



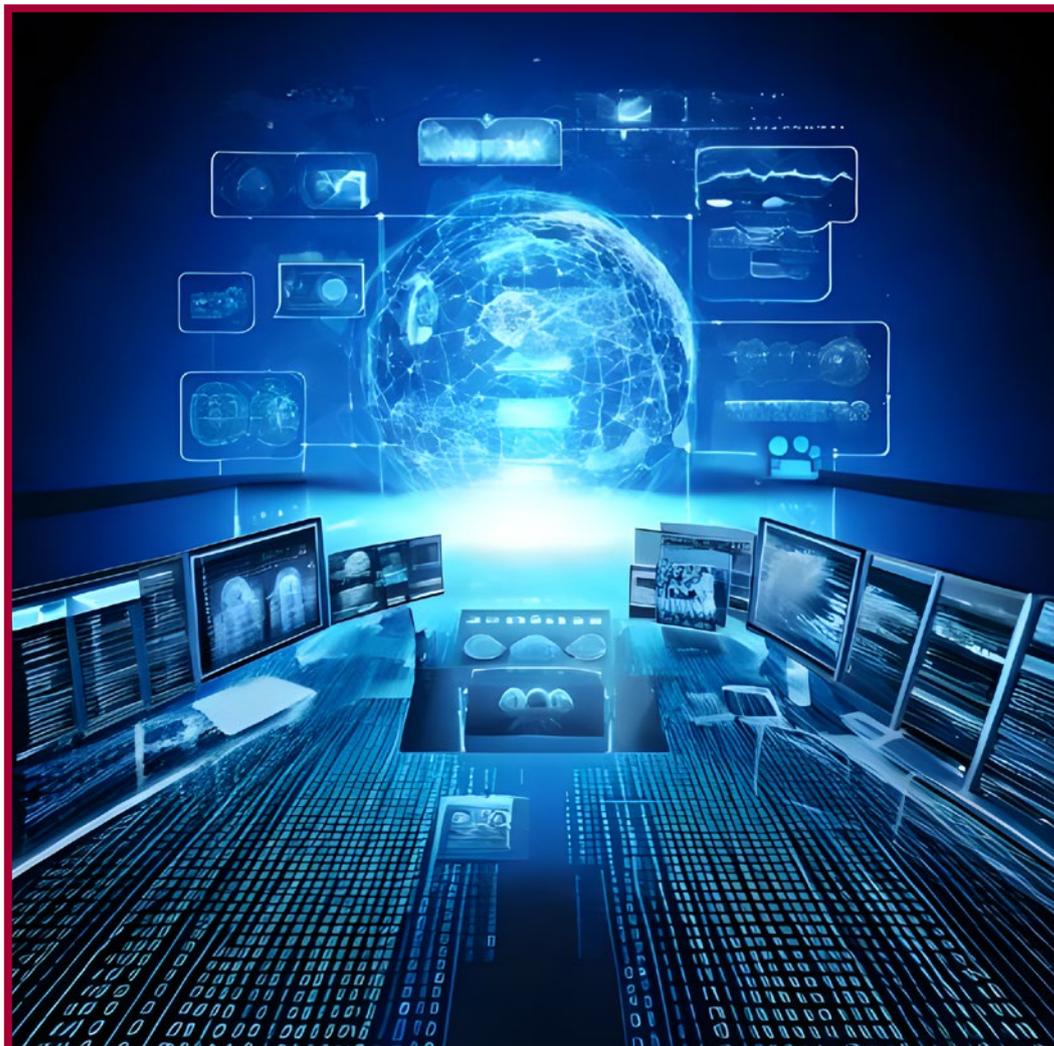
Manuel Schulz, Alexander Fay, Oliver Niggemann, Wenzel Matiaske, Detlef Schulz

# **dtec.bw-Beiträge**

## **der Helmut-Schmidt-Universität Universität der Bundeswehr Hamburg**

Forschungsaktivitäten im Zentrum für Digitalisierungs-  
und Technologieforschung der Bundeswehr dtec.bw

**Band 2 · 2024 · EnDig - KIIPS - KoDiA - OPAL**



Druckversion: ISBN 978-3-86818-329-0  
Digitale Version: ISBN 978-3-86818-317-7

DOI: <https://doi.org/10.24405/16768>



Der Sammelband wird veröffentlicht durch:  
dtec.bw – Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Herausgeber:  
Manuel Schulz, Alexander Fay, Oliver Niggemann, Wenzel Matiaske, Detlef Schulz

Adresse:  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg, Holstenhofweg 85, 22043 Hamburg

Telefonnummer: +49 40 / 6541 - 1  
E-Mail: [dtecbw@hsu-hh.de](mailto:dtecbw@hsu-hh.de)

Gestaltung Einband:  
Michael Bölke, Medienzentrum, Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Bildnachweis Titelseite:  
KI-generierter Inhalt (Adobe Firefly)

# **dtec.bw-Beiträge der Helmut-Schmidt-Universität/ Universität der Bundeswehr Hamburg**

## **Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw Band 2 • 2024**

**mit Beiträgen der Dachprojekte:  
Energie und Digitalisierung (EnDig),  
Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS),  
Kompetenzen für die digitale Arbeitswelt (KoDiA),  
Organisation – Personal – Arbeit – Leadership (OPAL)  
von der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg**

Hamburg, Dezember 2024

Herausgeber:

Prof. Dr. phil. habil. Manuel Schulz  
Prof. Dr.-Ing. Alexander Fay  
Prof. Dr. Oliver Niggemann  
Prof. Dr. Wenzel Matiaske  
Prof. Dr.-Ing. habil. Detlef Schulz

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
Holstenhofweg 85, 22043 Hamburg  
Tel.: +49 40 / 6541 - 1  
E-Mail: [dtecbw@hsu-hh.de](mailto:dtecbw@hsu-hh.de)  
Internet: <https://dtecbw.de>

Redaktionelle Überarbeitung:

Manuel Schulz  
Katharina Kreuzmann



# Inhaltsverzeichnis

## Vorworte

V 1	Vorwort der Herausgeber	VII
V 2	Vorwort des Präsidenten der Helmut-Schmidt-Universität/ Universität der Bundeswehr Hamburg	IX
V 3	Vorwort der Vizepräsidentin für Forschung der Helmut-Schmidt-Universität/ Universität der Bundeswehr Hamburg	XI

## Kapitel I Energie und Digitalisierung (EnDig)

V I	Vorwort zum Dachprojekt Energie und Digitalisierung (EnDig)	2
1.1	Trajektorienplanung und -optimierung für die automatisierte Reparaturanwendung durch robotergeführtes Kaltgasspritzen	3
1.2	A Practical Toolpath Planning Method for Cold Spray Additive Manufacturing of Rotational Symmetry Parts	8
1.3	Temperature and Stress Management in Cold Sprayed Deposits	12
1.4	CoupleIT! Coupling Energy Grids and Research Disciplines	17
1.5	Energy System-Oriented Identification of Hydrogen Storage Supply Scenarios: Method Development and Application within the Digi-HyPro Project	30
1.6	Experimental Development of the Set-Unit Concept on a Laboratory Scale for the H <sub>2</sub> -Chain: Production-Storage-Compression-Utilization (Digi-HyPro)	36
1.7	Dispatch Optimization of the Electricity and Heat of the Smart-Energy-Transform-Unit: A Residential Case Study	42
1.8	Customer-Oriented Sustainability Criteria for Value Chains in Local Production Networks of SMEs	47
1.9	Auf dem Weg in die Kreislaufwirtschaft – Die Bedeutung der Konsument*innen für eine lokale, nachhaltige Wertschöpfung	53
1.10	Leistungselektronik in kritischer Infrastruktur am Beispiel der medizinischen Bildgebung – Untersuchung der Elektromagnetischen Verträglichkeit und Optimierung der Resilienz	59
1.11	Elektrische Lastkraftwagen und Zugmaschinen in Hamburg im Jahr 2050: Hochlauf- und Energiebedarfsprognose	64
1.12	Textiles in the Digital Age: Navigating the Integration of the Digital Product Pass within the Fab City Framework	70
1.13	Die Fab City Hamburg auf dem Weg zu einer Open-Source-basierten Kreislaufwirtschaft: Förderliche und hinderliche Faktoren der Transformation	75
1.14	Progress on Computer-Aided and Experimental Determination of Thermophysical Properties of Gas Mixtures Containing Hydrogen	83
1.15	xbat: A Continuous Benchmarking Tool for HPC Software	91

## Kapitel II Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS)

V II	Vorwort zum Dachprojekt Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS)	99
2.1	Optimierung der Auslagerungsprozesse in automatisierten Kompaktlagern	100
2.2	Drone-Aided Electrical Current Sensing – A Breakthrough for Smart Grids	105
2.3	Ein formales Modell für Anwendungen Künstlicher Intelligenz in Automatisierungssystemen	111
2.4	A Cross-Domain Electromagnetic Compatibility Data Platform for the Digital Development of Automated Vehicles	119
2.5	Synergistic Integration of Physical and Digital Simulation of a Shoulder Exoskeleton	125
2.6	Erfolg braucht Taktik – nicht bloß Masse. Wie die Bundeswehr mit GhostPlay das Defense Metaverse zur Fähigkeitsentwicklung nutzt	130
2.7	Systematische Entwicklung einer Architektur für die virtuelle Inbetriebnahme eines automatisierten Produktionssystems	136
2.8	Production-Related Testing and Optimisation Environment for Hardware and Software in Aircraft Production of Today and Tomorrow	143
2.9	Knowledge Graphs for the Enhancement of Process Planning in Manufacturing	149
2.10	Fortschritte der interdisziplinären Forschung zur Digitalisierung in der Leichtbauproduktion: Anwendungen und Optimierungen in der LaiLa Modellfabrik	155
2.11	UAV-Based Hyperspectral Imaging of Iron(III) on Steel Infrastructure	164
2.12	Echtzeit-Optimierung für die Steuerung von energetisch flexiblen Anlagenverbänden	169
2.13	Chatbot-basierte Interaktion mit Ontologien: Nutzung von Large Language Models und domänenspezifischen Standards	177
2.14	Prävention und Beseitigung von Fehlerursachen im Kontext von unbemannten Fahrzeugen	182
2.15	Ermittlung optimaler Rendezvouspunkte für heterogene Fahrzeugverbände	190
2.16	Dezentrale Informationsverteilung in Verbänden autonomer Roboter	196
2.17	From Micromechanics to Optimal Sensor Positioning in SHM Applications – A Concept Applying a Substructure Technique	206

## Kapitel III Kompetenzen für die digitale Arbeitswelt (KoDiA)

V III	Vorwort zum Dachprojekt Kompetenzen für die digitale Arbeitswelt (KoDiA)	212
3.1	Entwicklung eines digitalen Mathematikförderprogramms für modulare Lerntherapie	213
3.2	hpc.bw: An Evaluation of Short-Term Performance Engineering Projects	223
3.3	Das dtec.bw-Forschungsprojekt „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“ – gestaltungsorientierte Forschung für, in und mit den Streitkräften	229
3.4	Komplexität als Kompetenzanforderung an Subjekte in organisationalen Kontexten unter Bedingungen der Digitalisierung	235

## Kapitel IV Organisation – Personal – Arbeit – Leadership (OPAL)

V IV	Vorwort zum Dachprojekt Organisation – Personal – Arbeit – Leadership (OPAL)	241
4.1	Entwicklung eines offenen Lernangebots zur Vermittlung digitaler Kompetenzen für Studierende – Ein Ergebnis des Projekts DigiTaKS*	242
4.2	Digitale Transformation und hybride Führungsidentitäten in der Sozialen Arbeit	248
4.3	Serious Play als edukativer Ansatz für die Weiterbildung von Fachpersonen	254
4.4	Zu den Funktionen und Folgen von Informationsmängeln in der Organisationsberatung	258
4.5	Working from Home: Opportunities for Transformational and Health-Oriented Leadership & Specific Challenges Arising from Remote and Hybrid Work	262
4.6	Are Videocalls Outdated? A Comparison with Virtual Reality Meetings as a Future Perspective	273
4.7	Hybride Führung meistern: Einblicke in die Onlineplattform für Führungskräfte DigiLAP	280
4.8	Praxisforschung zwischen Anspruch und Wirklichkeit. Über die Umsetzung von Partizipation und Wissenstransfer	286
4.9	Reconfiguring Teachers' Work through Automated Evaluation	291
4.10	Digitalisierung kommunaler Verwaltungen – ein drei Säulen Modell	297
4.11	Representativeness in Business Surveys and Linked Employer Employee Data: Evidence from the SOEP-LEE2 Data for Germany	308

**Vorworte zum Sammelband „dtec.bw-Beiträge der  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg“**

mit Beiträgen von

Vorwort der Herausgeber:

Prof. Dr.-Ing. habil. Detlef Schulz  
Prof. Dr.-Ing. Alexander Fay  
Prof. Dr. Oliver Niggemann  
Prof. Dr. phil. habil. Manuel Schulz  
Prof. Dr. Wenzel Matiaske

Vorwort des Präsidenten  
der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Prof. Dr. Klaus Bertram Beckmann

Vorwort der Vizepräsidentin für Forschung  
der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Prof. Dr. jur. Margarete Schuler-Harms

## Vorwort der Herausgeber zum dtec.bw-Sammelband

Seit der Herausgabe unseres ersten dtec.bw-Sammelbandes im November 2022 sind nunmehr zwei Jahre vergangen, und wir freuen uns, nach dem großen Interesse und der guten Resonanz, der geneigten Leserschaft nun unseren aktuellen zweiten Band vorlegen zu dürfen. In diesen zwei Jahren haben sich einige grundlegende Rahmenbedingungen unserer Forschungsarbeit verändert und weiterentwickelt. Gleichzeitig blicken wir aber auch auf viele neue Erkenntnisse, Ergebnisse, Produkte und Prototypen zurück.

Bezüglich unserer veränderten Rahmenbedingungen möchten wir auf folgende Aspekte noch einmal eingehen:

Der durch den völkerrechtswidrigen Angriff Russlands entstandene „Ukrainekrieg“ dauert bis heute unvermindert an. Auch wenn dieser Krieg bereits im Februar 2022, also noch vor der Herausgabe unseres ersten Sammelbandes, begonnen hat, so sind doch wesentliche inhaltliche „Nachschärfungen“ und Fokusverlagerungen der Forschungsarbeit unter dem Einfluss der sicherheitspolitischen Zeitenwende in vielen Einzelprojekten von dtec.bw in den vergangenen zwei Jahren erfolgt. Im vorliegenden Band werden Sie in einer Reihe von Beiträgen neue bzw. veränderte Schwerpunktsetzungen zur Sicherstellung und Weiterentwicklung der digitalen Souveränität Deutschlands und Europas nachverfolgen können. Auch an unserer Universität konnten die dafür erforderlichen sicheren Rahmenbedingungen parallel ausgebaut und verbessert werden.

Zweitens liegt nunmehr die positive Begutachtung und Evaluation des Deutschen Wissenschaftsrats zu unserer gemeinsamen Forschungsarbeit mit unserer Schwesteruniversität, der Universität der Bundeswehr München, im Rahmen von dtec.bw vor. Wir freuen uns über die insgesamt positive Bewertung und nehmen diese als Bestätigung aber auch als Ansporn für unsere zukünftige Arbeit an.

Drittens wurde die Förderung von dtec.bw zwischenzeitlich in den Deutschen Aufbau- und Resilienzplan (DARP) übernommen und damit die Voraussetzungen für eine Gesamtprojektverlängerung für die Jahre 2025/2026 ermöglicht. Die Projektleitenden der Einzelprojekte waren aufgefordert, sich jeweils an beiden Universitäten im Wettbewerb einem Antragsverfahren zu stellen, das inzwischen abgeschlossen ist und nunmehr die Projektverlängerungen ermöglicht.

Viertens hat im September 2024 die erste dtec.bw-Jahrestagung an der Universität der Bundeswehr München stattgefunden, in der alle dtec.bw-Einzelprojekte die Chance erhalten haben, sich einer interessierten Öffentlichkeit innerhalb wie außerhalb der Bundeswehr mit Vorträgen, Posterpräsentationen und Informationsständen im Rahmen einer Fachausstellung zu stellen. Die große Resonanz und die vielfältigen Interessen aktueller und zukünftiger Stakeholder haben uns in unserer Forschungsarbeit bestärkt, uns aber auch neue Impulse gegeben.

Und darüber hinaus haben sich beide Universitäten (fünftens) im Frühjahr 2023 auf eine gemeinsame Forschungsstrategie für das dtec.bw unter Beteiligung ihrer akademischen Selbstverwaltung geeinigt, die sich an einem weiten Verständnis des Konzeptes von „Human Security“ orientiert und allen Beteiligten ein breites aber durchaus zielgerichtetes Forschungsverständnis ermöglicht.

Und letztlich ist sechstens auch noch auf eine wichtige personelle Veränderung bei den Dachprojektleitenden einzugehen. Nach vielen erfolgreichen Jahren an unserer Universität in der Fakultät Maschinenbau ist unser Kollege Alexander Fay 2024 einem Ruf an die Ruhr-Universität Bochum gefolgt, um dort den Lehrstuhl für Automatisierungstechnik zu übernehmen. Er hat maßgeblich die Forschungsarbeit in unserem Dachprojekt „Künstliche Intelligenz und intelligente physische Systeme“ (KIIPS) vorangetrieben und geprägt. Dafür, aber auch für seine stets kollegiale wie hilfreiche Zusammenarbeit, seine kreativen Anregungen und innovativen Ideen möchten wir verbliebenden Dachprojektleitenden ihm an dieser Stelle besonders danken. Prof. Dr. Oliver Niggemann, Inhaber der Professur für Informatik im Maschinenbau seit 2019 an unserer Universität, hat sich bereit erklärt, die Aufgabe des Dachprojektleitenden KIIPS zu übernehmen.

Die bisherige Strukturierung der Forschungsarbeit im Anteil der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg am dtec.bw in vier thematisch aufgeteilte Dachprojekte hat sich somit bewährt und wird nunmehr auch für die nachgeschärfte und verlängerte Forschungsarbeit aufrechterhalten:

- Energie und Digitalisierung (EnDig),
- Künstliche Intelligenz und Intelligente physische Systeme (KIIPS),
- Kompetenzen für die digitale Arbeitswelt (KoDiA),
- Organisation – Personal – Arbeit – Leadership (OPAL).

Diese Dachprojekte fokussieren weiterhin in einem komplementären Verständnis die besonderen Aspekte von Energie, künstlicher Intelligenz, menschlicher Kompetenzen sowie Organisation von Arbeit und Führung. Innerhalb dieser Dachprojekte gibt es nach wie vor eine Vielzahl von Einzelprojekten zu unterschiedlichen Themen.

Der vorliegende zweite Band ermöglicht den Leserinnen und Lesern erneut, sich über die aktuellen Zwischenergebnisse zu informieren, und schafft Transparenz unserer Forschungsarbeit.

Die Themenauswahl wurde den Autoren und Autorinnen freigestellt, um auch im zweiten Band den unterschiedlichen fachlichen und organisatorischen Rahmenbedingungen sowie den domänenspezifischen Fächerkulturen angemessen Rechnung zu tragen.

Wir bedanken uns sehr herzlich bei allen beteiligten Autorinnen und Autoren und allen weiteren Beteiligten der Einzelprojekte für ihre engagierte Arbeit an diesem Sammelband.

Besonderer Dank gilt Frau Katharina Kreuzmann, wissenschaftliche Mitarbeiterin im Zentrum für technologiegestützte Bildung, für ihr Lektorat und ihre vielfältige organisatorische Unterstützung, Herrn Michael Bölke aus dem Grafikstudio unserer Universität für seine kreative Unterstützung bei der Gestaltung des Bandes, sowie allen beteiligten Mitarbeitenden unserer Universitätsbibliothek für ihre begleitende Unterstützung bei der Veröffentlichung.

Dieser Band wird sowohl als Druckversion als auch online über die Bibliothek der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg verfügbar sein:

<https://ub.hsu-hh.de/>.

Wir wünschen Ihnen viel Freude beim Lesen und freuen uns auf Ihre Kommentare und Rückmeldungen!

Hamburg, im November 2024

Detlef Schulz, Alexander Fay, Oliver Niggemann, Wenzel Matiaske, manuel schulz

## Vorwort des Präsidenten der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Im Juni 2020 kam die Nachricht, dass die beiden Universitäten der Bundeswehr ab dem August zu gleichen Teilen 500 Mio. € für vier Jahre erhalten sollten, um die coronabedingt eingebrochene Konjunktur zu stützen. Dabei sei die Stärkung der digitalen Souveränität Deutschlands voranzutreiben und gemeinsam mit der Industrie zu forschen, aber unter inhaltlicher und organisatorischer Führung der Universitäten. Als Finanzwissenschaftler und Universitätspräsident musste ich die Hände über dem Kopf zusammenschlagen: Konjunkturförderung durch universitäre Forschung über vier Jahre? Dazu kamen die vergaberechtlichen Herausforderungen sowie der damit verbundene hohe administrative Aufwand – dies mit Personal, das normalerweise ein Mittelvolumen im zweistelligen Millionenbereich jährlich verwaltet.

Und so machten wir uns mit Nachdruck ans Werk. Dabei ergab sich schnell, dass die inhaltliche Leitidee der digitalen Souveränität vorzüglich geeignet war, dem Projekt eine politisch wie wissenschaftlich sinnvolle Richtung zu geben. Der Überfall Russlands auf die Ukraine und das öffentliche Sichtbarwerden der Zeitenwende ließen die Entscheidung für dtec.bw nachgerade als prophetisch erscheinen. Und wir an den Universitäten der Bundeswehr erkannten, dass wir die erheblichen Herausforderungen meistern und daran wachsen konnten.

An der Helmut-Schmidt-Universität kam es uns auf drei Dinge an: Erstens, dass die Forschungsprojekte an den Stärken unseres Hauses ansetzen müssen, diese weiter befördern sollen und Interdisziplinarität groß zu schreiben haben. Zweitens auf eine wettbewerbliche Auswahl unter den Vorschlägen aus unserem Hause. Und drittens auf Compliance. Die ersten beiden Grundsätze führten rasch zu der Gründung unserer vier Dachprojekte, deren Leitungen eine Schlüsselrolle bei der Gestaltung unseres Teils von dtec.bw gespielt haben und spielen. Ein Dachprojekt kümmert sich um Digitalisierung, Automatisierung und künstliche Intelligenz, das zweite um nachhaltige Energieversorgung, das dritte um Digitalisierung in der Arbeitswelt und das vierte um Digitalisierung in der (Aus-)Bildung, hier unter besonderer Berücksichtigung der Bundeswehr. Darunter versammeln sich jeweils eine große Zahl innovativer Projekte, die miteinander im Wettbewerb stehen.

Ein herzlicher Dank gebührt an dieser Stelle den Gefährten der ersten Stunde, meinem damaligen Vizepräsidenten für Forschung Prof. Rolf Lammering, den vier Dachprojektleiter, Prof. Alexander Fay, Prof. Detlef Schulz, Prof. Wenzel Matiaske und Prof. Manuel Schulz, und der Verwaltung einschließlich ihres Kanzlers Axel Puckhaber.

Vor kurzem konnten die beiden Universitäten der Bundeswehr beim ersten dtec.bw-Jahreskongress auf dem Campus in Neubiberg zeigen, was aus diesen Anfängen geworden ist. Ich kannte vieles schon, war aber gleichwohl mehr als nur beeindruckt. Der vorliegende Band gesellt sich neben solche Leistungsschauen und die zahlreichen Vorträge, ganz im Sinne Vergils: „exegi monumentum aere perennius“. Denn nur wer schreibt und publiziert, der bleibt.

Es sei denn, eine Publikation ist aufgrund von Geheimhaltungserfordernissen ausgeschlossen. dtec.bw entwickelt sich zusehends in Richtung Dual-Use, und das ist erstens folgerichtig und zweitens seitens der Mittelgeber zu Recht erwünscht. Denn nur so kann ein wissenschaftlicher Beitrag zur Bewältigung der “Zeitenwende” sowie der neuen Aufgaben im Rahmen einer ressortübergreifenden Gesamtverteidigung und einem gesellschaftlichen Streben nach Human Security erzielt werden. Daraus mögen in einigen Fällen hinzunehmende Beschränkungen für Publikationen erwachsen. Als jemand, der Mitverantwortung für den Schutz der Wissenschaft trägt, muss ich allerdings darauf hinweisen, dass offene Kritik der Peers und die Schwarmintelligenz der Community die Kreativität und die Qualität von Wissenschaft erst ermöglichen.

Mittlerweile wurde dtec.bw um weitere zwei Jahre verlängert, was für jede der beiden Bundesuniversitäten noch einmal vier Fünftel der Förderung bedeutet, komprimiert auf zwei Jahre (2025-2026). Das ist erfreulich, zumal sich dadurch erste nachhaltige Entwicklungen durch dtec.bw zeigen. An der HSU/UniBw H haben wir einen Profilbildungsprozess begonnen, der sicherlich zum Teil von dtec.bw geprägt sein wird. Schon zuvor hatte der Akademische Senat im Zusammenhang mit dem dtec.bw und dann auf seiner Klausurtagung im Sommer 2023 hochschulstrategische Grundsätze einstimmig beschlossen, welche die Herausforderungen der “Zeitenwende” für unsere Universität mitbedenken. Strategisch sind wir besonders für die nähere Zukunft gut aufgestellt.

Der besondere Dank für die Verlängerung gebührt aber den Forschenden: Sowohl das Gutachten des Wissenschaftsrates, welches auf Vorschlag der beiden Universitätspräsidentinnen zum dtec.bw eingeholt wurde, als auch die Anerkennung der Peers in der wissenschaftlichen Gemeinschaft sprechen eine deutliche Sprache: Das dtec.bw ist als eine Veranstaltung von hoher Qualität akzeptiert, und nur deshalb kann es auch verlängert werden. Ich möchte die Gelegenheit ergreifen, allen Forscherinnen und Forschern in der dtec.bw-Gemeinschaft für diesen herausragenden Erfolg zu gratulieren.

Wir greifen jetzt nach der Verstetigung des dtec.bw. Wenn dieses Vorhaben gelingt, werden wir nicht über die gleichen Fördersummen reden können wie bisher, und auch die Governance ist noch unklar. Immerhin werden wir von unserer Seite mit sechs erfolgreichen Jahren dtec.bw ein sehr gutes Fundament gelegt haben. Worin dieses Fundament wissenschaftlich besteht und welche Kooperationen, Netzwerke und Ausgründungen auf dieser Grundlage erfolgt sind, das können Sie dem nachfolgenden Band entnehmen, welchen ich Ihrer besonderen Aufmerksamkeit empfehle.

Hamburg, im November 2024

Prof. Dr. Klaus Bertram Beckmann

Präsident der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

## Vorwort der Vizepräsidentin für Forschung der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Mit großer Freude begrüße ich den zweiten Sammelband zum dtec.bw mit Abhandlungen aus den 45 Projekten an der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg. Kolleginnen und Kollegen der Universität der Bundeswehr München sind an 16 dieser Projekte und damit auch an den Beiträgen aus diesen Projekten im vorliegenden Band beteiligt. Die in diesem Band versammelten Beiträge präsentieren den Fortschritt der Projekte im Digitalisierungs- und Technologiezentrum der Bundeswehr (dtec.bw) seit dem Erscheinen des ersten Sammelbandes Ende 2022. Erneut werden diese Beiträge unter die vier in Hamburg gebildeten Forschungsschwerpunkte („Dachprojekte“) eingeordnet und damit thematisch etwas sortiert.

Seit dem Erscheinen des ersten Sammelbandes hat das dtec.bw sich erfolgreich weiterentwickelt. Eine Stellungnahme des Wissenschaftsrats vom Juli 2023 bescheinigt dem dtec.bw und damit auch unseren Projekten, gut und teils auch herausragend zu arbeiten. Die Hinweise des Wissenschaftsrats zu Verbesserungen des dtec.bw im Rahmen einer Verlängerung und ggf. auch Verstetigung nehmen wir auf und sehen sie als Ansporn, dieses besondere Zentrum zu optimieren.

Am 16./17. September 2024 fand die erste Jahrestagung dtec.bw an der Universität der Bundeswehr in München-Neubiberg statt. Unter dem Motto „Forschung mit Mehrwert für alle Dimensionen“ richtete sich diese Tagung ausdrücklich an Vertreter und Vertreterinnen aus dem BMVg und der Bundeswehr. Die Gelegenheit, sich dort zu präsentieren, nahmen auch viele der Hamburger Projekte wahr. Als Mitglied der wissenschaftlichen Leitung geht mein herzlicher Dank für die gelungene Umsetzung eines interessanten Konzepts an die Geschäftsstelle des dtec.bw sowie – für das „Hosting“ auch der Hamburger Projekte – an die Präsidentin der UniBw München, Frau Professorin Dr.-Ing. Eva-Maria Kern, sowie den Direktor und Sprecher des dtec.bw und meinen Kollegen, Herrn Vizepräsidenten Professor Dr.-Ing. Geralt Siebert. Die zweite Jahrestagung im kommenden Jahr wird in der ersten Septemberwoche an der HSU/UniBw in Hamburg stattfinden.

Im Juni 2024 bewilligte der Bundesminister der Verteidigung die Verlängerung des dtec.bw an beiden Universitäten der Bundeswehr bis Ende des Jahres 2026. Ab 2025 werden einige der Hamburger Projekte zusammengeführt, sodass formal 43 Projekte aus Hamburg in die Verlängerung gehen, aber alle bisherigen Projekte inhaltlich fortgeführt werden. Einige Neuerungen wird es geben, vor allem die, dass alle Projekte ab 2025 (auch) einen Mehrwert für die Bundeswehr erbringen sollen. Diese Zielsetzung wird für ein verstetigtes dtec.bw noch an Bedeutung gewinnen und auch seine Governance deutlicher als bisher prägen. Bewährt hat es sich, die Forschung zu Digitalisierung und Technologie in der und für die Bundeswehr auch nach der Verstetigung den beiden Bundeswehruniversitäten anzuvertrauen. Hierin hat uns auch die Stellungnahme des Wissenschaftsrats bestärkt. In der Verantwortung der beiden Universitäten wird es liegen, die im dtec.bw liegenden Chancen und Impulse für die Ertüchtigung einer qualitativ hochwertigen Forschung vor allem, aber nicht nur auf dem Gebiet der Sicherheit und Verteidigung zu nutzen und die an den beiden Universitäten bestehenden Schätze zu heben. Welche Potentiale sich hierbei bieten, dokumentiert nicht zuletzt auch der vorliegende Sammelband.

Der Band erscheint zu einem Zeitpunkt, in dem die erste, bis Ende 2024 konzipierte Phase des dtec.bw ihren Abschluss findet. Herrn Oberst i.G. Professor Dr. Manuel Schulz, einem der vier Dachprojektleiter, und seinem Team des Zentrums für technologiegestützte Bildung danke ich sehr herzlich für die kompetente Betreuung dieses zweiten Sammelbandes. Ebenso danke ich den vier Dachprojektleitern für Ihre Einführungen in die jeweiligen Abschnitte sowie besonders herzlich allen Autorinnen und Autoren für ihre Berichte aus ihren jeweiligen Forschungsprojekten.

Hamburg, im November 2024

Prof. Dr. jur. Margarete Schuler-Harms

Vizepräsidentin für Forschung an der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg und wissenschaftliche Leiterin der Projekte des dtec.bw an der HSU/UniBw H.

# **Kapitel I**

## **Energie und Digitalisierung (EnDig)**

mit Beiträgen von

Deutsche Elektronen-Synchrotron DESY, Universität Hamburg

DLR - Institut für Vernetzte Energiesysteme

Eurotechnica GmbH

HafenCity Universität Hamburg

Helmholtz-Zentrum Hereon

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Hochschule Bochum

House of All UG

Karlsruhe Institute of Technology

Megware Computer Vertrieb & Service GmbH

Physikalisch-Technische Bundesanstalt Berlin

Ruhr-Universität Bochum

Technische Universität Hamburg

Universität der Bundeswehr München

Universität Hamburg

## Vorwort zum dtec.bw-Dachprojekt „Energie und Digitalisierung (EnDig)“

In den zwei Jahren seit dem ersten Band zu unseren dtec.bw-Projekten und 4,5 Jahren seit Projektbeginn wurde uns sehr deutlich vor Augen geführt, wie stark das Thema der Energieversorgung unsere Gesellschaft bewegen, vereinen und leider auch spalten kann. Dabei hat sich gezeigt, dass auch gut gemeinte Ideen und Strategien einer breiten und tiefgehenden Kommunikation bedürfen. Wir als Wissenschaftler können durch eine faktenbasierte Argumentation dazu beitragen. Deshalb finde ich es sehr erfreulich, dass nun dieser Band 2 zur Beschreibung, Diskussion und fachlichen Bewertung unserer Projektergebnisse erscheint.

Wir benötigen technische, organisatorische, ökonomische und rechtliche Lösungsansätze für eine klimaneutrale, bezahlbare und gesellschaftlich akzeptierte Energieversorgung in allen denkbaren Anwendungsbereichen. Die Vernetzung der Fachdisziplinen wird in den hier vertretenen Projekten in großen Teilen durch die Zusammensetzung der Projektteams erreicht, wobei insbesondere überregionale Wissenschafts- und Industriekooperationen den fachlichen Horizont erweitern und oft auch der Schlüssel zu neuartigen gemeinsamen Lösungsansätzen sein können. Viele dieser Lösungsansätze, beispielsweise die zukünftige Kopplung unserer Strom-, Gas- und Wärmenetze, lassen sich nur durch breite Digitalisierungsansätze umsetzen.

Unser Dachprojekt „Energie und Digitalisierung“ wirkt in den Zielbereichen Wasserstoff, Smart Grid, Klimaschutz, Erneuerbare Energien, klimaneutrale Luftfahrt, Hafenlogistik, maritime Technologien und digitale Produktion. Durch unsere gemeinsame und zielgerichtete Forschung kann die Wettbewerbsfähigkeit der deutschen Industrie durch Projekte zur Digitalisierung im Energie- und Produktionssektor gestärkt werden. Die schon lange diskutierte Wasserstoffwirtschaft mit ihren Möglichkeiten zur Sektorenkopplung und Energiespeicherung, die Umstellung auf „grüne“ Energieerzeugung und die Elektromobilität benötigen längere Anlaufzeiten als erwartet, auch hier können passende interdisziplinäre Forschungsansätze beschleunigend wirken.

Alle hier beteiligten Wissenschaftler tragen gemeinsam dazu bei, dass innovative Ideen, Konzepte, Prototypen und Demonstratoren zur klimaneutralen und bezahlbaren Energieversorgung unserer Gesellschaft gefunden, erprobt und später in der Breite angewendet werden.

Die vorliegenden Beiträge wurden im Zeitraum Februar bis Oktober 2024 erstellt, die vielen hier präsentierten Ergebnisse sind überzeugende Belege für die Wirksamkeit der gewählten wissenschaftlichen Themencluster-Ansätze in den Dachprojekten.

Viele der geplanten Vorhaben trafen in der Umsetzungsphase auf vielfältige Herausforderungen in den Bereichen Personalgewinnung, Beschaffungen und Vertragsgestaltung mit Kooperationspartnern. Dank der kontinuierlichen Unterstützung unserer Hochschulleitung und unserer Verwaltung konnten nahezu alle prozessualen Schwierigkeiten überwunden werden. Dafür bedanke ich mich hier im Namen aller Wissenschaftler des Dachprojekts herzlich!

Besonders erfreulich ist, dass die zukünftige Projektverlängerung um weitere zwei Jahre bereits genehmigt ist. Somit können die Forschungsthemen weiter vertieft werden, nicht zuletzt wurde die Forschungsperspektive vieler junger Wissenschaftler gesichert und erweitert.

Ich möchte mich bei allen beteiligten Wissenschafts-Autoren des Dachprojekts "Energie und Digitalisierung" herzlich dafür bedanken, dass Sie neben den vielen anderen Aufgaben in ihren laufenden Projekten hier Ihre spannenden Forschungsergebnisse präsentieren!

Dieser Band wird neben der Veröffentlichung als Druckversion auch online über die Bibliothek der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg zur Verfügung gestellt:

<https://ub.hsu-hh.de/>

Nun wünsche ich Ihnen viel Freude beim Lesen und freue mich auf die kommenden Diskussionen, auch in den kommenden zwei Jahren der Verlängerungsphase, denn die Zielstellungen in unserem Dachprojekt „Energie und Digitalisierung“ bleiben sehr vielfältig und können nur gemeinsam erreicht werden.

Hamburg, im November 2024

Detlef Schulz

Dachprojektleiter „Energie und Digitalisierung“

# Trajektorienplanung und -optimierung für die automatisierte Reparaturanwendung durch robotergeführtes Kaltgasspritzen

Marcel Lewke\*, Felix Gehlhoff  
 Institut für Automatisierungstechnik  
 Helmut-Schmidt-Universität/  
 Universität der Bundeswehr Hamburg  
 Hamburg, Deutschland  
 \*marcel.lewke@hsu-hh.de

Hongjian Wu\*, Alexander List, Frank  
 Gärtner, Thomas Klassen  
 Institut für Werkstofftechnik  
 Helmut-Schmidt-Universität/  
 Universität der Bundeswehr Hamburg  
 Hamburg, Deutschland  
 \*wuh@hsu-hh.de

Alexander Fay  
 Lehrstuhl für Automatisierungstechnik  
 Ruhr-Universität Bochum  
 Bochum, Deutschland  
 alexander.fay@rub.de

**Kurzfassung** – Das Kaltgasspritzen entwickelt sich zu einem Verfahren mit großem Potenzial für die Reparatur metallischer Bauteile, insbesondere für das Aufbringen von hitze- und oxidationsempfindlichen Materialien. In diesem Zusammenhang ermöglicht der Einsatz von Automatisierung und Robotik eine flexible Steuerung des Reparaturprozesses. Um einen optimalen Reparaturprozess zu gewährleisten, müssen die verschiedenen Anforderungen des robotergeführten Kaltgasspritzens bereits in der simulativen Planungsphase berücksichtigt werden. Herkömmliche Trajektorien zum Materialauftrag berücksichtigen jedoch oft nicht die bei Reparaturen zu beachtenden geometrischen Randbedingungen des Materialaufbaus, den effizienten Materialeinsatz und die zugrundeliegenden Einschränkungen der Roboterkinematik. In dieser Arbeit wird daher ein Konzept zur automatisierten Trajektorienplanung und anschließenden Trajektorienoptimierung zur Reparatur durch robotergeführtes Kaltgasspritzen beschrieben. Das Ziel ist es, eine optimierte Trajektorie zu erzeugen, die die Anforderungen des Kaltgasspritzens und der Roboterkinematik berücksichtigt, um eine qualitativ hochwertige Reparatur und einen effizienten Materialeinsatz zu gewährleisten. Dazu gehören die Minimierung des überschüssigen Materials und die Minimierung des Rucks bei der Roboterbewegung. Die Ergebnisse zeigen die erfolgreiche Anwendung der initialen Trajektorienplanung und der anschließenden Trajektorienoptimierung für die Bauteilreparatur durch Kaltgasspritzen.

**Stichworte** – Kaltgasspritzen, Reparatur, Robotik, Trajektorie, Optimierung

## I. EINLEITUNG

Der weltweite Ressourcenverbrauch und die Auswirkungen auf die Umwelt zählen zu den großen Herausforderungen unserer Zeit. Ressourceneffiziente Verfahren zur Herstellung und Reparatur von Bauteilen sind damit unerlässlich, um diese Herausforderungen zu bewältigen. Der Einsatz additiver Fertigungsverfahren ermöglicht dabei neue Wege der Reparatur, um lokal beschädigte Bereiche von Bauteilen mit minimalem Ressourceneinsatz wiederherzustellen. Das Verfahren des Kaltgasspritzens (KGS) erweist sich hierbei als vorteilhaft für Reparaturanwendungen. KGS ist ein Beschichtungsverfahren, bei dem pulverförmiges Material in einem erhitzten Gasstrom durch eine Lavaldüse auf Überschallgeschwindigkeit be-

schleunigt wird. Der Partikelauflaufprall auf die Substratoberfläche erfolgt im festen Zustand, d. h. die Temperatur des Materials bleibt dabei unterhalb der Schmelztemperatur des Pulvers. Dies bietet bei dem Auftrag wärme- und oxidationsempfindlicher Materialien erhebliche Vorteile gegenüber thermischen Verfahren [1]. Der geringe thermische Einfluss zeichnet KGS als geeignetes Verfahren für Reparaturanwendungen und zur additiven Fertigung aus, wie in verschiedenen Übersichtsstudien hervorgehoben wird [2], [3], [4]. Für die Reparatur komplexer Geometrien werden Industrieroboter benötigt, was im Folgenden als *robotergeführtes KGS* bezeichnet wird. Dabei ist die Handhabung des Spritzstrahls ein wesentlicher Aspekt des KGS, da die Pulverzufuhr nicht abrupt gestoppt oder dynamisch gesteuert werden kann [5]. Darüber hinaus müssen, insbesondere bei hochfesten Werkstoffen, alle Prozessparameter hinreichend erfüllt werden, um die notwendigen Materialeigenschaften durch KGS zu gewährleisten. Dazu gehören sowohl die primären Prozessparameter wie Prozessgas-temperatur und -druck als auch die sekundären Prozessparameter, die durch den handhabenden Industrieroboter gesteuert werden. Die sekundären Prozessparameter, die im Folgenden nach der Norm EN ISO 14917 [6] benannt sind, umfassen den Spritzwinkel, den Spritzabstand, die Spritzgeschwindigkeit und die Spritzstrahlüberdeckung (im Folgenden als Spritzspurabstand bezeichnet), welche entscheidend sind für einen erfolgreichen Materialauftrag [4]. Diese sekundären Prozessparameter beschreiben die Bewegung der KGS-Düse relativ zum Substrat und wirken sich direkt auf die Auftragseigenschaften aus. Für einen kontrollierten Materialauftrag ist somit eine systematische Trajektorienplanung unter Berücksichtigung der spezifischen Randbedingungen des KGS erforderlich. Um die verschiedenen Anforderungen und Einschränkungen des KGS effektiv zu bewältigen und schließlich eine optimale Trajektorie für die Reparatur zu generieren, ist der Einsatz von mathematischer Optimierung erforderlich.

In dieser Arbeit wird ein Konzept zur initialen Trajektorienplanung und zur Trajektorienoptimierung durch mathematische Optimierung für die Reparatur mittels robotergeführten KGS vorgestellt. Die Initialtrajektorie liefert eine akzeptable Startlösung und dient als Referenz für die nachfolgende Trajektorienoptimierung.

Das Konzept beginnt mit der Extraktion des zu füllenden Volumens durch den Vergleich von Soll- und Ist-Bauteil. Zur Erzeugung des extrahierten Volumens wird dieses in geeignete, adaptiv gekrümmte Schichten unterteilt und in Punktwolken für die Pfadplanung umgewandelt. Der Pfad wird dann in eine Trajektorie umgewandelt, indem eine geeignete Geschwindigkeitsverteilung für die Spritzgeschwindigkeit hinzugefügt wird, um die erforderliche lokal variierende Schichtdicke zu erzeugen. Nach dieser initialen Trajektorienplanung werden für die anschließende Trajektorienoptimierung Entscheidungsvariablen, Zielfunktion, Rand- und Nebenbedingungen und ein Materialauftragsmodell formuliert, um die Menge des aufgetragenen Materials entsprechend zu steuern.

Diese Arbeit gibt einen Überblick über die Arbeiten aus [7] und [8]. Der Abschnitt II stellt relevante Arbeiten aus der Literatur vor. Das Konzept zur initialen Trajektorienplanung und die Formulierung des Optimierungsproblems werden in Abschnitt III erläutert. Die Anwendung des Konzepts für einen Anwendungsfall und die dabei erzielten Ergebnisse werden in Abschnitt IV vorgestellt. Schließlich schließt Abschnitt V diese Arbeit mit einer Zusammenfassung ab und gibt einen Ausblick über zukünftige Forschungsaktivitäten.

## II. ÜBERSICHT ZU RELEVANTEN ARBEITEN

Für Anwendungen in der KGS-Reparatur schlugen die Autoren Chen et al. [9] eine adaptive Spiraltrajektorie vor. Es wird eine archimedische Spirale verwendet, um einen konstanten Abstand zwischen zwei benachbarten Spritzspuren zu gewährleisten. Die Trajektorie setzt sich aus zwei symmetrischen archimedischen Spiralen zusammen. Diese werden durch lineare Transformation an die Schadenskontur angepasst. Für jeden Zielpunkt wird die Spritzgeschwindigkeit an die entsprechende Tiefe der Kratergeometrie angepasst, indem eine ungefähre lineare Korrelation zwischen Spritzgeschwindigkeit und Schichtdicke angewandt wurde. Die Ergebnisse zeigen eine erfolgreiche Reparatur des Bauteils mit ausreichender Qualität hinsichtlich Porosität und Haftfestigkeit. Allerdings weist die vorgeschlagene Spiraltrajektorie Abweichungen vom orthogonalen Spritzwinkel zur lokalen Oberfläche auf. Im gegebenen Beispiel war die Trajektorie für den Anwendungsfall funktional, da die Oberfläche nur kleine Neigungswinkel aufwies. Bei möglicherweise größeren Neigungen muss jedoch der lokale Spritzwinkel berücksichtigt werden. Außerdem wurde die Trajektorie 30 Mal mit einem konstanten Abstand zwischen KGS-Pistole und Bauteil wiederholt, was zu einem stetig abnehmenden effektiven Spritzabstand führt. Es werden keine Angaben zur Berechnung der erforderlichen Anzahl von Wiederholungen gemacht. Um optimale KGS-Bedingungen aufrechtzuerhalten, wäre ein schichtweiser Ansatz erforderlich, der die erforderliche Anzahl von Schichten mit individuell angepasstem Spritzabstand berücksichtigt.

Spiralförmige Trajektorien werden auch von Wu et al. [10] verwendet. Die Autoren schlugen ein Konzept für einen vollständigen Reparaturprozess durch KGS vor. Dieses Konzept umfasst ein 3D-Scansystem zur Schadensanalyse und -erkennung, eine Schadensmusterdatenbank und ein robotergeführtes KGS-Reparatursystem. Die mit dem 3D-Scansystem gewonnenen Daten des Bauteils werden mit vorhandenen Schadensmustern aus der Datenbank abgeglichen, um Standardfehler zu identifizieren. Diese Informationen werden verwendet, um eine Reparaturstrategie festzulegen und eine geeignete KGS-Spiraltrajektorie für den Materialauftrag zu erstellen. Al-

lerdings bestehen bei der oben genannten Methode der Spiraltrajektorie die erwähnten Einschränkungen in Bezug auf den Spritzwinkel und den Spritzabstand.

Hinsichtlich der Trajektorienoptimierung schlugen Kout und Müller [11] eine Optimierung durch Gradientenabstieg zur Minimierung des Schichtdickenfehlers unter Einbeziehung eines Materialauftragsmodells vor. Dieses Konzept bestimmt zur Zielerreichung die erforderlichen Spritzpistolenpositionen und -ausrichtungen in Bezug auf das Substrat, anstatt sich initial auf die rechnerunterstützte Design-(CAD)-Geometrie zu konzentrieren. Obwohl die Ergebnisse gute Annäherungen an die gewünschten Beschichtungen erzielen, vernachlässigt dieser Ansatz die Trajektorienplanung des Roboters und die resultierende Roboterbewegung vollständig.

Ren et al. [12] untersuchten die Optimierung der Schichtdicke für thermisch gespritzte Schichten. Auf Grundlage eines Materialauftragsmodells wurde in dieser Arbeit die Nelder-Mead-Methode zur Lösung der Optimierung verwendet, was zu optimalen sekundären Prozessparametern für einen Zickzack-Pfad führte. Das Konzept wurde dann für einen Anwendungsfall implementiert, um die optimierte Trajektorie zu validieren. In dieser Arbeit wurden jedoch nur flache Substrate und keine komplexen Geometrien untersucht.

Bei Reparaturanwendungen gilt es das zu reparierende Volumen systematisch durch schichtweisen Aufbau wiederherzustellen. Da der beschädigte Bereich für die Reparatur vorbearbeitet werden muss, haben die Vertiefungen konkave Oberflächen, bei denen eine gleichmäßige Schichtdicke für eine effiziente Materialnutzung nicht ausreicht. Die Materialeffizienz ist definiert als die Differenz zwischen dem tatsächlich für die Reparatur benötigten Materialvolumen und der tatsächlich aufgetragenen Materialmenge. Anstatt sich auf die Optimierung des Pfades zur Aufrechterhaltung einer konstanten Spritzgeschwindigkeit zu konzentrieren, wird in dieser Arbeit die Spritzgeschwindigkeit als Entscheidungsvariable in die Optimierung einbezogen. Ziel ist die Nutzung der durch den Pfad und die Spritzgeschwindigkeit gegebenen Freiheitsgrade, um den angestrebten Materialauftrag bei der Reparatur zu erreichen und dabei die Einschränkungen der Roboterkinematik zu berücksichtigen. Aufgrund der stochastischen Natur nicht-deterministischer Optimierungsmethoden (z. B. evolutionäre Algorithmen) wird in dieser Arbeit ein deterministischer Ansatz verfolgt, um die Reproduzierbarkeit zu gewährleisten, die für industrielle Anwendungen entscheidend ist.

## III. KONZEPT ZUR INITIALEN TRAJEKTORIENPLANUNG UND TRAJEKTORIENOPTIMIERUNG FÜR DAS ROBOTERGEFÜHRTE KALTGASSPRITZEN

In diesem Abschnitt wird das Konzept zur initialen Trajektorienplanung und der anschließenden Trajektorienoptimierung vorgestellt. Die Implementierung des Konzepts erfolgte in FreeCAD (Version 0.20) [13] und RoboDK (Version 5.5.0) [14]. FreeCAD ist eine parametrische 3D-Open-Source-Software für CAD, die einen auf der Programmiersprache Python basierenden Interpreter enthält. Das vorgestellte Konzept wurde mit Python-Skripten implementiert. Es wurde eine Verbindung zur Roboterprogrammier- und Simulationssoftware RoboDK hergestellt. Dies ermöglicht eine Offline-Simulation der generierten Roboterprogramme, so dass die Programme getestet werden können, bevor sie in Robotersteuerungscode für den realen Prozess umgewandelt werden.

### A. Initiale Trajektorienplanung

In diesem Abschnitt wird die initiale Trajektorienplanung beschrieben. In Hinblick auf die Anforderungen an einen orthogonalen Spritzwinkel und die Materialeffizienz werden zunächst das zu reparierende Volumen bestimmt und geeignete adaptiv gekrümmte Schichten für den Materialauftrag definiert. Darüber hinaus bedingen konkave Oberflächen Unterschiede in der erforderlichen Schichtdicke. Es gibt Bereiche, an denen mehr Material benötigt wird, und andere, an denen weniger Material zum Wiederaufbau des Volumens erforderlich ist. Zur Einsparung des aufzutragenden Materials und unter dem Ziel für einen nahezu endkonturnahen Materialauftrag sind variable Schichtdicken vorgesehen. Zusammenfassend ergibt sich daraus der Ansatz eines Algorithmus zum adaptiven „Slicen“ von gekrümmten Schichten. Anschließend erfolgt die automatische Pfadplanung für jede einzelne Schicht unter Verwendung von Punktwolken und Normalenvektoren. Der Pfad entspricht der geometrischen Darstellung der Bewegung. Dieser Schritt erlaubt es, verschiedene Pfadalgorithmen zu integrieren. Danach erfolgt die Generierung der Trajektorie durch Hinzufügen einer Spritzgeschwindigkeitsverteilung zum generierten Pfad. In diesem Schritt wird der geplante Pfad in eine Trajektorie umgewandelt. Der relevante, zu berechnende Parameter, ist die Spritzgeschwindigkeit. Ausgehend von der gewünschten Schichtdicke an einem Punkt ergibt sich die erforderliche Spritzgeschwindigkeit. Die Spritzgeschwindigkeit ist somit der Parameter zur Gewährleistung variabler Schichtdicken. Dies ermöglicht es, die Menge des Materialauftrags pro Fläche durch eine kontrollierte Materialauftragszeit zu steuern.

Für eine detaillierte Beschreibung der initialen Trajektorienplanung wird auf die Arbeit in [7] verwiesen. Nach der erfolgreichen initialen Trajektorienplanung erfolgt die anschließende Trajektorienoptimierung.

### B. Trajektorienoptimierung

In diesem Abschnitt wird das Konzept der Trajektorienoptimierung vorgestellt. Um einen überschaubaren Suchraum für die mathematische Optimierung zu erhalten, basiert die Trajektorienoptimierung auf einer initialen Trajektorie. Diese Initialtrajektorie liefert eine akzeptable Startlösung und dient als Referenz für die nachfolgende Optimierung. In diesem Konzept erfolgt die Erzeugung der Initialtrajektorie entsprechend des Abschnitts III A. Das Slicing der gekrümmten Schichten wird zusätzlich durch eine adaptive Verkleinerung der Flächen verbessert, um zu verhindern, dass der Spritzstrahl Bereiche überfährt, in denen bereits ausreichend Material vorhanden ist. In dieser Arbeit wird eine spiralförmige Trajektorie verwendet, um unnötigen Materialauftrag außerhalb des relevanten Bereichs zu reduzieren, die Materialeffizienz zu erhöhen und den Aufwand für die Nachbearbeitung zu verringern.

Die Beschreibung des Problems beginnt mit der Darstellung der Startoberfläche der zu füllenden Kavität. Die Startoberfläche der Kavität in CAD wird von der Points Workbench von FreeCAD in eine diskrete Punktwolke umgewandelt, die die Oberfläche durch Punkte im dreidimensionalen Raum darstellt. Durch die Einstellung des euklidischen Abstands kann die räumliche Auflösung der Punktwolke angepasst werden. Darüber hinaus werden jedem Punkt der Punktwolke entsprechende Normalenvektoren zugewiesen, die die Orthogonalität zur darunter liegenden Oberfläche gewährleisten. Die gewünschte Endkontur der zu füllenden Kavität ist durch die Soll-Schließfläche der Kavität in CAD gegeben. Das

eingeschlossene Volumen zwischen der Kavitäten-Startoberfläche und der nominalen Kavitäten-Schließfläche entspricht somit dem mit KGS herzustellenden Volumen. Die Initialtrajektorie wird mit Werten für die Pose und die Spritzgeschwindigkeit als Referenz angenommen. Um die Bewegung der KGS-Pistole für die Trajektorienoptimierung vollständig darzustellen, werden die Pose und die Spritzgeschwindigkeit als Entscheidungsvariablen definiert. Zur Modellierung des Materialauftrags wird in dieser Arbeit eine Überlagerung von zweidimensionalen Gaußfunktionen definiert. Um das Ziel einer optimierten Trajektorie zu erreichen, die die Anforderungen des KGS und der Roboterkinematik berücksichtigt, ist eine Zielfunktion erforderlich. Im Rahmen dieser Arbeit werden die Ziele der Minimierung des Materialüberschusses an der Reparaturstelle und der Minimierung des Rucks der Roboterkinematik formuliert, um einen effizienten Materialeinsatz und eine qualitativ hochwertige Reparatur durch gleichmäßigen Materialauftrag zu gewährleisten. Schließlich werden alle Teilziele durch die Methode der gewichteten Summe zu einer multikriteriellen Optimierungsfunktion zusammengefasst, die es zu minimieren gilt. Nicht-negative skalare Gewichtungsfaktoren werden verwendet, um sicherzustellen, dass alle Teilziele in ähnlichen Größenordnungen liegen. Die Gewichtungsfaktoren dienen auch dazu, die Teilziele zu priorisieren. Es werden Rand- und Nebenbedingungen für die Roboterkinematik formuliert, um die Spritzgeschwindigkeit und die Beschleunigung zu limitieren. Zusätzlich wird die Bedingung formuliert, eine vollständige Füllung der Kavität zu gewährleisten. Darüber hinaus wird die Anforderung eines zur Oberfläche orthogonalen Spritzwinkels durch Verwendung der entsprechenden Normalenvektoren berücksichtigt. Der Spritzabstand wird adaptiv an bereits aufgetragenes Material angepasst.

Für eine detaillierte Beschreibung der Trajektorienoptimierung wird auf die Arbeit in [8] verwiesen. Folgend zur initialen Trajektorienplanung und der anschließenden Trajektorienoptimierung wird im Folgenden die Anwendung des Konzepts für ein Anwendungsbeispiel vorgestellt.

## IV. ERGEBNISSE

Um die Gültigkeit des Konzepts zu testen, wird eine Kavität in einem ebenen Bauteil als repräsentativer Anwendungsfall verwendet. Die Kavität hat eine Länge von 14,4 mm, eine Breite von 13,8 mm und eine Tiefe von 1,2 mm, woraus sich aus der geplanten Initialtrajektorie 3 Schichten für den Materialaufbau ergeben. Der Durchmesser des KGS-Spritzflecks wird zu 8 mm angenommen. Als Pulvermaterial wird die Aluminiumlegierung 6061 mit einer Dichte von  $2,7 \text{ g/cm}^3$  verwendet. Die Pulverfördertrate wird auf 20 g/min und der Spritzspurabstand auf 2 mm eingestellt. Der Auftragswirkungsgrad wird mit 70 % angenommen. Die Anzahl der Elemente in der Initialtrajektorie beträgt 162. Bei der Trajektorienoptimierung werden die Grenzen für die Spritzgeschwindigkeit und die Beschleunigung auf 50-500 mm/s bzw.  $\pm 5000 \text{ mm/s}^2$  festgelegt. Der Toleranzbereich für überschüssiges Material wird beispielhaft auf 0 mm festgelegt. Für die Optimierung werden dabei folgende Gewichtungsfaktoren gewählt: 1 für den Materialauftrag, 0,1 für den Pfad und  $1e-8$  für den Ruck. Die Ergebnisse der Optimierung sind in Abb. 1 dargestellt. Die Optimierung gewährleistet eine deutliche Glättung der Spritzgeschwindigkeiten durch Verringerung der Beschleunigung und des Rucks. Die final erzeugte Oberfläche nach dem Materialauftrag berücksichtigt die Anforderung, dass die Kavität vollständig gefüllt wird. Das Ergebnis für den

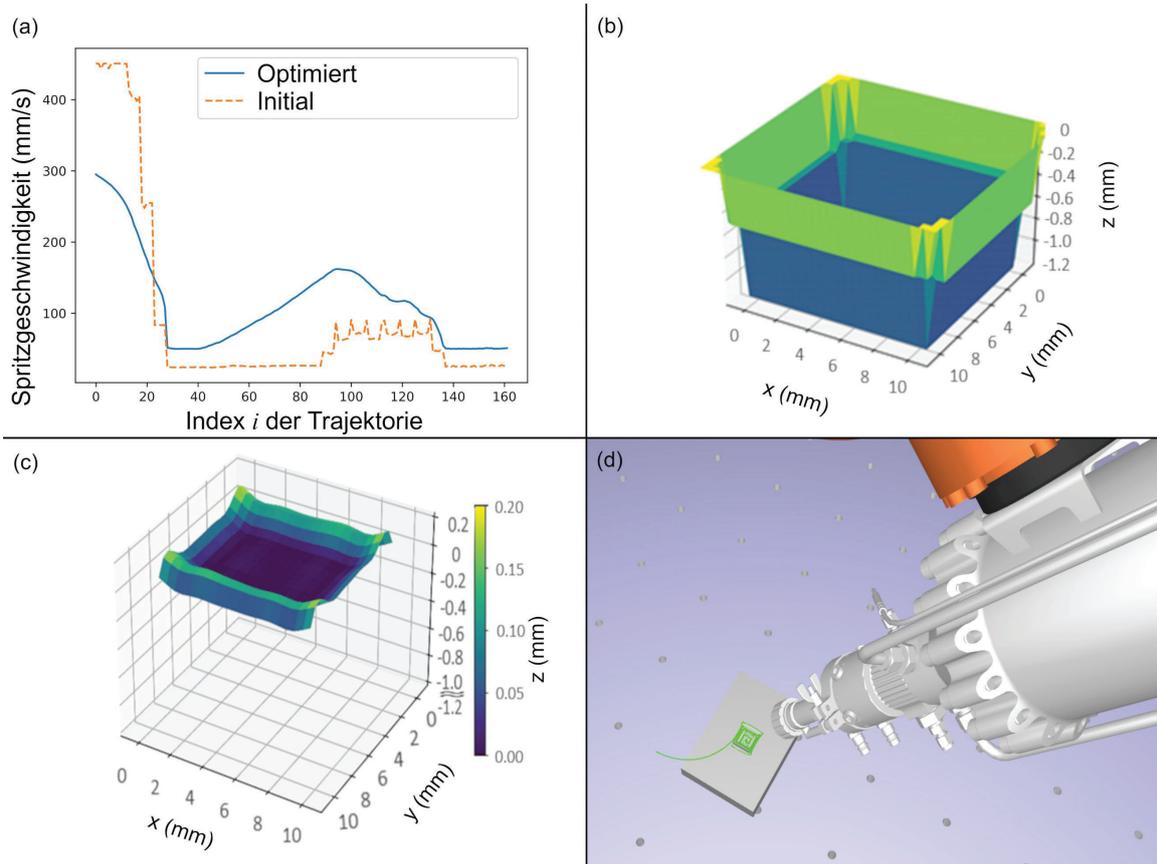


ABBILDUNG 1: SIMULATIONSERGEBNIS DER TRAJEKTORIENOPTIMIERUNG FÜR DIE EBENE KAVITÄT, (A) SPRITZGESCHWINDIGKEITEN VOR UND NACH DER OPTIMIERUNG, (B) KAVITÄT VOR DEM MATERIALAUFTRAG, (C) MATERIALAUFTRAG IN DER KAVITÄT NACH DER OPTIMIERUNG, (D) OPTIMIERTE TRAJEKTORIE IN DER ROBOTERSIMULATIONSUMGEBUNG ROBODK. BITTE BEACHTEN SIE DIE UNTERSCHIEDLICHEN MAßSTÄBE DER X-Y-EBENE UND DER Z-ACHSE IN (B) UND (C). NACH [8]

Materialauftrag zeigt eine mittlere Abweichung von 0,16 mm von der nominalen Kavitäten-Schließfläche, was einer hinreichend geringen Abweichung entspricht und mit dem nur leichten Überschuss eine hohe Materialeffizienz gewährleistet. Basierend auf der optimierten Trajektorie ermöglicht das vorgestellte Konzept nun die Generierung eines Roboterprogramms und dessen Übertragung auf die reale Robotersteuerung.

## V. ZUSAMMENFASSUNG UND AUSBLICK

Das robotergeführte Kaltgasspritzen ermöglicht eine flexible Steuerung und Automatisierung des Materialauftrags. Dies ermöglicht die Anwendung des Kaltgasspritzens als Reparaturverfahren für lokal beschädigte Bauteile. Für den erfolgreichen Einsatz als Reparaturverfahren des robotergeführten Kaltgasspritzens ist allerdings unerlässlich, die prozessbedingten Anforderungen zu berücksichtigen. Dies erfordert ein umfassendes Konzept, um den Reparaturprozess optimal zu planen, zu simulieren und auszuführen. In dieser Arbeit wird ein Konzept zur automatisierten initialen Trajektorienplanung und Trajektorienoptimierung mittels mathematischer Optimierung für die Reparatur durch robotergeführtes Kaltgasspritzen vorgestellt. Die mathematische Optimierung wird eingesetzt, um die Ziele und Einschränkungen des Reparaturprozesses im Hinblick auf die geometrischen Einschränkungen des Materialauftrags, den effizienten Materialeinsatz und die Einschränkungen der Roboterkinematik effektiv zu handhaben. In diesem Zusammenhang werden Entscheidungsvariablen, Zielfunktion, Rand- und Nebenbedingungen und ein Materialauftragsmodell definiert, um den Materialauftrag und die

Roboterbewegung entsprechend zu steuern. Die Ergebnisse zeigen am Beispiel eines Anwendungsfalls die erfolgreiche Umsetzung dieses Konzepts. Unter Berücksichtigung der Roboterkinematik und der sekundären Prozessparameter des Kaltgasspritzens wird eine optimierte Trajektorie erzeugt, die einen minimalen Ruck und einen minimalen Materialüberschuss gewährleistet. Die nächsten Schritte umfassen die Integration dieses Konzepts in ein umfassendes robotergeführtes Kaltgasspritz-Reparatursystem und die Durchführung von Laborexperimenten zur praktischen Validierung. Experimente und die anschließenden Analysen sind erforderlich, um die Einflüsse der Trajektorie auf die Mikrostrukturen und Eigenschaften des Materialauftrags zu untersuchen. Neben der Reparatur kann dieses Konzept auch für Anwendungen in der additiven Fertigung durch Kaltgasspritzen adaptiert werden.

## DANKSAGUNG

Die Autoren bedanken sich im Rahmen des Forschungsprojekts „CORE - Computergesteuerte Bauteilbearbeitung“ für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. Das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr – ist ein von den Universitäten der Bundeswehr Hamburg und München gemeinsam getragenes wissenschaftliches Zentrum und Bestandteil des Konjunkturprogramms der Bundesregierung zur Überwindung der COVID-19- Krise. Es unterliegt der akademischen Selbstverwaltung. Die Mittel, mit

dem das dtec.bw ausgestattet wurde, werden an beiden Universitäten der Bundeswehr zur Finanzierung von Forschungsprojekten und Projekten zum Wissens- und Technologietransfer eingesetzt. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

#### LITERATUR

- [1] H. Assadi, H. Kreye, F. Gärtner, and T. Klassen, "Cold spraying – A materials perspective," *Acta Mater.*, vol. 116, pp. 382–407, 2016, doi: 10.1016/j.actamat.2016.06.034.
- [2] R. N. Raoelison, C. Verdy, and H. Liao, "Cold gas dynamic spray additive manufacturing today: Deposit possibilities, technological solutions and viable applications," *Mater. Des.*, vol. 133, pp. 266–287, 2017, doi: 10.1016/j.matdes.2017.07.067.
- [3] W. Li, K. Yang, S. Yin, X. Yang, Y. Xu, and R. Lupoi, "Solid-state additive manufacturing and repairing by cold spraying: A review," *J. Mater. Sci. Technol.*, vol. 34, no. 3, pp. 440–457, 2018, doi: 10.1016/j.jmst.2017.09.015.
- [4] S. Yin *et al.*, "Cold spray additive manufacturing and repair: Fundamentals and applications," *Addit. Manuf.*, vol. 21, pp. 628–650, 2018, doi: 10.1016/j.addma.2018.04.017.
- [5] I. M. Nault, G. D. Ferguson, and A. T. Nardi, "Multi-axis tool path optimization and deposition modeling for cold spray additive manufacturing," *Addit. Manuf.*, vol. 38, 2021.
- [6] *Thermisches Spritzen – Begriffe, Einteilung*, 14917, Deutsches Institut für Normung e. V., Berlin, Aug. 2017.
- [7] M. Lewke, H. Wu, A. List, F. Gärtner, T. Klassen, and A. Fay, "Automated Path and Trajectory Planning for Automated Repair of Damaged Components by Cold Spray," in *Thermal Spray 2023: Proceedings from the International Thermal Spray Conference*, Québec City, Canada, 2023, pp. 436–442.
- [8] M. Lewke, H. Wu, A. List, F. Gärtner, T. Klassen, and A. Fay, "Trajectory Optimization for Repair by Robot-guided Cold Spray," in *Thermal Spray 2024: Proceedings from the International Thermal Spray Conference*, Milan, Italy, 2024, pp. 123–129.
- [9] C. Chen *et al.*, "A novel spiral trajectory for damage component recovery with cold spray," *Surf. Coat. Technol.*, vol. 309, pp. 719–728, 2017.
- [10] H. Wu, S. Liu, X. Xie, Y. Zhang, H. Liao, and S. Deng, "A framework for a knowledge based cold spray repairing system," *J. Intell. Manuf.*, vol. 33, pp. 1639–1647, 2022, doi: 10.1007/s10845-021-01770-7.
- [11] A. Kout and H. Müller, "Parameter optimization for spray coating," *Adv. Eng. Softw.*, vol. 40, no. 10, pp. 1078–1086, 2009, doi: 10.1016/j.advengsoft.2009.03.001.
- [12] J. Ren, Y. Sun, J. Hui, R. Ahmad, and Y. Ma, "Coating thickness optimization for a robotized thermal spray system," *Robot. Comput. Integr. Manuf.*, vol. 83, 2023, doi: 10.1016/j.rcim.2023.102569.
- [13] The FreeCAD Team, *FreeCAD (Version 0.20)*. [Online]. Available: <https://www.freecadweb.org/> (accessed: Nov. 3 2022).
- [14] RoboDK Inc., *RoboDK (Version 5.5.0)*. [Online]. Available: <https://robodk.com/> (accessed: Nov. 3 2022).

# A Practical Toolpath Planning Method for Cold Spray Additive Manufacturing of Rotational Symmetry Parts

Hongjian Wu \*, Alexander List, Frank Gärtner, Thomas Klassen

*Institute of Materials Technology  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany.  
\*wuh@hsu-hh.de*

Marcel Lewke

*Institute of Automation Technology  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany*

**Abstract** – Rotational symmetry parts are common and essential in industrial applications. Cold spray additive manufacturing (CSAM) is an attractive and rapidly developing solid-state material deposition process, providing an efficient and convenient method for producing such parts, as it allows for the rapid formation of high-quality, large-volume 3D objects. Since there is no highly reactive liquid phase involved in this process, the deposited material is free of oxides. As compared to conventional additive manufacturing methods, cold spraying enables to reduce the production costs and times.

In this work, a general implementation method for CSAM of rotating symmetry casing parts is presented. Here, the developed application can handle rotational symmetry parts of arbitrary geometry in the form of CAD files to generate precise toolpaths. Robot offline programming allows for process simulation, analysis, and optimization. Additionally, modelling of robot kinematics is employed to evaluate the effect of the planned toolpaths on the spraying process, ensuring efficient and precise manufacturing processes.

**Keywords** – Cold spray, additive manufacturing, toolpath planning, rotational symmetry part

## I. INTRODUCTION

Cold gas dynamic spraying, or cold spraying (CS) for short, is a rapidly evolving solid powder deposition technique. In this process, micron-sized powders are accelerated to supersonic velocities by a high-pressure gas flow through a convergent-divergent de-Laval nozzle. The particles remain at a temperature below the melting point. The impact on the solid surface results in severe localized plastic deformation and a temperature rise, causing local interface softening and enabling the formation of deposits by adiabatic shear instabilities [1]. CS was initially used for developing various functional surface properties, ranging from protection against harsh environments to specifically tailored surfaces in electronic applications [2]. Nowadays, in industry and research laboratories, serious attention is paid to topics of CS-based additive manufacturing (AM) as well as to applications in repair or remanufacturing techniques [3], [4], [5]. Unlike the usually used thermal AM techniques involving melting, cold spray additive manufacturing (CSAM) solely involves solid state deposition. This implies that oxidation as pronounced by fast reaction kinetics in the molten state and the formation of undesired phases or substructures associated with rapid solidification can be

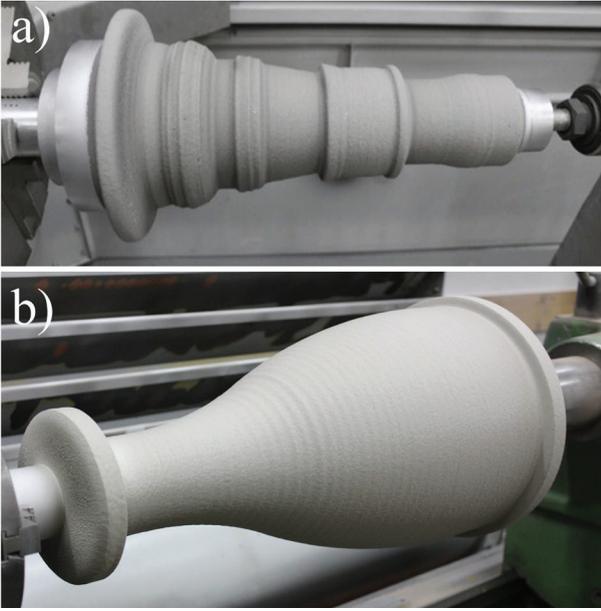
avoided. In addition, CSAM can reach high spraying rates of up to 10 kg/h and allows for fabrication of complex and larger parts assisted by industrial robots. Therefore, CSAM is considered as an effective and economic solution for manufacturing of heat sensitive materials (especially pure titanium, copper, aluminium, and their alloys) and rapid fabrication of large-volume parts.

Casing parts with rotational symmetry are commonly used in engineering structures, such as casings of aircraft engines, gas turbines, pumps, and wind power plants. These components usually have features such as specialized flanges or bosses on the periphery. Conventional manufacturing methods are performed through a combination of several turning and milling processes, which are often very intricate and costly, resulting in significant material waste. CSAM enables to create the contour of the component close to the final shape easily by introducing an external rotation axis driven by an individual electrical regulator. Many similar cases have been reported [6], [7], e.g. a fan shaft [8] (as shown in Figure 1a) and a rocket nozzle [9] (as shown in Figure 1b) fabricated using CSAM. However, the implementation methods are not clearly explained in the reports.

The present work aims to develop a general implementation strategy and toolpath planning method for CSAM of rotating symmetry casing parts. In the practical spraying process achieving precise control over kinematic parameters is essential for successful shape buildup. This involves assigning the coordinates of each point on the part surface and specifying associated conditions, such as spray angle and local scanning speed. All that could be correlated to associated robot movements as well as obtained layer deposition and guide to more profound toolpath development.

## II. GENERAL CONCEPTS

The general AM process involves two phases dealing with design and printing. These phases can be further divided into four steps: (i) digital modelling; (ii) pre-processing, involving simulation and preparation of the files for 3D printing; (iii) printing; (iv) post-processing, including removal of the support and separation of the parts. In order to introduce the basic idea of AM into the CS process, a generic CSAM implementation process is proposed based on existing generic technical configurations and layouts.



FIGURES 1: ADDITIVE MANUFACTURING (A) STAND-ALONE TURBOJET ENGINE FAN SHAFTS AND (B) ROCKET NOZZLES WITH COLD SPRAYING ACCORDING TO [8] and [9]

The first step of the process is to design an initial geometry and to create a CAD file of the printed object. After establishing the CAD model, the step of analysing the features of the object and formulating corresponding building strategies is completed by simulation tools, which mainly involve toolpath planning and generation as well as analyses and adjustments of robot motion programs. Currently, generic strategies for material buildup are proposed for rotational symmetry casings. The corresponding toolpaths are planned according to the axial cross-section of the part. A previously established numerical model to obtain 3D CS deposits [10] is applied to determine path parameters. A custom developed script is used to automatically generate the desired toolpaths. It takes CAD files as input and generates the toolpaths in the form of a 6-axis robotic arm program as output.

Before actually depositing material for building the part, it is essential to carefully check the generated robot trajectory program. This should ensure that no singularities arise in the robot joints, or a specific robot axis reaches its motion limits during the workpiece deposition process. If such situations occur, adjustments to the robot program or even re-planning of the toolpath are required. In addition, machine collisions must be avoided. Therefore, process simulation via a virtual workstation established in robot offline programming software is necessary to enable machine accessibility and collision detection. Additionally, it facilitates the analysis of robot kinematics, adjustment of toolpaths and buildup processes as well as the calculation of manufacturing time. After calibration and debugging of the manufacturing program in the actual spraying workshop, the spraying process can be implemented to build the initial shape. The details of the part and the needed tolerances require further machining (e.g. drilling, edge chamfering, and final dimensional finishing).

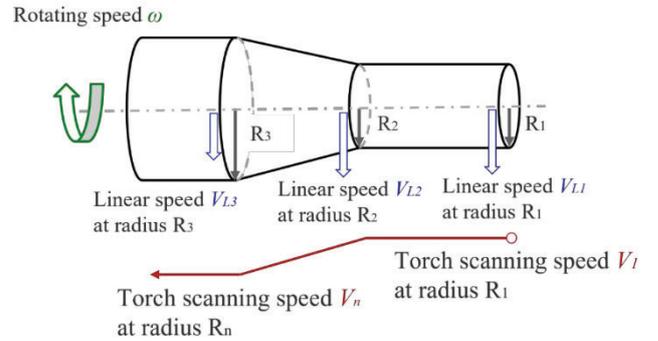
### III. BUILTUP STRATEGIES FOR ROTATIONAL SYMMETRY PARTS

#### A. Path Parameter Determination: Scanning Speed

CSAM is capable of producing various rotational structures including free-form cylinders and cones. In the depositing process of the object, the basic motion configuration of the system includes that the external axis rotates unidirectionally at rotating speed  $\omega$  and the robot performs a linear motion (nozzle scanning speed  $V$ ) in axis direction. The centre of the nozzle keeps moving in the same plane aligned to the axis of the rotating substrate. The outlet of the nozzle maintains the specified stand-off distance to the surface and torch orientation at all points. The gun needs to be offset in each pass to compensate for the reduction in stand-off distance due to increasing part dimensions under deposition, thus ensuring constant process parameters, especially for depositing thick shell parts. The settings of rotational speed and torch scanning speed depend on the variation of the radius of the rotating substrate. For a substrate with a constant radius (e.g. a cylinder), the speed settings remain constant. For a substrate with varying radius, the tangential speed of the relative linear speed  $V_L$  between torch and substrate increases with the increase in radius. The torch scanning speed should therefore decrease as the radius increases to maintain the same increase in deposit thickness on the substrate (as shown in Figure 2). The torch scanning speed can be described by

$$V_T = \frac{R_1}{R_T} \cdot V_1 \quad (1)$$

Here,  $V_1$  is the torch scanning speed at radius  $R_1$ , defined as the reference speed and radius for the calculation of other torch scanning speeds.  $R_T$  is the corresponding radius at a different target point along the path. The ratio of  $R_1$  and  $R_T$  is regarded as the adjustment factor of speed.



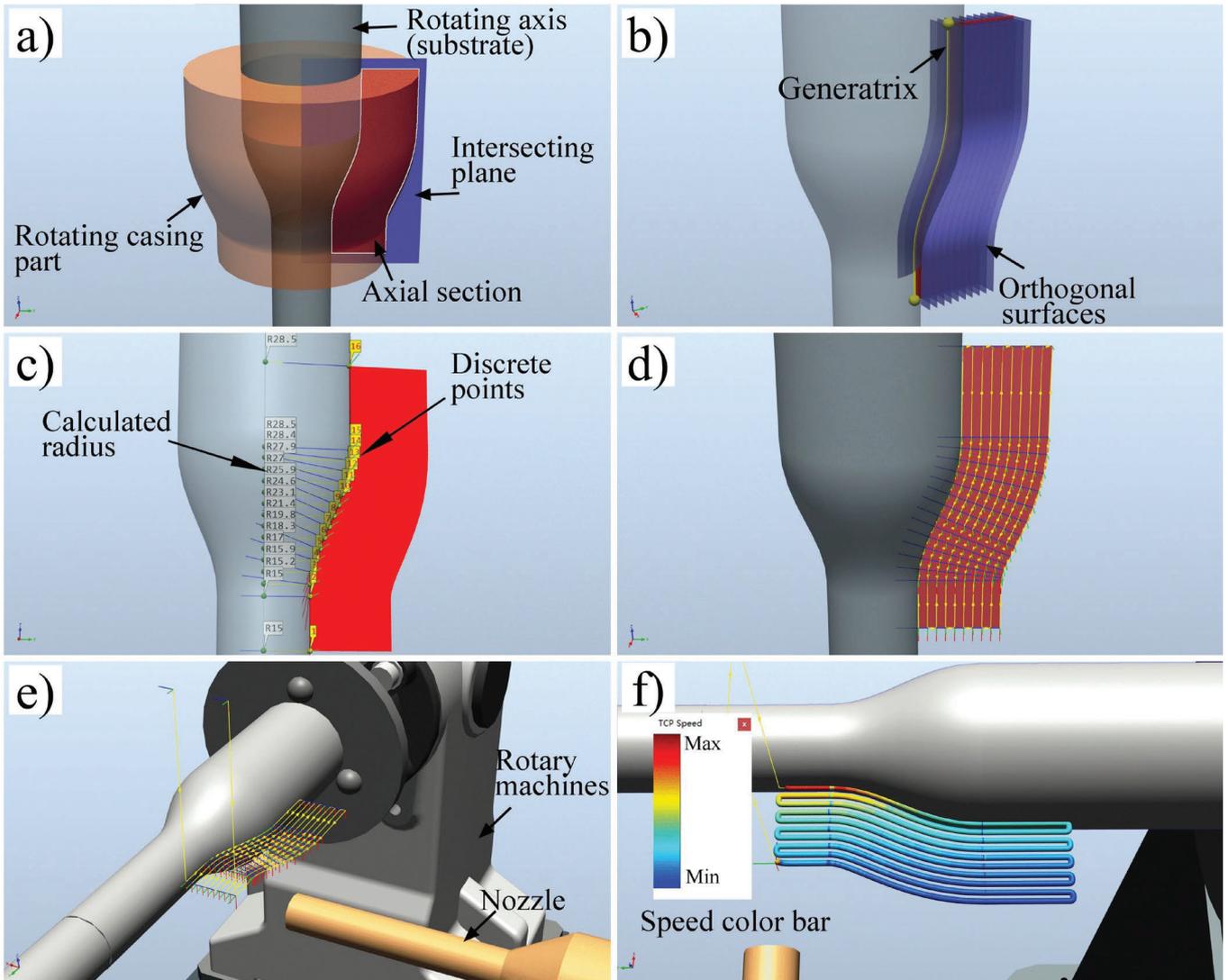
FIGURES 2: RELATION BETWEEN SETTING OF TORCH SCANNING SPEED AND SUBSTRATE RADIUS

#### B. Toolpath Planning for Rotational Symmetry Parts

Figure 3 illustrates the principle of toolpath planning and generation for casing objects with rotational symmetry. The first step is to obtain geometric data of the axial section of the rotating object. These can be obtained by using measured dimensions as input or by calculations from plane intersections in the CAD file of the object (as shown in Figure 3a). Then, the generatrix of the rotating substrate, i.e. the intersection line between the substrate and the obtained axial section, is selected to create an orthogonal surface (as shown in Figure 3b). This orthogonal surface is then off-set according to layer thickness to create a series of orthogonal surfaces that can completely cut the axial section. The values of layer thickness

can be calculated empirically or obtained through experiments for a certain powder/substrate material system and adjusted process parameter sets. Based on Boolean operations, intersecting curves between the axial section and the orthogonal surfaces can be obtained. Afterwards, each intersecting curve is discretized into points. In general, for straight-line segments only respective endpoints are needed. For curved segments, a tighter distribution of discrete points is required (as shown in Figure 3c). Each point contains information on index number, position, direction, and the radius of the substrate at the corresponding position. The orientation of each point is perpendicular to the surface of the substrate or the already buildup deposit. The axis radius information of each point is used for the

automatic speed calibration after the subsequent setting of the torch scanning speed. Finally, a continuous toolpath can be created by linking each target in sequence (as shown in Figure 3d). The corresponding robot motion code can be automatically created and simulated in the virtual environment of robot (as shown in Figure 3e). Based on the previously developed method of robot kinematic analyses and visualization [11], Figure 3f shows that the torch scanning speed is executed in a preset way, i.e. the torch scanning speed decreases as the radius increases. A script is developed based on the aforementioned principles and methods and can automatically generate toolpaths for various rotary casing objects, as depicted in Figure 4.



FIGURES 3: PRINCIPLES OF THE TOOLPATH PLANNING APPROACH FOR OBJECTS WITH ROTATIONAL SYMMETRY: (A) OBTAINING AXIAL SECTIONS, (B) CREATING OF ORTHOGONAL PLANES BASED ON THE GENERATRIX, (C) DISCRETE POINTS WITH THE INFORMATION OF THEIR INDEX NUMBER AND THE RADIUS OF THE SUBSTRATE AT THEIR CORRESPONDING POSITION, (D) ORDERING AND CONNECTION OF POINTS, (E) TOOLPATH SIMULATION, AND (F) TORCH SCANNING SPEED VARIATION (FOR INTERPRETATION OF THE REFERENCES TO COLOUR IN THIS FIGURE LEGEND, THE READER IS REFERRED TO THE WEB VERSION OF THIS ARTICLE)

#### IV. SUMMARY AND OUTLOOK

In this work, a general method for CSAM toolpath planning and implementation for rotationally symmetric casing parts is presented. This method comprehensively considers the characteristics of cold spray, the robot kinematics, and the geometry of the workpiece, enabling the automatic and accurate

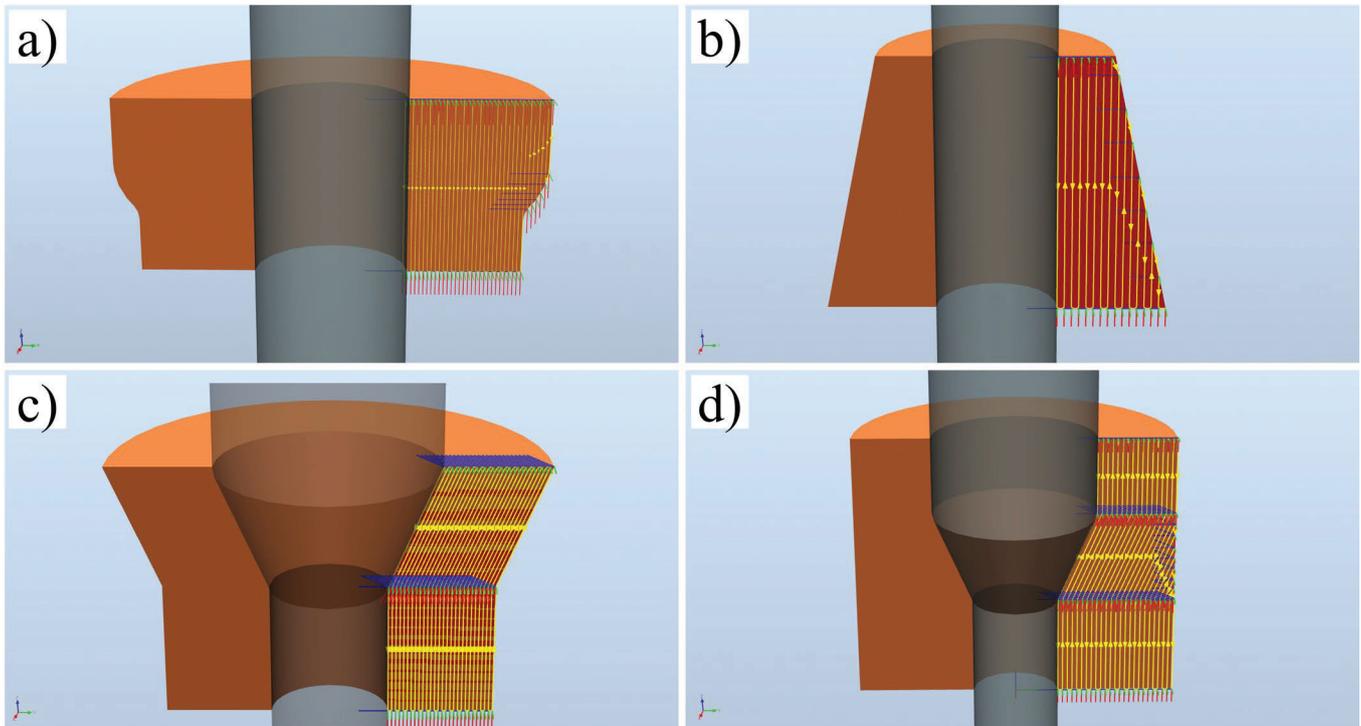
creation of a desired toolpath through customized scripts. Robot kinematics simulations have demonstrated the feasibility of this method. Further validation and optimization will be pursued through specific case studies. Additionally, efforts will be made to integrate and coordinate other sub-processes,

such as online monitoring and diagnostics as well as subtractive machining techniques to exert the full potential of cold spray additive manufacturing in shape building.

