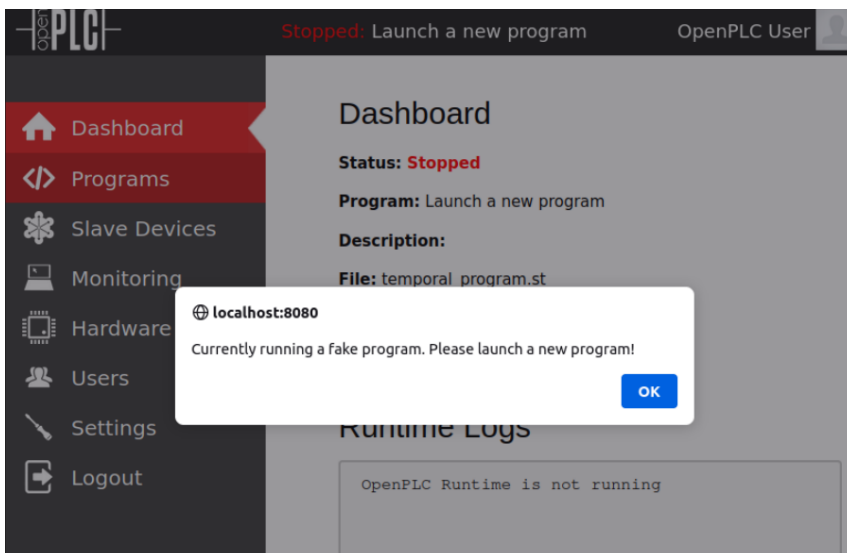


Fig. 5: OpenPLC Aqua alarms the user to launch a new program

3.3 Whitelisting Approach

In the OpenPLC Aqua, we added a function that verifies the ST file content, as well as the IP address each time a new ST file is uploaded into the OpenPLC. This bans any attempt to maliciously upload an ST file from suspicious users see figure 6.

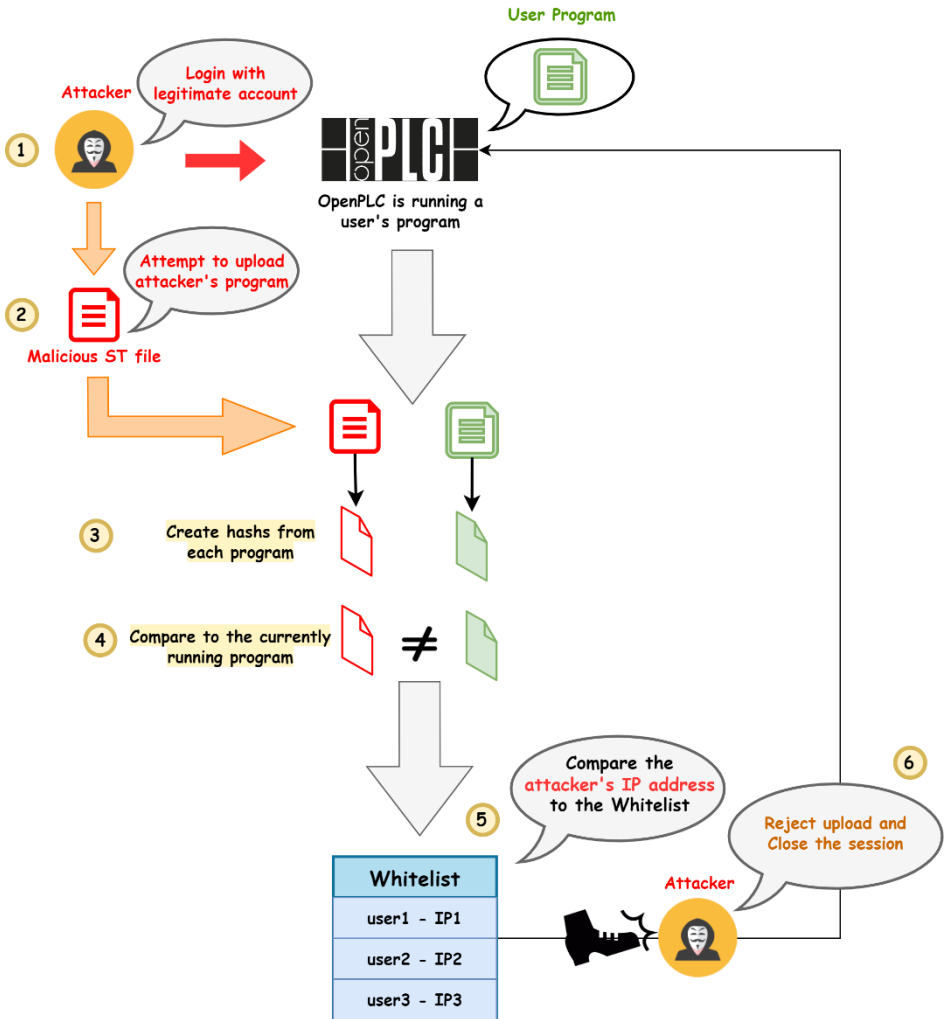


Fig. 6: Whitelisting approach in OpenPLC Aqua

At the beginning, if a trusted user creates a new account using the user-dashboard, a pair of his IP address and username will be automatically added to the whitelist. By assuming

that only the default-account is registered in the user-dashboard, and the user of the default-account attempts to upload a different program than the OpenPLC is currently running; the OpenPLC Aqua creates two hashes from both: the old program and the new one. After that, it compares the hashes of the programs and only if they are not matched, it checks the IP address of the user who attempts to upload. If his IP address is not listed in the whitelist, it then rejects the upload process, closes the session and the suspicious user is automatically logged out from the session. Otherwise, i.e., the IP address of the user is listed in the whitelist, the OpenPLC approves the upload process and the new program is then ready to start. This approach generates a solid security advantage to our project. For instance, even if the attacker could somehow access the OpenPLC with legitimate user-credentials and accounts, he would not be able to upload his malicious program.

3.4 SSL/TLS Communication Channel

Our investigations conducted on the last version of OpenPLC in [ALS1] proved that the software is prone to replay attacks. We showed that adversaries with appropriate attacking tools can reproduce certain traffic captures (HTTP packets) from previous communication sessions over the Internet to access and make malicious changes in target OpenPLC based systems. SSL/TLS (Secure Sockets Layer/Transport Layer Security) is a communication approach that provides an appropriate secure channels over the internet. It establishes an encrypted connection between a client (user) and server (OpenPLC). This encryption ensures that the data transmitted between them remains confidential and cannot be easily intercepted or revealed by unauthorized parties as the case in the existing OpenPLC software. Therefore, to introduce confidentiality, integrity, and end-point authentication in the OpenPLC Aqua software, we implemented the SSL/TLS handshake approach depicted in figure 7.

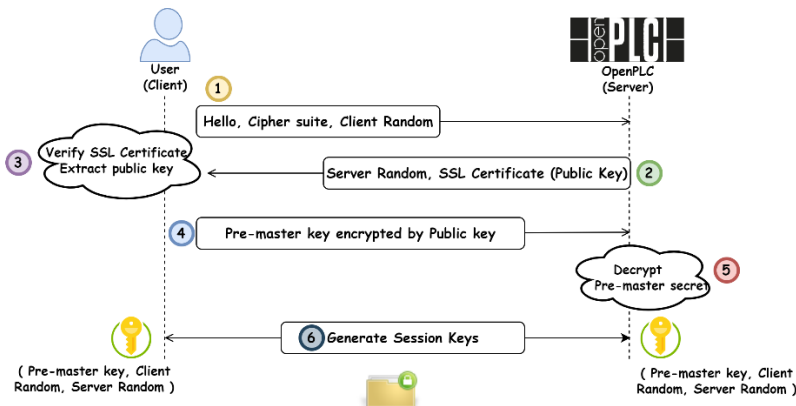


Fig. 7: SSL/TLS communication process approach

First 1) the client (user) sends a *Hello* message informing the server (OpenPLC) that he wants to establish a secure channel between the both parties and provides the OpenPLC with his cipher suites and the compatible SSL/TLS version. Then 2) the OpenPLC responds back to the user by sending its random, the chosen cipher suite, and the SSL certificate. After that 3) the client verifies the SSL certificate with the authority and extracts a public key from it. The client then 4) generates an encrypted pre-master key by using the extracted server public key. Afterwards, the OpenPLC 5) decrypts the transmitted pre-master key to verify the extracted key. Once the verification is finished, both client and server 6) create the same session key from the random they got from each other and the pre-master key. From now on, all data transmitted between both client and server is encrypted and decrypted using the master key.

4 Experimental Results

To assess our developed OpenPLC software, we conducted five different attack scenarios on four OpenPLC versions: OpenPLCV3, OpenPLC Neo, OpenPLC61850 and OpenPLC Aqua. All Our experimental results are listed in table 1.