#### ACKNOWLEDGEMENT

The authors acknowledge the financial support in the frame of the project “CORE – Computerized Refurbishment” funded by dtec.bw – Digitalization and Technology Research

Center of the Bundeswehr. Digitalization and Technology Research Center of the Bundeswehr is a scientific centre jointly supported by the Universities of the Federal Armed Forces in Hamburg and Munich and is part of the German government's economic stimulus program to overcome the COVID-19 crisis. It is subject to academic self-administration. The funds within dtec.bw for both universities of the Federal Armed Forces finance research projects and projects for knowledge and technology transfer. dtec.bw is funded by the European Union – NextGenerationEU.



FIGURES 4: TOOLPATHS FOR VARIOUS ROTATING CASING PARTS

#### REFERENCES

- [1] H. Assadi, H. Kreye, F. Gärtner, and T. Klassen, “Cold spraying—A materials perspective,” *Acta Materialia*, vol. 116, pp. 382–407, 2016.
- [2] F. Gärtner, T. Stoltenhoff, T. Schmidt, and H. Kreye, “The cold spray process and its potential for industrial applications,” *Journal of Thermal Spray Technology*, vol. 15, pp. 223–232, 2006.
- [3] H. Wu *et al.*, “Stable layer-building strategy to enhance cold-spray-based additive manufacturing,” *Additive Manufacturing*, vol. 35, p. 101356, 2020.
- [4] F. Gärtner, A. List, S. Nielsen, H. Wu, Z. Arabgol, L. Wiehler, et al. Computerized Refurbishment, dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw, 1(2022) 9–19.
- [5] M. Lewke, H. Wu, A. List, F. Gärtner, T. Klassen, and A. Fay, “Automated Trajectory Planning and Analytical Improvement for Automated Repair by Robot-Guided Cold Spray,” *J Therm Spray Tech*, Dec. 2023, doi: 10.1007/s11666-023-01697-w.
- [6] F. Ortega, A. Sova, M. D. Monzón, M. D. Marrero, A. N. Benítez, and P. Bertrand, “Combination of electroforming and cold gas dynamic spray for fabrication of rotational moulds: feasibility study,” *The International Journal of Advanced Manufacturing Technology*, vol. 76, no. 5, pp. 1243–1251, 2015.
- [7] J.-M. Lamarre and F. Bernier, “Permanent magnets produced by cold spray additive manufacturing for electric engines,” *Journal of Thermal Spray Technology*, vol. 28, pp. 1709–1717, 2019.
- [8] “Additive manufacturing Ti-6Al-4V.” Accessed: Jun. 26, 2024. [Online]. Available: <https://impact-innovations.com/en/applications/cold-spray-additive-manufacturing-ti-6al-4v/>
- [9] “Additive manufacturing rocket nozzle.” Accessed: Jun. 26, 2024. [Online]. Available: <https://impact-innovations.com/en/applications/rocket-nozzle/>
- [10] H. Wu *et al.*, “A new approach to simulate coating thickness in cold spray,” *Surface and Coatings Technology*, vol. 382, p. 125151, 2020.
- [11] H. Wu *et al.*, “Generic implementation of path design for spray deposition: Programming schemes, processing and characterization for cold spraying,” *Surface and Coatings Technology*, vol. 458, p. 129368, 2023.

# Temperature and Stress Management in Cold Sprayed Deposits

Zahra Arabgol\*, Sören Nielsen, Alexander List,  
Levke Wichler, Frank Gärtner, Thomas Klassen

*Institute of Materials Technology  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany  
\*arabgolz@hsu-hh.de*

Dhruvit Gabani\*, Jens Gibmeier  
*Institute for Applied Materials (IAM-WK)  
Karlsruhe Institute of Technology (KIT)  
Karlsruhe, Germany  
\*dhruvit.gabani@kit.edu*

**Abstract** – Material deposition in cold spraying occurs in solid state and thus avoids undesired effects of melting and solidification. However, residual stress conditions in cold sprayed coatings could limit possible part performance. The temperature distribution and thermal history of the cold sprayed components has significant influence on stress distribution and thus deposition and part quality.

The present study investigates the effect of substrate material and nozzle traverse speed (as a secondary parameter) on effective temperatures and residual stress distributions of titanium-grade 1 deposits. The results demonstrate that substrate material properties and nozzle traverse speeds have significant influence on residual stresses of the cold spray deposit. It is understood that coefficient of thermal expansion (CTE) difference of the coating and substrate materials has significant effect on residual stress state. On the other hand, the residual stresses change from more compressive to more tensile state as the temperature of the components increases by decreasing the nozzle traverse speed. These findings indicate that thermal parameters affect residual stresses substantially. Thus, by adjusting the kinematic parameters and reducing maximum reached local temperatures within the part, more favorable stress states of the finished component can be obtained. The attained knowledge is essential for the development of high-quality deposits and the selection of the best strategies for repair and additive manufacturing applications.

**Keywords** – Cold spray, residual stress, kinematic parameters, thermal property, repair

## I. INTRODUCTION

Cold spraying, also known as cold gas dynamic spraying, is a solid-state coating and additive manufacturing technique. In this process, micron-sized powder particles are accelerated to high velocities by an expanding gas stream in a converging diverging De-Laval type nozzle. The particles impact in the solid state onto the substrate at high velocity. Therefore, the formation of coatings occurs due to the kinetic energy of the particles. Extensive plastic deformation and related phenomena at the interfaces ensures bonding of the particles and allows for forming dense and internally well bonded coatings [1], [2]. Due to relatively low temperature in cold spraying, high oxide content, phase transformations, compositional changes and other problems associated with thermal spray methods are mostly alleviated [3], [4]. That makes cold spraying particularly suitable for spraying materials, which are sensitive to heat or oxidation such as Ti and its alloy [3]. With

these characteristics, cold spraying proves to be a suitable technique for processing of functional coatings. In addition, it also enables part repair and additive manufacturing of bulk components.

Cold gas spraying can be used in many sectors, including the aerospace, automotive, transportation, metal processing, electronics, marine, and ceramics and glass manufacturing industries [5]. Due to advantages by low temperature processing, cold gas spraying is currently in the focus of interest for repair and additive manufacturing applications. For both, several studies investigated the deposit quality and the effect of kinematic and process parameters on the deposit qualities for different combinations of metallic materials [6], [7], [8], [9].

In all applications the integrity of the entire component should be ensured. As one of the factors that influence the quality of the deposit, residual stresses can affect the adhesion, wear, fatigue life and overall performance of the cold sprayed components [4]. Therefore, it is essential to gain a better knowledge about the sources of residual stresses and the parameters that change their magnitude and states.

Residual stresses in cold sprayed components can be divided into two categories: 1) Mechanical, mainly compressive stresses that are generated by peening effects upon particle impact and consequent plastic deformation. 2) Thermal mainly tensile stresses that are generated by the thermal mismatch of substrate and coating due to different coefficients of thermal expansion (CTE), rapid cooling of the particles after impact (quenching stress), and/or temperature gradients in the multi-pass deposition process [4], [10], [11].

There are some studies in the literature that examine the effect of substrate materials, process and kinematic parameters on the development of residual stresses in cold sprayed components. Their findings show that the coating - substrate material combinations as well as the process parameters and kinematics influence the final residual stress states by thermal and mechanical properties and reached temperature distributions within the component [11], [12], [13]. According to Suhonen et al. the generated stresses in cold sprayed coatings can be either compressive or tensile, dependent on (i) the density of particles and their deformation behaviour upon impact, and (ii) the CTEs of the sprayed material and the substrate [11]. Schmitt et al. showed that slowing down the robot traverse speed could shift the typical compressive residual stresses to

tensile residual stresses for IN718 coatings on IN718 substrate [12]. Marzbanrad et al. have investigated the effect of substrate temperature and exposure time on residual stress formation in Al7075 coatings on AZ31B. They found that by controlling the heat input (nozzle speed), cooling rate (substrate heat transfer), and the number of colliding particles (feed rate), desirable tensile and/or compressive residual stress were produced in the substrate near the interface and at the coating surface. According to their results, the thermal mismatch between the Al7075 coating and AZ31B substrate caused tensile residual stress induced into the substrate [13].

However, there is no comprehensive study on residual stress evaluation in pure titanium deposits. Therefore, the present study investigates the effect of the substrate material and nozzle traverse speed (as kinematic parameter) on the residual stresses of titanium-grade 1 (Ti-Gr1) coatings and respectively obtained temperature developments over the components. The aim was to gain a deeper understanding about the sources and effective parameters on the magnitude and final state of residual stresses in titanium coatings. The findings should provide a guideline for tuning of residual stresses by adjusting the nozzle traverse speeds for using different coating - substrate combination.

There are different methods to measure residual stress, in this study the hole drilling method was used to analyse the residual stresses over the entire coating thickness [14], [15].

## II. MATERIALS AND METHODS

Feedstock powder of titanium-grade 1 (99.7%) from Eckart TLS GmbH, Germany was cold sprayed onto flat substrates ( $70 \times 50 \times 3 \text{ mm}^3$ ). Titanium-grade 2 (Ti-Gr2), steel AISI304, and commercially pure copper (CP) were selected as substrate materials. Powder particles were almost spherical in shape and the particle sizes were in the range of 32-45  $\mu\text{m}$ . Cold gas spraying was performed by using a system of type 5/11 from Impact Innovation GmbH, Germany. On each substrate, an area of  $50 \times 50 \text{ mm}^2$  was coated with Ti-Gr1 in thickness of about 1 mm. The process parameters and the materials specific data considered in this study are shown in Table I and II, respectively. In order to avoid additional stresses, the substrates were only cleaned and not pretreated prior to the coating process. Different heat input and thus temperatures during the spray process were attained by variation of the kinematic parameters via applying different nozzle traverse speed of 125 mm/s, 250 mm/s and 500 mm/s, respectively. To reach the same coating thickness of about 1 mm in all cases, the number of layers differ for each traverse speed (see Table II).

TABLE I: COLD SPRAY PROCESS PARAMETERS

Gas temperature ( $^{\circ}\text{C}$ )	Gas pressure (bar)	Line distance (mm)	Standoff distance (mm)
1100	45	2	40

TABLE III: MATERIAL PROPERTIES USED IN THIS STUDY (MATWEB.COM)

Material	Conductivity ( $\text{W/mk}$ )	Thermal Expansion ( $\mu\text{m/m-}^{\circ}\text{C}$ )	Density ( $\text{kg/m}^3$ )	Tensile Strength (MPa)
Titanium-Gr 1	16	9.7	4500	240
Titanium-Gr 2	16.4	9.7	4500	344
CP copper	385	20.2	8900	260
Steel- AISI304	16	18.9	8000	564

TABLE III: INVESTIGATED VARIABLES

Substrate materials	Titanium Grade 2		
	Steel AISI304		
	CP copper		
Nozzle traverse speed (mm/s)	125	250	500
Number of layers	2	4	8

The experimental analyses concerned thermal and residual stress measurements. The thermal history of the components over the duration of the coating process was measured by inserting a thermocouple into a 1 mm diameter hole, drilled into the substrate from backside to a depth of about 2 mm, as shown in Figure 1 [4]. Residual stresses through the depth of the coating was measured by incremental hole-drilling method (ASTM E837 standard), by introducing a blind hole in certain number of drilling steps (increments) and recording the strain relaxation after each step. The recorded strain is then converted into stress using a specific evaluation method [15]. Instead of standard centre hole-drilling, for present experiments an orbital drilling technique was employed [16]. An in-house hole-drilling device at the Karlsruhe institute of technology (KIT), which has a high-speed air turbine facilitated by an orbital motion, was used for incremental hole-drilling.



FIGURES 1: PLACEMENT OF THERMOCOUPLE FROM THE BACKSIDE OF THE SUBSTRATE. A DEPTH OF APPROX. 2 MM, APPROXIMATELY REFERS TO A DISTANCE OF 1MM TO THE SUBSTRATE SURFACE

## III. RESULTS AND DISCUSSION

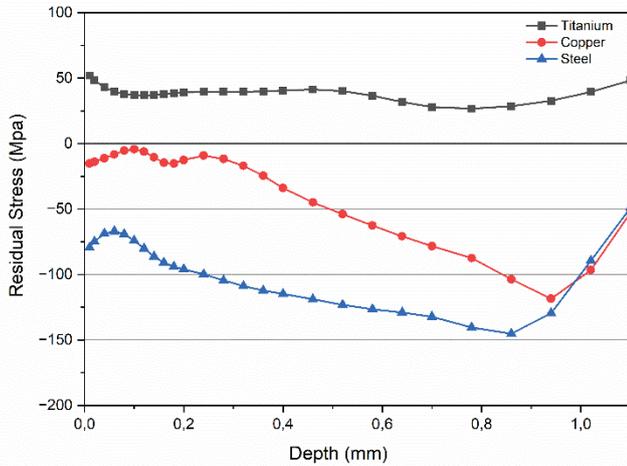
### A. Effect of Substrate Material on Residual Stress

Figure 2 displays the residual stresses throughout the thickness of Ti-Gr1 coatings on the different substrate materials according to Table II by using a nozzle traverse speed of 500 mm/s. Covering a depth of approximately 1.1 mm, the graphs show that the residual stresses are tensile in case of the titanium substrate and compressive in case of the steel and copper substrates.

The behaviour of residual stress of Ti-Gr1 coatings on the steel and copper substrate, could be explained by different coefficient of thermal expansion (CTE) between coating and substrate materials. The difference in CTE of the coating and substrate materials induces thermal stresses due to thermal misfit strain [4], [10], [11]. While the CTE of the substrate is greater than that of the coating, it means that the substrate expands and contracts more than the coating during the heating and cooling phases, respectively. This results in major thermal misfit strain and causes compressive thermal stresses dominating the ultimate residual stress state.

Despite the almost similar CTE, the coating stress levels on steel and copper substrates are different, which may be attributed to the mechanical and thermal properties of steel and copper. The lower thermal conductivity of the steel can cause

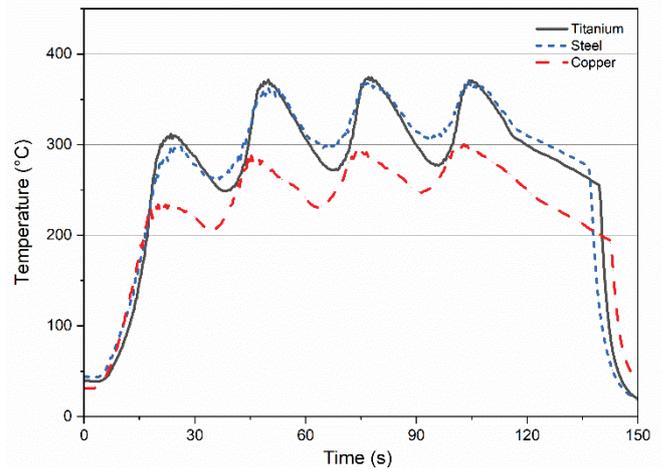
a steeper temperature gradient. On the other hand, the lower stiffness of copper may cause more bending of the coating-substrate bimetal than in case of the stiffer steel substrate. As a result, more stresses could be released. Associated different stress states during deposition can lead to higher compressive residual stresses in the overall Ti-Gr1 coating (according to Figure 2).



FIGURES 2: RESIDUAL STRESS THROUGHOUT THE THICKNESS OF TI GRADE 1 COATINGS ON DIFFERENT SUBSTRATE MATERIALS AS PROCESSED WITH A NOZZLE TRAVERSE SPEED OF 500 MM/S

In the case of the Ti-Gr1 coating on Ti-Gr2 substrate, the absence of CTE differences between the coating and the substrate materials avoids contributions from thermal misfit strains onto the stress development. Despite the expected compressive stress due to the peening effect of particles impact, the stress state in this case is tensile. This can be explained by temperature gradients over coating and substrate, and their history over the duration of the spray process. According to Figure 3, the overall temperature in case of the titanium substrate is higher than that by using a copper substrate. The lower thermal conductivity of Ti as compared to Cu restricts fast heat diffusion and keeps the coating itself and the substrate volumes close to the interface at higher temperatures. Less steep temperature gradients may be the reason for the tensile residual stresses in the case of the titanium substrate that even overcompensate the peening effect by the impacting particles. According to Boruah et al., the residual stresses in cold gas sprayed Ti-6Al-4V coatings on Ti-6Al-4V substrate deposited with similar process parameters (1100 °C gas temperature and 50 bar gas pressure) were also found to be tensile in the coatings [17].

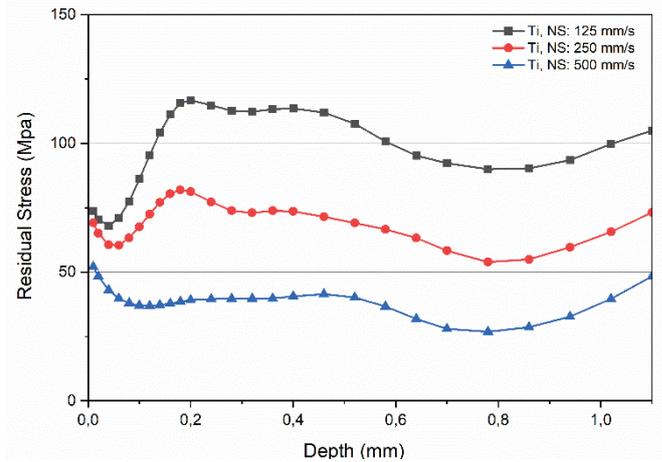
Apart from overall stress states, also differences over the depth within the coatings can be distinguished. The residual stresses of the Ti-Gr1 coating on copper and steel substrates show a gradient and get more compressive near the interface between coating and substrate. This could contribute to peening effects by particle impact being less compensated by tensile stresses due to thermal effects. During build-up of the first layers, the cold substrate as effective heat sink ensures fast heat extraction and thus lower coating temperatures.



FIGURES 3: TEMPERATURE DEVELOPMENT WITHIN THE SUBSTRATES AS OBTAINED BY APPLYING A NOZZLE TRAVERSE SPEED OF 500 MM/S DURING SPRAYING

### B. Effect of Nozzle Traverse Speed on Residual Stresses

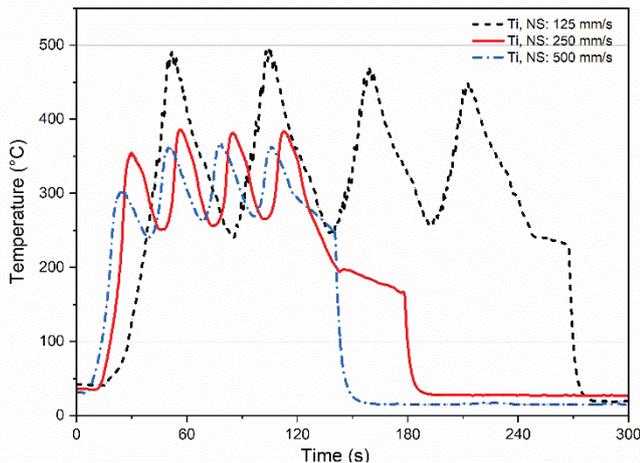
Figure 4 shows the effect of nozzle traverse speed on the residual stresses of Ti-Gr1 coating on Ti-Gr2 substrate. As the nozzle traverse speed decreases, the residual stress in the coating tends to more tensile states. A nozzle traverse speed of 500 mm/s causes the lowest tensile stress and nozzle traverse speed of 125 mm/s causes the highest tensile stress. These results correlate to the temperatures reached during deposition.



FIGURES 4: RESIDUAL STRESS THROUGHOUT THICKNESS OF TI GRADE 1 COATING ON TI GRADE 2 SUBSTRATE BY APPLYING NOZZLE TRAVERSE SPEEDS OF 125 MM/S, 250 MM/S, AND 500 MM/S

The thermal history of the parts during deposition (Figure 5) show that by decreasing the nozzle traverse speed, the temperature in the substrate close underneath the coating increases with decreasing traverse speeds. That is due to a longer duration of local heating on the component (increasing the thermal input). This indicates that a lower traverse speed increases local temperature and enhances the contribution of quenching effect by rapid cooling of the particles after impact and shifts the residual stress distributions towards a more tensile mode. This is consistent with reports from Schmitt et al. for NI718 coatings [12] and Marzbanrad et al. for Al7075 coatings [13].

According to this finding, it is important to adjust and find the optimum nozzle traverse speed to achieve the best coating quality in terms of residual stress states.



FIGURES 5: THERMAL HISTORY OF Ti-Gr2 SUBSTRATES BY APPLYING NOZZLE TRAVERSE SPEEDS OF 125 MM/S, 250 MM/S, AND 500 MM/S

### C. General

In this study, the effect of substrate materials and nozzle traverse speed were investigated. It can be assumed that impacts in the solid state at first instance contribute to compressive peening stresses. These peening stresses are to different extent superimposed by tensile thermal stresses. The heat input per unit area in combination with thermal and heat transfer parameters determines the effective surface temperature and temperature gradients during deposition and in consequence the thermal history and stress development within the component during cooling. Hence, having more knowledge about thermal related parameters and effective kinematic parameters on temperature and resulting stress distributions within the cold sprayed components is crucial for reaching optimum part performance.

The levels of thermal stresses depend on substrate material types and applied kinematics. The choice of substrate material by CTE mismatch and different temperature gradients influence the contributions by thermal tensile stresses. Tensile contributions scale with temperatures attained just underneath the coatings within the substrate.

Apart from substrate material, also the choice of kinematics determines coating stress states following similar rules. Higher locally reached temperatures close to the interface by lower travers speeds results in higher contributions by tensile stresses. Thus, these primary investigations allow supplying a first guideline to tune coating stresses by thermal management as attainable by kinematics during deposition.

## IV. SUMMARY AND CONCLUSION

This study investigates the residual stresses of cold sprayed Ti-Gr1 coatings and respective influences by substrate materials and local heat input by different nozzle traverse speeds as a kinematic parameter. The main conclusion of this study is summarized as follows:

- Solid state impacts by cold spraying in first instance cause compressive stresses by peening effects. The overall residual stresses are determined by the superposition of these compressive stresses with tensile stresses due to local temperature rises under the spray jets and respective temperature gradients.

- The tensile stresses can overcompensate the compressive stresses and thus determine the overall stress states.
- The type of substrate material determines tensile stress levels by the mismatch of thermal expansion coefficients with respect to the coating material and the thermal effusivity. A higher thermal effusivity and thus more efficient heat extraction reduces coating and surface temperatures and thus thermal stresses.
- Tensile stress contributions depend on applied kinematics during deposition. A higher nozzle travers speed leads to a shorter heating duration and thus lower locally reached temperatures during one pass. The lower coating temperatures result in reduced tensile residual stresses.
- The tensile stress contributions and thus the overall coating stress states can be influenced by thermal management during deposition. Thus, robot kinematics proves as key parameter to adjust the needed coating and part stress levels.

## ACKNOWLEDGEMENT

The authors acknowledge the financial support in the frame of the project "CORE – Computerized Refurbishment" funded by dtec.bw - Digitalization and Technology Research Center of the Bundeswehr. Digitalization and Technology Research Center of the Bundeswehr is a scientific centre jointly supported by the Universities of the Federal Armed Forces in Hamburg and Munich and is part of the German government's economic stimulus program to overcome the COVID-19 crisis. It is subject to academic self-administration. The funds within dtec.bw for both universities of the Federal Armed Forces finance research projects and projects for knowledge and technology transfer. dtec.bw is funded by the European Union – NextGenerationEU.

## REFERENCES

- [1] H. Assadi, F. Gärtner, T. Stoltenhoff, and H. Kreye, "Bonding Mechanism in Cold Gas Spraying", *Acta Materialia*, vol. 51, pp. 4379–4394, 2003.
- [2] T. Schmidt, F. Gärtner, H. Assadi, and H. Kreye, "Development of a generalized parameter window for cold spray deposition." *Acta Materialia*, vol. 54(3), pp. 729-742, 2006.
- [3] T. Stoltenhoff, H. Kreye, and H.J. Richter, "An analysis of the cold spray process and its coatings", *Journal of Thermal Spray Technology*, vol. 11(4), pp. 542-550, 2002.
- [4] Z. Arabgol, H. Assadi, T. Schmidt, F. Gärtner, and T. Klassen, "Analysis of thermal history and residual stress in cold sprayed coatings", *Journal of Thermal Spray Technology*, vol. 23, pp. 84-90, 2014
- [5] "Cold spray coatings, recent trends and future perspective", *Springer International Publishing AG*, pp 3-24, 2018.
- [6] S. Pathak and G. C. Saha, "Development of sustainable cold spray, coatings and 3D additive manufacturing components for repair/manufacturing applications: A Critical Review", *Coatings*, vol. 7(8)122, 2017.
- [7] F. Lang, J-C. Schmitt, S. Cabeza, T. Pirling, J. Fiebig, R. Vassen, and J. Gibmeier, "IN718 cold gas repair spray of large cavities-microstructure and residual stresses", *Proceedings of the 10th International Symposium on Superalloy 718 and Derivatives*, pp. 739-753, 2023.

- [8] S. Yin, P. Cavaliere, B. Aldwell, R. Jenkins, H. Liao, W. Li, and R. Lupoi, "Cold spray additive manufacturing and repair: Fundamentals and applications," *Additive Manufacturing*, vol. 21, pp. 628–650, 2018.
- [9] M. Faccoli, G. Cornacchia, D. Maestrini, G.P. Marconi, and R. Roberti, "Cold spray repair of martensitic stainless steel components", *Journal of Thermal Spray Technology*, vol. 23 (8), pp. 1270-1280, 2014.
- [10] D. Boruah, X. Zhang, and M. Doré, "Theoretical prediction of residual stresses induced by cold spray with experimental validation", *Multidiscipline Modeling in Materials and Structures*, ISSN: 1573-6105, 2019.
- [11] T. Suhonen, T. Varis, S. Dosta, M. Torrell, and J.M. Guilemany, " Residual stress development in cold sprayed Al, Cu and Ti coatings", *Acta Materialia*, vol 61, pp. 6329–6337, 2013.
- [12] J. Schmitt, J. Fiebig, S. Schrüfer, O. Guillon, and R. Vaßen, "Adjusting residual stresses during cold spray deposition of IN718", *Journal of Thermal Spray Technology*, vol 33, pp. 210-220, 2024.
- [13] B. Marzbanrad , E. Toyserkani , and H. Jahed, "Customization of residual stress induced in cold spray printing", *Journal of Materials Processing Technology*, vol. 289, 116928, 2021.
- [14] Y.Y. Santana, J.G. La Barbera-Sosa, M.H. Staia, J. Lesage, E.S. Puchi-Cabrera, D. Chicot, and E. Bemporad, "Measurement of residual stress in thermal spray coatings by the incremental hole drilling method, *Surface & Coatings Technology*, vol. 201, pp. 2090-2098, 2006.
- [15] E. Obelode and J. Gibmeier, "Residual stress analysis on thick film systems by the incremental hole-drilling method - Simulation and experimental results", *Experimental Mechanics*, vol. 53, pp. 965-976, 2013.
- [16] G. S. Schajer and P. S. Whitehead, "Hole drilling and ring coring, *Practical Residual Stress Measurement Methods*, chapter 2, 2013.
- [17] Boruah1, B. Ahmad, T. L. Lee, S. Kabra, A. Khadar Syed, P. McNutt, M. Doré, and X. Zhang, "Evaluation of residual stresses induced by cold spraying of Ti-6Al-4V on Ti-6Al-4V substrates", *Surface and Coatings Technology*, vol. 374, , pp. 591-602, 2019.

# CoupleIT!

## Coupling Energy Grids and Research Disciplines

Kazem Pourhossein\*, Edgar Diego Gomez Ancas†, Dalia Salem‡, Detlef Schulz  
*Electrical Power Systems  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany*

\*k.pourhossein@hsu-hh.de,

†diego.gomez@hsu-hh.de,

‡dalia.salem@hsu-hh.de

Simon Sassen\*, Andreas Fink  
*Economic and Social Sciences  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany*

\*simon.sassen@hsu-hh.de

Tim Rensmeyer\*, Oliver Niggemann  
*Computer Science in Mechanical  
Engineering  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany*

\*tim.rensmeyer@hsu-hh.de

Heiko Bornholdt\*, Mathias Fischer  
*Computer Networks  
University Hamburg  
Hamburg, Germany*

\*heiko.bornholdt@uni-hamburg.de

Rubina Steller\*, Petra Schweizer-Ries  
*Sustainable Development  
Bochum University of Applied Sciences  
Bochum, Germany*

\*rubina.steller@hs-bochum.de

*Abstract* – The dtec.bw project *CoupleIT! – IT-based sector coupling: Digitally controlled fuel cell and electrolyzer technologies for stationary and mobile applications* is an interdisciplinary approach to combine a wide range of competencies from disciplines as varied as electrical power systems, economic and social sciences, computer sciences and networks as well as sustainable development and social acceptance research. As such, this article is composed of individual contributions, constituting the main chapters that showcase general approaches and motivations but also concrete results.

This compendium article starts in with a delineation of the motivation behind research in so-called microgrids composed of fuel cell and electrolyzer components and a presentation of the microgrid architecture opted for in this project. Chapter two goes into more detail on the side of electrical engineering and the feasibility of a parallel operation of inverters in microgrids to achieve the ability for an upscaling. Chapter three highlights economic and technological factors for an economically viable and grid-maintaining deployment of a hydrogen-based energy system. In addition, degradation of Li-ion batteries is discussed against the background of their flexible operation in a microgrid and other scenarios. Chapter four grants a glimpse into the field of computer science and the possibility to use artificial intelligence and neural networks for a new way to simulate the behaviour of matter on atomic and molecular scales. This approach holds potential to increase the efficiency of fuel cells by improving the molecular design of fuel cell membranes used within this project. Chapter five elucidates the intricacies of secure communication within one but also between multiple microgrids, an important aspect for achieving a resilient system. Chapter six concludes this compendium by highlighting the human perspective seen from the field of psychological acceptance research nested in the broader context of sustainable development. Among other

things, areas of potential barriers to a public acceptance of hydrogen technology are identified and ways to overcome those barriers proposed.

This interdisciplinary round trip starts with electrical engineering (chapters one and two), economic and social sciences (chapter three), followed by computer sciences (chapter four) and computer networks (chapter five) whence the baton is passed for one last time to the field of sustainable development and psychological acceptance research (chapter six).

### CONTENTS

#### CHAPTER ONE

*PEM Fuel Cell-Based Microgrids*

by Kazem Pourhossein and Edgar Diego Gomez Ancas

#### CHAPTER TWO

*Parallel Operation of Inverters in Islanded Microgrids*

by Dalia Salem

#### CHAPTER THREE

*Economics of Hydrogen and Li-ion batteries in microgrids*

by Simon Sassen

#### CHAPTER FOUR

*Neural Networks in Molecular Simulations*

by Tim Rensmeyer

#### CHAPTER FIVE

*IT and Network Security in Microgrids*

by Heiko Bornholdt

#### CHAPTER SIX

*Shaping Hydrogen Transformation*

by Rubina Shirin Steller

## CHAPTER ONE

## PEM Fuel Cell-Based Microgrids: Advantages and Future Prospects

by Kazem Pourhossein and Edgar Diego Gomez Ancas

**Abstract** – Microgrids can provide reliable power supply by operating independently of the main grid during outages, ensuring continuous electricity for critical infrastructure and services. PEM (proton exchange membrane) fuel cell-based microgrids are gaining attention as a reliable, efficient, and environmentally friendly energy solution. In this paper, structural components, advantages, challenges, and future outlooks of PEM fuel cell microgrids are addressed.

**Keywords** – PEM fuel cells, microgrids, energy storage

## I. INTRODUCTION

Microgrids, as localized energy systems, have the potential to revolutionize the way we generate and distribute electricity. PEM fuel cell-based microgrids offer numerous benefits, including high efficiency, low emissions, and reliable power supply.

The concept of fuel cells dates back to the 19<sup>th</sup> century, but significant advancements in PEM fuel cell technology occurred in the mid-20<sup>th</sup> century, particularly during NASA's space programs. PEM fuel cells were used to provide clean and reliable power for space missions, proving their viability in demanding applications.

The application of PEM fuel cells in microgrids is a more recent development, driven by the increasing demand for clean energy solutions and advancements in fuel cell technology. Microgrids incorporating PEM fuel cells have been deployed in various sectors, including residential, commercial, and industrial, showcasing their versatility and effectiveness.

## II. STRUCTURE AND COMPONENTS OF PEM FUEL CELL-BASED MICROGRIDS

PEM fuel cells are the core power-generating units in these microgrids. They operate by converting hydrogen and oxygen into electricity, water, and heat through an electrochemical process. Key components of PEM fuel cells include the anode, cathode, electrolyte membrane, and catalyst layers.

Hydrogen supply is critical for the operation of PEM fuel cells whereas the necessary oxygen can be supplied by air. Hydrogen can be produced through various methods, including natural gas reforming and water electrolysis. Storage and distribution systems ensure a steady supply of hydrogen to the fuel cells.

Energy storage systems, such as batteries and supercapacitors, are integrated with PEM fuel cells to store excess energy for later use. These systems enhance the reliability and stability of the microgrid by providing backup power during peak demand or outages.

Power electronics, including inverters and converters, manage the flow of electricity between fuel cell units, storage systems, and loads. Advanced control systems monitor and optimize the operation of the microgrid, ensuring efficiency and stability.

A possible power electronics scheme for a laboratory-scale stand-alone PEM fuel cell system is depicted in Figure 1.

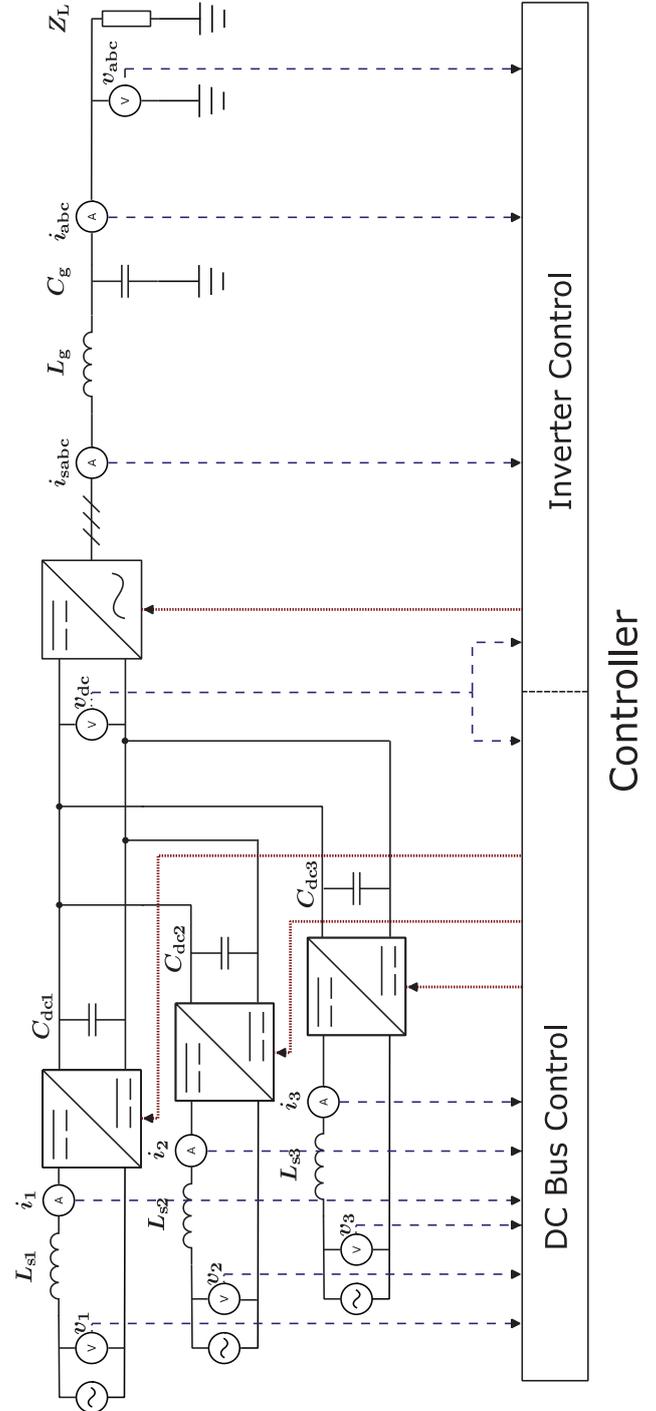


FIGURE 1: A LABORATORY-SCALE STANDALONE PEM FUEL CELL SYSTEM [1]

### III. ADVANTAGES OF PEM FUEL CELL-BASED MICROGRIDS

PEM fuel cells exhibit high electrical efficiency and can be integrated into Combined Heat and Power (CHP) systems, utilizing waste heat to further increase overall efficiency. When using pure hydrogen, PEM fuel cells produce only water and heat as byproducts, resulting in zero emissions of greenhouse gases and other pollutants. This makes them a clean and sustainable energy solution. PEM fuel cells provide a consistent and reliable power supply, with fewer moving parts leading to lower maintenance requirements as compared to more conventional systems based e.g. on internal combustion engines. Their durability and longevity make them a dependable choice for various applications. PEM fuel cell systems are modular and can be scaled to cover a broad spectrum of power demands, from residential to industrial applications. The operation of PEM fuel cells is silent, making them suitable for use in noise-sensitive environments such as residential areas, hospitals, and schools. By generating power locally, PEM fuel cell microgrids reduce dependence on centralized power plants and improve energy security. They enhance the resilience of the energy supply, particularly in remote areas or those prone to suffer from natural disasters. Performance analysis of existing projects highlights the efficiency, reliability, and environmental benefits of PEM fuel cell microgrids.

### IV. CHALLENGES AND LIMITATIONS

**A. Technical Challenges:** Technical challenges include hydrogen production and storage, durability, and lifespan of PEM fuel cells. Addressing these challenges requires ongoing research and development [2].

**B. Economic Considerations:** The cost of fuel cell technology remains a significant barrier to widespread adoption. Comparing the costs of PEM fuel cell microgrids to other energy solutions is essential for evaluating their economic viability.

**C. Regulatory and Policy Issues:** Current policies and regulations can either support or hinder the adoption of PEM fuel cell microgrids. Identifying and addressing regulatory barriers is crucial for promoting the growth of this technology.

### V. FUTURE OUTLOOK

**A. Technological Advancements:** Ongoing research and development are expected to improve efficiency, durability, and cost-effectiveness of PEM fuel cells. Innovations in fuel cell design and materials will drive technological progress.

**B. Market Trends and Predictions:** The market for PEM fuel cell microgrids is ready for growth, driven by an increasing demand for resilient and sustainable energy solutions. Emerging markets and applications will further expand the reach of this technology.

**C. Policy and Regulatory Support:** Governmental incentives and subsidies, as well as international collaborations and standards, will play a vital role in supporting the adoption of PEM fuel cell microgrids. Policy support is essential for overcoming economic and regulatory challenges.

### VI. CONCLUSION

PEM fuel cell-based microgrids represent a versatile and sustainable energy solution with significant advantages in efficiency, emissions, and reliability. With continued technological advancements and supportive policies, these systems will

play a significant role in the future energy landscape, contributing to a more resilient and cleaner power infrastructure.

### REFERENCES

- [1] E. D. Gomez Ancas, K. Pourhossein, D. Becker, and D. Schulz, "Detailed Controller Synthesis and Laboratory Verification of a Matching-Controlled Grid-Forming Inverter for Microgrid Applications", *Energies*, 2023, doi:10.3390/en16248079.
- [2] L. P. Van, K. D. Chi, and T. N. Duc, "Review of hydrogen technologies based microgrid: Energy management systems, challenges and future recommendations", *International Journal of Hydrogen Energy*, 2023, doi: 10.1016/j.ijhydene.2022.12.345.

## CHAPTER TWO

# Parallel Operation of Inverters in Islanded Microgrids

by Dalia Salem

**Abstract** – An important goal of the project is to use the gained insights to engage in the modelling of an upscaled and extended version of the microgrid. After successfully initiating, operating, and analysing the single-sourced microgrid in steady-state and dynamical operation, the next step includes the stable operation of multiple sources connected in parallel. In the here applied scenario, two inverters connected in parallel in an islanded microgrid are investigated. The main inverter is generalized droop-controlled and grid-forming, while the second inverter is phase-locked, loop-based, and grid-following. Both are operated isolated from the main grid and supply a local load.

**Keywords** – *Grid-forming inverter, grid-following inverter, (generalized) droop control, phase-locked loop*

### NOMENCLATURE

GFM	Grid-forming
GFL	Grid-following
$V_q$	Quadrature Voltage Component
PI	Proportional Integral
PLL	Phase-locked loop

### I. INTRODUCTION

The first experiment consists in establishing a microgrid with a single source feeding a local load. The microgrid can be operated either in islanded mode, i.e. with no support from the main grid, or in a mode where it follows the main grid and benefits from the latter's grid-supporting capabilities. The islanded mode was investigated first, and a grid-forming (GFM) inverter was set up that supplied a load bank. This system was tested in two operating conditions: steady-state and dynamic operation. The latter featured sudden load jump and load shedding as well as load imbalances in the three phases. The second experiment looked at two inverters connected in parallel. Their synchronization was evaluated and the interaction between the two controllers was analysed.

### II. FIRST EXPERIMENT: GRID-FORMING INVERTER IN AN ISLANDED MICROGRID

The laboratory experiments illustrate a solid operation of the microgrid under steady-state conditions, but it collapses upon sudden changes in operational condition. During steady-state operation, the microgrid generates and sets voltage, frequency, and phase angle of the coupling point even in isolation from the main grid. The same control structure can be implemented when the microgrid is connected to the main grid, in which case the inverter follows and supports the main grid.

#### A. Grid-Forming Inverter and Droop Control Technique

The experiment was about the ability of the GFM inverter to operate independently from the main grid. The inverter generates and sets voltage and frequency of the microgrid within

the accepted boundaries [1], when the microgrid is autonomously operated. The implemented controller has to take into account the technical limitations of the inverter, while being able to detect islanded and grid-connected mode and adapt its performance accordingly. The capabilities of the controller have to include black start, DC offset isolation, post-fault recovery, and parallel operation of inverters.

### III. SECOND EXPERIMENT: PARALLEL OPERATION OF GRID-FORMING AND GRID-FOLLOWING INVERTERS IN AN ISLANDED MICROGRID

#### A. Generalized Droop Control Technique

An inverter governed by a generalized droop control is more robust to sudden changes and disturbances than its counterpart under a classic droop control regime [1]. The former leverages the flexibility of a grid-forming inverter for parallel connection to other inverters in an islanded microgrid and allows of more individual control over each single phase of the system with the required state feedback control.

#### B. Grid-Following Inverter and Phase-Locked Loop-based Controller

A grid-following (GFL) inverter functioning as a current source, requires a reference voltage source to establish its terminal voltage for operation. The implemented phase-locked loop (PLL) control achieves rapid phase angle detection. The proportional integral (PI) controller within the PLL drives the quadrature voltage component ( $V_q$ ), thereby providing the angular speed values required to latch and lock onto the phase angle of the voltage vector at the point of common coupling (PCC) thus achieving synchronization.

### IV. SYNCHRONIZATION

Examining synchronization is a fundamental step for understanding the dynamic behaviour of microgrids. One key metric to judge the stability of a grid is its frequency. According to [2], the frequency of systems with no synchronous connection to a broader grid, e.g. supply systems on islands, shall stay within the following boundaries: 50 Hz  $\pm$  2 % (i.e. 49 Hz... 51 Hz) during 95% of a week; 50 Hz  $\pm$  15 % (i.e. 42.5 Hz... 57.5 Hz) during 100% of the time". Experimental parallel operation is carried out according to IEEE Std 2030.8-2018 [3].

Figure 1 illustrates a comparison between two conditions, the unsynchronized connection between the two inverters when connected in parallel compared to a healthy synchronized parallel connection achieved through harmonized interconnection between the two controllers.

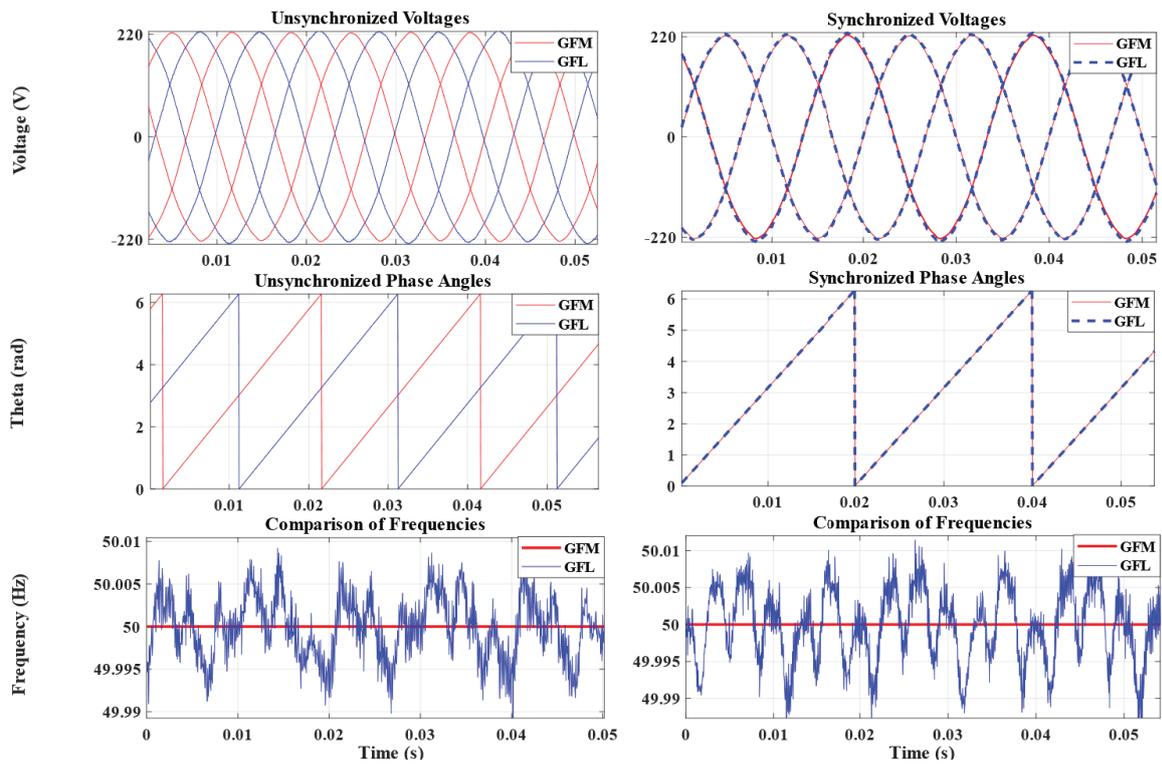


FIGURE 1: COMPARISON OF UNSYNCHRONIZED (LEFT) AND SYNCHRONIZED (RIGHT) OPERATION OF VOLTAGE (TOP), PHASE ANGLE THETA (MIDDLE), AND FREQUENCY (BOTTOM) OVER A PERIOD OF 50 MS

#### REFERENCES

- [1] P. Bhagwat, D. Groß, “Three-phase grid-forming droop control for unbalanced systems and fault ride through”, 28 July 2023
- [2] S. Fletcher, T. Bransden, J. Devlin, “Locked-Loop Based Linear Synchronization Scheme for Islanded Inverters and Microgrids”, FRAZER-NASH consultancy, Final Issue 2020.
- [3] “IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces”, IEEE Std 1547-2018, 6 April 2018.

## CHAPTER THREE

Optimal Operation of Li-ion Batteries in Microgrids  
and other Scenarios

by Simon Sassen

**Abstract** – The flexibility in energy provision and consumption that lithium-ion batteries and hydrogen-based energy systems offer, plays a key role in the widespread roll-out of renewable energies. Economic optimization of hydrogen systems must consider technical as well as economic objectives. In this contribution, three aspects of hydrogen systems are elucidated. First, the use of hydrogen in a household is investigated. The aim is to reduce the end consumer costs for electricity by converting hydrogen into electricity using a flexible price tariff. The used method is a linear optimization model for energy storage scheduling. Second, the further goal to reduce peak loads is added to the cost reduction incentive. To this end, the model of the first article is extended to include shiftable loads. Li-ion batteries and hydrogen systems consisting of electrolyser and fuel cell can act as consumers and producers of electricity. The employed method is a bicriteria mixed-integer linear optimization model. It has been shown that costs can be reduced when load smoothing is a priority. Third, an optimal operational management of Li-ion batteries is pursued. The key factors to find a trade-off between drawbacks from battery degradation and benefits from flexible operation are defined. A sequential decision problem is proposed to solve the non-linear characterization of degradation.

**Keywords** – Microgrid, energy scheduling, demand side management, peak load reduction, cyclic and calendar battery ageing, state of health

## NOMENCLATURE

PV	Photovoltaic
RE	Renewable energy
LP	Linear program
MILP	Mixed-integer linear program
NMC	Nickel manganese cobalt battery
SoH	State of Health
SoC	State of charge
DoC	Depth of cycle

## I. OPTIMAL ENERGY SCHEDULING FOR A MICROGRID WITH BATTERY AND HYDROGEN STORAGE

In [1] we investigate the suitability of a hydrogen system from the perspective of an end consumer. Although the feed-in tariff decreases from year to year, a photo voltaic (PV) system, despite its significant investment costs, is often still a viable option to reduce the electricity consumption from the public grid and thus the electricity costs to the end consumer. However, due to the fluctuation of renewable energy (RE), energy storage systems are necessary. However, while energy storage systems increase the degree of self-sufficiency, they do also increase the electricity generation costs. For example, the use of hydrogen as an energy storage medium by operating fuel cell (gas-to-power) and electrolysis units (power-to-gas)

introduces unavoidable energy conversion losses. Nonetheless, legislation that creates tax reduction or elimination schemes can help to compensate for monetary drawbacks in the case that hydrogen is converted back into electricity.

In here, we delineate a linear program (LP) that optimizes the deployment of battery, electrolysis and fuel cell units for within the frame of the fluctuating EPEX Spot exchange (see the example in Figure 1) [1]. It has been shown that a hydrogen storage system is usually not a profitable investment at least for a private household. Despite tax breaks for hydrogen reconversion into electricity (fuel cell), flexible trading on the electricity exchange and the increased degree of self-sufficiency mostly do not make up for the investment costs for a household. Instead, it is more worthwhile to sell the hydrogen produced directly instead of reconverting it into electricity [1].

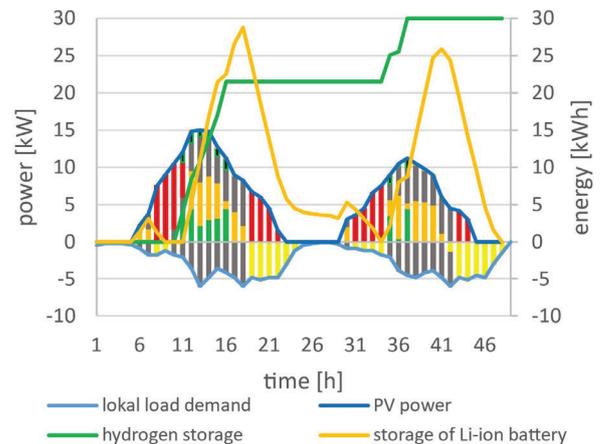


FIGURE 1: ENERGY SYSTEM WITH DEPLOYMENT PLANNING OF THE HYDROGEN SYSTEM AND THE LI-ION BATTERY [1]

## II. LOAD SMOOTHING OF A MICROGRID WITH TWO STORAGE SYSTEMS AND SHIFTABLE LOADS

In [2] we study the impact of the use of energy storage and demand-side management (DSM) on instabilities in a microgrid. These instabilities are defined here as the imbalance of supply and demand of electricity. The energy storage units are a battery and a hydrogen system, which can be used as electricity consumers or producers at any one time. The DSM is realized by shiftable loads, whereby a distinction must be made between interruptible and non-interruptible. The proposed mixed-integer linear program (MILP) optimization model addresses the scheduling of shiftable loads and energy storage. The objective is to minimize a weighted sum of the two target variables profit and imbalances. Since these are often conflicting goals, a factor  $a$  is used to determine how strongly profit-oriented and how strongly imbalance-minimizing the result should be. In order to reduce load peaks, the imbalances were additionally assessed with a penalization factor.

The higher the imbalance, the higher this factor. The result is a comparison between two scenarios. The first result is yield-oriented without shiftable loads. The result for the second scenario is imbalance-oriented with shiftable loads [2]. Figure 2

shows the deployment plan for electricity consumers and producers for the second scenario. It has been shown that in the second scenario, the maximum imbalances and costs are reduced slightly, while the minimum imbalances and the sum of the amounts of all imbalances are reduced considerably [2].

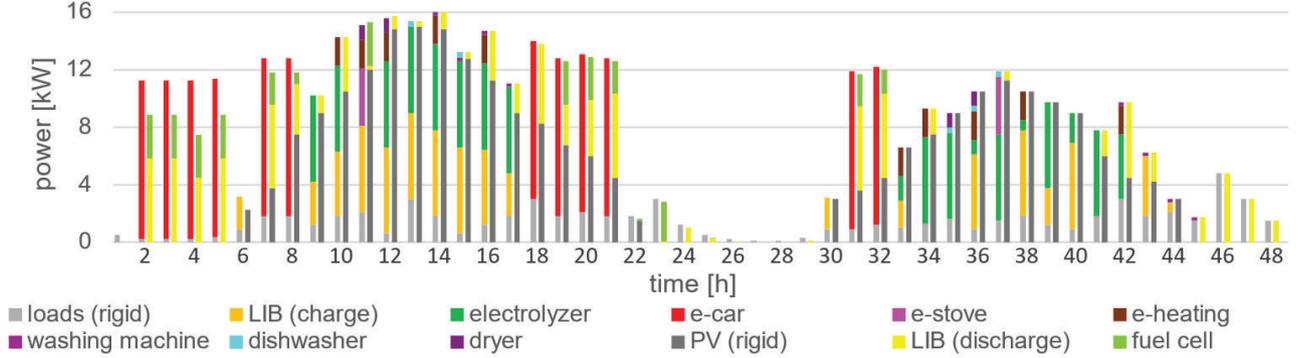


FIGURE 2: ARRANGEMENT OF RIGID AND FLEXIBLE ELECTRICITY PRODUCERS AND CONSUMERS [2]

### III. AN OUTLOOK INTO OPTIMAL DIMENSIONING OF A LI-ION BATTERY TAKING INTO ACCOUNT CALENDER AND CYCLIC DEGRADATION EFFECTS DURING DYNAMIC OPERATION

The initial cost of an energy system has a significant impact on the profitability of the entire system. It is therefore important to improve the amortization of the components.

Batteries are subject to calendar and cyclic degradation effects, which reduce the usable capacity or ‘state of health’ (SoH) of the batteries. The level of cyclic degradation depends on many influencing factors, which vary depending on the battery's operating strategy. Since a battery is usually used until a certain SoH is reached, the maximum number of cycles, and therefore battery life, is highly dependent on the operating strategy. For example, low power usage relative to total capacity results in less degradation than high power usage. However, increasing the capacity in order to keep this ratio low also in high power usage scenarios results in higher acquisition costs and calendar degradation that affects the entire capacity from the start. In addition, it may make economic sense to accept increased degradation when exchange rates fluctuate.

Although the degradation of Li-ion batteries has already been examined in an economic context in the literature, there is still a need for further research. The effects of dynamic operation of the battery on overall degradation, considering the large number of influencing factors, have not yet been sufficiently recorded in an economic context. The aim of ongoing work is to model and optimize the dynamic operation of batteries with different capacities. Both calendar and cyclic degradation effects, fluctuating electricity market prices, and optimal battery sizing are to be considered. The non-linearity of the degradation and the uncertainties are formulated and optimized in a sequential decision problem.

The used battery degradation model is based on [3]. According to this model, the total degradation  $D_{\text{cell}}$  is constituted by calendar and cyclic degradation (formula (1)). The calendar degradation is described as a function of the temperature ( $Temp$ ), the state of charge ( $SoC$ ) and the time ( $t$ ) (formula (2)). The cyclic degradation depends on the mean SoC ( $M_{SoC}$ ), the depth of cycle ( $DoC$ ), the  $Crate$  and the energy flow

( $AH_{\text{thp}}^y(N)$ ) as a function of the number of cycles performed ( $N$ ) (formula (3)) [3].

$$D_{\text{cell}} = D_{\text{cal}}(Temp, SoC, t) + D_{\text{cyc}}(DoC, M_{SoC}, Crate, N) \quad (1)$$

$$D_{\text{cal}}(Temp, SoC, t) = \alpha(Temp, SoC) * t^{0.5} \quad (2)$$

$$D_{\text{cyc}}(DoC, M_{SoC}, Crate, N) = \beta(DoC, M_{SoC}) * EXP\left(\frac{\rho_{AH}}{112} * Crate\right) * AH_{\text{thp}}^y(N) \quad (3)$$

In a first step, the degradation functions  $\alpha(Temp, SoC)$  and  $\beta(DoC, M_{SoC})$  of the nickel manganese cobalt battery (NMC) are calculated. In the next step, formulas (2) and (3) are expressed as a function of time and only of the charging, discharging power, and capacity during a period to enable scheduling. These formulas are to be employed in an optimization approach that examines efficient battery usage patterns for different application scenarios.

### REFERENCES

- [1] S. Sassen and A. Fink, Optimal Energy Scheduling for a Microgrid with Battery and Hydrogen Storage, Hamburg: in Neis 2022, Conference on Sustainable Energy Supply and Energy Storage Systems, 2022.
- [2] S. Sassen and A. Fink, Lastenglättung eines Microgrids mit zwei Speichersystemen und verschiebbaren Lasten, Hamburg: in TKB 2022, Hamburger Beiträge zum technischen Klimaschutz, 2022.
- [3] F. Alhaider, T. Klein and S. Gerhard, A Simplified Methodology to Quickly Build a Degradation Prediction Model for Lithium-Ion Batteries, Bremen: in NEIS 2018, 2018.

## CHAPTER FOUR

## Towards Practical Neural Network Accelerated High Accuracy Molecular Simulations at Scale

by Tim Rensmeyer

**Abstract** – Due to the computational complexity involved in first-principles simulations of molecular and material dynamics, the prediction of interatomic forces with neural networks has become an active area of research. An important open question in this domain is how to efficiently generate training datasets to train such neural networks in the first place. As a solution to this, we here propose an approach that utilizes Bayesian statistics and on-the-fly fine-tuning of pre-trained models.

**Keywords** – Molecular dynamics, neural networks, Bayesian learning

## I. INTRODUCTION

Even though the prediction of chemical and material properties such as catalytic activity from first principles has theoretically been possible since the discovery of the equations of quantum mechanics in the 1920s, it remains a challenging task to this day due to the computational complexity involved in solving these equations numerically [1]. This makes the prediction of the time evolution of molecular systems – Molecular Dynamics (MD) – challenging from first principles since the forces acting on each atom have to be calculated at each time step. Consequently, even with the workhorse quantum chemical simulation method Density Functional Theory (DFT), simulations are only possible with a limited time horizon and system size [2]. As a response to these challenges substitutional modelling of molecular dynamics via neural networks has become an active area of research [3]. State-of-the-art neural network architectures have by now achieved unprecedented accuracy and data efficiency in the prediction of atomic forces and energies on small molecule benchmarks [4] where the training data has been subsampled from molecular dynamics simulations that are temporally long enough for the molecule to have traversed its configuration space exhaustively. However, while such benchmarks are very useful for the determination of suitable neural network architectures, they do not represent a realistic scenario of how a neural network-based force field would be used in practice by computational chemists or material scientists. The reason for this is that if it is possible to simulate such an exhaustively long MD trajectory with classical quantum chemical simulations, the neural network model would not be needed in the first place. The application scenarios where a substitutional neural network model would be the most valuable are those where classical simulations do not allow for a sufficient time horizon to be simulated and where it is not known beforehand, what atomic configurations the system might traverse.

## II. SOLUTION APPROACH

To make neural network accelerated molecular dynamics simulations of fuel cell components (i.e. polyelectrolyte membranes, catalysts) possible, the professorship of computer science in mechanical engineering at the HSU Hamburg is developing a Bayesian neural network framework to integrate large-

scale public materials databases, uncertainty quantification and classical quantum chemical simulations into an efficient on-the-fly learning approach. We want to utilize the large-scale material databases of different molecular configurations with force labels to construct an informative Bayesian prior density over the weights of the parameters of the neural network. This prior density will be iteratively refined via Bayes' rule based on new data from the specific system that is supposed to be simulated. In this approach, the Bayesian neural network drives the molecular dynamics until an atomic arrangement is encountered for which the uncertainty in the prediction is above a certain threshold. At this point, a classical DFT simulation will be called into action to calculate the forces in that configuration. Based on these calculated forces, the Bayesian neural network will be updated and then resume driving the molecular dynamics simulation until again a configuration above the uncertainty threshold is met (Figure 1).

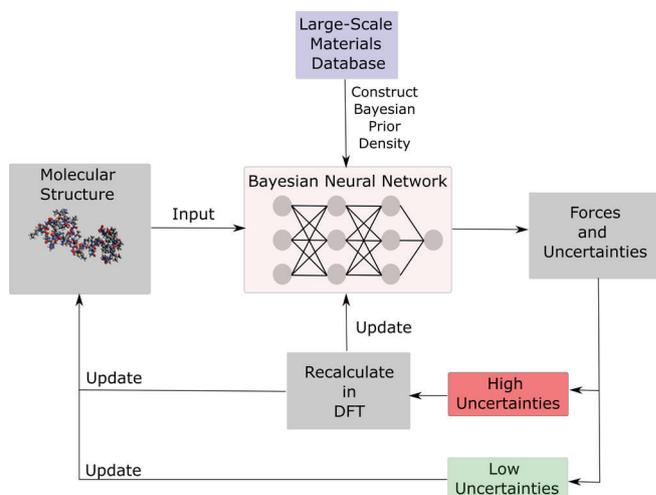


FIGURE 1: ILLUSTRATION OF THE ON-THE-FLY LEARNING APPROACH TO ACCELERATE MOLECULAR SIMULATIONS

The development of that framework requires three key steps.

First, a suitable Bayesian neural network model has to be created based on state-of-the-art neural network architectures. The practical viability of this network has to be verified with a focus on viability of sampling the Bayesian posterior, model accuracy, and quality of uncertainty quantification.

Secondly, a suitable Bayesian prior to incorporate large-scale databases has to be identified and empirically verified.

Lastly, the Bayesian neural network model and Bayesian prior have to be integrated into an on-the-fly learning workflow (Figure 1).

At his point, the first two steps have been achieved [5], while the last step is being worked on. In the future, we hope to utilize this learning framework to gain insights into the molecular processes of fuel cells at quantum mechanical accuracy.

#### REFERENCES

- [1] P. W. Atkins and R. S. Friedman, "Molecular Quantum Mechanics," *OUP Oxford*, 2011.
- [2] P. Marquetand and M. Gastegger, "Molecular dynamics with neural network potentials," *Machine Learning Meets Quantum Physics*, p. 233–252., 2020.
- [3] E. Kocer, W. K. Tsz and J. Behler, "Neural network potentials: A concise overview of methods," *Annual Review of Physical Chemistry*, p. 163–186, 2022.
- [4] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, Kozinsky and Boris, "E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials," *Nature Communications*, 2022.
- [5] T. Rensmeyer, W. Großmann, D. Kramer and O. Niggemann, "Bayesian Transfer Learning of Neural Network-Based Interatomic Force Models," *The 38th Annual AAAI Conference on Artificial Intelligence | Workshop on AI to Accelerate Science and Engineering*, 2023.

## CHAPTER FIVE

## Decentralized Security Monitoring as Enabler for Zero-Trust in Networked Microgrids

by Heiko Bornholdt

**Abstract** – The transition of power grids to an open, dynamic, and automated system necessitates more advanced security monitoring approaches. This paper proposes a decentralized security monitoring architecture for networked microgrids, implementing a zero-trust security model capable of detecting internal and external attacks. Furthermore, threat intelligence is shared among microgrids to increase the resilience of individual microgrids and the overall power grid.

**Keywords** – *Microgrid, cybersecurity, security monitoring, anomaly detection*

## I. INTRODUCTION

The ongoing opening of power grids, which results in increased participation and the need for more automation and management, not only presents opportunities for enhancing efficiency but also poses risks concerning grid availability.

Power grids transform from a closed static system, characterized by a few large electricity providers, into an open, dynamic, and automated system. This new system includes many decentralized energy sources, prosumers (entities that consume and produce energy), microgrids, and intelligent control technologies. The shift to an open system requires more sophisticated security monitoring approaches for microgrids [1]. Security monitoring refers to the continuous observation and analysis to detect suspicious behaviour or unauthorized changes in the system.

To this end, we define the following requirements for security monitoring in microgrids within modern power grids [2]: **high accuracy** for detecting and responding to threats and anomalies; **scalability** regarding the number of devices within a single microgrid and across multiple microgrids; and **decentralization**, since in an open system other participants cannot be trusted and centralized trusted authorities represent a single point of failure that must be avoided.

This paper proposes a decentralized security monitoring architecture for microgrids. By implementing a zero-trust security model, we make minimal assumptions about the trustworthiness of system entities, such as users, electric systems, or network components. We ensure that all entities and actions undergo strict authentication before processing. This approach enables the detection of threats originating from external sources and within the microgrid while also protecting the security monitoring system itself. Additionally, security monitoring information is shared across microgrids to enhance the overall system's ability to detect and respond dynamically to security threats and anomalies.

## II. SYSTEM ARCHITECTURE

Figure 1 depicts our architecture for a decentralized security monitoring of networked microgrids enabling zero-trust

capabilities. The power grid is a hierarchically organized network with large regional network operators (Figure 1 left box, top). Within an operator's network, microgrids are run by third parties (right box). A microgrid is divided into a power network (red arrows) and a communication network (green arrows). Depending on the specific grid configuration, the power network contains several electrical systems that provide, store, or consume energy. These individual systems are managed locally by a first-level controller or monitored by a smart meter. The communication network includes a router connected to the Internet, all first-level controllers, smart meters, and the second-level controller. The second-level controller orchestrates the microgrid by collecting information from and sending commands to first-level controllers and smart meters to optimize the grid's operation towards a defined goal, such as efficiency, reliability, or sustainability. The router's connection to the Internet allows the second-level controller to communicate with a third-level controller of the operator's grid for coordination. To enable zero-trust in this context, we add the following functions to the system.

**Network monitoring sensor:** This sensor is connected to the router on a monitoring port configured to send a copy of all traffic passing through the router. This enables the sensor to analyse data without impacting network performance. All information between first-, second-, and third-level controllers and smart meters is sensed and acts as the primary source of security monitoring. This also includes messages received from the Internet sent by the third-level controller or third parties, monitoring information from smart meters such as live consumption, commands from the first-level controller, and directives issued by the second-level controller. This information is then forwarded to the security data fabric for further analysis and action.

**Security data fabric:** The fabric centrally manages and analyses security data received from individual sources and provides a unified view and control. In this context, information is stored, cleaned, transformed, normalized, formatted, or combined with other information to handle missing data. These data wrangling and data engineering tasks are necessary for the next step, the security analysis.

**Security analysis:** Here, information from the fabric is evaluated to identify vulnerabilities and potential threats. This includes model inference, which applies trained machine learning models to real-time data to detect complex attack patterns and subtle anomalies beyond traditional methods. Anomaly detection uses statistical models and algorithms to identify unusual patterns in data that may signify security threats, providing critical early warnings. Signature-based detection identifies malicious activities by comparing network behaviour against a database of known threat signatures, effectively

recognizing established threats. Once data is analysed, the risk is assessed.

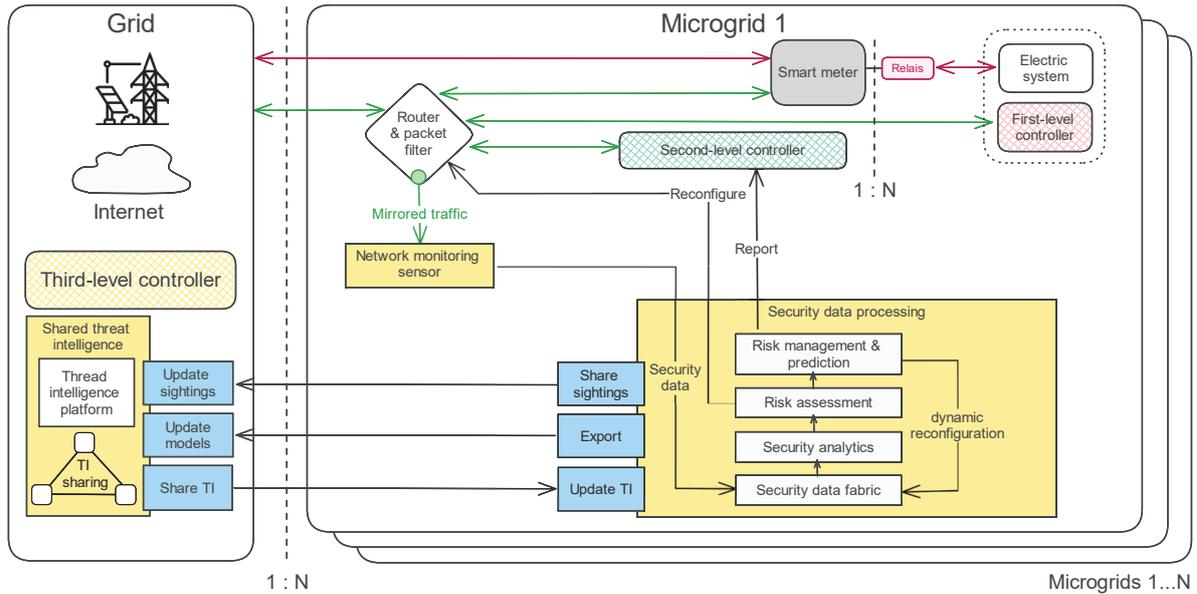


FIGURE 1: ARCHITECTURE OF DECENTRALIZED SECURITY MONITORING IN NETWORKED MICROGRIDS (TI = THREAT INTELLIGENCE)

**Risk assessment:** Risk assessment quantifies and prioritizes risks by evaluating the potential impact and likelihood of occurrence of identified threats. This process helps determine which vulnerabilities could cause the most damage, guiding the prioritization of mitigation efforts. Two measures can be taken in response to this task:

- The packet filter embedded in the router is reconfigured. This packet filtering can include blocking traffic from specific IP addresses, ports, or protocols identified as sources of malicious activity and permitting only traffic that meets established security policies. This filtering is conducted to prevent unauthorized access and potential attacks and to isolate vulnerable systems from others.
- A report is sent to the second-level controller, which includes information about the identified threat and the countermeasures conducted at the router's packet filter. Using this information, the controller can respond by counteracting on the power network side, for example, by shutting down affected systems and allowing unaffected systems to adapt to this change.

**Risk management and prediction:** Furthermore, risk management involves implementing strategies and timely measures to mitigate identified risks, continuously monitor their effectiveness, and adapt as necessary. Risk prediction is an aspect of this, forecasting potential future threats and vulnerabilities to address them proactively before they can impact the microgrid. This task may involve reconfiguring the data fabric, for example, to incorporate new protocols or update processing algorithms to improve the unified information view.

**Shared threat intelligence:** Individual microgrids will share their sightings of potential threats with other microgrids to further improve their threat detection accuracy, allowing them to protect against new attacks and proactively harden the overall grid. To facilitate this, the grid provides a platform for shared threat intelligence, where microgrids can provide their

data and update their local threat intelligence database. The zero-trust model helps by not blindly trusting potentially malicious or infected microgrids, as each microgrid's contribution to the platform is verified first.

### III. CHALLENGES AND OPPORTUNITIES

We present an architecture for distributed security monitoring across networked microgrids. Individual microgrids can share their sightings to a centralized threat intelligence platform. Shared information is authenticated and verified before being locally applied, enabling a zero-trust security model. In this context, providing an incentive model that prevents participants from becoming free riders and not contributing to the system will be challenging. In the future, we want to implement a proof of concept to show how our system performs in different scenarios, such as an attack on a single system, multiple systems, or multiple microgrids.

### REFERENCES

- [1] M. Uddin, H. Mo, D. Dong, S. Elsayah, J. Zhu, and J. M. Guerrero, "Microgrids: A review, outstanding issues and future trends," *Energy Strategy Reviews*, vol. 49, p. 101127, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2211467X23000779>
- [2] A. Dagar, P. Gupta, and V. Niranjana, "Microgrid protection: A comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 149, p. 111401, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032121006869>

## CHAPTER SIX

## Shaping Hydrogen Transformation: Socio-Psychological Insights in Hydrogen Acceptance

by Rubina Shirin Steller

**Abstract** – Achieving sustainable diffusion of hydrogen technology (HT) necessitates societal engagement and responsiveness to stakeholder concerns. The socio-psychological project at the Integrative Institute for Sustainable Development, part of the dtec.bw CoupleIT! project, emphasizes early stakeholder engagement and systematically analyses factors influencing hydrogen acceptance. Stakeholder interviews with industry leaders in Hamburg identify important perceived prerequisites, barriers, chances, risks, and institutional responsibilities in the context of a hypothetical project involving the implementation of smart integrated grids (SmInT-Grids). Representative surveys are employed to assess theoretical frameworks affecting HT acceptance. Our research underscores the necessity of validating acceptance measures, integrating participation variables, and broadening stakeholder engagement to enhance understanding and practical application of HT acceptance, ensuring a more inclusive and effective transition to sustainable energy systems.

**Keywords** – Hydrogen technology, acceptance, stakeholder interviews, socio-psychological factors, discourse analysis

## NOMENCLATURE

HT	Hydrogen technology
RET	Renewable energy technologies
SmInT	Smart-InTegrated

## I. INTRODUCTION

The energy transition requires solutions that are not only technically and economically feasible but also socially accepted. Innovations such as the development of a SmInT-Grid evolve within a complex economic, social, cultural, and political landscape, influenced by the diverse factors inherent in this environment [1]. Acceptance research examines the transition from CO<sub>2</sub>-producing systems to renewable energy technologies (RET), focusing on socio-political, community, and market dimensions. These dimensions capture various aspects of stakeholder acceptance, reflecting broader social climates, local community attitudes, and market behaviours [2]. Despite increasing research in this field, end-users are often viewed as obstacles rather than integral contributors to system design, overlooking the essential role of social acceptance dynamics in energy transitions [3].

Within the research project, we emphasize early stakeholder engagement. We systematically analyse socio-psychological factors influencing hydrogen technology (HT) acceptance and investigate how researchers perceive and discuss barriers, drivers, and implications of HT diffusion, highlighting their understanding, focus areas, and potential biases in acceptance research. Engaging stakeholders aims to understand

their perspectives and attitudes, creating a participatory platform for effective communication and deriving recommendations to support sustainable transformation processes.

## II. INSIGHTS INTO RESULTS

1) *A Systematic Review of Socio-Psychological Influencing Factors*

Our systematic literature review of 2023, spanning studies from 2004 to 2023 on acceptance research, reveals a predominant focus on upstream phases, with significantly less attention given to downstream research. Upstream research involves the early stages of technology development, including conceptualization and initial implementation, while downstream research pertains to later stages such as widespread adoption and post-implementation effects [4]. However, public perceptions and acceptance are critical throughout the entire technology lifecycle [5], underscoring the necessity for longitudinal studies to capture temporal variations in the acceptance of HT. Significant socio-psychological factors impacting HT acceptance were identified, with environmental attitudes being the most extensively studied. Social trust, showing a positive correlation with HT acceptance, was the second most researched factor. There is a notable interconnection between trust and risk perception, where higher perceived risk leads to diminished acceptance. Additionally, trust and participation in decision-making processes positively correlate with overall acceptance. Other influential factors include affect, knowledge, and social norms. However, our analysis indicates a focus on enhancing acceptance, potentially overshadowing the broader aim of understanding acceptance dynamics [6].

2) *A Stakeholder Analysis on SmInT-Grids*

A comprehensive stakeholder analysis is integral to our research project, involving online face-to-face interviews averaging 59 minutes. We interviewed heads of companies and research institutes, energy suppliers, network operators, and regulatory authorities in Hamburg. This analysis provides insights into stakeholders' perceptions regarding the integration of SmInT-Grids, a previously underexplored area in acceptance research. Our stakeholder analysis revealed various perspectives, of which we present a brief snapshot (due to pending peer-reviewed journal publications, more detailed results cannot be disclosed at this time). Key prerequisites for the potential implementation of a smart integrated grid include effective stakeholder management and knowledge exchange, ensuring investment security, gaining access to suitable construction sites, and establishing an efficient regulatory framework. Stakeholder management, from the perspectives of interview partners, involves regular meetings, information dissemination, and resource management. Transparency in project conditions and investment incentives from governmental bodies were perceived as essential for investment security.

Stakeholders were also asked to consider potential application areas for SmInT-Grids. They identified potential applications such as grid balancing, load peak management, and supplying areas with inadequate grid connections. The grids were considered promising for medium-term balancing, seasonal energy storage, and neighbourhood solutions. However, criticisms included the viability of island grids in urban areas with centralized power systems, high costs, and complexity in maintaining voltage and frequency stability.

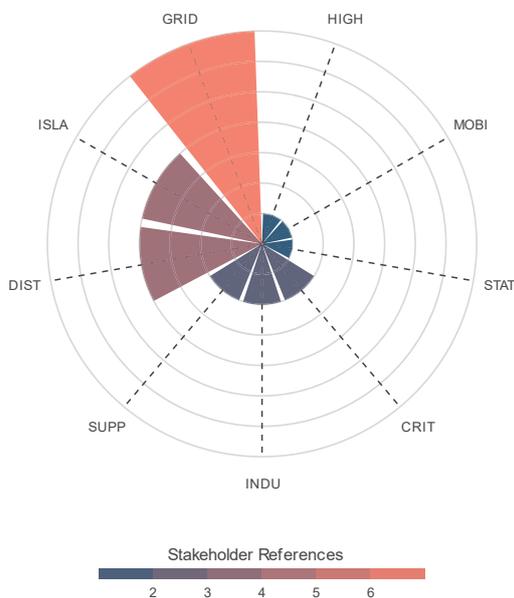


FIGURE 1: SUGGESTED APPLICATION AREAS OF *SMINT-GRIDS* BY STAKEHOLDERS. CRIT = CRITICAL INFRASTRUCTURE (E.G. HOSPITALS, PUBLIC TRANSIT), DIST = DISTRICT ENERGY SOLUTIONS FOR RESIDENTIAL AREAS, GRID = GRID BALANCING/LOAD PEAK MANAGEMENT, HIGH = AREAS WITH HIGH ENERGY INPUT/OUTPUT, INDU = INDUSTRY, ISLA = AREAS WITHOUT GRID CONNECTION (E.G. ISLANDS), MOBI = MOBILE POWER SUPPLY FOR EVENTS, STAT = STATIONARY BACKUP POWER INSTEAD OF DIESEL GENERATORS, SUPP = POWER SUPPLEMENT FOR CORPORATE SUSTAINABILITY GOALS

### 3) A Survey Study on Institutional Influencing Factors

Additionally, we employ structural equation models (SEM) to assess theoretical frameworks explaining relationships among variables affecting HT acceptance, focusing on institutional influencing factors and using data from our representative survey study with  $N = 1,000$  participants. Targeting a broad audience, these surveys provide data for statistical analysis and model testing.

## III. IMPLICATIONS

The insights gathered from our analyses provide directions for future research in HT acceptance and actionable insights for stakeholders involved in the development and implementation of HTs. Key implications include:

- Validation of Acceptance Measures: Rigorous validation is essential for precise measurement of HT acceptance facets.
- Integration of Participation Variables: Often overlooked in transformative processes, participation variables should be prioritized, focusing on informational contexts and institutional influences.

- Exploration of Societal Narratives and Media Frames: Investigating societal trends and media frames is crucial as they impact technology deployment and uptake.
- Broadening Stakeholder Engagement: Understanding HT acceptance requires exploring acceptance beyond early adopters and the general public, particularly in co-production contexts.
- Focus on Longitudinal Studies: Emphasizing longitudinal studies is crucial to track evolving stakeholder perceptions and continuously adapt strategies.
- Integration of Feedback: Continuously gather and integrate feedback from stakeholders throughout the project.

## REFERENCES

- [1] Stephens, J. C., Peterson, T. R., Wilson, E. J. Socio-political evaluation of energy deployment (SPEED): A framework applied to smart grid. *UCLA Law Rev* 2013;61:1930. DOI: <https://doi.org/10.1016/j.techfore.2007.12.003>.
- [2] Wüstenhagen, R., Wolsink, M., Bürer, M. J. Social acceptance of renewable energy innovation: An introduction to the concept. *Energy Policy* 2007;35(5):2683–2691. DOI: <https://doi.org/10.1016/j.enpol.2006.12.001>.
- [3] Martin, A., Agnoletti, M. F., & Brangier, E. Users in the design of Hydrogen Energy Systems: A systematic review. *Int J Hydrog Energy* 2020;45(21):11889–11900. DOI: <https://doi.org/10.1016/j.ijhydene.2020.02.163>.
- [4] Lopez Jaramillo, O., Stotts, R., Kelley, S., Kuby, M. Content analysis of interviews with hydrogen fuel cell vehicle drivers in Los Angeles. *Transp Res Rec* 2019;2673(9):377–388. DOI: <https://doi.org/10.1177/0361198119845355>.
- [5] Bentsen, H. L., Skiple, J. K., Gregersen, T., Derempouka, E., Skjold, T. In the green? Perceptions of hydrogen production methods among the Norwegian public. *Energy Res Soc Sci* 2023;97:102985. DOI: <https://doi.org/10.1016/j.erss.2023.102985>.
- [6] Steller, R.S., Recklies, E.A., Schweizer-Ries, P. Shaping Transformation: Discourse Analysis and Systematic Review of Socio-Psychological Factors in Hydrogen Technology Acceptance. Unpublished manuscript 2024, Bochum University of Applied Sciences.

## ACKNOWLEDGEMENT

The research presented in all six chapters is being conducted in the project *CoupleIT! – IT-based sector coupling: Digitally controlled fuel cell and electrolyzer technologies for stationary and mobile applications*. This research is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

# Energy System-Oriented Identification of Hydrogen Storage Supply Scenarios: Method Development and Application within the Digi-HyPro Project

Jelto Lange\*, Martin Kaltschmitt  
*Institute of Environmental Technology  
 and Energy Economics  
 Hamburg University of Technology  
 Hamburg, Germany  
 \*j.lange@tuhh.de*

Julián Puszkiewicz\*, Myriam Covarrubias  
 Guraneros, Lukas Fleming, Thomas  
 Friedrich Johannes Kaufmann, Phillip  
 Krause, Jan Warfsmann, Eike Wienken,  
 Lukas Wildner, Thomas Klassen, Julian  
 Jepsen  
*Faculty of Mechanical Engineering,  
 Institute of Material Engineering  
 Helmut Schmidt University/  
 University of the Federal Armed Forces  
 Hamburg, Germany  
 \*puszkiewicz@hsu-hh.de*

Nies Reininghaus\*, Astrid Pistor, Car-  
 los Munoz Robinson, Michael Kröner,  
 Alexander Dyck  
*Urban and Residential Technologies  
 Department Institute of Networked  
 Energy Systems  
 Oldenburg, Germany  
 \*Nies.Reininghaus@dlr.de*

**Abstract** – Effective hydrogen storage is vital for the widespread adoption of hydrogen in energy systems, as it enables flexibility across various sectors. However, assessing such energy storage systems' suitability in future energy system configurations presents several challenges. One such challenge is the identification of representative operational scenarios for experimental testing of storage systems. Against this background, this paper presents an approach to derive such operational scenarios with the help of energy system modelling and optimization. Using the open-source energy system model and data set of Europe, PyPSA-Eur, cost-optimal future energy system configurations are identified, allowing the derivation of operational scenarios for energy storage facilities from the operation of the overall energy system. For this purpose, the methodology provides a way to identify a representative storage system from the entirety of corresponding storages in the energy system. Further, it allows determining representative time series sections using a segment identification algorithm, providing a basis for experimental technology testing. For an exemplary application of this methodology, further post-processing is implemented to consider the feasibility limits of subsystem components. The results showcase the effectiveness of the approach, offering a transparent and reproducible framework for defining operational scenarios for storage testing aligned with future energy system requirements.

**Keywords** – Energy and hydrogen storage, energy system modelling and optimization, operational scenarios, technology development

## NOMENCLATURE

$C_{H_2}$ in kWh	Hydrogen storage capacity
$D$	Euclidian distance
$i$	Index of technology implementation
$I$	Total number of energy system-wide technology implementations
$m$	Time series segment length

$n_{cyc}$	Number of equivalent annual storage cycles
$\bar{n}_{cyc}$	System-average of the number of equivalent annual storage cycles
$\dot{Q}_{H_2,cha}$ in kW	Hydrogen flow while charging
$t$	Time step index
$\Delta t$ in h	Time step width
$T$	Total number of time steps
$X$	Time series segment
$\hat{X}$	Z-normalized time series segment
$Y$	Time series segment
$\mu$	Arithmetic mean
$\sigma$	Standard deviation

## I. INTRODUCTION

Hydrogen is pivotal to the comprehensive de-fossilization of energy systems [1]. To achieve its widespread system integration and maximize its potential for energy system flexibilization, hydrogen storage plays a significant role [2]. Against this background, the Digi-HyPro project [3] is investigating a hydrogen storage system that uses various components to enable far-reaching and flexible sector coupling by connecting different energy system sectors (e.g. electrolyzer and fuel cell for electricity grid coupling, hydrogen compressors for mobility sector coupling). In this context, a key challenge lies in the evaluation of such systems and the experimental testing of their suitability to meet the requirements of the energy system (of the future), as both aspects are in dire need of information on possible operating scenarios that must reflect the energy system's requirements as representatively as possible.

For the preparation of corresponding information, recourse is frequently made to empirical values, or a wholly assumption-based estimation is carried out (e.g. [4], [5]). However, assumption-based investigations can often lead to arbitrary results; without a reasoned derivation of the operational requirements, the gained insights are also limited in their meaningfulness. Furthermore, the use of empirical values also has its disadvantages, as it implicitly transfers the conditions of the (fossil-based) energy past to a potential energy system of the future; a de-fossilized energy system will, however, be comprised differently and, therefore, likely yield substantially different requirements to be met by future-proof technologies. Consequentially, it is crucial to consider the superordinate energy system development when evaluating technology and deriving representative operational scenarios. The former has now been adapted by various studies, assessing techno-economic performance and market relevance under consideration of the energy system evolution and requirements (e.g. comparing electrolyzer, fuel cell and batteries [6], various energy [7] or electricity storages [8], and different above- and underground hydrogen storage systems [9]). However, regarding the experimental testing of new technology concerning its suitability to meet future energy system requirements, there is still a lack of appropriate methodologies.

This paper tries to close this identified research gap by providing an approach to derive operational scenarios – especially for energy storage systems – for experimental technology testing based on energy system modelling and optimization for future system configurations. For this purpose, in section II, first the fundamental procedure is described. Subsequently, in section III, the underlying assumptions are discussed, and the system to which the approach is exemplarily applied is briefly outlined. In section IV, the results are then presented and discussed. Finally, a conclusion is drawn in section V.