Attack Scenario	OpenPLC V3	OpenPLC Neo	OpenPLC61850	OpenPLC Aqua
Authentication Attack	✓	✗	✓	✗
Man-in-the-Middle Attack	✓	✗	✓	✗
Control Logic Injection Attack	✓	✓	✓	✗
Replay Attack	✓	✗	✓	✗
Access Attack	✓	✓	✓	✗

Tab. 1: The success of various attacks against different OpenPLC software versions

As we noticed from the table, our OpenPLC Aqua is secure against all the attacks we conducted and show more resistance compared to the previous OpenPLC software.

5 Conclusion

This work defined an alternative design to secure OpenPLC based systems and SCADA networks by introducing and integrating four security features in the new OpenPLC Aqua software. The new software protects the OpenPLC and its systems from different attacks that were feasible in the last software versions. For instance, encrypting and encoding the user credentials prevent adversaries from performing authentication attacks. Furthermore,

the OpenPLC Aqua locks the *Webserver* allowing only legitimate users with root permissions to access. In addition, the ST file copy is removed from the *Webserver* once the user deletes its corresponding program from the user-dashboard. We also secure all the communication session between the user and OpenPLC by implementing SSL/TLS protocols. Finally, to ensure that only authorized and trusted users can upload/change the program in the OpenPLC, we introduced a whitelisting approach which prevents malicious infections. The experimental results proved that OpenPLC Aqua was able to significantly enhance the security of OpenPLC project by preventing different attack scenarios.

6 Demo and Open-Source Code

Please refer to [ALS4] for the demo of our experiments, and the GitHub repository [ALS5] for detailed information of OpenPLC Aqua.

Bibliography

- [ALS1] Alsabbagh, W.; Kim, Ch.; Langendörfer, P.: Good Night and Good Luck: A Control Logic Attack on OpenPLC. Accepted at IECON 2023 – 49th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 2023. <https://ww.easychair.org/publications/preprint/Swt2>.
- [ALS2] Alsabbagh, W.; Langendörfer, P.: A Flashback on Control Logic Injection Attacks against Programmable Logic Controllers. *Automation* 2022, 3, 596–621. <https://doi.org/10.3390/automation3040030>.
- [ALS3] Alsabbagh, W.; Amogbonjaye, S.; Urrego, D.; Langendörfer, P.: A Stealthy False Command Injection Attack on Modbus based SCADA Systems. 2023 IEEE 20th Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2023, pp. 1-9, doi: 10.1109/CCNC51644.2023.10059804.
- [ALV1] Alves, Th.; Buratto, M.; De Souza, F. M.; Rodrigues, T. V.: OpenPLC: An open source alternative to automation. IEEE Global Humanitarian Technology Conference (GHTC 2014), San Jose, CA, USA, 2014, pp. 585-589, doi: 10.1109/GHTC.2014.6970342.
- [ALV2] Alves, Th.; Morris, Th.: OpenPLC: An IEC 61,131–3 compliant open source industrial controller for cyber security research, *Computers & Security*, Volume 78, 2018, Pages 364-379, ISSN 0167-4048, <https://doi.org/10.1016/j.cose.2018.07.007>.
- [ALV3] Alves, Th.; Morris, Th.; Yoo, S.: Securing SCADA Applications Using OpenPLC With End-To- End Encryption. (2017) In Proceedings of the 3rd Annual Industrial Control System Security Workshop (ICSS 2017). Association for Computing Machinery, New York, NY, USA, 1–6. <https://doi.org/10.1145/3174776.3174777>.
- [ROO1] Roomi, M. M.; Ong, W. S.; Mashima, D.; Hussain S. M. S.: OpenPLC61850: An IEC 61850 compatible OpenPLC for Smart Grid Research, *SoftwareX*, vol. 17, Jan. 2022, doi: 10.1016/j.softx.2021.100917.

- [TIEG] Tiegelkamp, M.; John, K.: IEC 61131-3: Programming Industrial Automation Systems; Springer: Berlin/Heidelberg, Germany, 2001; Volume VI, p. 376.
- [DESO] De Sousa, M.; Carvalho, A.: An IEC 61131-3 compiler for the MatPLC. EFTA 2003. 2003 IEEE Conference on Emerging Technologies and Factory Automation. Proceedings (Cat. No.03TH8696), Lisbon, Portugal, 2003, pp. 485-490 vol.1, doi: 10.1109/ETFA.2003.1247746.
- [HUSS] Hussain, S. M. S.; Ustun, T. S.; Kalam, A.: A Review of IEC 62351 Security Mechanisms for IEC 61850 Message Exchanges. In IEEE Transactions on Industrial Informatics, vol. 16, no. 9, pp. 5643-5654, Sept. 2020, doi: 10.1109/TII.2019.2956734.
- [ROO2] Roomi, M. M.; Ong, W. S.; Hussain, S. M. S.; Mashima, D.: IEC 61850 Compatible OpenPLC for Cyber Attack Case Studies on Smart Substation Systems. in IEEE Access, vol. 10, pp. 9164-9173, 2022, doi: 10.1109/ACCESS.2022.3144027.
- [ALS4] <https://www.youtube.com/watch?v=knVTQfUdNfU>.
- [ALS5] <https://github.com/rnrn0909/OpenPLC-Aqua.git>.

Autorenverzeichnis

A

Abecker, Andreas, 1473
Abedjan, Ziawasch, 919
Achmann, Michael, 821
Ackermann, Marcel R., 909, 919
Adler, Simon, 461
Ahlborn, Juliane, 363
Ahmad, Raia Abu, 919, 925, 931
Allwein, Florian, 1793
Alsabbagh, Wael, 2085
Altschaffel, Robert, 2041
Annaias, Yves, 1683
Arlinghaus, Julia, 437
Arndt, Hans-Knud, 759
Artiushenko, Viktor, 381
Auer, Sören, 919, 925

B

Babalou, Samira, 1035
Bähne, Rosemarie, 1221
Bahrini, Mehrdad, 677
Balke, Stefan, 841
Balling, Julian, 1833
Barenkamp, Marco, 1407
Barlieb, Christophe, 241
Baroud, Ibrahim, 1607
Barton, Thomas, 147
Bauer, Christian, 1551
Bauer, Laura, 1895
Bauer, Luca T., 1097
Bauer, Thomas, 1909
Bäumer, Frederik Simon, 1815, 1827
Baumer, Sven, 399
Baumgärtel, Hartwig, 1725
Bawa, Vanshika, 1607

Bayard, Guido, 1801
Becker, Alexa, 893
Becker, Ilja, 1739
Becker, Jörg, 1017
Beckmann, Helmut, 967, 991, 1003
Behrendt, Fabian, 431
Behrens, Grit, 1267
Behrens, Janek, 1435
Bender, Christian, 1693
Berger, Arne, 893
Berger, Bernhard J., 211
Berger, Thomas, 1515
Berger-Konen, Caroline, 371
Berges, Marc, 395
Bergmann, Malina, 1181
Bergner, Nadine, 419
Bernardi, Ansgar, 1515, 1639
Berndl, Mario, 1535
Berndt, Julian, 1221
Betzendahl, Jonas, 395
Beyer, Andrea, 391
Beyer, Florian, 1207
Biehler, Rolf, 389, 453
Biemann, Chris, 123
Binder, Frank, 1833
BirgittaKönig-Ries, 1035
Bleier, Arnim, 919
Böck, Ronald, 1591
Bodach, Ronny, 597
Bogdan, Martin, 1645
Böhmman, Tilo, 123
Bökle, Sebastian, 1551
Boll, Alina, 489
Böll, Marvin, 2025

Bookhahn, Marian, 353
Borisova, Ekaterina, 925, 931
Boukhers, Zeyd, 909, 919
Brandt-Pook, Hans, 1815
Bräuer, Paula, 777
Bressel, Fabian, 1221
Brockmann, Carsten, 1851, 1871
Brömme, Josephine, 1181
Brunner, Dominik, 2015
Bruns, Julian, 1473
Brützke, Paul, 1017
Bühl, Johannes, 461
Burghardt, Frank, 231, 295
Burghardt, Manuel, 831
Bürgl, Kim, 1665
Burmeister, Carsten, 223
Busch, Matthias, 377
Buschermöhle, Ralf, 1375, 1399,
1421
Butara, Maryna, 1029

C

Carl, K. Valerie, 523
Castro, Leyla Jael, 909, 919
Christoforaki, Maria, 905
Class, Christina B., 439
Cuno, Silke, 1599
Czarnecki, Christian, 147, 1851

D

Dartmann, Guido, 1503
Decker, Reinhold, 107
Dede, Jens, 1583
Degenkolb, Borislav, 549
Dehghan, Robert, 1483
deMeer, Jan, 2071
Demirözer, Betül, 991
Derksen, Felix, 1945

Dessí, Danilo, 909
Dieckhoff, Christina, 241
Dietrich, Ute, 253, 289
Dietze, Stefan, 909, 931
Dodrimong, Léon Niclas, 289
Döhmman, Indra Spiecker genannt,
675
Dokic, Dusan, 1255
Dörpinghaus, Jens, 1945
Drechsler, Rolf, 73
Dröge, Martin, 341
Durner, Dominik, 1591
D'Souza, Jennifer, 909, 925, 931