## II. METHODOLOGY

The general approach follows the idea to directly derive the requirements for storage operation (i.e. when, how often, how long and how fast to charge or discharge) from the necessities that arise in a large-scale energy system context. The challenge in this regard is to arrive at a basis for experimental technology testing on component level coming from data sets of large-scale energy systems. The implemented methodology for this is described in the following. Figure 1 provides a graphical summary of the approach. Under the presented approach, the identification of a potential future (cost-optimal) energy system configuration is the prerequisite for deriving any operational requirements for energy storage on the component level. By optimizing the overall energy system design and operation including the storage technology under consideration, information is gained on how the technology is operated in an overall cost-optimal system configuration. Taking this operation as a basis for experimental testing allows to test the storage technology under operational conditions for which it would be part of a cost-optimal system and, thus, likely the most competitive.

Here, we use the open-source energy system model and data set of Europe PyPSA-Eur [10], [11], [12] for the modelling and optimization task. PyPSA-Eur uses various openly available data sources to model the energy demands in all of Europe's energy sectors (i.e. electricity, transport, heating, industry, agriculture, etc.). Under the strict requirement that all energy demands are always met, the model is used to identify

the system configuration that would manage the supply with the lowest annualised costs. Here, further boundary conditions or constraints can be defined (e.g. the requirement of a CO<sub>2</sub>-neutral energy supply). The model reflects temporal variations (e.g. time series of the availability of weather-dependent renewable energies) and spatial differences (i.e., location-specific demand and generation data, which in turn is aggregated in distinct geographical nodes) and, thus, takes into account a spatiotemporal resolution. While the highest temporal resolution amounts to 1 h time steps, the largest amount of spatially distinct model nodes exceeds 1,000. Generally, a trade-off has to be made between modelling accuracy and computational complexity. For a more in-depth explanation of the model, please refer to the original literature (e.g. [10], [11], [12]).

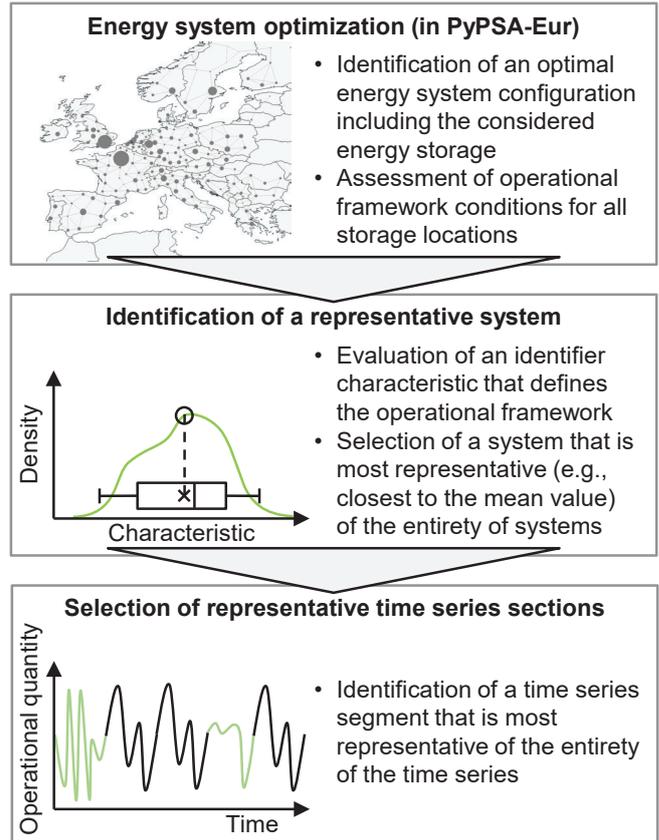


FIGURE 1: FLOW DIAGRAM OF THE APPROACH CONSIDERED FOR DERIVING OPERATIONAL FRAMEWORK CONDITIONS ON THE COMPONENT LEVEL FROM DATA FROM THE SUPERORDINATE ENERGY SYSTEM

Based on the energy system optimization, an overall system configuration results that includes various potentially relevant time series of the operation of the considered storage technology. Here, the nodal time series of the state of charge are of particular interest as they summarize the operational requirements for each storage system. However, if particular emphasis is placed, e.g. on the charging process, one could also define the time series for storage charging as the central operational characteristic. Depending on the geographical resolution implemented in PyPSA-Eur, the results include several hundreds of locations and, thus, time series to evaluate. In order to make this amount more processable, a particularly representative system needs to be determined. A key operational characteristic (e.g. the number of equivalent annual storage cycles) must be defined and evaluated to classify the different locations or system implementations. This information can then be used to determine the most representative system, e.g.

by selecting the storage for which this key characteristic is closest to the average value for all locations.

Finally, the last step is to assess the time series of the identified representative storage system concerning particularly relevant sections. These sections should be temporal segments of the total storage operation that optimally represent the overall time series. Thus, a method needs to be implemented to identify patterns that are especially regular and often repeated throughout the total temporal data. So, starting with the whole time series (e.g. one year of data), this is intended to identify shorter periods that allow for experimental implementation but are still as representative as possible of the total data.

To find such segments, each time series section of a certain length  $m$  is compared with every other possible section of the same length (e.g. for an overall time series including ten time steps and a segment length of five steps, this would result in five comparisons per segment). By calculating the z-normalized Euclidean distance between all these pairs of segments, the section with the smallest differences to a certain number of best-fitting segments can be determined. This way, the time series section with the most pronounced representativeness of the entire time series is identified. The z-normalized Euclidean distance  $D$  between each pair of segments is calculated using equation (1)

$$D(\hat{X}, \hat{Y}) = \sqrt{(\hat{x}_1 - \hat{y}_1)^2 + \dots + (\hat{x}_m - \hat{y}_m)^2} \quad (1)$$

with z-normalization according to equation (2).

$$\hat{X} = \frac{X - \mu_X}{\sigma_X} \quad (2)$$

Here,  $\hat{X}$  and  $\hat{Y}$  are two z-normalized segments of length  $m$  calculated from the original segments  $X$  and  $Y$  using their mean values  $\mu_X$  and  $\mu_Y$  as well as their standard deviations  $\sigma_X$  and  $\sigma_Y$  respectively. An efficient implementation of this segment comparison is used according to [13]. Finally, the identified time series section is adapted to the real considered system size for experimental testing (e.g. by scaling the maximum state of charge to the available storage capacity).

In conclusion, this provides information on the necessary operation of a storage technology under investigation that is derived in its entirety from the overall energy system operation for which the storage is to be tested. Thus, the approach allows a generalizable and reproducible definition of operational framework conditions without the need for explicitly assuming any key characteristics of storage operation. Depending on the particular storage system under consideration, some additional post-processing steps might be relevant for a final transfer of the identified time series sections to experimental technology testing; however, the main foundation stays the same. In the following sections the method will be applied exemplarily showing results for each methodological step.

### III. METHOD APPLICATION

This section sets the frame for the application of the developed methodology by describing the considered technology and framework assumptions (section A), defining relevant method parameters (section B) and specifying additional post-processing steps (section C).

#### A. Considered Technology and Energy-System Assumptions

The identification of operational scenarios will be applied to a hydrogen storage system operating in a future energy supply configuration. The storage system is to be operated both in connection with the electricity grid (i.e. via electrolyzers and fuel cells) and with the transportation sector (i.e., by supplying high-pressure hydrogen via hydrogen compressors). We apply the methodology to find a representative time series for charging and discharging such hydrogen storage, resulting from an optimized interaction with the energy system.

To define a potential future energy system configuration, carbon-neutrality of the electricity and transportation sector is set as a constraint for the optimization in PyPSA-Eur. It is assumed that 15% of transportation end energy consumption is met by hydrogen, resulting in a demand for land-based transportation that is in line with industry expectations [14]. The remaining 85% of transportation energy demand is met by using electricity, representing the general tendency for widespread electrification in the mobility sector [15]. Techno-economic data (e.g. efficiencies, specific CAPEX) for the energy system optimization are taken from literature values with 2035 as the reference year [16]. Regarding the energy infrastructure – essential for energy system development and operation – a 25% increase of power grid capacities is set as the upper boundary, while the model can endogenously develop hydrogen pipelines. This substantial potential for infrastructure expansion reflects the ongoing planning, development, and construction of such infrastructure throughout Europe. The analysis covers the entire European continent as implemented in PyPSA-Eur, realizing a spatial aggregation to 181 nodes (considered sufficient to adequately account for regional differences [17]). Further, the system operation is optimized for one year, considering a temporal resolution of 3 h time steps. Lastly, 2013 is used as a baseline for weather data as the year showed wind speeds and solar radiation values throughout Europe that were close to their long-term average values [18], [19]. Thus, power production from photovoltaics and wind will neither be over- nor underestimated.

#### B. Method Parameters

We define the state of charge of the hydrogen storage as the operational quantity to be evaluated. Further, the equivalent number of annual storage cycles is chosen as the key characteristic for identifying a representative system. The storage that exhibits annual storage cycles closest to the average of all storage cycle values is chosen as the most representative system. The average storage cycle value  $\bar{n}_{\text{cyc}}$  is calculated according to equation (3)

$$\bar{n}_{\text{cyc}} = \frac{\sum_{i=1}^I n_{\text{cyc},i}}{I} \quad (3)$$

based on the cycle value of each system as per equation (4)

$$n_{\text{cyc},i} = \frac{\sum_{t=1}^T \dot{Q}_{\text{H}_2, \text{cha}, i, t} \Delta t}{C_{\text{H}_2, i}} \quad (4)$$

Here,  $i$  stands for each optimized hydrogen storage implementation of the entirety of nodes  $I$  for which the optimization results in an installation of a corresponding hydrogen storage facility. The flow of hydrogen that is charged to the storage  $\dot{Q}_{\text{H}_2, \text{cha}, i, t}$  in each time step  $t$  is multiplied by the time step width  $\Delta t$  to calculate the total amount of stored hydrogen over the period  $T$  (i.e., one year). This amount of stored hydrogen is related to the storage capacity  $C_{\text{H}_2, i}$  for each implementation

node  $i$ . Accordingly, the number of annual equivalent storage cycles indicates how often the storage tank would have been fully charged and discharged if an operation with full charging cycles had been implemented. For the subsequent evaluation of the time series of the chosen representative storage system using the segment-identification algorithm, week-long segments are considered (i.e.  $m = 168$ ) assessing the z-normalized Euclidian distance to the 50 most similar time series.

C. Additional Post-Processing

While we apply the methodology to identify representative time series for hydrogen storage operation, an experimental setup might include further system components (like in the Digi-HyPro project, where a metal hydride hydrogen storage is linked to an electrolyzer as well as a fuel cell [3]). Thus, beyond the identification of representative time series on the component level, there might be further limits regarding operational feasibility in experiments that need to be considered. This necessitates an additional post-processing for the particular system under consideration here (depicted in the box in Figure 2). This additional processing, which we discuss briefly in the following, includes steps to derive operational requirements for an electrolyzer and a fuel cell from the identified hydrogen storage operation.

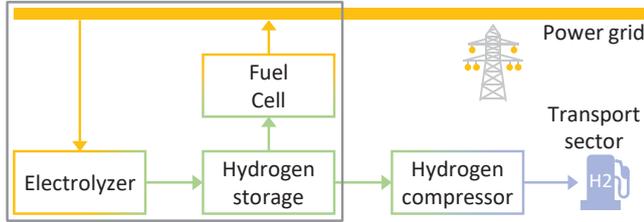


FIGURE 2: SCHEMATIC DEPICTION OF THE CONSIDERED SYSTEM AND EXPERIMENTAL SETUP (GREY BOX)

The experimental characteristics to be considered are given in Table I and include both minimum and maximum operational hydrogen flow rates of the electrolyzer and the fuel cell.

TABLE I: SYSTEM PARAMETERS.

Parameter	Unit	Min. value	Max. Value
Storage capacity	kg <sub>H2</sub>	0	2
Electrolyzer hydrogen flow rate (in operation)	kg <sub>H2</sub> /h	0.027	0.18
Fuel cell hydrogen flow rate (in operation)	kg <sub>H2</sub> /h	0.07	0.38

Besides the scaling of the state of charge to the available storage capacity, minimum and maximum flow rates of both the electrolyzer and the fuel cell need to be considered so that the corresponding charging and discharging of the storage can be managed by the available subsystems. To achieve this, first the time series of the state of charge is used to calculate the necessary hydrogen withdrawal and uptake to facilitate the identified changes of the state of charge. Secondly, these resulting flow rates are compared to the technical limitations of the available electrolyzer and fuel cell (see Table I). Finally, any exceedance of feasibility limits is excluded and the remaining longest time series section is defined as the basis for technology testing. Figure 3 illustrates this additional post-processing for a hypothetical case of a flow rate exceedance for the fuel cell.

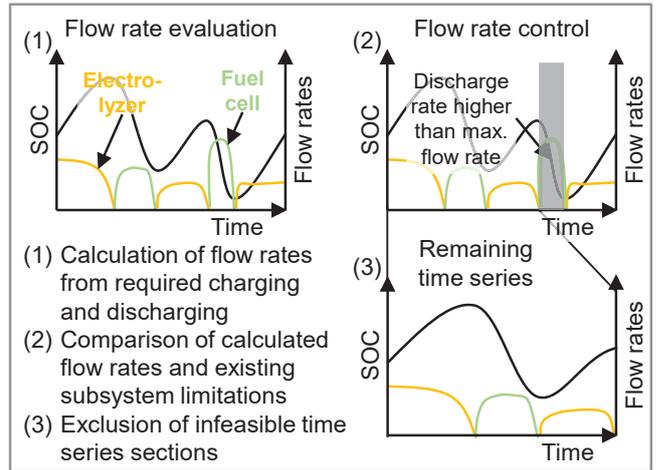


FIGURE 3: ADDITIONAL POST-PROCESSING STEPS FOR DERIVING FUEL CELL AND ELECTROLYZER OPERATION (SOC: STATE OF CHARGE)

IV. RESULTS AND DISCUSSION

The following results give an overview of the different steps for an energy system-oriented identification of operational scenarios for testing a hydrogen storage system. Since the energy system optimization is the basis for the subsequent steps, Figure 4 shows the PyPSA-Eur results concerning the optimized hydrogen storage capacity of the considered storage technology and its regional distribution throughout the European energy system.

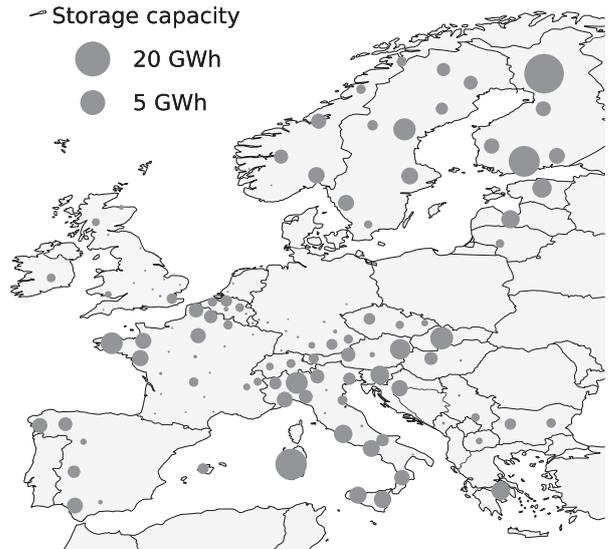


FIGURE 4: OPTIMIZED DEVELOPMENT AND REGIONAL DISTRIBUTION OF (ABOVEGROUND) HYDROGEN STORAGE SYSTEMS AND THE LOCALLY IMPLEMENTED STORAGE CAPACITY.

Visibly, there are different regions, where the considered hydrogen storage could play a substantial role in the further development of the overall system. Simultaneously, there are various regions where it would not be implemented under cost-optimal conditions (e.g. in Northern Central Europe, where sufficient underground storage potentials render aboveground hydrogen storage uncompetitive).

Using the entirety of locations as the basis for system evaluation, each system's equivalent annual storage cycles are calculated by applying equation (3). The total distribution of storage cycles yields the probability density, as shown in Figure 5. It can be seen that the annual equivalent storage cycles are

generally relatively high (i.e. almost no seasonal energy storage in the considered hydrogen storage systems), while there is still a significant variance in terms of cycle amount. 50% of the systems exhibit between 163 and 354 annual equivalent storage cycles. The mean value amounts to roughly 257 under the considered energy system framework. The hydrogen storage that achieves equivalent annual storage cycles closest to this system-wide average operates at roughly 263 cycles (at a node in Norway). This hydrogen storage system is consequently considered the most representative. However, it should be noted that the reduction and restriction to a single system result in a noticeable loss of information, as the selection of a single system does not optimally reflect the substantial variance.

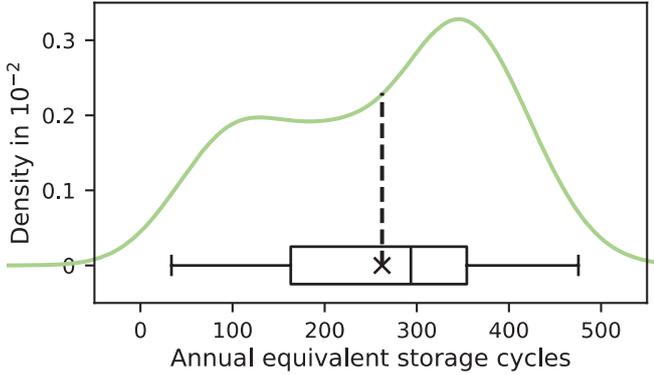


FIGURE 5: PROBABILITY DENSITY OF ANNUAL EQUIVALENT STORAGE CYCLES OF THE (ABOVEGROUND) HYDROGEN STORAGE SYSTEMS IN A CARBON-NEUTRAL EUROPEAN ENERGY SYSTEM (THE CROSS 'X' DENOTES THE MEAN VALUE IN THE BOXPLOT)

Applying the segment-identification algorithm to the time series of the state of charge of the identified hydrogen storage system yields the evaluation shown in Figure 6. The most representative week-long pattern (i.e.  $m = 168$ ) lies in the middle of the year, while the 50 most similar segments are fairly evenly distributed throughout the total time series. The identified most similar patterns show some stronger differences to the final selection (especially from hour 7,500 to hour 8,200). However, their greater similarity compared to the non-selected parts of the overall time series is also evident. Thus, the algorithm for identifying the most frequent patterns found that most of the storage operation takes place approximately between 15% and 75% of the state of charge. The full charge and discharge cycles, which do occur but are much less frequent, therefore do not optimally represent the most common mode of operation.

Finally, the identified time series for the state of charge is translated to the necessary fuel cell and electrolyzer operation in the additional post-processing step. Figure 7 shows the resulting operation of the considered fuel cell, hydrogen storage, and electrolyzer setup over the week-long period identified previously (top) and the remaining time series after post-processing (bottom).

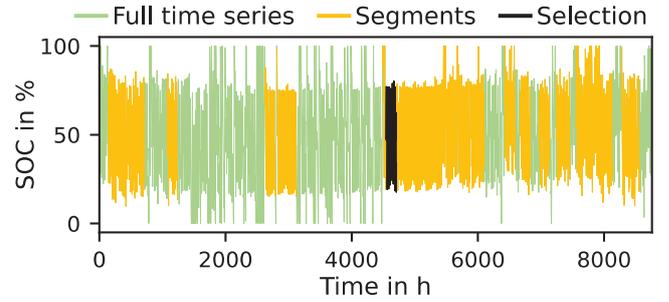


FIGURE 6: TIMES SERIES OF THE STATE OF CHARGE OF THE MOST REPRESENTATIVE HYDROGEN STORAGE INCLUDING IDENTIFIED SEGMENTS AND THE SELECTED MOST REPRESENTATIVE SECTION (SOC: STATE OF CHARGE)

It can be seen that when converting to flow rates, the respective technical limits of the subsystems would be exceeded at four points in time. Either the flow rate exceeds the maximum hydrogen output of the electrolyzer, or it falls short of the minimum flow requirement of the fuel cell or the electrolyzer. Nevertheless, a sufficiently long time series remains after post-processing, which adequately represents the characteristics of the identified week-long section. Thus, the final time series still offers valuable insights into how a corresponding hydrogen storage system might be operated in a future energy system configuration.

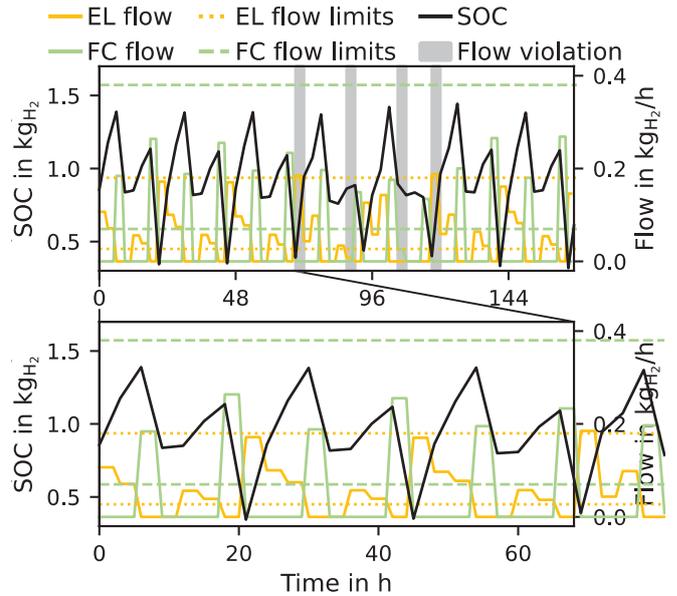


FIGURE 7: FINAL TIME SERIES (TOP) AND POST-PROCESSED TIME SERIES SECTION THAT CAN ACT AS A FINAL TESTING INPUT (BOTTOM)

## V. CONCLUSION

Due to the uncertain operational framework conditions of future technology, which may already have to be developed and tested today, there is an urgent need for reliable and comprehensible ways of defining representative operational scenarios. This article proposes an approach for deriving such operational framework conditions for energy storage systems from the necessities that arise in the overall energy system operation. By performing this derivation on a cost-optimized energy system configuration, the storage technology under consideration can be tested for operational conditions, for which it would be cost-optimal in the respective energy system and, therefore, the most competitive.

For deriving the corresponding operational framework conditions, the developed approach supports both the selection of a suitable representation of the storage technology under investigation from a large number of systems in the higher-level energy supply system and the selection of representative time series segments from long overall time series that would be unsuitable for technology testing due to their length (e.g. one year). The selection is made in such a way that the remaining operational scenarios represent the entirety of storage operation of the considered technology in the energy system as well as possible.

In an exemplary application of the approach, we show that it can be used to effectively derive operational scenarios that are highly representative of storage operation in the higher-level energy system while still offering a feasible basis for technology testing.

Thus, the proposed approach shows a significant improvement over entirely assumption-based procedures, as the identified operation scenario can be directly linked to the respective energy system configuration and the resulting operational requirements. While assumptions still need to be made regarding the development of the energy system (e.g. regarding the remaining CO<sub>2</sub>-emissions during system operation or the development of various energy system sectors), no explicit assumptions need to be formulated for energy storage operation. Furthermore, it is reproducible under which energy system conditions which operational scenarios arise for energy or hydrogen storage, resulting in a transparent and well-supported methodology for identifying operational scenarios for energy storage technology testing.

#### ACKNOWLEDGEMENT

This research paper is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] Breyer, C., Khalili, S., Bogdanov, D., Ram, M., Oyewo, A.S., Aghahosseini, A., Gulagi, A., Solomon, A.A., Keiner, D., Lopez, G., Ostergaard, P.A., Lund, H., Mathiesen, B.V., Jacobson, M.Z., Victoria, M., Teske, S., Pregger, T., Fthenakis, V., Raugei, M., Holttinen, H., Bardi, U., Hoekstra, A., Sovacool, B.K.: On the History and Future of 100% Renewable Energy Systems Research, *IEEE Access*, 10, 2022, S. 78176–78218
- [2] Neumann, F., Brown, T.: The near-optimal feasible space of a renewable power system model, *Electric Power Systems Research*, 190, 2021, S. 106690
- [3] dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr: *Digi-HyPro – Digitalisierte Wasserstoffprozesskette für die Energiewende*, <https://dtecbw.de/home/forschung/hsu/projekt-digi-hypro>, 2020. Zuletzt geprüft: 13.03.2024
- [4] Abdin, Z., Khalilpour, K., Catchpole, K.: Projecting the levelized cost of large scale hydrogen storage for stationary applications, *Energy Conversion and Management*, 270, 2022, S. 116241
- [5] Martínez de León, C., Ríos, C., Molina, P., Brey, J.J.: Levelized Cost of Storage (LCOS) for a hydrogen system, *International Journal of Hydrogen Energy*, 52, 2024, S. 1274–1284
- [6] Parzen, M., Neumann, F., van der Weijde, A.H., Friedrich, D., Kiprakis, A.: Beyond cost reduction: improving the value of energy storage in electricity systems, *Carbon Neutrality*, 1, 2022, 1, S. 1–21
- [7] Parzen, M., Fioriti, D., Kiprakis, A.: *The Value of Competing Energy Storage in Decarbonized Power Systems*, 2023
- [8] Gøtske, E.K., Andresen, G.B., Victoria, M.: Cost and Efficiency Requirements for Successful Electricity Storage in a Highly Renewable European Energy System, *PRX Energy*, 2, 2023
- [9] Lange, J., Schulthoff, M., Puszkiel, J., Sens, L., Jepsen, J., Klassen, T., Kaltschmitt, M.: Aboveground hydrogen storage – Assessment of the potential market relevance in a Carbon-Neutral European energy system, *Energy Conversion and Management*, 306, 2024, S. 118292
- [10] Hörsch, J., Hofmann, F., Schlachtberger, D., Brown, T.: PyPSA-Eur: An open optimization model of the European transmission system, *Energy Strategy Reviews*, 22, 2018, S. 207–215
- [11] Neumann, F., Zeyen, E., Victoria, M., Brown, T.: The Potential Role of a Hydrogen Network in Europe, *arXiv*, 2022
- [12] Victoria, M., Zeyen, E., Brown, T.: Speed of technological transformations required in Europe to achieve different climate goals, *Joule*, 6, 2022, 5, S. 1066–1086
- [13] Law, S.: STUMPY: A Powerful and Scalable Python Library for Time Series Data Mining, *Journal of Open Source Software*, 4, 2019, 39, S. 1504
- [14] Wang, A., Jens, J., Mavins, D., Moultak, M., Schimmel, M., van der Leun, K., Peters, D., Buseman, M.: *European Hydrogen Backbone - Analysing future demand, supply, and transport of hydrogen*, Utrecht, 2021
- [15] Plötz, P.: Hydrogen technology is unlikely to play a major role in sustainable road transport, *Nature Electronics*, 5, 2022, 1, S. 8–10
- [16] lisazeyen, euronion, Markus Millinger, Fabian Neumann, Max Parzen, Tom Brown, Lukas Franken, *martavp, lukasnacken: PyPSA/technology-data: - Technology Data v0.6.0 (v0.6.0)*, Zenodo, 2023
- [17] Frysztacki, M.M., Hörsch, J., Hagenmeyer, V., Brown, T.: The strong effect of network resolution on electricity system models with high shares of wind and solar, *Applied Energy*, 291, 2021, S. 116726
- [18] European Centre for Medium-Range Weather Forecasts, *ECMWF: Wind and solar energy resources*, <https://climate.copernicus.eu/esotc/2022/wind-solar-energy-resources>, 2022. Zuletzt geprüft: 09.05.2023
- [19] European Centre for Medium-Range Weather Forecasts, *ECMWF: Clouds and sunshine duration*, <https://climate.copernicus.eu/esotc/2022/clouds-and-sunshine-duration>, 2022. Zuletzt geprüft: 09.05.2023

# Experimental Development of the Set-Unit Concept on a Laboratory Scale for the H<sub>2</sub>-Chain: Production-Storage-Compression-Utilization (Digi-HyPro)

Julián Puszkiel\*, Torsten Brinkmann, Myriam Covarrubias Guraneros, Lukas Fleming, Merle Höne Thomas Friedrich Johannes Kaufmann, Phillip Krause, Jan Warfsmann, Eike Wienken, Lukas Wildner, Thorsten Wolf, Thomas Klassen, Julian Jepsen  
*Institute of Hydrogen Technology and Institute of Membrane Research Helmholtz-Zentrum Hereon Geesthacht, Germany*  
 \*julian.puszkiel@hereon.de

Jelto Lange\*, Martin Kaltschmitt  
*Institute of Environmental Technology and Energy Economics (IUE) Hamburg University of Technology Hamburg, Germany*  
 \*j.lange@tuhh.de

Nies Reininghaus\*, Astrid Pistor, Carlos Muñoz Robinson, Michael Kröner, Alexander Dyck  
*Institute of Networked Energy Systems German Aerospace Center Oldenburg, Germany*  
 \*Nies.Reininghaus@dlr.de

**Abstract** – The Digi-HyPro (Digitalized Hydrogen Process Chain for the Energy Transition) project's conceptual development of the SET-Unit investigates and facilitates the connection between the electric, gas, and mobility grid. This application report describes the experimental design of the Smart Energy Transition unit (SET-Unit), contemplating the bottom-up and top-down approaches. For the bottom-up approach, the design of core devices such as metal hydride-based hydrogen storage (MHS) and compressor (MHC) systems are shown. The gas separation system (GSS) concept is based on a hybrid process composed of membrane and pressure swing adsorption (PSA) for the gas grid coupling. Commercial anion exchange membrane electrolyzer (AEM-EL) and polymer exchange membrane fuel cell (PEM-FC) are assembled for the power grid connection. For the top-down approach, the first experimental SET-Unit composed of AEM-EL–MHS–PEM-FC in the nominal power range between 5 and 10 kW<sub>el</sub> and its control strategy for the optimal hydrogen and heat coupling is presented. All experimental development is carried out in the facilities of the Helmholtz-Zentrum Hereon in the frame of a cooperation agreement with the Helmut Schmidt University/University of the Federal Armed Forces.

**Keywords** – Green hydrogen, hydrides, storage, compression, purification

## NOMENCLATURE

AEM-EL	Anion Exchange Membrane Electrolyzer
CM	Cooling Module
EL	Electrolyzer
FC	Fuel Cell
FCM	Fuel Cell Module
FLC	Fuzzy Logic Controller
FSM	Finite State Machine
GDL	Gas Diffusion Layer
GSS	Gas Separation System
HCD	Humidity Control Device
HER	Hydrogen Evolution Reaction

HTF	Heat Transfer Fluid
HTTF1	Hydrogen Tank Testing Facility One
MHC	Metal Hydride Compressor
MHS	Metal Hydride Storage
MPC	Model Predictive Controller
OER	Oxygen Evolution Reaction
PED	Power Electronic Device
PEM-FC	Proton Exchange Membrane Fuel Cell
RHCs	Reactive Hydride Composites
WPD	Water Purification Device

## I. INTRODUCTION

Hydrogen is a very light gas at standard pressure and temperature and is difficult to store. For compact storage, e.g. in vehicles, high pressures of up to 700 bar are necessary. Alternatively, hydrogen energy can be liquefied in a costly process whereby approx. 15-30% of the energy content of the hydrogen must be spent on liquefaction [1].

An energy-efficient and compact solution is hydrogen storage in metal hydrogen compounds, so-called metal hydrides. Compared to high-pressure hydrogen storage, twice the volumetric capacity (> 50 kg H<sub>2</sub>/m<sup>3</sup>) is achieved at much lower pressures, typically a few 10 bar [2]. Especially for stationary applications in urban areas, where land is limited and expensive and high restrictions regarding safety apply, metal hydride storage has a decisive advantage over the other storage methods. The storage is also stable over the long term, as there is almost no diffusion of gaseous hydrogen through the container wall due to the low storage pressure, and the metal hydrides do not cause any other side reaction. The hydrogen is stored 100% reversibly - completely independent of the time scale [1]. Interstitial metal hydrides benefit from absorbing and releasing hydrogen within a few seconds on a small scale [3]. However, in larger tanks, the transport of the heat of the

reaction becomes rate-determining [4]. Therefore, metal hydride tanks must be optimally designed systematically and knowledge-based for the desired service.

In the Digi-HyPro project (Digitalized Hydrogen Process Chain for the Energy Transition), the digital and experimental coupling of all required components to connect the gas, power, heat, and mobility grid has been developed. The concept of the so-called SET-Unit can be seen in Figure 1. Our previous work [5] presented the bottom-up and top-down strategies for developing the SET-Unit concept. On the one hand, the bottom-up approach consists of units and subsystem-level model developments and their experimental setup designs and optimizations. On the other hand, the top-down approach consists of the conceptual design and development of the whole SET-Unit and the final conception considering an optimized automatic control strategy.

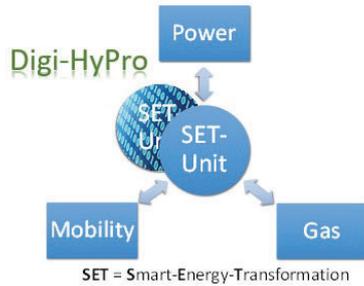


FIGURE 1: CONCEPTUAL VISUALIZATION OF THE SET-UNIT

Figure 2 shows a simplified diagram of the SET-Unit concept, including a water EL, water FC, MHS, MHC, and GSS. The system is designed to use and deliver electrical power during low and high-demand hours. Furthermore, mitigating and stabilizing the intermittency of power generation from renewable sources is possible. Hydrogen generated by the EL can be stored compactly and safely in the MHS or sent to the MHC to provide the mobility sector. The hydrogen can be separated from the mixed gas grid (natural gas + hydrogen) by the GSS. For power delivery, hydrogen can be converted in the FC.

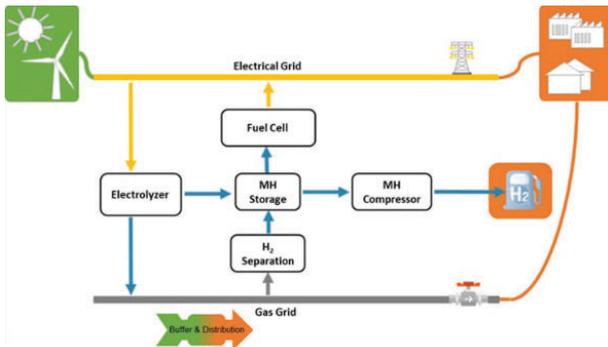


FIGURE 2: SIMPLIFIED DIAGRAM OF THE SET-UNIT WITH THE HYDROGEN-POWER COUPLING

This application report describes the conceptual and experimental development and assembly of the MHS system, MHC system, GSS, commercial AEM-EL and PEM-FC. Furthermore, the first experimental conception of the SET-Unit in the nominal power range between 5 and 10 kW<sub>el</sub> and its control strategy for the optimal hydrogen and heat coupling have been planned. All experimental development is carried out in the facilities of Helmholtz-Zentrum Hereon in the frame of a cooperation agreement with Helmut Schmidt University/University of the Federal Armed Forces.

## II. DEVELOPMENT AND ASSEMBLY OF THE SET-UNIT COMPONENTS: BOTTOM-UP

This section is about the experimental development and assembly of the components included in the SET-Unit concept as a part of the bottom-up approach.

### A. AEM-EL

Figure 3 (centre) shows an AEM-EL rack capable of producing green hydrogen and oxygen through electrochemical conversion of water while consuming electrical energy from renewable sources. This type of EL's water electrolysis occurs at low temperatures of  $\leq 50$  °C and in an alkaline environment created by a 1 wt.-% K(OH) electrolyte solution.

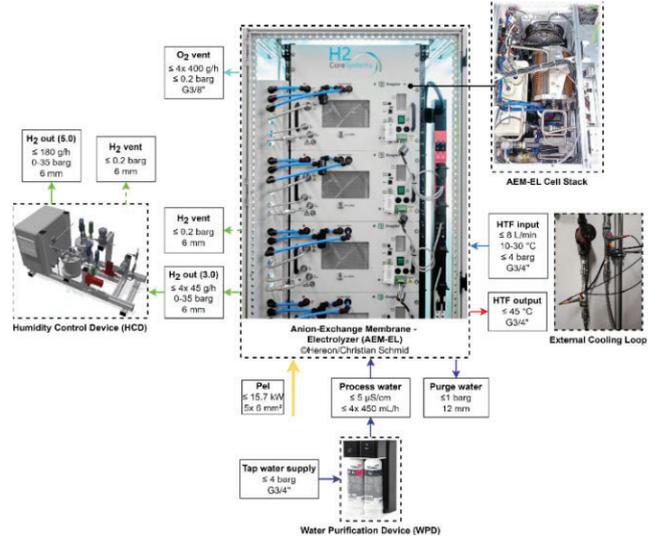
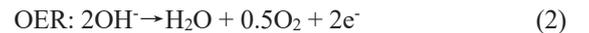
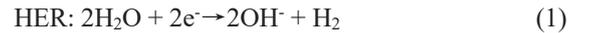


FIGURE 3: AEM-EL SYSTEM SETUP WITH EXTERNAL CONNECTIONS

The hydrogen and oxygen evolution reactions (HER and OER) are given in Equations (1) and (2). Noticeably, the hydroxide ions from the added electrolyte solution are not consumed for the overall electrochemical reactions [6].



The AEM-EL is based on four individual electrolyzer modules with approximately a net power of 2.4 kW<sub>el</sub> per module and 9.6 kW<sub>el</sub> in total. While in full load operation of the four electrolyzer modules, it will produce up to 0.18 kg/h hydrogen and 1.6 kg/h of oxygen with a water consumption of ca. 1.8 L/h.

As the hydrogen produced can contain up to 1,000 ppm water vapor, subsequent hydrogen drying is required to achieve a metal hydride-compatible hydrogen purity of 5.0 ( $\leq 10$  ppm water content). The used process-water and produced hydrogen must be treated for purity before and after the water-splitting process. The electrolysis water is provided by a water purification device (WPD), which can purify tap water through a combined water treatment process with several filtration, desalination, and ion exchange stages to achieve a final purity, measured by the electrical conductivity of  $\leq 5$  µS/cm. For the hydrogen drying and purification process, a commercially available temperature swing adsorption dryer is used on the one hand, and on the other hand, a project-specific humidity control device (HCD) based on a water-affine polymer membrane to separate the residual water from the hydrogen

content of the hydrogen gas mixture produced is used. The latter HCD measures the water content via dew point transmitters and the feed's mass flows, temperatures, and pressures, as well as retentate and permeate connections.

When the AEM-EL modules are operated, a certain amount of waste heat is generated due to the losses during water electrolysis. In particular, the anode and cathode activation overpotentials and the ionic conductivity of the membrane play a crucial role in this regard.

Due to the low-temperature electrolysis, the temperature of the recoverable waste heat is limited to 45 °C with a heat output of around 2.8 kW<sub>therm</sub> in full-load operation. As the EL modules are liquid-cooled, a water-ethylene glycol mixture is used as heat transfer fluid (HTF). Depending on the given operating load profile and the HTF inlet temperature, the volume flow is adjusted between 4 to 8 L/min.

At a later stage of the project, a thermal-hydraulic coupling of the AEM-EL with the other heat-consuming or heat-supplying devices is planned to optimize the overall system efficiency of the SET-Unit.

### B. PEM-FC

A water-cooled PEM-FC system from the Swedish company Powercell Group is used. It has a working range pressure between 3-6 bar to ensure proper hydrogen diffusion through the gas diffusion layer (GDL). It consumes 0.36 kg/h of hydrogen with a purity of  $\geq 99.99\%$  (3.5 grade) at a maximum net power and up to 260 m<sup>3</sup>/h of ambient air for chemical reaction and ventilation.

The HER and OER are given in Equations (3) and (4). Hydrogen molecules dissociate and adsorb at the anode, oxidize to protons, and generate flowing electrons, which provide power. Then, protons diffuse through the membrane. Oxygen molecules adsorb at the cathode and react with the protons, producing water [7].



Figure 4 shows the PEM-FC system, which comprises of three main modules: fuel cell module (FCM), power electronics device (PED), and cooling module (CM).

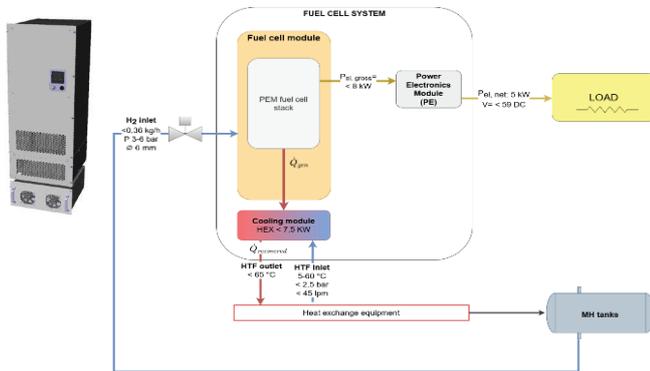


FIGURE 4: PEM-FC SYSTEM SETUP WITH EXTERNAL CONNECTIONS

A PEM-FC stack is the main component of the FCM producing up to 8 kW gross electrical power and an unregulated DC output voltage. Next, voltage regulation takes place in the PED delivering a maximum of 59 VDC and  $\leq 5$  kW Net Power to the end-user.

In the CM a brazed plate heat exchanger constitutes the interface between an external heating/cooling loop (for the option to recover the heat by the system) and an internal heating/cooling loop, both working together to regulate the temperature of the stack ( $\leq 70$  °C) with a volume flow of up to 45 L/min.

The chosen HTF for the external heating/cooling loop will be a 50/50 water-glycol mixture. For the inlet of the CM, the operative temperature of the HTF is in the range between 5 to 60 °C and pressure up to 2.5 bar. The outlet of the HTF should be a maximum of 60 °C.

For the first stage of experiments, the PEM-FC system is tested in basic operation together with an air-cooled thermostat (LAUDA) with 8 kW of heating and 11 kW of cooling power for analysing steady and dynamic thermal behaviour of the PEM-FC system.

The PEM-FC system's second step is to be thermally coupled with the MHS by using the recovered heat in the external heating/cooling loop to release hydrogen from the MHS system.

### C. MHS

#### 1) MHS material

The material of choice for this project is the commercially available interstitial AB<sub>2</sub>-metal alloy HydralloyC5 (Ti<sub>0.95</sub>Zr<sub>0.05</sub>Mn<sub>1.46</sub>V<sub>0.45</sub>Fe<sub>0.09</sub>), which has several advantages for a hydrogen storage system. For instance, unlike other solid-state hydrogen storage materials, e.g. MgH<sub>2</sub> [8] or reactive hydride composites (RHCs) [9], it is stable against air and humidity and can be stored in ambient conditions without any protective atmosphere before activation for ease of handling. Furthermore, HydralloyC5 possesses fast kinetics (filling of a hydrogen tank possible within minutes) and a high volumetric capacity of  $> 60$  kg/H<sub>2</sub> m<sup>3</sup> while still retaining the acceptable gravimetric capacity for a stationary system of about 1.6 wt%. Another advantage of HydralloyC5 is the thermodynamic properties of this material. As mentioned before, the AEM-EL for the SET-Unit can provide an operation temperature range between 20 and 30 °C and an operation pressure between 30 and 35 bar. In contrast, the used PEM-FC has an operative temperature between 45 and 50 °C and needs a hydrogen back pressure between 3 and 6 bar for the operation. These operative conditions of the EL and FC are highlighted in Figure 5 as green rectangles and compared to the temperature-dependent equilibrium pressure of HydralloyC5.

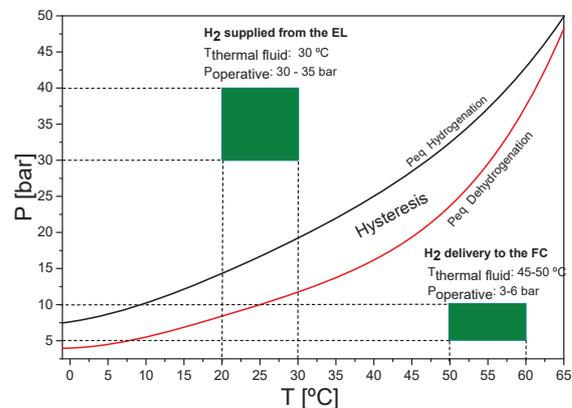


FIGURE 5: TEMPERATURE-DEPENDENT EQUILIBRIUM PRESSURE FOR THE USED MATERIAL HYDRALLOYC5 AND THE OPERATION PARAMETERS OF THE FUEL CELL AND ELECTROLYZER

The working conditions of HydralloyC5 are, therefore, at a "sweet spot" for an intermediate storage system: It can absorb hydrogen from the electrolyzer and feed the fuel cell without the need for an additional compressor system, which is known to be one of the most expensive parts during operation [10], or heat pump.

## 2) MH-tank and System

A digital twin was developed to design a metal hydride storage tank for a large-scale application using measured data from the lab scale (a few grams of material). With this digital twin, a large-scale storage vessel consisting of a standard pipe (168.3 mm diameter), an external heat exchanger, and an internal passive heat exchanger structure made from aluminium was designed to reduce the hydrogenation/dehydrogenation times and to provide the PEM-FC with the required flow mainly. The main focus was on the internal structure to compensate for the poor thermal conductivity of the storage material. Using a parametric analysis, this structure was optimized to reduce mass, dead volume, and costs [5].

This internal structure was then implemented inside a smaller prototype storage vessel that could store up to 250 g of hydrogen and compared to the same storage vessel with the same volumetric storage capacity without any internal heat exchanger structure. Those measurements showed that the internal structure can drastically improve storage kinetics.

The results of the prototype were used to improve the design of the internal structure, with a main focus on manufacturability and assembly on different scales. This overall design approach can be seen in Figure 6.

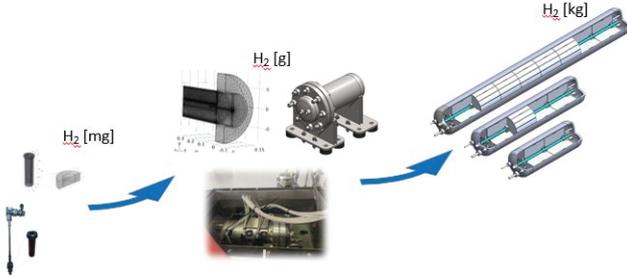


FIGURE 6: DESIGN APPROACH FOR THE MHS

Different sizes of the storage vessel (2 x 15 kg and 1 x 30 kg of storage material) were developed to build a MHS system, as shown in the flow diagram in Figure 7. This MHS system is designed to be coupled with the AEM-EL and PEM-FC, investigate the hydrogen and heat coupling, and evaluate different control strategies, as mentioned before.

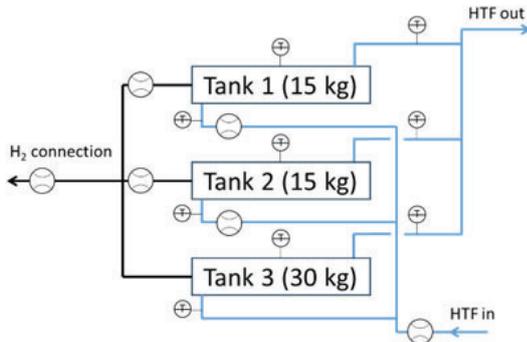


FIGURE 7: FLOW DIAGRAM OF MHS

## D. MHC

### 1) MHC material

The goal of the compressor is to use hydrogen provided by the AEM-EL or MHS at around 30 bar and supply hydrogen at pressures of up to 350 bar for heavy-duty vehicles like trains and trucks. An MHC based on two stages was developed to achieve such a compression ratio.

The hydride-forming alloy is the core of the MHC. Thus, four candidate materials have been identified to reach the target pressures at an assumed temperature operation range of the compressor between 10 and 90 °C. Two material candidates for each stage are listed in Table I, along with their respective thermodynamic properties.

TABLE I: PROPERTIES OF THE SELECTED HYDRIDE FORMING ALLOYS

Stage	1 <sup>st</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	2 <sup>nd</sup>
Material	<i>MmNi<sub>4.15</sub>Fe<sub>0.85</sub></i> (AB <sub>5</sub> )	<i>Hydralloy C5</i> (AB <sub>2</sub> )	<i>TiCr<sub>1.9</sub></i> (AB <sub>2</sub> )	<i>TiCrMn<sub>0.7</sub>Fe<sub>0.2</sub>V<sub>0.1</sub></i> (AB <sub>2</sub> )
$\Delta H_{Des}$ [kJ/molH <sub>2</sub> ]	25.3	26.5	26.19	20.6
$\Delta S_{Des}$ [J/molH <sub>2</sub> K]	105	108.3	122	107
$\Delta H_{Abs}$ [kJ/molH <sub>2</sub> ]	24.7	21.5	27	17.7
$\Delta S_{Abs}$ [J/molH <sub>2</sub> K]	107	95.5	125.4	101
H <sub>2</sub> capacity [wt.%]	0.8 - 1.1	1.1 - 1.6	0.8 - 1.9	1.0-1.95
Literature sources	[11], Lab data	Lab data	[12],[13]	[14-15]

Besides the thermodynamic properties, other considerations in the material selection are to be considered. Material kinetics and the hydrogen capacity of the material play a role in the outlet flow of the compressor. The environmental impact of mining, purification, and refining the ores required to produce such materials is also crucial, and materials based on energy-intensive raw metals like pure lanthanum or vanadium should be avoided [16-17].

### 2) MHC system

In addition to examining individual components, the MHC, as part of the SET-Unit, is also being investigated as a system. System modeling and simulation are being driven forward to create a digital twin of the compressor and thus be able to carry out a wide range of optimization and scaling considerations. This includes the investigation of different operating modes, the dimensioning of the individual compressor components, such as the tanks of the two stages with each other, the scaling of the overall system size, and, finally, the investigation of the interaction of the compressor system with other modules of the SET-Unit. Parallel to these developments, two MHCs of different sizes are being experimentally developed, constructed, and evaluated. The smaller compressor contains 1-2 kg of MH material per stage and can be used to measure the performance of the material present for the first time in a larger scale than a few grams. Figure 8 shows a flow diagram of the 1-2 kg MHC batch system.

It is designed as a simple two-stage compressor with easily dismantled components to exchange the MH materials relatively quickly. It has comprehensive measurement technology for recording all hydrogen and heat transfer medium flows, pressures, and temperatures. The construction is complete, and the first investigations are ongoing work. Figure 9 shows the build setup.

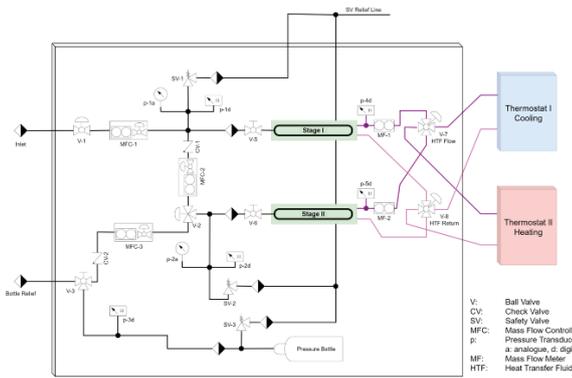


FIGURE 8: FLOW DIAGRAM OF 1-2 kg MHC BATCH SYSTEM

A further compressor on a significantly larger scale with approx. 30 kg of MH material per stage is currently being developed. It is being realized as a container structure and will be one of the few MH compressors in this application-related size range in the world. This compressor system is equipped with extensive control options and innovative concepts in its components, such as the heat exchanger in the metal hydride tank, which is crucial for its performance.

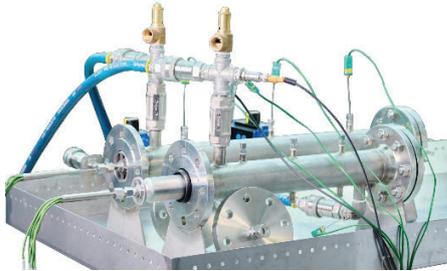


FIGURE 9: 1-2 kg MHC BATCH DEVELOPED SETUP

E. GSS

A so-called hybrid process that uses membrane and pressure swing adsorption (PSA) techniques is considered the most promising configuration. Figure 10 presents the concept of the GSS and the hybrid process. The bulk separation of hydrogen is accomplished in the membrane stage, while the following PSA process is employed to achieve the required H<sub>2</sub> purity. The hybrid process was developed using process simulation tools [5]. The model consists of a H<sub>2</sub> selective membrane stage combined with a 4-bed PSA system. Moreover, it is planned to build the GSS system in a container structure.

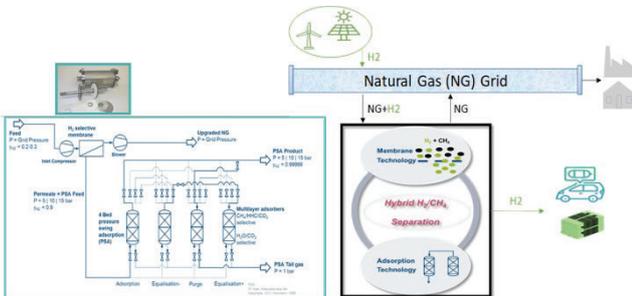


FIGURE 10: GAS SEPARATION SYSTEM CONCEPT (ADAPTED FROM [5])

III. DEVELOPMENT OF THE FIRST SET-UNIT COUPLING: TOP-DOWN

This section describes the first approach for the experimental development of the SET-Unit, including the coupling

of the AEM-EL, PEM-FC, and MHS as a top-down strategy. The coupling includes the hydrogen management and the heat exchange, as well as the conception of the control strategy to optimize the system's operation.

A. Infrastructure: Assembly of the Set-Unit

Figure 11 – left shows the building and layout at the Helmholtz-Zentrum Hereon, where the first SET-Unit approach has been developed. The spatial distribution of the current and planned test facilities, including the existing and future pipelines, is also presented (Figure 11 – right). As seen, the MHS, the AEM-EL, and an open container for energy storage are placed in one room (R106). The PEM-FC with all the required peripheral devices, a thermostat (LAUDA), the needed control cabins, and the hydrogen management (HTTF1: hydrogen tank testing facility one) are placed outside of this room in room 101. The AEM-EL–MHS–PEM-FC SET-Unit is hydrogen and thermally coupled. The hydrogen management is done through the HTTF1 and the developed control cabins. The thermal coupling is designed to improve the overall system's efficiency by using the waste heat from the AEM-EL and PEM FC through the energy storage container (open container) and the thermostat.

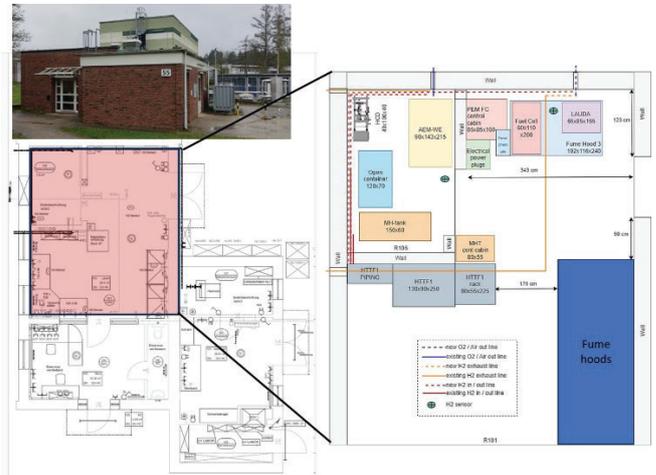


FIGURE 11: ESPACIAL DISTRIBUTION OF THE SYSTEM IN THE LABORATORY

B. Control Strategy Concept

Optimizing the control strategy of the SET-Unit is challenging because of the several system variables and market conditions that affect its performance. The approach here focuses on the thermal balance of one of the core elements: the MHS. The temperature control system is significantly responsible for the hydrogen absorption rate from the AEM-EL, the storage capacity of the MHS, the hydrogen flow and pressure to the PEM-FC, and, thus, the whole system's efficiency. As seen in Figure 13, a finite-state machine (FSM) and two types of controllers are used: a model predictive controller (MPC) and a fuzzy logic controller (FLC). The FSM ensures that each device operates within its defined states, guaranteeing safety by requiring certain conditions to be met before transitioning to the next state. This methodical approach is crucial in maintaining operational integrity and preventing accidents by strictly adhering to predefined safety protocols. The MPC is based on the system behaviour, and it determines in advance the control indication difference to the ideal operating parameters, such as pressure, temperature, and hydrogen flow, depending on external demand. The MPC also considers the objectives of the operating strategy as prefixes and the dynamic

demand, such as maximizing hydrogen storage/supply, electricity consumption/supply (peak shaving), or thermal energy storage/supply. Implementing dual fuzzy logic controllers (FLC) in the system presents an approach to managing thermal and flow dynamics complexities within the SET-Unit. The first FLC's role is balancing the internal and external heat contributions, ensuring optimal utilization of the buffer storage, and maintaining the desired flow rates of the heat transfer fluid. This is achieved by monitoring pressure, temperature, and hydrogen flow deviations of critical system performance parameters. The second FLC complements the first by assessing the operational status of the system's components, guiding the procedural operations such as the activation or deactivation of additional tanks. This evaluation considers the ideal operational parameters and incorporates empirical data, environmental factors, and market trends, which are essential for adaptive and responsive control. The control strategy's effectiveness, developed through simulation, will undergo rigorous validation, emphasizing its potential for stand-alone applications and its comparative advantage over other control methods. This comprehensive control of an intelligent energy transformation unit based on hydrogen is unique to date and promises to maximize potential synergy approaches. [17-19].

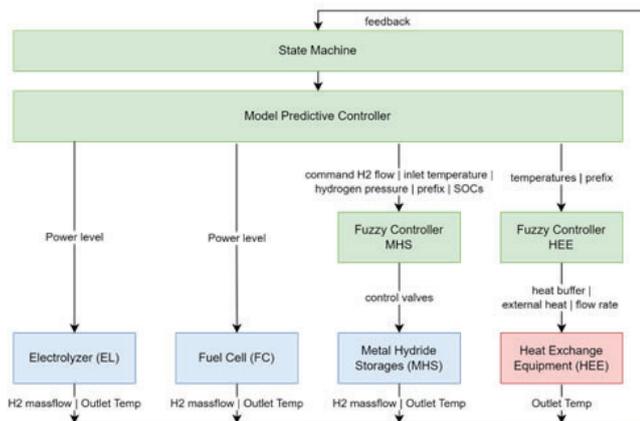


FIGURE 13: CONTROL STRATEGY FOR THE INTEGRATED SYSTEM

#### IV. CONCLUSION

This application report provides an overview about the approaches used to develop the first SET-Unit approach. The scope covers from the development and assembly in the laboratory of the different single components, i.e., MHS, MHC, AEM-EL, PEM-FC and GSS, to the hydrogen and thermal coupling of the AEM-EL–MHS–PEM-FC in their nominal scale of 5-10 kWel and employing advanced control strategies. As shown, the experimental SET-Unit is in a high degree of development and the concretization for the first proof of concept is going to be achieved soon. Following to this initial proof of concept, it is planned to scale-up the SET-Unit to a scale of 20-30 kWel and the coupling of devices such as MHC and GSS are envisioned.

#### ACKNOWLEDGEMENT

This research paper is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] L. Pasquini, K. Sakaki, E. Akiba, M. D. Allendorf, E. Alvares, J. R. Ares, et al., "Magnesium- and intermetallic alloys-based hydrides for energy storage: modelling, synthesis and properties," *Progress in Energy*, vol. 4, p. 032007, 2022, doi: <https://iopscience.iop.org/article/10.1088/2516-1083/ac7190/pdf>
- [2] J. Warfsmann, J.A. Puszkiel, M. Passing, P.S. Krause, E. Wienken, K. Taube, et al., "Applying wash coating techniques for swelling-induced stress reduction and thermal improvement in metal hydrides," *Journal of Alloys Compounds*, vol. 950, p. 169814, 2023, doi: <https://doi.org/10.1016/j.jallcom.2023.169814>
- [3] J. Puszkiel, J.M. Bellosta von Colbe, J. Jepsen, S.V. Mitrokhin, E. Movlaev, V. Verbetsky, et al., "Designing an AB<sub>2</sub>-Type Alloy (TiZr-CrMnMo) for the Hybrid Hydrogen Storage Concept," *Energies*, vol. 13, p. 2751, 2020, doi: <https://doi.org/10.3390/en1312751>
- [4] A. Mazzucco, M. Dornheim, M. Sloth, T. R. Jensen, J. O. Jensen, M. Rokni, "Bed geometries, fueling strategies and optimization of heat exchanger designs in metal hydride storage systems for automotive applications: A review", *International Journal of Hydrogen Energy*, vol. 39, p. 17054-17074, 2014, doi: <https://doi.org/10.1016/j.ijhydene.2014.08.047>
- [5] J. Puszkiel, M. Covarrubias Guraneros, L. Fleming, T. F. J. Kaufmann, P. Krause, J. H. Warfsmann, et al., "Hydrogen in stationary applications: Coupling the electricity, gas and mobility sectors (Digi-HyPro)", In: *dtec.bw – Helmut-Schmidt-Universität*, 2022, doi: <https://openhsu.uh.hsu-hh.de/handle/10.24405/14527>
- [6] P. Cavaliere, "Water Electrolysis for Hydrogen Production, Chapter: Anion Exchange Membrane Water Electrolysis," *Springer Nature Switzerland*, p. 287-307, 2023, doi: [https://doi.org/10.1007/978-3-031-37780-8\\_7](https://doi.org/10.1007/978-3-031-37780-8_7)
- [7] R. O'Hayre, S.-W. Cha, W. Colella, F. B. Prinz, "Fuel Cell Fundamentals", 3<sup>rd</sup> ed., 2016, ISBN: 9781119113805, Online ISBN:9781119191766, DOI:10.1002/9781119191766
- [8] Y. Shang, Claudio Pistidda, Gökhan Gizer, Thomas Klassen, Martin Dornheim, "Mg-based materials for hydrogen storage," *J. Magnes. Alloy*, vol. 9, p. 1837-1860, 2021, doi: <https://doi.org/10.1016/j.jma.2021.06.007>
- [9] J. Puszkiel, A. Gasnier, G. Amica, F. Gennari, "Review: Tuning LiBH<sub>4</sub> for hydrogen storage: Destabilization, Additive, and Nanoconfinement approaches", *Molecules*, vol. 25(1), p. 163, 2019, doi: <https://doi.org/10.3390/molecules25010163>
- [10] C. Blazquez-Diaz, "Techno-economic modelling and analysis of hydrogen fuelling stations," *International Journal of Hydrogen Energy*, vol. 44(2), p. 495–510, 2019, DOI: 10.1016/j.ijhydene.2018.11.001.
- [11] A. Sathesh and P. Muthukumar, "Performance investigations of a single-stage metal hydride heat pump," *International Journal of Hydrogen Energy*, vol. 35, p. 6950–6958, 2010.
- [12] C. Corgnale, J. Hattrick-Simpers, M. Sulic et al., "Metal Hydride Material Development for High Efficiency and Low Cost Hydrogen Compressors", DOE Technical Report, DOE-GWE-17076, 2018.
- [13] J. R. Johnson, "Reaction of hydrogen with the high temperature (C14) form of TiCr<sub>2</sub>", *Journal of Less-Common Metals*, vol. 73, p. 345–354, 1980.
- [14] J. Vanhanen, M. T. Hagström, and P. D. Lund, "Combined hydrogen compressing and heat transforming through metal hydrides," *International Journal of Hydrogen Energy*, vol. 24, p. 441–448, 1999.
- [15] M.T. Hagström, J.P. Vanhanen, and P.D. Lund, "AB<sub>2</sub> metal hydrides for high-pressure and narrow temperature interval applications," *Journal of Alloys Compounds*, vol. 269, p. 288–293, 1998.
- [16] P. Koltun and A. Tharumarajah, "Life Cycle Impact of Rare Earth Elements," *ISRN Metallurgy*, vol. 2014, p. 1–10, 2014.
- [17] P. Nuss and M. J. Eckelman, "Life cycle assessment of metals: a scientific synthesis," *PloS one*, vol. 9, p. e101298, 2014.
- [18] J.K. Peckol, "Introduction to Fuzzy Logic," 2021 John Wiley & Sons Ltd, DOI:10.1002/9781119772644
- [19] J. Carter, F. Chiclana, A.S. Khuman, T. Chen, "Fuzzy Logic Recent Applications and Developments," *Springer Nature Switzerland AG* 2021, doi: <https://doi.org/10.1007/978-3-030-66474-9>

# Dispatch Optimization of the Electricity and Heat of the Smart-Energy-Transform-Unit: A Residential Case Study

Carlos Muñoz\*, Nies Reininghaus, Astrid Pistor, Michael Kroener, Alexander Dyck, Martin Vehse  
*Institute of Networked Energy Systems, German Aerospace Center (DLR), Oldenburg, Germany*  
 \*carlos.munozrobinson@dlr.de

Jelto Lange\*, Martin Kaltschmitt  
*Institute of Environmental Technology and Energy Economics, Hamburg University of Technology, Hamburg, Germany*  
 \*j.lange@tuhh.de

Julián Puzskiel\*, Myriam Covarrubias Guraneros, Lukas Fleming, Thomas Friedrich Johannes Kaufmann, Phillip Krause, Jan Warfsmann, Eike Wienken, Lukas Wildner, Thomas Klassen, Julian Jepsen  
*Faculty of Mechanical Engineering, Institute of Material Engineering, Helmut Schmidt University/ University of the Federal Armed Forces, Hamburg, Germany*  
 \*puszkicj@hsu-hh.de

**Abstract** – With the integration of renewable energy production into grids, hydrogen storage is an effective solution for coping with the fluctuating nature of the resources and reliably providing energy demands. Metal hydride storage is seen as a key technology due to its low operating pressure and temperatures near ambient, while it has a significant volumetric capacity (for room temperature hydrides: 50-110 kg/m<sup>3</sup>) compared to pressurized (40 kg/m<sup>3</sup> under 700 bar and room temperature) or even liquefied hydrogen (70 kg/m<sup>3</sup> at – 253 °C and 1 bar). One potential application with metal hydride storage lies in the flexibilization of residential energy demand. Excess photovoltaic generation from a house can power an electrolyser to produce hydrogen, which is then stored in the metal hydride storage. When power and heat are needed in the building, the hydrogen is released into a fuel cell. This case study investigates the dispatch optimization of a metal hydride storage system within a residential household energy system. The interaction of the electrolyser, metal hydride storage, and fuel cell, all components of a container solution called Smart Energy Transform Unit, was studied during summer and winter. Results show that in an exemplary period in winter, from 21 December 2021 to 28 December 2021, the total electricity demand is 98% covered by supply from the grid due to the low photovoltaic generation, which also yields a low hydrogen production; the total heat demand is 90% covered by the heat pump and the thermal storage as a buffer. During an exemplary period in summer, from 20 June 2021 to 27 June 2021, the system is self-sufficient, as hydrogen was stored during the day due to the high yield of photovoltaic generation, and hydrogen is used in a fuel cell at night to provide energy demands. In addition, heat pump operation during summer is small due to the heat provided by the electrolyser, the fuel cell, and the thermal buffer storage. The PV system, together with the Smart Energy Transform Unit, covers 99% of the total electric demand during this period in summer, while for the total heat demand, a coverage of 85% is observed, and the heat pump covers 15%.

**Keywords** – Sector integration, metal hydride storage, energy system modelling, optimization

## NOMENCLATURE

Digi-HyPro	Digitalized Hydrogen Process Chain for the Energy Transition
MH	Metal hydride
PV	Photovoltaic
SET-Unit	Smart-Energy-Transform-Unit

## I. INTRODUCTION

Hydrogen storage plays a key role in balancing fluctuations between energy production from renewable sources [1] and energy demands. It can provide reliable energy storage and a continuous power supply without the dependency on fossil fuels [2]. However, there are challenges associated with the storage of hydrogen. As hydrogen has a very low density of 0,089 kg/m<sup>3</sup> at a pressure of 1 atm and a temperature of 0 °C (Standard Test Conditions) [1], storing large amounts under these conditions is challenging. High pressures of 350 and 700 bar are usually used for compressed hydrogen storage, which is then employed in mobility applications [3]. It can also be stored in cryogenic liquid form which requires a temperature of – 253 °C, but around one-third of the hydrogen energy content is required for liquefaction [4]. An alternative for hydrogen storage is based on chemically bonding the hydrogen to a powdered metal, forming a metal hydride (MH). In the MH, hydrogen absorption is an exothermic reaction, i.e. it releases heat. The desorption process, hydrogen discharge, is endothermic, i.e. it requires heat. This type of storage in MHs can operate at much lower pressures (e.g. 10 bar), and the volumetric capacity in terms of hydrogen atoms per cubic meter of material can be four times higher than compressed hydrogen and twice that for cryogenic hydrogen [5].

In the context of the project Digi-HyPro (Digitalized Hydrogen Process Chain for the Energy Transition), the goal is to develop a container solution where hydrogen is stored in a MH tank, and released when needed to supply energy demands [6]. This container solution, called Smart-Energy-Transform-Unit (SET-Unit), allows a flexible transformation of energy where hydrogen is the main energy carrier, thereby allowing the coupling of electricity, gas, heat, and mobility grid. One potential application of the SET-Unit is to provide electricity and heat for a household in form of a backup power generation unit or to increase self-consumption of renewable energies.

In this paper, the dispatch optimization of the SET-Unit for a residential energy system will be presented. This application is depicted in Figure 1 for a single-family household. Excess electricity from the solar PV system can drive the electrolyser for the production of hydrogen, which can be stored in the MH-storage. The electricity demand can be supplied by providing the hydrogen to a fuel cell. Any electricity needed can also be supplied from the grid at a defined cost. Also, excess electricity from the PV system can be sold to the grid at a certain feed-in tariff. For heat supply and heat storage, a heat pump and a thermal storage are present, respectively. In addition, heat is produced during operation of the electrolyser, the fuel cell, and during the flow of hydrogen into the MH-storage (absorption process). On the other hand, heat is consumed during the release of hydrogen from the MH-storage (desorption process) and in the household.

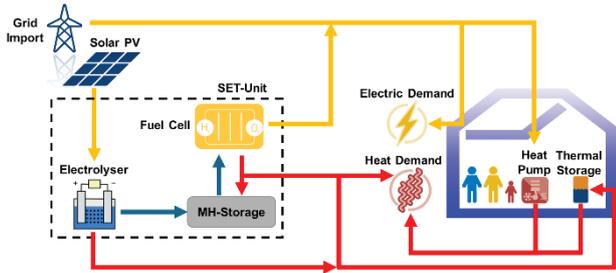


FIGURE 1. COMPONENTS OF THE SMART ENERGY TRANSFORM (SET)-UNIT IN A RESIDENTIAL APPLICATION. THE COLOURED ARROWS INDICATE ELECTRICITY (YELLOW), HYDROGEN (BLUE), AND HEAT (RED) FLOW

As there are multiple components and multiple flows (electricity, heat, and hydrogen) in the residential energy system shown in Figure 1, there are multiple paths to operate the system with the aim to supply the heat and electricity demands. Considering the boundary conditions of the residential energy system, there is a need to find an optimal dispatch strategy with the minimum operating cost. So, the main research question of this case study is: what is the optimal dispatch strategy of the SET-Unit to provide electricity and heat demands for a residential household during seven days in an exemplary winter and summer period?

The subsequent sections of this investigation are structured as follows. Method (Section II) defines all data required and steps taken to model the SET-Unit for a residential energy system. Results and Discussion (Section III) presents the main findings of the optimized dispatch strategy during summer and winter. Conclusion (Section IV) provides a summary of the relevant findings of this investigation, as well as future research topics.

## II. METHOD

An energy system model that considers the SET-Unit components interacting with the household was first built. For this purpose, the python package oemof.solph from the Open Energy Modelling Framework (oemof) was employed [7]. oemof is an open-source framework that is used to model energy systems. It is used to perform dispatch optimization, i.e. to supply the energy demands at the minimum cost by optimizing the use of the resources. It has been used in multiple investigations, including the modelling of the Italian energy system [8], a fuel cell electric vehicle providing electricity and heat to a neighbourhood [9], and designing a district heating system in a sub-urban region in northern Germany [10].

The energy system model in oemof.solph can be seen in Figure 2. Five main classes were employed: bus, sink, source, transformer, and generic storage. Sources were used to import electricity from the grid and from the PV system. Sinks were employed to supply the electricity and heat demands, and if there was excess electricity from PV. This excess PV generation can be sent to the excess electricity sink at a defined feed-in tariff for PV. An electricity bus, heat bus, and hydrogen buses had been implemented as well. Transformers, which convert one type of energy to another, were employed for the electrolyser, MH absorption process (storing hydrogen and releasing heat), MH desorption process (releasing hydrogen and need of heat), fuel cell, and heat pump. The MH and thermal storage were modelled using the generic storage class.

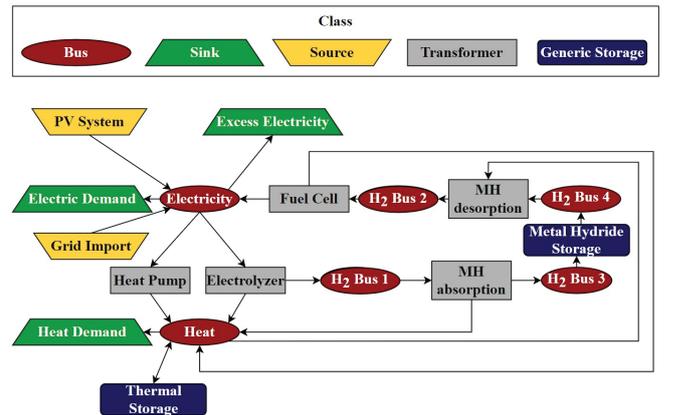


FIGURE 2. ENERGY SYSTEM MODEL IN OEMOF CONSIDERING THE SET-UNIT

The following values were used in the model. They are designed to be as close as possible to the hardware used to build the SET-Unit in the project Digi-HyPro.

### A. Electrolyser and Fuel Cell

It was assumed that the electrolyser has 9,6 kW of nominal power consumption, 62% efficiency, and 30% of waste heat utilization capacity. For the fuel cell, 5 kW of installed electrical capacity, 40% efficiency, and 60% waste heat utilization capacity were assumed. No operating costs for the electrolyser and the fuel cell were considered.

### B. MH Storage

A total of 2 kg hydrogen capacity in the MH storage was modelled. Additional details on the material used in the project Digi-HyPro can be found in [6]. The amount of heat required during absorption was assumed to be the same as that produced during desorption, with a value of 0,445 kWh of heat

per kWh of hydrogen. The initial state of charge was assumed to be zero.

### C. Thermal Storage

Stratified thermal storage was modelled with a nominal storage capacity of 35 kWh (3 m<sup>3</sup> of water). A hot water temperature of 60 °C and a cold-water temperature of 30 °C was assumed. The initial state of charge was assumed to be 100%.

### D. Heat Pump

A heat pump with a heat production capacity of 4 kW was considered. The Coefficient of Performance (COP) was assumed to be 3.

### E. Costs and Incentives

An electricity cost of 0,4627 EUR/kWh was assumed for obtaining electricity from the grid. On the other hand, excess PV generation can be sold as an incentive with a feed-in tariff of 0,086 EUR/kWh.

### F. PV Supply and Household Demands

The PV system had a nominal capacity of 10 kW<sub>peak</sub>, with an hourly profile extracted from the Photovoltaic Geographical Information System (PVGIS) [11]. A typical German single-family house of three members was assumed with an annual heat demand of 21,4 MWh and an electricity demand of 5 MWh [12]. From these demands, hourly load profiles were generated using the demandlib package from oemof [13].

The previous boundary conditions were applied to the energy system model. Minimum operating loads for the electrolyser, fuel cell and heat pump were excluded. A total of seven days during winter and seven days during summer were analysed, with an hourly time resolution. The simulation objective is to find an optimal dispatch strategy that satisfies the energy demands of the household with the minimum operating cost.

## III. RESULTS AND DISCUSSION

### A. Winter Dispatch

The electricity and heat dispatch of the residential energy system during seven days of winter can be seen in Figure 3 and Figure 4, respectively, ranging from 21 December 2021 to 28 December 2021. Figure 3 shows the electric power in the y-axis (left). A convention was defined so that an electric power value greater than zero represents consumed electricity, which is true for the electric demand (cyan), heat pump (blue), electrolyser (red), and excess electricity (green). Electric power provided by the PV system (orange), grid import (purple), and fuel cell (brown) is considered electricity produced and it has a negative value. The state of charge of the MH storage can be seen in grey, ranging from zero to one, in the secondary y-axis (right).

In Figure 4 the y-axis (left) depicts the heat power. Similar to the power convention, the heat convention was defined so that if the heat power value is greater than zero is considered heat consumed which is the case for the heat demand (magenta) and desorption process (pink). A negative heat power value is considered heat supply which applies for the heat pump (olive), fuel cell (brown), and absorption process (blue). The secondary y-axis (right) depicts the state of charge of the thermal storage in grey and it is scaling from zero to one.

In Figure 3 it can be seen that on 22 December, PV generation reached 4 kW, which was sufficient to operate the electrolyser, and hydrogen was stored in the MH storage. Again, on 26 December, a significant PV generation reaching 5 kW enable the electrolyser to produce hydrogen, which then increased the state of charge of the MH storage to 0,15. On 27 December, the strategy was to release all the hydrogen stored into the fuel cell for electricity production. Electricity from the grid was mainly needed during the seven days due to the low PV generation. Because of that, there was also no excess PV supply to the grid. 98% of the total electricity demand in the seven days in winter was covered by electricity from the grid.

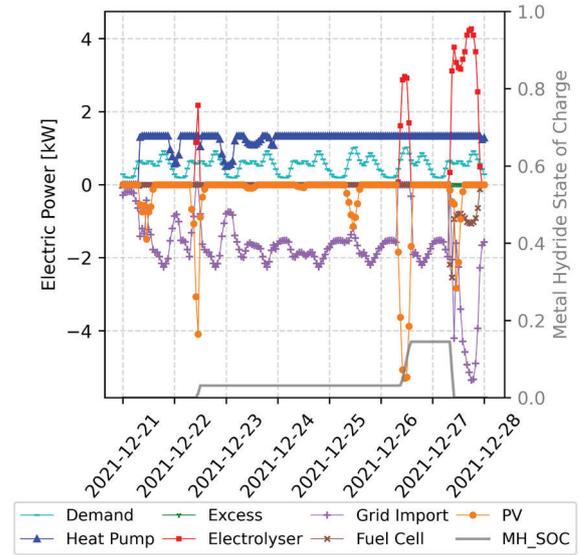


FIGURE 3. ELECTRICITY DISPATCH DURING SEVEN DAYS IN WINTER

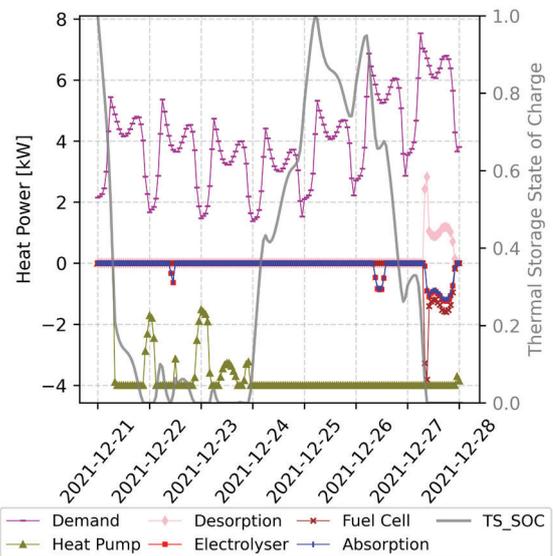


FIGURE 4. HEAT DISPATCH DURING SEVEN DAYS IN WINTER

In terms of heat, shown in Figure 4, the energy stored in the thermal storage was enough to cover the heat demand without operating the heat pump for the first twelve hours; after that, the heat pump started operating. The heat pump operated most of the given period at nominal capacity (4 kW) to provide the heat demand. As the heat demand was seen to increase after 25 December and reaching a maximum of almost

8 kW between 27 December and 28 December, the resulting strategy was to operate the heat pump at nominal capacity to supply the heat demand as well as to charge the thermal storage before the maximum peak. With this strategy, both the thermal storage and heat pump cope with the expected increase in demand. The thermal storage was fully charged between 25 December and 26 December and reached a state of charge of 90% (0,9) between 26 December and 27 December. Afterwards, it supplied part of the heat demand until it was fully discharged between 27 December and 28 December. During this period, as the heat pump capacity was insufficient to provide the heat, the resulting strategy was to provide additional heat by operating the MH's electrolyser, fuel cell, and absorption process. In addition, the desorption process also required additional heat. This simultaneous operation of the electrical devices to provide heat is why in Figure 3 during this period there is electricity grid import to satisfy the electrolyser consumption, and on the other hand, electricity production from the fuel cell while consuming hydrogen. In reality, it might not be feasible to operate the electrolyser and fuel cell at the same time, as well as simultaneously charging and discharging the MH storage. However, with the given boundary conditions, there was no other way for meeting the household heat demand during this simulation for the optimizer. During these seven days in winter, the heat pump and thermal storage covered 90% of the total heat demand.

### B. Summer Dispatch

Figure 5 and Figure 6 show the heat and electricity dispatch, respectively, during summer from 20 June 2021 to 27 June 2021. The aforementioned conventions regarding the positive and negative values of the electricity and heat flows are also applied.

A much higher PV generation compared to winter results in a longer operational time of the electrolyser, which can be seen in Figure 5. In addition, some of the excess PV generation (green) was sold to the grid. On 20 June it can be observed that part of the PV generation supplies the electrolyser, which produces hydrogen, and it is absorbed in the MH storage (increase in the state of charge). The remaining PV generation supplies the electric demand and is additionally sold as excess to the grid. At night, when there is no PV generation, the fuel cell supplies the electric demand by consuming the hydrogen coming from the MH desorption process. This operational pattern during day and night of the SET-Unit is also seen for the rest of the days. This results in an oscillating behaviour of the state of charge of the MH storage, i.e. absorption of hydrogen during day and desorption of hydrogen during night. To add, there was a small grid import and heat pump consumption, which makes the SET-Unit practically self-sufficient during summer. 99% of the total electric demand during these seven days in summer was covered by the PV system and the SET-Unit. Because the optimization period for this model is only seven days, no hydrogen is stored for longer self-sufficiency (e.g. winter). Therefore, the electrolyser was not operated at full load as seen in Figure 5. As a consequence, the optimization strategy was to sell excess PV generation to the grid with the aim of reducing operating costs. To see seasonal energy shifting, longer time periods have to be simulated.

The heat dispatch in summer shown in Figure 6 depicts small operation times of the heat pump during the seven days. The reason is due to the low heat demand together with heat provided by other components. When the electrolyser was operated during the day, the heat from it and the heat from the

absorption process covered the heat demand and some heat increased the state of charge of the thermal storage. During the fuel cell operation at night, the desorption heat required was covered by the heat produced from the fuel cell, while the thermal storage covered the heat demand. As in the electric dispatch curve in Figure 5, the SET-Unit showed a self-sufficient behaviour to cover the heat demand. 85% of the total heat demand was provided by the heat from the SET-Unit and the thermal storage, and the heat pump supplied 15%.

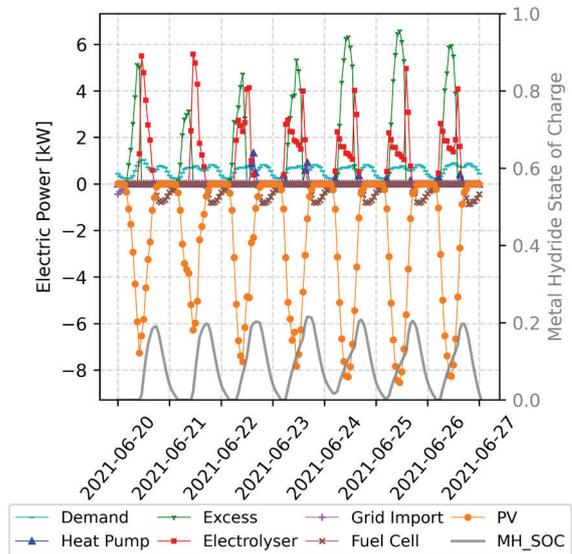


FIGURE 5. ELECTRICITY DISPATCH DURING SEVEN DAYS IN SUMMER

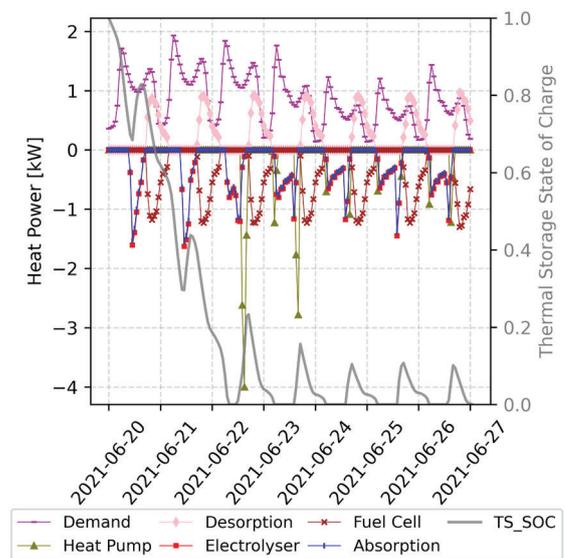


FIGURE 6. HEAT DISPATCH DURING SEVEN DAYS IN SUMMER

## IV. CONCLUSION

This case study aimed to investigate the integration of the Smart-Energy-Transform-Unit (SET-Unit) of the Digi-HyPro project into a residential household. A heat pump and thermal storage are considered to be provided in a single-family house. In this application, surplus electricity from a residential PV system powers an electrolyser, and the hydrogen produced is stored in a metal hydride storage. Then, hydrogen can be converted to electrical power by a fuel cell when power is needed.

The coverage of the residential electricity and heat demands using the SET-Unit and the interaction of the different components in summer and winter are also investigated. The main research question was: what is the optimal dispatch strategy of the SET-Unit to provide electricity and heat demands for a residential household for seven days at a stretch in winter and summer?

Results show that hydrogen production (and storage) are low during winter due to the low generation of PV. Running the electrolyser and storing hydrogen is feasible when the PV generation is larger than 4 kW. Electricity from the grid is also needed during the seven days of winter to supply the electric demand. Regarding the heat demand, the heat pump is seen to operate at full load most of the time, and as the heat demand towards the end of the seven days is expected to increase, the strategy was to operate the heat pump at full load towards the end and charge the thermal storage as much as possible. On the seventh day, as the heat pump and thermal storage capacity were insufficient to cover the high heat demand, the strategy was to obtain heat from the electrolyser and fuel cell by operating these components.

Furthermore, the results show that the SET-Unit serves as a daily buffer for the energy system during summer. During the day, hydrogen is stored due to high PV generation, and at night, hydrogen is released for electricity production. In addition, the heat produced in the electrolyser and fuel cell, and heat stored in the thermal storage is enough to cope with most of the heat demand. This results in a small heat pump operational time and small electric grid import. The residential energy system including the SET-Unit is shown to be practically self-sufficient during summer.

The framework of oemof for modelling and optimizing an energy system with the SET-Unit has the potential to aid in the investigation of further research topics. One topic involves seasonal shifting, where energy is stored in summer and in winter it can be used. A similar shift was seen in this investigation when during summer the metal hydride storage was charged during the day and discharged during night, but this shift was rather intraday. Therefore, investigating a seasonal period will involve finding an optimal size of the metal hydride storage and performing a full year dispatch optimization. Another topic is the price competitiveness of the SET-Unit relative to other components and systems. For this latter topic, capital expenditures and operating expenditures of different alternatives to metal hydride storage will be evaluated.