E

Eberz-Eder, Daniel, 1529, 1639
Edeh, Natasha, 2041
Egloffstein, Marc, 409, 1041
Eichhorn, Christian, 325
Eichhorn, Domenik, 1129
Eleks, Marian, 561
Enders, Matthias, 1645
Engels, Jens Ivo, 1563

F

Fähndrich, Johannes, 659
Fastnacht, Julian, 1105
Faust, Anna, 341
Fazlic, Lejla Begic, 1503
Feddoul, Leila, 1035
Federer, Marika, 1121
Fehring, Florian, 1267
Feichtinger, Moritz, 863
Felgentreu, Jessica, 371
Fellmann, Michael, 1879
Fey, Goerschwin, 211
Fischer, Martin R., 313
Flegr, Salome, 313

Fleischer, Yannik, 389, 453
Flierl, Philipp, 1633
Förster, Anna, 1583
Fournier, Lisa, 1983
François, Peter A., 1923
Franke, Javier, 1435
Franken, Jonas, 1563
Fransecky, Tanja von, 129
Freese, Maria, 437
Freitag, Christopher, 1289
Freye, Merle, 677
Fuchs, Erich, 1633
Fuchs-Kittowski, Frank, 1149, 1221,
1535
Funk, Burkhardt, 1777

G

Gärtner, Tanya, 477
Gautschi, Thomas, 81
Geissler, Ines, 497, 511
Gembariski, Paul Christoph, 1303
Gerighausen, Heike, 1207
Gilsbach, Judith, 925
Gläser-Zikuda, Michaela, 447
Goeken, Matthias, 977
Golla, Burkhard, 1207
Görzig, Philline Thalia, 377
Goscinska, Daria, 947
Gösling, Henning, 1359
Götte, Gesa, 39
Götz, Mathias, 795
Greiser, Steffen, 1391
Grellmann, Silko, 1057
Grellner, Janosch, 1149
Grimm, Marvin, 353
Grimm, Rüdiger, 675
Grosch, Dorian, 1049
Groth, Christian, 241

Gschwend, Thomas, 81
Gücük, Gian-Luca, 1963
Gudenkauf, Stefan, 1435
Guldner, Achim, 1277, 1289, 1503
Günther, Maik, 1801

H

Haak, Liane, 1399
Haase, Andreas, 223
Haase, M., 1623
Hafemeister, Tim, 1535
Hafner, Alexander, 1751
Hagen, Simon, 1303
Hagenmeyer, Veit, 937
Hajinejad, Nassrin, 1049
Hamel, Lucas, 117
Handke, Stefan, 1057
Hänel, Martin, 385
Hannoun, Nico, 1149
Hartenstein, Sandro, 45
Hartmann, Matthias, 1085
Hastik, Canan, 137
Haugwitz, Jan-Michael, 137
Haupt, Benedikt, 893
Häußermann, Tim, 1725
Hein, Andreas, 1335
Heinbach, Christoph, 1359
Heinke, Florian, 573
Helfer, Dominic, 597
Hennecke, Martin, 415
Hennig, Christine, 901, 905
Hennig, Silvia, 1739
Herbst, Alexander, 677
Herchel, Sabrina, 635
Hergersberg, Pauline, 265
Herklotz, Markus, 89
Hesse, Mario, 1693
Heuschkel, Marie-Luise, 573

Heyer, Gerhard, 1833
Hilbring, Désirée, 1493
Hillmann, Stefan, 377
Hinsche, Maria, 439
Hiyama, Atsushi, 325
Hlawitschka, Mario, 1683
Hoffmann, Anna, 1455
Höfinghoff, Max, 1399
Höfinghoff, Maximilian, 1375, 1421
Hofmann, Florian, 447
Hofstedt, Petra, 1739
Holstein, Katharina, 1615
Holz, Felix, 1879
Honekamp, Wilfried, 649
Höper, Lukas, 453
Hopf, Lisa, 439
Hoppe, Christoph, 1323
Hoppe, I., 1623
Hornung, Gerrit, 675
Hoyer, Jonas, 1035
Hüsing, Sven, 453
Hussaini, Mortesa, 1653

I

Ifenthaler, Dirk, 1041
Inami, Masahiko, 325
Iovine, Ivan, 347

J

Janzen, Sabine, 1255
Jaraus, Wolfgang, 1615
Jetzinger, Franz, 399
Joachim, Silvia, 415
John, Ulrich, 1709
Johnson, Sidney Leroy, 45
Junger, Dennis, 1197, 1289
Jussen, Ilka, 1309