In addition, this case study excluded to model the detailed thermodynamics that are taking place in every component and the resultant temperature profiles. These considerations are included in current investigations of the Digi-HyPro project, specifically in the building of digital twins for system simulation.

#### ACKNOWLEDGEMENT

This research paper is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] T. Amirthan and M. Perera, "The role of storage systems in hydrogen economy: A review," *Journal of Natural Gas Science and Engineering*, vol. 108, p. 104843, 2022, doi: 10.1016/j.jngse.2022.104843.
- [2] C. Tarhan and M. A. Çil, "A study on hydrogen, the clean energy of the future: Hydrogen storage methods," *Journal of Energy Storage*, vol. 40, p. 102676, 2021, doi: 10.1016/j.est.2021.102676.
- [3] E. Rivard, M. Trudeau, and K. Zaghbi, "Hydrogen Storage for Mobility: A Review," *Materials (Basel, Switzerland)*, vol. 12, no. 12, 2019, doi: 10.3390/ma12121973.
- [4] A. T. Wijayanta, T. Oda, C. W. Purnomo, T. Kashiwagi, and M. Aziz, "Liquid hydrogen, methylcyclohexane, and ammonia as potential hydrogen storage: Comparison review," *International Journal of Hydrogen Energy*, vol. 44, no. 29, pp. 15026–15044, 2019, doi: 10.1016/j.ijhydene.2019.04.112.
- [5] Fraunhofer IFAM Dresden, "Applications of Metal Hydrides: Comparison of volumetric hydrogen storage densities," 2024. Accessed: Apr. 3 2024. [Online]. Available: <https://www.ifam.fraunhofer.de/en/Aboutus/Locations/Dresden/HydrogenTechnology/hydrides/applications-of-metal-hydrides.html>
- [6] J. Puszkiel *et al.*, "Hydrogen in stationary applications: Coupling the electricity, gas and mobility sectors (Digi-HyPro)," 2022.
- [7] U. Krien, P. Schönfeldt, J. Launer, S. Hilpert, C. Kaldemeyer, and G. Pleßmann, "oemof.solph—A model generator for linear and mixed-integer linear optimisation of energy systems," *Software Impacts*, vol. 6, p. 100028, 2020, doi: 10.1016/j.simpa.2020.100028.
- [8] M. G. Prina *et al.*, "Multi-objective investment optimization for energy system models in high temporal and spatial resolution," *Applied Energy*, vol. 264, p. 114728, 2020, doi: 10.1016/j.apenergy.2020.114728.
- [9] T. Tiedemann *et al.*, "Supplying electricity and heat to low-energy residential buildings by experimentally integrating a fuel cell electric vehicle with a docking station prototype," *Applied Energy*, vol. 362, p. 122525, 2024, doi: 10.1016/j.apenergy.2023.122525.
- [10] M. Kersten, M. Bachmann, T. Guo, and M. Kriegel, "Methodology to design district heating systems with respect to local energy potentials, CO<sub>2</sub>-emission restrictions, and federal subsidies using oemof," 2021, doi: 10.5278/ijsepm.6323.
- [11] T. Huld, R. Müller, and A. Gambardella, "A new solar radiation database for estimating PV performance in Europe and Africa," *Solar Energy*, vol. 86, no. 6, pp. 1803–1815, 2012, doi: 10.1016/j.solener.2012.03.006.
- [12] Statistisches Bundesamt (Destatis), "Umweltökonomische Gesamtrechnungen: Private Haushalte und Umwelt," 2021. Accessed: Oct. 27 2022. [Online]. Available: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Umwelt/UGR/private-haushalte/Publikationen/Downloads/haushalte-umwelt-pdf-5851319.html>
- [13] U. Krien, "demandlib," Release 0.1.9, Open Energy Modelling Framework (oemof), Mar. 2023. Accessed: Mar. 27 2024. [Online]. Available: [https://demandlib.readthedocs.io/\\_/downloads/en/stable/pdf/](https://demandlib.readthedocs.io/_/downloads/en/stable/pdf/)

# Customer-Oriented Sustainability Criteria for Value Chains in Local Production Networks of SMEs

Dominik Saubke, Julia Christina Markert, Lucas Hermann,  
Maximilian Klinkhammer, Pascal Krenz, Lisa Stoltenberg,  
Jens P. Wulfsberg

*Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg*  
Hamburg, Germany

\*info@productionnextdoor.com

**Abstract** – Today’s production sector is shaped by global supply chains and the division of labour. However, in light of recent crises showcasing the vulnerabilities of these global supply chains, a new research interest in local production has been sparked. Local value chains make it possible for the customer to be closer to where their goods are produced, which may increase people’s awareness of manufacturing circumstances. For that reason, the project Production Next Door (ProNeD) introduced the idea for the customer to be able to choose between different sustainability criteria that are used to determine the value chain through which their product will be manufactured. That way the customer is involved in the entire value creation process and may participate more actively in sustainable choices. This paper addresses the challenge of forming criteria that are relevant and meaningful to the customer but also applicable to small and medium-sized Enterprises (SME), as well as founded in standards and thus measurable. To this end, a consumer survey as well as literature on sustainability reporting is analysed to systematise key indicators for assessing sustainability within a company that meet the requirements of the European Union and its directives. This general approach is intended to also familiarise small and medium-sized enterprises (SME) with the upcoming European sustainability directives.

**Keywords** – *Local production, sustainability, small and medium-sized enterprises*

## NOMENKLATUR

SME	Small and Medium-Sized Enterprises
ProNeD	Production Next Door
VCP	Value Chain Parameter
KPI	Key Performance Indicators

## I. INTRODUCTION

Sustainability has been one of the most widely discussed topics of the last few years. With the manufacturing industry being one of the contributors of climate change with the use of resources, cause of emissions and transport of materials, just to name a few, sustainability has been quite prevalent in manufacturing research. One approach recently discussed to become more sustainable is to produce more locally to avoid long transport and to enable smaller value creation circles. This would encourage, e.g. repairing products to extend their usage. Local production close to the customer, however, may

also empower consumers by including and involving them in the value creation process.

There are various ways with different depths of involvement, in which this can be and sometimes already is done. In FabLabs or shared workshops, e.g. the consumer may actually also act as the producer [1], [2], becoming part of the value creation process. In this paper, the focus will be on ways that allow a moderate level of customer involvement in the value creation process of products that they order from a production network of micro and small enterprises (MSE). The type of network this will be targeting, and the corresponding process steps are detailed in a previous publication [3], will be briefly described in Section II.

The idea is to empower the customer by giving them the possibility to choose the underlying principles of the value chain. These could be economic principles, such as finding the cheapest or fastest value chain. Considering the growing awareness for sustainability, however, principles such as finding the value chain with the shortest transport distance could also be relevant to the customer.

This begs the question of which principles or value chain parameters (VCP) are measurable with reasonable effort and are important and relevant to the customer. To answer this, multiple sources were used to identify potential value chain criteria or parameters. First, a customer survey was conducted within the project to find criteria important to customers. Then, existing norms and standards on sustainability in SMEs were analysed to find the ones that could be used to measure the previously found criteria.

This paper is structured as follows. Firstly, background information regarding the network application Production Next Door (ProNeD) is presented. Next, we describe the methodology and approach as well as the sources for identifying potentially relevant VCPs. In Section IV these VCPs and their rating in a consumer survey are described. Section V classifies and details the VCPs chosen for implementation. Section VI introduces applicable standards and norms for sustainability criteria in SMEs and, thus, assigns trackable and measurable KPIs to the VCPs. Lastly, a brief summary and outlook will follow.

## II. USE CASE: LOCAL NETWORK PRONeD

This paper will focus on criteria fitting the use case ProNeD, or Production Next Door. ProNeD is a digital plat-

form solution designed for the demand-driven local production of goods. The concept and development are part of the Digital Urban Production research project at Helmut Schmidt University/University of the Federal Armed Forces in Hamburg. The platform is currently in the development phase, with operations scheduled to begin in 2024. Initially, the focus will be on furniture production. To this end, a production network in the furniture industry has been established in the Hamburg metropolitan area. The platform integrates advanced planning technology, which automatically contacts the partners in the value network who are best suited to fulfil a value-adding task. The value chain is reassembled for each order, dynamically responding to current demand and the capacities of the producers. Additionally, during the product development phase, global developer communities are involved through the platform.

ProNeD has the goal of benefiting both, the producers and the customers. The platform will use existing producer capabilities but bring them together more effectively through a planning tool developed specifically for easy use with low initial investment. Furthermore, the cooperation with other companies is meant to build a community and increase visibility with local potential customers. Further potential benefits of this type of collaboration are detailed in [4].

Customers can comfortably purchase high-quality, sustainable, customised products via an online platform. They also receive information on where and by whom their piece of furniture was produced. Besides the option to customise size, colour and material, they are also offered the option of choosing between different VCPs. This is meant to empower consumers to actively influence the principles with which the producers for their order are chosen.

### III. METHODOLOGY AND APPROACH

First, potential customer-relevant VCPs were brainstormed by a group of experts. These parameters were then presented to participants of a digital consumer survey. The survey contained 20 questions in total and was executed by a market research company with 300 consumers residing in Germany [5]. Through screening questions, it was ensured that the participants had bought furniture within the last three years, would consider purchasing furniture online or from a carpentry and had a household net income of €3,000 or more per month [5]. Using the results of the survey, as well as insights gained from interviewing potential producers, a few parameters were chosen to be offered to customers in the ordering process. To use the parameters, however, they need to be quantified so that value chains can be rated according to them. Therefore, 43 existing norms and standards on sustainability in SMEs were analysed according to KPIs and measurement variables.

### IV. VALUE CHAIN PARAMETERS

Relevant and suitable parameters for the ProNeD use case must be defined to offer customers a choice of VCPs. This was done in three steps. First, possible parameters were found, then rated by customers, and then modified to fit the platform requirements. This should ensure that customer opinions, which are important for success, are taken into account at the earliest possible stage, thus enabling prioritisation during the platform design phase.

#### A. Parameter Identification and Consumer Rating

As a first step, parameters that might potentially be relevant for consumers were brainstormed in a multidisciplinary workshop. Seven criteria were selected for further investigation. These criteria were presented to participants of the aforementioned consumer survey with a general description of the ProNeD scenario. The participants were asked to rank potential VCPs according to importance: “Imagine you could choose between different manufacturing criteria in the above scenario. Which of these would be of most interest to you? Sort the possible answers according to how important they would be to you personally. To do this, click on the fields in the order in which you would like to sort them. Start with the most important aspect for you.” Figure 1 shows the relative number of times each suggested parameter was chosen as the most relevant one.

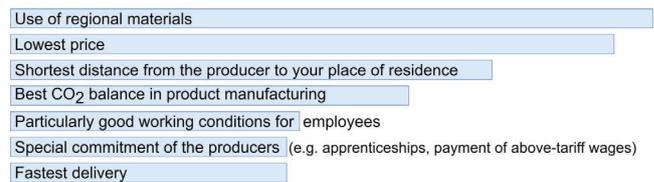


FIGURE 1: MOST RELEVANT VCPs ACCORDING TO THE CUSTOMER SURVEY

Consequently, consumers value the use of regional raw materials and a low price the most. Closeness to the producer and a good CO<sub>2</sub> balance follow with some distance in importance. Some still chose the remaining three factors as the most important, but there is a significant gap compared to the leading criteria. There is, however, no parameter that was generally viewed as unimportant by the survey participants.

#### B. Discussion of the Identified VCPs

Out of the identified parameters three were chosen to be offered to the customers in the first launch of the ProNeD platform. Some were discarded altogether, while others will be reconsidered at a later, more advanced time. An overview is provided in Table 1.

TABLE 1: OVERVIEW OF THE VCP OPTIONS

Parameter	Included	Discarded	Later
Use of regional materials	x		
Lowest price		x	
Shortest distance from the producer to place of residence			x
Best CO <sub>2</sub> balance in product manufacturing	x (modified)		
Particularly good working conditions for employees	x		
Special social commitment of the producers	x		
Fastest delivery	x		

The “*use of regional materials*” is not a parameter that is inherent to a producer, but rather a choice producers make when purchasing their raw materials. This parameter will thus be included in the customisation options of materials rather than in the VCP options. The parameter “*lowest price*”, identified in the consumer survey, was discarded for the purpose of the use case. The reason for this is that producers in the craft

sector (such as carpenters) will very likely not participate in a platform that encourages price competition as this is already an issue that they have to deal with [5]. The distance parameter will not be implemented in the first phase of ProNeD due to technical reasons, which will be detailed in the next chapter. Since the initial network will consist of only regional producers anyway, it should be sufficient to include this in future development stages with a wider network. The “*best CO<sub>2</sub> balance*” will be modified slightly in the beginning to accommodate the difficulties of measuring this for a single product. The chapter on sustainability standards will investigate this more in depth. The “*particularly good working environment*” and “*social commitment*” are very similar parameters, so they will be combined for easier application. The “*fastest delivery*” was among the lowest-rated parameters in terms of importance for the customers. In the technical planning system, however, the production time is calculated anyway because producer later in the value chain need to know, when they will be able to start their production process. Therefore, this parameter will be offered in the first implementation phase as an extra choice. It may later on be replaced by the shortest distance parameter.

## V. PARAMETERS FOR THE INITIAL IMPLEMENTATION

As shown in Table 1, five of the seven initial VCPs will be included in the initial release of the platform. They will be presented to customers in different ways, however, and have been slightly modified for practicality and marketing purposes. This is also due to the fact that the VCPs depend on different factors, meaning some are producer-specific and thus relatively static, while others change according to the customer or the product configuration. The customers can also decide to not choose a VCP, in which case the platform will randomly select one.

### A. Material-Related VCPs

The term “*use of regional materials*” was widened to “*use of sustainable materials*” since not everything can be sourced locally. Furthermore, as mentioned, this will not be selectable in the VCP options but instead in the customisable material options. Thus, this will also have a direct influence on the price of the product since the raw material prices for sustainable and ethical sourced materials are often higher.

### B. Order-Specific VCPs

These VCPs are hinged on the actual order and may thus change with every new order. This makes such parameters more difficult to include technically because the VCP ratings of the possible value chains need to be recalculated for every order to select the best one depending on the current order. One example is the distance between the customer and the producer, since this changes depending on the requested delivery address. Hence, this will be implemented at a later stage in the project. For the initial launch, no order-specific VCP is included.

### C. Producer-Specific VCPs

The advantage of producer-specific VCPs is that they only need to be pre-calculated once before the first order because they rarely change. To do so, every network producer needs to be evaluated and assigned a numerical value for their level of fulfilment of the producer-specific parameters. Then, ratings for each of these criteria for all possible value chains of every product can be calculated in advance. Thus, when a new order is placed, the value chain with the highest rating for the chosen parameter can be found immediately and production can be

started. A planning tool specifically for this application was also developed within the ProNeD project. For further information on the technology and mechanism of said planner, refer to [6].

The producer-specific VCPs chosen for the initial implementation are “*best CO<sub>2</sub> balance in product manufacturing*”, “*particularly good working conditions for employees*”, and “*special social commitment of the producers*”.

The “*best CO<sub>2</sub> balance in product manufacturing*” was also modified to be more inclusive of other environmentally relevant emissions and measures to minimise the use of resources such as electricity and water. The wording was thus changed to “*production with the lowest emissions and environmentally friendly production*”.

The “*particularly good working conditions for employees*” and “*special social commitment of the producers*” will be, as mentioned before, combined to “*social commitment in the region and family-friendliness*”. The difference between these parameters is not pronounced enough to differentiate them in a meaningful way.

The “*fastest delivery*” is a hybrid parameter, since it depends on the time of the order and the current open production capacities, but also on the speed at which producers can manufacture the product. This may vary significantly depending on personnel, but especially the available machines. Since the initial platform will not be capable of tracking producer capacities, this parameter is reliant on the production time predictions made by each producer. Thus, for the platform launch, this parameter can be considered producer-specific.

## VI. APPLYING TRACKABLE KPIS AND MEASURABLE VARIABLES TO THE VCPS

An essential aspect of the selection of VCPs is, of course, making them transparent and assessable. This transparency is particularly important for the customer so that they understand precisely what is behind the parameter. On the other hand, the exact evaluation basis is of particular interest to the producers in order to increase the company's own value in the selection of parameters by the customer and the associated formation of the value chains; ergo to be considered more often in the selection of value chains.

### A. Regulations and Standards for SMEs

To establish a solid foundation for this analysis, the current standards for fulfilling the sustainability reporting obligation have been applied analogously to the European reporting obligation. Directive 2014/95/EU from 22 October 2014, states: “The disclosure of non-financial information is a key element in managing the transition to a sustainable global economy by combining long-term profitability with social justice and environmental protection. In this context, the disclosure of non-financial information helps to measure, monitor and manage the business performance of companies and their impact on society” [7]. In this context, the term sustainability controlling is relevant. Controlling, as an umbrella term, describes the planning, control, as well as management and steering of economic processes, where key figures serve as an essential tool [8]. Accordingly, sustainability controlling includes measures for the effective management of the non-financial performance of companies and organizations, taking into account Environmental, Social and Governance (ESG) aspects. Companies and their responsibility towards society have gained significant importance [9]. This responsibility is reflected in

the documentation and disclosure of a company's individual impact on the environment in the form of a sustainability report. With the entry into force of the CSR Directive 2014/95/EU in 2017, non-financial reporting has gained increasing importance [10]. CSR stands for “Corporate Social Responsibility” [11]. The requirements for sustainability reports include at least the topics of social and employee matters, environmental issues, human rights and the fight against corruption and bribery [12].

In line with the recommendations of the CSR committees, existing frameworks that cover all the required aspects should be used for guidance. In this context, the Global Reporting Initiative (GRI) and the German Sustainability Code (DNK) should be mentioned [10]. Furthermore, the EMAS (Eco-Management and Audit Scheme) is particularly suitable for use in SMEs. It is an environmental management system introduced by the European Commission as an instrument for assessing the environmental performance of companies [13], whereby EMAS refers to Regulation (EC) 1221/2009 and came into force on 11 January 2010 [14]. The European Federation of Financial Analyst Societies (EFFAS) created a framework in 2010 that provides guidelines for incorporating ESG (environmental, social, governance) factors into financial analyses and company valuations, which are utilized in sustainability reporting [15]. The International Organization for Standardization (ISO) is a non-governmental, independent organization. It brings together expertise from its international members to promote knowledge exchange across national and regional borders, with the aim of collaboratively developing market-relevant international standards. These standards are designed to foster innovation and address global challenges [16]. Notable among the environmental and sustainability standards are ISO 14001 and ISO 26000. ISO 14001 is an internationally recognized standard for environmental management systems, providing a framework for organizations to design and continually improve their environmental performance. These standards were also considered to incorporate appropriate detailed descriptions into the VCPs.

Based on this overview, a literature review was initiated to analyse the examined standards with regard to various publications on different topics. A total number of 43 different guidelines were identified and included in the detailed analysis. The list comprises 13 titles from the GRI standard, 14 publications from the EFFAS standard, 11 publications from the EMAS standard and 5 international standards. During the initial research, it quickly became evident that both the ISO standards and the guidelines of the GRI standard place a clear focus on industrial companies. As this work is a platform for designing value chains in a network of small and medium-sized enterprises, these standards were therefore not taken into further consideration. Nevertheless, for the reason of completeness, they are presented as a whole to emphasise the accuracy of the research. Table 2 provides an overview of the sources considered. The title and a heading have been included in each considered publication or guideline in one of the stated standards. the headlines of the publications have been partially shortened for the sake of clarity.

TABLE 2: OVERVIEW OF THE VCP OPTIONS

Standard	Titel	Source
GRI 201	economic performance	[17]
GRI 205	anti-corruption	[17]
GRI 301	materials	[17]
GRI 302	Energy and water	[17]
GRI 303	water and wastewater	[17]
GRI 305	emissions	[17]
GRI 306	waste	[17]
GRI 403	health and safety at work	[17]
GRI 404	training and further education	[17]
GRI 405	diversity and equal opportunities	[17]
GRI 406	non-discrimination	[17]
GRI 414	social assessment of suppliers	[17]
GRI 415	political influence	[17]
EFFAS S06	supplier agreements /supply chain	[18]
EFFAS E13	eco-design	[18]
EFFAS V04	innovation	[18]
EFFAS E01	energy efficiency	[18]
EFFAS E04	waste scope I	[18]
EFFAS E05	waste scope II	[18]
EFFAS E02	greenhouse gas emissions	[18]
EFFAS S03	maturity of the labour force	[18]
EFFAS S10	diversity	[18]
EFFAS S02	training and qualification	[18]
EFFAS S07	certification of facilities	[18]
EFFAS G01	contributions to political parties	[18]
EFFAS V01	risks of legal disputes	[18]
EFFAS V02	corruption risks	[18]
EMAS EN	energy efficiency	[19]
EMAS M	material efficiency	[19]
EMAS W	water efficiency	[19]
EMAS A	waste	[19]
EMAS B	biodiversity	[19]
EMAS EM	emissions	[19]
EMAS Go	overall organisation	[19]
EMAS BS	procurement	[19]
EMAS PV	product responsibility and sales	[19]
EMAS Pm	human resources management	[19]
EMAS Km	communication and marketing	[19]
ISO 14001	environmental management systems EMS	[20]
ISO 14006	EMS-guidelines for [...] product design	[21]
ISO 14031	environmental performance evaluation	[22]
ISO 14050	environmental management - terms	[23]
ISO 26000	guideline on social responsibility	[24]

### B. Specification of the Object

When analysing the standards, it is evident that they place a clear focus on comparability and trackability. This is understandable as they serve as the basis for reporting to a regulatory body, in this case the representative of the elected authority in which a company operates. The document therefore usually contains indicator sets that make various aspects of sustainability trackable. In general, key indicators provide quantitative information used to represent structures and processes within a company or its various areas. They summarize fundamental data into meaningful metrics. In this work, the terms key figures and performance indicators are used synonymously. These metrics are collections of numbers that often result from differentiating or summarising the outcomes of various activities or linking specific information [25]. Statistically, key figures are divided into absolute and relative figures. Absolute key figures indicate the quantitative scope of a specified quantity and provide information on how many elements this quantity consists of. In contrast, relative key figures, also known as ratios, are created by meaningfully linking facts in the form of a quotient [26]. Measurement variables, on the other hand, are to be understood as much softer, they are rather descriptions that appear in the documents and paraphrase the

selected KPIs or make them easier to understand, but at the same time also express the companies' internal endeavours to consistently improve their own sustainability.

### C. Performing the Analysis

Starting with the two leading VCPs selected by the customer survey for the initial implementation, a search was now conducted for KPIs and measurement variables in the selected standards. The following have been selected:

*VCP - I. Social commitment in the region and family-friendliness.*

*VCP - II. Production with the lowest emissions and environmentally friendly production.*

For the description of the *VCP - I. Social commitment in the region and family-friendliness*, seven different parameters were selected as more detailed descriptors; these are listed in detail in Table 3, and each is coded in column 2.

TABLE 3: OVERVIEW OF THE VCP-I. OPTIONS

Overall	Code	Description
VCP - I. Social commitment in the region and family-friendliness	VCP-I / A	The company acts as a training organisation, providing comprehensive training and instruction in the relevant specialist area. The company assumes responsibility for imparting the necessary technical knowledge and practical skills required for a successful state-recognised vocational qualification.
	VCP-I / B	The company is officially involved in the regional training company in the subject area and is involved in structural, organisational or other processes to promote training across company boundaries.
	VCP-I / C	The company or entrepreneur volunteers for the guild or the Chamber of Crafts to strengthen their own speciality regionally.
	VCP-I / D	An officially authorised person working in the company acts voluntarily as an ambassador for the craftsman sector to represent the own speciality (e.g. carpenter) in this role and actively contribute to the promotion and recognition of it in society.
	VCP-I / E	The company actively contributes to the visibility of its speciality. This is achieved by offering school internships and familiarisation days for future trainees. Such measures enable prospective adults to gain an early insight into the professional requirements and activities of the speciality.
	VCP-I / F	The company is actively committed to promoting equal opportunities by, for example, pursuing the targeted promotion of diversity in the workplace, participating in initiatives such as Girls' Day, supporting socially disadvantaged people and integrating refugees.
	VCP-I / G	The company is involved regionally or nationally on a voluntary or social basis in various areas of social life. Such commitment can manifest itself in a variety of ways, for example by sponsoring jerseys for local sports clubs or actively supporting community projects by utilising the company's resources.

The following aspects, seen in Table 4, were selected to quantify *VCP - II Production with the lowest emissions and environmentally friendly production*. The tabular presentation follows the categorisation in Table 3.

TABLE 4: OVERVIEW OF THE VCP-II. OPTIONS

Overall	Code	Description
VCP - II. Production with the lowest emissions and environmentally friendly production.	VCP-II / A	The company integrates the methodologies of the 7-R principles into its business model and, by analogy, carries out value-adding activities such as repairs or maintenance when these are appropriate to its field of activity.
	VCP-II / B	The company independently takes comprehensive measures to protect the environment, developing context-specific concepts to reduce the use of resources in the value chain as well as application measures to implement these in its own organisation.
	VCP-II / C	The company develops its own concepts for ecological energy supply, for example, green electricity is purchased and measures such as the use of a photovoltaic system are part of the company's overall concept for ecological energy utilisation
	VCP-II / D	The company has a cross-value chain strategy for holistic ecological energy procurement, energy utilisation and energy feed-in. It also includes cooperation with other companies along its own value chain.
	VCP-II / E	The company is familiar with the methodologies of the 7-R principles, and is integrating the methods through, for example, an on-site energy evaluation to avoid transport, will be a sensible option if no better overall impact on increasing sustainability can be achieved through other R methodologies [27].
	VCP-II / F	As part of the internal value chain, the company pays meticulous attention to avoiding potentially environmentally harmful materials in its own value creation process and takes the utmost care when handling such materials if they are indispensable.
	VCP-II / G	The company pursues a holistic strategy to reduce the total number of journeys and transport routes required in the context of operational value creation in a structured manner, which will thus have a positive influence on the company's overall impact on the environment (traffic, CO <sub>2</sub> emissions).

The seven descriptive parameters A to G for each VCP listed in Tables 3 and 4 were defined through a comparative evaluation of the relevant norms and standards that had been selected. The decision was based on the premise of maximising customer impact and selecting criteria that are easily understandable and broadly accepted by customers.

## VII. SUMMARY AND OUTLOOK

As part of this scientific work, a systematic approach was discussed and implemented to identify customer-oriented sustainability criteria for value creation in local production networks of small and medium sized Enterprises (SME). The aim is to involve customers in value creation decisions at an early stage. The goal of selecting value chain parameter (VCP) for the application context and quantifying them more precisely was achieved with the approach described. The application context comprises a new type of platform concept capable of arranging value chains in a demand-oriented manner and selecting the optimum form according to requirements. As part of a pilot project, a technology offshoot of this system, Production Next Door (ProNeD), is being installed on a platform and put online. The selected value chain parameters (VCP) will therefore be trialled in reality in the near future. For the context of Production Next Door (ProNeD), four value chain parameters (VCP) were selected through a survey of potential customers. Through a successful quantification of the two producer-specific parameters by means of thorough research in the areas of sustainable corporate orientation (ISO standards)

and common EU directive-compliant sustainability reporting standards, both value chain parameters (VCP) were presented in more detail with seven descriptors. This is also a point of criticism: A finer granularity would possibly have allowed for an even better differentiation and, thus, evaluation of the companies. In general, more value chain parameters (VCP) should be implemented in the future to achieve a greater impact of customer choice. In addition, a transfer to the global reporting initiative (GRI) standard could be of interest to create a generally valid list that can be used in different application contexts. The next step is to implement the already mentioned realisation of the selected customer-oriented sustainability criteria in the new shop environment.

#### ACKNOWLEDGEMENT

This paper is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge [project Digital Urban Production]. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] L. Stoltenberg, P. Krenz, and T. Redlich, "Prosuming for a sustainable future? Effects of different forms of participatory value creation with respect to sustainability," *Discov Sustain*, vol. 5, no. 1, 2024, doi: 10.1007/s43621-024-00279-2.
- [2] B. Lange, "Offene Werkstätten und Postwachstumsökonomien: kollaborative Orte als Wegbereiter transformativer Wirtschaftsentwicklungen?," *Zeitschrift für Wirtschaftsgeographie*, vol. 61, no. 1, pp. 38–55, 2017, doi: 10.1515/zfw-2016-0029.
- [3] P. Krenz, L. Stoltenberg, D. Saubke, and J. Markert, "Production Next Door – A business model for local, sustainable production," in *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg*, Band 1, Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw, D. Schulz, A. Fay, W. Matiaske, and M. Schulz, Eds., Hamburg: Helmut-Schmidt-Universität Bibliothek, 2022, pp. 64–69.
- [4] J. C. Markert and P. Krenz, "Local Production Networks of SMEs: The Future of Producing Locally?," in *SDG - Forschung, Konzepte, Lösungsansätze zur Nachhaltigkeit, Global collaboration, local production: Fab City als Modell für Kreislaufwirtschaft und nachhaltige Entwicklung*, M. Moritz, T. Redlich, S. Buxbaum-Conradi, and J. P. Wulfsberg, Eds., Wiesbaden: Springer Gabler, 2024, pp. 281–295.
- [5] L. Stoltenberg, P. Krenz, J. Markert, and D. Saubke, "Acceptance of Local Production among Regional Stakeholders. Results of Qualitative and Quantitative Research of Producers and Consumers," in *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg*, Band 1, dtec.bw, Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw, D. Schulz, A. Fay, W. Matiaske, and M. Schulz, Eds., Hamburg: Helmut-Schmidt-Universität Bibliothek, 2022, pp. 70–75.
- [6] E. Aghajanzadeh et al., *Planning And Optimising Value Chains In Production Networks Of MSEs: A Lightweight Planner For Parallel Processes*, in press.
- [7] P. Office, "RICHTLINIE 2014/95/EU DESEUROPÄISCHEN PARLAMENTES UND DES RATES - vom 22. Oktober 2014 - zur Änderung der Richtlinie 2013/34/EU im Hinblick auf die Angabe nichtfinanzieller und die Diversität betreffender Informationen durch bestimmte große Unternehmen und Gruppen -," 2014. [Online]. Available: <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32014L0095>
- [8] B. Wolfgang and P. Ulrich, Eds., *Praxishandbuch Controlling*. Wiesbaden: Springer Gabler, 2016.
- [9] Michael S. Abländer, *Handbuch Wirtschaftsethik*, 2nd ed. Berlin: J.B. Metzler ein Teil von Springer Nature, 2021.
- [10] Bernhard Schwager, *CSR und Nachhaltigkeitsstandards: Normung und Standards im Nachhaltigkeitskontext*. B: Springer Gabler, 2022.
- [11] B. Sheehy, "Defining CSR: Problems and Solutions," *J Bus Ethics*, vol. 131, no. 3, pp. 625–648, 2015, doi: 10.1007/s10551-014-2281-x.
- [12] Y. Zwick and K. Jeromin, Eds., *Mit Sustainable Finance die Transformation dynamisieren: Wie Finanzwirtschaft nachhaltiges Wirtschaften ermöglicht*. Wiesbaden, Germany, Heidelberg: Springer Gabler, 2023. [Online]. Available: <http://www.springer.com/>
- [13] E. g. Umweltmanagement, *Was ist EMAS: Umweltmanagementsystem EMAS*. [Online]. Available: <https://www.emas.de/was-ist-emas> (accessed: Jan. 22 2024).
- [14] E. g. Umweltmanagement, *Rechtliches: Umweltmanagementsystem EMAS*. [Online]. Available: <https://www.emas.de/rechtliches> (accessed: Jan. 24 2024).
- [15] R. Erchinger, R. Koch, and R. B. Schlemminger, *ESG (E)-Kriterien - die Schlüssel zum Aufbau einer nachhaltigen Unternehmensführung: Eine Eignungsanalyse ausgewählter Standardkriterien*. Wiesbaden, Heidelberg: Springer Gabler, 2022. [Online]. Available: <http://www.springer.com/>
- [16] ISO - International Organization for Standardization, *ISO - About us*. [Online]. Available: <https://www.iso.org/about-us.html> (accessed: May 23 2024).
- [17] Global Sustainability Standards Board (GSSB), *GRI - GRI Standards English Language*. [Online]. Available: <https://www.globalreporting.org/how-to-use-the-gri-standards/gri-standards-english-language/> (accessed: Jul. 15 2024).
- [18] EFFAS, Home - *EFFAS*. [Online]. Available: <https://effas.com/> (accessed: Jan. 22 2024).
- [19] Bundesfinanzministerium, *Umweltmanagementsystem EMAS - Bundesfinanzministerium - Ministerium*. [Online]. Available: <https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Ministerium/Umweltmanagementsystem/umweltmanagementsystem-emas.html> (accessed: Jan. 22 2024).
- [20] ISO 14001 - Environmental management systems – Requirements with guidance for use: International Standard, International Organization for Standardization.
- [21] ISO 14006 - Environmental management systems - Guidelines for incorporating ecodesign: International Standard, International Organization for Standardization.
- [21] ISO 14031 - Environmental management - Environmental performance evaluation - Guidelines: International Standard, International Organization for Standardization.
- [22] ISO 14031 - Environmental management - Environmental performance evaluation - Guidelines: International Standard, International Organization for Standardization.
- [23] ISO 14050 - Environmental management - Vocabulary: International Standard, International Organization for Standardization.
- [24] ISO 26000 - Guidance on social responsibility: International Standard, International Organization for Standardization.
- [25] R. Ewert, A. Wagenhofer, and A. Rohlfing-Bastian, *Interne Unternehmensrechnung*, 9th ed. Berlin, Heidelberg: SpringerGabler, 2023. [Online]. Available: <http://www.springer.com/>
- [26] W. Gladen, *Kennzahlen- und Berichtssysteme: Grundlagen zum Performance Measurement*, 2nd ed. Wiesbaden: Gabler Verlag, 2003.
- [27] D. Reike, W. J. Vermeulen, and S. Witjes, "The circular economy: New or Refurbished as CE 3.0? — Exploring Controversies in the Conceptualization of the Circular Economy through a Focus on History and Resource Value Retention Options," *Resources, Conservation and Recycling*, vol. 135, pp. 246–264, 2018, doi: 10.1016/j.resconrec.2017.08.027.

# Auf dem Weg in die Kreislaufwirtschaft – Die Bedeutung der Konsument\*innen für eine lokale, nachhaltige Wertschöpfung

Lisa Stoltenberg\*, Pascal Krenz, Tobias Redlich

Fakultät für Maschinenbau, New Production Institut  
Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg  
Hamburg, Deutschland

\*Lisa.Stoltenberg@hsu-hh.de

**Kurzfassung** – Produktion und Konsum leisten einen hohen Beitrag an den vom Menschen verursachten Umweltschäden und strapazieren so die planetaren Grenzen. Aus diesem Grund wurden sie in den Fokus internationaler Bemühungen um mehr Nachhaltigkeit gerückt und als eine der 17 Nachhaltigkeitsziele der UN aufgenommen. Produktion und Konsum verursachen zahlreiche Schäden sowohl an der Umwelt als auch an den Menschen und strapazieren so die planetaren Grenzen. Daher wird versucht Stoffkreisläufe zu etablieren, durch die Materialien und Ressourcen möglichst lange wiederverwendet werden können. Dabei wird sich auf die Etablierung und Schließung von Kreisläufen konzentriert. Gleichzeitig wird die industrielle Wertschöpfung als Ausgangspunkt betrachtet. Dadurch wird jedoch übersehen, dass die konsequente Verkleinerung von Kreisläufen notwendig ist, um den Effekt der Schließung von Kreisläufen noch zu verstärken. Die Verkleinerung der Kreisläufe kann z. B. durch die Reduktion der Wertschöpfungsstufen und die Dislozierung der wertschöpfenden Akteure erreicht werden. Dieser Beitrag stellt zwei Möglichkeiten vor, dies zu erreichen: Die lokale Produktion und die Einbeziehung der Nutzer\*innen in die Wertschöpfung.

**Stichworte** – Circular Economy, Lokale Produktion, Prosuming

## I. EINLEITUNG

Konsum und Produktion erfolgen bislang vor allem im Rahmen linearer Prozesse. Die Herstellung eines Produkts bedarf dabei zahlreicher Wertschöpfungsstufen, in denen jeweils Ressourcen verbraucht und Emissionen produziert werden. Nach der Nutzung wird ein Produkt i. d. R. entsorgt und die gesamte erbrachte Wertschöpfungsleistung mit ihm. Die Verkettung dieser Herstellungsprozesse und der Nutzung ist häufig mit globalen Transportwegen verbunden, wodurch hohe Mengen an CO<sub>2</sub> ausgestoßen werden. Die Entsorgung der Produkte erzeugt Abfälle, wiederum Treibhausgase und verschmutzt Gewässer und Landschaften und strapazieren somit fortwährend die planetaren Grenzen [1]. Beispielsweise wurde laut Angaben der EU für den Textilverbrauch eines durchschnittlichen EU-Bürgers/ einer durchschnittlichen EU-Bürgerin in einem Jahr 400 m<sup>2</sup> Land verbraucht, 9 m<sup>3</sup> Wasser und 391 kg Rohstoffe, sowie 270 kg CO<sub>2</sub> ausgestoßen [2]. So gehen etwa 20 % der globalen Wasserverschmutzung auf die Textilproduktion zurück, viele der für die Textilproduktion genutzten Stoffe sind giftig und gefährden so die Gesundheit von Mensch, Tier und Pflanzenwelt. Bisherige Anstrengungen

diese Effekte zu minimieren, setzen in der Regel auf die Steigerung von Effizienz [3], um weiterhin Wachstum zu gewährleisten, wie Stichworte wie "Grünes Wachstum" oder „grüne Technologien“ zeigen [4].

Allerdings werden diese Wirkungen durch den immer weiter steigenden Konsum und Reboundeffekte, wie eine (Mehrnutzung oder intensivere Nutzung) infragegestellt [5], [4], [6].

## II. KREISLAUFWIRTSCHAFT ALS LÖSUNG FÜR DIE FOLGEN VON KONSUM UND PRODUKTION

Aus diesem Grund werden Optionen gesucht, um die Umweltschäden durch Produktion und Konsum signifikant zu verringern. Hierzu setzt sich allmählich der Ansatz durch, Wertschöpfung zukünftig in geschlossener Stoffkreisläufen stattfinden zu lassen. Dabei existiert bislang keine einheitliche Definition einer solchen Kreislaufwirtschaft bzw. Circular Economy (CE) [7], [8].

Das europäische Parlament definiert die Kreislaufwirtschaft beispielsweise als „ein Modell der Produktion und des Verbrauchs, bei dem bestehende Materialien und Produkte so lange wie möglich geteilt, geleast, wiederverwendet, repariert, aufgearbeitet und recycelt werden. Auf diese Weise wird der Lebenszyklus der Produkte verlängert.“ [9].

Die Circular Economy Initiative Deutschland beschreibt die Kreislaufwirtschaft als „notwendigen systemischen Strukturwandel von einer linearen, ressourcenintensiven hin zu einer zirkulären ressourcenschonenden Wertschöpfung.“

Die Idee der Kreislaufwirtschaft zielt also in der Regel darauf ab, die Lebenszeit der Produkte und Ressourcen weitgehend zu verlängern, indem Materialien und Ressourcen nach der Nutzung aufbereitet und wiederverwendet werden sowie Produkte oder Bauteile repariert, aufgewertet oder mit neuer Funktion fortwährend in neue Nutzungsprozesse überführt werden [10]. Damit stellt dies einen alternativen Ansatz zu linearen Wertschöpfungsmodellen dar, in denen die in einem Produkt enthaltene Wertschöpfungsleistung, die durch Arbeit, Energie und Ressourcen aufgebracht wurde, i. d. R. mit dem Ende der Nutzung verloren geht. Durch die Kreislaufförmigkeit soll diese auch nach der Nutzung weiterhin erhalten bleiben, um Ressourcen zu sparen und Emissionen zu reduzieren. Dies würde es ermöglichen die Wertschöpfungsleistung auch möglichst lange nach der Nutzung zu erhalten [10]. Dadurch,

verringert sich nicht nur der Einsatz neuer Ressourcen und Materialien, auch wird davon ausgegangen, dass die Aufbereitung von Ressourcen energieeffizienter ist als die Verwendung neuer Ressourcen [10]. Kreislaufwirtschaft zielt folglich auf die Schließung von Stoffkreisläufen ab [11] und bildet so eine Alternative zu den aktuell noch dominierenden linearen Wertschöpfungsmodellen [11].

Um diese Kreislaufförmigkeit zu erreichen, werden verschiedene Strategien zur Weiter- und Wiederverwendung von Materialien und Ressourcen empfohlen (9R-Strategien), die entsprechend der jeweiligen Phase des Produktlebenszyklus bestehen.

Tabelle I: Das 9R Framework [7]

Produktlebensphase	Zirkularitätslevel	Strategie
Ressourcenentnahme	0	Refuse
	1	Rethink
	2	Reduce
Nutzung	3	Reuse
	4	Repair
	5	Refurbish
	6	Remanufacture
	7	Repurpose
Nutzungsende	8	Recycle
	9	Recover

Entsprechend dieser Grundlagen wird auf europäischer Ebene als auch in den Nationalstaaten als auch auf regionaler Ebene derzeit an Aktionsplänen und Maßnahmen zur Etablierung einer Kreislaufwirtschaft gearbeitet. Vor diesem Hintergrund wurden verschiedene Pläne verglichen, darunter z. B. ein Aktionsplan für Deutschland der Circular Economy Initiative [12] sowie ein Maßnahmenkatalog für die Textilproduktion von der EU [13] als auch eine Analyse zu CE-Maßnahmen in verschiedenen Städten [14], um einen Überblick über bisherige Anstrengungen zu gewinnen. Die geleisteten und angestrebten Maßnahmen beziehen sich dabei auf verschiedene gesellschaftliche Handlungsfelder:

- **Wissen und Informationspolitik** (z. B. Aufklärung, Schaffung von Netzwerken, Forschung, Transparenz)
- **Technologieförderung und -verbreitung** (z. B. Recyclingtechnologien, Digitalisierung)
- **Anreize für Unternehmen und Geschäftsmodelle** (z. B. Förderung von Innovationen und Start-Ups, Steuerliche Vorteile)
- **Richtlinien und Verordnungen** (z. B. Recyclingvorgaben, Evaluationsstandards, Abfallvermeidungsvorgaben)
- **Angebote und Services zur Umsetzung der 9R-Strategien** (z. B. Sharing-Angebote, Reparaturservices, Rücknahmeservices etc.)
- **Wertschöpfungsprinzipien** (z. B. lokale Produktion, Prosuming, Nachhaltiges Design)

Dabei fiel auf, dass die Schwerpunkte der Circular Economy Initiative auf der Ausarbeitung von Regularien und Anordnungen sowie auf der Technologieförderung liegen. Die Städte, die Petit-Boix und Leipold [14] untersuchten, legten Wert auf die Bereitstellungen von Services und Angeboten sowie Infrastrukturen. In dem Rahmenplan für die Textilproduktion seitens der EU war kein klarer Schwerpunkt zu erkennen.

Auffällig ist dabei, dass sich die Aktionspläne und Handlungsempfehlungen häufig auf industrielle Wertschöpfungs-systeme beziehen. So beschreibt die CE Initiative für Deutschland als ersten von zehn Handlungsschwerpunkten, dass aufbauend auf den Ansätzen der Industrie 4.0 ein ressourcenproduktives, datengetriebenes zirkuläres Wirtschaftsmodell zu entwickeln ist [12]. Dabei wird eine sektorspezifische, hoch technologisierte Zusammenarbeit von Industrieunternehmen beschrieben, die gemeinsam an Innovationen arbeiten [12]. Die Wissenschaftliche Gesellschaft für Produktionstechnik (WGP) skizziert die Umsetzung dieser Idee einer industriellen Kreislaufwirtschaft mittels des Konzeptes der Update-Fabrik. Im Rahmen dieser sollen Produktionsstätten dahingehend gestaltet werden, dass dort nicht nur linear produziert wird, sondern auch Materialien wieder zurückgenommen und aufgearbeitet oder Bauteile neu verbaut werden [15]. Es sollen also produzierende Unternehmen umgestaltet werden, so dass sie in ihren Fertigungsprozessen die 9R-Strategien berücksichtigen. Zwar werden unter dem letzten Handlungspunkten der Circular Economy Initiative Deutschland auch andere Wertschöpfungsmodelle genannt (Offene Werkstätten, Reparatur Initiativen etc.). Diese werden im Konzept aber nicht im Bereich der primären Wertschöpfung verordnet, sondern vorrangig im Bereich Weiterbildung. Insgesamt wird dort bereits in der Einleitung auf die Schaffung von neuen Geschäftsmodellen und die Schaffung einer „neuen zirkulären Industrie“ verwiesen, die sich vor allem auf neue Dienstleistungen fokussieren soll [12].

Auch das von der EU verabschiedete „Recht auf Reparatur“, weist die Logik der industriellen Wertschöpfung auf. So verweist es vor allem auf Ansprüche von Verbraucher\*innen, Reparaturdienstleistungen von Dritten in Anspruch nehmen zu können [16]. Defekte Geräte sollen also zu den Herstellern oder Reparaturdienstleistern gebracht werden, um dort von diesen erneuert/überarbeitet zu werden. Auch hier werden wieder die herstellenden Unternehmen oder Dienstleister als die für die Reparatur zuständigen Akteure betrachtet. Ebenso beinhaltet der Aktionsplan der EU für eine Kreislaufwirtschaft in erster Linie Maßnahmen, die auf die Regulierung von Unternehmenspraktiken abzielen, wie etwa Anforderungen an Verpackungen und Materialien [9]. Neue Regelungen beziehen sich folglich vor allem auf eine Optimierung und ressourcensparendere Durchführung der Produktionspraktiken. In all diesen beschriebenen Fällen wird somit die Rolle der Industrieunternehmen manifestiert. Diese werden als die zentralen wertschöpfende Akteure betrachtet. Neue Wertschöpfungsprinzipien werden zwar vereinzelt genannt, aber nicht näher ausgeführt.

Folglich fokussieren sich aktuelle Definitionen und Maßnahmen für die Kreislaufwirtschaft auf das Erzeugen von Kreisläufen, wobei gleichzeitig weiterhin in der Logik industrieller Wertschöpfung gedacht wird. Regularien beziehen sich auf Produktionsunternehmen, die ihre Prozesse optimieren und in diese die 9R-Strategien integrieren sollen. Etablierte Akteure, Lieferketten und Branchen werden als Kreis gedacht,

ermöglicht durch neue Produktions- und Informationstechnologien, die eine effiziente Rückführung von Bauteilen und Produktteilen in die Kreisläufe ermöglichen ggf. ergänzt um Akteure, die die Ketten schließen (z. B. Reparaturdienstleister).

Da die Transformation zu Wertschöpfungskreisläufen ausgehend von der aktuellen globalen, industriellen linearen Wertschöpfung gedacht wird, wird aber das eigentlich Entwicklungspotenzial der Kreislaufwirtschaft häufig übersehen bzw. nicht betrachtet.

Dieses liegt nicht allein in der Bildung von Kreisläufen, sondern vielmehr in ihrer konsequenten Minimierung, die durch die Reduktion der Anzahl von Wertschöpfungsstufen und der Verringerung der räumlichen Dislozierung der Prozesse, die zwischen den Nutzungsphasen eines Produktes liegt, erzielt wird.

Effiziente Kreisläufe entstehen folglich dann, wenn man sie so eng wie möglich um den Ort der Nutzung bzw. den Nutzer\*innen selbst legt [17]. Eine solche Minimierung kann aber nur funktionieren, wenn von der bisherigen Art industrieller Wertschöpfung Abstand genommen wird, welche auf eine globale, arbeitsteilige Herstellung von Produkten ausgerichtet ist und alternative, offenere Wertschöpfungsformen betrachtet werden, die nah am Ort der eigentlichen Produktnutzung angesiedelt sind (lokale Produktion) und eine Einbindung der Verbraucher:innen über möglichst viele Phase der Produktentstehung vorsieht (Prosuming).

### III. VERKLEINERUNG VON KREISLÄUFEN

Wie gesehen, zielen Anstrengungen zu einer Kreislaufwirtschaft bislang vor allem auf die Kreisförmigkeit, nicht aber auf die Verkleinerung von Kreisläufen ab. Allerdings erscheint eine Verkleinerung der Kreisläufe aus mehrfacher Hinsicht sinnvoll.

#### 1) Reduktion der Wertschöpfungsstufen

Die Herstellung eines Produkts ist in der Regel mit zahlreichen Wertschöpfungsschritten verbunden. Jeder dieser Schritte benötigt Arbeit, Zeit, Energie und Ressourcen [6], [10]. Ziel sollte also sein, diese eingebrachte Energie und die aufgewandten Ressourcen möglichst in ihrem ursprünglichen Zustand zu erhalten. Dies wird dann erreicht, wenn für die Erfüllung von Nutzerbedürfnissen Kreisläufe mit möglichst wenig Wertschöpfungsstufen gestaltet werden, denn mit jeder Stufe wird erneut Energie oder Materialien eingesetzt. So ist das Wiederverwenden eines Produkts (Reuse) i.d.R. mit weniger Energie und Materialaufwand verbunden, als die Wiederaufbereitung (Remanufacturing / Refurbishment). Die Wiederaufbereitung eines Produktes ist effizienter als das Recycling seiner Bauteile und Neuherstellung eines Produktes.

Das Ziel ist also eine möglichst weitgehende Minimierung der Anzahl von Wertschöpfungsstufen, die ein Produkt nach einer Nutzungsphase durchläuft, ehe es wieder in die Nutzung überführt wird. Je weniger Wertschöpfungsstufen enthalten sind, desto kleiner der Kreislauf.

#### 2) Reduktion der Dislozierung der Akteure

Wie bereits festgestellt, findet Wertschöpfung heutzutage vielfach im Rahmen globaler, industrieller Prozesse statt. Die Produktion erfolgt dabei i.d.R.: an dislozierten Standorten, an denen jeweils Teilschritte durch verschiedene Akteure umge-

setzt werden. Eine nachhaltige Kreislaufwirtschaft muss darauf ausgerichtet sein, die Dislozierung der Prozesse zu reduzieren, um die Kreisläufe zu verkleinern. Die Verkleinerung von Kreisläufen würde bedeuten diese Dislozierung zu reduzieren, die Akteure also stärker räumlich zusammenzuführen. Die damit einhergehende Reduktion der Transportwege ist dabei eng mit der Reduktion von Emissionen verbunden, denn die Logistik ist aktuell der zweitgrößte Verursacher von Treibhausgasemissionen [18]. Gleichzeitig ist weniger Transportinfrastruktur nötig, so dass weniger Straßen, Brücken, Logistikzentren etc. gebaut werden, wodurch wiederum Emissionen, Energie und Ressourcen gespart werden [11].

Eine Verkleinerung der Kreisläufe könnte also in einer Reduktion der Wertschöpfungsstufen und einer Reduktion der Dislozierung der Akteure bestehen. Beides würde zu mehr Ressourceneffizienz bzw. Emissionseinsparungen führen. Im Folgenden sollen zwei Möglichkeiten dargestellt werden, die das Potenzial zur Verkleinerung von Kreisläufen beinhalten.

### IV. VERKLEINERUNG VON KREISLÄUFEN DURCH LOKALE PRODUKTION

Die Verkleinerung der Kreisläufe kann, wie gesehen, durch die Reduktion der Anzahl an Wertschöpfungsstufen in der Wertschöpfungskette als auch durch die räumliche Nähe der Akteure erzeugt werden. Beides findet im Rahmen lokaler Produktion statt [19].

Ziel der lokalen Produktion ist es, Produkte möglichst an dem Ort herzustellen, an dem sie auch benötigt werden [20] und hierbei möglichst regionale Materialien und Ressourcen zu verwenden [19]. Lokale Produktion ist dabei kein neues Phänomen. Während es die meisten Jahrhunderte hindurch üblich war am Ort des Bedarfs zu produzieren, entstand mit zunehmender Entdeckung der Welt und technischer Neuerungen ein globaler Markt. Im Zuge des vergangenen Jahrhunderts wurden schließlich mehr und mehr Produktionsstätten vom Unternehmenssitz und den Konsumorten weg versetzt [20], um günstigere Produktionsbedingungen zu nutzen, aber auch um durch Zentralisierung von Skaleneffekten zu profitieren [21]. Dies führte letztendlich jedoch dazu, dass Produkte an Orten produziert werden, an denen sie häufig gar nicht konsumiert werden, da die lokalen Absatzmärkte nicht ausreichend groß sind bzw. lukrativ.

Auch heute findet ein hoher Anteil von Wertschöpfung noch lokal statt, z. B. durch Handwerksbetriebe, aber auch kleine urbane Unternehmen [20]. Dennoch wird ein Großteil der Konsumprodukte durch globale, arbeitsteilige Wertschöpfungsprozesse bereitgestellt.

Lokale Produktion findet somit am Ort des Bedarfs statt und führt so zu einer räumlichen Nähe sowohl zwischen verschiedenen Produzenten als auch zu den Konsument\*innen [22], [23]. Durch diese Nähe können Transportwege gespart werden und damit transportbezogene Emissionen [23].

Dieser Effekt wird noch verstärkt, wenn zusätzliche lokale Ressourcen verwendet werden. Fallen auch hier Transportwege weg, können zusätzliche Emissionen gespart werden.

Auch das Teilen von Betriebsmitteln und Ressourcen ist einfacher, je mehr Akteure sich in räumlicher Nähe zueinander befinden. Entsprechend können sich Produzenten z. B. Betriebsmittel teilen, was wiederum den Bedarf an Materialien und Rohstoffen reduziert [23].

Durch die räumliche Nähe der Akteure kann auch die Umsetzung der Strategien Reuse, Repair, Sharing erleichtert werden [24]. Häufig muss ein defektes Produkt zum Hersteller zurückgeschickt werden. Dieses wird anschließend wiederum versendet, um an einem weit entfernten Standort repariert oder durch ein neues ersetzt zu werden. Die Reparaturdauer ist damit sehr hoch, ebenso der Aufwand für die Nutzer\*innen. Entsprechend selten wählen Konsument\*innen diesen Weg. Die räumliche Nähe zum Hersteller bzw. zu einem Dienstleister könnte diese Prozesse jedoch verkürzen und somit für die Nutzer\*innen wieder attraktiver werden lassen. Das Rückführen von Materialien und Bauteilen ist (z. B. im Rahmen von Refurbishment-Prozessen) einfacher und schneller möglich, wenn die Akteure sich in räumlicher Nähe zueinander finden [25].

Lokale Produktion kann folglich durch die räumliche Nähe der Akteure die Kreisläufe verkleinern und verkürzen.

Neben diesen Effekten kann lokale Produktion zu einer Förderung der kleineren, nicht global agierenden Unternehmen führen und auch die Resilienz eines Standorts stärken und weisen damit neben ökologischen auch soziale Vorteile auf.

#### V. EINBEZIEHUNG DER NUTZER\*INNEN ZUR VERKLEINERUNG VON KREISLÄUFEN

Eine weitere Möglichkeit zur Verkleinerung der Kreisläufe bietet die Mitwirkung der Nutzer\*innen an der Wertschöpfung.

Die Beteiligung der Nutzer\*innen an der Wertschöpfung, die häufig auch als Prosuming (Komposition aus Produzieren und Konsumieren) kann dabei auf vielfältige Weise erfolgen [26]. Nutzer\*innen können sowohl in Produktentwicklungsprozessen einbezogen werden (z. B. als sogenannte Lead User [27]), sie können im Produktionsprozess mitwirken (z. B. indem sie ihr Produkt aus Modulen zusammenstellen) oder ein Produkt fertigstellen (indem sie z. B. ein Möbelstück eigenständig montieren oder ihr neues Smartphone oder Tablet einrichten [28], [29]). Auch die Nutzungsverlängerung zu bewirken, etwa durch Reparieren, Umarbeiten oder Upcycling (also ganz im Sinne der 9R-Strategien), zählt zum Prosuming [30]. Prosuming kann aber auch bedeuten, dass Nutzer\*innen ein Produkt vollständig in Eigenregie herstellen (z. B. wenn sie ein Kleidungsstück selbst nähen) [30]. Auch gemeinschaftliche Produktherstellung zählt zu diesem Phänomen: Nutzer\*innen, die sich zusammenschließen, um gemeinsam neue, alternative Wertschöpfungsmodelle zu entwickeln (z. B. solidarische Landwirtschaft). Nutzer\*innen an der Wertschöpfung zu beteiligen, kann auch bedeuten, dass diese sich um die Weiterverwendung von Produkten selbst kümmern, wie etwa das Weiterverkaufen, Spenden oder das Teilen von Produkten [31].

Prosuming hebt folglich die Trennung zwischen Konsum und Produktion dabei nicht nur im Sinne des Arbeitsprozesses auf, sondern auch in räumlicher Hinsicht. Produktionsort und Konsumort liegen noch enger zusammen, bzw. können sogar gleich sein. So ist der denkbar kleinste Kreislauf folglich, wenn ein\*e Nutzer\*in, aus seinem/ihren selbsthergestellten Rohstoffen ein Produkt erzeugt (z. B. ein selbstgekochtes Essen aus Zutaten, die auf dem eigenen Balkon oder Garten gewachsen und geremtet wurden). In diesem Fall fallen (so gut wie alle) Transportwege weg (sieht man von dem Kauf des Equipments: Blumensamen, Blumentöpfe, Erde etc. ab). Ein ebenfalls kleiner Kreislauf ergibt sich, wenn ein\*e Nutzer\*in

z. B. ein Möbelstück selbst herstellt, bestenfalls sogar aus Holzresten.

Doch auch bereits nur eine teilweise Beteiligung der Nutzer\*innen an der Wertschöpfung kann bereits zu einer Verkleinerung der Kreisläufe beitragen. Wenn z. B. eine Nutzerin in Lage ist ein Produkt selbständig zu reparieren oder umzuarbeiten (z. B. unterstützt durch intelligente Selbstdiagnoseeinrichtungen im Produkt, Reparaturkits usw.), anstatt es wegzuworfen oder zum Hersteller zurückzuschicken. Durch das Selberrmachen werden Transportwege reduziert, ebenso Kommunikationsaufwand und damit Transaktionskosten.

Nun ist es weder möglich noch wünschenswert die Herstellung von Produkten wieder grundsätzlich an die Nutzer\*innen zu übertragen, da zahlreiche der modernen Produkte zu komplex sind, um außerhalb von industriellen Anlagen und ohne Fachkräfte hergestellt zu werden. Ist ein Produkt jedoch in einem solchen Rahmen fertiggestellt, können Nutzer\*innen die Lebensdauer durch die R-Strategien (insbesondere reuse, repair, refurbish und repurpose) beeinflussen. Hierfür müssen durch die Strategie des Re-Designs jedoch entsprechende Möglichkeiten in der Produktentwicklung und Produktion berücksichtigt werden, so dass der\*die Nutzer\*in das Produkt auch entsprechend selbstständig aufarbeiten oder weiterverwenden kann (z. B. durch ein reparaturfähiges Design, selbstdiagnose-Einrichtungen, Reparatur-Kits, Ersatzteilverfügbarkeit, Modularisierung etc.) [32].

Zudem hat auch die Beteiligung der Nutzer\*innen und das Verkleinern der Kreisläufe einen sozialen Effekt, der über die ökologische Wirkung einer Kreislaufwirtschaft hinausgeht.

So wird davon ausgegangen, dass die Mitwirkung an Wertschöpfungsprozessen zu einer Ermächtigung der Nutzer\*innen führt. Sie können selbständig Ideen entwickeln und umsetzen und werden so unabhängiger von Märkten [33], [34], [35]. Besonders das gemeinschaftliche Arbeiten an einem oder mehreren Produkten, wie etwa im Falle der solidarischen Landwirtschaft soll zudem positive Effekte auf die Solidarität innerhalb der Gemeinschaft haben [36].

#### VI. FAZIT

Wie gesehen existieren bereits einige Initiativen, um Konsum- und Produktionsauswirkungen zu reduzieren. Die Maßnahmen und Konzepte setzen dabei vor allem darauf zirkuläre Wertschöpfungsprozesse zu erreichen, durch die die Lebensdauer von Produkten und Produktkomponenten verlängert wird. Diese werden dabei jedoch i. d. R. noch von einer globalen, arbeitsteiligen, industriellen Wertschöpfung aus gedacht. Dies hat zur Folge, dass Kreisläufe zwar zirkularisiert, nicht aber verkleinert werden.

Um eine konsequente Verkleinerung der Wertschöpfungskreisläufe zu erzielen, muss deren Gestaltung konsequent von dem\*r späteren Nutzer\*in aus erfolgen. Wertschöpfung muss so nah wie möglich an ihn herangeführt werden: Hierzu eignen sich Ansätze der lokalen Produktion und des Prosumings.

Um Kreisläufe konsequent zu verkleinern eignen sich z. B. lokale Produktionsstrukturen und Prosuming.

Diese Ansätze kommen in den Konzepten für eine Kreislaufwirtschaft auch bereits vor. So nennt die EU in den Maßnahmen zur Kreislaufwirtschaft für die Textilproduktion die lokale Produktion. Allerdings wird dieser Punkt als eine unter zahlreichen Maßnahmen genannt und nicht näher ausgeführt. Im Entwurf zur nationalen Kreislaufstrategie für Deutschland

kommt lokale Produktion gar nicht vor. So wird lediglich beschrieben, dass „Rohstoffflüsse [...] teilweise lokalisiert bzw. regionalisiert“ werden sollen, wobei dies durch das Wiederaufbereiten vor Ort geschehen soll. Die Förderung lokaler (Erst-)Produktionsstätten wird hingegen nicht erwähnt. Auch wird in der Analyse von Petit-Boix/Leipold lokale Produktion lediglich für den Lebensmittelbereich in Erwägung gezogen, nicht aber für andere Branchen [14].

Die Beteiligung der Kund\*innen als Maßnahme für die Kreislaufwirtschaft wird auch von der EU, z. B. in den Handlungsempfehlungen für die Textilbranche vorgesehen (unter dem Stichwort: „Pro-sumerism“<sup>1</sup> zu finden) [13]. In ihrem Aktionsplan für die Kreislaufwirtschaft schreibt sie weiterhin: „Neue Geschäftsmodelle, die auf einer engeren Beziehung zum Kunden, kundenspezifischer Massenproduktion, der Wirtschaft des Teilens und der kollaborativen Wirtschaft basieren [...] werden nicht nur den Übergang zur Kreislaufwirtschaft, sondern auch die Entmaterialisierung der Wirtschaft beschleunigen und die Abhängigkeit Europas von Primärrohstoffen verringern.“ [37]. Auch in Petit-Boix/Leipolds Recherche kommt Prosuming in Form von „sharing initiatives“ vor [14]. Die Bedeutung, die Nutzer\*innen in Form von Prosuming jedoch für die Wertschöpfung, insbesondere für die Verkleinerung von Kreisläufen haben, wird darin jedoch nicht gekennzeichnet. Industrielle Unternehmen werden weiterhin als die zentralen wertschöpfenden Akteure, auch im Rahmen einer Kreislaufwirtschaft, betrachtet.

Lokale Produktion und vor allem Prosuming werden folglich für die Etablierung von Kreislaufwirtschaften bereits berücksichtigt. Allerdings scheinen sie eher eine untergeordnete Rolle zu spielen, zumindest aber keine bedeutungsvolle. Ebenso wird der Zusammenhang zu einer Verkleinerung der Kreisläufe bislang nicht deutlich. Dies ist nicht verwunderlich, wenn man bedenkt, dass Kreislaufwirtschaft aktuell noch in der Logik industrieller Wertschöpfung gedacht wird.

Lokale Produktion und die Beteiligung der Nutzer\*innen an Wertschöpfungsprozessen sollten aufgrund ihres Potenzials Kreisläufe nicht nur herzustellen, sondern auch zu verkleinern stärkere Beachtung in den Konzepten und Maßnahmen für eine Kreislaufwirtschaft finden, da die räumliche Nähe der Akteure und die darauf resultierende einfacherer Zusammenarbeit, nicht nur Ressourcen zu sparen, sondern auch Emissionen zu reduzieren und neue, wertvolle Formen von Wertschöpfungsprozessen erzeugen kann. Hierzu müsste Kreislaufwirtschaft jedoch von Anfang an stärker vom Ort der Nutzung und der Nutzer\*innen heraus gedacht werden.

## VII. AUSBLICK

Wie gezeigt, beinhalten lokale Produktion und Prosuming dabei ein vielversprechendes Potenzial um Wertschöpfungskreisläufe zu verkleinern und dabei auch noch soziale Nachhaltigkeitsziele (Stärkung regionaler Akteure, Demokratie und Solidarität) zu fördern.

Weitere Forschungsanstrengungen müssen nun klären, wie sich lokale Produktion und Prosuming sinnvoll im Rahmen einer Kreislaufwirtschaft einbetten lassen. Dabei muss

insbesondere berücksichtigt werden, welcher Grad der Einbeziehung der Nutzer:innen sinnvoll erscheint und für welche Produktgruppen eine lokale Produktion geeignet ist. Zudem darf nicht übersehen werden, dass die Einbeziehung der Nutzer:innen immer auch eine Leistung ihrerseits darstellt, die nicht als selbstverständlich hingenommen werden darf. Andernfalls verlagert sich lediglich die Ausbeutung, die aktuell durch Outsourcing stattfindet.

## DANKSAGUNG

This research is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr [project: Production Next Door]. dtec.bw is funded by the European Union – NextGenerationEU.

## LITERATUR

- [1] J. Rockström, W. Steffen and K. Noone, "A safe operating space for humanity," *Nature*, vol. 461, pp. 472–475, 2013, doi: 10.1038/461472a.
- [2] Europäisches Parlament, "Umweltauswirkungen von Textilproduktion und -abfällen. Letzter Zugriff: 24.10.202. Abrufbar unter: <https://www.europarl.europa.eu/topics/de/article/20201208STO93327/umweltauswirkungen-von-textilproduktion-und-abfaellen-infografik>.
- [3] Europäisches Parlament, "Änderung unserer Produktions- und Verbrauchsmuster: neuer Aktionsplan für Kreislaufwirtschaft ebnet Weg zu klimaneutraler und wettbewerbsfähiger Wirtschaft mit mündigen Verbrauchern", Brüssel, 11.03.2020, 2020.
- [4] R. Loske, "Neue Formen kooperativen Wirtschaftens als Beitrag zur nachhaltigen Entwicklung," *vz-nrw*, 2009.
- [5] M., Golde, *Rebound-Effekte: Empirische Ergebnisse und Handlungsstrategien*. Hintergrund, 2016.
- [6] N. Paech, "Nach dem Wachstumsrausch: eine zeitökonomische Theorie der Suffizienz," *Zeitschrift für Sozialökonomie*, vol. 47, 2016.
- [7] J. Kirchherr, D. Reike D und M. Hekkert, "Conceptualizing the circular economy: An analysis of 114 definitions," *Resources, Conservation and Recycling*, 2016, pp. 127:221–232, 2016, doi:10.1016/j.resconrec.2017.09.005.
- [8] F. Moser, "Kreislaufwirtschaft und nachhaltige Entwicklung," in, *Handbuch des Umweltschutzes und der Umweltschutztechnik, Band 2: Produktions- und produktintegrierter Umweltschutz*, H. Brauer (Hrg.), Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1059–1153, 1996.
- [9] Europäisches Parlament, "Kreislaufwirtschaft: Definition und Vorteile". Letzter Zugriff: 24.10.2024. Abrufbar unter: <https://www.europarl.europa.eu/topics/de/article/20151201STO05603/kreislaufwirtschaft-definition-und-vorteile>.
- [10] F. Müller, R. Kohlmeyer und F. Krüger et al., "Leitsätze einer Kreislaufwirtschaft", Umweltbundesamt, Dessau-Roßlau,
- [11] M. Larsson, *Circular Business Models*, Springer International Publishing, Cham, 2018.
- [12] S. Kadner, J. Kobus, E.G. Hansen et al., "Circular Economy Roadmap für Deutschland", acatech - Deutsche Akademie der Technikwissenschaften (Hg.), 2021.

<sup>1</sup> Die Bezeichnung „Prosuming“ geht, auf Alvin Toffler zurück (vgl. Toffler 1980). Dieser versuchte damit zu beschreiben, wie sich das Verhältnis zwischen Produzenten und Konsumenten über die Laufe der Zeit verändert hatte. Im Prosuming steckt dabei die Aufhebung der Trennung zwischen Produktion und Konsum, die im Zeitalter der Industrialisierung vorherrschte. Dadurch werde der Nutzer/die Nutzerin eines Produkts, gleichzeitig auch zu dessen Produzent/in. Von Bedeutung für Toffler ist dabei, dass Nutzer:innen die

Produkte für ihren eigenen Bedarf produzieren. Auf diese Weise, so Toffler, würden Produkte wieder weniger zu einer Ware mit einem Tauschwert, sondern vielmehr wieder für ihren direkten Gebrauch hergestellt.

- [13] European Environment Agency, "Textiles in Europe's circular economy", Briefing no. 10/2019, 2019, 2020. Letzter Zugriff: 24.10.2024. Abrufbar unter: [https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://circulareconomy.europa.eu/platform/sites/default/files/briefing-textiles-in-europe-s-circular-economy\\_1.pdf&ved=2ahUKewjIhoyn7aaJAXuBN-sEHbG\\_LJcQFnoECBUQAQ&usg=AOvVaw3L3JIIFC-QIWbqp4ruR5F](https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://circulareconomy.europa.eu/platform/sites/default/files/briefing-textiles-in-europe-s-circular-economy_1.pdf&ved=2ahUKewjIhoyn7aaJAXuBN-sEHbG_LJcQFnoECBUQAQ&usg=AOvVaw3L3JIIFC-QIWbqp4ruR5F).
- [14] A. Petit-Boix A und S. Leipold, "Circular economy in cities: Reviewing how environmental research aligns with local practices," *Journal of Cleaner Production* vol. 195, pp. 1270–1281, 2018 doi: 10.1016/j.jclepro.2018.05.281.
- [15] V. Schulze, J.C. Aurich und P. Groche et al., "Update Factory für ein industrielles Produkt-Update: Ein Beitrag zur Kreislaufwirtschaft," Positionspapier Dezember 2021, WGP, Aachen, 2021.
- [16] Europäisches Parlament, "Recht auf Reparatur: Welche Maßnahmen ergreift die EU, um Reparaturen attraktiver zu machen?," Letzter Zugriff: 24.10.2024. Abrufbar unter: <https://www.europarl.europa.eu/topics/de/article/20220331STO26410/recht-auf-reparatur-massnahmen-reparaturen-attraktiver-zu-machen>.
- [17] K. Webster, J. Blériot und C. Johnson, "Effective business in a circular economy", *A new dynamic*, vol 1. Ellen MacArthur Foundation (Hrg.), Cowes, Isle of Wight, 2013.
- [18] Statista, "Verteilung der CO2-Emissionen weltweit nach Sektor, ". Letzter Zugriff: 24.10.2024. Abrufbar unter: <https://de.statista.com/statistik/daten/studie/167957/umfrage/verteilung-der-co-emissionen-weltweit-nach-bereich/>.
- [19] P. Krenz, L. Stoltenberg, J. Markert et al., "The Phenomenon of Local Manufacturing: An Attempt at a Differentiation of Distributed, Re-distributed and Urban Manufacturing," in: *Towards Sustainable Customization: Proceedings of the 8th Changeable, Agile, Reconfigurable and Virtual Production Conference (CARV2021) and the 10th World Mass Customization and Personalization Conference (MCPC2021)*, A.L. Andersen (Hrg.), Aalborg, Denmark, Springer International Publishing AG, Cham, pp. 1014–1022, 2022.
- [20] D. Läßle, "Produktion zurück in die Stadt?," in *Polarisierte Städte: Soziale Ungleichheit als Herausforderung für die Stadtpolitik*, M. Kronauer, W. Siebel (Hrg.), Campus, Frankfurt, 2013.
- [21] O. Büchmann und R. Kniese, „Auslaufmodell Outsourcing?," *Wirtsch Inform Manag.* vol. 6, pp. 7–21, 2014, doi: 10.1365/s35764-014-0418-4.
- [22] S. Sassen, "Cities Today: A New Frontier for Major Developments," *The ANNALS of the American Academy of Political and Social Science*, vol. 626, pp. 53–71, 2009, doi: 10.1177/0002716209343561.
- [23] A.C. Erbstößer, "Produktion in der Stadt," in *Berliner Mischung 2.0.*, Technologiestiftung Berlin (Hrg.), 2006.
- [24] S. Prendeville, G. Hartung, E. Purvis et al., "Makespaces: From Redistributed Manufacturing to a Circular Economy," in *Sustainable design and manufacturing 2016*, R. Setchi, R.J. Howlett, Y. Liu Y et al. (Hrg.), vol 52. Springer, Cham, pp. 577–588, 2016.
- [25] C. Kohtala, "Addressing sustainability in research on distributed production: an integrated literature review," *Journal of Cleaner Production*, vol. 106, pp. 654–668. 2015, doi 10.1016/j.jclepro.2014.09.039.
- [26] B. Blätzel-Mink, K.U. Hellmann K-U, "Prosumer revisited: Zur Aktualität einer Debatte," VS Verl. für Sozialwiss, Wiesbaden, 2010.
- [27] E. von Hippel, "Lead Users: A Source of Novel Product Concepts," in *Management Science*, vol. 32, pp- 791–805, 1986, doi: 10.1287/mnsc.32.7.791.
- [28] H. Klemisch und M. Boddenberg, "Genossenschaftliche Prosumermodelle", vz-nrw (Verbraucherzentrale Nordrhein-Westfalen e. V.), 2016.
- [29] H. Hanekop, V. Wittke, "Kollaboration der Proumenten," in *Prosumer revisited: Zur Aktualität einer Debatte*, B. Blätzel-Mink B und K.W. Hellmann (Hrg.) 1. Aufl. VS Verl. für Sozialwiss, Wiesbaden, pp 96–113, 2010.
- [30] B. Blätzel-Mink, "Varieties of Prosuming – konzeptionelle Überlegungen und empirische Befunde zur veränderten Rolle von Konsument\_innen," in *Entgrenzungen des Konsums: Dokumentation der Jahreskonferenz des Netzwerks Verbrauchersforschung*, P. Kenning, J. Lamla (Hrg.), Springer Fachmedien Wiesbaden, Wiesbaden, pp 17–31, 2017.
- [31] B. Blätzel-Mink, "Prosuming im online-gestützten Gebrauchsgüterhandel und Nachhaltigkeit", in *Prosumer revisited: Zur Aktualität einer Debatte*, B. Blätzel-Mink K.U. Hellmann (Hrg.), 1. Aufl. VS Verl. für Sozialwiss, Wiesbaden, pp 117–130, 2010.
- [32] M. Moreno, O. Ponte, F. Charnley, "Taxonomy of design strategies for a circular design tool," *Product Lifetimes And The Environment*, Conference Proceedings of PLATE 2017, 8-10 November 2017, Delft, the Netherlands, vol. 9. IOS Press, Amsterdam, Netherlands, 2017.
- [33] B. Corona B, L. Shen L, D. Reike et al., "Towards sustainable development through the circular economy—A review and critical assessment on current circularity", in *Metrics*, vol. 151, 2019.
- [34] N. Scharf, T. Wachtel, S.E. Reddy et al., "Urban Commons for the Edible City—First Insights for Future Sustainable Urban Food Systems from Berlin, Germany," in *Sustainability*, vol. 11, 2019, doi: doi.org/10.3390/su11040966.
- [35] E. Unterfrauner, M. Hofer, B. Pelka et al., "A New Player for Tackling Inequalities? Framing the Social Value and Impact of the Maker Movement," *Social Inclusion*, vol. 8, pp. 190–200, 2020, doi: 10.17645/si.v8i2.2590.
- [36] T. Scholz, "Plattform Kooperativismus: Wie wir uns die Sharing Economy zurückholen können". Letzter Zugriff: 24.10.2024. Abrufbar unter: <https://www.rosalux.de/publikation/id/8813/plattform-kooperativismus>.
- [37] Europäisches Parlament, "Wie will die EU bis 2050 eine Kreislaufwirtschaft erreichen?,". Letzter Zugriff: 24.10. 2024. Abrufbar unter: <https://www.europarl.europa.eu/topics/de/article/20210128STO96607/wie-will-die-eu-bis-2050-eine-kreislaufwirtschaft-erreichen#aktionsplan-der-eu-fr-die-kreislaufwirtschaft-10>.

# Leistungselektronik in kritischer Infrastruktur am Beispiel der medizinischen Bildgebung – Untersuchung der Elektromagnetischen Verträglichkeit und Optimierung der Resilienz

Simon Podendorf\*, Manikandan Sankar,  
Ahmet Buran, Stefan Dickmann  
*Professur Grundlagen der Elektrotechnik  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland*  
\*simon.podendorf@hsu-hh.de

Eden Tulu, Marcus Stierner  
*Professur Theoretische Elektrotechnik  
und Numerische Feldberechnung  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland*

Sönke Brandt, Martin Rasch,  
Klaus F. Hoffmann  
*Professur für Leistungselektronik  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland*

**Kurzfassung** — Hohe Anforderungen an die elektronischen Komponenten von Systemen der medizinischen Bildgebung können im Rahmen derer Energieversorgung nur durch den Einsatz von moderner Leistungselektronik erfüllt werden. Mit der Verwendung von Wide-Bandgap-Leistungshalbleitern, z. B. in Stromrichtern von Röntgensystemen, gehen damit schnellere Schaltvorgänge sowie hohe Schaltfrequenzen einher. Dies macht eine Untersuchung der Auswirkungen auf die Elektromagnetische Verträglichkeit (EMV) erforderlich. Das betrifft sowohl die Einhaltung von normativen EMV-Grenzwerten als auch die Wechselwirkung der parasitären Eigenschaften von Komponenten innerhalb eines Systems. Im Forschungsprojekt DiMoLEK werden messtechnische Untersuchungen durchgeführt und es wird ein Ansatz zur Modellierung der parasitären Eigenschaften von passiven Komponenten entwickelt. Die Simulation von Stromrichtern mit anschließender Analyse der Wechselwirkungen parasitärer Eigenschaften ermöglicht die Optimierung der (elektromagnetischen) Resilienz. Dies kann beispielsweise durch eine Anpassung des Designs der Komponenten und der Modulationsverfahren erreicht werden.

**Stichworte** — Elektromagnetische Verträglichkeit, Leistungselektronik, kritische Infrastruktur, Wide-Bandgap-Halbleiter, Resonanzwandler

## NOMENKLATUR

CM	Gleichtakt
DM	Gegentakt
DUT	Device-under-Test
EMV	Elektromagnetische Verträglichkeit
ESB	Ersatzschaltbild
GaN	Galliumnitrid
IGBT	Insulated-Gate-Bipolar-Transistor
MOSFET	Metalloxid-Halbleiter-Feldeffekttransistor
PFM	Pulsfrequenzmodulation
SiC	Siliziumkarbid
ZVS	Nullspannungsschalten

## I. EINLEITUNG

Eine zuverlässige Gesundheitsversorgung ist für die Gesellschaft unverzichtbar und kann daher zu der kritischen Infrastruktur gezählt werden. Moderne Geräte der medizinischen Bildgebung sind ein wichtiger Bestandteil der Diagnostik und sind daher von großer Bedeutung. Die hohen Anforderungen an die elektronischen Komponenten der Energieversorgung von komplexen Geräten, wie z. B. Röntgensystemen, können nur durch den Einsatz von moderner Leistungselektronik erfüllt werden. Die Verwendung von Wide-Bandgap-Leistungshalbleitern auf der Basis von Siliziumkarbid (SiC) und Galliumnitrid (GaN) in Stromrichtern ermöglicht schnellere Schaltvorgänge und höhere Schaltfrequenzen im Vergleich zu Standard-Leistungsbaulementen wie dem Insulated-Gate-Bipolar-Transistor (IGBT). Dies erhöht tendenziell die Störemissionen von Stromrichtern und erfordert eine Untersuchung der Auswirkungen auf die Elektromagnetische Verträglichkeit (EMV) [1]. Dabei kann zwischen der Einhaltung von normativen EMV-Grenzwerten und der gegenseitigen Beeinflussung der parasitären Eigenschaften von Komponenten innerhalb eines Systems unterschieden werden. Im Forschungsprojekt „Digitales Lebenszyklus-Monitoring, Härtung und Optimierung der Resilienz von Leistungselektronik in kritischer Infrastruktur (DiMoLEK)“ wird ein Ansatz zur Modellierung der parasitären Eigenschaften passiver leistungselektronischer Komponenten entwickelt. Durch das Einbinden dieser realitätsnahen Modelle in Simulationen wird unter anderem eine Optimierung der Pulsfrequenzmodulation eines Resonanzwandlers ermöglicht. Zur Überprüfung der Störemissionen wird eine Testumgebung, angelehnt an die EMV-Normen für medizinische Geräte, in Betrieb genommen und eingesetzt.

## II. RESONANZWANDLER

In der Energieversorgung von Röntgensystemen kommen unter anderem Resonanzwandler zur Regelung unterschied-

licher Ströme und Spannungen zum Einsatz. Diese bringen hochperformante Eigenschaften mit. Die in dieser Forschungsarbeit untersuchte Schaltung entspricht einem LLC-Resonanzwandler, siehe Abbildung 1. Dabei steht ein „L“ für die Serieninduktivität  $L_r$ , ein „L“ für die Parallelinduktivität  $L_m$  und „C“ für die Kapazität aus  $C_1$  und  $C_2$ . Im Gegensatz zu einem konventionellen Resonanzwandler wird auf den Gleichrichter mit kapazitiver Glättung am Ausgang der Schaltung verzichtet.

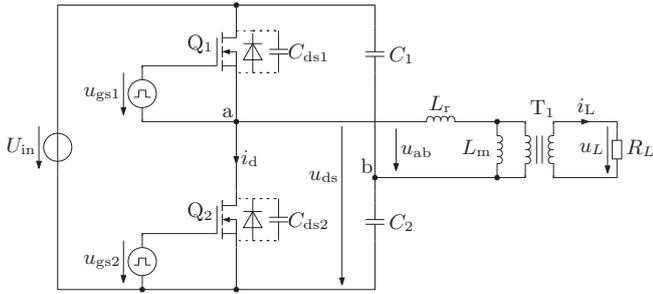


Abbildung 1: Schaltung des Halbbrücken-LLC-Resonanzwandlers

Der Wechselrichter, bestehend aus den SiC-MOSFETs  $Q_1$  und  $Q_2$ , wird eingangsseitig durch die Gleichspannung  $U_{in}$  gespeist. Der Resonanzwandler ermöglicht die Steuerung der Ausgangsspannung  $u_L$  durch Variation der Schaltfrequenz  $f_{sw}$ . Dies wird als Pulsfrequenzmodulation (PFM) bezeichnet. Das Intervall der Schaltfrequenz wird unter anderem durch die untere resultierende Resonanzfrequenz (Leerlauf) festgelegt.

$$f_{rm} = \frac{1}{2\pi\sqrt{L_{rm}C_r}} \quad (1)$$

des LLC-Resonanzkreises festgelegt. Die Resonanzkapazität  $C_r$  ist aufgeteilt auf die Parallelschaltung aus  $C_1$  und  $C_2$ , so dass

$$C_r = C_1 + C_2. \quad (2)$$

Die gesamte Induktivität  $L_{rm}$  wird aus

$$L_{rm} = L_r + L_m \quad (3)$$

berechnet. Für  $C_1$ ,  $C_2$  und  $L_r$  werden diskrete Bauelemente auf der Leiterplatte eingesetzt. Die Magnetisierungsinduktivität des Transformators  $T_1$  bildet die Induktivität  $L_m$ . Der Lastwiderstand  $R_L$  wird direkt an die Sekundärseite des Transformators angeschlossen. Die MOSFETs  $Q_1$  und  $Q_2$  werden durch die Treiberspannungen  $u_{gs1}$  und  $u_{gs2}$  alternierend ein- und ausgeschaltet. Die Ausgangsspannung des Wechselrichters  $u_{ab}$  entspricht dann einer rechteckförmigen Spannung mit der Amplitude  $\pm U_{in}/2$ . Ein Vorteil der Resonanzwandler ist die Verwendung von Nullspannungsschalten (zero-voltage switching, ZVS) zur Reduzierung der Schaltverluste. Dies erfordert eine Ansteuerung der MOSFETs mit einer Schaltfrequenz oberhalb der resultierenden Resonanzfrequenz.

### A. Messung im Zeitbereich

Das Messergebnis der Drain-Source-Spannung  $u_{ds}$  in Abbildung 2 zeigt hochfrequente Schwingungen nach den Schaltvorgängen. Es wird eine Eingangsspannung  $U_{in} = 300$  V angelegt und eine Schaltfrequenz  $f_{sw} = 50$  kHz verwendet. Der Tastgrad wird auf nahezu 50 % eingestellt. Die Schwingungen resultieren aus der Wechselwirkung der parasitären (nicht idealen) Eigenschaften der Bauelemente und der Leiterplatte. Mögliche Folgen der hochfrequenten Schwingungen sind die Beschädigung der MOSFETs durch Überspannungen und die Erzeugung von elektromagnetischen Wellen.

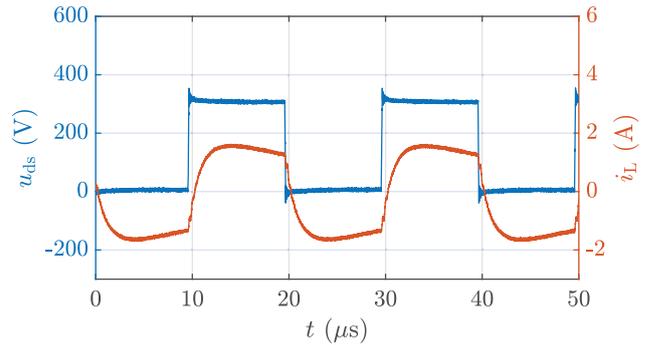


Abbildung 2: Messung am LLC-Resonanzwandler mit einem Tastgrad nahezu 50 %, entsprechend einer Totzeit von 750 ns

## III. VARIATION DES TASTGRADES

Als Totzeit wird die Periode der Treiberspannungen  $u_{gs1}$  und  $u_{gs2}$  bezeichnet, in der beide MOSFETs gleichzeitig ausgeschaltet sind. Damit stehen der Tastgrad und die Totzeit in einem direkten Zusammenhang. Die Totzeit verhindert das Kurzschließen der Halbbrücke, das auftritt, wenn beide MOSFETs gleichzeitig eingeschaltet sind [2]. Dies führt zur Zerstörung der MOSFETs und muss daher verhindert werden. Die Totzeit wird entsprechend der Schaltcharakteristik der MOSFETs gewählt und beeinflusst außerdem das ZVS. Die minimale Totzeit ist die Zeit, die der Aufladevorgang der effektiven Kapazität parallel zu Drain und Source bei minimalem Magnetisierungsstrom benötigt [3]. In Abbildung 3 sind die Messergebnisse der Drain-Source-Spannung  $u_{ds}$  und des Laststroms  $i_L$  bei einer doppelten Totzeit von 1500 ns im Vergleich zu Abbildung 2 gezeigt. Die Messergebnisse zeigen einen deutlichen Einfluss der Totzeit auf die Schwingungen nach den Schaltvorgängen von  $u_{ds}$ . Die stark ausgeprägten Schwingungen bei einer größeren Totzeit können die Störemissionen enorm erhöhen [4].