K

Kadi, Ahmad, 1639
Kahl, Anja, 253
Kahle, Reinhard, 749
Kammler, Friedemann, 1303
Karimanzira, Divas, 1493
Karmakar, Saurav, 909, 919, 931
Kaschuba, Reinhard, 1645
Kaubukowski, Kenn, 1683
Kelle, Robert, 1335
Keller, Johannes, 925
Kellermann, Stefan, 583
Kemmler, Kai Julian, 689
Khorasani, Sima Attar, 919
Kienle, Andrea, 765
KI-Kompetenzen und -Funktionen,
Selbsteinschätzung zu, 445
Kilias, David, 1615
Kim, Chaerin, 2085
Kirdan, Erkin, 1995, 2051
Klassen, Gerhard, 1097
Klauk, Stephanie, 805, 883
Klein, Lauritz, 1615
Kleinertz, Rainer, 805, 883
Klinge, Alexander, 1739
Klinge, Xenia, 377
Klinker, Gudrun, 325
Knackstedt, Ralf, 1351
Knauer, Ricardo, 353
Knauer, Uwe, 1615
Koch, Henning, 425
Koddebusch, Michael, 1017
Koelmann, Holger, 1017
Kögler, Kristina, 409
Köhler, Eileen, 607
Köhler-Bußmeier, Michael, 2003
Kohlhase, Michael, 395
Kohls, Martin, 1029

Kohn, Matthias, 677
Korn, Goy-Hinrich, 1375, 1391, 1399,
1421
Kortum, Henrik, 561
Kraft, Angelie, 905, 919
Kratsch, Christina, 197
Krause, Stefanie, 461
Krcmar, Helmut, 1335
Kreideweiß, Sebastian, 1105
Kretzschmar, Vanessa, 1683
Kreupl, Franz, 1833
Kriegel, Jessica, 489
Krieger, Felix, 1777
Krüger, René, 1221
Krump, Gerhard, 873
Kruse, Theresa, 395
Kučević, Emir, 1963
Kuehnel, Stephan, 1895
Kuhn, Jochen, 313
Kukushkin, Maksim, 1645
Kunau, Gabriele, 765
Kuntke, Franz, 1529, 1575
Kuscher, Philipp, 1041

L

Labudde, Dirk, 573, 583, 611, 635
Lam, Hendrik, 223
Lang, Sebastian, 381
Lange, Frank, 137
Lange, Lucas, 549
Langendörfer, Peter, 2085
Lässig, Jörg, 1121
Latif, Atif, 919
Laue, Ralf, 1877
Legat, Christoph, 2015, 2025
Lehmann, Christian, 295
Lehmann, Joel, 1725
Lehnhoff, Sebastian, 937

Lehtosalo, Suvi, 537
Leible, Stephan, 1963
Lenk, Steve, 1121
Leskow, Justus, 1391
Lewandowski, Tom, 1963
Lieber, Matthias, 919
Liebl, Bernhard, 831
Liedtke, Patrick, 759
Lilliestam, Johan, 937
Limani, Fidan, 909, 919, 931
Link, Hendrik, 715
Linner, Thomas, 241
Lippold, Judith, 265
Lobe, Elisabeth, 1115
Lober, Andreas, 1725
Löffler, Felicitas, 1035
Löffler, Sven, 1739
Lohr, Dominic, 395
Lorenz, Anna-Lena, 905
Lorenz, C., 1623
Louis, Sebastian Gerling, 123
Lübbe, Alexander, 147
Lußky, Phoenix, 1105

M

Maaß, Wolfgang, 1255, 1809
Mainzer, Klaus, 735
Maiworm, Adrian, 841
Majuntke, Verena, 149
Maltzan, Stephanie von, 905
Maoro, Falk, 1815
Maria, Marco Di, 1351
Marmor-Drews, Carol, 1029
Martini, Daniel, 1515, 1639
Marx, Eric, 419
Matzka, Stephan, 353
Mauch, Marianne, 1035
McClelland, Jennifer, 1207

McLeod, Andrew, 815
Mechler, Lars, 659
Meier, Klaus-Jürgen, 1763
Meier, Peter, 873
Meister, Matthias, 1955
Meister, Vera G., 147
Menz, Patrick, 1615
Mertzen, Daniela, 95
Michaeli, Tilman, 399
Miller, Dominik, 223
Mittelbach, Konstantin, 1289
Möller, Bonke, 223
Möller, Frederik, 1309, 1323
Morgenstern, Martin, 649
Möstl, Mona, 957
Moualeu-Ngangue, Dany, 1407
Müller, Dennis, 395
Müller, Lydia, 1665, 1675
Müller, Meinard, 785, 805, 873, 883
Münzberg, Alexander, 1515
Müssig, Daniel, 1121
Mutschke, Peter, 909, 919