## IV. MODELLIERUNG DER PARASITÄREN EIGENSCHAFTEN

Sollen die nach den Schaltvorgängen auftretenden hochfrequenten Schwingungen simuliert werden, müssen die parasitären Eigenschaften des Resonanzwandlers berücksichtigt werden. In diesem Abschnitt wird die Methode nach [5] zur

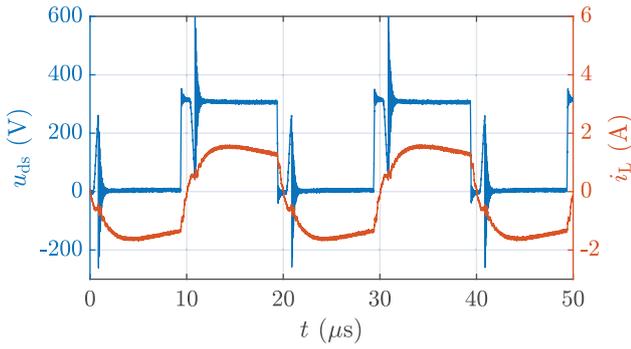


Abbildung 3: Messung am LLC-Resonanzwandler mit einer Totzeit von 1500 ns

Modellierung der parasitären Eigenschaften passiver Bauelemente am Beispiel der Spule  $L_r$  vorgestellt. Im Projekt „DiMOLEK“ wird diese Methode zur Erstellung detaillierter Modelle erweitert [6]. Die Methode basiert auf der Nachbildung von gemessenen Impedanzen der Bauelemente mit elektrischen Ersatzschaltbildern (ESB). Die Näherung der Impedanzen erfolgt mit rationalen Funktionen in Partialbruchdarstellung,

$$F(s) = g + se + \sum_{k=1}^N \frac{r_k}{(s - p_k)} \quad (4)$$

die mit Hilfe des Vector Fitting Algorithmus berechnet werden [7], [8], [9]. Dabei ist  $s = \sigma + j\omega$  die komplexe Variable,  $r_k$  und  $p_k$  sind die  $k$ -ten Residuen bzw. Polstellen. Für ein komplex konjugiertes Paar von Polstellen und Residuen lautet die genäherte Funktion

$$F(s) = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} = \frac{as + b}{s^2 + sc + d} \quad (5)$$

mit

$$\begin{aligned} a &= r_1 + r_2 \\ b &= -(r_1 p_2 + r_2 p_1) \\ c &= -(p_1 + p_2) \\ d &= p_1 p_2. \end{aligned} \quad (6)$$

Diese rationale Funktion kann durch das ESB in Abbildung 4 modelliert werden.

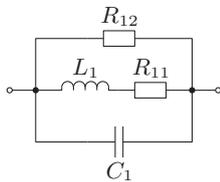


Abbildung 4: Ersatzschaltbild einer realen Spule

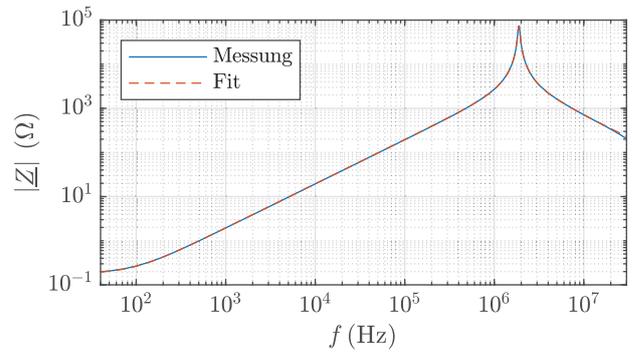
Die Impedanz des ESB ist:

$$Z_1(s) = \frac{1}{C_1} \frac{(s + \frac{R_{11}}{L_1})}{s^2 + (\frac{1}{R_{12}C_1} + \frac{R_{11}}{L_1})s + (\frac{R_{11}}{R_{12}L_1C_1} + \frac{1}{L_1C_1})}. \quad (7)$$

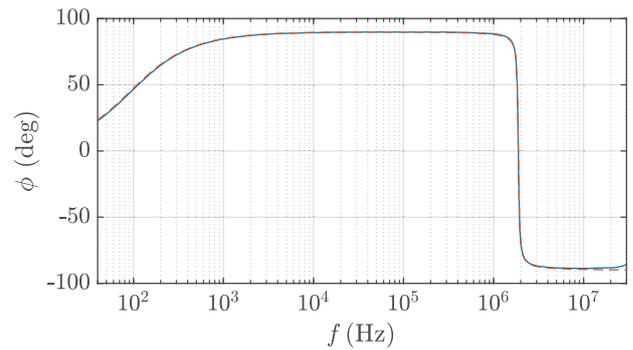
Die Parameter der Bauelemente im ESB können durch einen Koeffizientenvergleich mit Gleichung (5) und anschließendem Lösen des Gleichungssystems berechnet werden. Daraus folgen die Bestimmungsgleichungen:

$$\begin{aligned} C_1 &= 1/a \\ R_{12} &= \frac{c - b/a}{a} \\ L_1 &= \frac{a^3}{a^2d + b^2 - abc} \\ R_{11} &= \frac{bL_1}{a}. \end{aligned} \quad (8)$$

Die gemessene Impedanz der Spule  $L_r$  in Abbildung 5 zeigt eine Parallelresonanz und kann daher durch den Parallelschwingkreis in Abbildung 4 modelliert werden. Die Induktivität  $L_1$  repräsentiert die gewollte Induktivität von  $L_r$ . Die Widerstände  $R_{11}$  und  $R_{12}$  berücksichtigen Wicklungs- bzw. Kernverluste.  $C_1$  modelliert eine parasitäre Kapazität der Wicklung. Der Fit und die gemessene Impedanz stimmen bis zu einer Frequenz von  $f = 30$  MHz überein.



(a) Betrag



(b) Phasenwinkel

Abbildung 5: Impedanzmessung und Fit der Spule  $L_r$

Für die Kondensatoren  $C_1$  und  $C_2$  ist die Vorgehensweise analog, mit dem Unterschied, dass diese durch einen Reihenschwingkreis modelliert werden. In dieser Forschungsarbeit ist der Lastwiderstand  $R_L$  konstant. Zur Vereinfachung können der Transformator und der Lastwiderstand in einem ESB

zusammengefasst werden. Eine Funktion nach Gleichung (4) nähert dann die primärseitig gemessene Impedanz des Transformators  $T_1$  mit sekundärseitig angeschlossenem Lastwiderstand.

### V. SIMULATION IM ZEITBEREICH

In diesem Abschnitt wird die Simulation des LLC-Resonanzwandlers im Zeitbereich durchgeführt. Aus einer realitätsnahen Simulation kann die EMV des Resonanzwandlers durch Auswahl der idealen Totzeit optimiert werden. Eine Optimierung der PFM, basierend auf einer Simulation, ist dabei weniger aufwändig als Messungen an der realen Schaltung im Labor.

#### A. Simulation der idealen Schaltung

In Abbildung 6 ist eine Simulation des LLC-Resonanzwandlers im Zeitbereich mit den idealen passiven Bauelementen  $C_1$ ,  $C_2$ ,  $L_r$ ,  $L_m$  und  $T_1$  mit einer Totzeit von 750 ns dargestellt. Für die MOSFETs wird das vom Hersteller bereitgestellte SPICE-Modell genutzt. Die simulierte Schaltung entspricht Abbildung 1. Die Schwingungen nach den Schaltvorgängen können mit den idealen Bauelementen nicht simuliert werden.

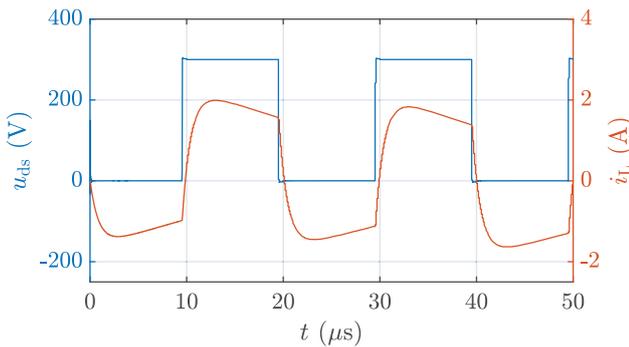


Abbildung 6: Simulation des LLC-Resonanzwandlers mit idealen Bauelementen mit einer Totzeit von 750 ns

#### B. Simulation der nicht idealen Schaltung

Die Simulation der Schaltung mit nicht idealen Bauelementen in Abbildung 7 und einer Totzeit von 1500 ns zeigt Schwingungen nach den Schaltvorgängen. Die Schwingungen und die Überspannung von  $u_{ds}$  sind erheblich geringer im Vergleich zur Messung. Die Simulation ist trotzdem geeignet zur Auswahl der idealen Totzeit, da dessen Einfluss in der Simulation berücksichtigt wird.

### VI. INBETRIEBNAHME EINER EMV-TESTUMGEBUNG

Es wird eine EMV-Testumgebung für leitungsgebundene Störemissionen in Betrieb genommen mit der präzise EMV-Messungen durchgeführt werden können. Der Messaufbau ist angelehnt an die Norm CISPR 11 [10], um reproduzierbare Messergebnisse zu ermöglichen. Die Messungen werden abweichend von der Norm im Frequenzbereich von 150 kHz bis

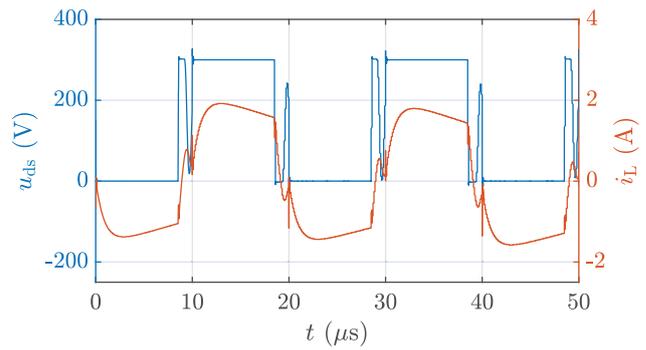


Abbildung 7: Simulation des LLC-Resonanzwandlers mit einer Totzeit von 1500 ns unter Berücksichtigung der parasitären Eigenschaften der Bauelemente

100 MHz durchgeführt. In Abbildung 8 ist der Messaufbau skizziert. Der LLC-Resonanzwandler dient als Messobjekt (Device-under-Test, DUT) und wird mit einer Netznachbildung auf einer Referenzebene platziert. Die Netznachbildungen werden zwischen dem Netzteil und das DUT angeschlossen. Das Netzteil, der Signalgenerator und der Spektrumanalysator befinden sich außerhalb der Absorberkammer. Diese stellt eine Testumgebung dar, die elektromagnetisch von der Umgebung abgeschirmt ist. Die Treibersignale der MOSFETs werden durch den Signalgenerator über geschirmte Leitungen bereitgestellt. Der Spektrumanalysator wird über einen CM/DM-Schalter an die Netznachbildung angeschlossen, damit zwischen Gleichtaktstörungen (common mode, CM) und Gegentaktstörungen (differential mode, DM) separiert werden kann.

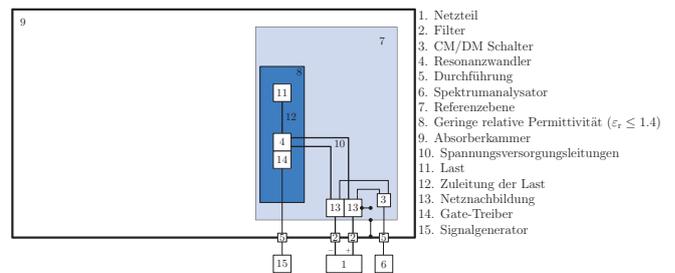


Abbildung 8: EMV-Messaufbau für leitungsgebundene Störemissionen

In Abbildung 9 sind die Messergebnisse der leitungsgebundenen Störemissionen des LLC-Resonanzwandlers als DUT gezeigt. Die Störemissionen sind in CM und DM separiert. Der Einfluss der Totzeit auf die Störemissionen ist deutlich, insbesondere im Frequenzbereich ab 1 MHz. Daher kann gefolgert werden, dass die korrekte Wahl der Totzeit erheblich dazu beitragen kann, die Störemissionen zu reduzieren. Dies ist ohne Erhöhung der Schaltverluste und damit einer Reduzierung des Wirkungsgrades möglich.

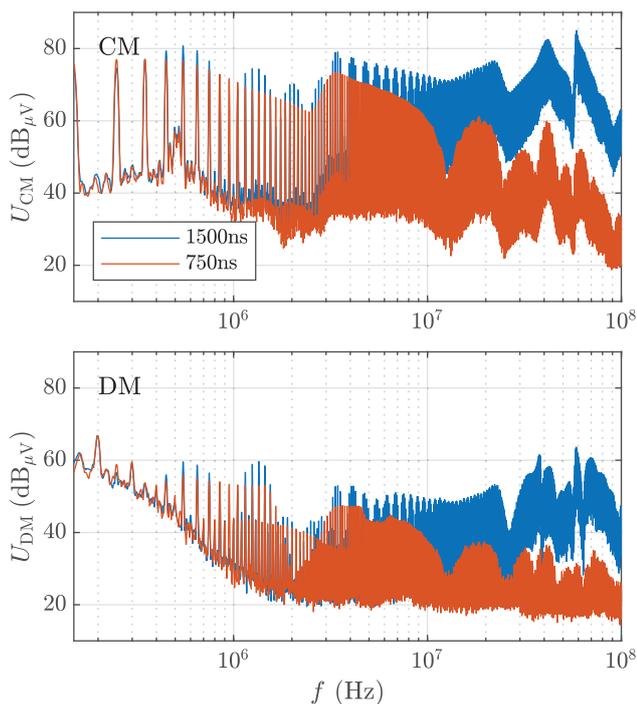


Abbildung 9: Leitungsgebundene Störemissionen des LLC-Resonanzwandlers

## VII. ZUSAMMENFASSUNG UND AUSBLICK

Die Verwendung von Wide-Bandgap-Leistungshalbleitern in der Energieversorgung von Systemen der medizinischen Bildgebung führt zu Herausforderungen der EMV. In dieser Forschungsarbeit wird die EMV am Beispiel eines Resonanzwandlers mit SiC-MOSFETs untersucht und optimiert. Es wird eine Methode zur Modellierung der parasitären Eigenschaften von passiven Komponenten vorgestellt. Die Impedanzen der Bauelemente lassen sich auf diese Weise sehr exakt nachbilden. Dies ermöglicht realitätsnahe Simulationen. Die auftretenden Amplituden der Schwingungen werden jedoch nicht vollständig simuliert. Zur Durchführung von reproduzierbaren EMV-Messungen wird eine EMV-Testumgebung in Betrieb genommen, die an geltende Normen angelehnt ist. Die EMV-Messungen zeigen, dass die Störemissionen des Resonanzwandlers durch die Wahl des optimalen Tastgrades reduziert werden. Dabei wird der Wirkungsgrad nicht gesenkt. Der optimale Tastgrad kann anhand der Simulationen bestimmt werden.

In der Zukunft können detaillierte Modelle des LLC-Resonanzwandlers entwickelt werden. Diese könnten parasitäre Eigenschaften der Leiterplatte und den Ausbreitungspfad der elektromagnetischen Störemissionen einschließen. Dies ermöglicht eine tiefere Analyse der EMV, detailliertere Simulationen und damit weiteres Optimierungspotential. Es kann untersucht werden, inwiefern die erzielten Erkenntnisse dieser Forschungsarbeit auf den Resonanzwandler zur Hoch-

spannungserzeugung im Röntgensystem übertragbar sind.

## DANKSAGUNG

Die Forschungsarbeit „Leistungselektronik in kritischer Infrastruktur am Beispiel der medizinischen Bildgebung – Untersuchung der Elektromagnetischen Verträglichkeit und Optimierung der Resilienz“ des Forschungsprojektes „DiMoLEK“ wird durch das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] B. Zhang und S. Wang, "A Survey of EMI Research in Power Electronics Systems With Wide-Bandgap Semiconductor Devices," in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 1, pp. 626–643, March 2020, doi: 10.1109/JESTPE.2019.2953730.
- [2] S. Yin, K. J. Tseng, C. F. Tong, R. Simanjorang, C. J. Gajanayake, und A. K. Gupta, "A novel gate assisted circuit to reduce switching loss and eliminate shoot-through in sic half bridge configuration," in *2016 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2016, pp. 3058–3064.
- [3] R. Beiranvand, B. Rashidian, M. R. Zolghadri, und S. M. H. Alavi, "Optimizing the normalized dead-time and maximum switching frequency of a wide-adjustable-range llc resonant converter," *IEEE Transactions on Power Electronics*, vol. 26, no. 2, pp. 462–472, 2011.
- [4] M. Sankar, S. Podendorf, K. F. Hoffmann und S. Dickmann, "An EMC- Investigation on Dead Time Distortion of SiC Based Half Bridge LLC Resonant Inverter," in *2024 International Symposium on Electromagnetic Compatibility – EMC Europe*, 2024.
- [5] G. Antonini, "Spice equivalent circuits of frequency-domain responses," *IEEE Transactions on Electromagnetic Compatibility*, vol. 45, no. 3, pp. 502–512, 2003. DOI: 10.1109/TEMPC.2003.815528.
- [6] S. Podendorf, O. Woywode und S. Dickmann, "A Vector Fitting Procedure for the High-Frequency Model Generation of Inductive Components," in *2024 International Symposium on Electromagnetic Compatibility – EMC Europe*, 2024.
- [7] B. Gustavsen und A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 1052–1061, 1999. DOI: 10.1109/61.772353.
- [8] B. Gustavsen, "Improving the pole relocating properties of vector fitting," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1587–1592, 2006. DOI: 10.1109/TPWRD.2005.860281.
- [9] D. Deschrijver, M. Mrozowski, T. Dhaene, und D. De Zutter, "Macro-modeling of multiport systems using a fast implementation of the vector fitting method," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 6, pp. 383–385, 2008. DOI: 10.1109/LMWC.2008.922585.
- [10] "Industrial, scientific and medical equipment - Radio-frequency disturbance characteristics - Limits and methods of measurement," (CISPR 11:2015, modified +A1:2017), Norm.

# Elektrische Lastkraftwagen und Zugmaschinen in Hamburg im Jahr 2050: Hochlauf- und Energiebedarfsprognose

Edvard Avdevcicus\*, Amra Jahic, Detlef Schulz

Professur für Elektrische Energiesysteme

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Hamburg, Deutschland

\*edvard.avdevcicus@hsu-hh.de

**Kurzfassung** – Eine umfassende Analyse der künftigen Elektrifizierungsszenarien im Verkehrssektor ist entscheidend, um Auswirkungen auf das Stromnetz rechtzeitig zu erkennen und zu bewältigen. Dieser Bericht prognostiziert die Anzahl der elektrischen Lastkraftwagen und Zugmaschinen, die bis 2050 in Hamburg im Einsatz sein werden, sowie den resultierenden Energieverbrauch in den einzelnen Stadtteilen bis 2050. Drei Szenarien werden erarbeitet, um unterschiedliche Entwicklungen und Annahmen zur Elektrifizierung und Energienutzung im Güterverkehr abzubilden. Außerdem wird eine Lineare Programmierung (LP) zur Optimierung der Verteilung in der Stadt genutzt, um die effizienteste und nachhaltigste Verteilung der elektrischen Lasten in den verschiedenen Stadtteilen zu ermitteln. Die Ergebnisse können als Grundlage für Strategien zur Integration der elektrischen Lastkraftwagen und Zugmaschinen in das Stromnetz in Hamburg dienen.

**Stichworte** – Lastkraftwagen, Zugmaschinen, Hochlauf, Ladeinfrastruktur, Lastprofil

## NOMENKLATUR

LKW	Lastkraftwagen
ZGM	Zugmaschinen
$F_{Typ,S}$	Anzahl nach Verkehrstyp $Typ$ aus Fläche in einem Stadtteil $S$
$U_{Typ,S}$	Anzahl nach Verkehrstyp $Typ$ aus Unternehmen in einem Stadtteil $S$
$B_{Typ,S}$	Anzahl nach Verkehrstyp $Typ$ aus Satellitenbildern in einem Stadtteil $S$
$A_{Typ,Ansatz,S}$	Differenz der Anzahlen bei verschiedenen Ansätzen nach Verkehrstyp: $F-U$ , $F-B$ , $U-B$
$G_{Ansatz,Typ,S}$	Gewichtungen des Optimierungsproblems für Ansatz, Verkehrstypen und Stadtteil

## I. EINLEITUNG

Die Elektrifizierung von Lastkraftwagen (LKW) und Zugmaschinen (ZGM) spielt eine zentrale Rolle bei der Optimierung der Logistik und der effizienten Nutzung des Stromnetzes. Besonders wichtig ist die Entwicklung einer umfassenden Infrastruktur für Ladestationen in stark frequentierten Regionen wie Hamburg. Diese Maßnahme zielt darauf ab, nicht nur

die logistischen Prozesse zu verbessern, sondern auch das Energiesystem nachhaltiger zu gestalten und grüne Technologien in das gesamte Energie- und Logistiknetz zu integrieren [1], [2], [3], [4], [5], [6], [7], [8], [9]. Durch die Analyse des Potenzials für den Umstieg von konventionellen zu Elektrofahrzeugen werden die Auswirkungen auf den Energieverbrauch sowie auf die Netzstabilität in Hamburg untersucht.

Langfristige Prognosen, die die Entwicklung des Verkehrs und der Elektrifizierung berücksichtigen, sind ein wichtiges Instrument für die Planung und Umsetzung von Maßnahmen der Regierung und der Industrie [10], [11]. Die Prognosen tragen dazu bei, die Umweltauswirkungen des Verkehrssektors zu verringern und die Belastung der Stromnetze wirksam zu steuern. Die Integration von Optimierungstechniken in diese Analyse ergänzt sie zusätzlich und macht sie bei der Bewertung möglicher Entwicklungen vollständiger und präziser.

Bei diesem Bericht handelt es sich um eine Metastudie, die darauf abzielt, die zu erwartende Anzahl elektrischer LKW und ZGM abzuschätzen. Diese Analyse umfasst eine Bewertung des aktuellen Zustands der Lkw- und ZGM-Flotte im Bundesland Hamburg und eine Prognose ihrer Entwicklung auf der Grundlage verschiedener Szenarien aus einer Reihe von Studien [1], [2], [3], [4], [5], [6], [7], [8], [9], die Mobilitätstrends, Innovationen und Änderungen der Verkehrs- und Umweltschutzgesetzgebung berücksichtigen. Unter Beachtung der im Rahmen der Energiewende entwickelten Pläne zur Vollerlektrifizierung enthält der Bericht auch eine detaillierte und umfassende Prognose des Energieverbrauchs des LKW- und ZGM-Verkehrs bis zum Jahr 2050.

Dieser Bericht ist in mehrere Kapitel unterteilt: Kapitel II behandelt die Klassifizierung des Güterverkehrs, Kapitel III analysiert und optimiert die Verteilung auf Hamburgs Stadtteile basierend auf drei Szenarien (Industrie- und Gewerbeflächen, Firmenadressen, Satellitenbilder), Kapitel IV prognostiziert den Anstieg der Fahrzeuganzahl und Kapitel V untersucht den Energieverbrauch bis 2050. Die Ergebnisse liefern wertvolle Informationen zur langfristigen Nachhaltigkeit des Hamburger Verkehrssektors.

## II. KLASSIFIZIERUNG

Für die Beförderung von Gütern im Straßenverkehr werden sowohl auf lokaler als auch auf internationaler Ebene Fahrzeuge verschiedener Typen und Kapazitäten eingesetzt. Die Klassifizierung nach den Richtlinien 2007/46/EG [12] und

97/27/EG [13] ist ein Klassifizierungssystem, das in der Europäischen Union (EU) und einigen anderen Ländern verwendet wird, um Klassen von Fahrzeugen für den Gütertransport zu definieren. Die Klassifizierung unterscheidet drei Klassen von Gütertransportfahrzeugen:

- Klasse N1 mit einer zulässigen Gesamtmasse von bis zu 3,5 Tonnen.
- Klasse N2 mit einer zulässigen Gesamtmasse von 3,5 bis 12 Tonnen.
- Klasse N3 mit einer zulässigen Gesamtmasse über 12 Tonnen.

Die folgenden Fahrzeuge gehören je nach Typ zur Kategorie N:

- Lastkraftwagen (LKW): Kraftfahrzeuge zur Güterbeförderung oder mit besonderer Zweckbestimmung.
- Vans: LKW mit in den Aufbau integriertem Führerhaus.
- Sattelzugmaschinen: Zugmaschinen (ZGM), die ausschließlich oder überwiegend zum Ziehen von Sattelanhängern bestimmt und gebaut sind.
- Straßenzugmaschinen: ZGM, die ausschließlich oder überwiegend zum Ziehen von Anhängern mit Ausnahme von Sattelanhängern bestimmt und gebaut sind.

Die Klasse N1 besteht mehrheitlich aus Vans (Transporter, Lieferwagen usw.) und leichten LKW. Weniger verbreitet sind Fahrzeuge der Klasse N1, die als Straßenzugmaschinen (Traktoren usw.) bekannt sind. Die Klasse N2 und N3 werden im Allgemeinen von LKW und Sattelzugmaschinen für den Transport mittlerer und großer Gütermengen dominiert. Dementsprechend sind mittelschwere LKW und ZGM in N2 und schwere LKW und ZGM in N3 eingeordnet. In diesem Bericht werden daher zwei Hauptgruppen von Transporten der Kategorie N unterschieden: N1-Verkehr und N2/N3-Verkehr.

### III. VERTEILUNG NACH STADTTTEILEN

In Deutschland müssen alle Fahrzeuge und Anhänger bei der örtlichen Kfz-Zulassungsstelle registriert werden, wobei jeder Bezirk ein eigenes Nummernschild erhält [14], [15], [15], [17]. Die Anzahl der zugelassenen Kraftfahrzeuge und Anhänger variiert stark abhängig von Bevölkerungsdichte, Wirtschaftsaktivitäten und Verkehrsinfrastruktur. Für die Analyse der Fahrzeugverteilung in Hamburg werden insbesondere Industriegebiete, Firmenadressen und Satellitenbilder berücksichtigt, um detaillierte Einblicke in Verkehrsmuster und logistische Infrastrukturen der Stadt zu ermöglichen. Diese Studie stützt sich auf die amtliche Statistik des Kraftfahrt-Bundesamtes (KBA) [14] und des Senats der Freien und Hansestadt Hamburg [17].

Die Stadtteile oder Bezirke, in denen die N-Fahrzeuge zugelassen sind, entsprechen oft nicht den Standorten, wo diese überwiegend betrieben werden (z. B. Betriebshöfe). Aus diesem Grund wird in diesem Bericht der Standort der Fahrzeugzulassung nach Hamburger Bezirken nicht berücksichtigt. Stattdessen ist die Grundlage dieser Arbeit die Verteilung von N-Fahrzeugen in 104 Hamburger Stadtteilen nach drei Gruppen von Faktoren des Standorts von Transporteinheiten und ihrer Logistikinfrastruktur in der Stadt:

- Gewerbe- und Industrieflächen in jedem Stadtteil – Hierbei werden Daten aus dem Immobilienmarktbericht Hamburg 2020 verwendet, die Informationen über die Verteilung und Nutzung dieser Flächen liefern [18].
- Anzahl bestimmter Unternehmenstypen in jedem Stadtteil – Dies umfasst Daten des Kraftfahrt-Bundesamtes (KBA, FZ5 und FZ23, 2023) [14], der Handelskammer Hamburg [19] und des Hamburger Telefonbuchs mit Firmenadressen [20], um Einblicke in die Konzentration und Art der Unternehmen in den Stadtteilen zu gewinnen.
- Geparkte N-Fahrzeuge – Diese werden anhand von Satellitenbildern [21] gezählt, wobei die Anzahl der LKW und ZGM auf Parkplätzen durch visuelle Berechnungen ermittelt wird. Die Parkplätze von LKW-Händlern sind dabei ausgenommen.

Die drei betrachteten Faktoren – Gewerbe- und Industrieflächen, Firmenadressen und Satellitenbilder – geben jeweils einen anderen Einblick in die Verteilung von N-Fahrzeugen in Hamburg und tragen zu einem umfassenderen Verständnis der präzisen Situation bei.

Die Gesamtfläche, die für gewerbliche und industrielle Aktivitäten in Hamburg vorgesehen ist, beträgt 69 qkm, was ungefähr 9,2 % der Gesamtfläche der Stadt ausmacht. Die Stadtteile Wilhelmsburg, Waltershof und Billbrook haben den größten Anteil an Industrie- und Gewerbeflächen (jeweils über 4 qkm) und sind Teil des Bezirks Hamburg-Mitte [18]. Ferner sind 74 % des gesamten Stadtteils Billbrook für solche Nutzungen vorgesehen. Der Stadtteil Wilhelmsburg ist das größte Gebiet, was die Verfügbarkeit von Flächen für industrielle und gewerbliche Zwecke angeht.

Anhand der Daten werden die Logistik-, Handels- und Dienstleistungsunternehmen in Hamburg nach ihren Adressen aus dem Hamburger Branchenbuch [20] untersucht, wobei mehr als 15.000 Unternehmen in über 21 Kategorien, darunter Baugewerbe, Handel, Lagerei und Landwirtschaft, erfasst wurden. Hamburg-Altstadt und Neustadt haben die meisten Unternehmen, mit mehr als 700 registrierten Betrieben. Die Lage großer Betriebshöfe und anderer Industrie- und Gewerbeflächen, die für Ladestationen geeignet sind, ermöglicht eine Einschätzung der Entwicklung der Lademöglichkeiten für N-Fahrzeuge in den Hamburger Stadtteilen. Die Verteilung der Fahrzeuge zeigt, dass in den Kategorien Arbeitnehmende und Dienstleistungen mehr als 54 % der N1-Fahrzeuge dominieren, während N2- und N3-Fahrzeuge ebenfalls in den Dienstleistungssektoren über 57 % ausmachen [14], [19].

Die Analyse der Satellitenbilder basiert auf einer gründlichen manuellen Auswertung der Google Earth-Datenbanken, wobei die einzelnen Stadtteile separat analysiert wurden [21]. Eine visuelle Analyse von Satellitenbildern von Hamburg zum 1. April 2023 ergab, dass sich, im Vergleich mit anderen Stadtteilen, die meisten N1-Fahrzeuge in Bahrenfeld und die meisten N2/N3-Fahrzeuge in Wilhelmsburg befinden.

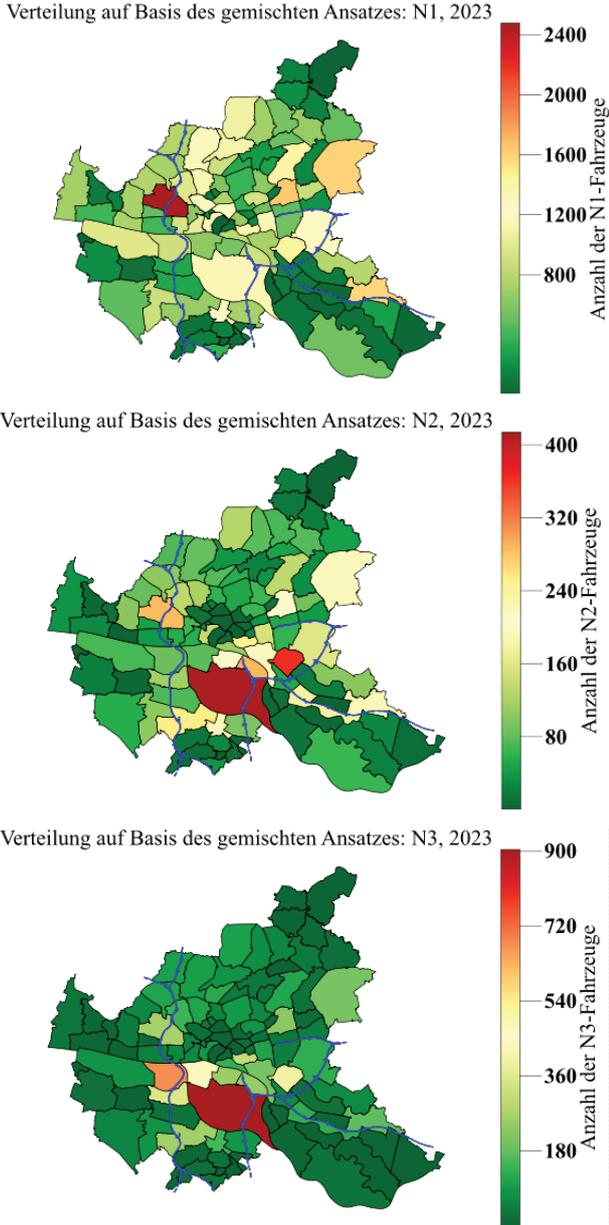


ABBILDUNG 1: ANZAHL DER N1-, N2- UND N3-Fahrzeuge IN DEN HAMBURGER STADTTETLEN IM JAHR 2023

Das Ergebnis der Untersuchung der einzelnen drei Szenarien zeigt ein unterschiedliches Bild der Verteilung von LKW und ZGM in der Stadt. Um zu entscheiden, welcher Ansatz der beste ist, wird eine Optimierung durch lineare Programmierung mit dem Optimierer Gurobi [22] durchgeführt. Die optimale Anzahl von N-Fahrzeugen für jeden  $Typ$ , wobei  $Typ \in \{N1, N2, N3\}$ , und in jedem Stadtteil  $S$ , wobei  $S \in \{1, 2, \dots, 104\}$ , wird auf der Grundlage der drei vorgestellten Ansätze berechnet. Dabei werden die Anzahl der Fahrzeuge, verteilt nach dem Faktor Fläche  $F_{Typ,S}$ , der Standort der Unternehmen  $U_{Typ,S}$  und die Satellitenbilder  $B_{Typ,S}$  berücksichtigt:

$$\min_{Ziel} \rightarrow A_{Typ,F-U,S} + A_{Typ,F-B,S} + A_{Typ,U-B,S}, \quad (1)$$

$$A_{Typ,F-U,S} = G_{F-U,Typ,S} \cdot (F_{Typ,S} - U_{Typ,S}), \quad (2)$$

$$A_{Typ,F-B,S} = G_{F-B,Typ,S} \cdot (F_{Typ,S} - B_{Typ,S}), \quad (3)$$

$$A_{Typ,U-B,S} = G_{U-B,Typ,S} \cdot (U_{Typ,S} - B_{Typ,S}), \quad (4)$$

$$G_{F-U,Typ,S} + G_{F-B,Typ,S} + G_{U-B,Typ,S} = 100\%, \quad (5)$$

wobei die Differenz zwischen der Anzahl der N-Fahrzeuge pro Stadtteil auf der Grundlage des Ansatzes  $A_{Typ,Ansatz} \in \{A_{Typ,F-U}, A_{Typ,F-B}, A_{Typ,U-B}\}$  und den vom Optimierer berechneten Gewichten  $G_{F-U,Typ,S}, G_{F-B,Typ,S}, G_{U-B,Typ,S}$  herangezogen wird.

Für das Jahr 2023 gibt es in Hamburg insgesamt, bezogen auf die Gewichtsklassen, 59.978 Fahrzeuge der Klasse N1, 7.584 N2-Fahrzeuge und 8.690 N3-Fahrzeuge [15]. Anhand der berechneten Verteilungsgewichte aus (5) wird die Anzahl der N-Fahrzeuge proportional auf die Stadtteile verteilt und ist in Abbildung 1 dargestellt. Die höchsten Konzentrationen aller N-Fahrzeuge zusammengenommen finden sich in den Stadtteilen Wilhelmsburg (über 2.600 Einheiten), Billbrook (ca. 2.200 Einheiten) und Rothenburgsort (ca. 1.100 Einheiten). Alle drei Stadtteile sind Teil des Bezirks Hamburg-Mitte. N1-Fahrzeuge sind jedoch im gesamten Stadtgebiet am häufigsten anzutreffen und konzentrieren sich in Gewerbe- und Industriegebieten, insbesondere in Hamburg-Bahrenfeld im Bezirk Altona (über 2.400 N1-Fahrzeuge) und Hamburg-Rahlstedt im Bezirk Wandsbek (ca. 1.600 Einheiten). Die Konzentration mittelschwerer und schwerer N2/N3-Fahrzeuge ist in Wilhelmsburg am höchsten (ca. 400 N2-Fahrzeuge und 900 N3-Fahrzeuge). Im Allgemeinen korrelieren diese Werte mit dem Verkehr in der Stadt und in dem verkehrsreichsten Gebiet des Hamburger Hafens [4], [16].

#### IV. HOCHLAUF IN HAMBURG BIS 2050

Der Trend im Güterverkehr ist ungebrochen und wird auch in Zukunft stetig wachsen [1], [2], [3]. Ein deutliches Wachstum ist bei den Transporten der Kategorie N zu verzeichnen. In diesem Bericht werden Prognosen für die Entwicklung von LKW und leichten Nutzfahrzeugen (Klasse N1) sowie größeren LKW und ZGM (Klassen N2 und N3) vorgestellt, wobei Szenarien aus verschiedenen Studien [1], [2], [3], [4], [8] berücksichtigt werden. Die Anzahl der elektrifizierten LKW und ZGM-Fahrzeuge wird ebenfalls berücksichtigt [1], [2].

##### A. Hochlauf des N1-Fahrzeugs

Das durchschnittliche Szenario für die Anzahl der N1-Fahrzeuge wird auf der Grundlage von [1], [3], [4] für Deutschland und proportional zu Hamburg berechnet. Auch der Anteil des elektrifizierten Fahrzeuge wird im Rahmen der deutschen Zahlen berücksichtigt. Die Prognose wird in drei Szenarien für die N1-Fahrzeuge dargestellt:

- Durchschnittliches Szenario (SN1) und elektrifizierter Teil (E-SN1),
- Optimistisches Szenario (SN1 +15 %) und elektrifizierter Teil (E-SN1 +15 %),
- Konservatives Szenario: (SN1 -15 %) und elektrifizierter Teil (E-SN1 -15 %).

Der Durchschnittswert aus den Prognosen der verwendeten Berichte [1], [3], [4], einschließlich konservativer und optimistischer Szenarien, ist in Abbildung 2 dargestellt.

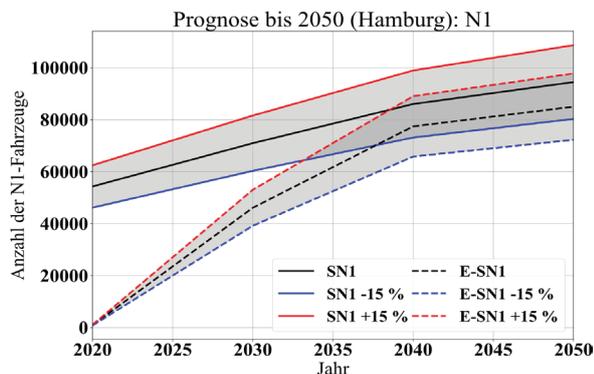


ABBILDUNG 2: PROGNOSE FÜR DIE ANZAHL DER N1-FAHRZEUGE BIS 2050 IN HAMBURG

Bis 2050 werden 90 % des gesamten N1-Fahrzeugs in Hamburg elektrifiziert sein [2]. Zu diesem Zeitpunkt wird die Zahl der elektrifizierten N1-Fahrzeuge im Durchschnitt bei über 85.000 liegen. Alle drei Szenarien zeigen einen Anstieg der Anzahl aller N1-Fahrzeuge und es ist ein klarer Aufwärtstrend zu beobachten. Bei dem konservativen Szenario wird die Gesamtzahl der N1-Fahrzeuge in Hamburg im Jahr 2050 über 80.000 Einheiten liegen, im besten Fall bei 108.000. N1-Fahrzeuge sind vorwiegend im gesamten Hamburger Stadtraum anzutreffen und werden von mittleren und schweren LKW und ZGM deutlich unterlegen.

### B. Hochlauf des N2/N3-Fahrzeugs

Das durchschnittliche N2/N3-Szenario wird identisch zu N1 berechnet und berücksichtigt eine Reihe von Berichten [1], [3], [4], [8] im Kontext von Deutschland und proportional zu den Hamburger Zahlen. In den deutschen Zahlen ist auch der Anteil der elektrifizierten Fahrzeuge berücksichtigt. Die Prognose wird in drei Szenarien für N2/N3-Fahrzeuge dargestellt:

- Durchschnittliches Szenario (SN2/N3) und elektrifizierter Teil (E-SN2/N3),
- Optimistisches Szenario (SN2/N3 +15 %) und elektrifizierter Teil (E-SN2/N3 +15 %),
- Konservatives Szenario (SN2/N3 -15 %) und elektrifizierter Teil (E-SN2/N3 -15 %).

Der Durchschnittswert aus den Prognosen, einschließlich konservativer und optimistischer Szenarien, ist in Abbildung 3 dargestellt. Der Anteil der elektrifizierten N2/N3-Fahrzeuge wird bis 2050 geringer sein als der der N1-Fahrzeuge, da sich die Technologie für schwere Wasserstoff-LKW und ZGM ebenfalls weiterentwickelt [2], [8]. Dennoch wird der Anteil der elektrifizierten LKW und ZGM der Klasse N2/N3 bis 2050 die Mehrheit bilden und voraussichtlich bei 69 % liegen [2]. Die durchschnittliche Anzahl der elektrischen N2-Fahrzeuge in Hamburg wird im Jahr 2050 bei ca. 2.400 und die der N3-Fahrzeuge bei ca. 9.100 liegen. Die Gesamtzahl aller elektrischen N2/N3-Fahrzeuge zeigt eine steigende Tendenz, mit einem Durchschnitt von ca. 11.500 Einheiten bis 2050. Im konservativen Szenario wird die Gesamtzahl bei 10.000 Einheiten liegen, im optimistischen Szenario jedoch bei 13.300. Dementsprechend wird die Nachfrage hoch bleiben, wobei der Schwerpunkt auf der Elektrifizierung liegt.

Ausgehend von den Ergebnissen der vorgestellten Szenarien wäre eine erhebliche Erhöhung der Energiereserve erforderlich, um alle LKW und ZGM von N1 und N2/N3 zu bedienen.

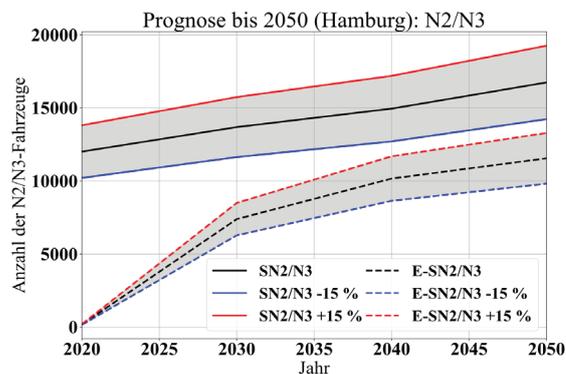


ABBILDUNG 3: PROGNOSE FÜR DIE ANZAHL DER N2/N3-FAHRZEUGE BIS 2050 IN HAMBURG

## V. ENERGIEBEDARF BIS 2050

Die Elektrifizierung von LKW- und ZGM-Depots und das Laden unterwegs benötigen Energiereserven, da die wachsende Anzahl von N-Fahrzeugen die Belastung des Stromnetzes erhöht. Generell gibt es bereits einen klaren Trend zum Nachladen für alle Verkehrsarten [3], [4], [5]. Es wird prognostiziert, dass bis 2030 rund 94 % aller N1-Fahrzeuge nachts aufgeladen werden, während dies für die restlichen 6 % während der Arbeitszeit möglich ist, wenn sich das Fahrzeug an einem Umladeort oder im Transit befindet. Für die Klassen N2 und N3 werden 87 % der Fahrzeuge nachts aufgeladen und 13 % während der Arbeitszeit [5]. Die Erstellung von Lastprofilen beginnt mit der Verwendung von Kurven für eine einzelne nominale Einheit der Klasse N1 und der Klassen N2/N3. Auf der Grundlage statistischer Daten, einschließlich der Aktivität, der durchschnittlich zurückgelegten Strecke und des Energieverbrauchs, wurden Profile sowohl für das Laden im Depot als auch unterwegs erstellt. Auf diese Weise lässt sich die potenzielle Belastung des Stromnetzes vorhersagen.

### A. Energiebedarf der N1-Fahrzeuge

N1-Fahrzeuge haben in der Regel eine geringere Batteriekapazität als N2/N3-Fahrzeuge. Für sie sind Ladestationen im Leistungsbereich von 7 bis 50 kW ausreichend [9]. Herkömmliche Ladestationen für Personenkraftwagen sind also auch für N1-Fahrzeuge geeignet.

N1-Fahrzeuge können an verschiedenen Standorten in der Stadt aufgeladen werden, wie Lagerhäusern, Depots, bei Paketzustellern und Handwerksbetrieben [4], [5], [23]. Die Residenz des Fahrers wird auch von einigen Einzelhändlern und technischen Diensten genutzt [5]. Ausgehend von der Verkehrsaktivität [4], [5] der Klasse N1 ist der erwartete Energiebedarf für das Aufladen von Fahrzeugen der Klasse N1 im Jahr 2050 in Hamburg in Abbildung 4 dargestellt.

Dabei wurde der experimentell ermittelte durchschnittliche Energieverbrauch von 11 kWh pro Tag berücksichtigt [11]. In den Abend- und Nachtstunden, ab ca. 18 Uhr, ist die Auslastung am höchsten. Tagsüber wird in der Regel zwischen 5 und 14 Uhr geladen, allerdings ist die Belastung in dieser Zeit deutlich geringer, da die meisten Verkehrsmittel aktiv genutzt werden [5]. Der Zeitrahmen stimmt auch mit anderen Studien überein [16], [25]. Hamburgs durchschnittliche Spitzenlast wird im Jahr 2050 voraussichtlich 100 MW überschreiten, mit einem konservativen Szenario von ca. 85 MW und einem Maximum von ca. 115 MW.

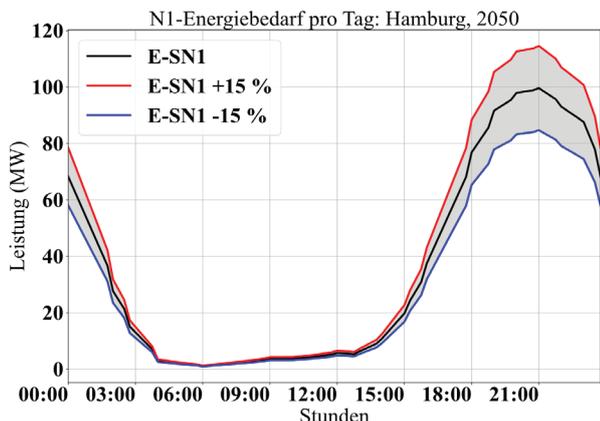


ABBILDUNG 4: NETZBELASTUNG DURCH DAS AUFLADEN VON N1-Fahrzeugen

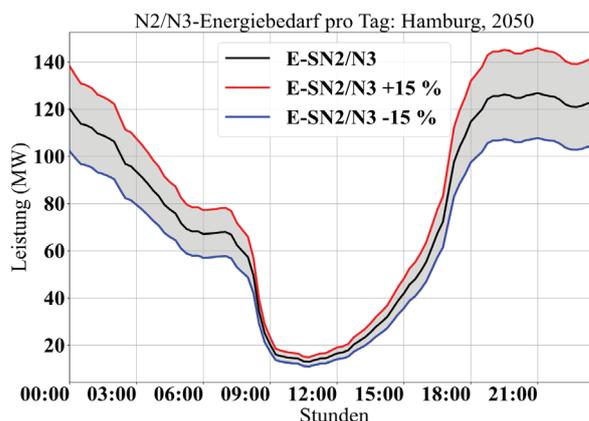


ABBILDUNG 5: NETZBELASTUNG DURCH DAS AUFLADEN VON N2/N3-Fahrzeuge

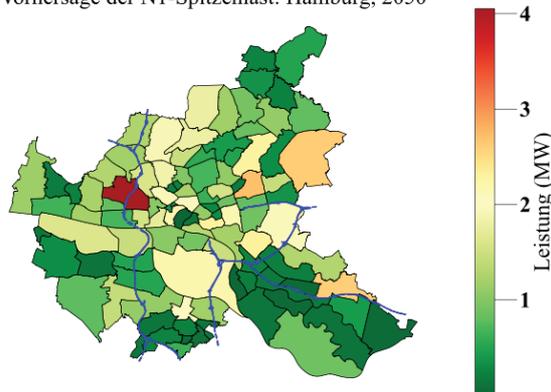
### B. Energiebedarf der N2/N3-Fahrzeuge

Die Ladeinfrastruktur für N2/N3-Fahrzeuge ermöglicht das Aufladen mittlerer und schwerer LKW und ZGM auf unterschiedliche Weise: Laden mit geringer Leistung (bis 150 kW), Schnellladen (bis 350 kW) und Laden mit hoher Leistung (500 kW oder mehr) sowie über Oberleitungen (ca. 130 kW) [2], [4], [6]. Generell hängt die Wahl des Ladestandorts für N2/N3-LKW und N2/N3-ZGM vom Einsatzgebiet ab [4]. Wie bereits erwähnt, laden 87 % der N2/N3-Fahrzeuge hauptsächlich auf dem Betriebs Hof und 13 % während der Arbeitszeit unterwegs [5]. Externe Dienstleisterstandorte eignen sich für N2-Fahrzeuge in der Lebensmittelindustrie, für die Anlieferung und Abholung durch Spediteure, für den Nahverkehr zwischen Logistikzentren und Produktionsstätten, für Bauunternehmen und für den Flughafenbedarf. Be- und Entladeflächen, geeignet für den Nahverkehr mit N2/N3-Fahrzeugen zwischen Logistikzentren und Produktionsstandorten. Diese Industrie- und Gewerbebezonen sind in der Regel in den Stadtteilen dichter konzentriert.

Berichte des Bundesministeriums für Verkehr und digitale Infrastruktur (BMVI) [4] und Verkehrsberichte der Hamburger Hafenbehörde [16] zeigen, dass N2/N3-Fahrzeuge im Hafengebiet recht aktiv sind. Wilhelmsburg ist in dieser Hinsicht das am stärksten frequentierte Stadtteil. Mittlere und schwere LKWs und ZGMs transportieren verschiedene Lebensmittel, Getränke und Non-Food-Artikel [24]. Im Handelssektor macht der Transport mit N2/N3 von Lebensmitteln 85 % und von Getränken 14 % aller Waren aus [25]. Lebensmittel werden hauptsächlich zu Einzelhandelsgeschäften und Lagerhäusern transportiert. Daten zur Fahreraktivität, einschließlich des durchschnittlichen Lastprofils der N2/N3-Einheit, für diese drei Gruppen werden vom National Laboratory System des US-Energieministeriums [25] bereitgestellt und auf die Zählung der Fahrertätigkeit angewandt.

Unter Berücksichtigung der gemessenen Energieverbrauchsprofile [10], [11] und der Fahreraktivität [4], [25] ist der erwartete Energiebedarf für das Aufladen von N2/N3-Fahrzeugen im Jahr 2050 in Abbildung 5 dargestellt. Die Spitzenlast tritt vor allem in den Abend- und Nachtstunden (18:00–03:00 Uhr) auf. Die durchschnittliche nächtliche Spitzenlast wird 125 MW betragen. Im konservativen Szenario werden es ca. 110 MW sein und im Maximalszenario ca. 140 MW.

Vorhersage der N1-Spitzenlast: Hamburg, 2050



Vorhersage der N2/N3-Spitzenlast: Hamburg, 2050

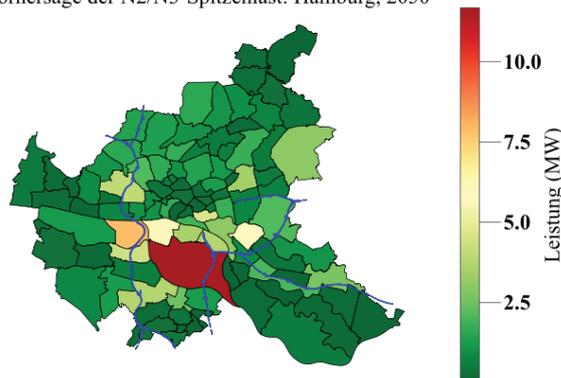


ABBILDUNG 6: SPITZENLASTEN AUS DEM N-VERKEHR IN HAMBURGER STADTEILEN IM JAHR 2050

### C. Energiebedarf nach Stadtteilen

Dieser Bericht bewertet die Belastung im Detail für einzelne Stadtteile. Abbildung 6 zeigt eine Prognose auf Basis der potenziellen Entwicklung des N-Segments in Hamburg bis zum Jahr 2050. Ein höherer Energiebedarf wird in den Bezirken Hamburg-Mitte, Hamburg-Nord und Wandsbek erwartet. Zu einer Last für die Stromnetze wird es in Stadtteilen wie Wilhelmsburg (bis zu 12 MW), Waltershof (bis zu 8 MW), Billbrook (bis zu 6 MW) und Bahrenfeld (bis zu 4 MW) ansteigen. Es besteht bereits jetzt Bedarf an Maßnahmen zur Erhöhung der Anzahl und Kapazität der Ladestationen, da in den kommenden Jahrzehnten erhebliche Reserven benötigt werden. Besonderes Augenmerk sollte auf den Ausbau der Reserven in Hamburg-Mitte, insbesondere im Hafengebiet [16], [26], sowie auf die Überwachung der Umspannwerksreserven in ganz Hamburg gelegt werden.

## VI. SCHLUSSFOLGERUNGEN

Vor dem Hintergrund der Energiewende und aktueller Herausforderungen ist die rechtzeitige Erkennung und Lösung von Problemen im Energiesystem besonders wichtig. Dieser Bericht enthält eine detaillierte Prognose der Anzahl der LKW- und ZGM-Fahrzeuge und des Energieverbrauchs zu den entsprechenden Spitzenlasten in bestimmten Stadtteilen Hamburgs bis zum Jahr 2050. Zur Abbildung der möglichen Entwicklung und Elektrifizierung der LKW- und ZGM-Segmente sowie des Energieverbrauchs wurden drei Szenarien entwickelt. Durch die Analyse verschiedener Quellen und die Anwendung von Optimierungsverfahren wurde die Verteilung der elektrischen Belastung von Elektro-LKW und Elektro-ZGM ermittelt, die in Hamburg im Jahr 2050 durchschnittlich bis zu 225 MW, im konservativen Szenario bis zu 195 MW und im optimistischen Szenario bis zu 255 MW betragen könnte, was als Orientierung für die weitere Netzentwicklung dient. Für die Zukunft ist geplant, LKW und ZGM aus anderen Regionen, insbesondere aus dem Hafengebiet, zu analysieren. Dies ermöglicht eine präzisere Beurteilung des Trends der N-Segment-Entwicklung.

## DANKSAGUNG

Dieser Beitrag im Forschungsprojekt eMob wird durch dtec.bw-Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der European Union – NextGenerationEU finanziert.

## LITERATUR

- [1] „Öko-Institut e.V.: StratES - Szenarien für die Elektrifizierung des Straßengüterverkehrs (2024)“ [Online]. Erhältlich: <https://www.oeko.de/fileadmin/oekodoc/StratES-Szenarien-Elektrifizierung-Strassengueterverkehr.pdf>. [Zugriff am 6. Juni 2024].
- [2] „Agora Energiewende: Klimaneutrales Deutschland 2045“ [Online]. Erhältlich: <https://www.agora-energiewende.de/publikationen/klimaneutrales-deutschland-2045-1>. [Zugriff am 6. Juni 2024].
- [3] „Transport and Environment (T&E): A roadmap for electric truck charging infrastructure deployment (2020)“ [Online]. Erhältlich: [https://te-cdn.ams3.cdn.digitaloceanspaces.com/files/2020\\_02\\_RechargeEU\\_trucks\\_paper.pdf](https://te-cdn.ams3.cdn.digitaloceanspaces.com/files/2020_02_RechargeEU_trucks_paper.pdf). [Zugriff am 10. Juni 2024].
- [4] „Wissenschaftliche Beratung und Begleitung des BMVI zur Weiterentwicklung der Mobilitäts- und Kraftstoffstrategie (MKS): Use Cases für klimafreundliche Nutzfahrzeuge“ [Online]. Erhältlich: <https://opus4.kobv.de/opus4-flwv/frontdoor/index/index/docId/3213>. [Zugriff am 7. Juni 2024].
- [5] D. Tyrer: Analysis to identify the EV charging requirement for vans, Climate Change Committee, United Kingdom, 2022, CID: 20.500.12592/vbrnj6.
- [6] „European Electric Vehicle Charging Infrastructure Masterplan (2023) ACEA,“ [Online]. Erhältlich: <https://www.acea.auto/publication/european-electric-vehicle-charging-infrastructure-masterplan/>. [Zugriff am 6. Juni 2024].
- [7] A. Kindl, E. Niemeier, S. Wilhelm, L. Benda, J. Nöppert, M. Winter, A. Steindl und R. Demtschenko, R.: Wissenschaftliche Beratung des Bmvi Zur Mobilitäts- und Kraftstoffstrategie. Schlussbericht zur Studie „Fahrradparken an Bahnhöfen, ORLIS Repositorium, Germany, 2019. <https://orlis.difu.de/items/54bef5d9-7c97-41c8-a275-f8d2d656d6a5>
- [8] „Bundesministerium für Digitales und Verkehr (Hg.): Gleitende Langfristverkehrsprognose 2021-2022“ [Online]. Erhältlich: <https://henry.baw.de/server/api/core/bitstreams/e6ad7dfa-2320-41c3-ba06-ab5493596e3d/content>. [Zugriff am 7. Juni 2024].
- [9] „European Electric Vehicle Charging Infrastructure Masterplan (2023) ACEA,“ [Online]. Erhältlich: <https://www.acea.auto/publication/european-electric-vehicle-charging-infrastructure-masterplan/>. [Zugriff am 6. Juni 2024].
- [10] Z. Gao, Z. Lin, O. Franzese: Energy Consumption and Cost Savings of Truck Electrification for Heavy-Duty Vehicle Applications, Transportation Research Record, 2017, 2628(1), 99-109, doi: <https://doi.org/10.3141/2628-11>.
- [11] W. Cieslik, W. Antczak: Research of Load Impact on Energy Consumption in an Electric Delivery Vehicle Based on Real Driving Conditions: Guidance for Electrification of Light-Duty Vehicle Fleet, Energies 2023, 16, 775, doi: <https://doi.org/10.3390/en16020775>.
- [12] „Richtlinie 2007/46/EG des Europäischen Parlaments und des Rates vom 5. September 2007 zur Schaffung eines Rahmens für die Genehmigung von Kraftfahrzeugen und Kraftfahrzeuganhängern sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge,“ [Online]. Erhältlich: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32007L0046>. [Zugriff am 19. April 2023].
- [13] „Richtlinie 97/27/EG des Europäischen Parlaments und des Rates vom 22. Juli 1997 über die Massen und Abmessungen bestimmter Klassen von Kraftfahrzeugen und Kraftfahrzeuganhängern und zur Änderung der Richtlinie 70/156/EWG,“ [Online]. Erhältlich: <https://op.europa.eu/de/publication-detail/-/publication/ae116d9b-10cd-4780-87cb-aad6832ea80a>. [Zugriff am 19. April 2023].
- [14] „Kraftfahrt-Bundesamt (KBA): Systematisierung von Kraftfahrzeugen und ihren Anhängern,“ [Online]. Erhältlich: [https://www.kba.de/DE/Statistik/Verzeichnisse/systematische\\_verzeichnisse\\_inhalt.html](https://www.kba.de/DE/Statistik/Verzeichnisse/systematische_verzeichnisse_inhalt.html). [Zugriff am 6. Juni 2024].
- [15] „Pressemitteilung Nr. 08/2023: Der Fahrzeugbestand am 1. Januar 2023“ [Online]. Erhältlich: [https://www.kba.de/DE/Presse/Pressemitteilungen/Fahrzeugbestand/2023/pm08\\_fz\\_bestand\\_pm\\_komplett.html](https://www.kba.de/DE/Presse/Pressemitteilungen/Fahrzeugbestand/2023/pm08_fz_bestand_pm_komplett.html). [Zugriff am 8. Juni 2024].
- [16] „Straßenverkehrsbericht 2018: Hafenerkehr erfassen, verstehen und verändern (2019)“ [Online]. Erhältlich: [https://www.hamburg-port-authority.de/fileadmin/user\\_upload/Strassenverkehrsbericht2018.pdf](https://www.hamburg-port-authority.de/fileadmin/user_upload/Strassenverkehrsbericht2018.pdf). [Zugriff am 8. Juni 2024].
- [17] „Schriftliche Kleine Anfrage des Abgeordneten Dennis Thering (CDU) vom 16.04.19 und Antwort des Senats (2019),“ [Online]. Erhältlich: [https://www.buergerschaft-hh.de/parldok/dokument/66450/entwicklung\\_der\\_zulassungszahlen\\_von\\_pkw\\_lkw\\_und\\_krafttraedern\\_in\\_hamburg\\_im\\_1\\_quartal\\_2019.pdf](https://www.buergerschaft-hh.de/parldok/dokument/66450/entwicklung_der_zulassungszahlen_von_pkw_lkw_und_krafttraedern_in_hamburg_im_1_quartal_2019.pdf). [Zugriff am 6. Juni 2024].
- [18] „Immobilienmarktbericht Hamburg 2020: Gewerbe- und Industrie-flächen“ [Online]. Erhältlich: <https://www.hamburg.de/contentblob/13894842/8f3b529e265b0f91e0726750e8c25064/data/d-immobilienmarktbericht-2020.pdf>. [Zugriff am 6. Juni 2024].
- [19] „Hamburger Wirtschaftszahlen: Kammerzugehörige Unternehmen nach Bezirken“ [Online]. Erhältlich: <https://www.ihk.de/hamburg/produktmarken/beratung-service/konjunktur-statistik/hamburger-wirtschaft-zahlen/bezirke-3676930>. [Zugriff am 6. Juni 2024].
- [20] „Das Telefonbuch: Alles in einem“ [Online]. Erhältlich: <https://www.telefonbuch.de/>. [Zugriff am 6. Juni 2024].
- [21] „Google Maps“ [Online]. Erhältlich: <https://www.google.com/maps>. [Zugriff am 6. Juni 2024].
- [22] „Gurobi 11.0: Every solution, globally optimized“ [Online]. Erhältlich: <https://www.gurobi.com/>. [Zugriff am 19. Juni 2024].
- [23] „KEP-Studie 2022: Analyse des Marktes in Deutschland“ [Online]. Erhältlich: <https://www.forschungsinformationssystem.de/servlet/is/Entry.555924.Display/>. [Zugriff am 6. Juni 2024].
- [24] „Lebensmittelhandel: Anteile der Warengruppen an Fläche, Umsatz und Ertrag eines Supermarkts in Deutschland im Jahr 2021 (in Prozent)“ [Online]. Erhältlich: <https://www.handelsdaten.de/lebensmittelhandel/anteile-warengruppen-flaechen-umsatz-ertrag-supermarkt>. [Zugriff am 6. Juni 2024].
- [25] M. Muratori, B. Borlaug: Perspectives on Charging Medium- and Heavy-Duty Electric Vehicles, United States, 2022, <https://www.osti.gov/servlets/purl/1840708>.
- [26] „Metadaten: Verkehrszählstellen Hamburg“ [Online]. Erhältlich: <https://www.geoportal.de/Metadaten/7E0B7C8A-20D5-4BFB-A9DA-C892638F74C8>. [Zugriff am 6. Juni 2024].

# Textiles in the Digital Age: Navigating the Integration of the Digital Product Pass within the Fab City Framework

Sarah Prien

*Open Lab Circular Textiles Hamburg*

*House of All gUG*

Hamburg, Deutschland

sarah@wearall.clothing

**Abstract** – This paper critically explores the challenges surrounding the implementation of the Digital Product Pass (DPP) within the context of a Fab City model, focusing on the textile industry exemplified by the use case Fab City Hamburg. The DPP, initiated by the European Union as part of its Circular Economy Action Plan, aims to enhance product traceability and promote sustainable consumption by providing detailed information about a product's lifecycle. However, its adoption in a Fab City setting, characterized by local production, community-based practices, and a focus on reducing consumption, encounters several obstacles. Financial constraints, particularly affecting small and medium-sized enterprises, pose a significant barrier to DPP implementation. Standardization of data models is another critical issue, given the diversity of products and manufacturing processes in Fab Cities. Ensuring transparency and accessibility of product information, influencing consumer behaviour towards sustainability, and addressing the scalability of upcycling processes are also major challenges. Critics argue that the DPP's emphasis on recycling may undermine efforts to prevent waste generation and that it lacks focus on systemic changes needed for genuine sustainability. The potential for greenwashing and the need for a user-friendly solution that accommodates the unique characteristics of Fab City production are also highlighted. This exploration underscores the complexities of integrating digital technologies with sustainability goals in a Fab City model, emphasizing the need for supportive policies and innovative approaches to overcome these challenges.

**Keywords** – *Cosmo-localism, commons-based future, digitalisation, digital product pass, textile industry*

## NOMENKLATURA

DPP	Digital Product Passport
EU	European Union
CEAP	Circular Economy Action Plan
ESPR	Eco Design Sustainable Products Regulation
CSDDD	Corporate Sustainability Due Diligence Directive

## I. INTRODUCTION

In a world with finite resources and growing environmental issues, the traditional linear consumption model is unsustainable. The circular economy, which relies on reuse and recycling, offers a regenerative solution. This approach aims not only to reduce harmful products but also to establish systems that benefit the environment and society. This text explores two digital concepts addressing the sustainability crisis: the Digital Product Pass (DPP) and the Fab City initiative. It critically examines their feasibility in the context of a Fab City exemplifying a case study from the research project Fab City by dtec.bw in Hamburg. The project part dedicated to textiles is carried out by Sarah Prien, her company House of All and her team Through qualitative research methods they endeavour to figure how Hamburg could become a Fab City in the textiles sector and evaluate local manufacturing capacities and the local scene.

## II. CIRCULAR ECONOMY AND SUSTAINABLE PRODUCT LIFECYCLE MANAGEMENT

The circular economy emphasizes keeping products and their components in circulation through reuse and recycling. Sustainable product lifecycle management is central to this approach, aiming to minimize environmental impacts while promoting economic viability and social justice. In the textile industry, significant challenges exist. Currently, only about 1% of textiles are converted into new textiles, and 26% are recycled, leaving a substantial "Circularity Gap" [1].

## III. DIGITAL PRODUCT PASS (DPP) BY THE EU

The European Union's Circular Economy Action Plan (CEAP) [2] includes implementing the Digital Product Pass (DPP) to address the Circularity Gap. The DPP is a structured collection of product-related data with a unique identifier, designed to enhance industry digitalization, regulatory frameworks, business opportunities, and circular economy efficiency.

The DPP aims to influence consumer behaviour towards sustainable purchasing by making sustainability aspects of a product lifecycle visible. Traceability of product information is crucial for maintaining products in a circular economy. Without consistent and precise flows of information about resources, products, and processes, lifecycle gaps cannot be closed.

### A. Technical Structure and EU Requirements for DPP

A DPP consists of three components: the DPP system, the digital twin of the product, and the DPP data. The system integrates standards and protocols into the IT architecture, covering data carriers, unique identifiers, access rights, interoperability, data transmission, storage, processing, authentication, reliability, and security. It must be referred to as a Digital Product Pass Ecosystem as pointed out by King et al. (2023) with intense implications [3]. The European Commission has adopted amendments to the EU Regulation for Ecodesign Sustainable Products (ESPR) in April 2024 to facilitate this [4]. Article 5 sets out the requirements for the design, from which the information requirements for the DPP should contain.

The textile industry has been working on a proposal for the Digital Product Passport (DPP) that meets the requirements of the European Commission. The overarching project for textiles DPP associated with this is CIRPASS. This initiative aims to identify the most suitable proposal, which then paves the way for defining the next steps in the entire industry. Especially the textile industry, one of the most environmentally harmful industries, plans to implement DPP regulations by 2027 and complete implementation by 2030.

### B. Application

The product is entered into an existing database along with all information about the material, manufacturing instructions including relevant processes and required resources, instructions for recycling the product, and ideally usage instructions, such as care tips for textiles. The full list of requirements is shown in Figure 1. Target groups are therefore varied, so different access rights must be configured. This database, usually maintained by the brand, must now be linked to a DPP system to make the data accessible via a unique identifier, the physically affixed data carrier, when the product "travels".

### C. Challenges and Market Solutions

The Digital Product Pass (DPP) emerges as a critical tool within the European Union's Circular Economy Action Plan (CEAP), aiming to bridge the gap between linear and circular economies, notably within the textile industry. This initiative is part of a broader regulatory framework that includes the Corporate Sustainability Due Diligence Directive and the Extended Producer Responsibility Act. These regulations mandate companies to identify and mitigate adverse human rights and environmental impacts across their global value chains, signalling a shift towards more sustainable and responsible corporate behaviour. However, the implementation of these regulations requires time and systemic changes that cannot occur overnight, highlighting the complexity of transitioning from established systems to more sustainable practices.

The DPP is envisioned as a mechanism to influence consumer behaviour towards sustainable purchasing and responsible product ownership by making sustainability aspects of a product lifecycle visible. Yet, the directive's effectiveness is contingent upon the ability of sustainability criteria to motivate purchases, a factor that studies suggest are secondary considerations for many consumers [5]. This raises concerns about the potential for greenwashing, where only specific, positive aspects of a product's lifecycle are highlighted, omitting critical information that could influence consumer choices.

Extract from Ecodesign Requirements Framework (December 2023)

Regulation establishing a framework for setting ecodesign requirements for sustainable products and repealing Directive 2009/125/EC

In order to address environmental impacts which arise in any of the stages of products' life cycle, and based on the product parameters in Annex I, the Commission shall establish ecodesign requirements to improve the following product aspects when relevant to the product group concerned

- (a) durability;
- (b) reliability;
- (c) reusability;
- (d) upgradability;
- (e) repairability;
- (f) the possibility of maintenance and refurbishment;
- (g) the presence of substances of concern;
- (h) energy use and energy efficiency;
- (i) water use and water efficiency;
- (j) resource use and resource efficiency;
- (k) recycled content;
- (l) the possibility of remanufacturing;
- (m) recyclability;
- (n) the possibility of the recovery of materials;
- (o) environmental impacts, including carbon footprints and environmental footprints;
- (p) expected generation of waste.

FIGURE 1: REQUIREMENTS TO BE INSERTED IN THE DPP

Moreover, the scope of the Corporate Sustainability Due Diligence Directive is specifically targeted at large EU companies (with more than 1,000 employees and a worldwide turnover of more than €450 million) and non-EU companies doing significant business in the EU (with a turnover of more than €450 million in the EU). This targeted approach underscores the directive's intention to leverage the influence of larger corporations to drive systemic change towards sustainability. However, it also raises questions about the inclusivity of such measures and the potential for smaller companies to navigate the complexities of compliance without significant resources.

The directive also introduces obligations for companies to undertake risk-based due diligence, adopt a climate transition plan, and report on their implementation and effectiveness. These requirements are designed to ensure that companies actively engage in identifying and mitigating adverse environmental and human rights impacts throughout their supply chains. However, the practical implications of these obligations, particularly for smaller companies, remain unclear, highlighting the need for further guidance and support to ensure widespread compliance.

In the realm of product passports, the market currently offers various solutions, each designed to address different problems. This diversity leads to inconsistencies in data and standards, complicating the understanding of benefits and the establishment of global standards for international product distribution. Current solutions are proprietary, adding to their cost, and their implementation requires high-capacity processes that

may not be affordable for smaller entities. Notably, only one solution met sustainability criteria using the Higgs index, as evaluated by Teresa Dianelly Flores Ramirez [6] in her thesis. Her work underscores the necessity for standardization of data models to render the DPP feasible.

#### IV. FAB CITY INITIATIVE: A BOTTOM-UP APPROACH

##### A. Concept and Vision

The Fab City [7] initiative represents a global movement aimed at transforming urban environments into self-sufficient entities by 2054 that are locally productive and globally connected. This concept is called cosmo-localism and holds the promise of shifting towards a more sustainable future [8]. Originating from the global Fab Lab Network, Fab City seeks to address social and environmental challenges within cities by empowering citizens and shifting the urban paradigm towards sustainability. This involves a significant cultural shift, requiring collaboration among local government, civic organisations, startups, universities, and other entities.

##### 1) Key Components of the Fab City Initiative [9]:

- **Fab City Network:** A federated network of cities committed to becoming Fab Cities, facilitating knowledge sharing and best practices.
- **Full Stack:** Represents the comprehensive framework that guides cities in implementing the Fab City challenge. It outlines the systemic change from centralized production and consumption to distributed production and self-sufficiency, emphasizing the importance of data in driving urban transformation.
- **Manifesto:** Defines the principles and values that underpin the Fab City Global Initiative. Commitment to these principles is essential for cities seeking to become part of the Fab City movement [10].

##### 2) Principles and Implementation

At the heart of the Fab City concept is the transition from a linear economy to a circular one, where more production occurs within the city, along with recycling materials and meeting local needs. The local inventiveness. This shift is encapsulated in the transition from "Products In Trash Out" (PITO) to "Data In Data Out" (DIDO), indicating that a city's imports and exports would predominantly consist of data, including information, knowledge, design, and code.

##### B. Fashion Localism in a Fab City

Fashion holds universal appeal, making textiles indispensable in our lives. The prevailing model of mass production, however, creates a disconnection between consumers and the manufacturing processes. Clothing becomes a mere commodity, detached from its cultural significance and basic necessity status. When individuals engage in designing and crafting their own attire, they acquire profound insights into the materials, processes, and labour involved. This intimate relationship fosters a deeper appreciation for clothing, encouraging practices such as mending and adapting garments to suit evolving preferences or requirements. Such behaviours not only conserve resources but also instil a sense of ownership and pride in one's wardrobe, underscoring the emotional value of clothing [11]. Consequently, this approach contributes to enhanced sustainability.

Personal involvement in creating or altering clothing empowers individuals, granting them autonomy over their fashion decisions and diminishing reliance on mass-produced, single-use items. This empowerment resonates with broader sustainability objectives by advocating for a culture of reuse, repair, and reduction, diverging from the fast-fashion ethos of unceasing consumption and discarding. Future production strategies could markedly lessen dependence on cotton and fibres from the Global South by repurposing existing garments and employing locally sourced fibres for upcycling, repair, and redesign. Community-based mending not only carries social worth but also cultivates social solidarity through collective responsibility for garment care, potential style transformations, and sustainable wear [11].

##### C. Case Study: Textile Manufacturing in Hamburg

Hamburg's textile manufacturing sector stands at a crossroads, facing significant challenges and opportunities in its quest to evolve into a Fab City model. This section delves into the current landscape of textile manufacturing, shedding light on the digital capacity gap and the intricacies of upcycling.

##### 1) Research Methodology

House of All collaborates with Helmut Schmidt University/University of the Federal Armed Forces on the Fab City project to evaluate the opportunities and challenges of transforming textile manufacturing in Hamburg into a Fab City model. House of All runs the living lab Open Lab Circular Textiles where aforementioned practices of care and local manufacturing are explored on an experimental scale. The study employed qualitative methods, including semi-structured interviews with 15 tailors from various ethnic backgrounds. Over six months beginning in December 2022, qualitative interviews were conducted with local craftsmen and designers to gauge current practices [7] and gather insights into their daily practices, challenges, and perceptions of sustainability and digital tools. Additionally, observations were made during visits to tailor workshops and designer ateliers to complement the interview data.

##### 2) Current Practices and Capabilities

The interviews underscored the pressing need for digital capacity building and highlighted the absence of circularity in the daily routines of most tailors. The majority of tailors, focused on community-centric production, lacked familiarity with sustainability or circular design, prioritizing mending and on-demand garment production.

- **Digital Literacy and Sustainability Awareness:** Most tailors demonstrated limited knowledge of sustainability concepts and circular design, focusing instead on traditional garment repair. The concept of circular design seemed to be perceived as a luxury, reflecting broader issues of access to sustainable education and resources.
- **Analog Processes in Garment Production:** The production process remained largely analogue, with minimal use of digital tools beyond basic communication devices. This was attributed to a lack of training and the perception that digital tools are unnecessary for their work. Many designers still make the patterns by hand and prefer this way of working, stating that they have greater control over silhouettes.

### 3) Non-Scalable Upcycling: A Challenge Explained

Upcycling, a cornerstone of circular economies, faces scalability challenges. As highlighted by two of the interviewed designers, the process of upcycling is time-consuming and requires a tailored approach, making it less amenable to large-scale production and digitization [13]. The complexity arises from the need to handle cuttings individually, often necessitating one garment at a time, which contrasts sharply with the efficiencies of traditional production methods. Additionally, the variability in fabric remnants complicates the upcycling process, making it more akin to an artful, manual task rather than a scalable industrial operation.

## V. COMPARATIVE ANALYSIS OF DPP AND FAB CITY

The Digital Product Pass (DPP) is intended to complement the existing global system of brands and mass manufacturing, including the fast-fashion industry, through recycling mechanisms. However, there is a recognized need for a system that encourages local ecosystems to produce and manage their own values. Despite its potential advantages, the DPP infrastructure presents challenges for small and medium-sized enterprises due to financial constraints, especially in implementing these systems.

### A. Ecodesign Regulation and Recycling

The Ecodesign Regulation focuses on the enduring durability of products, with recycling considered only as a final resort. However, starting from January 2025, the expectation is that the internalization of recycling costs will make recycling the preferred option, potentially accelerating mass production.

#### 1) Critique of the Current Overproduction-Centric Model

The prevailing fashion system is criticized for generating excessive demand for disposable goods, prioritizing production over consumption reduction. The DPP's emphasis on recycling overlooks the initial 'R' of the 9R model — Refuse — and fails to challenge the production of new clothing. The efficacy of technological advancements remains uncertain, as demonstrated by Renewcell's bankruptcy despite significant investment, underscoring the limitations of technology alone in addressing sustainability challenges. Thus, according to De-Growth models, reducing consumption becomes essential. As Kate Fletcher argues, we have a limited timeframe to halve the impact of resource consumption by 75% [14].

#### 2) Fab City Initiative and Local Production

In contrast to the EU's top-down approach, the Fab City initiative supports a bottom-up strategy that starts with "Rethink" to tackle sustainability issues. It advocates for local textile production using available materials, promoting proximity between production and usage. This approach suggests extended product lifecycles and potential for consumption reduction and enhanced resource efficiency through community-based practices like mending and Do It Together (DIT). Many local designers' unique selling point is the personalized touch of their products, offering bespoke options. Implementing a DPP in this context would require entering information on a case-by-case basis, posing challenges for transparency. Determining necessary, available, and helpful information becomes crucial. If production scales down to meet local needs, the complexity of system implementation versus labelling basic information, such as materials and recyclability, may be

reconsidered. Technologies like Infrared Scanning could replace RFID tags, improving recycling efficiency.

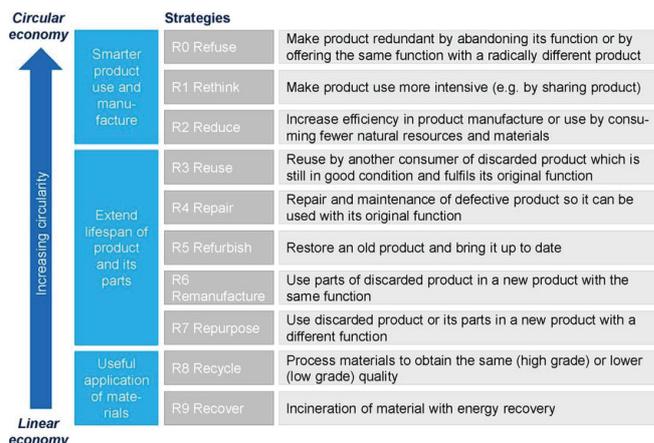


FIGURE 2: THE 9 R MODEL [18]

The debate on fashion sustainability contrasts analogue and digital processes. While digitalization, exemplified by the DPP, aims to streamline operations and increase transparency, it also raises concerns about increased consumption and environmental impact. Analogue processes, such as sewing, mending, and personal clothing creation, are advocated for their potential to reduce consumption, conserve resources, and foster a deeper emotional connection with clothing. This approach aligns with the principles of the "craft of use," emphasizing engagement with clothing to extend its lifecycle and minimize waste [15].

## VI. CHALLENGES AND CONSIDERATIONS

The Digital Product Pass (DPP) endeavours to optimize the fashion supply chain and augment transparency, albeit potentially exacerbating consumption and complexity. Furthermore, it perpetuates the existing global system reliant on unstable supply chains, which are underpinned by the exploitation of workers in the Global South and Far East. Enhancing working conditions in these contexts is a formidable challenge, given the current stability of these conditions.

Conversely, rendering the sustainability facets of a product's lifecycle discernible to consumers might sway their purchasing behaviour. The DPP equips consumers with information to select more sustainable alternatives, catalysing demand for environmentally friendly products and practices within the fashion industry.

In juxtaposition, the Fab City initiative advocates for local production, mitigating consumption through communal practices, thereby fostering resilient supply chains. The Fab City model, grounded in a bottom-up approach, seems more viable for sustainability by cultivating sustainable local ecosystems. The proliferation of Fab Cities is evident, with a recent addition of a Fab Region in Germany.

The concept of upcycling, involving the design and manufacture of clothing in small batches or singular pieces, primarily executed by independent labels, though challenging to scale, would likely dominate in a Fab City. Prioritizing waste prevention generates both economic and social benefits [16]. Engaging in the commons implies the existence of alternative economic models that eliminate brand rivalry and involve

individuals in the design and production of clothing through community-supported concepts [17]. The extent to which a digital product pass could be beneficial and implemented within these newly formed commons-based infrastructures, inherently sustainable [8], remains to be determined.

Future endeavours should bolster research and community-driven initiatives aimed at constructing a digital ecosystem compatible with the specific demands of textile-centric cosmo-local production. Implementing a Digital Product Pass in harmony with Fab City principles offers numerous prospects but also presents hurdles. These include the necessity for uniform data models to ascertain the viability of the DPP, the imperative for digital capability enhancement among designers, manufacturers, and consumers, and the need for supportive policies to surmount financial and logistical obstacles to local production. Crucially, it demands a user-friendly solution that is dynamic, openly accessible, and provides an index of materials available in the Fab City, ensuring compatibility with extant systems. Moreover, transitioning to a more sustainable and localized fashion industry necessitates a fundamental systemic overhaul, accompanied by substantial incentives for local craftspeople, whose numbers have dwindled in recent years.

#### ACKNOWLEDGEMENT

This research paper is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] Ellen MacArthur Foundation (2017): Ellen MacArthur Foundation, A new textiles economy: Redesigning fashion's future.
- [2] A new Circular Economy Action Plan - For a cleaner and more competitive Europe: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0098>.
- [3] Melanie R.N. King, Paul D. Timms, Sara Mountney: A proposed universal definition of a Digital Product Passport Ecosystem (DPPE): Worldviews, discrete capabilities, stakeholder requirements and concerns, *Journal of Cleaner Production*, Volume 384, 2023, 135538, ISSN 0959-6526, <https://doi.org/10.1016/j.jclepro.2022.135538>. (<https://www.sciencedirect.com/science/article/pii/S0959652622051125>)
- [4] <https://www.consilium.europa.eu/media/69109/st16723-en23.pdf> (accessed 01.07.2024)
- [5] Dr. Corinna Fischer, Katja Moch, Siddharth Prakash, Dr. Jenny Teufel, Dr. Immanuel Stieß, Sarah Kresse, Barbara Birzle-Harder (2019): Nachhaltige Produkte – attraktiv für Verbraucherinnen und Verbraucher?, Umweltbundesamt.
- [6] Flores Ramirez, Teresa Dianelly (2022/2023): Exploring the Feasibility and Limitations of Digital Product Passports in the Textile Industry: A Critical Assessment of Current Models. <https://atlasofthefuture.org/project/fab-city/>
- [7] Michel Bauwens, Vasilis Niaros (2017): Changing Societies through Urban Commons Transitions. P2P Foundation and Heinrich Boll Foundation. <http://commonstransition.org/changing-societies-through-urban-commons-transitions/> (accessed 1.7.2024).
- [9] <https://www.fabcity.hamburg/en/fabcity/fcglobal/>
- [10] <https://fab.city/resources/manifesto/>
- [11] Fletcher, K. (2014): Sustainable Fashion and Textiles: Design Journeys (2nd ed.). Routledge. <https://doi.org/10.4324/9781315857930>
- [12] <https://www.wearall.clothing/blog/forschung-2/lokalproduktion-im-sinne-einer-fab-city-1>
- [13] <https://www.theupcyclingtoolkit.com/challenges-solutions-of-upcycling>
- [14] Fletcher, K.; Tham, M. (2019): Earth Logic Fashion Action Research Plan. London: The J J Charitable Trust.
- [15] Fletcher, K. (2016): Craft of Use: Post-Growth Fashion (1st ed.). Routledge. <https://doi.org/10.4324/9781315647371>
- [16] Felipe Schmidt Fonseca (2022): Reuse Commons – a toolkit to weave generous cities; DOI: 10.5281/zenodo.7432153
- [17] Silke Helfrich, David Bollier (2019): Frei, fair und lebendig – Die Macht der Commons, transcript Verlag.
- [18] Potting, José & Hekkert, M.P. & Worrell, Ernst & Hanemaaijer, Aldert. (2017). Circular Economy: Measuring innovation in the product chain. p.5.

# Die Fab City Hamburg auf dem Weg zu einer Open-Source-basierten Kreislaufwirtschaft: Förderliche und hinderliche Faktoren der Transformation

Merle Ibach\*, Jörg Knieling  
 Fachgebiet Stadtplanung und Regionalentwicklung  
 HafenCity Universität Hamburg  
 \*merle.ibach@hcu-hamburg.de

**Kurzfassung** – Die Fab City Initiative Hamburg ist Teil der globalen Fab-City-Bewegung, die auf die Herausforderungen von Ressourcenknappheit, sozialer Ungleichheit und Städtewachstum reagiert und eine nachhaltige lokale Produktion anstrebt. Diese Studie untersucht die Transformation von Hamburg zu einer Open-Source-basierten Kreislaufwirtschaft (OSCE) und die Rolle von Transformationspionieren (TP) in diesem Prozess. Mittels qualitativ-interpretativer Interviews wurden förderliche und hinderliche Faktoren identifiziert, die sich für die TP bei der Umsetzung der OSCE stellen. Die Ergebnisse zeigen, dass TP durch Innovationskraft, intrinsische Motivation und interdisziplinäre Netzwerke entscheidend zur Transformation beitragen, jedoch vor erheblichen Herausforderungen stehen, wie bürokratischen Hürden, dem Fehlen eines gemeinsamen Verständnisses zur Kreislaufwirtschaft sowie universellen Indikatoren, durch die der Fortschritt hin zu einer OSCE messbar gemacht werden kann. Die Multi-Level-Perspektive verdeutlicht, dass die Transformation sowohl durch Unterstützung aus der Nische als auch durch strukturelle Veränderungen im Regime gefördert werden kann. Die Studie unterstreicht, dass die OSCE-Implementierung eine breite Einbeziehung verschiedenster Akteursgruppen erfordert, um systemische Barrieren zu überwinden und eine nachhaltige Governance zu etablieren. Schlüsselkompetenzen dafür sind strategisches Handeln, Innovationsfähigkeit und Kommunikationsfähigkeit. Förderlich ist außerdem, bürokratische Prozesse zu digitalisieren und zu vereinfachen, um den administrativen Aufwand zu reduzieren.

**Stichworte** – Circular Economy, lokale Produktion, Fab City, Transition Governance, Transformationspioniere

## NOMENKLATUR

FCI	Fab City Initiative
OSCE	Open-Source-basierte Kreislaufwirtschaft
TP	Transformationspioniere

## I. EINLEITUNG

Die Fab-City-Bewegung sieht sich als eine Antwort auf die gegenwärtigen gesellschaftlichen Herausforderungen, wie sozialer Ungleichheit, die begrenzte Verfügbarkeit und die Klimaschädlichkeit fossiler Rohstoffe, die Abhängigkeiten von globalisierten Lieferketten und das Anwachsen der städtischen

Bevölkerung, was mit einem erhöhten Ressourcenbedarf einhergeht. Sie zielt darauf ab, ein Modell für eine nachhaltige und kreislauffähige Stadt zu entwickeln [1]. Die Fab-City-Bewegung ist zu einem globalen Netzwerk von 49 Städten und Regionen angewachsen, die eine nachhaltige lokale Produktion fördern wollen [2]. Im Fokus der Fab City Initiative (FCI) Hamburg steht daher die Transformation von einer linearen Wirtschaft zu einer datengestützten Kreislaufwirtschaft [3], [4]. Ein zentrales Merkmal der FCI Hamburg ist der Einsatz digitaler Fertigungstechniken, wie 3D-Druckern, Lasercuttern und CNC-Fräsen, mit denen eine kleinteilige Produktion von Konsumgütern des alltäglichen Bedarfs direkt vor Ort stattfinden kann [5]. Ihr Ziel ist es, eine Open-Source-basierte Kreislaufwirtschaft (OSCE) auf städtischer Ebene zu etablieren, in der (fast) alle Konsumgüter vor Ort produziert werden können [6], [7].

Die Transformationsforschung geht davon aus, dass eine gesellschaftliche Transformationen zwar nicht direkt steuerbar ist, jedoch kann sie durch eine transformative Governance beeinflusst, gefördert und beschleunigt werden [8], [9], [10]. Modellen einer nachhaltigen Transition [11], [12], [13], [14] folgend, sind wesentliche Treiber einer solchen umfassenden Transformation die Nischenakteure, die durch Innovationen Impulse in das bestehende sozio-technische System hineingeben [11]. In Anlehnung an den WBGU [15] werden sie als Pioniere des Wandels und hier im Weiteren als Transformationspioniere (TP) bezeichnet. TP tragen durch ihr Handeln dazu bei, Innovationen voranzutreiben und neue Wege für eine nachhaltige Entwicklung zu erschließen [12], [16]. Sie verfügen über vielseitige Kompetenzen und Fähigkeiten, die es ihnen ermöglichen, sich mit weiteren Akteuren zusammenzuschließen [15]. Als Treiber zukünftiger gesellschaftlicher Praktiken erleben sie in ihrer Tätigkeit aber auch Hemmnisse und Hürden und sind im besonderen Maße mit Widerständen konfrontiert, die etwa durch die Pfadabhängigkeit bestehender Normen und Werte im Regime manifestiert sind [12].

Die vorliegende Studie untersucht die Rolle und den Handlungsspielraum von Transformationspionieren bei der Umsetzung einer OSCE in Hamburg. Die Forschungsfrage zielt darauf, die förderlichen und hinderlichen Faktoren ausfindig zu machen, welche die Transformationspioniere betreffen. Dazu wurden qualitativ-interpretative Interviews [17] mit zuvor identifizierten Transformationspionieren aus dem Umfeld der Fab City-Initiative geführt, darunter Akteure aus der organisierten Zivilgesellschaft, Verwaltung und Politik, Wirtschaft

und Wissenschaft [18], [19]. In den Interviews wurde gefragt, wie die Pioniere ihre Rolle bei der Förderung einer OSCE wahrnehmen, welche Bedingungen, Kompetenzen und Instrumente [20] ihnen zur Verfügung stehen und welche Faktoren sie als förderlich oder hinderlich einschätzen [21]. Unter Einbezug des theoriegeleiteten Modells der Multi-Level-Perspektive [11], [22] konnten für die FCI Hamburg umfassende Erkenntnisse zu den Rahmenbedingungen der OSCE, der Zusammenarbeit der Akteure und dem eigenen Handlungsspielraum der TP abgeleitet werden.

## II. FRAMEWORK: DIE MULTI-LEVEL-PERSPEKTIVE UND DIE ROLLE VON TRANSFORMATIONSPIONIEREN

Die Forschung im Bereich der Nachhaltigkeit hat zunehmend deutlich gemacht, dass die gesellschaftlichen Herausforderungen nicht alleine durch inkrementelle Verbesserungen und technologische Korrekturen bewältigt werden können [18], [21], [23], [24], [25]. Vielmehr erfordern sie einen transformativen Wandel, der die zugrundeliegenden Probleme systemisch angeht. Die FCI Hamburg sieht daher ihre Aufgabe darin, eine umfassende Transformation voranzutreiben, die sowohl technologische Innovationen fördert als auch soziale, wirtschaftliche und strukturelle Veränderungen anstößt und dabei verschiedenste Akteursgruppen mit einbezieht [1].

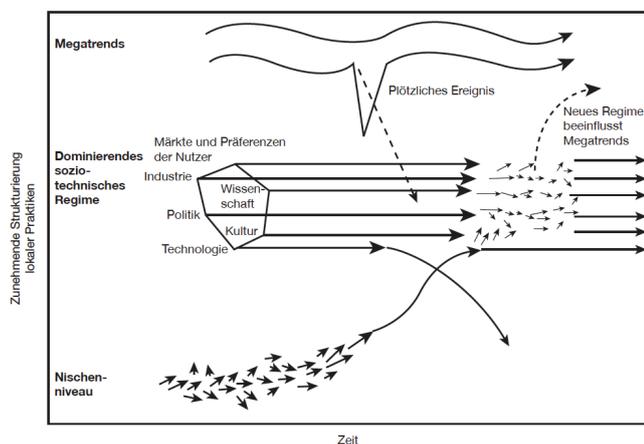


ABBILDUNG 1: WBGU HAUPTGUTACHTEN "WELT IM WANDEL" 2011, ABB. 3.4.1, V. GRIN, ROTMANS & SCHOT 2010, ADAPTIERT V. GEELS & SCHOT 2010

Für Transformationsprozesse beschreibt die Multi-Level-Perspektive (MLP) [11] drei Ebenen, die bei einem Systemwandel beteiligt sind. Im Zentrum des Wandels steht die Ebene des Regimes. Sie beschreibt die gegenwärtige Richtung der Politik, technologischer Entwicklungen und die Ausrichtung von Märkten. Das Regime wird bestimmt durch bestehende Normen, Werte und Gesetze und ist daher u. a. von Pfadabhängigkeiten, Lock-in-Strukturen und Routinen geprägt. Um strukturelle Veränderungen zu ermöglichen, müssen diese überwunden werden [12], [23], [26]. Ein transformativer Wandel geschieht jedoch selten durch einen singulären Auslöser, sondern resultiert aus einem Zusammenspiel verschiedener Faktoren, die sowohl von oben („Landscape“) als auch von unten („Nischen“) auf das Regime einwirken [11]. Ein zentraler Akteur in diesem Prozess sind die Nischenakteure [11], oder eben Transformationspioniere [15]. Sie stoßen Innovationen an, die als Nischeninnovationen beginnen und Einfluss auf die bestehenden Strukturen des Regimes nehmen können.

Transformationspioniere sind Individuen oder kleinere Gruppen, „die mit ihrem Handeln in Richtung eines gesamtgesellschaftlichen Wandels zur Nachhaltigkeit wirken“ [21]. In der FCI Hamburg sind sie in offenen Werkstätten, aber auch in Verbänden oder der Politik zu finden. Sie agieren wert- und zielorientiert und tragen dazu bei, Innovationen voranzutreiben und neue Wege für eine nachhaltige Entwicklung zu erschließen [15], [27], [28]. Vor allem aber sind Transformationspioniere in fast allen Akteursgruppen zu finden, die für die Umsetzung einer OSCE in Hamburg relevant sind. Angelehnt an Heyen et al. [29] werden für diese Untersuchung sechs grundlegende Akteursgruppen entsprechend ihrer Handlungen definiert, die an einer gesellschaftlichen Transformation beteiligt sind: (1) Wissenschaft und Forschung, (2) Politik und Verwaltung, (3) organisierte Zivilgesellschaft, (4) Medien und Presse, (5) Wirtschaft und Industrie und (6) Intermediäre (Zusammenschlüsse verschiedener Sektoren und Handlungsfelder).

Für eine umfassende Transformation ist es jedoch entscheidend, dass die Innovationen aus der Nische in den Mainstream des Regimes gelangen und sich dort verfestigen [16], [24]. Dieser Übergang wird durch Lernprozesse, Netzwerkbildung, Empowerment und gezielte Förderungsstrategien unterstützt [21]. Wenn die Nischen eine kritische Masse erreichen, können sie auf das Regime einwirken und zu Veränderungen beitragen [13]. Dabei spielen auch unterstützende Rahmenbedingungen eine wichtige Rolle, wie politische Vereinbarungen auf übergeordneter Ebene, etwa die EU-Richtlinien oder die Klimapolitik der UN, die positiv auf Veränderungsprozesse einwirken.[13] Darüber hinaus kann ein plötzlicher Schock von außen, wie eine Finanzkrise, eine Pandemie oder eine Umweltkatastrophe, ein Gelegenheitsfenster öffnen und den Wandel beschleunigen [18]. In solchen Momenten können neue Wege und Lösungen entstehen, die zu nachhaltigen Veränderungen führen [13], [24].

## III. METHODE: QUALITATIVE FALLSTUDIENANALYSE

Anhand von Interviews mit den Lab-Manager:innen der Fab City Initiative und dem Mapping von Akteurs- und Interessenkonstellationen [45] wurden Personen identifiziert, die sich als besonders transformativ herausgestellt haben und so als Transformationspioniere der OSCE benannt werden konnten.

TABELLE I: Transformationspioniere einer OSCE in Hamburg

Organisation	Akteursgruppe	Handlungsebene
BUKEA	Verwaltung/ Politik	Lokal
BWI	Verwaltung/ Politik	Lokal
Fab City Initiative	Intermediäre	Lokal
House of All	Wirtschaft	Lokal
Insel e. V.	Organisierte Zivilgesellschaft	Lokal
Handwerkskammer	Intermediäre	Lokal
HiiCCE	Forschung/ Wissenschaft	Lokal
HSU	Forschung/ Wissenschaft	Lokal
Anstiftung	Organisierte Zivilgesellschaft	National
OSEG	Organisierte Zivilgesellschaft	National

Im Umfeld der FCI wurden Transformationspioniere in vielfältigen Bereichen und Akteursgruppen identifiziert. Sie sind in der Politik und Verwaltung aktiv, entwickeln in der Wissenschaft und Forschung innovative Lösungen, treten in der organisierten Zivilgesellschaft als Impulsgeber auf und bilden als Intermediäre Schnittstellen zwischen verschiedenen Akteursgruppen [30]. Für diese Untersuchung liegt der Fokus auf der lokalen und nationalen Ebene, da im Kontext der Transformation in Hamburg internationale Akteure zu lokalen Herausforderungen wenig beitragen können. Bezogen auf die Akteursgruppen ist auffällig, dass mit Blick auf die oben definierten sechs Gruppen in dieser Untersuchung Akteure aus dem Bereich Medien und Presse fehlen und damit eine Gruppe, die maßgeblich öffentliche Diskurse beeinflussen kann.

Die vorliegende Studie basiert auf einer qualitativen Fallstudienanalyse. Insgesamt wurden zehn leitfadengestützte Interviews mit Expert:innen aus der organisierten Zivilgesellschaft, Wissenschaft und Forschung, Verwaltung und Politik sowie Wirtschaft und Intermediäre durchgeführt. Die Interviews fanden zwischen April und Juni 2024 statt. In den Interviews wurde gefragt, wie die Akteure ihre Rolle bei der Umsetzung einer OSCE wahrnehmen: Welche geopolitischen Entwicklungen sie für wichtig erachten, welche Potenziale oder auch Hürden sie darin für ihr Wirken sehen [15], [31], welche Kompetenzen und Instrumente ihnen zur Verfügung stehen [20] und welche Faktoren sie insgesamt als förderlich, aber auch als Hürden für ihre Arbeit einschätzen [21]. Für die Auswertung der Interviews wurde die Qualitative Inhaltsanalyse nach Mayring und Frenzl verwendet [17].

Das Ziel dieser Untersuchung ist eine vertiefte Analyse der Transformationspioniere, um relevante Faktoren im Transformationsprozess zu identifizieren. Dabei soll herausgefunden werden, welche Faktoren sich förderlich und welche hinderlich auf die Umsetzung einer OSCE in Hamburg auswirken. Unter Einbezug der Multi-Level-Perspektive [11] konnten Erkenntnisse zu den Rahmenbedingungen (A), Zusammenarbeit mit Organisationen im Regime (B) und dem eigenen Handlungsspielraum der TP (C) abgeleitet werden.

#### IV. ANALYSE: FÖRDERLICHE UND HINDERLICHE FAKTOREN EINER OSCE IN HAMBURG

Generell benennen die TP sehr ähnliche Faktoren, die für ihre Arbeit an einer OSCE relevant sind. Bei der Bewertung der Faktoren zeigen sich jedoch sektorale Unterschiede der Akteursgruppen.

##### A. Rahmenbedingungen: Faktoren der geopolitischen Landschaft

Die Rahmenbedingungen, die auf die Umsetzung einer OSCE einwirken, werden von den TP in allen Akteursgruppen ähnlich eingeschätzt. Globale Produktions- und Lieferketten werden als ineffizient und anfällig für Störungen angesehen, während soziale und ökologische Krisen das öffentliche Bewusstsein für die Notwendigkeit einer nachhaltigen Transformation geschärft haben. Formale Regulierungen auf EU-Ebene unterstützen diesen Bewusstseinswandel und schaffen eine breitere Unterstützung für die Kreislaufwirtschaft.

###### 1) Globale Produktion

Die global verzweigte Produktion und der intensive Handel werden als nicht mehr zeitgemäß angesehen. Externe Schocks wie die Covid-19-Pandemie, Erdbeben oder der russische Angriffskrieg auf die Ukraine zeigen die Anfälligkeit

globaler Lieferketten. Es wird darauf hingewiesen, dass im Zuge der weltweiten Arbeitsteilung ein Großteil des Expertenwissens in das Ausland verlagert wird und auch Recyclingprozesse außerhalb Hamburgs stattfinden. Die Notwendigkeit, lokale Produktionskapazitäten und Fachwissen zu stärken, wird daher betont. Die Makerszene in Hamburg wird als positives Beispiel genannt, da während der Covid-19-Pandemie medizinische Ausrüstung kurzfristig und lokal bereitgestellt werden konnte.

###### 2) Ökologische Krisen und gesellschaftliches Bewusstsein

Die Pandemie und Lieferengpässe haben die hohe Fragilität des globalisierten Systems offengelegt. Teilweise wird argumentiert, dass die Gesellschaft über ihre Verhältnisse lebt und dass die hohen CO<sub>2</sub>-Emissionen und der Ressourcenverbrauch durch Importe aus dem Ausland nicht tragbar seien. Trotz der negativen Auswirkungen von Krisen haben diese zugleich zu einer höheren Akzeptanz für die Kreislaufwirtschaft geführt. Die Müllproblematik, insbesondere durch die Fashion-Industrie und die Verwendung schwer rezyklierbarer Kunststoffe, sowie die Belastung der Umwelt durch Chemikalien werden als dringende Probleme erkannt. Es zeigt sich ein wachsendes Bewusstsein für die Notwendigkeit, Gegenstände zu reparieren und langlebige Produkte zu bevorzugen.

###### 3) Gesetze und regulatorische Entwicklungen

Das Recht auf Reparatur, das Lieferkettengesetz und die Pflicht zur Nachhaltigkeitsberichterstattung werden als Beispiele für einen formalisierten Bewusstseinswandel auf EU-Ebene genannt. Diese Regulierungen erhöhen den Druck auf die Politik, nationale Strategien zur Kreislaufwirtschaft zu formulieren. In Hamburg wurde etwa 2023 eine Mehrwegangebotspflicht eingeführt, welche die Gastronomie verpflichtet, für Außerhausbestellungen Mehrweggeschirr bereitzustellen. Die Rückverlagerung der Produktion in die Stadt steht jedoch unter erschwerten Genehmigungsbedingungen. So werden insbesondere von Akteuren aus der Wirtschaft und der organisierten Zivilgesellschaft hohe Mietpreise und Raumnot als Hindernisse benannt sowie eine Zunahme des administrativen Aufwands. Sobald eine Fläche anders verwendet wird als vertraglich festgeschrieben, muss etwa eine Flächenumnutzung beantragt werden. Eine Ausweisung als urbane Quartiere, die eine Mischnutzung aus Wohnen und Gewerbe zulassen, wird mehrfach als Vorschlag genannt. Die Akteure in Politik und Verwaltung sehen in den Regelungen auch eine wichtige Rahmensetzung, um soziale und ökologische Anforderungen zu berücksichtigen.

##### B. Faktoren in der Zusammenarbeit mit dem Regime

Bezogen auf die Zusammenarbeit mit anderen Organisationen nennen die TP aus den verschiedenen Akteursgruppen ähnliche Faktoren, die für die Umsetzung einer OSCE von Bedeutung sind. Während Intransparenz, interdisziplinäre Zusammenarbeit, Konkurrenz, langwierige Prozesse und Zielkonflikte Herausforderungen darstellen, zeigen sich positive Ansätze für die Zusammenarbeit, wenn es beispielsweise gemeinsame Interessen und Strategien gibt oder einzelne unterstützende Personen einen Anschluss an das Regime ermöglichen.

###### 1) Intransparenz

Ein wiederkehrendes Thema in den Interviews war die Intransparenz innerhalb der bestehenden Strukturen. Hauptsächlich wurde die verfügbare Datenlage bemängelt, insbesondere im Bereich von Materialströmen und Produktionsprozessen in

Hamburg. Durch Patente geschützte Baupläne oder Betriebsgeheimnisse in der Industrie erschweren zusätzlich die Reparatur von Produkten und führen zu geplanter Obsoleszenz. Gerade im Hinblick auf wichtige Industriezweige, wie das Hamburg Aviation Luftfahrtcluster der Metropolregion, scheint die Umsetzung einer OSCE viele Kompromisse zu erfordern. Auch fehlt den Kooperationspartnern der FCI ein gemeinsames Verständnis von Kreislaufwirtschaft. Das fehlende Leitbild erschwert eine zielgerichtete Verwendung von finanziellen Fördermitteln. Etablierte Labs scheinen mehr Mitsprache zu haben als neue Initiativen. Die Kommunikation der Behörden stellt sich aus Sicht der TP teils intransparent dar, so dass Informationen nur selektiv nach außen gegeben werden und für die TP verfügbar sind.

## 2) *Interdisziplinäre Zusammenarbeit*

Die interdisziplinäre Zusammenarbeit stellt eine große Herausforderung dar. Von allen lokalen Akteuren wird bemängelt, dass es an einem regelmäßigen Austausch fehlt, insbesondere der Gründungsmitglieder der FCI. Schwierigkeiten bestehen vor allem an den Schnittstellen zwischen verschiedenen Organisationsformen: So dürfen staatliche Institutionen wie Universitäten keine gebrauchten Fabrikationsgeräte an offene Werkstätten verkaufen oder verschenken, und traditionelle Handwerksbetriebe reiben sich an den Arbeitsweisen der bottom-up-Initiativen der offenen Werkstätten und Fablabs. Repair-Cafés werden kritisiert, da sie rechtlich nicht abgesichert seien, zudem zeigen nur wenige Handwerksbetriebe Interesse an den Workshop-Formaten. Trotz dieser Herausforderungen gibt es auch positive Beispiele: So unterstützen einzelne Personen in Behörden die Initiativen und fördern den Anschluss an die Strukturen und Prozesse des Regimes. In einigen Branchen, wie im Tischlereihandwerk, ist die Bereitschaft zur interdisziplinären Zusammenarbeit stärker ausgeprägt.

## 3) *Konkurrenz*

In Hamburg laufen verschiedene Initiativen und Projekte zur Kreislaufwirtschaft parallel, was zu interner Konkurrenz führt. Die Stadt fördert beispielsweise die lokale Fertigung im Textilbereich, baut aber parallel zum „Open Lab Textil“ der FCI ein eigenes „Future Fashion Lab“ auf. Dies wird auch als Grund für die unzureichende Vernetzung der Akteure gesehen, da sie bei finanzieller Förderung im Wettbewerb stehen. Zusätzlich befürchten Handwerksbetriebe, dass der Selbstbau von Produkten in offenen Werkstätten eine zusätzliche Konkurrenz zum traditionellen Handwerk schafft. Trotzdem gibt es auch positive Entwicklungen: Durch ein gemeinsames Interesse an der Reduktion von Plastikabfällen, die in Hamburg auf besonderes zivilgesellschaftliches und politisches, aber auch unternehmerisches Interesse trifft, konnte eine Zusammenarbeit zwischen zunächst konkurrierenden Akteuren, in diesem Fall Betrieben (marktorientiert) und offener Zivilgesellschaft (ehrenamtlich), aufgebaut werden.

## 4) *Langwierige Prozesse*

Die lange Dauer von Organisations- und Verwaltungsprozessen stellt eine weitere Herausforderung für die Zusammenarbeit der Akteure dar. Hier kommt erschwerend Hamburgs politische Stellung als Stadtstaat hinzu, wodurch die Behörden den Ministerien in anderen Bundesländern entsprechen. So dauert es oft lange, bis etwa Kooperationsverträge mit öffentlichen Trägern abgeschlossen sind (bis zu einem halben Jahr), was bei kurzen Laufzeiten geförderter Projekte von nur wenigen Jahren oder Monaten zu erschwerten Arbeitsabläufen und Frustration führen kann. Geplante Projekte können dadurch

stellenweise nicht realisiert werden. Als positiv zeigt sich dagegen der Einbezug von Verbänden, wie der Handwerkskammer, und das Aufsetzen gemeinsamer Positionspapiere. Gleichzeitig haben die langwierigen Prozesse durchaus den Vorteil, dass Ansätze der OSCE in die Strategieentwicklung der Verbände einfließen können. Dies trägt dazu bei, dass eine langfristige Transformation in Richtung Kreislaufwirtschaft angestoßen werden kann.

## 5) *Konflikte bei Zielen und Interessen*

Ein wesentliches Hindernis für die Umsetzung einer OSCE ist das Fehlen einer gemeinsamen Zielvorstellung, beziehungsweise einer gemeinsamen Strategie zur Erreichung einer Kreislaufwirtschaft auf lokaler Ebene. Obwohl die FCI das generelle Ziel formuliert hat, Hamburg bis 2054 in Richtung Kreislaufwirtschaft zu transformieren, gibt es weder unter den Akteuren der FCI noch bei den Akteuren in Politik und Verwaltung ein gemeinsames Verständnis, was im Einzelnen mit Kreislaufwirtschaft gemeint ist und wie der Umsetzungsprozess aussehen kann. Gerade durch den Hamburger Hafen und das Aviation-Cluster zeichnet sich die Stadt als globaler Logistikstandort aus, so dass eine lokale Kreislaufwirtschaft der bisherigen Entwicklung eher entgegenläuft. Worin sich jedoch alle Akteure einig sind, ist das Fehlen eindeutiger und allgemeingültiger Indikatoren und Standards, durch die der Fortschritt hin zu einer OSCE messbar gemacht werden kann. Im Gegensatz zum Klimaschutz, wo CO<sub>2</sub>-Äquivalente als Maßstab dienen, fehlen solche Standards in der Kreislaufwirtschaft. Dies führt zu Konflikten bei den Zielen und Interessen der beteiligten Akteure und erschwert eine zielgerichtete Nutzung der verfügbaren Ressourcen.

## C. *Faktoren für Handlungsspielräume innerhalb der Nische*

Bei den Faktoren, welche die TP im eigenen Handlungsspielraum erleben, gibt es ebenfalls ähnliche Einschätzungen der Akteursgruppen. Auf der einen Seite verfügen die Transformationspioniere über ein breites Fachwissen, eine hohe intrinsische Motivation und klare Visionen für eine nachhaltige Zukunft von Wirtschaft und Gesellschaft. Ihre Netzwerke, ihre Agilität und ihre Vermittlungsfähigkeiten ermöglichen es ihnen, Innovationen erfolgreich voranzutreiben und verschiedene Interessengruppen zu mobilisieren. Jedoch stehen sie auch vor Herausforderungen, etwa einem hohen bürokratischem Aufwand, fehlenden personellen und finanziellen Ressourcen sowie einem Mangel an Erfahrung in Bezug auf eine OSCE.

### 1) *Fachwissen*

Die Akteure der offenen Zivilgesellschaft und der Wirtschaft verfügen über ein breites Fachwissen, insbesondere im Bereich der technologischen Entwicklung und dem Umgang mit digitalen Fertigungsmaschinen. Sie haben umfassende Erfahrungen im Recycling, der Arbeit mit Wertstoffhöfen und der Förderung von Mehrwegsystemen gesammelt. Zudem besitzen sie Know-how im Bereich Circular Design, was ihnen ermöglicht, innovative Lösungen für nachhaltige Produkte und Dienstleistungen zu entwickeln. Auch das Wissen von Handwerksbetrieben über Produktions- und Bauprozesse trägt zur Umsetzung von Projekten bei und kann insbesondere beim Rückbau zum Einsatz kommen. Trotz dieser Stärken gibt es noch Bereiche, in denen es etwa in Verwaltung und Politik an Expertise mangelt, beispielsweise zu Kreislaufwirtschaft und Materialströmen, weshalb die Behörde für Umwelt, Klima, Energie und Agrarwirtschaft (BUKEA) etwa ein Referat für

Kreislaufwirtschaft eingerichtet hat. Die kontinuierliche Weiterbildung wird daher als ein wesentlicher Bestandteil für die Umsetzung der OSCE gesehen.

### 2) *Persönliche und strategische Eigenschaften*

Die Transformationspioniere sind durch ihre hohe intrinsische Motivation charakterisiert. Sie sind agil und können sich schnell an veränderte Situationen anpassen. Besonders häufig wurden Mut und Risikobereitschaft als wesentliche Eigenschaften genannt, die im Transformationsprozess wichtig sind. Diese Eigenschaften ermöglicht es den Akteuren, innovative Wege zu beschreiten und komplexe Herausforderungen anzugehen. So wurde etwa eine temporäre Micro-Factory im Jupiter-Haus eingerichtet, an der die Idee lokaler, digitalgestützter Fertigung erprobt und veranschaulicht wird. Eine weitere Voraussetzung sehen die TP daher in der Vermittlungskompetenz und Empathie. Sie sind in der Lage, ihre Visionen auf verschiedene Zielgruppen anzupassen und so andere Menschen wie auch Behörden zu sensibilisieren und für die Anliegen der OSCE zu gewinnen. Dennoch müssen sich insbesondere die Akteure im öffentlichen Dienst gelegentlich einem Realitätsabgleich unterziehen und einen Ausgleich zwischen ihrem persönlichen Engagement und der Praktikabilität in Bezug auf die vorhandenen Regimestrukturen finden.

### 3) *Vision und Gestaltungswille*

Die TP aller Akteursgruppen verfolgen Visionen für eine nachhaltige Zukunft von Wirtschaft und Gesellschaft, insbesondere in Bezug auf den Bereich der Kreislaufwirtschaft. Sie denken zukunftsorientiert und sind bereit, für ihre Werte einzustehen. Die Visionen reichen hier von einem Circular Economy Council für Hamburg oder einem Reparaturbonus bis hin zu Kleidung als Grundrecht, Labs mit freiem, von der sozialen Herkunft unabhängigen Zugang, in jedem Stadtteil oder eine Neuausrichtung von Wertstoffhöfen zu Zwischenlagern und Knotenpunkten von Stoffkreisläufen. Auch die TP in Politik und Verwaltung betonen, dass sie über einen gewissen Gestaltungsfreiraum verfügen, um zukunftsfähige Ideen umzusetzen. Beispielsweise wurden Veranstaltungen, die von der Stadt ausgerichtet wurden, auf Mehrweg umgestellt. Hier könne die Stadt als Vorbild fungieren. Außerdem wird die FCI Hamburg seit Beginn durch die Innovationsförderung der BWI unterstützt.

### 4) *Netzwerken*

Die Transformationspioniere sind aktiv in verschiedenen Netzwerken involviert. Sie betreiben eine umfangreiche Öffentlichkeitsarbeit, die zu einem positiven Rücklauf von Interessierten führt. Die Vernetzung ist auch für Verbände von großer Bedeutung. Durch einen gut gepflegten Datenbestand aller Handwerksbetriebe in Hamburg können diese gezielt informiert werden, was zur Verbreitung und Akzeptanz der Ziele der FCI beitragen kann. TP besitzen auch Vermittlungsfähigkeiten, um Brücken zwischen verschiedenen Interessengruppen zu bauen und so die Zusammenarbeit zu stärken. Jedoch ist etwa im Bereich Handwerk die Bereitschaft zur Zusammenarbeit noch verhalten.

### 5) *Bürokratischer und administrativer Aufwand*

Die Erstellung von Berichten sowie das Beantragen und die Abrechnung von Förderungen sind zeitintensive Aufgaben. Sie binden die personellen und zeitlichen Ressourcen der

TP. Hierin erleben die TP eine der größten Hürden in ihrem Arbeitsalltag, wenn sie auf öffentliche Förderung angewiesen sind. Auch von neuen EU-Richtlinien, wie dem „Right to Repair“<sup>1</sup> oder dem „Digital Product Passport“<sup>2</sup> gehen erhöhte administrative Anforderungen aus, um die notwendigen Informationen bereitstellen zu können. Dennoch sehen die TP in der Umweltpolitik der EU eine grundsätzliche und wichtige Weichenstellung für einen Systemwechsel. Akteure aus Politik und Verwaltung betonen, dass das Problem nicht die Verordnungen oder Vorgaben an sich sind, sondern der teils schwerfällige Regelvollzug, der den bürokratischen Aufwand weiter erhöht. In Hamburg, wie auch in ganz Deutschland, gäbe es einen großen Nachholbedarf in Bezug auf die Digitalisierung administrativer Prozesse.

### 6) *Ressourcen*

Die TP benötigen eine Vielzahl von Ressourcen, um ihre Arbeit erfolgreich umzusetzen. Dazu gehören staatliche Förderungen, die jedoch, wie oben beschrieben, oft mit einem erheblichen bürokratischen Aufwand verbunden sind. Die Wartung von Maschinen, die Koordination der Projekte sowie die Kommunikation mit Projektpartnern oder in die Öffentlichkeit sind ebenfalls zeit- und ressourcenintensiv. Geeignete Räumlichkeiten für ihre Projekte zu finden, kann eine weitere Herausforderung darstellen. Gleichzeitig stellen insbesondere die Labs Ressourcen, wie Maschinen, Materialien und Wissen, zur Verfügung. Förderlich ist, wenn die TP auf die Unterstützung durch Verbände und die Verwaltung zählen können, sofern diese ein generelles Interesse daran zeigen, innovative Projekte und Kooperationen zu fördern.

## V. ERGEBNISSE UND DISKUSSION

Während es zwischen den verschiedenen Akteursgruppen keine wesentlichen Unterschiede bei der Identifikation der relevanten Faktoren gab, wurden diese jedoch von diesen unterschiedlich bewertet. Ein Großteil der Nischenexperimente, durch die eine gesellschaftliche Transformation angestoßen werden soll, ist auf eine öffentliche Förderung angewiesen und damit meist temporär begrenzt. Dabei wird der administrative Aufwand als Restriktion deutlich, da zeitliche und personelle Ressourcen gebunden werden. Während die Mehrheit der Akteure den bürokratischen Aufwand und die langwierigen Prozesse in der Zusammenarbeit als Belastung empfindet, liegt darin für die Akteure in Politik und Verwaltung die Gewähr für solide Prozesse und eine nachhaltige Transformation. Die Akzeptanz innerhalb des Regimes ist eine Voraussetzung dafür, dass Innovationen langfristig implementiert und eine Transformation tatsächlich erfolgen kann [32], [33], [34]. Dennoch weisen auch Politik und Verwaltung darauf hin, dass viele Prozesse durch den Ausbau der Digitalisierung vereinfacht und beschleunigt werden könnten.

Die Untersuchung deutet weiter darauf hin, dass die Einbeziehung von Akteuren aus allen Akteursgruppen für die Entwicklung einer umfassenden transformativen Governance von entscheidender Bedeutung ist [19], [35], [36], [37]. Während sich frühere Studien oft auf zivilgesellschaftliche Nischenakteure konzentriert haben [12], [21], [24], [38], unterstreichen die Ergebnisse dieser Fallstudie das Potenzial von Transformationspionieren in allen Akteursgruppen. Die befragten Akteure verfügen über transformative Schlüsselkompetenzen,

<sup>1</sup> <https://www.europarl.europa.eu/news/en/press-room/20240419IPR20590/right-to-repair-making-repair-easier-and-more-appealing-to-consumers>, abgerufen 14.05.2024.

<sup>2</sup> [https://commission.europa.eu/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/sustainable-products/ecodesign-sustainable-products-regulation\\_en#the-new-digital-product-passport](https://commission.europa.eu/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/sustainable-products/ecodesign-sustainable-products-regulation_en#the-new-digital-product-passport), abgerufen 14.05.2024.

darunter wert- und zielorientiertes Handeln, intrinsische Motivation, strategische Kompetenz, Innovationsfähigkeit, Gemeinwohlorientierung, Kommunikationsfähigkeit und Fachwissen und nehmen Einfluss auf individuelle Lebensstile und kulturelle Normen [15]. Damit tragen sie in vielfältiger Weise und in ihren ganz unterschiedlichen institutionellen Kontexten dazu bei, dass Aspekte der Kreislaufwirtschaft Teil des gesellschaftlichen Transformationsprozesses werden [39].

Die Vielfalt der Akteure erfordert aber gleichzeitig ein hohes Maß an Koordination und Kooperation. Die Analyse hat verdeutlicht, dass gerade bei der Zusammenarbeit noch Verbesserungspotenzial besteht. Während TP ihren eigenen Handlungsspielraum als weitreichend und einflussreich erleben und auf der persönlichen Ebene eine hohe intrinsische Motivation benennen, wird die interdisziplinäre Zusammenarbeit mehrheitlich als mühsam und schleppend bewertet. Hier wurden insbesondere eine intransparente Kommunikation und eine Konkurrenz zwischen den Akteuren, u. a. aufgrund knapper finanzieller Ressourcen sowie Interessens- und Zielkonflikte genannt. Fehlende Absprachen, unterschiedliche Verständnisse von Kreislaufwirtschaft sowie das Fehlen eines gemeinsamen Leitbildes und einer Strategie erschweren die Zusammenarbeit. Dagegen werden das Vorhandensein von Fördermöglichkeiten und der Anschluss an transformationsbereite und -offene Akteure im Regime als positiv genannt. Letzteres ist eine maßgebliche Voraussetzung, damit die Transformation gelingen kann [16], [31].

Trotz zahlreicher laufender Aktivitäten der Fab City-Akteure in Hamburg gibt es derzeit keinen umfassenden Governance-Rahmen, der die Akteure in einem strukturierten Prozess zusammenführt [4], [8], [40]. Die zahlreichen Nischenexperimente geben wichtige Impulse, wie etwa Kreisläufe von Textilien oder Plastik nachhaltig und lokal umgesetzt [41], [42] und Hamburg zu einem Pionier der Kreislaufwirtschaft werden könnte. Diese Vision wird von Akteuren aus Wirtschaft, Forschung und organisierter Zivilgesellschaft ebenso gestützt wie von Akteuren in Politik und Verwaltung. Insbesondere die Einbindung der verschiedenen politischen und administrativen Akteure und Ebenen in die Fab City Hamburg-Initiative ist aber bislang noch unzureichend [43]. Zwar treten einzelne Behörden in Hamburg bereits als finanzielle Förderer oder Kooperationspartner auf, jedoch konnte die Zusammenarbeit nicht verstetigt werden. Dies führt zu einer unsicheren Planungslage und hindert weniger intrinsisch motivierte Akteure daran, sich auf eine Transformation ihrer Arbeits- und Produktionsabläufe einzulassen.

Weitere Herausforderungen für eine kreislauforientierte, transformative Governance sind Hamburgs Stellung als Stadtstaat, was Transformationsprozesse innerhalb der Verwaltung erschweren kann, sowie als Logistikstandort, der in globale Handelsketten integriert ist. Das Vorgehen und Engagement der Stadt ist daher wesentlich für das Etablieren und Gelingen einer langfristigen Transformation hin zu einer OSCE [44]. Zum einen können die transformationsbereiten Akteure in Politik und Verwaltung unmittelbar Einfluss auf die Normen und Werte im Regime nehmen, maßgebliche Förderungen und Regelungen veranlassen sowie darüber hinaus als Vorbild agieren. Zum anderen kann die Stadt ein kreatives Umfeld für die zahlreichen Initiativen und Nischenexperimente bereitstellen, das Kooperation der verschiedenen Akteure anstelle von Konkurrenz in den Vordergrund rückt.

## VI. AUSBLICK

Im weiteren Verlauf der Studie soll zum einen die Rolle von Politik und Verwaltung näher untersucht werden. Hierfür werden die Interviews mit einem dezidierten Fokus auf deren Einbindung und derzeitige Aktivitäten ausgewertet und weitere ergänzende Befragungen durchgeführt. Zum anderen sollen Maßnahmen für eine transformative Governance ausgearbeitet werden. Dafür werden in einem nächsten Schritt relevante Instrumente identifiziert, die zur Umsetzung einer OSCE in Hamburg zur Verfügung stehen. Diese Instrumente werden mithilfe eines Mappings ausgewertet und diskutiert, wobei auch mögliche Lücken einer transformativen Steuerung aufgedeckt werden sollen. Daraus sollen Handlungsvorschläge für eine transformative Governance einer nachhaltigen und lokalen Kreislaufwirtschaft in Hamburg abgeleitet werden.

## DANKSAGUNG

Diese empirische Fallstudie ist Teil einer größeren Untersuchung zu einer transformativen Governance für die Kreislaufwirtschaft in Hamburg und wird an der HafenCity Universität (HCU) in Kooperation mit der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg (HSU) durchgeführt. Besonderer Dank geht an alle Interviewpartner:innen, die ihre Zeit und Erfahrung für die Studie zur Verfügung gestellt haben. Diese Forschungsarbeit wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] M. Moritz, T. Redlich, S. Buxbaum-Conradi, and J. P. Wulfsberg, Eds., *Global collaboration, local production*. Wiesbaden: Springer Fachmedien Wiesbaden, 2024.
- [2] Fab City Global Initiative, *Sustainable Cities Network: Building a Better Future*. [Online]. Available: <https://fab.city/network-sustainable-cities/> (accessed: Jun. 21 2024).
- [3] S. Buxbaum-Conradi, "What Kind of City is the Fab City?," in *SDG - Forschung, Konzepte, Lösungsansätze zur Nachhaltigkeit, Global collaboration, local production*, M. Moritz, T. Redlich, S. Buxbaum-Conradi, and J. P. Wulfsberg, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2024, pp. 27–41.
- [4] S. Buxbaum-Conradi, J. Koppe, M. Moritz, T. Redlich, and J. P. Wulfsberg, *Fab City Hamburg: A living lab approach to explore new forms of open, distributed manufacturing in an urban context*: Helmut-Schmidt-Universität, 2022.
- [5] M. Moritz, T. Redlich, and S. Buxbaum-Conradi, "Einführung: Wie Fab Cities zu einer nachhaltigen Entwicklung beitragen," in *SDG - Forschung, Konzepte, Lösungsansätze zur Nachhaltigkeit, Global collaboration, local production*, M. Moritz, T. Redlich, S. Buxbaum-Conradi, and J. P. Wulfsberg, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2024, pp. 1–10.
- [6] T. Diez and A. Posada, "The fab and the smart city," in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, Barcelona Spain, 2013, pp. 447–454.
- [7] T. Diez, J. Millard, and Menichinelli, Massimo, Sorivelle, Marie Nicole, "The Fabrication City," in *Innovative Solutions for Creating Sustainable*, Albert, Ed., 2019, pp. 329–356.

- [8] A. Wildhack, J. Koppe, and J. Knieling, "Transitioning to a Fab City: A Governance Perspective," in *SDG - Forschung, Konzepte, Lösungsansätze zur Nachhaltigkeit, Global collaboration, local production*, M. Moritz, T. Redlich, S. Buxbaum-Conradi, and J. P. Wulfsberg, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2024, pp. 71–87.
- [9] O. Treib, H. Bähr, and G. Falkner, "Modes of governance: towards a conceptual clarification," *Journal of European Public Policy*, vol. 14, no. 1, pp. 1–20, 2007, doi: 10.1080/135017606061071406.
- [10] J. Rückert-John and M. Schäfer, *Governance für eine Gesellschaftstransformation*. Wiesbaden: Springer Fachmedien Wiesbaden, 2017.
- [11] F. W. Geels, "Technological transitions as evolutionary re-configuration processes: a multi-level perspective and a case-study," *Research Policy*, vol. 31, 8-9, pp. 1257–1274, 2002, doi: 10.1016/S0048-7333(02)00062-8.
- [12] J. Grin, J. Rotmans, and J. Schot, *Transitions to Sustainable Development. : New Directions in the Study of Long Term Transformative Change*: Taylor & Francis Group, 2010.
- [13] D. LOORBACH, J. M. Wittmayer, H. Shiroyama, J. Fujino, and S. Mizuguchi, *Governance of Urban Sustainability Transitions*. Tokyo: Springer Japan, 2016.
- [14] J. Farla, J. Markard, R. Raven, and L. Coenen, "Sustainability transitions in the making: A closer look at actors, strategies and resources," *Technological Forecasting and Social Change*, vol. 79, no. 6, pp. 991–998, 2012, doi: 10.1016/j.techfore.2012.02.001.
- [15] WBGU, *Welt im Wandel: Gesellschaftsvertrag für eine große Transformation*, 2nd ed. Berlin: WBGU, 2011.
- [16] K. Kristof, *Models of change: Einführung und Verbreitung sozialer Innovationen und gesellschaftlicher Veränderungen in transdisziplinärer Perspektive*. Zürich: Vdf-Hochschulverl., 2010.
- [17] P. Mayring and T. Fenzl, "Qualitative Inhaltsanalyse," in *Handbuch Methoden der empirischen Sozialforschung*, N. Baur and J. Blasius, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2014, pp. 543–556.
- [18] A. Benz, *Governance — Regieren in komplexen Regelsystemen*. Wiesbaden: VS Verlag für Sozialwissenschaften, 2004.
- [19] J. Newig and O. Fritsch, "Environmental governance: participatory, multi-level – and effective?," *Env Pol Gov*, vol. 19, no. 3, pp. 197–214, 2009, doi: 10.1002/eet.509.
- [20] J. Fröhlich, J. Knieling, and T. Kraft, *Informelle Klimawandel-Governance: Instrumente der Information, Beteiligung und Kooperation zur Anpassung an den Klimawandel*, Neopolis working papers, No.15. Hamburg: HCU, Stadtplanung und Regionalentwicklung, 2014.
- [21] T. Engel, K. Klindworth, and J. Knieling, "Transformationspioniere für eine klimafreundliche Stadt," *ÖW*, no. 1, pp. 30–34, 2019, doi: 10.14512/OEW340130.
- [22] K. Hillman, M. Nilsson, A. Rickne, and T. Magnusson, "Fostering sustainable technologies: a framework for analysing the governance of innovation systems," *Sci. and Pub. Pol.*, vol. 38, no. 5, pp. 403–415, 2011, doi: 10.3152/030234211X12960315267499.
- [23] J. Beyer, *Pfadabhängigkeit: Über institutionelle Kontinuität, anfällige Stabilität und fundamentalen Wandel*. Frankfurt: Campus, 2006.
- [24] J. Schot and F. W. Geels, "Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy," *Technology Analysis & Strategic Management*, vol. 20, no. 5, pp. 537–554, 2008, doi: 10.1080/09537320802292651.
- [25] J. Köhler, F.W. Geels, F. Kern, J. Markard, E. Onsongo, A. Wieczorek, F. Alkemade, F. Avelino, A. Bergek, F. Boons, F. Fünfschilling, D. Hess, G. Holtz, "An agenda for sustainability transitions research: State of the art and future directions," *Environmental Innovation and Societal Transitions*, vol. 31, pp. 1–32, 2019, doi: 10.1016/j.eist.2019.01.004.
- [26] W. Kahlenborn, J. Clausen, S. Behrendt, and E. Göll, "3.1 Pfadabhängigkeiten," in *Neue Ökologie, Auf dem Weg zu einer Green Economy*, W. Kahlenborn, J. Clausen, S. Behrendt, and E. Göll, Eds., Bielefeld, Germany: Transcript Verlag, 2019, pp. 21–31.
- [27] T. Redlich, P. Krenz, S.-V. Basmer, S. Buxbaum-Conradi, S. Wulf, and J. P. Wulfsberg, "The Impact of Openness on Value Co-creation in Production Networks," *Procedia CIRP*, vol. 16, pp. 44–49, 2014, doi: 10.1016/j.procir.2014.01.007.
- [28] T. Redlich, M. Moritz, S. Buxbaum-Conradi, P. Krenz, S. Heubischl, and S. V. Basmer-Birkenfeld, "OpenLabs - Collaborative Industrialization with Distributed and Open Source Microfactories," *AMM*, vol. 794, pp. 470–477, 2015, doi: 10.4028/www.scientific.net/AMM.794.470.
- [29] Heyen, D.A., Brohmann, B., Libbe, J., Riechel, R., Trapp, J.H., "Stand der Transformationsforschung unter besonderer Berücksichtigung der kommunalen Ebene," 2018. [Online]. Available: [https://www.bbsr.bund.de/BBSR/DE/forschung/programme/exwost/Forschungsfelder/2017/stadtbau-transformationstrategie/synthesepapier-transformationforschung.pdf?\\_\\_blob=publicationFile&v=6](https://www.bbsr.bund.de/BBSR/DE/forschung/programme/exwost/Forschungsfelder/2017/stadtbau-transformationstrategie/synthesepapier-transformationforschung.pdf?__blob=publicationFile&v=6)
- [30] T. Moss, "Intermediaries and the Governance of Sociotechnical Networks in Transition," *Environ Plan A*, vol. 41, no. 6, pp. 1480–1495, 2009, doi: 10.1068/a4116.
- [31] WBGU, *Der Umzug der Menschheit: die transformative Kraft der Städte*, 2nd ed. Berlin: Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen, 2016.
- [32] J. Wittmayer and K. Hölscher, "Transformationsforschung - Definitionen, Ansätze, Methoden," Dessau-Roßlau 103, Nov. 2017.
- [33] N. Frantzeskaki, K. Hölscher, M. Bach, and F. Avelino, Eds., *Co-creating Sustainable Urban Futures*. Cham: Springer International Publishing, 2018.
- [34] C. J. Termeer, A. Dewulf, and G. R. Biesbroek, "Transformational change: governance interventions for climate change adaptation from a continuous change perspective," *Journal of Environmental Planning and Management*, vol. 60, no. 4, pp. 558–576, 2017, doi: 10.1080/09640568.2016.1168288.
- [35] F. Avelino and J. M. Wittmayer, "Shifting Power Relations in Sustainability Transitions: A Multi-actor Perspective," *Journal of Environmental Policy & Planning*, vol. 18, no. 5, pp. 628–649, 2016, doi: 10.1080/1523908X.2015.1112259.
- [36] L.-B. Fischer and J. Newig, "Importance of Actors and Agency in Sustainability Transitions: A Systematic Exploration of the Literature," *Sustainability*, vol. 8, no. 5, p. 476, 2016, doi: 10.3390/su8050476.
- [37] P. Kivimaa, W. Boon, S. Hyysalo, and L. Klerkx, "Towards a typology of intermediaries in sustainability transitions: A systematic review and a research agenda," *Research Policy*, vol. 48, no. 4, pp. 1062–1075, 2019, doi: 10.1016/j.respol.2018.10.006.
- [38] R. Greer and D. Loorbach, *Governing the transition to a circular economy: Key dynamics, paradoxes, and implications for strategizing*. Rotterdam: Erasmus University Rotterdam, 2022.
- [39] F. Savini, "The economy that runs on waste: accumulation in the circular city," *Journal of Environmental Policy & Planning*, vol. 21, no. 6, pp. 675–691, 2019, doi: 10.1080/1523908X.2019.1670048.
- [40] J. M. Wittmayer and D. LOORBACH, "Governing Transitions in Cities: Fostering Alternative Ideas, Practices, and Social Relations Through Transition Management," in pp. 13–32.
- [41] J. Singh and I. Ordoñez, "Resource recovery from post-consumer waste: important lessons for the upcoming circular

- economy,” *Journal of Cleaner Production*, vol. 134, pp. 342–353, 2016, doi: 10.1016/j.jclepro.2015.12.020.
- [42] P. Smith, J. Baille, and L.-S. McHattie, “Sustainable Design Futures: An open design vision for the circular economy in fashion and textiles,” *The Design Journal*, vol. 20, sup1, S1938-S1947, 2017, doi: 10.1080/14606925.2017.1352712.
- [43] A. Smith, “Translating Sustainabilities between Green Niches and Socio-Technical Regimes,” *Technology Analysis & Strategic Management*, vol. 19, no. 4, pp. 427–450, 2007, doi: 10.1080/09537320701403334.
- [44] A. Kronsell and D. Mukhtar-Landgren, “Experimental governance: the role of municipalities in urban living labs,” *European Planning Studies*, vol. 26, no. 5, pp. 988–1007, 2018, doi: 10.1080/09654313.2018.1435631.
- [45] M. S. Reed, A. Graves, N. Dandy, H. Posthumus, K. Hubacek, J. Morris, C. Prell, C.H. Quinn, L.C. Stringer, “Who’s in and why? A typology of stakeholder analysis methods for natural resource management,” (in eng), *Journal of environmental management*, vol. 90, no. 5, pp. 1933–1949, 2009, doi: 10.1016/j.jenvman.2009.01.001.

# Progress on Computer-Aided and Experimental Determination of Thermophysical Properties of Gas Mixtures Containing Hydrogen

Robert Hellmann, Ru Wang, Ruben Harten, Philipp Marienhagen, Karsten Meier\*

*Institut für Thermodynamik*

*Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg*

Hamburg, Germany

\*karsten.meier@hsu-hh.de

Christian Günz, Mathis Manzel, Christof Gaiser\*

*Physikalisch-Technische Bundesanstalt*

Berlin, Germany

\*christof.gaiser@ptb.de

Sebastian Schulz, Arne Pietsch, Joachim Zipfel, Philip Jaeger\*

*Eurotechnica GmbH*

Bargteheide, Germany

\*philip.jaeger@eurotechnica.de

**Abstract**—In the project H2MIXPROP, highly accurate data for several thermophysical properties of gaseous mixtures containing molecular hydrogen are obtained by state-of-the-art theoretical approaches and experimental methods. Such data are required for many technical applications in the transition of the energy supply system to renewable energy sources, in which hydrogen is expected to play a prominent role. This contribution describes theoretical results for cross second virial coefficients of several binary mixtures, the development and validation of a path integral Monte Carlo code for the simulation of quantum gases, and the current status of the experimental tasks.

**Index Terms**—Burnett method, computer simulation, hydrogen, mixture, virial coefficient, vibrating-wire viscometer, viscosity

## I. INTRODUCTION

In the transition of the energy supply system to renewable energy sources, thermophysical properties of mixtures containing hydrogen ( $H_2$ ) are required in a multitude of technical applications, e.g., in the injection of hydrogen into the natural gas grid or the storage of hydrogen in subsurface caverns. In our project, we investigate mixtures of hydrogen with methane ( $CH_4$ ), nitrogen ( $N_2$ ), oxygen ( $O_2$ ), water ( $H_2O$ ), carbon dioxide ( $CO_2$ ), hydrogen sulfide ( $H_2S$ ), ethane ( $C_2H_6$ ), and propane ( $C_3H_8$ ). We predict virial coefficients, viscosity, thermal conductivity, and diffusion coefficients of hydrogen-rich gaseous mixtures on the basis of quantum-chemical *ab initio* calculations of the intermolecular interaction energies. Moreover, we apply the path integral Monte Carlo technique to calculate thermodynamic properties at high pressures in

the supercritical region and in the liquid region. In order to validate the theoretically determined data, experiments for selected mixtures are carried out with newly developed Burnett instruments at Helmut-Schmidt-Universität (HSU) in Hamburg and Physikalisch-Technische Bundesanstalt (PTB) in Berlin, which complement each other with their operating conditions. Moreover, viscosity measurements are carried out at HSU with two different viscometers. Eurotechnica consults PTB and HSU with regard to laboratory safety for the handling of hydrogen and contributes measurements of the density for selected mixtures. For the use in practical applications, the data determined in the project will be represented by correlating equations. Beyond the data for practical applications, the project will yield valuable insights about the possibilities to replace time-consuming, expensive, and possibly even dangerous experiments (e.g., with hydrogen/oxygen mixtures, so-called oxyhydrogen gas) increasingly by computer simulations in the future. Especially at the high temperatures relevant for combustion processes, experiments are virtually impossible, and computer simulations are the only source for the thermophysical properties at these conditions.

## II. FIRST-PRINCIPLES DETERMINATION OF SECOND VIRIAL COEFFICIENTS

In the group at HSU, pair potentials, which describe the interaction energy between two molecules as a function of their separation and mutual orientation, were developed for the interactions between unlike molecules in all eight binary mixtures and pure oxygen. Interaction energies were computed

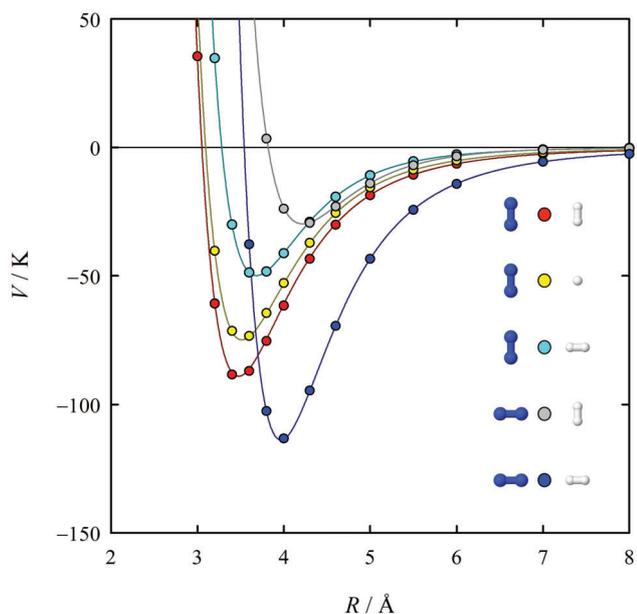


Figure 1.  $\text{N}_2\text{-H}_2$  interaction energy (in kelvin due to reduction by the Boltzmann constant) as a function of the separation (in angstrom) between the centers of mass of the two molecules for five distinct angular configurations. Symbols represent quantum-chemically calculated interaction energies, while solid lines represent the fitted potential function

by means of the supermolecular approach using standard quantum-chemical program packages. These calculations were carried out for a large number of interparticle separations and mutual orientations in order to obtain a complete description of the intermolecular interactions. Suitable mathematical functions, in this case site–site functions, were fitted to the calculated interaction energies. The number parameters that had to be fitted in order to obtain a satisfying fit for each potential function involving  $\text{H}_2$  varied from 46 for the  $\text{N}_2\text{-H}_2$  and  $\text{O}_2\text{-H}_2$  systems to 109 for the  $\text{C}_3\text{H}_8\text{-H}_2$  system. For the  $\text{O}_2\text{-O}_2$  interaction, the three possible spin couplings between the  $\text{O}_2$  molecules yield three distinct pair potentials, which each required 41 fit parameters.

Figure 1 exemplarily depicts the  $\text{N}_2\text{-H}_2$  pair potential in the well region as a function of the separation of the two molecules for different mutual orientations. It can be seen that the fitted potential function reproduces the quantum-chemically calculated interaction energies very accurately.

One of the most important thermodynamic properties of a gas is its second virial coefficient  $B$ , which provides the lowest-order correction to the ideal gas law in the so-called virial expansion:

$$\frac{p}{\rho_m RT} = 1 + B(T)\rho_m + C(T)\rho_m^2 + \dots, \quad (1)$$

where  $p$ ,  $\rho_m$ ,  $R$ ,  $T$ , and  $C$  denote the pressure, the molar density, the molar gas constant, the temperature, and the third virial coefficient, respectively.  $B(T)$  is a property that depends only on interactions between two molecules, with statistical

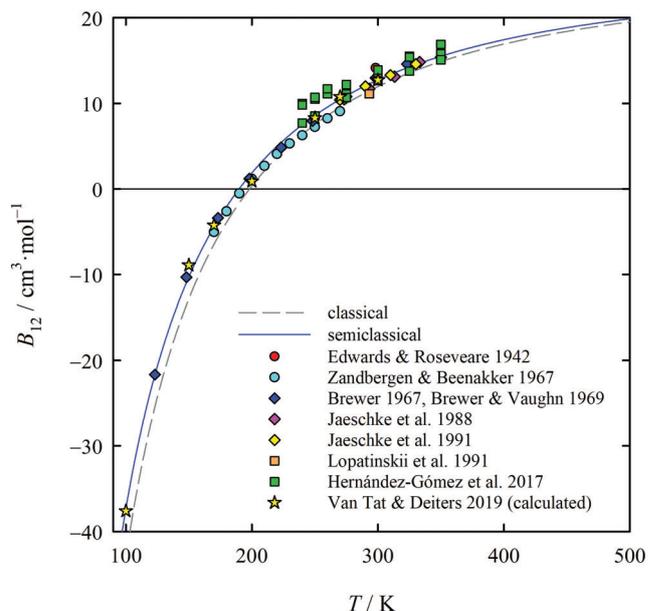


Figure 2. Classically and semiclassically calculated values for the cross second virial coefficient  $B_{12}$  of the  $\text{N}_2\text{-H}_2$  system as a function of temperature. Experimental data and theoretically calculated values from the literature are shown for comparison

thermodynamics providing the link between  $B(T)$  and the pair potential. In a binary gas mixture of components 1 and 2, the second virial coefficient is given by the following exact relation:

$$B(T) = x_1^2 B_{11}(T) + 2x_1 x_2 B_{12}(T) + x_2^2 B_{22}(T), \quad (2)$$

where  $x_1$  and  $x_2$  are the mole fractions,  $B_{11}$  and  $B_{22}$  denote the second virial coefficients of the pure gases, and  $B_{12}$  is the so-called cross second virial coefficient, which is directly related with the unlike-species pair interactions.

We used the Mayer-sampling Monte Carlo method [1] implemented in an in-house code to extract the cross second virial coefficients of the eighth binary systems and the second virial coefficient of pure oxygen from the respective pair potentials both classically and semiclassically using the quadratic Feynman–Hibbs (QFH) [2] approach. Figure 2 exemplarily shows the results for the  $\text{N}_2\text{-H}_2$  system together with experimental and calculated values from the literature. The agreement with the literature data is excellent, but these data only extend to temperatures of at most 350 K, whereas the present values extend to 2000 K. We note here that none of the other seven binary systems investigated in this project has previously been studied as extensively as the  $\text{N}_2\text{-H}_2$  system, so that the data situation is improved even more substantially for these systems by the present calculations.

The results for pure oxygen and for the  $\text{H}_2\text{O-H}_2$  and  $\text{H}_2\text{S-H}_2$  systems are already published [3], [4]. In the next step of this part of the project, we will use kinetic theory to

calculate the low-density viscosities, thermal conductivities, and binary diffusion coefficients of the eight mixture systems.

### III. PATH INTEGRAL MONTE CARLO SIMULATIONS OF QUANTUM GASES

As a preparation for the calculation of thermodynamic properties of supercritical and liquid hydrogen at moderate and high pressures in the extension of the project H2MIXPROP, a code for path integral Monte Carlo (PIMC) simulations was developed. The code is written in C and was optimized with respect to vectorization and parallelization with OpenMP.

Using the methodology for the derivation of expressions for thermodynamic properties in statistical ensembles developed by Lustig [5], we derived equations for the calculation of ten different thermodynamic properties in the canonical and isothermal-isobaric ensembles within the path integral framework. This was done for three different types of estimators: primitive estimators, analytical virial estimators [6] and finite-difference virial estimators [7]. Many of the thermodynamic properties, such as the speed of sound or Joule–Thomson coefficient, can now be calculated in PIMC simulations for the first time.

The intermolecular interactions are described in the simulations by state-of-the-art recent pair and nonadditive three-body potentials for neon [8], [9] and helium [10], [11]. To validate the code, we performed simulations in the supercritical region of both noble gases and compared the results for the speed of sound to the recent data measured by Dietl and Meier in our group [12]. Because the impact of quantum effects increases for lower temperatures, we chose the 80 K isotherm for validation, as this is the lowest temperature measured by Dietl and Meier, which is near the boiling point of the cooling agent nitrogen. As an example, the results for the density  $\rho$  and speed of sound  $w$  of supercritical helium for five state points on the 80 K isotherm are shown in Figure 3 as relative deviations from the equation of state (EoS) of Ortiz Vega et al. [13] as a function of pressure  $p$ . The calculated densities are more precise than the available experimental data and in better agreement with the equation of state. For the speed of sound, our calculated data are in excellent agreement within in their expanded uncertainty of up to 0.15% with the highly accurate experimental data of Dietl, which shows that the derived equations and developed code are correct.

In contrast to experiments, which usually only focus on a single thermodynamic property, we are now able to calculate a range of different properties within a single simulation. It is further possible to calculate properties in regions of the phase diagram where experiments are difficult and costly, e.g., at temperatures below the boiling point of nitrogen. We plan to extend our PIMC code to simulations of hydrogen, which requires the quantum-mechanical treatment of the rotation of molecules. The developed framework would then provide a feasible alternative to experimental techniques at very high

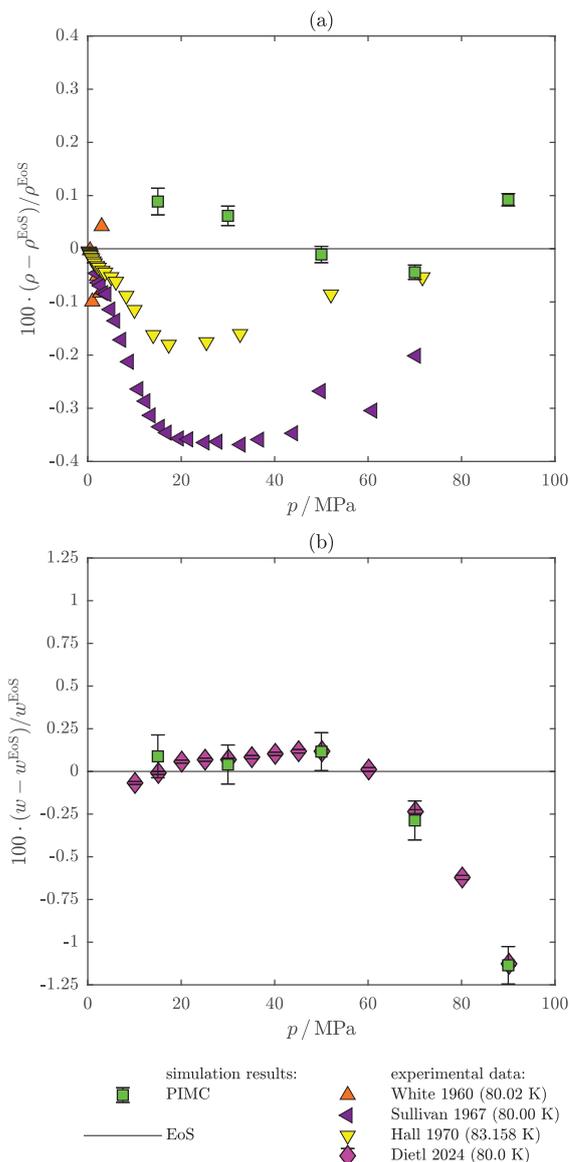


Figure 3. (a) and (b) Relative deviations of PIMC results for the density (top) and speed of sound (bottom) of helium and experimental data from the EoS as a function of pressure at the isotherm 80 K

pressures, where the laboratory safety becomes increasingly difficult.

### IV. VIRIAL COEFFICIENTS FROM THE BURNETT METHOD

The equation of state (EOS) of a gas plays an important role in both scientific and industrial applications. The Burnett technique provides a simple experimental method to determine the equation of state, in which neither the mass nor the volume of gas has to be measured.

The principle of a Burnett apparatus is sketched in Figure 4. A series of expansions in two volumes on the isotherm are carried out, so that the required properties temperature and

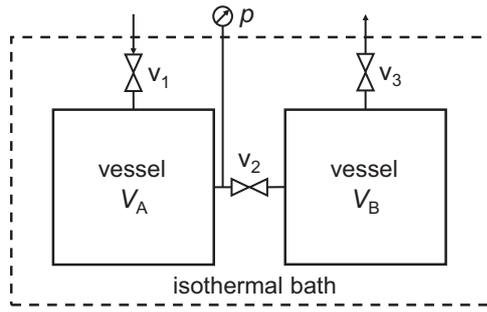


Figure 4. Working principle of a Burnett apparatus

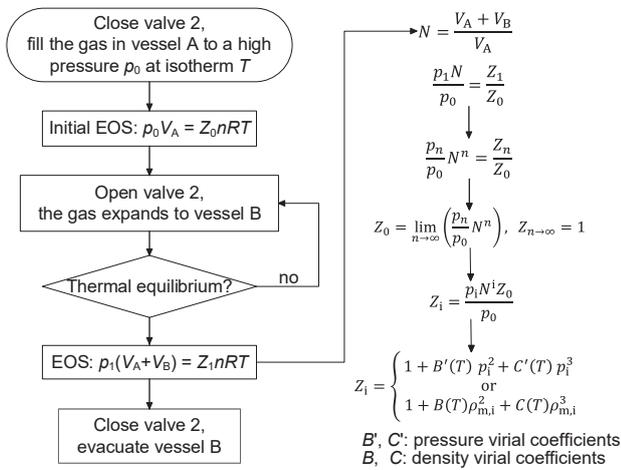


Figure 5. Flow chart of Burnett expansion series

pressure can be easily measured. Then the compressibility factor  $Z$  is determined, as shown in Figure 5.

### A. Design of a New Burnett Apparatus at HSU

In the group at HSU, a new Burnett apparatus for the measurement of the second and third virial coefficients of pure gases and gas mixtures containing hydrogen was developed and set up in the laboratory. It is designed for the temperature range from 233 K to 343 K with pressures up to 7 MPa.

As shown in Figure 6, the two chambers  $V_A$  and  $V_B$  are dug out from both ends of a pressure vessel, which is made of stainless steel 1.4980. This material is an austenitic precipitation-hardened stainless steel, has a high creep resistance, and is compatible with hydrogen in the working range of the apparatus. The three main valves for supply ( $v_1$ ), expansion ( $v_2$ ) and discharge ( $v_3$ ) of the gas are retrofitted with motors, which are controlled and operated automatically by the data acquisition computer. The hydrogen mixture is separated from the absolute pressure transducer by a differential pressure transducer, in which a membrane with gold plating on the hydrogen side acts as a separator. The branch of the valve

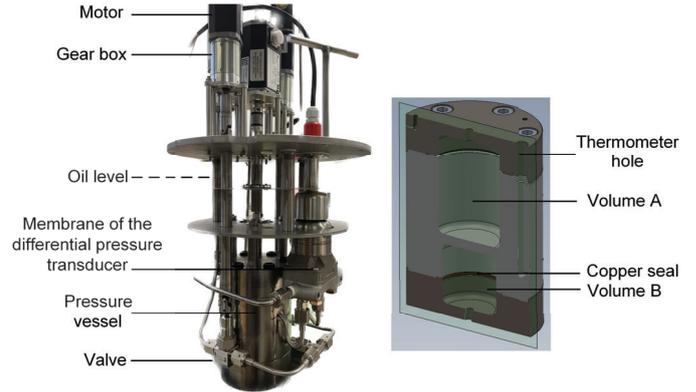


Figure 6. Assembly inside the oil bath and a cross section of the pressure vessel

system with the absolute pressure transducer is filled with nitrogen. The pressure vessel, the three valves  $v_1$ ,  $v_2$ , and  $v_3$ , and the differential pressure transducer are thermostated in a circulating liquid bath thermostat with silicone oil as thermostating liquid. The thermostat ensures a temperature stability of 0.3 mK. Since the uncertainty of the measured virial coefficients is mostly influenced by the uncertainty of the pressure measurement, the absolute pressure transducer is calibrated before and after each expansion series with a highly accurate gas piston gauge. In order to avoid zero point shifts of the differential pressure transducer during a measurement campaign, the zero point is calibrated by short-cutting both sides. The temperature is measured by a long-stem standard platinum resistance thermometer calibrated on the ITS-90 in the wall of the pressure vessel with an uncertainty of 2.5 mK by a resistance bridge system.

The operation of the apparatus is fully automated by a software written in the graphic programming environment LabVIEW. It enables the automatic calibration of the pressure sensors and automatic execution of an expansion series.

The apparatus at HSU was validated using argon. The validated volume ratio of the two measurement chambers is 1.4599. The experimental pressure ratios of a series of expansion with argon at  $T = 313.12$  K is shown in Figure 7. Figure 8 presents the absolute deviation of each pressure ratio from the obtained fourth-order virial equation of state fit to the data in Figure 7. The value of the second virial coefficient  $B$  obtained in this manner is shown in Figure 7 and compared with an *ab initio* calculated value from the literature [14].

For the apparatus at HSU, more advanced calibration schemes are being evaluated to significantly reduce the uncertainty in the pressure measurement, which currently dominates the uncertainty budget. Measurements on hydrogen-containing gas mixtures will be performed after the validation measurements with argon, nitrogen, and helium have reached the desired level of accuracy.

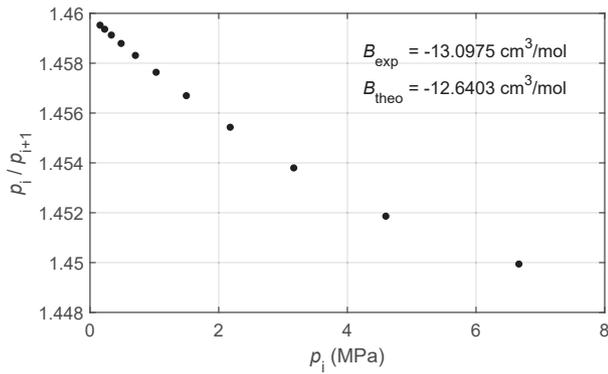


Figure 7. Results of an expansion series of argon at 313.12 K

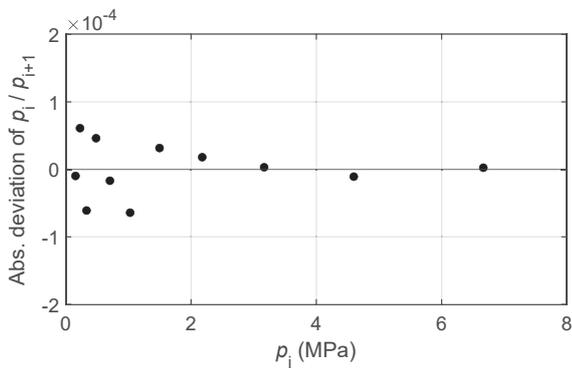


Figure 8. Absolute deviations of pressure ratios from the fit in Figure 7

### B. Implementation of a Burnett Apparatus at PTB

The apparatus at PTB combines Dielectric-Constant Gas Thermometry with Burnett expansion experiments using highly accurate pressure transducers and capacitance bridges. The combination allows for determining the density and the dielectric virial coefficients separately from the expansion experiments as well as their combinations by means of the so-called DCGT virial coefficients. Virial coefficients obtained in this manner are cross validated to some extent since the main sources of uncertainty are fairly different for both methods. Since usually large experimental setups and time consuming measurement procedures are needed, a more compact and automated setup has been designed in the past and has been used for experimental studies on noble gases [15], [16]. During these studies, weaknesses of the setup, e.g., temperature gradients, valve types used for the expansion, and the ratio of volumes at measurement and ambient temperature, have been identified. Within the frameworks of this project and on the basis of these findings, another improved setup has been designed and is now under construction. In Figure 9, a picture of the new apparatus at PTB (DCGT-H2) can be seen. In contrast to the first design draft, stainless steel chambers were fitted into the copper block to ensure a tight seal between the CF50 flanges of the capacitor mounts and the chamber at 7 MPa. Sufficient thermal conduction between the steel



Figure 9. Current state of the new DCGT-H2 apparatus at PTB

chamber and the copper block is obtained by a tight fit and use of suitable vacuum grease. Meanwhile, tests regarding the hydrogen compatibility of the measurement equipment have been conducted using the old setup described in refs. [15], [16]. First measurements showed that the two selected Digiquartz Sensors (manufacturer: Paroscientific, type: 9000-2K and 42K-101) are not suitable. During the time of one isothermal measurement, negative drifts of the vacuum value in the order of 0.1 MPa appeared after exposition to hydrogen. The sensors approach the initial zero values very slowly under vacuum, but this takes time in the order of weeks or even months. The exact responsible mechanism is unknown at this point. The drift follows no strict or simple rules, preventing a mathematical compensation. However, it is likely that the hydrogen either penetrates the vacuum reference of the sensor or permeates into the used materials, changing their mechanical properties. New pressure sensors were purchased from Keller (PAA-23SX, 6.0/6.6 MPa) and Wika (CPT9000, 10 MPa) and will be tested with regard to hydrogen stability, resolution, and uncertainty. As a backup, classical DCGT measurements were performed with pure hydrogen utilizing a differential pressure transducer and a pressure balance. These first results of combined dielectric and density virial coefficients showed reasonable uncertainties and agreement with established literature values. Furthermore, the laboratory safety requirements were evaluated for the new system. Fortunately, the efforts in this regard can be reduced significantly in comparison to the old system, since the room for the new apparatus is equipped with a permanent ventilation.

## V. MEASUREMENTS ON THE NITROGEN/HYDROGEN SYSTEM WITH A MODIFIED VIBRATING-WIRE VISCOMETER

The vibrating-wire viscometer at HSU was originally developed by Seibt [17] at Universität Rostock in the group of Prof. Eckhard Vogel and was applied there to measure the viscosity of gaseous nitrogen, methane, ethane, propane, *n*-butane, and isobutane [17], [18]. After the retirement of Prof. Vogel, it was transferred to HSU, rebuilt in our laboratory, and used to measure the viscosity of neon by Kochan-Eilers [19]. The experimental setup consists of a single-sinker densimeter and three vibrating-wire sensors (Figure 10) in a single pressure vessel and enables simultaneous measurements of the density and viscosity of a gas. It covers the temperature range between 293 K to 498 K at pressures up to 30 MPa. The density measurements are characterized by a relative expanded uncertainty (coverage factor  $k = 2$ ) of 0.1%, except in the low density range, where it is larger, while the relative expanded ( $k = 2$ ) uncertainty in viscosity is estimated to be 0.25% to 0.3%.

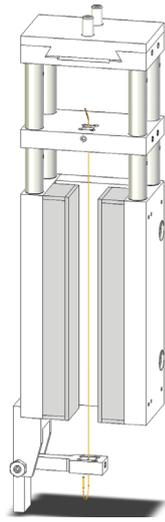


Figure 10. Technical drawing of a redesigned vibrating-wire sensor

The working principle of the vibrating-wire technique and the design of the apparatus were already described by Hellmann et al. [20]. For measurements of hydrogen and mixtures containing hydrogen, some components of the valve system, of the vibrating-wire sensors, and a differential pressure transducer, which couples the measured gas with the nitrogen in the pressure measurement system, were replaced by components in which the parts in contact with hydrogen are made of materials compatible with hydrogen. The differential pressure transducer (Emerson, type 3051) is replaced by a new transducer of the same type, in which the stainless steel membrane, which separates the measured gas from the nitrogen, is gold plated on the side in contact with hydrogen. The gold plating prevents hydrogen from diffusing into and through the membrane. Since the copper-beryllium alloy of the

pressure vessel could not be qualified as a certified material for pressure vessels, the outer stage of the double-mantle thermostat, which surrounds the pressure vessel, is employed as shatter protection. This required a redesign of the double-mantle thermostat.

In previous measurement campaigns, the temperature control of the thermostat of the viscometer was sometimes unstable because of heat losses at the top cover of the thermostat. Therefore, an additional temperature-controlled heating coil was applied near the top cover to decouple the pressure vessel in the thermostat thermally from the laboratory environment. In addition, a stand-alone setup with a single vibrating-wire sensor in a new pressure vessel was built. The pressure vessel is thermostatted in a circulating liquid oil bath of a calibration thermostat with a temperature stability of 5 mK (Fluke Calibration 7341). The pressure vessel was designed by Eurotechnica and made of the stainless steel 1.4980. The thermal equilibrium in the pressure vessel and gas is attained much faster than in the rather complex combined densimeter/viscometer setup due to the convective heat transfer in the liquid bath thermostat.

Moreover, the operation of the viscometer has been automated. A pressure controller (WIKA, type CPC8000) is used to control the pressure in the nitrogen branch of the valve system. The manual valves for discharging gas from the high pressure system have been replaced by pneumatic valves, which are controlled by the data acquisition computer. By using a computer-controlled relay switch system, the data acquisition computer can automatically switch between the three different vibrating-wire sensors.

The results of the calibration measurements for the determination of the wire radius with helium at 298.15 K are shown in Figure 11. The wire radius was adjusted iteratively until the extrapolated viscosity value at zero density agreed with the highly accurate theoretical value at 298.15 K by Cencek et al. [21].

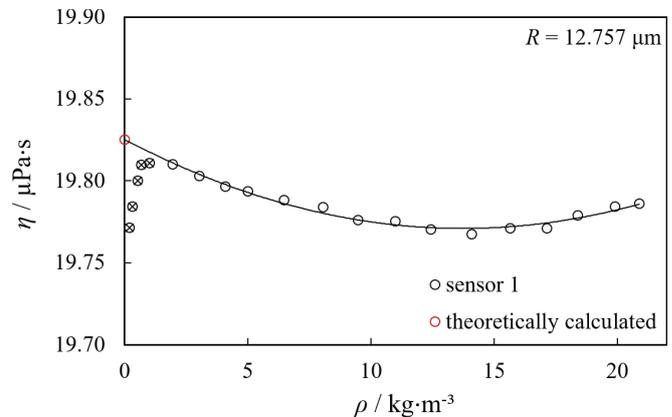


Figure 11. Calibration measurements with helium at 298.15 K as a function of density

After calibration, measurements on nitrogen at 293.15 K were performed to validate the calibration as shown in Figure 12. The relative deviations of our data agree with the correlation of Lemmon and Jacobsen [22] within its expanded ( $k = 2$ ) uncertainty, which confirms the calibration.

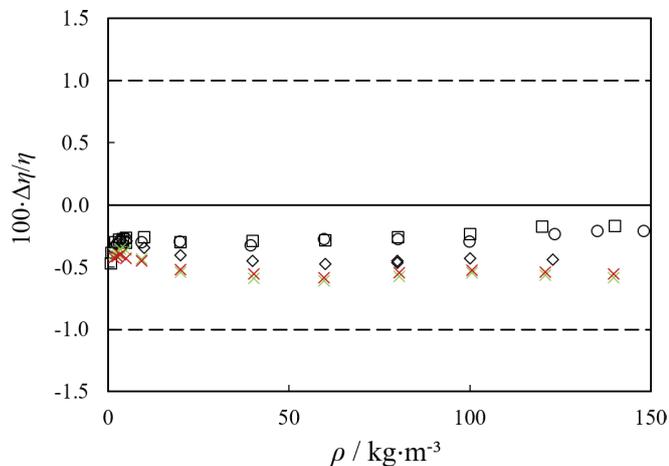


Figure 12. Relative deviations of our results for nitrogen at 293.15 K from the correlation of Lemmon and Jacobsen [22]. The maximum density corresponds to a pressure of 12 MPa.  $\times$ , sensor 1;  $\times$ , sensor 2;  $\square$ , Seibt [17];  $\circ$ , Kochan-Eilers [19];  $\diamond$ , Herrmann [18], -- expanded uncertainty of the correlation of Lemmon and Jacobsen [22]

A mixture of hydrogen and nitrogen was prepared gravimetrically at HSU according to the procedure proposed by McLinden [23] with a molar composition of  $x_{\text{H}_2} = 0.25447$  and  $x_{\text{N}_2} = 0.74553$  and an uncertainty of  $U = 0.00012$  in the mole fractions. The results of the first measurements at 293.15 K and 313.15 K are compared with the extended corresponding states model implemented in REFPROP 10.0 [24] in Figures 13 and 14.

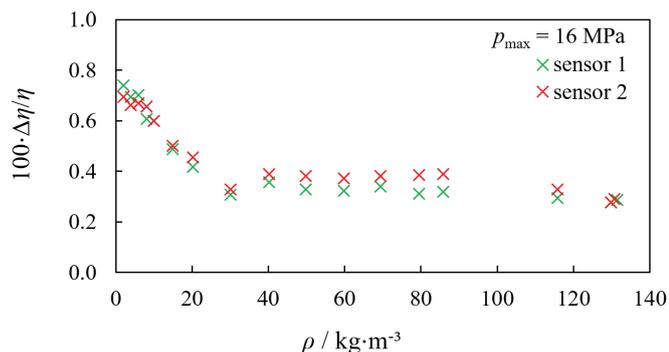


Figure 13. Relative deviations of our viscosity measurements at 293.15 K from the extended corresponding states model implemented in REFPROP 10.0 [24]

Figure 15 shows the extrapolated viscosity at zero-density of our measurements at 293.15 K compared to the data of Kestin and Yata [25]. While the values obtained by our sensors differ by 0.03%, the interpolated value of Kestin and Yata is

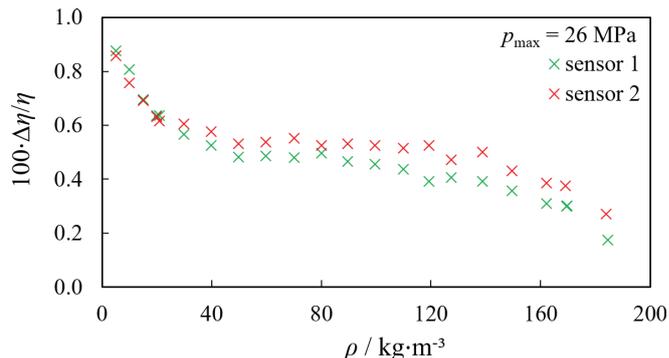


Figure 14. Relative deviations of our viscosity measurements at 313.15 K from the extended corresponding states model implemented in REFPROP 10.0 [24]

about 0.3% higher, but the agreement is within the mutual uncertainties.

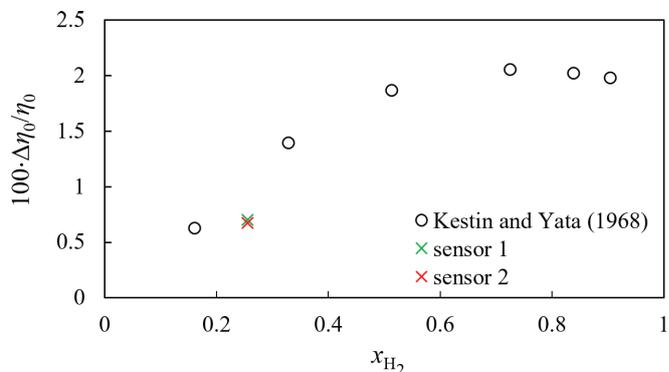


Figure 15. Relative deviation of the viscosity at zero density from the extended corresponding states model implemented in REFPROP 10.0 [24] at 293.15 K

Due to the improvements in the experimental setup, the vibrating-wire viscometer can now be used for automated viscosity measurements of hydrogen containing gas mixtures. In the extension of the project, further binary systems containing hydrogen will be measured in the temperature range accessible by our viscometer.

## VI. OUTLOOK

In the theoretical part of the project, the potential energy surfaces for the eight investigated binary mixtures with hydrogen and pure oxygen were developed, and the cross second virial coefficients for the mixtures and the second virial coefficient of pure oxygen were calculated. Moreover, a code for PIMC simulations of neon and helium was developed and used to calculate thermodynamic properties of both noble gases at high pressures in the supercritical region and in the liquid region of neon with very small uncertainties comparable to those achieved by the most accurate experimental techniques. Using the methodology devised by Lustig, expressions for the calculation of ten different properties in the  $NpT$  ensemble were derived, which enable the calculation of several of these

properties, such as the speed of sound and Joule–Thomson coefficient, in PIMC simulations for the first time. In the remainder of the first phase of the project, the transport properties of the eight binary systems will be calculated, and the PIMC simulations will be extended to the grand canonical ensembles as well as to the liquid region of helium, where Bose–Einstein statistics must be applied.

On the experimental side, the Burnett apparatuses at HSU and PTB were set up in the laboratories and automated using data acquisition software coded in LabVIEW. First validation measurements with argon were carried out at HSU. Second virial coefficients derived from the expansion series data showed already good agreement with theoretical reference values. To reduce the uncertainty further, more sophisticated procedures for the calibration of the pressure sensors will be developed before measurements on hydrogen containing mixtures will be performed. With the vibrating-wire viscometer modified for measurements on hydrogen containing mixtures, two isotherms of a mixture of about 25 mol% hydrogen and 75 mol% nitrogen were measured. The results of these measurements agree well with literature data at low pressure, but deviate by up to 1% from an extended corresponding states model. These measurements will be continued and extended to binary hydrogen mixtures with methane, ethane, and carbon dioxide. Moreover, Eurotechnica will perform additional density measurements with a single-sinker densimeter, which serve as a check of the measurements carried out at HSU and also extend the temperature and pressure range of the measurements at HSU.

#### ACKNOWLEDGMENT

This research is funded by dtec.bw - Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union - NextGenerationEU.