N

Nahrstedt, Bastian, 265
Nake, Leonard, 1895
Naumann, Stefan, 1277, 1289, 1473,
1503
Nehring, Jan, 377
Neubig, Stefan, 1335
Neumann, Frank, 197, 353
Neumuth, Thomas, 919, 925
Neuroth, Heike, 95
Nguyen, Thi, 1591
Niedermeier, Michael, 1633
Niehaves, Björn, 1097
Niehus, Alexander, 1515
Nitze, André, 1739

Noori, Faryal, 1633
Nowotka, Dirk, 73
Nsonga, Baldwin, 1683

O

Oberländer, Lea, 73, 81
Obermaier, Frederik, 149
Odebrecht, Carolin, 341
Orlov, Denis, 1575
Osborne, Tobias, 1129
Otten, Lisa, 977
Otto, Wolfgang, 931
Ötvös, Bettina, 371

P

Palombo, Raphael, 1097
Pancratz, Nils, 403, 771
Panitz, Michael, 129
Pauken, Cedric, 1071
Peachkah, Danooosh, 1973
Pelz, Peter, 137
Pentzold, Christian, 893
Perakis, Lymperis, 1833
Peters, Isabella, 777
Pfuhl, Helen, 73
Pidun, Tim, 1057
Pistorius, Elena, 611
Plattfaut, Ralf, 1923
Plecher, David A., 313, 325
Pleger, Michael, 1245
Plomin, Jana, 1049
Podworny, Susanne, 389, 453
Ponciano, Claire, 57
Ponciano, Jean-Jacques, 57
Ponzetto, Simone, 81
Poser, Nick, 1129
Povalej, Roman, 623
Prechel, Kerstin, 457

Prieß, Malte, 457
Putzke, Johannes, 117

R

Rache, Alessa, 1725
Rahm, Erhard, 549
Rampelt, Florian, 425
Ramstetter, J., 1623
Rauch, Sebastian, 1181
Raupach, Maximilian, 1035
Rebholz, Dominik, 1335
Rebstadt, Jonas, 561
Reers, Volker, 1135
Rehm, Georg, 909, 919, 925, 931
Reichwald, Julian, 1725
Reinkensmeier, Jan, 937
Reinosch, Nils, 1515
Rettig, Rasmus, 185
Reuter, Christian, 1529, 1563, 1575
Richter, Manuela, 117
Richter, Sabine, 611
Riebel, Jasmin Antonia, 49
Riedel, Tanja, 1207
Rodner, Erik, 353
Roeder, Torsten, 851
Romeike, Ralf, 447
Rothe, Felix, 597
Rother, Alinka, 1049
Ruhi, Zurana Mehrin, 1809
Runne, Miriam, 1615

S

Sack, Harald, 909
Sackmann, Stefan, 1895
Schaefer, Ina, 1129
Schaffner, Stefan, 1607
Schaffert, Markus, 57
Schauer, Kai, 299

Schenkendorf, René, 461
Schiering, Ina, 1245
Schildein, Sabine, 623
Schiller, Carina, 129
Schiller, Fabian, 795
Schimmer, Thomas, 117
Schimmler, Sonja, 901, 905, 909, 919,
925
Schindler, Josef, 1995, 2051
Schleiss, Johannes, 381, 425
Schlichte, Uta Katja, 1029
Schlick, Sabine, 1029
Schlosser, Lena, 967
Schmailzl, Anton, 1633
Schmelzer, Robert, 1323
Schmid, Thomas, 1645
Schmidt, Sebastian, 1483
Schmidt, Thomas, 795
Schmitz, Andreas, 1071, 1085
Schmolenzky, Pascal, 883
Schneemann, Carsten, 95
Schneider, Daniel, 919, 925
Schneider, Kerstin, 461
Scholl, Margit, 165
Schönthier, Ludwig, 1057
Schoormann, Thorsten, 1303, 1323
Schröder-Puls, Lisa, 1029
Schuktomow, Regina, 165
Schüller, Katharina, 425
Schulte, Carsten, 453
Schultenkämper, Sergej, 1815
Schulz, Konstantin, 391
Schulze, Martin, 1535
Schulze, Paul, 1535
Schumacher, Marcel, 1375, 1399,
1421
Schwär, Simon, 873
Schwarz, Lucas, 777