#### REFERENCES

- [1] J. K. Singh, D. A. Kofke, “Mayer sampling: calculation of cluster integrals using free-energy perturbation methods,” *Phys. Rev. Lett.* **92**, 220601 (2004).
- [2] R. P. Feynman, A. R. Hibbs, “Quantum mechanics and path integrals,” McGraw-Hill, New York (1965).
- [3] R. Hellmann, “*Ab initio* potential energy surfaces for the O<sub>2</sub>–O<sub>2</sub> system and derived thermophysical properties,” *J. Chem. Phys.* **159**, 104303 (2023).
- [4] R. Hellmann, “Cross second virial coefficients of the H<sub>2</sub>O–H<sub>2</sub> and H<sub>2</sub>S–H<sub>2</sub> systems from first principles,” *J. Chem. Eng. Data* **68**, 2212–2222 (2023).
- [5] R. Lustig, “Microcanonical Monte Carlo simulation of thermodynamic properties,” *J. Chem. Phys.* **109**, 8816–8828 (1998).
- [6] M. F. Herman, E. J. Bruskin, and B. J. Berne, “On path integral Monte Carlo simulations,” *J. Chem. Phys.* **76**, 5150–5155 (1982).
- [7] T. M. Yamamoto, “Path-integral virial estimator based on the scaling of fluctuation coordinates: Application to quantum clusters with fourth-order propagators,” *J. Chem. Phys.* **123**, 104101 (2005).
- [8] R. Hellmann, C. Gaiser, B. Fellmuth, T. Vasylytsova, and E. Bich, “Thermophysical properties of low-density neon gas from highly accurate first-principles calculations and dielectric-constant gas thermometry measurements,” *J. Chem. Phys.* **154**, 164304 (2021).
- [9] R. Hellmann and G. Garberoglio, in preparation (2024).
- [10] P. Czachorowski, M. Przybytek, M. Lesiuk, M. Puchalski, and B. Jeziorski, “Second virial coefficients for <sup>4</sup>He and <sup>3</sup>He from an accurate relativistic interaction potential,” *Phys. Rev. A* **102**, 042810 (2020).
- [11] J. Lang, G. Garberoglio, M. Przybytek, M. Jeziorska, and B. Jeziorski, “Three-body potential and third virial coefficients for helium including relativistic and nuclear-motion effects,” *Phys. Chem. Chem. Phys.* **25**, 23395–23416 (2023).
- [12] T. Dietl and K. Meier, in preparation (2024).
- [13] D. O. Ortiz Vega, K. R. Hall, J. C. Holste, A. H. Harvey, and E. W. Lemmon, “An equation of state for the thermodynamic properties of helium,” Technical Report No. NISTIR 8474, National Institute of Standards and Technology, (2023).
- [14] E. Vogel, B. Jäger, R. Hellmann, and E. Bich, “*Ab initio* pair potential energy curve for the argon atom pair and thermophysical properties for the dilute argon gas. II. Thermophysical properties for low-density argon,” *Mol. Phys.* **108**, 3335–3352 (2010).
- [15] C. Günz, C. Gaiser, and M. Richter, “Concept of a new technique for the combined determination of dielectric and density virial coefficients,” *Meas. Sci. Technol.* **28**, 027002 (2017).
- [16] C. Günz “Combined Dielectric-Constant Gas Thermometry and Expansion Experiments - Virial Coefficients of Argon,” Thesis PTB-Bericht Th **12**, doi 10.7795/110.20220106 (2021).
- [17] D. Seibt, “Schwingdrahtviskosimeter mit integrierter Ein-Senkörper-Dichtemessverfahren für Untersuchungen an Gasen in größeren Temperatur- und Druckbereichen,” *Fortschritt-Berichte, Reihe 6*, Nr. 571, VDI-Verlag, Düsseldorf (2008).
- [18] S. Herrmann, “Viskositäts- und Dichtemessungen an *n*-Butan und Isobutan in größeren Temperatur- und Druckbereichen,” *Fortschritt-Berichte, Reihe 6*, Nr. 615, VDI-Verlag, Düsseldorf (2015).
- [19] U. Kochan-Eilers, “Measurements of the viscosity of the noble gas neon with a vibrating-wire viscometer and correlations of the viscosity of gases in the limit of zero density,” Dissertation, Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg (2021).
- [20] R. Hellmann, R. Wang, R. Harten, K. Meier, C. Günz, M. Manzel, C. Gaiser, S. Schulz, A. Pietsch, J. Zipfel, P. Jaeger, “Progress on computer-aided and experimental determination of thermophysical properties of gas mixtures containing hydrogen,” in *Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw*, Band 1, Hrsg.: D. Schulz, A. Fay, M. Schulz, and W. Matiaske (dtec.bw Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, Hamburg, 2022).
- [21] W. Cencek, M. Przybytek, J. Komasa, J. B. Mehl, B. Jeziorski, K. Szalewicz, “Effects of adiabatic, relativistic, and quantum electrodynamic interactions on the pair potential and thermophysical properties of helium,” *J. Chem. Phys.* **136**, 220601 (2004).
- [22] E. W. Lemmon, R. T. Jacobsen, “Viscosity and thermal conductivity equations for nitrogen, oxygen, argon, and air,” *Int. J. Thermophys.* **25**, 21–69 (2010).
- [23] M. O. McLinden “Gravimetric preparation of gas mixtures,” private communication (National Institute of Standards and Technology, Boulder, 2007).
- [24] Lemmon, E. W.; Bell, I.H.; Huber, M. L.; McLinden, M. O. NIST Reference fluid thermodynamic and transport properties database (Refprop version 10.0), NIST Standard Reference Database 23; National Institute of Standards and Technology, Standard Reference Data Program: Gaithersburg, 2018.
- [25] J. Kestin, J. Yata, “Viscosity and diffusion coefficient of six binary mixtures,” *J. Chem. Phys.* **69**, 4780–4791 (1968).

# xbat: A Continuous Benchmarking Tool for HPC Software

Nico Tippmann\*, Axel Auweter  
MEGWARE Computer Vertrieb und  
Service GmbH

Chemnitz, Deutschland

\*First author: nico.tippmann@megware.com

Amartya Das Sharma<sup>†</sup>, Willi Leinen  
Chair for High Performance Computing,  
Department of Mechanical and Civil Engineering,  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
<sup>†</sup> das-sharma@hsu-hh.de

Philipp Neumann  
Deutsches Elektronen-Synchrotron (DESY)  
Universität Hamburg  
Hamburg, Deutschland

Andreas Fink  
Institute of Computer Science,  
Department of Economics and Social Sciences  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland

**Abstract** — Benchmarking the performance of one’s application in high performance computing (HPC) systems is critically important for reducing runtime and energy costs. Yet, accessing the plethora of relevant metrics that impact performance is often challenging, particularly for users without hardware experience. In this paper, we introduce the novel benchmarking tool xbat developed by MEGWARE GmbH. xbat requires no setup from the user side, and it allows the user to run, monitor and evaluate their application from the tool’s web interface, consolidating the entire benchmarking process in an approachable, intuitive workflow. We demonstrate the capabilities of the tool using benchmark applications of varying complexity and show that it can manage all aspects of the benchmarking workflow in a seamless manner. In particular, we focus on the open-source molecular dynamics research software `ls1 mardyn`, and the closed-source optimisation package `Gurobi`. Both packages present unique challenges. Mixed-integer programming solvers, such as those integrated in the `Gurobi` software, exhibit significant performance variability, so that seemingly innocuous parameter changes and machine characteristics can affect the runtime drastically, and `ls1 mardyn` comes with an auto-tuning library `AutoPas`, that enables the selection of various node-level algorithms to compute molecular trajectories. Focusing on these two packages, we showcase the practicality, versatility and utility of xbat, and share its current and future developments.

**Keywords** — benchmarking, high performance computing, numerical simulation, mixed-integer programming, performance analysis

## I. INTRODUCTION AND MOTIVATION

Gaining insight into an application’s performance and its use of the underlying hardware can be a complex and error prone process. Furthermore, benchmarking directly on the cluster poses additional challenges to regular users in obtaining relevant data, and requires processing and comparing a potentially large number of performance indicators. Tools like `JUBE`<sup>1</sup> can

<sup>1</sup><https://github.com/FZJ-JSC/JUBE>

provide an automated, reproducible environment to users, but require extensive configuration and do not collect any data on their own, leaving it to the users. Additionally, different levels of instrumentation also complicate selecting the appropriate tool, as they use compiler instrumentation (like `VTune Profiler`<sup>2</sup>) or require access to the binary (like `MAQAO`<sup>3</sup>). A more general approach is using Hardware Performance Monitoring solutions like `LIKWID`<sup>4</sup> or `Perf` [1] that provide application independent access to low-level metrics. However, they are less user friendly and typically monitor only a limited number of specified metrics throughout the application’s runtime. In an effort to decrease the barrier of entry, `ClusterCockpit`<sup>5</sup> offers cluster-wide job-specific monitoring for administrators and users through a web-based user interface to assess general hardware utilisation. Nevertheless, due to limits in portability and scope, high performance computing (HPC) would greatly benefit from an easy-to-use, universally applicable, and comprehensive benchmarking tool.

## II. XBAT

xbat, the extended benchmarking automation tool, builds upon the idea of `ClusterCockpit` [2] by automating the entire job configuration, submission and benchmarking process through a web-based interface. Since it maps the job configuration process, it is possible to easily generate different job scripts and thus conduct parameter or scalability studies similar to `JUBE`. Additionally, the in-depth data collection performed by xbat provides both inexperienced and experienced

<sup>2</sup><https://www.intel.com/content/www/us/en/developer/tools/oneapi/vtune-profiler.html>

<sup>3</sup><https://www.maqao.org>

<sup>4</sup><https://github.com/RRZE-HPC/likwid>

<sup>5</sup><https://www.clustercockpit.org/>

users easy access to a plethora of performance metrics that can serve as a basis for performance analysis and optimisation.

Integration of *xbat* into existing cluster installations has been made as simple and straightforward as possible, requiring only minimal configuration. As depicted in Figure 1, this is achieved by encapsulating most of the tool’s components into a docker/podman composable infrastructure for deployment on a management or login node. A collector daemon (*xbatd*) exists on each compute node, which is controlled by custom Slurm<sup>6</sup> prolog and epilog scripts so that the daemon only collects data on selected jobs. The ability to connect to existing authentication providers (IPA/LDAP/PAM) provides seamless access to all users of the cluster.

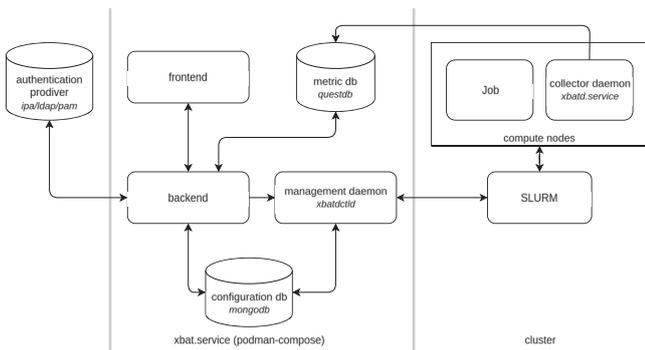


FIGURE 1: XBAT ARCHITECTURE

*xbat*’s philosophy is to provide the most in-depth monitoring possible that can be applied to any application. In order to make this possible, we leverage the LIKWID performance tool suite as the main source for low-level performance metrics. Other data is acquired via the *proc* filesystem or through libraries like *nvml*<sup>7</sup>. In total, up to 120 different metrics<sup>8</sup> are collected for the CPU, memory, I/O, interconnect, energy, GPU and FPGA. All measurements are performed at the lowest possible resolution. This enables *xbat* to aggregate these measurements and provide them dynamically at thread, core, numa, socket, device, node and job level. The measurement interval is adjustable by the users – with 5 seconds currently being the lowest possible interval.

Apart from job configuration and submission, the web interface serves as the central access points for all benchmarks. A built-in collaboration feature allows sharing configurations and results between individual users and groups. As shown in Figure 2, users are provided with basic information about their job, such as the runtime, output and configuration, and hardware and software characteristics of the executing node. The measurements are displayed within a highly extensive, customisable and interactive graph system that also shows statistics on the collected data. All graphs presented in this paper are exported directly from *xbat* as it features exporting graphs as PNG or SVG, in addition to exporting the underlying

data as CSV and JSON. Furthermore, users may explore the memory-boundedness (whether the performance is limited by the memory bandwidth) or compute-boundedness (whether the performance is limited by the peak floating point operations per second (FLOPS)) of their application via the integrated, automatic roofline modelling and compare different jobs with one another.

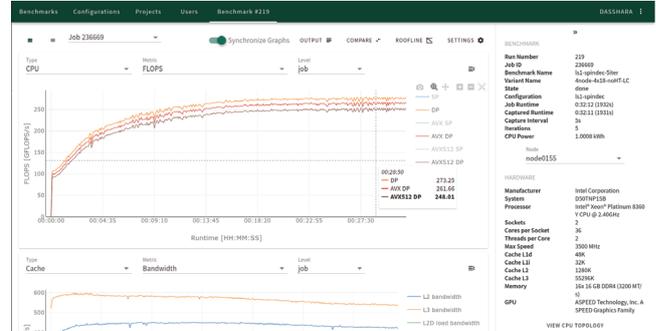


FIGURE 2: PARTIAL XBAT INTERFACE FOR BENCHMARK RESULTS

### III. COLLABORATION BENEFITS

For the purposes of collaboration, *xbat* was deployed onto HSU’s own Intel Icelake based HPC system HSUper<sup>9</sup>. Over several months, HSU provided beta testing for the tool, utilising both open- and closed-source software and running real research problems and provided feedback and suggestions to MEGWARE. This collaboration has been fruitful for both parties; *xbat* is now a more complete software with minimal bugs and more features than before, and HSU has access to an excellent easy-to-use benchmarking software that has already started proving its worth in relevant research.

### IV. SOFTWARE CORRECTNESS AND OVERHEAD

LIKWID, the main data source for *xbat*, has already established itself as a reliable tool [3]. However, *xbat*’s way of instrumenting LIKWID, especially in the context of data collection at (hardware) thread level and subsequent aggregation, requires additional validation. For this purpose, we compare the performance of the industry standard benchmarks STREAM<sup>10</sup> and HPL<sup>11</sup> with the values gathered by *xbat*, depicted in Figures 3 and 4. These benchmarks are typical examples for memory-bound and compute-bound applications, respectively. All benchmarks in this work use HSUper. Each node of HSUper contains two sockets equipped with Intel Xeon Platinum 8360Y (36 cores, 2.4 gigahertz) processors, hence 72 physical cores (or 144 virtual CPUs with hyper-threading). From the results, we see that both STREAM and HPL are consistent with *xbat* within a margin of less than two percent.

We measured the typical baseline performance of a system by capturing performance data for a job that simply calls

<sup>6</sup><https://slurm.schedmd.com>

<sup>7</sup><https://developer.nvidia.com/management-library-nvml>

<sup>8</sup>depending on availability

<sup>9</sup><https://www.hsu-hh.de/hpc/en/cbrz/>

<sup>10</sup><https://www.cs.virginia.edu/stream/>

<sup>11</sup><https://netlib.org/benchmark/hpl/>

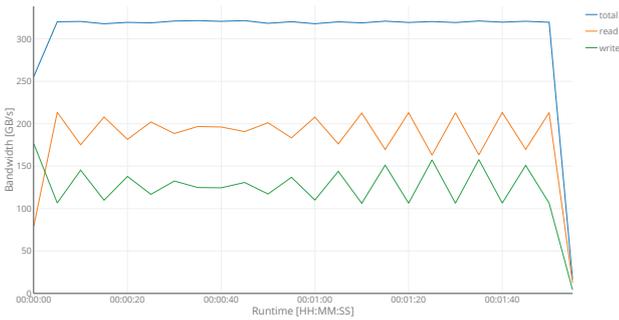


FIGURE 3: STREAM BENCHMARK BANDWIDTH [GB/s]

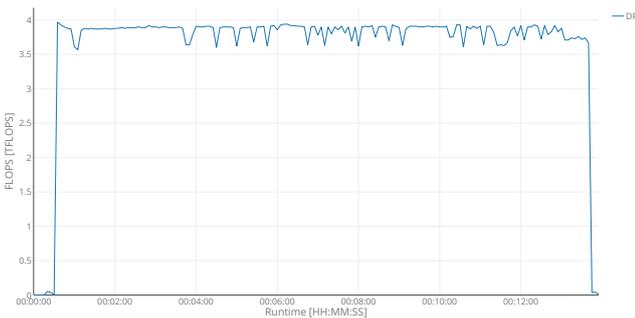


FIGURE 4: HPL BENCHMARK [TFLOPS]

sleep, causing the CPU to idle. The background noise amounts to roughly 225 FLOPS, 13 megabyte per second memory bandwidth and an average of 1.86 percent CPU usage at HSUPER, which is caused by the operating system and the daemon. According to our measurements, the overhead of the daemon of xbat on an application's performance (such as HPL) is less than 0.5 percent, which can mainly be attributed to the very low overhead of LIKWID [3]. Therefore, we can conclude that the impact on the performance measurements can be neglected in HPC.

## V. SHOWCASE

### A. Gurobi

#### 1) Introduction

Gurobi<sup>12</sup> is a closed-source commercial solver which can be applied to different types of optimisation problems such as mixed-integer linear programming (MILP).

A major challenge in evaluating the performance of a MILP solver is performance variability, since the performance is intrinsically not deterministic (i.e. seemingly performance-neutral changes such as a changed order of the constraints, or using a different random seed or different hardware can lead to a significant change in the performance) [4], [5], [6]. Benchmarking MILP software is further discussed in [7], [8], [9]. Gurobi provides general guidelines about good hardware setups<sup>13</sup>. The main conclusion is that the prediction of the Gurobi performance on a given hardware is difficult,

however better hardware (high clock speed, larger, low-latency, and high-bandwidth memory, more cores) usually accelerates the computation. However, better hardware is usually more expensive. Therefore, it is desirable to identify the hardware characteristics that noticeably limit Gurobi's performance.

#### 2) Experimental Setup

We focus on the performance of Gurobi (Version 11.0.1, build v11.0.1rc0 for Linux with 64 bit) on a single cluster node. Our test problem is the time-dependent traveling salesman problem *ta064* from [10].

One run of Gurobi always uses one HSUPER node exclusively and we do not limit the hardware resources via SLURM. Hence, the Gurobi solver can use up to 32 software threads at a time while the (number of) used hardware threads can change over time.

Strong scaling experiments for the test problem *ta064* on various hardware architectures have been presented previously by the authors in the hpc.bw newsletters [11], [12], [13], illustrating the influence of performance variability while using various seed values. The influence of concurrent optimisation is shown in [14] whereas various time-dependent traveling salesman problems from [10] are solved in [15]. These tests illustrated that a moderate number of threads, e.g. between 4 and 32 threads, usually leads to the best performance and it is often not advisable to use all cores of hardware with many cores (e.g. 144 threads on a HSUPER node). These observations correspond to Gurobi's choice to impose a soft limit of 32 threads in the Gurobi solver. This is because the Gurobi solver is not yet highly tuned for hardware with many cores, since a typical Gurobi customer uses the tool on hardware with 4 to 16 cores<sup>14</sup>.

In the following, we focus mainly on Gurobi with default parameters<sup>15</sup> and only vary the seed value to assess performance variability. We use xbat to provide further insights into the observations presented in the previous paragraph.

#### 3) Results

Figure 5 shows the amount of performed floating point operations per second in GFLOPS (i.e.  $10^9$  FLOPS) at node level for Gurobi with default parameters, i.e. a seed value of 0. xbat hides zero traces by default. Hence, this figure shows that Gurobi is performing only double precision (DP) and no single precision (SP) operations. The amount of performed DP FLOPS oscillates significantly over time, whereas almost no operations are vectorised using, e.g. AVX or AVX512. The Gurobi log file reveals that the peak is associated with a heuristic procedure that finds a better primal solution.

Figure 6 shows the percentage of vectorised operations for several seed values and reveals that mostly less than 10 percent of the performed operations are vectorised.

Figure 7 is similar to Figure 5 except that no vectorised operations are included, and the DP operations are shown at thread level. The legend is cut off after thread 19. Omitted

<sup>12</sup><https://www.gurobi.com>

<sup>13</sup><https://support.gurobi.com/hc/en-us/sections/360010047491-Hardware>

<sup>14</sup><https://www.nvidia.com/en-us/on-demand/session/gtc24-s62555/>

<sup>15</sup><https://www.gurobi.com/documentation/11.0/refman/parameters.html>

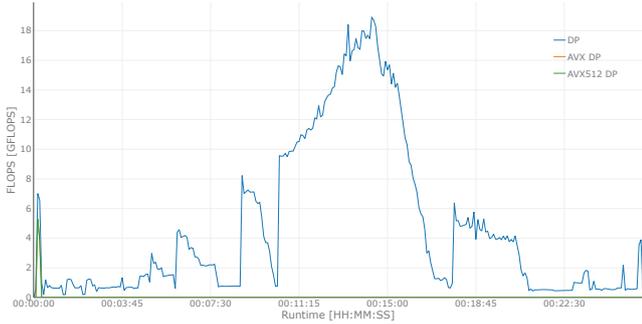


FIGURE 5: FLOPS AT NODE LEVEL FOR GUROBI WITH DEFAULT VALUES.

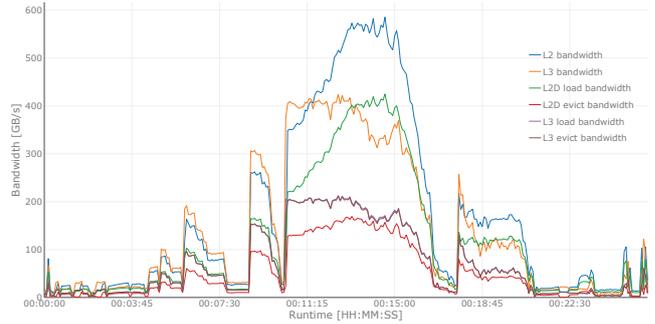


FIGURE 8: CACHE BANDWIDTHS AT NODE LEVEL FOR GUROBI WITH DEFAULT VALUES.

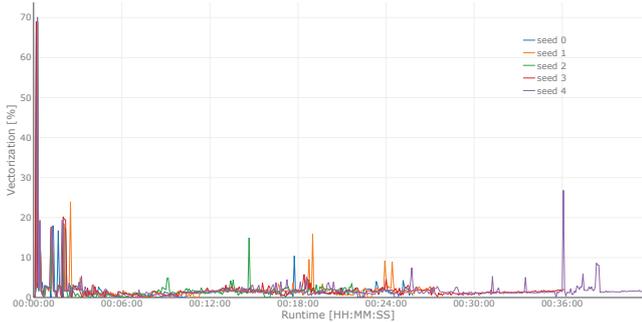


FIGURE 6: PERCENTAGE OF VECTORISED FLOPS FOR GUROBI WITH VARIOUS SEED VALUES.

time. Gurobi reads more data than it writes (within main memory) and peaks in the memory bandwidths correspond to peaks in the L3 bandwidths.

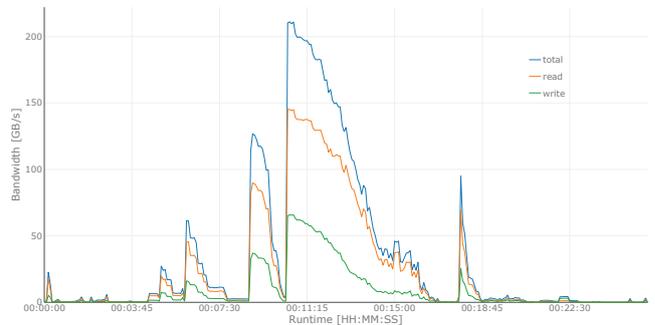


FIGURE 9: MEMORY BANDWIDTH AT NODE LEVEL FOR GUROBI WITH DEFAULT VALUES.

threads such as thread 15 performed no FLOPS during the computation, and 102 out of 144 threads performed FLOPS. This figure illustrates that the majority of the FLOPS are often performed by only a few threads, sometimes just one.

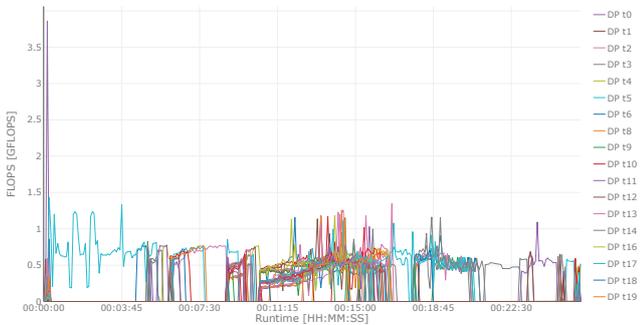


FIGURE 7: FLOPS AT THREAD LEVEL FOR GUROBI WITH DEFAULT VALUES.

Figure 10 illustrates the memory usage in gigabyte for various seed values and demonstrates that the memory usage is increasing over time. However, only a fraction of the available memory (256 gigabyte on a normal HSUPER node and 1 terabyte on a fat node) is needed to solve this MILP.

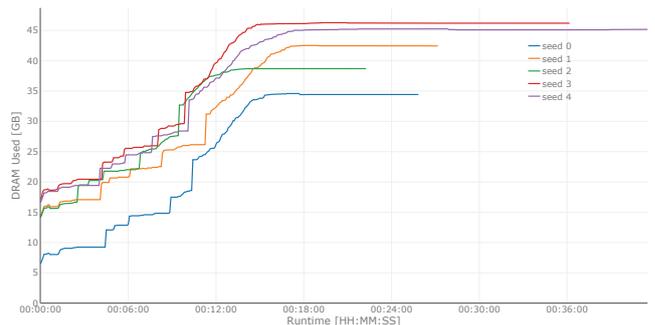


FIGURE 10: DRAM USED FOR GUROBI WITH VARIOUS SEED VALUES.

Figure 8 shows the L2 and L3 cache bandwidths in gigabyte per second. The total bandwidth is given by the sum of the corresponding load and evict bandwidths. This shows that the bandwidth oscillates significantly over time. The L2 bandwidth is seen to match (and even exceed) the L3 bandwidth. A comparison with Figures 5 and 7 reveals that the most FLOPS are performed when a large L2 bandwidth is used.

Figure 9 shows the memory bandwidths in gigabyte per second corresponding to the data from Figure 8. The total bandwidth is given by the sum of the read and write bandwidths and these bandwidths also oscillate significantly over

Our results show that the Gurobi solver usually does not need the total available memory and FLOPS, and the cache and memory bandwidths oscillate significantly over time. Hence, to better utilise the computing resources, it seems to be advisable to solve several MIP instances simultaneously on hardware with many cores<sup>14</sup> such as an HSUPER node (possible with

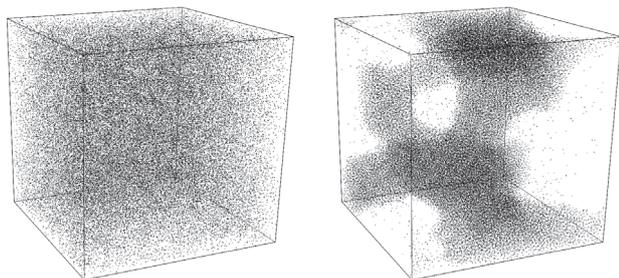
fewer than 32 threads), instead of one Gurobi instance with all available threads. The ideal case would be that the bandwidth of the simultaneously running instances do not peak at the same time.

## B. *ls1 mardyn + AutoPas*

### 1) Introduction

*ls1 mardyn*<sup>16</sup> is an open-source molecular dynamics simulation software, used by chemical engineers to simulate small rigid molecules at large scales [16], [17] using the Lennard-Jones (LJ) potential [18]. Usually, it uses a linked-cell data structure to store particles. However, it also supports the node-level library *AutoPas*<sup>17</sup>, which not only supports various data structures, traversal algorithms and data layouts, but also switches between them at runtime to best suit the needs of evolving simulations [19]. The benefits of this dynamic adaptation (termed *autotuning*) have been researched [20], and we attempt to partially replicate these findings using *xbat*.

For this purpose, we consider the spinodal decomposition scenario. In this scenario, the temperature of a low-density volume of liquid is abruptly reduced below critical, causing the liquid to cluster and form local inhomogeneities. This scenario was chosen due to the evolution of the system from homogeneous to inhomogeneous, causing local high-density volumes of particles. A visualization is shown in Figure 11.



(a) INITIAL STATE (b) LATER STATE  
FIGURE 11: THE EVOLUTION OF A SPINODAL DECOMPOSITION

### 2) Experimental Setup

For the scenario, we use the default values from the *ls1 mardyn* example setup for this experiment (particle mass  $m = 1$ , Lennard-Jones (LJ) particle size  $\sigma = 1$ , LJ particle energy parameter  $\epsilon = 1$ , density  $\rho = 0.29$ , temperatures  $T_{equilibrium} = 1.40$ ,  $T_{decomposition} = 0.70$ , all in LJ reduced units) except for the liquid volume, which is a  $150 \times 150 \times 150$  box, giving us 976,563 particles. We benchmark the latter half of the experiment, when the equilibration of the liquid is complete, and the decomposition is triggered. The simulation is run for 80000 steps.

Four experiments are run, each on 4 nodes, with 4 MPI ranks on each node. The experiments are run with hyperthreading turned on and off (18 and 36 OpenMP threads per MPI

<sup>16</sup><https://github.com/ls1mardyn/ls1-mardyn>

<sup>17</sup><https://github.com/AutoPas/AutoPas>

TABLE I: AVERAGE RUNTIME OF THE *LS1 MARDYN* EXPERIMENTS

Experiment type	Avg. runtime [s]
HT, LC	2079.4
HT, AP	2487.2
no HT, LC	1674.2
no HT, AP	1166.6

rank, resulting in 1 and 2 threads per core respectively) and with *AutoPas* enabled and disabled (using the inbuilt linked cell structure in *ls1 mardyn* for the latter). The goal is to observe the differences in runtime that hyperthreading and *AutoPas* create. Runs with and without hyperthreading, and with and without *AutoPas*, are designated *HT*, *no HT*, *AP* and *LC* respectively. The experiments were each run 5 times. The average runtime is shown in Table I. The remaining plots show the output from the fastest run of each experiment.

### 3) Results

Firstly, we report the runtimes of our experiments in Table I. Non-hyperthreaded runs are faster than their hyperthreaded counterparts (1.24 times faster for *LC*, 2.13 times faster for *AP*), and non-hyperthreaded *AutoPas* is the fastest available configuration (1.43 times faster than *no HT, LC*).

The following figures shed more light on this behaviour. We see in Figure 12 that for all experiments, the FLOPS gradually increase, as local particle densities increase due to clustering of particles, leading to more particle-particle interactions as time increases. We see that the *no HT, LC* run has a higher consistent FLOP rate than the runs with *AutoPas*, and yet the *AutoPas* run ends sooner. We also see the hyperthreaded runs performing worse. Figure 13 provides a possible explanation for the underperformance of hyperthreading; the stall count (stalls caused by data traffic in cache hierarchy) for hyper-threaded runs is often double that of regular runs, throttling the performance considerably.

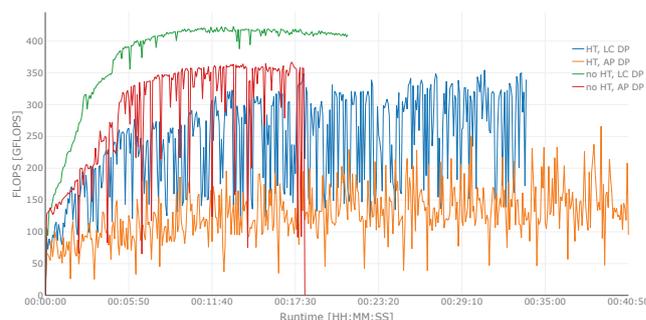


FIGURE 12: FLOP RATES FOR *LS1 MARDYN* EXPERIMENTS

The faster execution of *AutoPas* is theorised to be due to the better vectorisation, and Figure 14 supports that claim, as we see a consistent  $\approx 10\%$  better vectorisation for *AutoPas* code. We can see in Figure 15 (memory bandwidth as a function of time) that *AutoPas* always uses less bandwidth than vanilla *ls1 mardyn* does, which points towards more efficient use of memory and consequently less time spent waiting for data. The *LC* runs have worse vectorisation, and hence need more



FIGURE 13: TOTAL STALL COUNT FOR LS1 MARDYN EXPERIMENTS

instructions (more FLOPs) and more data to compute the same result, hence the slower runtime. However, more research is needed to further investigate this claim.

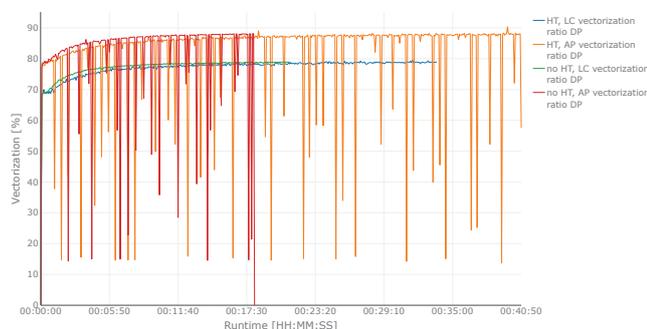


FIGURE 14: VECTORISATION PERCENT OF INSTRUCTIONS FOR LS1 MARDYN EXPERIMENTS



FIGURE 15: MEMORY BANDWIDTH FOR LS1 MARDYN EXPERIMENTS

## VI. CONCLUSION

In this work, we showcased the novel benchmarking tool `xbat`. We explained the underlying functionality, its capabilities and limitations, and with this background knowledge we demonstrated its use. We first used the tool to run the `STREAM` and `HPL` benchmarks to verify the correctness of the calculations, and then we ran an idle job to check the overhead that the tool causes. After observing favourable results, we used the tool to benchmark the software packages `Gurobi` and `ls1 mardyn + AutoPas`, in order to show how the tool may be used in real-life applications.

We can conclude that `xbat` is a powerful, yet simple tool that helps investigate behaviour of a software, so that users can optimise the hardware usage. Moreover, finding an optimal hardware setup for a given software can be greatly simplified. The tool is useful for both beginners and veterans of HPC and lends itself to a variety of use cases. Despite being still under development, `xbat` intends to go fully open source in 2024. Further features are planned to augment its capabilities, such as extending the scope of the data collection, increasing the focus on supporting analysis efforts and identifying performance bottlenecks. The latter may be used in conjunction with version control to provide automated feedback about performance changes during continuous integration.

## ACKNOWLEDGMENT

The HPC cluster `HSUper` has been provided by the projects `hpc.bw` and `MaST`. The projects `hpc.bw` and `MaST` are funded by `dtec.bw` – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. `dtec.bw` is funded by the European Union – NextGenerationEU. Development of `xbat` was partially funded by EuroHPC project `MICROCARD`.

## REFERENCES

- [1] A. C. De Melo, “The New Linux ‘perf’ tools,” in *Slides from Linux Kongress*, vol. 18, pp. 1–42, 2010.
- [2] T. Röhl, J. Eitzinger, G. Hager, *et al.*, “LIKWID Monitoring Stack: A flexible framework enabling job specific performance monitoring for the masses,” in *2017 IEEE International Conference on Cluster Computing (CLUSTER)*, pp. 781–784, IEEE, 2017.
- [3] J. Treibig, G. Hager, and G. Wellein, “LIKWID: A Lightweight Performance-Oriented Tool Suite for x86 Multicore Environments,” in *2010 39th international conference on parallel processing workshops*, pp. 207–216, IEEE, 2010.
- [4] E. Danna, “Performance variability in mixed integer programming,” in *Workshop on Mixed Integer Programming, Columbia University, New York*, vol. 20, 2008.
- [5] T. Koch, T. Achterberg, E. Andersen, *et al.*, “MIPLIB 2010: mixed integer programming library version 5,” *Math. Program. Comput.*, vol. 3, pp. 103–163, 2011.
- [6] A. Lodi and A. Tramontani, “Performance variability in mixed-integer programming,” in *Theory driven by influential applications*, pp. 1–12, INFORMS, 2013.
- [7] T. Achterberg and R. Wunderling, “Mixed integer programming: Analyzing 12 years of progress,” in *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*, pp. 449–481, Springer, 2013.
- [8] H. D. Mittelman, “Benchmarking Optimization Software - a (Hi)Story,” in *SN Oper. Res. Forum*, vol. 1, 2020.
- [9] S. J. Maher, T. K. Ralphs, and Y. Shinano, “Assessing the Effectiveness of (Parallel) Branch-and-bound Algorithms,” *arXiv:2104.10025*, 2021.
- [10] A. Fink and S. Voß, “Solving the continuous flow-shop scheduling problem by metaheuristics,” *European J. Oper. Res.*, vol. 151, no. 2, pp. 400–414, 2003.
- [11] P. Neumann and M. Mayr, “hpc.bw newsletter 01/2023,” 2023.
- [12] P. Neumann and M. Rathmann, “hpc.bw newsletter 03/2023,” 2023.
- [13] P. Neumann and M. Rathmann, “hpc.bw newsletter 01/2024,” 2024.
- [14] P. Neumann, M. Rathmann, H. Preuß, *et al.*, “hpc.bw newsletter 02/2023,” 2023.
- [15] P. Neumann and M. Rathmann, “hpc.bw newsletter 04/2023,” 2023.
- [16] C. Niethammer, S. Becker, M. Bernreuther, *et al.*, “ls1 mardyn: The Massively Parallel Molecular Dynamics Code for Large Systems,” *Journal of chemical theory and computation*, vol. 10, no. 10, pp. 4455–4464, 2014.
- [17] N. Tchipev, S. Seckler, M. Heinen, *et al.*, “TweTriS: Twenty trillion-atom simulation,” *The International Journal of High Performance Computing Applications*, vol. 33, no. 5, pp. 838–854, 2019.

- [18] J. Lennard-Jones and B. M. Dent, "Cohesion at a crystal surface," *Transactions of the Faraday Society*, vol. 24, pp. 92–108, 1928.
- [19] F. A. Gratl, S. Seckler, N. Tchipev, *et al.*, "AutoPas: Auto-Tuning for Particle Simulations," in *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, (Rio de Janeiro), IEEE, 2019.
- [20] S. Seckler, F. Gratl, M. Heinen, *et al.*, "AutoPas in ls1 mardyn: Massively parallel particle simulations with node-level auto-tuning," *Journal of Computational Science*, vol. 50, 2021.

## **Kapitel II**

### **Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS)**

mit Beiträgen von

CTO 21strategies

Airbus Operations GmbH

CTC GmbH (An Airbus Company)

Frankfurt University of Applied Sciences

Fraunhofer-Institut für Materialfluss und Logistik IML

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Leibniz Universität Hannover

Robert Bosch GmbH

Ruhr-Universität Bochum

Technische Universität Bergakademie Freiberg

Technische Universität Braunschweig

Universität der Bundeswehr München Uni-

versität Hamburg

voraus robotic GmbH

## **Vorwort zum dtec.bw-Dachprojekt „Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS)“**

Künstliche Intelligenz ist in aller Munde und zieht mit großer Geschwindigkeit in viele Lebensbereiche ein. Aktuell ist insbesondere bei Werkzeugen zur Erstellung von Texten und Bildern auf Basis von Large Language Models eine rasante Entwicklung zu beobachten. Intensiv wird – auch im Rahmen von dtec.bw – erforscht, wie diese und andere Methoden der Künstlichen Intelligenz für den Entwurf und Betrieb technischer Systeme eingesetzt werden können. Aufgrund der statistischen Natur der zugrundeliegenden Algorithmen sind die Ergebnisse, die viele Algorithmen liefern, weder vorhersagbar noch vollständig richtig. Wie können solche Algorithmen in technischen Systemen für die Diagnose, die Optimierung oder die Steuerung eingesetzt werden, ohne Sicherheit und Zuverlässigkeit zu gefährden? Auch das sind Forschungsfragen, die in den Projekten unter dem Dach von „Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS)“ erforscht werden. Darüber hinaus geht es in KIIPS um eine große Bandbreite von Methoden und Technologien der Informations- und Kommunikationstechnik, der Elektronik und Automatisierungstechnik, die für verbesserte oder ganz neue Anwendungen adaptiert, rekombiniert und erweitert werden. Zahlreiche Technologie- und Forschungspartnerschaften ermöglichen eine interdisziplinäre Zusammenarbeit und vielfältige Anwendungsbezüge. Die Anwendungsbereiche reichen von autonomen Fahrzeugen (zu Land, zu Wasser und in der Luft) über verbesserte Produktionsverfahren von Flugzeugteilen und die Überwachung von Brücken bis zu Nutzungsmöglichkeiten von 5G-Kommunikation.

Die Beiträge zu diesem Sammelband zeigen die Breite der Projekte und der Arbeitsschwerpunkte darin, insbesondere aber auch, wie innovative Methoden bereits in Funktionsmustern und Pilotanwendungen praktisch erprobt werden. Die Jahre der kommenden Fortsetzung von dtec.bw werden ermöglichen, diese Methoden weiterzuentwickeln und die Anwendungsmöglichkeiten auszubauen und die Methoden umfangreich zu validieren.

Allen beteiligten Autoren des Dachprojekts „KIIPS“ sei herzlich gedankt für diese Beiträge zu diesem Sammelband.

Der Staffelstab der Leitung des Dachprojekts „KIIPS“ ist inzwischen von Herrn Prof. Dr. Alexander Fay an Herrn Prof. Dr. Oliver Niggemann übergeben worden. Wir möchten bei dieser Gelegenheit allen Projektleitungen und allen in den Projekten Mitarbeitenden, Studierenden und Forschungs- und Anwendungspartnern für ihr Engagement danken.

Allen Leserinnen und Lesern wünsche ich anregende Lektüre. Die Autorentteams freuen sich auf Fragen und Anregungen von Ihnen!

Hamburg, im November 2024

Prof. Dr. rer. nat. Oliver Niggemann und Prof. Dr.-Ing. Alexander Fay

Dachprojektleiter „Künstliche Intelligenz und Intelligente Physische Systeme (KIIPS)“

# Optimierung der Auslagerungsprozesse in automatisierten Kompaktlagern

Nicolas Fauvé, Florian Jaehn  
*Institut für Management Science und Operations Research  
Helmut-Schmidt-Universität  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
nicolas.fauve@hsu-hh.de*

Julian Golak, Yagmur Gül,  
Philipp Koch, Simone Neumann  
*Institut für Nachhaltige Logistik- und Mobilitätssysteme/  
Institut für Operations Management  
Universität Hamburg  
Hamburg, Deutschland*

Alice Kirchheim  
*Fraunhofer-Institut für Materialfluss und Logistik IML  
Dortmund, Deutschland*

Lara Nehrke  
*Institut für Beschaffung und Produktion  
Helmut-Schmidt-Universität  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland*

**Zusammenfassung**—Um einen schnellen Zugriff auf die gelagerten Artikel zu gewährleisten, gibt es verschiedene Arten von Lagersystemen. Eine Reihe von Faktoren, darunter die physische Größe und das Gewicht der zu lagernden Artikel, die Nutzungshäufigkeit und die verfügbaren Ressourcen, bestimmen, welcher Systemtyp für ein bestimmtes Lager am besten geeignet ist. In diesem Beitrag soll ein neuer Typ von Lagersystemen untersucht werden: Kompaktlagersysteme. Kompaktlagersysteme zielen darauf ab, eine möglichst hohe Raumausnutzung bei begrenztem Lagerplatz zu erreichen. Zwar ist solch ein Lager platzsparend, allerdings ist die Entnahme von Artikeln aus Kompaktlagersystemen komplex. Durch die dichte Anordnung sind die Ladeeinheiten oft nicht direkt zugänglich, da sie von anderen Einheiten verdeckt werden. Diese müssen erst verschoben oder umpositioniert werden, um genügend Platz für den Zugriff auf das gewünschte Objekt zu schaffen. Dies führt zu komplexen Entscheidungsprozessen. Da das Umpositionieren von Ladeeinheiten Energie verbraucht, trägt eine effiziente Entscheidungsfindung direkt zu einem energieeffizienten Lagersystem bei. Dies erfordert die Entwicklung von algorithmischen Unterstützungssystemen, die in der Lage sind, optimale Entnahmestrategien in angemessener Zeit zu berechnen. In diesem Bericht geben wir einen Überblick über die Annahmen und das strukturelle Optimierungsproblem, das diesen Entnahmestrategien zugrunde liegt. Darüber hinaus bieten wir Einblicke in die Berechenbarkeit und die algorithmischen Ansätze, die zur Ableitung effizienter Lösungen verwendet werden.

**Index Terms**—Logistik, Algorithmen, Lagersysteme,

## I. EINFÜHRUNG

In den letzten Jahren ist die Anzahl der Bestellungen stark gestiegen. Dies hat zu einem erhöhten Bedarf an Lagern geführt, insbesondere in urbanen Gebieten, wo Platz knapp und teuer ist. Traditionelle Lagersysteme, bei denen Ladeeinheiten oder Artikel einfach in Regalen gelagert werden, sind zwar simpel und ermöglichen eine effiziente Auslagerung, brauchen jedoch viel Platz, da zwischen den Regalen genügend Raum für Menschen oder Roboter zum Manövrieren gelassen werden muss. Dies motiviert die Untersuchung von Kompaktlagersystemen. In Kompaktlagersystemen werden die

Ladeeinheiten (also die Boxen oder Behälter, im Folgenden als Unit Loads (ULs) bezeichnet) so dicht wie möglich aneinandergereiht, um eine optimale Nutzung des Platzes zu gewährleisten. Diese Systeme erfordern eine komplexere Ein- und Auslagerung, die von autonomen Robotern übernommen wird. AutoStore<sup>1</sup> ist ein Beispiel für ein solches Kompaktlagersystem. Dabei gibt es keine Regale mit festen Gängen, sondern die einzelnen ULs, die einheitlich sind und die eingelagerten Artikel enthalten, werden direkt übereinander und nebeneinander gestapelt. Die Ein- und Auslagerung erfolgt über temporäre Gänge, die durch das Anheben der ULs entstehen. Wir betrachten ein Lagersystem, bei dem die ULs von der Seite ausgelagert werden können. In einem dichten Layout sind die ULs oft nicht direkt zugänglich, da sie von anderen Einheiten verdeckt werden. Um auf einen bestimmten UL zugreifen zu können, müssen möglicherweise andere Einheiten zuerst verschoben oder umpositioniert werden. Dies führt zu komplexen Entscheidungsprozessen. Da das Umpositionieren von ULs Energie verbraucht, ist eine effiziente Entscheidungsfindung elementar für die Energieeffizienz des Lagersystems.

Dieses neuartige Kompaktlagersystem mit Zugriff von der Seite wirft eine Vielzahl von Problemstellungen auf, die alle zu komplex sind, um allein durch menschliche Intuition gelöst zu werden. Dies motiviert die Entwicklung algorithmischer Entscheidungssysteme, die in der Lage sind, gute oder optimale Entscheidungen in kurzer Rechenzeit zu treffen. Um solche algorithmischen Ansätze zu entwickeln und zu analysieren, benötigen wir eine Beschreibung, Intuition und Einblick in die einzelnen Teilprobleme, die betrachtet werden müssen. In diesem Artikel greifen wir eines dieser Probleme auf und zeigen dem Leser die Herausforderung und mögliche Lösungsansätze.

Bisher wurden solche Kompaktlagersysteme von der akademischen Fachliteratur kaum betrachtet und es gibt nur zwei

<sup>1</sup><https://de.autostoresystem.com>

Publikationen, die genau dieses Problem adressieren. Fauvé und Neumann [1] führen diese Systeme ein und diskutieren Modellierungs- und Kostenfaktoren. Von Fauvé et al. [2] werden sowohl bereits auf dem Markt existierende Kompaktlager wie AutoStore und Ocado betrachtet, als auch grundsätzliche technische und organisatorische Rahmenbedingungen solcher Lagersysteme. Da AutoStore und Ocado lediglich von oben auf das Lager zugreifen und gemäß dem LiFo-Prinzip (Last in – First out) funktionieren, wird gezeigt, dass auch Systeme mit seitlicher Ein- und Auslagerung möglich sind, die flexiblere und effizientere Durchlaufzeiten bieten.

## II. GRUNDLAGEN

In diesem Abschnitt bieten wir eine Betrachtung der grundlegenden Problemstrukturen, die dem seitlichen Zugriff auf Lagertypen zugrunde liegen.

Wir untersuchen ein Lagersystem, in dem ULs vertikal gestapelt werden, sodass die gesamte Anordnung die Form eines Quaders annimmt (siehe Abbildung 1). Diese ULs müssen eine maximale Stapelhöhe einhalten, um eine effiziente Raumnutzung zu gewährleisten.

Autonome Fahrzeuge, sogenannte *Shuttles*, sind für den Transport der ULs verantwortlich. Diese Shuttles können die ULs eigenständig anheben und bewegen. Sie werden mithilfe von Aufzügen auf die gewünschte Höhe im Quader angehoben und können dann in den Quader einfahren, sofern ausreichend Platz vorhanden ist. Dieser notwendige Platz wird durch integrierte Hebesysteme geschaffen, die einen Teil des Stapels anheben und so einen tunnelähnlichen Raum erzeugen, in dem sich die Shuttles bewegen können (siehe Abbildung 2). Die Shuttles nutzen diesen Raum, um die ULs aus dem Quader herauszuholen. Zusätzlich bewegen sich die Shuttles innerhalb des Quaders auf einem Schienensystem, das oben auf den ULs angebracht ist. Dieses System verhindert, dass sich die Shuttles innerhalb des Quaders drehen.

Zur Vereinfachung betrachten wir in unserem Modell nur eine einzelne Spalte des Quaders. Wir nehmen außerdem an, dass die Shuttles nur von einer Seite in das System einfahren können. Diese Annahme erleichtert die Darstellung erheblich, da Probleminstanzen als zweidimensionale Gitter modelliert werden können. Eine Verallgemeinerung der Probleminstanz auf den gesamten Quader ist zwar direkt möglich, jedoch sehr aufwendig. In der Praxis können solche Probleme auftreten, wenn der Platz auf allen anderen Seiten des Quaders eingeschränkt ist. In diesen Fällen reduziert sich das Problem auf die Steuerung der Auslagerung einzelner Spalten, was unserer vereinfachten Modellbetrachtung entspricht.

## III. BESCHREIBUNG DES AUSLAGERUNGSPROZESSES

Im Folgenden beschreiben wir die Modellierung des Auslagerungsprozesses, also den Ablauf der Bearbeitung eines Kundenauftrags. Dabei erläutern wir die einzelnen Schritte, die notwendig sind, um die angeforderten Waren aus dem Lager zu entnehmen.

Kundenbestellungen umfassen Artikel, die in den ULs gelagert sind. Um diese Bestellungen effizient zu bearbeiten,

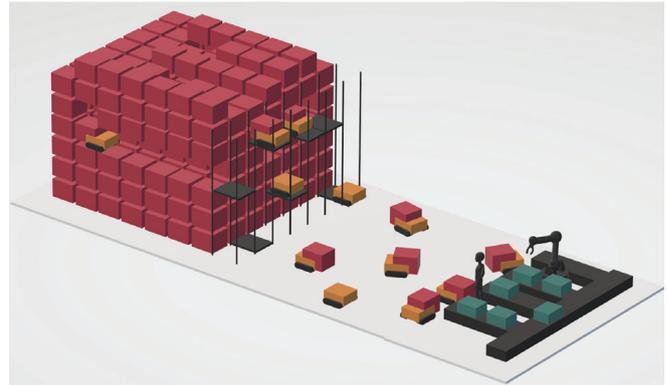


Abbildung 1: ILLUSTRATION DES LAGERLAYOUTS MIT DEN GESTAPELTEN ULS SOWIE DEN SEITLICHEN AUFZÜGEN UND HEBESYSTEME. DIE SEITLICHEN AUFZÜGE HEBEN DIE SHUTTLES AUF DIE RICHTIGE HÖHE, UND DIE HEBESYSTEME SCHAFFEN DURCH ANHEBEN DER ULS PLATZ FÜR DAS SHUTTLE IM QUADER.

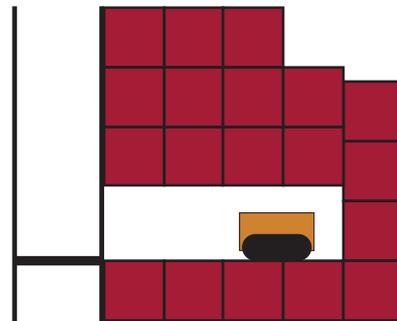


Abbildung 2: ILLUSTRATION DES VEREINFACHTEN LAGERLAYOUTS MIT DEN GESTAPELTEN ULS UND ZUGANG VON LINKS.

wird im Voraus festgelegt, welche spezifischen ULs ausgelagert werden sollen. Diese Vorabentscheidung trägt dazu bei, Konflikte zu vermeiden und stellt sicher, dass die Aufträge effektiv aufgeteilt und bearbeitet werden können.

Nehmen wir an, es gibt eine Liste der auszulagernden ULs. Jede UL hat eine bestimmte Höhe und eine spezifische Position im Stapel innerhalb des Lagers. Der Auslagerungsprozess dieser ULs wird in mehrere Schritte oder Zyklen unterteilt. In jedem Zyklus wird entschieden, welche ULs angehoben werden müssen, um die gewünschten Artikel zu erreichen.

Für jeden dieser Zyklen muss der Bediener festlegen, welche ULs angehoben werden sollen und in welcher Reihenfolge die auszulagernden ULs ausgelagert werden. Dies wird durch die Anzahl, die in jedem Stapel während eines Zyklus angehoben wird, dargestellt. Nach der Auslagerung einer UL wird diese zu einem externen Ablageort gebracht und nicht wieder in das Lager zurückgeführt. In jedem Zyklus muss gewährleistet sein, dass die ULs, die ausgelagert werden sollen, tatsächlich erreichbar sind. Dabei ist zu beachten, dass das Shuttle nur auf bereits vorhandenen ULs fahren kann. Es ist daher notwendig sicherzustellen, dass keine ULs zwischen dem Einstiegspunkt des Shuttles und den gewünschten ULs liegen, die nicht ebenfalls bewegt oder entnommen werden.

Wir stellen fest, dass unterschiedliche Entnahmereihenfolgen der ULs zu verschiedenen Konfigurationen führen. ULs, die sich oberhalb einer entnommenen UL befinden, fallen sozusagen nach unten, was bedeutet, dass in den nachfolgenden Zyklen andere Konfigurationen von ULs vorliegen. Dies führt dazu, dass genaue Entscheidungen über die Reihenfolge der Entnahme erforderlich sind, um eine Optimalität entsprechend einer Zielfunktion zu gewährleisten. Diese Beobachtung impliziert, dass jede Entscheidung über die Reihenfolge der Entnahme direkten Einfluss auf die zukünftige Lagerkonfiguration und damit auf die Zugänglichkeit der verbleibenden ULs hat. Um eine optimale Entnahmestrategie zu entwickeln, müssen alle möglichen Konfigurationsänderungen berücksichtigt und die Reihenfolge der Entnahme entsprechend geplant werden.

#### IV. ZIELFUNKTIONEN

Unser Ziel ist es, den Energieverbrauch des Lagersystems zu minimieren, wobei der größte Teil des Energieverbrauchs durch das Anheben der ULs entsteht, um den benötigten Platz zu schaffen. Die Energie, die zum Anheben einer UL erforderlich ist, kann mithilfe der Formel für potenzielle Gravitationsenergie berechnet werden:  $E = m \cdot g \cdot h$ . Hierbei steht  $E$  für die potenzielle Gravitationsenergie,  $m$  für die Masse des Objekts,  $g$  für die Erdbeschleunigung und  $h$  für die Höhe, auf die das Objekt angehoben wird.

Diese Formel verdeutlicht, dass die Energie, die zum Anheben eines Objekts benötigt wird, direkt proportional zu dessen Gewicht ist, wobei das Gewicht das Produkt aus Masse und Gravitationsbeschleunigung darstellt. Um die Berechnungen zu vereinfachen, nehmen wir an, dass alle ULs das gleiche Gewicht haben. Dies führt dazu, dass der gesamte Energieaufwand direkt proportional zur Anzahl der angehobenen ULs ist. Um die Analyse weiter zu vereinfachen, normalisieren wir die Gewichte der ULs.

Die Eingabe für den Optimierungsalgorithmus besteht aus einer Darstellung des Lagersystems, nämlich einer Liste der Höhen der jeweiligen Stapel und einer Liste der auszulagernden ULs. Die Ausgabe, oder Lösung, wird durch eine Sequenz von Zyklen dargestellt, in denen festgelegt wird, welche ULs angehoben werden, um den erforderlichen Platz für das Shuttle zu schaffen, und welche ULs aus dem Kundenauftrag entnommen werden. Diese Lösung ist gültig, wenn alle ULs des Kundenauftrags entnommen wurden und jede dieser ULs zugänglich ist. Das Ziel des Optimierungsalgorithmus besteht darin, eine zulässige Lösung zu finden, die den Energieverbrauch minimiert.

#### V. BEISPIEL ZUR VERANSCHAULICHUNG DES PROBLEMS

In diesem Abschnitt wird das Problem anhand eines konkreten Beispiels illustriert, siehe Abbildung 3 und 4. Wir betrachten eine Instanz des Problems der Auslagerung von ULs aus einem Kompaktlager mit fünf Stapeln, deren Höhen wie folgt definiert sind: Der erste Stapel hat eine Höhe von 4, der zweite eine Höhe von 3, der dritte eine Höhe von 4, der vierte eine Höhe von 4 und der fünfte eine Höhe von 5.

Weiterhin liegt uns ein Kundenauftrag vor, der die ULs  $\{b_1, b_2, b_3, b_4, b_5, b_6, b_7\}$  umfasst. Jede UL  $b$  ist einem spezifischen Stapel und einer bestimmten Höhenposition zugeordnet. Die Positionen der ULs sind wie folgt:  $b_1$  befindet sich im Stapel 1 auf Höhe 3,  $b_2$  im Stapel 3 auf Höhe 3,  $b_3$  im Stapel 4 auf Höhe 4,  $b_4$  im Stapel 4 auf Höhe 1,  $b_5$  im Stapel 5 auf Höhe 5,  $b_6$  im Stapel 5 auf Höhe 2 und  $b_7$  im Stapel 5 auf Höhe 1 (siehe Abbildung 3a). Zur besseren Übersicht werden die ULs durch Koordinaten identifiziert, das heißt, zu Beginn ist  $b_1$  mit  $(1,3)$  identifiziert.

Im ersten Zyklus wird die UL  $b_6$  entnommen, die sich an Position  $(5,2)$  befindet. Dazu müssen alle ULs angehoben werden, die sich links, links oberhalb und oberhalb von  $b_6$  befinden. Die ULs, die sich links befinden, werden um etwas mehr als die Höhe eines ULs angehoben, damit der entstandene Platz ausreicht, damit das Shuttle in den Tunnel einfahren und das gewünschte UL entnehmen kann. Bei der Kalkulation der Kosten ignorieren wir den marginalen Höhenunterschied und nehmen an, dass die Energie dem Anheben der Anzahl der ULs entspricht. Die ULs, die sich über  $b_6$  befinden, werden ebenfalls um eine Höhe angehoben, da aus Sicherheitsgründen ein Mindestabstand zwischen den ULs eingehalten werden muss. Daher entspricht auch hier die Energie der Anzahl der angehobenen ULs. Das sind insgesamt 14 ULs, was einen Energieaufwand von 14 Energieeinheiten bedeutet. Nach diesem Zyklus ist die Höhe von Stapel 5 um 1 reduziert (siehe Abbildung 3b).

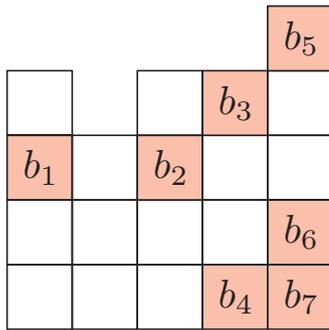
Im zweiten Zyklus werden die ULs  $b_4$  und  $b_7$  entnommen, die sich an Positionen  $(4,1)$  und  $(5,1)$  befinden. Dafür müssen alle ULs um die Höhe eines ULs angehoben werden, die sich oberhalb von  $b_4$  und  $b_7$  und links der beiden ULs befinden. Insgesamt werden somit 17 ULs angehoben, was einen Energieaufwand von 17 Energieeinheiten für diesen Zyklus bedeutet. Nach diesem Zyklus sind die Stapel 4 und 5 um eine (weitere) Höhe reduziert (siehe Abbildung 4a).

Im dritten Zyklus stellen wir fest, dass sich die restlichen vier ULs alle auf einer Höhe befinden. Daher können wir alle restlichen ULs in einem Zyklus auslagern. Dafür müssen alle ULs angehoben werden, die sich über und zwischen den auszulagernden ULs befinden. Insgesamt werden somit drei ULs angehoben, was einen Energieaufwand von drei Energieeinheiten für diesen Zyklus bedeutet.

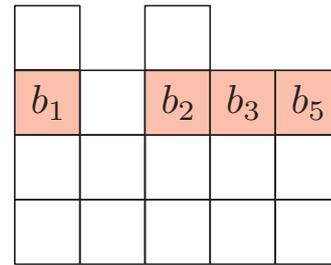
Nach dem dritten Zyklus sind alle ULs ausgelagert, und somit ist das Problem gelöst. Zusammengefasst wurden alle ULs des Kundenauftrags entnommen, und der gesamte Energieverbrauch beträgt 22 Energieeinheiten (siehe Abbildung 4b).

#### VI. ALGORITHMISCHE LÖSUNGSANSÄTZE

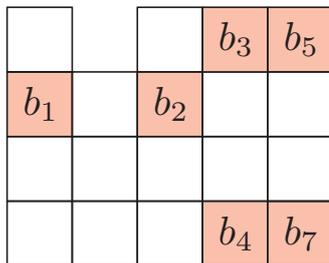
In diesem Abschnitt diskutieren wir verschiedene Ansätze zur Entwicklung von Algorithmen zur Lösung des Problems der effizienten Auslagerung in Kompaktlagern. Dabei betrachten wir sowohl einfache heuristische Methoden als auch komplexe Ansätze, die auf mathematischer Optimierung basieren. Ziel ist es, die unterschiedlichen Strategien zu beleuchten, ihre Vor- und Nachteile zu erörtern und aufzuzeigen, wie sie zur Verbesserung der Lagerlogistik beitragen können.



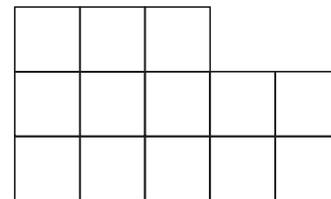
(a) ANFANGSINSTANZ, BEVOR EINE UL ENTNOMMEN WIRD.



(a) ULS  $b_4$  UND  $b_7$  WERDEN IN EINEM ZYKLUS AUSGELAGERT MIT ENERGIEAUFWAND VON 17.



(b) UL  $b_6$  WIRD IM ERSTEN ZYKLUS MIT ENERGIEAUFWAND VON 14 AUSGELAGERT.



(b) ULS  $b_1, b_2, b_3$  UND  $b_4$  WERDEN IN EINEM ZYKLUS AUSGELAGERT MIT ENERGIEAUFWAND VON 3. ALLE AUSZULAGERNDEN ULS SIND SOMIT AUSGELAGERT.

Abbildung 3: ILLUSTRATION DES ENTNAHMEPROZESSES: TEIL 1. IN ROT HERVORGEHOBN DIE AUSZULAGERNDEN ULS.

Abbildung 4: ILLUSTRATION DES ENTNAHMEPROZESSES: TEIL 2. IN ROT HERVORGEHOBN DIE AUSZULAGERNDEN ULS.

Im Manuskript von Fliedner, Golak, Gül und Neumann zeigen die Autoren durch eine Verbindung zu einer Klasse von Partitionierungsproblemen, dass das Problem NP-vollständig ist [3]. NP-Vollständigkeit ist ein Konzept in der Informatik, das verwendet wird, um die Schwierigkeit bestimmter Probleme zu beschreiben. Ein Problem ist NP-vollständig, wenn es zwei Bedingungen erfüllt: Erstens muss es in der Klasse NP liegen, was bedeutet, dass eine vorgeschlagene Lösung für das Problem in polynomieller Zeit überprüft werden kann. Zweitens muss jedes andere Problem in NP auf dieses Problem in polynomieller Zeit reduzierbar sein, was bedeutet, dass eine effiziente Lösung für dieses Problem auch effiziente Lösungen für alle anderen Probleme, die in NP liegen, ermöglichen würde.

Die Implikationen für Algorithmen sind erheblich. Wenn ein Problem als NP-vollständig klassifiziert wird, gibt es keine bekannte Methode, um es schnell (in Polynomialzeit) für alle Fälle zu lösen. Das bedeutet, dass Algorithmen, die solche Probleme lösen, in der Regel sehr zeitaufwendig sind und ihre Laufzeit exponentiell mit der Größe des Eingabeproblems zunimmt. In der Praxis bedeutet dies oft, dass für NP-vollständige Probleme nur für kleine Instanzen Lösungen gefunden werden können oder dass Approximationsalgorithmen und heuristische Methoden verwendet werden müssen, um

brauchbare Lösungen in vernünftiger Zeit zu finden.

Für die Entwicklung optimaler Lösungen können Methoden wie mathematische Programmierung oder dynamische Programmierung eingesetzt werden. Diese Ansätze zielen darauf ab, exakte Lösungen zu finden, sind jedoch oft rechnerisch sehr intensiv und daher für große Probleminstanzen nicht immer praktikabel. Mathematische Programmierung umfasst Techniken wie lineare Programmierung, gemischt-ganzzahlige Programmierung und nichtlineare Programmierung, die darauf abzielen, eine optimale Lösung innerhalb der gegebenen Beschränkungen zu finden. Dynamische Programmierung hingegen zerlegt das Problem in kleinere Teilprobleme und löst diese rekursiv, was jedoch ebenfalls hohe Rechenressourcen erfordert.

Alternativ können Ansätze entwickelt werden, die keine exakte Lösung finden, aber dennoch praktikable und effiziente Lösungen bieten. Einfache heuristische Methoden wie der Greedy-Algorithmus sind leicht zu verstehen und zu implementieren. Der Greedy-Algorithmus trifft in jedem Schritt die lokal beste Entscheidung, was oft schnell zu guten, wenn auch nicht immer optimalen Lösungen führt. Diese Methode ist besonders nützlich, wenn schnelle Entscheidungen gefragt sind und die Problemgröße groß ist.

Komplexere Heuristiken, die unter den Rahmen von Meta-

heuristiken fallen, bieten eine weitere Möglichkeit, effektive Lösungen zu finden. Simulated Annealing und genetische Algorithmen sind zwei prominente Beispiele. Simulated Annealing ist von den physikalischen Prozessen des Abkühlens inspiriert und erlaubt gelegentliche Verschlechterungen der Lösung, um lokale Minima zu vermeiden und so bessere globale Lösungen zu finden. Dieser Ansatz hat sich in vielen Optimierungsproblemen als sehr effektiv erwiesen. Genetische Algorithmen basieren auf den Prinzipien der natürlichen Selektion und Evolution. Sie arbeiten mit einer Population von Lösungen, die durch Kreuzung und Mutation neue Lösungen erzeugen. Diese neuen Lösungen werden basierend auf ihrer Fitness bewertet, und die besten Lösungen werden ausgewählt, um die nächste Generation zu bilden. Genetische Algorithmen sind besonders nützlich für komplexe Probleme mit großen Lösungsräumen, da sie durch ihre stochastischen Elemente eine breite Exploration des Lösungsraums ermöglichen.

Zusätzlich zu diesen Ansätzen könnten auch hybride Methoden und maschinelles Lernen in Betracht gezogen werden, um die Lösungsqualität weiter zu verbessern. Hybride Methoden kombinieren verschiedene Optimierungstechniken, um die Stärken jeder Methode zu nutzen und ihre Schwächen zu minimieren. Maschinelles Lernen kann verwendet werden, um Muster in den Daten zu erkennen und darauf basierend bessere Heuristiken zu entwickeln.

Insgesamt bieten diese verschiedenen algorithmischen Ansätze eine breite Palette an Möglichkeiten zur Lösung des Problems der effizienten Auslagerung in Kompaktlagern. Die Wahl des geeigneten Ansatzes hängt von den spezifischen Anforderungen und Beschränkungen des jeweiligen Lagersystems ab.

## VII. SCHLUSSFOLGERUNGEN UND ZUKÜNFTIGE FORSCHUNG

In diesem Abschnitt präsentieren wir weitere Forschungsthemen, die eine algorithmische Untersuchung erfordern, um die Effizienz und Effektivität von Lagersystemen weiter zu verbessern.

Ein bedeutendes Forschungsthema im Bereich der Lagerlogistik ist das sogenannte Housekeeping, insbesondere während Zeiten, in denen keine Aufträge eingehen, wie beispielsweise nachts. Das Hauptziel besteht darin, das Lager effizient zu reorganisieren, um eine vordefinierte Zielkonfiguration zu erreichen. Hierbei stellt sich die Frage, wie das Lager umorganisiert werden kann, sodass zukünftige Auslagerungen möglichst energie- und zeiteffizient durchgeführt werden können. Dies erfordert die Entwicklung von Algorithmen, die in der Lage sind, optimale oder zumindest sehr gute Umorganisationen des Lagers zu berechnen. Diese Algorithmen müssen eine Vielzahl von Faktoren berücksichtigen, darunter die aktuelle Lagerkonfiguration, die erwarteten zukünftigen Anforderungen sowie die physikalischen Beschränkungen des Lagersystems. Die Herausforderung besteht darin, eine Balance zwischen den unmittelbaren Kosten der Umorganisation und den langfristigen Einsparungen bei den Auslagerungsvorgängen zu finden.

Ein weiteres relevantes Forschungsthema ist die Bestimmung einer optimalen Zielkonfiguration des Lagers. Dies kann durch die Einführung von Stochastizität erreicht werden. Wenn für jeden Container eine Wahrscheinlichkeitsverteilung für seine zukünftige Nutzung bekannt ist, kann die Güte einer Konfiguration quantifiziert werden. Das Ziel besteht darin, eine Lagerkonfiguration zu finden, die die erwarteten zukünftigen Zugriffe minimiert.

Diese Forschungsthemen bieten spannende Möglichkeiten für die Weiterentwicklung der Lagerlogistik. Die Entwicklung neuer Algorithmen und Modelle zur Verbesserung der Lagerorganisation und -effizienz kann zu erheblichen Einsparungen und Leistungssteigerungen führen. Es bleibt jedoch noch viel Arbeit zu tun, um diese Herausforderungen zu bewältigen und praktische Lösungen zu finden, die in realen Lagerumgebungen angewendet werden können. Die kontinuierliche Zusammenarbeit zwischen Forschern und Praktikern wird entscheidend sein, um die entwickelten Theorien und Modelle in die Praxis zu überführen und deren Wirksamkeit zu demonstrieren.

## ACKNOWLEDGMENT

Die Autor:innen bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. Das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr – ist ein von den Universitäten der Bundeswehr Hamburg und München gemeinsam getragenes wissenschaftliches Zentrum und Bestandteil des Konjunkturprogramms der Bundesregierung zur Überwindung der COVID-19- Krise. Es unterliegt der akademischen Selbstverwaltung. Die Mittel, mit dem das dtec.bw ausgestattet wurde, werden an beiden Universitäten der Bundeswehr zur Finanzierung von Forschungsprojekten und Projekten zum Wissens- und Technologietransfer eingesetzt. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] N. Fauvé and S. Neumann, "Storage and Retrieval in Fully Automated Grid-Based Storage Systems." *International Conference on Operations Research*. Cham: Springer International Publishing, 2022.
- [2] N. Fauvé, F. Jaehn, M. Azizpour, B. Hein, A. Kirchheim, and S. Neumann, "Technische und methodische Analyse von automatisierten Kompaktlagern." *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg*, 2022.
- [3] M. Fliedner, J. Golak, Y. Gül, and S. Neumann, "Order Retrieval in Compact Storage Systems" *Arbeitspapier*, 2024.

# Drone-Aided Electrical Current Sensing – A Breakthrough for Smart Grids

Khaled Osmani\*, Detlef Schulz

*Department of Electrical Engineering*

*Helmut Schmidt University/University of the Federal Armed Forces*

Hamburg, Germany

\*alosmani.k@hsu-hh.de

**Abstract** – This paper serves as an overview of the current progress concerning the DNeD (Digitalized, legally safe and low-emission airborNe inspection and grid data acquisition using automated Drones) project. DNeD aims to provide a full solution of an intelligent current sensor, integrable inside a dedicated drone-embedded sensor box. The research began with a brief survey of non-invasive current sensing technologies as a guide for the selection of the optimum current sensing physics. Afterwards, an all-optical sensor (MICATU RG235) was tested under laboratory conditions. The results from these tests were systematically tabulated. Concurrently, a study on magnetic field-based current sensing was performed using Finite Element Analysis (FEA) in COMSOL. This dual approach allows for a comprehensive evaluation of different sensing technologies. Ultimately, the study concludes that magnetic-based sensing, utilizing Hall effect sensors (e.g. DRV5055), is by far the most suitable approach. Additionally, the design of the sensor-box housing considered crucial mechanical factors such as robustness and weight. The design process utilized CAD models developed in SolidWorks, which form the backbone for real sensor-box manufacturing. Overall, the circuitry suggestions consisted of the development of analogue modulation electronic circuits, incorporating low pass filters, voltage dividers, and additional protective circuitry to ensure the reliability and accuracy of the sensor data. A prototype was developed, featuring Bluetooth as the communication protocol between the sensor-box and a remote system running MATLAB. Preliminary results indicate that the prototype effectively transfers dummy analogue voltage signals over distances up to 30 metres, demonstrating the feasibility of the design for practical applications. This comprehensive study not only addresses the mechanical and physical design aspects of embedding an intelligent sensor in a drone but also explores various current sensing technologies and their practical implementation. The findings provide a robust framework for future development and optimization of intelligent sensor systems in aerial applications.

**Keywords** – Drones, intelligent sensors, smart grids, remote surveillance

## NOMENCLATURE

CAD	Computer Aided Design
DNeD	Digitalized, legally safe and low-emission airborNe inspection and grid data acquisition using automated Drones
FEA	Finite element analysis
MFs	Magnetic Fields
TL	Transmission line

## I. INTRODUCTION

The increasing demand for electrical energy necessitates a more flexible and robust electric grid to meet the evolving power requirements. The load demands of the grids are nowadays governed by new technological sectors such as electric mobility: the rise of electric vehicles and their continuous and systematic need for recharging creates a significant new load on the grid. Higher energy consumption is additionally manifested by the electrification of industrial processes as industries experience a transition from fossil fuel-based to electric power systems. The substantial amounts of power for data centres and cloud computing also contribute to the amplification of electric demands. Moreover, the increased adoption of electric cooling/heating systems in residential and commercial areas contributes to the increased power demand as well. As a result, critical challenges for power infrastructures are created. One significant obstacle is the excessive electrical load demand, straining the capacity of the existing grids. In other terms, electrical grids are pushed closer to their full potential. Such a strain leads to the burnout of electrical transmission networks as they are forced to full-capacity operation more frequently. Hence, the likelihood of failures is augmented due to the accelerated wear and tear resulting from the constant high demands. With these new challenges, conventional electric grids have become less reliable and more prone to breakdowns and frequent blackouts. Furthermore, resultant voltage sags and swells with their destructive effects on loads lead to further economic costs and inefficiencies. The new trend of smart grids is partially induced by such consequences. In the light of the escalating demands, there is a need to modernize and add intelligence to conventional electrical grids to ensure a stable and reliable power supply. Hence, smart grids can achieve better load management, real-time monitoring, and better reactions to prevent power disruptions. The digitalization of grids allows for a dynamic response towards modern energy consumption.

There are strategies for the transition to smart grids that aim to achieve this goal: smart energy meters, distribution automation, and microgrids [1]. On the one hand, smart meters provide real-time data, allowing thus for better energy consumption monitoring and energy costs savings. Additionally, smart meters enable two-way communications which improves load balancing and reduces downtime from outages. The disadvantages of smart meters are their high cost of installation, their need for extra infrastructure, and data privacy threats. Distribution automation on the other hand, ameliorates conventional grids by automating faults detection and service

restoration but requires a significant initial investment in infrastructure (e.g. sensors implementation, control systems, skilled labour, etc.). From a different perspective, microgrids enhance the overall grid’s resiliency by providing localized power generation and continuous critical loads supply during outages. In effect, microgrids reduce downtime and transmission losses due to decentralized power generation. Even so, microgrids are implemented with high complexities, along with their dependencies for additional control, monitoring equipment, and technical challenges in coordinating multiple interconnected systems [1]. By addressing the identified research gap in the race towards smart grids, the “Digitalized, legally safe and low-emission airborNe inspection and grid data acquisition using automated Drones (DNeD)” project aims to employ a drone as a messenger for an intelligent non-invasive electric sensor on overhead transmission lines. In other terms, a sensor-box, embedding an intelligent sensor consisting of a payload for a drone, can be deployed on transmission lines, allowing thus for a remote monitoring of electrical current passing through the transmission line (TL). Eventually, data about the TL, acquired by the deployed sensor box, can be dynamically processed and visualized by a human operator at any time. Therefore, the suggested DNeD project adds a digital trace over conventional analogue electric grids, enhancing the grids’ operation and maintenance by:

- enabling remote monitoring of overhead TL by providing real-time data (e.g. instantaneous measured current),
- establishing a mathematical framework for predictive data maintenance as maintenance needs are anticipated with leveraged data analytics,
- facilitating predictive extrapolation of the grid’s status so that human operators can better anticipate potential issues and plan pre-emptive actions accordingly,
- enhancing the overall grid’s performance as fault occurrences are minimized, hence improving reliability and efficacy of electrical distribution.

The rest of this paper consists as follows: Section 2 presents a short overview about the physics of non-invasive current sensing. Section 3 outlines a laboratory test of MICATU RG235 with tabulated measurements outcomes. Section 4 highlights a Finite Element Analysis (FEA) of Magnetic Fields (MFs) radiated around electrical transmission lines with complex stranded geometries. Section 5 presents the Computer Aided Design (CAD) of the suggested sensor box. Section 6 showcases the first trials of the proposed hardware implementation of the DNeD sensor. Section 7 highlights the suggested future work and conclusions.

## II. PHYSICS OF NON-INVASIVE CURRENT SENSING

In conjunction with the DNeD project description, the intelligent current sensor should be non-invasive as it will never have any physical connection with the TL. Since the drone shall deploy the sensor box onto the TL, the embedded current sensor within is hence expected to captivate data of the electrical current flow in the TL, at a distance. The design should also take into consideration the limitations caused by the sensor box itself (i.e. produced MFs can be limited in front of the embedded sensor due to the physical barrier of the sensor box). With that being said, the surveyed physics for remote current sensing during the realization of the DNeD project are mainly classified as either magnetic or optical based. For magnetic

based sensors, they utilize Ampere’s circuital law, relying on the generated magnetic flux in the circuit: this law relates the MF around a closed loop, to the flowing current within, as expressed in the integral form of Formula (1).

$$\oint B \cdot dl = \mu_0 \iint j \cdot dS = \mu_0 I \quad (1)$$

With  $\oint B \cdot dl$  representing the line integral of the magnetic field  $B$  around a closed loop,  $\mu_0$  the free space permeability,  $j$  the current density, and  $I$  the total electric current passing through the loop [2]. With the proportionality of the MF around a conductor loop with respect to the current through the loop, electrical currents can be measured non-invasively. An example of MF-based sensors with the above-described physics can be tunnel magnetic junction, magneto-resistor with a differential Rogowski coil, three-axis MF vertical sensor array, hall effect sensors with coupling matrix, vertical MF sensors array based on air-core coils, and circular array of MF sensors. Alternatively, optical-based current sensing utilizes the Faraday Magneto-Optical effect: it describes the rotation of the light polarization’s plane, when it passes through a magneto-optical material, under the influence of an active MF. In other terms, the rotation angle of the polarized light is proportional to the strength of the magnetic field and the length of the material through which the light travels, as expressed in Formula (2).

$$\varphi = \mu \cdot V \cdot \int H dl \quad (2)$$

Such that  $\mu, V$  represent the magneto-optical material’s relative permeability and Verdet constant respectively, and  $H$  the magnetic field intensity generated by the measured current [2]. Accordingly, an applied current can be calculated (thus measured experimentally) indirectly after knowing its resulting magnetic field, which proportionally modifies the amount of the polarization’s rotation. For the described optical physics an example consists of a fibre optic sensor with a polarization splitter and a Faraday rotator: incident light is divided into two polarization states, by means of the polarization beam splitter, where the final state of the polarization reflects the flowing current proportionally. The selection of the sensor’s physics took into consideration key factors related to the actual sensor’s power consumption, sensitivity, resolution, percentage of error, shape, size, etc. with a general comparative assessment as presented in Table 1.

TABLE I: COMPARATIVE ASSESSMENT OF MAGNETIC AND OPTICAL SENSORS.

<i>Sensor’s physics</i>		
	<i>Optical</i>	<i>Magnetic</i>
Pros	No flowing electrons Dynamic current range Fail-safe	High sensitivity Ease of integration Compact size
Cons	Susceptible to optical interference Limited historical data Higher cost	Drift over time Temperature sensitive Magnetic interference

Both physics of Table 1 share the remote ability for current sensing (i.e. without a direct contact with the conductor) hence their applicability in the DNeD project.

In regard to optical based sensors the considerations were taken with respect to their high immunity versus electromagnetic interference and accurate readings in environments with high electrical noise. Additionally, optical based sensors minimize the risk of fire-sparking with an enhanced safety against heat induced failures. Along with their fail-safe mechanism optical sensors sustain the safety of the sensor box. Therefore, it is decided to begin with an optical sensor as the data source of TL's electrical currents in the DNeD project.

### III. MICATU RG235 LABORATORY TEST

As per previous recommendations, the all-optical RG235 sensor from MICATU, Inc. is tested in the high voltage laboratory of the Helmut Schmidt University/University of the Federal Armed Forces. The conducted tests to investigate its suitability of application inside the DNeD's sensor box began with a high-voltage setup. Forced voltage values ranged from 1.48 kV to 20.464 kV. Additionally, RG235 was subjected to different current values to be measured from 10 A to 60 A, as presented in the montage of Figure 1 [3].

Despite that at this stage of research no intention exists to involve voltage measurement in the DNeD's sensor, the work in [3] represented a general background for any future work. As can be seen from Figure 1, the SE-CUR-CLAMP-150-DC represents a current measurement reference, to record any error in measurement in a drone-emulated (i.e. subsection of the RG235 to systematic and repeated vibrations) environment. With the error estimation Formula (3), Table 2 represents the measurements outcome of the RG235.

$$\% \text{ Error} = \frac{|(\text{Reference value} - \text{measured value})|}{\text{Reference value}} \times 100 \quad (3)$$

TABLE II: TABULATED VOLTAGE AND CURRENT MEASUREMENT DATA OF THE RG235 UNDER LABORATORY TESTS [3].

VOLTAGE MEASUREMENTS				
Ref. RMS voltage [kV]	Measured RMS voltage [kV] (no vibrations)	Measured RMS voltage [kV] (under vibrations)	Error [%] (no vibrations)	Error [%] (under vibrations)
1.542	1.530	1.500	0.77	2.72
2.573	2.563	2.646	0.39	2.84
3.762	3.719	3.851	1.14	2.37
5.199	5.228	5.026	0.56	3.33
6.385	6.469	6.553	1.32	2.63
7.946	7.969	8.161	0.29	2.71
8.641	8.692	8.414	0.59	2.63
9.875	10.035	10.129	1.62	2.57
10.318	10.241	10.596	0.75	2.69
12.847	13.049	12.516	1.57	2.58
14.726	14.920	14.965	1.32	1.62
16.203	16.273	16.058	0.43	0.89
18.912	19.120	18.611	1.10	1.59
19.624	19.996	19.419	1.90	1.04
CURRENT MEASUREMENTS				
Ref. RMS current [A]	Measured RMS current [A] (no vibrations)	Measured RMS current [A] (under vibrations)	Error [%] (no vibrations)	Error [%] (under vibrations)
10	10.11	9.71	0.89	2.90
20	20.31	20.77	1.55	3.84
30	30.62	31.01	2.07	3.37
40	40.46	41.31	1.14	3.28
50	49.66	52.12	0.68	4.24
60	60.15	60.88	0.25	0.18

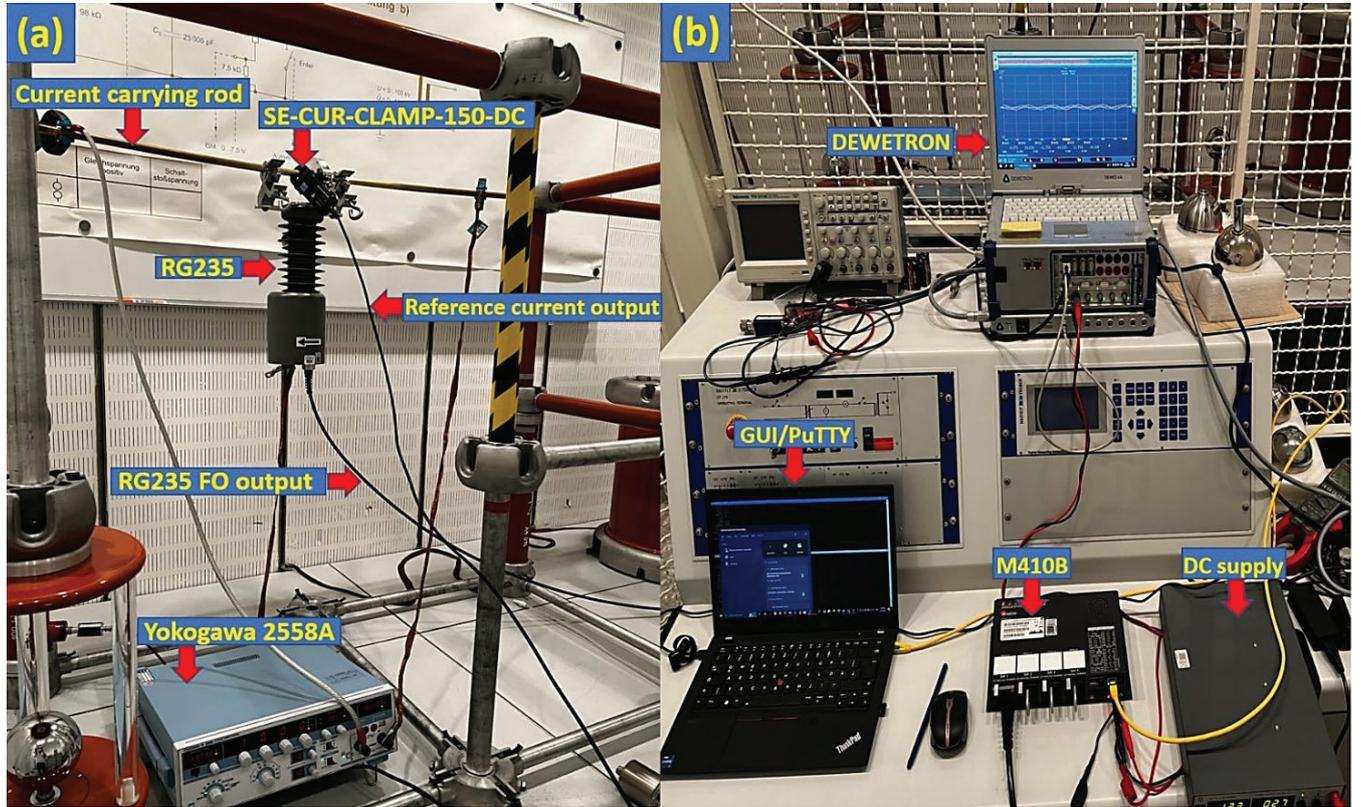


FIGURE 1: THE RG235 CURRENT MEASUREMENTS TEST SETUP: A) CIRCUITRY, B) OUTPUT ASSEMBLY [3].

A reliable performance with minimal interference risks was demonstrated by the RG235. This was mainly due to its fibre optical output. In addition, the RG235 supports internet

protocol addressing enabling thus a dynamic remote access to the grid's operational status. Despite the safe and accurate measurements offered by the RG235, its bulky design limits its implementation in the DNeD's sensor box. Although it can be disassembled in a way that its internal Faraday's electro-optical current sensor is solely installed in the destined sensor box, that still poses fragility given the harsh working conditions of the DNeD project. At this point in the investigation, the idea to change the physics of the sensing technique had formed. Accordingly, the new approved sensing physics are magnetic based (i.e. done through magnetic based sensors such as hall effect or tunnel magnetoresistance).