Schweihoff, Julia, 1309
Schweikart, Maximilian, 1129
Schwerin, Reinhold von, 1751
Schwierzy, Julian, 1483
Scorna, Ulrike, 431
Sebrak, Sebastian, 1851
Seebacher, Uwe, 2015
Seegerer, Stefan, 447
Seipolt, Arne, 1375, 1399, 1421
Selmanagić, André, 363
Selzer, Annika, 467, 469, 705
Semmann, Martin, 123
Seuring, Liv, 1515
Siebel, Nils T., 277
Siegert, Ingo, 377, 381
Siewert, Hellen, 1221
Simecek, Tim, 325
Simic, Dejan, 1963
Simmet, Christan, 1535
Sirvend, Bastian, 1827
Söllner, Matthias, 357, 385
Sorge, Christoph, 675
Spiecker, Claus, 129
Spiekermann, Daniel, 607
Srivastava, Rajiv, 1515
Stahlkopf, Alexander, 1827
Stecker, Benjamin, 1815
StefanDietze, 919
Stein, Anthony, 1653
Stein, Hannah, 1255, 1809
Steinbichl, Stefanie Nina, 1149
Steinert, Mirijam, 635
Steinmann, Lena, 73
Stocker, Markus, 925
Stoll, Patrick, 957
Stolzenburg, Frieder, 461
Streck, Thilo, 1515
Stromeyer, Luise, 231

Stroscher, Jan-Philipp, 715
Stuckenschmidt, Heiner, 73, 81
Sudeikat, Jan, 2003
Sultanow, Eldar, 1851

T

Taffa, Tilahun A., 919
Thiéé, Lukas-Walter, 1777
Thiel, Jannis, 1267
Thielemann, Pascal, 1057
Thielert, Bonito, 1615
Thiessen, Nadja, 1563
Thomas, Oliver, 561, 1359
ThomasBaar, 265
Thürkow, F., 1623
Tiedemann, Lucas, 1169
Tiemann, Michael, 1983
Timm, Ingo J., 467, 469
Tippelskirch, Hubertus von, 165
Tobolla, Marinho, 1851
Transchel, Fabian, 461
Trapp, Matthias, 1529
Troost, Christian, 1515

U

Udelhofen, Stefan, 1983
Ullrich, André, 1851
Upadhyaya, Sharmila, 909
Usbeck, Ricardo, 919, 925, 931

V

Vahdatnia, Ali, 1973
Venkateswaran, Siddarth, 1591
Vladova, Gergana, 1851
Vo, Gia Minh, 403
Vogel, Dennis, 1879
Vogl, Jonathan, 1493
Volkman, Dirk, 623

W

Wachter, Lukas, 369
Waedt, Karl, 1995, 2041, 2051
Wagner, Maik H., 1793
Walchshäusl, Sebastian, 325
Walgenbach, Franz, 369
Wallsberger, Raphael, 353
Walter1, David, 1351
Wambsganss, Thiemo, 357
Wappler, Markus, 1121
Warnemünde, Sebastian, 1615
Weber, Florian, 357
Weber, Sebastian, 1277, 1289, 1503
Weber, Wiebke, 89
Wegner, Vanessa, 1003
Weidlich, Anke, 937
Weigert, David, 431
Wein, Amanda, 937
Wein, Eva, 1221
Weininger, Florian, 241
Weis, Martin, 1551
Weiß, Christof, 805, 883
Weiß, Katharina, 107
Weltz, Maximilian, 185
Wentzel, Bianca, 919
Westing, Max, 1277, 1289
Wewetzer, David, 1583

Wiegrefe, Daniel, 1675, 1683
Wimmer, Maria A., 1071, 1085
Winkelmann, Stephanie, 1309
Winkler, Till J., 947
Wittmann, Jochen, 1169, 1181
Wohlgemuth, Volker, 295, 1197, 1289
Wolf, Armin, 1599
Wolff, Christian, 795, 821
Wolff, Ian, 129
Wolschewski, Anastasia, 925
Woods, Daniel W., 537
Woywod, Kathrin, 95
Wudka, Björn, 295

Y

Yatagha, Romarick, 2051

Z

Zalkow, Frank, 785
Zapilko, Benjamin, 909
Zengerle, Tobias, 313
Zimmermann, Alexander, 1633
Zivkovic, Marco, 1563
Zloch, Matthäus, 909
Zoll, Michaela, 1169
Zürn, Birgit, 437
Zwanzig, Dorian, 253