#### IV. FINITE ELEMENT ANALYSIS (FEA) OF MFs AROUND COMPLEX STRANDED TRANSMISSION LINES IN COMSOL

Based on the observations in [3], and in order to precisely select magnetic-based sensors, the study in [4] aimed for numerical data acquisition about current-induced MFs. In other words, the simulation in [4] guides the selection process of the magnetic based sensors' sensitivity. The geometric structure of overhead TLs (i.e. sub-filaments are redundantly stranded over each other) makes the MFs quantization around them hard to compute. Basically, for a single straight wire, the resulting MF due to current passage can be straightforwardly calculated with the numerical form of Biot-Savart law in Formula (4).

$$B = \frac{\mu_0 I}{2\pi R} \quad (4)$$

Such that  $\mu_0$  represents the free-space permeability,  $I$  the current, and  $R$  the distance from the wire to the point of observation [4]. This is not the case for stranded TLs: the stranded structure of overhead TLs, creates superimposing MF vectors of instantaneously varying norms/orientations. Therefore, it is hard to manually calculate the exact resulting MF around a stranded TL, at a specified point of observation in the space

near the TL. For this reason, the Finite Element Analysis (FEA) in [4], executed by means of COMSOL Multiphysics 6.1 was conducted. It approximates as accurate as possible, the resulting MFs around stranded TLs, composed from 7-filaments (as shown in Figure 2) to 91-filaments. Table 3 encapsulates the resulting MFs norms after current injection from 10 A to 110 A in the 7-filament TL.

TABLE III: RESULTING MFs DUE TO THE APPLICATION OF DIFFERENT CURRENT VALUES IN THE 7-FILAMENT TL [4].

Total current [A]	Current in each filament	$P(x; y; z)$			MF norm [mT]
		x [mm]	y [mm]	z [mm]	
10	1.428	16	9	188	$1.02 \times 10^{-1}$
		21	12	188	$8.13 \times 10^{-2}$
20	2.857	13	9	188	$2.37 \times 10^{-1}$
		19	13	188	$1.65 \times 10^{-1}$
30	4.285	14	11	188	$3.29 \times 10^{-1}$
		19	14	188	$2.46 \times 10^{-1}$
40	5.714	14	9	188	$4.76 \times 10^{-1}$
		18	16	188	$3.24 \times 10^{-1}$
50	7.142	13	14	188	$5.92 \times 10^{-1}$
		18	16	188	$4.13 \times 10^{-1}$
60	8.571	15	9	188	$6.65 \times 10^{-1}$
		18	15	188	$5.05 \times 10^{-1}$
70	10	13	10	188	$8.09 \times 10^{-1}$
		18	15	188	$5.84 \times 10^{-1}$
80	11.428	14	9	188	$9.32 \times 10^{-1}$
		17	15	188	$6.78 \times 10^{-1}$
90	12.857	13	10	188	1.00
		19	14	188	$7.40 \times 10^{-1}$
100	14.285	13	11	188	1.00
		17	16	188	$8.32 \times 10^{-1}$
110	15.714	13	10	188	1.00
		16	16	188	$9.28 \times 10^{-1}$

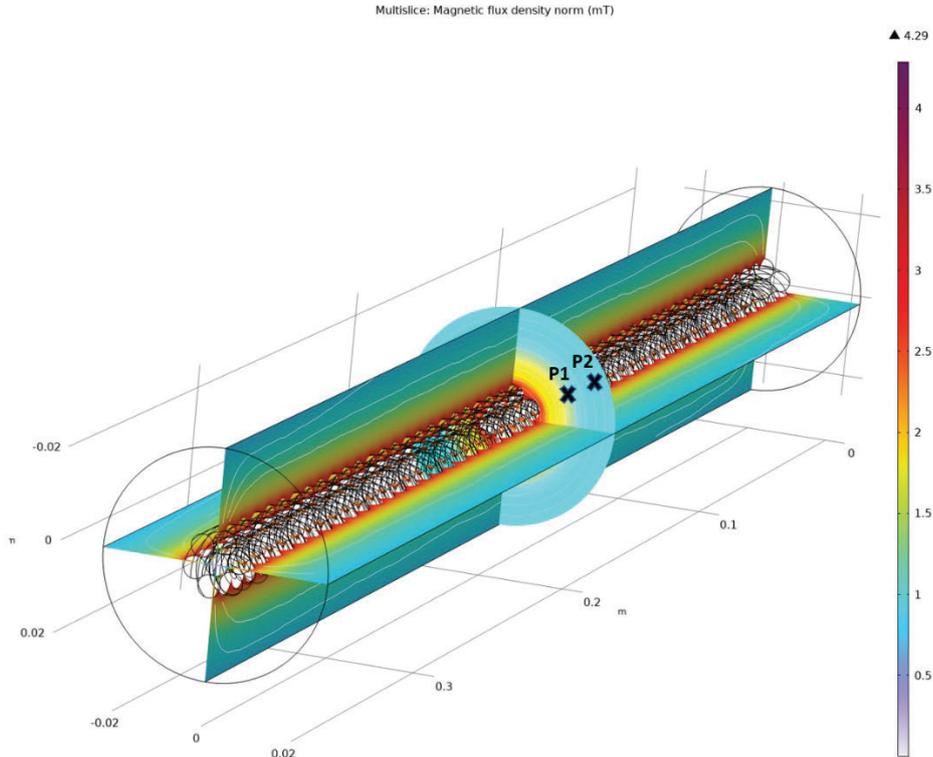


FIGURE 2: RADIATION AROUND THE SEVEN-FILAMENT TL, INJECTED WITH 100 A, WITH P1 AND P2<sup>86×10<sup>-4</sup></sup> AS THE POINTS OF OBSERVATION [4].

As an exemplary case of the 7-filament TL, the average resulting MF is valued at  $5.74 \times 10^{-1}$  mT. Accordingly, for this particular TL, with the two points of observation P1 and P2 as depicted in Figure 2, a magnetic based sensor with a [0 ; 1 mT] sensitivity range, should be able to detect MFs resulting from currents of 10 A to 110 A. Additionally, this study evaluated the MF sensors sensitivity ranges for larger currents (i.e. up to 1450 A in the 91-filament TL). In conclusion, the suggested MF based sensors can be designed for a large current range, which makes the application of the sensor box of Figure 3 more reliable.

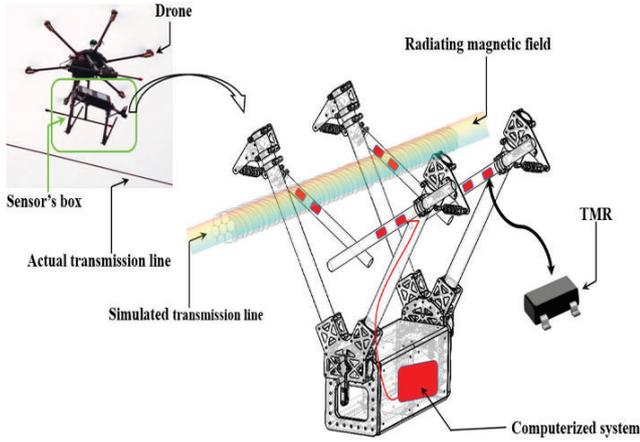


FIGURE 3: THE DNEd's PROJECT OVERVIEW [4].

## V. SENSOR BOX CAD

With the recommended sensor workflow in [4], the sensor box should be manufactured based on certain design criteria to withstand various working conditions (e.g. environmental, operational integrity, etc.). With that being said, the sensor box should be lightweighted, robust, and non-magnetic. This is to avoid any interference with the MFs around the TLs. For these reasons, this part mainly accounts for the sensor box appropriate material and geometric design. Beginning with the sensor box material, it should have a high ingress protection to protect against water and dust penetration. Moreover, the sensor box should have a low magnetic permeability (Formula (5)) to allow the penetration of the MF into its internal space (where the MF sensor is installed) without any interactions.

$$\mu = \mu_0(1 + \chi_v) \quad (5)$$

Such that  $\mu$  represents the magnetic permeability of the material,  $\mu_0$  the vacuum permeability, and  $\chi_v$  the volume magnetic susceptibility [5]. Following the same perspective, the sensor box material must adhere to a robust structure. Since it is destined to be repeatedly laid onto/off the TL, it must present a high compressive strength, as approximated by the point on the engineering stress-strain curve ( $\sigma_e^*$ ,  $\varepsilon_e^*$ ) defined in (6) and (7).

$$\sigma_e^* = F^*/A_0 \quad (6)$$

$$\varepsilon_e^* = l^* - l_0/l_0 \quad (7)$$

With  $\sigma_e^*$  representing the engineering stress,  $\varepsilon_e^*$  the engineering strain,  $F^*$  the load applied before the material's crashing,  $A_0$  the material's original area,  $l^*$  the material's length before crashing, and  $l_0$  the material's original length. Likewise, the optimum material should adapt to a high compressive

strength (i.e. to withstand momentary axial loads without deforming/breaking) and a high tensile strength (i.e. ability to stay solid under pulling/stretching). Ultimately, in addition to a high ingress protection, low weight, high compressive/tensile strengths, the material should present high tensile modulus (i.e. rigidity) as represented in Formula (8).

$$E = \sigma/\varepsilon \quad (8)$$

With  $\sigma$  representing the compressive strength and  $\varepsilon$  the axial strength. By taking into account the mathematical model depicted by the Formulas (5)-(8), the chosen material for the sensor box construction is Carbon Fiber. From another side, the box' geometry must be easily integrable inside the drone. It should also attach securely to TL and resist environmental forces such as rain and wind. Subsequently, the designed sensor box consisted of a rectangular base and triangular fasteners, guaranteeing hence a stable attachment over the TL, and a secure fitting inside the drone as portrayed in Figure 4.

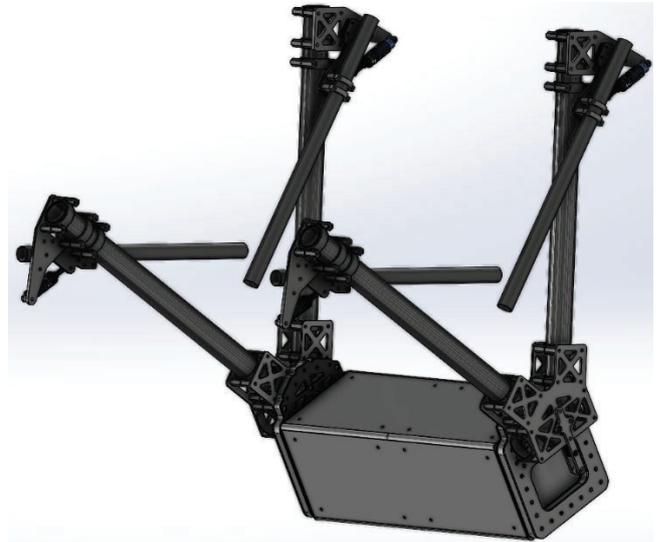


FIGURE 4: SENSOR BOX 3D FULL DESIGN WITH CARBON FIBRE AS CONSTRUCTING MATERIAL.

## VI. FIRST TRIALS OF THE SUGGESTED HARDWARE

After the acknowledgement of the suggested current sensing physics (i.e. magnetic based), in addition to the recommended sensor box materials (i.e. carbon fibre) and geometry (i.e. Figure 4), the hands-on prototype realization can be started. First and foremost, initial tests can be manifested through a programmable DC voltage source as the source of information for flowing current, instead of applying a real MF sensor. Since an MF sensor would produce a voltage proportional to the sensed MF, the trial here is to only utilize dummy voltages to proceed with further tasks (e.g. communication, etc.). Such a DC voltage is to be processed via a small computerized system (e.g. Arduino Due), to be then processed wirelessly (i.e. via Bluetooth) to a MATLAB environment. This montage can be graphically expressed as in Figure 5. The variable DC voltage source instead of a real MF sensor (e.g. DRV5055) bypasses any electronic-based problems, mainly related to impedance and grounding. Therefore, the implementation began with code programming in the Arduino IDE as



FIGURE 5: BLOCK DIAGRAM REPRESENTATION OF THE DNeD SENSOR FIRST-STEPS HARDWARE IMPLEMENTATION.

well as in MATLAB, to send/receive the sensed voltages with the corresponding analogue to digital conversions. The actual montage is hence composed of a real Arduino Due board, two Bluetooth modules (HC-05), a DC power supply, and variable DC voltages to be inputted onto the analogue inputs of the Arduino, as shown in Figure 6. From the receiver side (i.e. a com-

## ACKNOWLEDGEMENT

This research paper is part of the project DNeD (“Digitalisierte, rechtssichere und emissionsarme flugmobile Inspektion und Netzdatenerfassung mit automatisierten Drohnen”, engl. “Digitalized, legally safe and low-emission airborNe inspec-

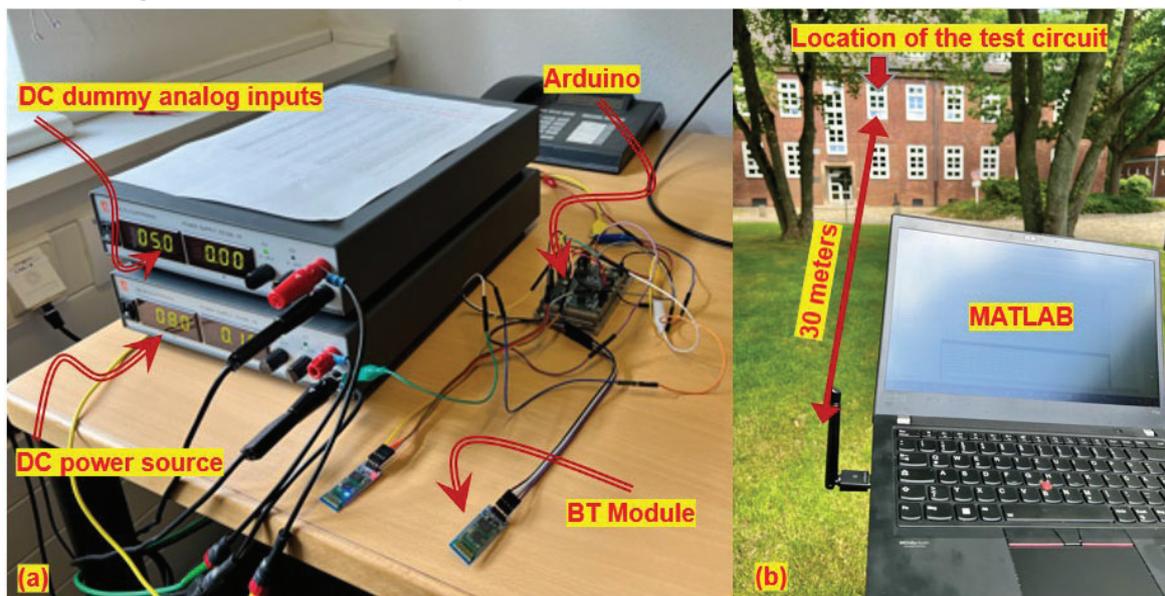


FIGURE 6: DNeD'S PROTOTYPE FIRST TRIAL: A) TEST CIRCUIT, B) RECEIVER SIDE.

puter with installed MATLAB), a Bluetooth range extender is needed to ensure a proper data transmission through longer distances (i.e. mimicking the distance between the sensor box installed on the overhead TL and the base station). The novelty behind codes implementation (instantaneous curve drawing of the sensed voltage norms), in both Arduino IDE as well as MATLAB, reflects the uniqueness behind the sensor workflow. The first trials revealed a successful data transfer between the test circuit (Figure (6, a)) and the receiver (Figure (6, b)), over a range of 30 metres.

## VII. CONCLUSION AND FUTURE WORK

The first outcomes of the test circuit with the receiver part showed encouraging results. The dummy voltages were able to be drawn instantaneously in MATLAB, over a distance of 30 metres with physical obstacles (e.g. wall, window). Next steps consist of implementing a real MF sensor such as the DRV5055 instead of the fixed DC supply, followed by filters and additional protection circuitry to ensure safe voltage values inputs at the Arduino analogue pins. The printed circuit board design shall be implemented inside the real Carbon Fiber sensor box to check how the range of 30 metres can be affected and react accordingly. Upon the full completion of the overall sensor design, a digital trace can be significantly added over TLs, allowing an actual sense of remote monitoring and analysis for electric grids.

tion and grid data acquisition using automated Drones”) and was funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

## REFERENCES

- [1] S. S. Mavuri, J. Nakka and A. Kotla, "Interconnected Microgrids: A Review and Future perspectives," 2022 IEEE 2nd International Conference on Sustainable Energy and Future Electric Transportation (SeFeT), Hyderabad, India, 2022, pp. 1-7, doi: 10.1109/SeFeT55524.2022.9908988.
- [2] A. Ghasemi, "Chapter 1 – Fundamentals of ferrites“, *Magnetic Ferrites and Related Nanocomposites, Micro and Nano Technologies*, 2022, pp. 301-367.
- [3] K. Osmani, M. F. Meyer, F. Grumm and D. Schulz, "Toward Smarter Grids: Experimental Investigation of an “All-Optical” Electrical Sensor Under Laboratory Conditions," in *IEEE Access*, vol. 11, pp. 108418-108432, 2023, doi: 10.1109/ACCESS.2023.3321219.
- [4] K. Osmani and D. Schulz, "Modeling Magnetic Fields around Stranded Electrical Transmission Lines via Finite Element Analysis (FEA)," *Energies*, vol. 17, no. 4, pp. 801–801, Feb. 2024, doi: https://doi.org/10.3390/en17040801.
- [5] C. J. Bonin and F. J. Bonetto, "A Low-Cost and Accurate Method for Measuring Weak Magnetic Susceptibilities," in *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-10, 2022, Art no. 6002510, doi: 10.1109/TIM.2022.3165274.

# Ein formales Modell für Anwendungen Künstlicher Intelligenz in Automatisierungssystemen

Marvin Schieseck, Philip Topalis, Felix Gehlhoff

*Institut für Automatisierungstechnik*

*Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg*

Hamburg, Deutschland

marvin.schieseck@hsu-hh.de

Alexander Fay

*Lehrstuhl für Automatisierungstechnik*

*Ruhr Universität Bochum*

Bochum, Deutschland

alexander.fay@rub.de

**Zusammenfassung**—Die Integration von Künstlicher Intelligenz (KI) in Automatisierungssysteme wird durch mangelnde Standardisierung im Bereich der Dokumentationen von KI-Anwendungen behindert. Diese Arbeit schlägt ein formales Modell vor, das auf Standards und Ontologien basiert, um eine klare und strukturierte Dokumentation von KI-Anwendungen in Automatisierungssystemen zu ermöglichen. Das vorgeschlagene Informationsmodell für Künstliche Intelligenz in Automatisierungssystemen nutzt Ontologie-Design-Patterns, um verschiedene Aspekte von Automatisierungssystemen und KI-Software abzubilden und zu verknüpfen. Evaluiert durch ein praktisches Beispiel, zeigt das Modell seine Effektivität bei der Verbesserung der Dokumentation und bei der Unterstützung der nachhaltigen Integration von KI in Automatisierungssysteme.

**Index Terms**—Künstliche Intelligenz, Ontologien, Semantic Web, OWL, Modellierung

## I. EINLEITUNG

In den letzten Jahren hat die Integration von Anwendungen der Künstlichen Intelligenz (KI) in technische Systeme wie Automatisierungssysteme erheblich an Dynamik gewonnen [1]. Zahlreiche Anwendungsfälle zeigen bereits, dass durch die Einbindung von KI neue Lösungen ermöglicht oder die Effizienz bestehender Lösungen gesteigert werden konnte [2]. Beispiele hierfür finden sich in den Bereichen Wartung, Qualitätskontrolle und Planung [3]. Trotz dieser vielversprechenden Einsatzmöglichkeiten bleibt die Verbreitung von KI in verschiedenen Industriezweigen relativ gering [4].

Häufig werden KI-Anwendungen im Rahmen von Forschungsprojekten entwickelt und lassen sich wirtschaftlich nicht in die Praxis umsetzen. Dies liegt teilweise daran, dass bestimmte Aspekte von KI-Software – wie die Qualität und Verwaltung der Daten, das Management der Rechenressourcen, die Vertrauenswürdigkeit und die Dokumentation der Software – im Vergleich zu herkömmlicher Automatisierungssoftware besonders herausfordernd sind und die wirtschaftliche Integration und den Betrieb erschweren [4]–[7].

Obwohl einige Aspekte, wie die Datenqualität und Datenverwaltung sowie die Vertrauenswürdigkeit, in den letzten Jahren umfangreich erforscht wurden, wird die Softwaredokumentation von KI-Anwendungen und deren

Auswirkungen auf das Gesamtsystem in verschiedenen Szenarien oft vernachlässigt [6], [7]. Gleichzeitig gibt es einen globalen Trend zur Schaffung neuer gesetzlicher Vorschriften für KI, wie am Beispiel der Europäischen Union mit dem EU Artificial Intelligence Act [8] zu sehen ist. Diese erfordert eine umfassende Analyse der Risiken von KI-Anwendungen und empfiehlt deren Dokumentation. Es kann angenommen werden, dass mit der fortschreitenden Entwicklung solcher Regulierungen eine klare und strukturierte Dokumentation von KI-Anwendungen immer wichtiger werden wird.

Zusätzliche Herausforderungen bei der Dokumentation ergeben sich aus der inhärenten Komplexität sowohl von Automatisierungs- als auch von KI-Systemen. Beide Systeme bestehen in der Regel aus mehreren Hardware- und Softwarekomponenten, die stark miteinander vernetzt und voneinander abhängig sind [9]. Darüber hinaus haben Automatisierungssysteme oft lange Lebenszyklen, die sich über Jahrzehnte erstrecken. Wenn einzelne technische Komponenten innerhalb dieser Systeme aktualisiert oder ersetzt werden, ist eine umfassende Dokumentation erforderlich. Diese kann anschließend genutzt werden, um festzustellen, wie sich Änderungen, wie der Austausch von Sensoren oder Steuerungen auf die KI-Anwendung auswirken. Daher ist eine detaillierte Dokumentation unerlässlich für die Implementierung und den Betrieb von KI-Anwendungen in Automatisierungssystemen.

Es bedarf einer Lösung zur Dokumentation von KI-Anwendungen in Automatisierungssystemen, welche Informationen über die Komponenten des Automatisierungssystems, die KI-Softwarekomponenten sowie den zugrunde liegenden technischen Prozess und deren Interdependenzen und Beziehungen erfasst. Insbesondere wird ein formales Modell zur Beschreibung von KI-Anwendungen in Automatisierungssystemen benötigt [10]. Deshalb wird in diesem Artikel ein formales Modell für KI-Anwendungen in Automatisierungssystemen vorgeschlagen, das auf Standards basiert und Ontologien nutzt.

## II. ANFORDERUNGSANALYSE

Die Softwareentwicklung für KI-Anwendungen ist komplexer als für traditionelle Software. Dies umfasst Aspekte wie unklare Systemgrenzen, nicht deklarierte

Datenabhängigkeiten, Konfigurationsprobleme und Veränderungen der externen Umgebung, um nur einige zu nennen. Bemerkenswert ist, dass der Code für KI-Modelle und deren Training oft nur aus wenigen Zeilen Programmcode besteht und somit einen kleinen Teil des Gesamtcodes ausmacht. Der Großteil einer KI-Anwendung bzw. ihres Programmcodes beschäftigt sich hingegen mit Aufgaben wie Automatisierung, Testen, Ressourcenmanagement, Prozessmanagement, sowie Bereitstellung, Sammlung, Speicherung, Übertragung und Verifizierung von Daten. [7]

Basierend auf den dargestellten Herausforderungen und Aspekten ergeben sich folgende Anforderungen (A) für ein formales Modell für KI-Anwendungen in Automatisierungssystemen:

**(A1) Beschreibung von Interdependenzen:** Das Modell muss Informationen und Interdependenzen zwischen Komponenten des Automatisierungssystems, KI-Softwarekomponenten und dem technischen Prozess erfassen können. Dies ist entscheidend, da KI-Anwendungen in Automatisierungssystemen typischerweise über mehrere Komponenten im technischen System verteilt sind und daher verschiedene Abhängigkeiten abdecken müssen [9].

**(A2) Semantisch eindeutig und standardisiert:** Das Modell muss semantisch eindeutig sein, damit Experten aus verschiedenen Domänen Informationen klar verstehen, dokumentieren und kommunizieren können. Um dies zu erreichen, sollte der Inhalt und die Terminologie auf Standards basieren, da diese von Expertengruppen entwickelt werden und innerhalb spezifischer Gemeinschaften einen semantischen Konsens darstellen [11].

**(A3) Formalisierte Darstellung:** Das Modell muss in der Lage sein, die Informationen formalisiert darzustellen, um Wiederverwendbarkeit zu gewährleisten und Mehrdeutigkeiten zu minimieren [10]. Die Formalisierung in Verbindung mit einem herstellerunabhängigen Austauschformat macht das Modell maschinenlesbar und erleichtert die Integration in die vielfältige Tool-Landschaft von Automatisierungssystemen und Softwaretechnik.

**(A4) Erweiterbarkeit und Anpassungsfähigkeit:** Das Modell muss erweiterbar und anpassungsfähig sein, um Änderungen in Standards, Vorschriften und Gesetzgebungen, insbesondere im sich schnell entwickelnden Bereich der KI, berücksichtigen zu können [10]. Da KI-Regulierungen länderübergreifend variieren und häufig aktualisiert werden, wird das Modell zahlreiche Anpassungen erfordern. Im Gegensatz dazu können die Teile des Modells, die Informationen über Automatisierungssysteme beschreiben, über Jahrzehnte stabil bleiben.

### III. STAND DER TECHNIK UND VERWANDTE ARBEITEN

Es wurden drei relevante Forschungsbereiche bezüglich der Dokumentation von KI-Anwendungen identifiziert: 1.) *Datenblätter*, 2.) *Ontologien* und 3.) *grafische Modellierungsansätze*. Im Folgenden werden relevante Arbeiten der jeweiligen Bereiche vorgestellt und hinsichtlich der Anforderungserfüllung analysiert.

#### A. Informelle Dokumentation mit Datenblättern und Karten

Ein Bereich der Forschung konzentriert sich auf die Erstellung von *Datenblättern*. Diese sind kurze Dokumente mit vordefinierten Kategorien, um relevante Informationen über die KI-Anwendung zu erfassen. Um die relevanten Informationen zu erhalten, verwenden die Autoren eine Reihe vordefinierter Fragen, die von den Entwicklern der KI-Anwendung beantwortet und niedergeschrieben werden müssen. [12]–[15]

Zum Beispiel schlagen Gebru et al. [12] einen Ansatz namens *Datasheets for Datasets* vor, um den Erstellungsprozess, die Verteilung, die Wartung sowie eventuelle Annahmen und Risiken eines Datensatzes zu reflektieren und zu dokumentieren. Inspiriert von den *Datasheets for Datasets* schlagen Mitchell et al. [13] ein Framework namens *Model Cards* vor, um transparente Berichterstattung über trainierte Machine-Learning-Modelle zu fördern. *Model Cards* sind kurze Dokumente, die Benchmarks zur Leistung des Modells unter verschiedenen Bedingungen dokumentieren. Arnold et al. [14] schlagen *FactSheets* vor, um das Vertrauen in KI-Dienste zu erhöhen, indem verschiedene Eigenschaften der KI-Dienste dokumentiert werden, einschließlich der Frage, ob die zur Entwicklung der Dienste verwendeten Daten durch *Datasheets* begleitet werden. Ähnlich schlagen Lavin et al. [15] einen Prozess vor, um die Dokumentation robuster, zuverlässiger und verantwortungsvoller Machine-Learning-Systeme durch sogenannte *Technology Readiness Level Cards* (TLR-Karten) sicherzustellen.

Die Ansätze mit *Datenblättern und Karten* beschränken sich lediglich auf die Beschreibung der KI-Komponenten und der Daten. Eine Dokumentation des technischen Systems, des technischen Prozesses sowie deren Interdependenzen sind nicht vorgesehen (A1). Trotz der Organisation in vordefinierte Kategorien und Fragen erfolgen die Antworten größtenteils im Freitextformat (A3). Folglich hängen Form und Genauigkeit von der Sorgfalt und Erfahrung der Befragten ab [13]. Die Blätter und Karten orientieren sich nicht an Vokabularen und Begriffen von Standards oder Vorschriften (A2). Ihre einfache Struktur erleichtert jedoch eine einfache Erweiterung und Anpassung an verschiedene Anwendungsfälle oder Domänen (A4).

#### B. Formale Dokumentation mit Ontologien

Das zweite Forschungsgebiet konzentriert sich auf die formale Dokumentation von KI-Anwendungen mithilfe von Ontologien anstelle von Datenblättern. Dieses Forschungsgebiet lässt sich in zwei Hauptbereiche unterteilen: 1.) Dokumentation des Entwicklungsprozesses und 2.) Dokumentation der Modelle. [16], [17]

Der erste Bereich konzentriert sich auf Ontologien zur Unterstützung der Erstellung und Dokumentation des zugrunde liegenden Entwicklungsprozesses, wie von Sinha et al. [16] zusammengefasst. Zum Beispiel schlagen Panov et al. [18] eine Ontologie namens *OntoDM-KDD* vor, um den Wissensentdeckungsprozess basierend auf dem *Cross*

*Industry Standard Process for Data Mining* (CRISP-DM) darzustellen. Eine weitere Ontologie, die *Data Mining Optimization Ontology* (DMOP) von Keet et al. [19], vereinfacht Entscheidungsfindung und optimiert die Leistung von Data-Mining-Prozessen. Der zweite Bereich fokussiert auf Ontologien, die die Dokumentation von Modellen und Algorithmen unterstützen, insbesondere im Bereich des maschinellen Lernens (ML), wie von Sinha et al. [17] durchgeführt. Zum Beispiel dient die *Exposé*-Ontologie, vorgeschlagen von Vanschoren und Soldatova [20], dazu, Data-Mining-Experimente zu modellieren. Esteves et al. [21] schlagen eine Ontologie namens *MEX Vocabulary* vor, die entwickelt wurde, um ML-Experimente zu beschreiben. Publio et al. [22] entwickeln die *ML-Schema*-Ontologie zur Darstellung und Interaktion von Informationen über ML-Algorithmen, Datensätze und Experimente. Diese auf Ontologien basierenden Ansätze für ML dienen als Grundlage für große Community-Plattformen wie OpenML<sup>1</sup> zum Informationsaustausch über ML-Ansätze [23].

Die auf *Ontologien* basierenden Ansätze sind von Natur aus formalisiert (A3). Die meisten Ansätze verwenden die *Web Ontology Language*<sup>2</sup> (OWL) als herstellerunabhängiges Format (A3). Weder die Ontologien, die sich auf Data-Mining-Prozesse konzentrieren, noch diejenigen, die sich auf ML-Algorithmen konzentrieren, berücksichtigen den technischen Prozess, das technische System, die KI-Komponenten und deren Interdependenzen (A1). Darüber hinaus orientieren sich die Ontologien nicht an Standards oder Vorschriften (A2). Die meisten Ontologien sind zudem untereinander semantisch inkonsistent. Sie bestehen aus großen, monolithischen Blöcken mit diversen Klassen und Beziehungen, was in der Praxis Herausforderungen mit sich bringt (A4), obwohl sie in der Theorie einfach erweiterbar und anpassbar sind.

### C. Semiformale Dokumentation mit grafischen Modellen

Das dritte Forschungsgebiet konzentriert sich auf den Einsatz von grafischen Modellierungsansätzen. Ein grafisches Modell kann den Kern eines Problems auf eine semiformale und verständliche Weise für verschiedene Stakeholder und Experten darstellen. Ein spezifischer Ansatz für die grafische Darstellung von KI-Systemen wurde von Kaymakci et al. [24] vorgestellt, der sich darauf konzentriert, die Beziehungen zwischen Datenquellen, Datenzielen und Datenverarbeitungskomponenten darzustellen. Auf dieser Grundlage wurde die *Graphical Modeling Language for Artificial Intelligence Applications in Automation Systems* (GML-AIAAS<sup>3</sup>) [25] vorgeschlagen, die KI-Systeme innerhalb von Automatisierungssystemen darstellen kann. Unter Verwendung der Systemtechnik-Prinzipien, wie sie von Haberfellner et al. [26] beschrieben wurden, kategorisiert diese Modellierungssprache das gesamte System in drei primäre Elementtypen: 1.) Systemkomponenten,

2.) Systemfunktionen und 3.) Systembeziehungen. Die Beziehungen (z. B. Kommunikation, Zuweisung und Produkt-/Material-/Energiefluss) etablieren die Verbindungen zwischen den Systemkomponenten (z. B. Sensor, Aktor, Controller, Cloud) und den Systemfunktionen (z. B. Trainieren, Aufzeichnen, Speichern, Inferieren). Um diese Elementtypen eindeutig zu repräsentieren, definiert die GML-AIAAS eine Reihe von Symbolen und eine Syntax. Zusätzlich existiert für GML-AIAAS ein Metamodell<sup>3</sup> zur weiteren Formalisierung.

Der Ansatz von Schieseck et al. [25] ermöglicht es, Interdependenzen zwischen technischem System, technischem Prozess und KI-Funktionen durch eine definierte Symbolik und Syntax darzustellen (A1). Durch die Verwendung des bereitgestellten Metamodells kann die Symbolik in ein herstellernertrales Format wie XML oder JSON transformiert und formalisiert werden (A3). Allerdings orientiert sich dieser Ansatz nicht an Standards (A2). Aufgrund der einfachen Struktur des Metamodells kann der Inhalt leicht erweitert und angepasst werden (A4).

### D. Diskussion der verwandten Arbeiten

Zusammenfassend erfüllt keiner der bestehenden Ansätze vollständig die spezifizierten Anforderungen zur Bewältigung des identifizierten Handlungsbedarfs. Die Datenblätter sind nicht formal und es mangelt an semantischer Eindeutigkeit. Die bestehenden Ontologien berücksichtigen nicht die Interdependenzen zwischen dem technischen System und dem KI-System und sind untereinander inkonsistent. Die grafischen Modelle sind nicht hinreichend formalisiert. Keiner der bestehenden vorgestellten Ansätze berücksichtigt die Terminologie oder Struktur der existierenden Standards für KI. Daher besteht Bedarf an einem neuen und anforderungsgerechten Ansatz, welcher im folgenden Abschnitt vorgestellt wird.

## IV. FORMALE BESCHREIBUNG VON INDUSTRIELLEN KI-ANWENDUNGEN

In diesem Abschnitt wird ein neues Informationsmodell für Künstliche Intelligenz in Automatisierungssystemen (*engl. Artificial Intelligence in Automation Systems*, kurz AIAS<sup>4</sup>) vorgestellt.

Es basiert auf Ontologien, da diese eine Möglichkeit bieten, semantisch angereicherte und formalisierte Informationsmodelle zu erstellen. Darüber hinaus ermöglicht die Verwendung von Ontologien in Kombination mit formalen Beschreibungssprachen des Semantic Web, wie z. B. OWL, die Nutzung von Reasoning, Querying und das Anwenden von Regelsprachen. Um sicherzustellen, dass das Informationsmodell leicht erweitert und angepasst werden kann, wird es nicht als eine einzige monolithische Ontologie ausgeführt. Stattdessen werden kleine und unabhängige Ontologien verwendet, welche als „Ontology Design Patterns“ (ODPs) bezeichnet werden [27]. Diese ODPs sind

<sup>1</sup><https://www.openml.org/>

<sup>2</sup><https://www.w3.org/OWL/>

<sup>3</sup><https://github.com/schiesem/GML-AIAAS>

<sup>4</sup><https://github.com/schiesem/aias-information-model>

auf der obersten Ebene über eine „Alignment-Ontologie“ miteinander verbunden, die als übergreifende Struktur des AIAS-Informationsmodells dient. Um die Semantik weiter zu verbessern, basiert jedes ODP auf einem Standard [11]. Die verwendeten ODPs werden in zwei Kategorien eingeteilt, um verschiedene Aspekte zu beschreiben:

- **Technisches System und Technischer Prozess:** Diese Kategorie umfasst ODPs, die Begriffe und Beziehungen enthalten, um die Struktur des technischen Systems und die zugrunde liegenden technischen Prozesse zu beschreiben.
- **Künstliche Intelligenz und Daten:** Diese Kategorie besteht aus ODPs, die Begriffe und Beziehungen enthalten, um die KI-Komponenten und deren Funktionalitäten zu beschreiben.

#### A. Beschreibung des technischen Systems und Prozesses

Die erste Anforderung (A1) betont die Notwendigkeit, die technischen Systemaspekte zu beschreiben, einschließlich der technischen Komponenten und ihrer zugrunde liegenden Prozesse. Die Konzepte, die in der formalisierten Prozessbeschreibung nach VDI 3682 [28] definiert sind, bieten einen standardisierten Rahmen zur Beschreibung verschiedener technischer Prozesse und der zugehörigen Systemstruktur. Nach der VDI 3682 beinhaltet ein Prozess die Umwandlung von Eingaben in Ausgaben, die von einem Prozessoperator mithilfe einer technischen Ressource ausgeführt wird. Diese technische Ressource führt dann den Prozess durch. In einem weiteren Sinne kann eine technische Ressource als fähig angesehen werden, spezifische Funktionen (z. B. Prozesse) zu implementieren, und kann als abstrakte Beschreibung der Systemhardware dienen. Es ist jedoch notwendig, die VDI 3682 um zwei technische Aspekte zu ergänzen, die von der VDI 3682 selbst nicht im Detail abgedeckt werden. Erstens fehlt es der VDI 3682 an Spezifität hinsichtlich verschiedener Arten von technischen Ressourcen. Die *ECLASS*<sup>5</sup> Spezifikation adressiert diese Lücke, indem sie semantisch eine Vielzahl von Klassen und Individuen technischer Ressourcen definiert, darunter Sensoren, Aktoren, Steuerungen, Computer sowie umfassendere Entitäten wie Rechenzentren oder Cloud-Services. Zusätzlich bietet die *Semantic Sensor Network Ontology*<sup>6</sup> (SSN) eine alternative Möglichkeit zur Spezifikation von Sensoren und Aktoren und stellt ein umfassendes Vokabular zur Beschreibung von Sensor- und Aktuator-Fähigkeiten bereit, falls erforderlich.

Zweitens spezifiziert VDI 3682 nicht die Kommunikation zwischen verschiedenen technischen Ressourcen. Ein weit verbreiteter Standard zur Beschreibung solcher Kommunikationen zwischen technischen Systemen wird durch die ISO 7498 [30] bereitgestellt. Dieser Standard basiert auf einem siebenstufigen Kommunikationsmodell namens *Open Systems Interconnection* (OSI) Modell. Unter Verwendung dieses OSI-Modells kann eine Kommunikationsverbindung

über die Schichten hinweg, vom physikalischen Level bis hin zum Anwendungslevel, spezifiziert werden.

#### B. Beschreibung der Künstlichen Intelligenz und Daten

Zusätzlich zu den technischen Aspekten betont Anforderung 1 (A1) die Notwendigkeit, die KI-Komponenten und die verwendeten Daten zu beschreiben. Die ISO 22989 [31] bietet geeignete Konzepte und Terminologie für KI. Dieser Standard ist um drei übergeordnete Aspekte strukturiert: 1.) KI-Komponenten und KI-Funktionen, 2.) KI-Algorithmen und 3.) Daten. Aufgrund der Komplexität der ISO 22989 ist es hilfreich, sie in verschiedene Betrachtungsweisen für jeden Aspekt zu unterteilen.

Im ersten Aspekt werden verschiedene Arten von KI (z. B. symbolisch oder subsymbolisch), Aufgaben (z. B. Klassifikation, Clustering, Regression, Generierung), Systemdesigns (z. B. Cloud, Edge, Hybrid), deren Komponenten und Funktionen (z. B. Datenaufbereitung, Training, Validierung, Evaluation, Inferenz) definiert.

Der zweite Aspekt beschreibt die Semantik und Beziehungen von maschinellen Lernmodellen, maschinellen Lernalgorithmen und Lernarten in Bezug auf die zuvor definierten Funktionen. Zusätzlich werden Modellparameter und Hyperparameter festgelegt.

Der dritte Aspekt beschreibt die Semantik und Beziehungen zwischen Daten und Datensätzen. Neben der Definition der Daten selbst werden auch die Funktionen des Speicherns von Daten in einem Datenspeicher und der Abruf von Daten aus einer Datenquelle beschrieben. Darüber hinaus werden verschiedene Arten von Datensätzen definiert, wie z. B. Trainings-, Evaluierungs-, Produktions- und Testdaten.

#### C. Beschreibung der Alignment-Ontologie

Bis zu diesem Punkt beschreiben die ODPs die jeweiligen Elemente unabhängig voneinander. Auf der einen Seite gibt es ODPs, welche die technischen Prozesse, die technischen Systeme und die Kommunikation beschreiben, wie zum Beispiel die VDI 3682, ISO 7498 und *ECLASS*. Auf der anderen Seite gibt es ODPs, die die KI-Systeme, die KI-Komponenten mit ihren Funktionen und die Daten beschreiben, wie zum Beispiel ISO 22989.

Jedoch ist es zur Erstellung eines Informationsmodells, das in der Lage ist, KI-Anwendungen innerhalb von Automatisierungssystemen zu beschreiben, notwendig, Verbindungen zwischen den beiden Kategorien und ihren jeweiligen ODPs herzustellen. Um dies zu erreichen, wird eine Alignment-Ontologie erstellt, welche die Inhalte der jeweiligen ODPs miteinander verbindet. Diese Alignment-Ontologie bildet das eigentliche AIAS-Informationsmodell.

Die AIAS-Ontologie wird erstellt, indem alle zuvor genannten ODPs über `owl:import` Anweisungen importiert werden. Durch diesen Ansatz bleiben alle importierten Ontologien in ihrem eigenen Namensraum (z. B. VDI3682:, ISO22989, ISO7489:), während die

<sup>5</sup><https://eclass.eu/>

<sup>6</sup><https://www.w3.org/TR/vocab-ssn/>

Alignment-Ontologie ihren eigenen Namensraum definiert, der als *AIAS*: bezeichnet wird.

Das Metamodell<sup>3</sup> aus dem grafischen Modellierungsansatz in [25] bildet die grundlegende Struktur des *AIAS*-Informationsmodells und dient als Grundlage zur Beschreibung von KI in Automatisierungssystemen. Dieses Metamodell umfasst die Kernklassen *Funktion*, *Komponente* und *Relation*. Ein Teil dieser Kernklassen sowie einige Unterklassen werden sukzessive durch importierte Klassen aus den ODPs der Standards ersetzt, erweitert oder gleichgestellt. Dieser iterative Prozess ermöglicht die Integration spezifischen Domänenwissens und gewährleistet die Integration von etablierten Standards. Das zentrale Konzept der *AIAS*-Alignment-Ontologie mit den zugrundeliegenden Elementen ist in einem Klassendiagramm in Abb. 1 abgebildet.

Da die jeweiligen Klassen aus den ODPs importiert wurden, werden auch weitere Beziehungen, die in den ODPs beschrieben sind, in *AIAS* importiert und integriert. Beispielsweise werden Beziehungen zwischen Training, Modell, Inferenz und Daten eingebunden. Dadurch wird eine detailliertere Modellierung von KI-Anwendungen in Automatisierungssystemen ermöglicht.

## V. ANWENDUNG AM FALLBEISPIEL

Zur Evaluierung des *AIAS*-Informationsmodells wurde ein exemplarisches industrielles Fallbeispiel einer KI-Anwendung modelliert. Der Anwendungsfall war Teil des Forschungsprojekts EKI<sup>7</sup> und wurde in Zusammenarbeit mit einem Industriepartner spezifiziert und implementiert. Die Modellierung erfolgte durch Instanziierung und Verbindung der anwendungsfallspezifischen Informationen. Auf diese Weise wurde ein anwendungsfallspezifischer Wissensgraph

<sup>7</sup><https://dtecbw.de/home/forschung/hsu/projekt-eki>

erzeugt, welcher die instanzspezifischen Informationen semantisch einheitlich dokumentiert. Dieser Graph ist in Abb. 2 abgebildet.

### A. Vorstellung des Fallbeispiels

Der Anwendungsfall des Fallbeispiels ist ein Stanzprozess, der von einer Stanzmaschine durchgeführt wird. Das Hauptgeschäftsziel des Anwendungsfalls besteht darin, die Wartungskosten für die Stanzmaschine zu minimieren.

Die Stanzmaschine wird verwendet, um Teile aus blanken Metallblechen für die Automobilindustrie herzustellen. Sie besteht aus einem Traggestell, einer festen Untermatrize und einer beweglichen Obermatrize. Die Obermatrize wird von einem Elektromotor betrieben, der auf dem Traggestell montiert ist, wobei die Antriebskraft über einen Antriebsriemen übertragen wird. Die Position der Obermatrize wird von einem Positionsgeber überwacht. Im Laufe der Zeit führt der Verschleiß des Antriebsriemens zu einer ungenauen Steuerung der Obermatrize, wodurch die hergestellten Teile von den strengen Toleranzanforderungen abweichen. Derzeit wird der Antriebsriemen in regelmäßigen Abständen ausgetauscht, was zu hohen Wartungskosten führt. Daher besteht das Hauptgeschäftsziel darin, die Wartungskosten zu senken, indem der Zustand des Antriebsriemens kontinuierlich klassifiziert wird.

Experimente des Industriepartners zeigen, dass ein abgenutzter Antriebsriemen Schwingungen in der Positionierung der Obermatrize verursacht. Mit diesem Erkenntnisgewinn können Daten, die während des Stanzprozesses von Sensoren aufgezeichnet werden, genutzt werden, um den Zustand des Antriebsriemens zu inferieren.

Die Sensordaten werden über ein Bussystem an einen Controller übertragen. Der Controller kommuniziert über Ethernet mit einem Edge-Gerät, während das Edge-Gerät

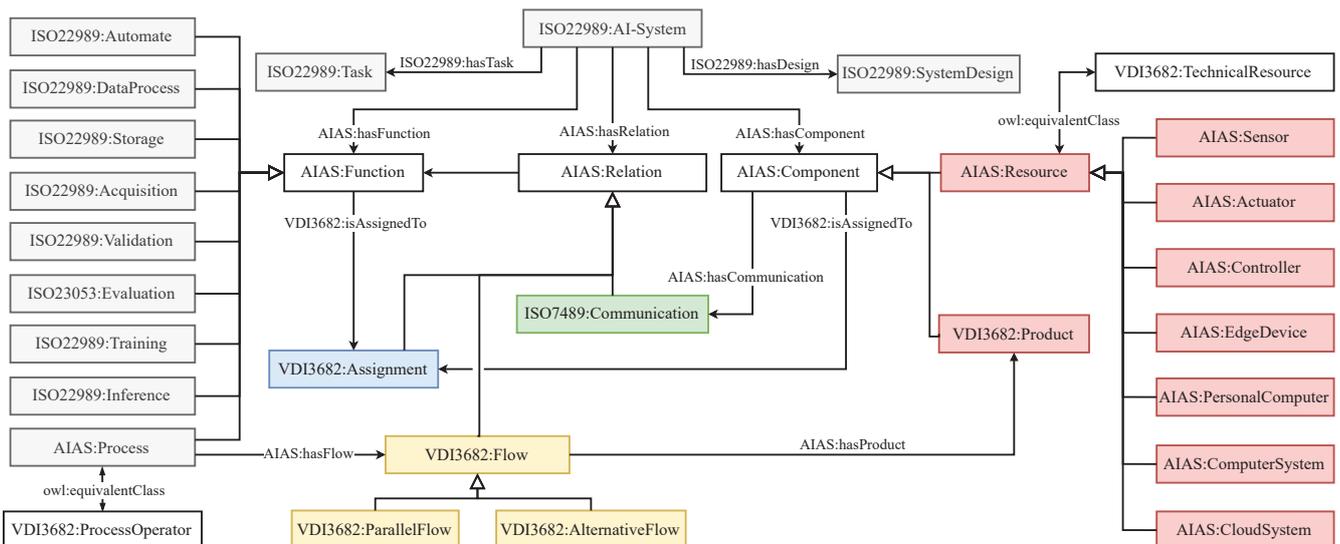


Abbildung 1: Kernkonzept der Alignment-Ontologie *AIAS*, welches die Verbindung verschiedener Standards durch Ontologie-Design-Patterns (ODPs) zeigt. Die detaillierten Inhalte der ODPs sind nicht dargestellt, um die Verständlichkeit zu erhöhen. Rot kennzeichnet technische Ressourcen oder Produkte. Blau kennzeichnet Zuweisungen. Grün kennzeichnet Kommunikationen. Grau kennzeichnet Systemfunktionen und -eigenschaften. Abbildung in Anlehnung an [29].

selbst über das Internet mit einer Cloud kommuniziert. In der Cloud wird das Modell trainiert und für die Inferenz bereitgestellt. Ein künstliches neuronales Netzwerk (KNN) wird als Modell verwendet, um die Sensordaten zu analysieren und den Zustand des Riemens zu klassifizieren.

### B. Durchführung der Modellierung

Der Anwendungsfall wurde erfolgreich unter Verwendung der AIAS-Alignment-Ontologie und der Software *Protégé*<sup>8</sup> modelliert. Zum Beispiel wurde die Kommunikation zwischen dem Edge-Gerät und dem Controller mit einer Instanz der Klasse `ISO7489:Communication` modelliert, die mit Instanzen der entsprechenden technischen Ressourcen, `AIAS:EdgeDevice` und `AIAS:Controller`, verbunden ist. Ebenso wurden das Training in der Cloud und die Bereitstellung des trainierten KNNs mit Instanzen der Klasse `VDI3682:Assignment` modelliert, die jeweils mit Instanzen der Funktionen `ISO22989:Training` und `ISO22989:Inference` verbunden sind. Während des Modellierungsprozesses wurden Elemente des ISO 22989 ODPs verwendet, um weitere Details bereitzustellen, wie Instanzen der Klassen `ISO22989:Modelparameter`, `ISO22989:Hyperparameter` und `ISO22989:Data`. Ein vereinfachter Auszug des resultierenden Wissensgraphs ist in Abb. 2 dargestellt.

<sup>8</sup><https://protege.stanford.edu/>

### C. Diskussion der Ergebnisse

Das AIAS-Informationsmodell ermöglicht die Beschreibung von KI-Anwendungen in Automatisierungssystemen durch die Fokussierung auf die Interdependenzen zwischen Komponenten des Automatisierungssystems, KI-Komponenten und technischen Prozessen, wie am Beispiel der Modellierung des Anwendungsfalls demonstriert (A1). Die Verwendung von Begriffen aus Standards wie der ISO 22989, ISO 7489 oder VDI 3682 gewährleistet eine klare und eindeutige Semantik (A2). Darüber hinaus können bei Bedarf die Semantiken der Komponenten des Automatisierungssystems durch den Import von ECLASS oder UNSPSC in AIAS weiter angereichert werden. Die Nutzung von Ontologien in Kombination mit OWL gewährleistet die formalisierte Erfassung von Informationen (A3). OWL als standardisierte Repräsentationssprache, gepflegt und definiert durch das W3C, macht den Ansatz maschinenlesbar und herstellerunabhängig (A3). Die AIAS-Ontologie bestehend aus mehreren unabhängigen ODPs und bietet Erweiterbarkeit und Anpassungsfähigkeit (A4). Weitere ODPs können einfach importiert und integriert werden, um detailliertere Beschreibungen zu ermöglichen. Aktualisierungen von Standards erfordern lediglich die Aktualisierung des relevanten ODPs, nicht der gesamten AIAS-Ontologie.

Zusammenfassend erfüllt das AIAS-Informationsmodell die genannten Anforderungen (A1-A4) und ermöglicht die Modellierung und Dokumentation von KI-Anwendungen für Automatisierungssysteme.

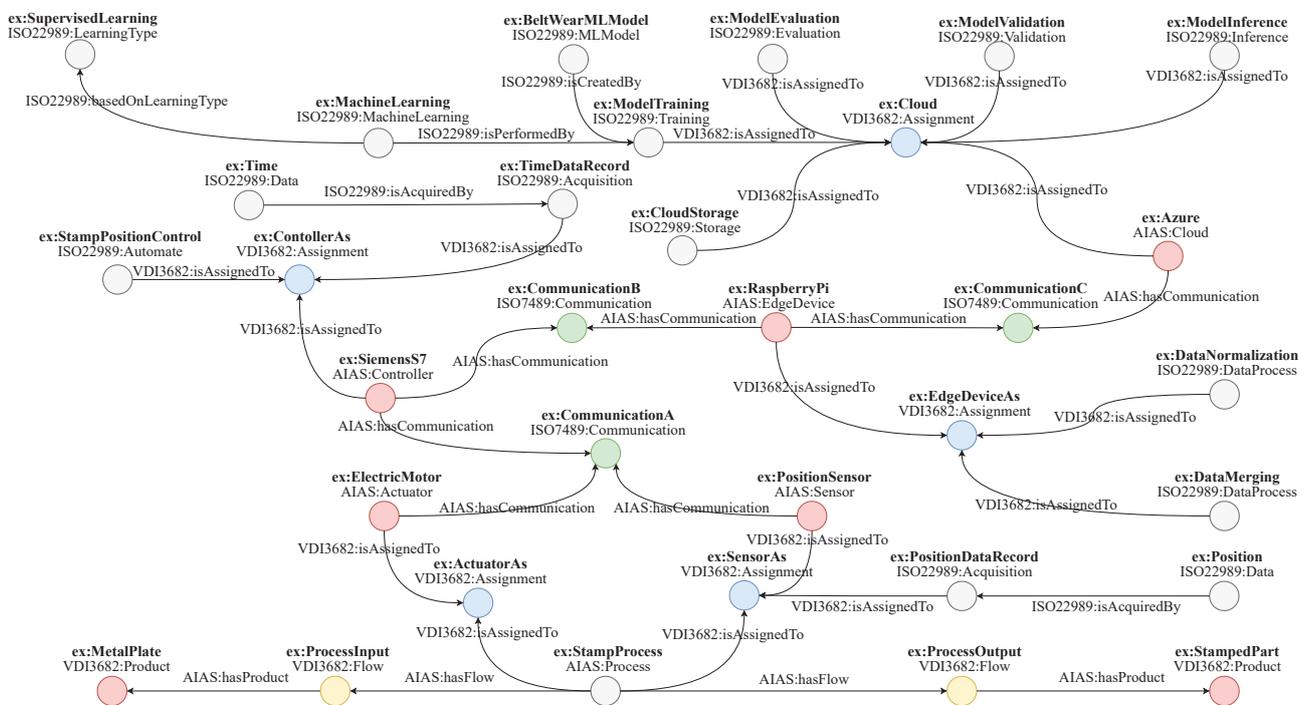


Abbildung 2: Ausschnitt aus dem resultierenden Wissensgraph des Fallbeispiels. Rote Knoten sind technische Ressourcen oder Produkte. Blaue Knoten entsprechen den Zuweisungen. Grüne Knoten sind Kommunikationsbeziehungen. Grauen Knoten sind Systemfunktionen und deren Detaillierung.

AIAS ermöglicht zudem Abfragen basierend auf Querying zur einfachen Extraktion spezifischer Informationen. Darüber hinaus ist die Formulierung komplexer Sachverhalte durch Regeln möglich sowie die Generierung neuen Wissens mithilfe von Reasoning. Trotz dieser Vorteile gibt es einige Nachteile des Ansatzes. Zunächst ist ein detailliertes Verständnis des AIAS-Informationsmodells, der AIAS-Ontologie sowie ihrer ODPs erforderlich, bevor mit der Modellierung begonnen werden kann. Ein weiterer Nachteil ist die Abhängigkeit von Werkzeugen wie Protégé, deren Nutzung ebenfalls Einarbeitungszeit erfordert.

## VI. ZUSAMMENFASSUNG UND AUSBLICK

Dieser Artikel stellt ein formales Modell zur Beschreibung von Anwendungen Künstlicher Intelligenz (KI) in Automatisierungssystemen vor. Das Konzept ermöglicht die Erstellung eines Wissensgraphen, der die Komponenten des Automatisierungssystems, die KI-Elemente, die technischen Prozesse und ihre Wechselwirkungen umfasst und dokumentiert.

Das formale Modell wurde unter Verwendung von Ontologien implementiert, wobei die Strategie verfolgt wurde, mehrere kleinere Ontologie-Design-Patterns (ODPs) auf Basis von Standards zu nutzen, anstatt eine einzelne große und monolithische Ontologie zu konstruieren. Diese modularen ODPs werden kombiniert, um eine erweiterbare und anpassbare Alignment-Ontologie namens AIAS-Ontologie zu bilden. Speziell wird ein ODP basierend auf dem ISO 22989 Standard genutzt, um Informationen über KI-Elemente zu beschreiben, während ein ODP basierend auf der VDI 3682 für Informationen über das Automatisierungssystem, technische Prozesse und Komponenten verwendet wird. Zusätzlich wird ein ODP, abgeleitet aus der ISO 7489, erstellt, um die Kommunikation innerhalb des KI- bzw. Automatisierungssystems darzustellen. Der Beitrag zeigt die Vorteile von AIAS anhand eines exemplarischen Anwendungsfalls im Bereich eines Stanzprozesses auf.

Zusammenfassend ermöglicht das Modellieren mit AIAS die formale Beschreibung von KI-Anwendungen für industrielle Anwendungsfälle. Folglich hat AIAS das Potenzial, die Dokumentation zu standardisieren und die Integration, den Betrieb und die Wartung von KI-Anwendungen in Automatisierungssystemen zu optimieren.

Für die Zukunft ist geplant, AIAS um weitere Subklassen zu erweitern, um den Umfang der Beschreibung von KI-Elementen zu erhöhen. Darüber hinaus beabsichtigen wir, Regeln und Einschränkungen zu erstellen, um den Dokumentationsprozess von KI-Anwendungen weiter zu verbessern. Zum Beispiel könnten Regeln verwendet werden, um automatisch den Typ eines KI-Systems gemäß gesetzlicher Vorschriften zu klassifizieren. Zusätzlich zur Erweiterung von AIAS ist die Entwicklung eines spezialisierten Modellierungstools geplant, um den Modellierungsprozess zu optimieren und die Einstiegshürden für die Nutzung zu senken.

## DANKSAGUNG

Diese Forschungsarbeit aus dem Projekt EKI<sup>7</sup> wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] "Harnessing the AI Revolution in Industrial Operations," World Economic Forum, Technischer Bericht, Oct. 2023. [Online]. Available: [https://www3.weforum.org/docs/WEF\\_Harnessing\\_the\\_AI\\_Revolution\\_in\\_Industrial\\_Operations\\_2023.pdf](https://www3.weforum.org/docs/WEF_Harnessing_the_AI_Revolution_in_Industrial_Operations_2023.pdf)
- [2] R. S. Peres, X. Jia, J. Lee *et al.*, "Industrial Artificial Intelligence in Industry 4.0-Systematic Review, Challenges, and Outlook," *IEEE Access*, Band 8, 2020, doi: 10.1109/ACCESS.2020.3042874.
- [3] "Artificial Intelligence Use cases," Technischer Bericht TR 24030, 2024.
- [4] M. Merkel-Kiss und J. Von Garrel, "Systematische Literaturanalyse zum KI-Einsatz und KI-basierten Geschäftsmodellen in produzierenden kleinen und mittleren Unternehmen [Systematic Literature Analysis on the Use of AI and AI-based Business Models in Small and Medium-sized Manufacturing Companies]," *Zeitschrift für Arbeitswissenschaft*, Band 77, Nr. 3, pp. 453–468, 2023, doi: 10.1007/s41449-022-00323-9.
- [5] M. W. Hoffmann, R. Drath, und C. Ganz, "Proposal for Requirements on Industrial AI solutions," in *Machine Learning for Cyber Physical Systems*. Springer Berlin Heidelberg, 2021, pp. 63–72, doi: 10.1007/978-3-662-62746-4\_7.
- [6] A. Paleyes, R.-G. Urma, und N. D. Lawrence, "Challenges in Deploying Machine Learning: A Survey of Case Studies," *ACM computing surveys*, Band 55, Nr. 6, pp. 1–29, 2022, doi: 10.1145/3533378.
- [7] D. Sculley, G. Holt, D. Golovin *et al.*, "Hidden Technical Debt in Machine Learning Systems," in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, Reihe NIPS'15. MIT Press, 2015, p. 2503–2511.
- [8] European Parliament, "P9 TA(2024)0138 Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts," Jan. 2024. [Online]. Available: <https://artificialintelligenceact.eu/wp-content/uploads/2024/02/AIA-Trilogue-Coreper.pdf>
- [9] C. P. Filho, E. Marques, V. Chang *et al.*, "A Systematic Literature Review on Distributed Machine Learning in Edge Computing," *Sensors*, Band 22, 2022, doi: 10.3390/s22072665.
- [10] "Modellierung- und Simulationsbedarfe der intelligenten Fabrik," Forschungsbeirat der Plattform Industrie 4.0 / acatech - Deutsche Akademie der Technikwissenschaften, Technischer Bericht, Dez. 2021. [Online]. Available: <https://www.acatech.de/publikation/modellierungs-und-simulationsbedarfe-der-intelligenten-fabrik-expertise/download-pdf?lang=de>
- [11] C. Hildebrandt, A. Köcher, C. Küstner *et al.*, "Ontology Building for Cyber-Physical Systems: Application in the Manufacturing Domain," *IEEE Transactions on Automation Science and Engineering*, Band 17, Nr. 3, pp. 1266–1282, 2020, doi: 10.1109/TASE.2020.2991777.
- [12] T. Gebru, J. Morgenstern, B. Vecchione *et al.*, "Datasheets for datasets," *CoRR*, 2018, doi: arXiv:1803.09010v1.
- [13] M. Mitchell, S. Wu, A. Zaldivar *et al.*, "Model Cards for Model Reporting," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, Reihe FAT\* '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 220–229, doi: 10.1145/3287560.3287596.
- [14] M. Arnold, R. K. E. Bellamy, M. Hind *et al.*, "FactSheets: Increasing Trust in AI Services Through Supplier's Declarations of Conformity," *IBM Journal of Research and Development*, Band 63, Nr. 4/5, pp. 6:1–6:13, 2019, doi: arXiv:1808.07261.
- [15] A. Lavin, C. M. Gilligan-Lee, A. Visnjic *et al.*, "Technology readiness levels for machine learning systems," *Nature Communications*, Band 13, Nr. 1, p. 6039, 2022, doi: 10.1038/s41467-022-33128-9.
- [16] P. K. Sinha, S. B. Gajbe, S. Debnath *et al.*, "A Review of Data Mining Ontologies," *Data Technologies and Applications*, Band 56, Nr. 2, pp. 172–204, 2022, doi: 10.1108/DTA-04-2021-0106.
- [17] P. K. Sinha, S. B. Gajbe, K. Chakraborty *et al.*, "A Review of Machine Learning Ontologies," *International Journal of Information*

- Dissemination and Technology*, Band 11, Nr. 4, pp. 158–164, 2021, doi: 10.5958/2249-5576.2021.00027.3.
- [18] P. Panov, L. Soldatova, und S. Džeroski, “OntoDM-KDD: Ontology for Representing the Knowledge Discovery Process,” in *Discovery Science*. Springer Berlin Heidelberg, pp. 126–140, doi: 10.1007/978-3-642-40897-7\_9.
- [19] C. M. Keet, A. Ławrynowicz, C. Amato *et al.*, “The Data Mining OPTimization Ontology,” Band 32, pp. 43–53, doi: 10.1016/j.webssem.2015.01.001.
- [20] J. Vanschoren und L. Soldatova, “Exposé: An Ontology for Data Mining Experiments,” in *International workshop on third generation data mining: Towards service-oriented knowledge discovery (SoKD-2010)*, 2010, pp. 31–46.
- [21] D. Esteves, D. Moussallem, C. B. Neto *et al.*, “MEX Vocabulary: A Lightweight Interchange Format for Machine Learning Experiments,” in *Proceedings of the 11th International Conference on Semantic Systems*, 2015, pp. 169–176, doi: 10.1145/2814864.2814883.
- [22] G. C. Publio, D. Esteves, A. Ławrynowicz *et al.*, “ML-Schema: Exposing the Semantics of Machine Learning with Schemas and Ontologies,” 2018, doi: arXiv:1807.05351.
- [23] J. Vanschoren, J. N. Van Rijn, B. Bischl *et al.*, “OpenML: Networked Science in Machine Learning,” *ACM SIGKDD Explorations Newsletter*, Band 15, Nr. 2, pp. 49–60, 2014, doi: 10.1145/2641190.2641198.
- [24] C. Kaymakci, S. Wenninger, und A. Sauer, “A Holistic Framework for AI Systems in Industrial Applications,” in *Innovation Through Information Systems*. Springer International Publishing, 2021, doi: 10.1007/978-3-030-86797-3\_6.
- [25] M. Schieseck, P. Topalis, und A. Fay, “A Graphical Modeling Language for Artificial Intelligence Applications in Automation Systems,” in *2023 IEEE 21st International Conference on Industrial Informatics (INDIN)*, Lemgo, DE, 2023, doi: 10.1109/INDIN51400.2023.10217890.
- [26] R. Haberfellner, P. Nagel, M. Becker *et al.*, *Systems Engineering*. Springer, 2019, doi: 10.1007/978-3-030-13431-0.
- [27] A. Gangemi und V. Presutti, “Ontology Design Patterns,” in *Handbook on ontologies*. Springer, 2009, pp. 221–243, doi: 10.1007/978-3-540-92673-3\_10.
- [28] *Formalised Process Descriptions*, Verein Deutscher Ingenieure e.V. (VDI) Standard 3682, 2015.
- [29] M. Schieseck, P. Topalis, L. Reinpold *et al.*, “A Formal Model for Artificial Intelligence Applications in Automation Systems,” in *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Padova, IT, 2024.
- [30] *Open Systems Interconnection*, International Organization for Standardization (ISO) Standard 7498, 1996.
- [31] *Artificial Intelligence Concepts and Terminology*, International Organization for Standardization (ISO) Standard 22989, 2022.

# A Cross-Domain Electromagnetic Compatibility Data Platform for the Digital Development of Automated Vehicles

Unai Aizpurua, Tim Brandt, Mohammed ElSayed, Michael Hagel, Erik Kampert, Shounak S. Kulkarni, Stefan Schenke, Marita Wendt, Stefan Dickmann\*, Klaus F. Hoffmann, Marcus Stiemer

*Faculty of Electrical Engineering  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Germany  
\*stefan.dickmann@hsu-hh.de*

Jan Wagner, Peter Thoma<sup>‡</sup>  
*Computer Science and Engineering Department  
Frankfurt University of Applied Sciences  
Frankfurt am Main, Deutschland  
<sup>‡</sup>peter.thoma@fb2.fra-uas.de*

Hermann Aichele, Salvador Cuevas Rosa,  
Roland Eidher  
*Robert Bosch GmbH  
Stuttgart, Deutschland*

**Abstract** — Maintaining compliance to electromagnetic compatibility (EMC) standards becomes an increasing challenge for the automotive industry in the course of the ongoing automation of vehicles. Novel extended design procedures and test standards are required to ensure safety of automated driving functions, particularly in an adverse electromagnetic (EM) environment. In order to keep the number of tests within a reasonable and practical limit, an evaluation framework based on virtual design methods and information drawn from legacy experiments and simulations can support the automotive industry. With this digital framework appropriate technological solutions can be identified during the pre-compliance phase and efficient experimental designs can be generated to ensure EMC compliance. Furthermore, such a framework paves the way for digital EMC twins of automated vehicles (AVs) considering the complex interrelations of AV's (sub-)systems to accurately predict the behaviour of new AV functions in various EM environments. To this purpose, a cross-domain platform is being developed in this work as the backbone of such a virtual framework. It supports the handling, storage and processing of various datasets from EMC test campaigns, including (intentional) electromagnetic interference ((I)EMI) tests, as well as simulations of automotive devices-under-test (DUTs). The platform allows for the establishment of interconnections between various data sources and deeper analyses based on artificial intelligence (AI) methods to deduce EMC information for new developments, whilst maintaining traceability.

**Keywords** — Digital twin, autonomous driving, electromagnetic compatibility, electromagnetic interference, power electronics, artificial intelligence

## I. INTRODUCTION

A digital twin of AVs with a particular emphasis on EMC requirements can greatly support the development of AVs from the design and pre-compliance phase to production.

Even beyond the life cycle of a vehicle's model line it can predict the EMC-impact of redesigns, environmental changes such as new wireless communication standards or new (I)EMI sources, or a new area of application, which frequently occurs in a military context. It enables performing virtual tests on complex (sub-)systems, consisting of multiple components and devices each of which possess individual EM susceptibility and interference behaviour. These virtual tests could reduce the amount of, or ideally substitute, real laboratory tests that would be difficult to perform, among others because of the vast number of system state combinations, which would be required to be examined. Furthermore, digital twins of AVs enable model-based predictions on layout changes of such components whilst considering predictions of their impact on EMC at higher system levels. Their overall goal is to speed-up implementing and confirming EMC compliance of an automotive product, reducing the amount of lab tests by minimising test redundancy and allowing to focus on verification and validation by merely performing the most crucial component, device and system tests. Employing a digital twin should yield the same confidence and fidelity in the electromagnetic resilience (EMR) of the automotive DUT and the systems it is part of as performing real lab tests.

A digital twin is as good as the data and models it comprises, as well as the platform that enables the cross-disciplinary exchange between models and data of different types and, thus, provides a holistic EMC evaluation (Fig. 1, [1]). Hence, this paper presents a general approach to storing data of various origins, labelling them, documenting their interrelations and enriching them through their use as an input to more refined models. The presented cross-domain

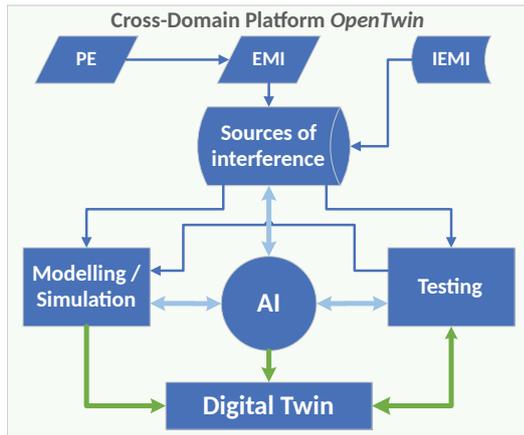


FIGURE 1: PROCESS OVERVIEW FOR THE IMPROVEMENT OF AUTOMOTIVE EM IMMUNITY AND THE AI-SUPPORTED EVALUATION OF THE SYNERGIES BETWEEN MODELLING, NUMERICAL SIMULATIONS AND LAB TESTS.

platform thus provides a general framework to render various data into EMC-relevant information by using both physics-based models and techniques from machine learning (ML). The process of data enrichment is transparently traced via metadata, which are stored with the raw and augmented datasets. The cross-domain platform provides the backbone for digital twins and may simultaneously serve as a data-lake. This includes methods for the transfer of experimental knowledge that was previously extracted from legacy data via ML to new projects, so that the gained know-how helps finding good starting points for further technical development.

Both the development and implementation of the cross-domain platform are jointly achieved by academic and industrial partners. Whereas the former contribute with state-of-the-art methodologies and their theoretical foundations, they benefit from the vast experience of the latter partners. This includes the exchange of information about EMC requirements, standards and best practice, the definition and scaling of relevant EMC scenarios, as well as insights into existing digital processes involving a large number and variation of use cases. In this work, special emphasis is given to the generation of example data from various sources, which will be processed by the cross-domain platform to demonstrate the deduction of EMC-relevant information. These data include both studies on conducted and radiated (I)EMI and comprises both measured and simulated data, with the underlying setups motivated by industrial use cases.

Similarly, both industry and academic researchers benefit from the cross-domain platform, which is based on an open access framework. In an industrial context, the platform allows for fusing data from various business units, for further processing and reordering them and for their storage together with metadata for full traceability, such that a company's know-how is manifested and made accessible and usable for new developments. An academic use of the platform comprises a

systematic bookkeeping of research results, combined with full access of subsequent researcher generations.

This paper starts with a brief summary of the cross-domain platform *OpenTwin* to be developed (Section II). In the following Section III and Section IV the generation of exemplary experimental data on conducted and radiated electromagnetic interference (EMI) and intentional electromagnetic interference (IEMI) is presented, respectively. Examples of ML-based processing and refining of data that have been computed from different physics-based models are then provided in Section V and Section VI.

## II. THE OPENTWIN CROSS-DOMAIN PLATFORM

The *OpenTwin* cross-domain platform developed in this work will, in its final state, be used to generate digital twins for the EMC behaviour of new concepts of AVs. It thus provides a general framework for the systematic construction of digital twins based on annotated legacy data with various levels of refinement, stemming from different sources.

Although the term *digital twin* has no unique definition, it is generally agreed that it consists of a physical asset, a virtual representation (model) and a data transfer between them [2], [3]. A digital twin has the potential to cover the entire physical asset's life cycle represented by its digital models, which can be data-driven, physics-driven or a hybrid. The physics-driven models are either analytical or realised by numerical simulation and may be capable of handling multi-scale or multi-physics behaviour of the physical asset.

At its root, the digital twin contains data about the various employed models as well as additional data describing the asset, ideally in a machine-readable format. From a data management perspective, this collection of data is diverse, large in volume and requires fast access to support high-throughput data-processing methods, e.g. based on AI. Most of the data for a digital twin originate from measurements, simulations or data post-processing methods. The models and data defining a digital twin and emerging from the information gathered on *OpenTwin* will be used for the EMC assessment of the current state of systems in various EM environments as well as that of future redesigns. Already before the platform has been supplied with sufficient data that have been transformed on an adequate level of refinement, the implemented data and services will be used for the development of EMC concepts for novel AVs.

Fig. 2 shows the data model, that is proposed in this work, for handling any kind of data stemming from the sources mentioned above, to enable the core aspects of a digital twin. Essentially, the metadata of a data generation process are separated from the collected data, which are stored in the *Quantity Container*. This allows for optimising this entity for storing large heterogeneous data. Unidirectional references from the *Quantity Container* to the corresponding metadata via abstract unique identifiers guarantee a highly efficient editing ability. The referenced *Series Metadata* entity contains metadata about the data acquisition process, which could be, e.g. a measurement series, and all information about the acquired

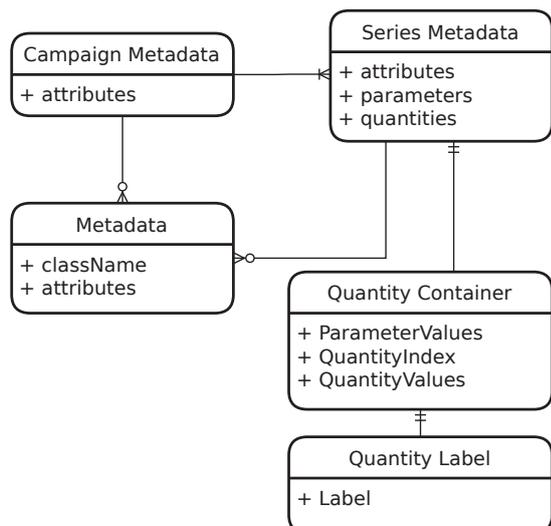


FIGURE 2: CONCEPTUAL DATA MODEL FOR DIGITAL TWINS.

quantities and the parameters that were varied during the acquisition process. A top-level entity, *Campaign Metadata*, acts as an instance to describe the entire dataset. Generically constructed *Metadata* entities allow nested metadata to be added. Finally, a *Quantity Label* entity enables the annotation of individual data points.

The concentration of all metadata in particular entities allows changes to be made without having to access the large data point collections. Here, this data model is implemented and tested in *OpenTwin* [4]. Its features include the ability to handle numerical simulations of multiple physics domains, the visualisation of results, a pipeline editor for highly customised Python-based data processing and a distributed version control of models. The metadata entities are also included in the version control, providing implicit version control of the stored data points, but with reduced storage overhead. Combining modelling with data handling of simulated, measured and post-processed data in a single version-controlled software enables traceability of model modifications and corresponding data changes. This forms the basis for approval and certification processes in a product's life cycle.

With respect to hardware resources, *OpenTwin's* microservice architecture allows for distributed deployment. This functionality permits the deployment of modules which require high computational power, such as simulation solvers, on specialised computing nodes, whereas large storage functionality can be deployed on other nodes. In cases in which the distribution of computational tasks conflicts with security requirements, methods of secure distributed computation can be considered [5].

### III. CONDUCTED EMI FROM AUTOMOTIVE DC-DC-CONVERTERS

Automotive power electronics (PE) components have been identified as important sources of potential EMI in a diversified board net of electric vehicles (EVs) [1]. Hence, laboratory

adaptations of modern DC-DC-converters based on wide-bandgap semiconductors interfacing the  $\approx 450$  V EV battery and the 12 V side with an intermediate stage of 48 V have been evaluated in detail in this project [6], [7]. The collected data is exemplary considered as a first data source for *OpenTwin*.

In particular, the fast-switching processes of Silicon Carbide (SiC) and Gallium Nitride (GaN) transistors result in transients that cause EMI at the fundamental frequency and its harmonics at a few hundred kilohertz and higher frequencies at several megahertz. Although substantial radiated emissions might arise from harmonics beyond 30 MHz, affecting nearby components and devices susceptible to radiative coupling, the focus of this work is on conducted emissions, as automotive PE are a main source of conducted EMI in automated EVs. As detailed in previous work, common mode (CM) and differential mode (DM) interferences have been measured both at the input and output lines of high- and low-stage DC-DC-converters [7]. Their expected impact on automotive DUTs is highest at the 48 V intermediate stage, as well as at the legacy user 12 V board net.

EMI data collected with laboratory test setups within this project include oscilloscope time series of input and output voltage, voltage over individual transistors [6], individual power line or CM/DM voltages at the line impedance stabilisation networks (LISNs) of the standards-compliant setup, as well as induced radio frequency (RF) currents in the converter's input and output lines. The latter voltages and currents have also been recorded with spectrum analysers, providing more accurate average and peak emission amplitudes and a higher frequency resolution, but lacking insights into the individual transient processes and their periodicity, which the time series contribute. As an example, results for the evaluation of a SiC-based DC-DC-converter with an input voltage of 450 V, an output voltage of 48 V, an output current of 48 A, a switching frequency of 100 kHz and for a SiC-chip temperature of 90 °C are presented in Fig. 3. The variation of transient currents with time is highlighted in Fig. 3a, which are averaged out in the recorded spectra in Fig. 3b for the same parameter.

Such time- and frequency-domain datasets have been collected for various operating conditions of the DC-DC-converters, including the variation of parameters such as input and output voltage (SiC: 300 V–450 V to 48 V, GaN: 48 V to 12 V–14 V), output current (SiC: 12 A–48 A, GaN: 24 A–220 A), switching frequency (SiC: 50 kHz–100 kHz, GaN: 250 kHz and 500 kHz), transistor chip temperature (SiC: 90 °C–150 °C) etc. A controlled automated evaluation of this vast number of datasets could uncover the underlying relationships between specific transients, their impact on the resulting EMI spectra and their dependence on the varied parameters. Additionally, the availability of corresponding time- and frequency-domain datasets on *OpenTwin* enables testing analysis methodologies that can relate both datasets to each other, for a large number of parameter variations [8]. For those cases in which the results can be made equivalent, EMC testing could be considerably accelerated by

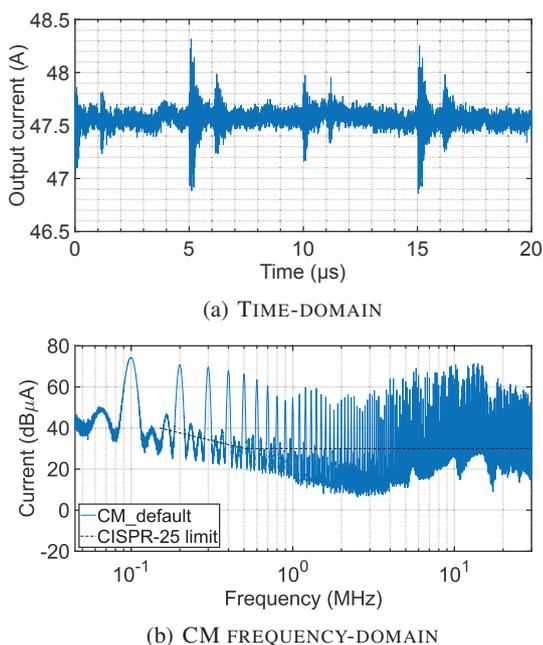


FIGURE 3: SELECTED TIME SERIES AND SPECTRUM OF CONDUCTED EMI BY AN AUTOMOTIVE SIC-BASED DC-DC-CONVERTER, WITH  $V_{in} = 450$  V,  $V_{out} = 48$  V,  $I_{out} = 48$  A,  $f_{sw} = 100$  kHz AND  $T_{chip} = 90$  °C.

performing time-domain measurements only.

#### IV. SUSCEPTIBILITY OF AUTOMOTIVE DUTS TOWARDS RADIATED IEMI

Another subset of information required for the development of an AV digital twin includes data related to the electromagnetic susceptibility (EMS) of automotive components and (sub-)systems. Beyond tests with EMI signals following automotive standards for both measurements and specific threshold levels, this work also includes radiated IEMI tests involving RF sources and testing areas with restricted availability and access, such as the Bundeswehr Research Institute for Protective Technologies and CBRN Protection (WIS<sup>1</sup>). Work is ongoing to reproduce selected IEMI tests with commercially available laboratory equipment, but these will likely remain limited in the available peak EM field strength for a working volume that can accommodate the typical size of an automotive (sub-)system [9].

In previous work, IEMI tests with broadband and mesoband pulses on an acceleration sensor indicated that such a sensor is resistant to field strengths up to 180 kV/m for the former and 40 kV/m for the latter pulse type, respectively [10]. Here, results of radiated IEMI tests on the electronic control unit (ECU) for the acceleration sensors previously tested, using a high-power pulsed S-band source (3 GHz, continuous wave,

0.7  $\mu$ s RF pulse width, 2 MW peak output power), are presented as an example. On an open area test site (OATS), the peak electric field for a distance of 30 m between the DUT and the RF source is  $\approx 8$  kV/m. The DUT includes a cable tree, load boxes and periphery that represent the ECU's network and connections when embedded within a vehicle. The main coupling path of the S-band IEMI signal is known to be through the ECU's cable tree.

In addition to detailed information about the (I)EMI signals in the time- and frequency domain obtained from calibration measurements, e.g. with D-dot sensors, the behaviour of a DUT during EM exposure needs to be captured. Ideally this is realised with a similar high time-resolution to be able to synchronise the DUT's behaviour with the (I)EMI source. Data that can be obtained from a DUT strongly depend on its functionality, hence they likely require a custom solution per DUT to be embedded in *OpenTwin* for the development of an AV digital twin.

The typical analysis of the impact of (I)EMI signals on an airbag ECU involves the monitoring of the least significant bit (LSB). The LSB represents the smallest increment of the digital representation of a quantised analogue input signal and determines the resolution of the converted digital signal. In Fig. 4 the LSB of the ECU's data stream is monitored during the exposure to a pulsed S-band signal. In Fig. 4a it can be observed, that the raw LSB data is strongly affected by the IEMI exposure in the time range central in the plot. Upon evaluating the obtained LSB values and comparing them with a threshold defined by standards, Fig. 4b displays that the limits are exceeded for this IEMI exposure. On *OpenTwin* the full datasets including multiple repetitions of the IEMI exposure can be automatically analysed in more depth, including examining the variation of the LSB during a single IEMI pulse, as highlighted in Fig. 4c.

#### V. SIMULATION-BASED ANALYSIS OF THE SUSCEPTIBILITY OF AUTOMOTIVE CABLES TO (I)EMI

An important source of information to predict an AV's EMC behaviour in various EM environments are numerical simulations. These can comprise electric circuit simulations with, e.g. LTspice [11] or computational electromagnetics (CEM)-simulations based on well-known numerical methods including finite-differences time-domain (FDTD), finite element method (FEM), and method of moments (MoM). The latter forms the basis of FEKO [12], the EM simulation software used to obtain the results presented in this section.

The automotive wire harness is known to be a key coupling path for RF signals received from both the vehicle's internal and external environment [13]. An (I)EMI signal coupling into it can affect both the ECUs and other controllers that process signals from their peripherals as well as the functional safety of connected sensors and the integrity of the transmitted signals. Whereas the coupling into such a harness can be analytically determined for relatively simple cable layouts, more complex, realistic configurations require either extensive experimental efforts or CEM based simulations [14]. Recently, in order

<sup>1</sup>WIS is a departmental research institute of the German Federal Government and is subordinate to the Federal Office of Bundeswehr Equipment, Information Technology and In-Service Support (BAAINBw).

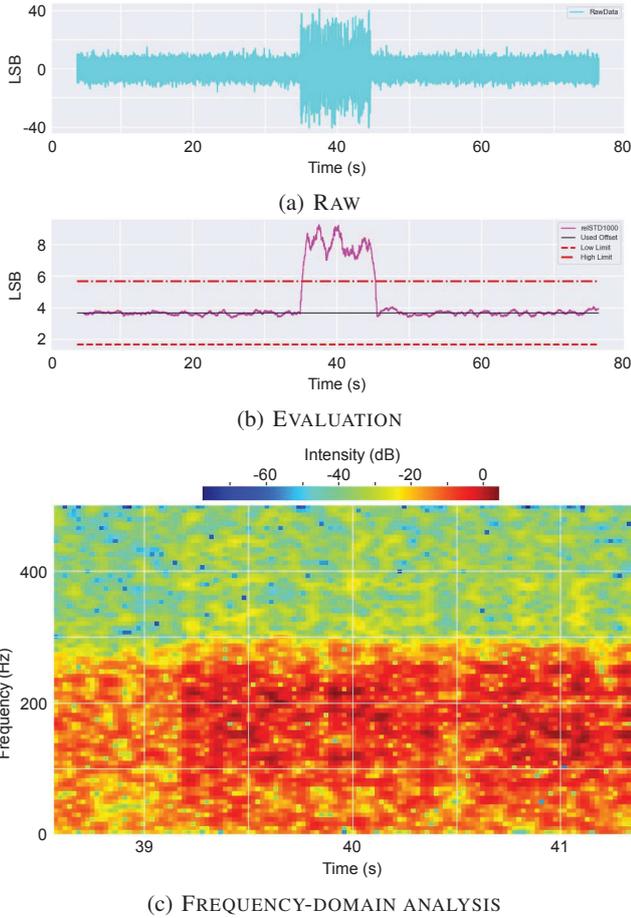


FIGURE 4: RAW AND EVALUATED LSB DATA FROM AN AIRBAG ECU EXPOSED TO AN S-BAND IEMI SIGNAL.

to accelerate the evaluation of cable layouts with regard to EM susceptibility, a method based on the reciprocity of an object’s RF signal receiving and transmitting properties has been proposed [15]. Instead of simulating an incoupling signal, the reciprocal and simpler determination of an excited cable’s radiation pattern via CEM is carried out. The cable layout to be assessed is excited by a known continuous wave (CW) signal from a voltage source, as schematically shown in Fig. 5a. The resulting three-dimensional EM field can be determined from the simulation at numerous positions around the layout, as shown in Fig. 5b. Through reciprocity such patterns also represent the EMS of the cable layout to a specific RF signal arriving from a certain direction.

When combining the results for various RF frequencies, the susceptibility of a given layout to broadband pulses, e.g. from IEMI sources, can be evaluated. *OpenTwin* is ideally suited to store these simulation data and to enable the automatic superposition of the results for CW signals into more complex, broadband signals. In a subsequent step, *OpenTwin* hosts ML algorithms that can identify hidden patterns in the relation between radiation profiles and cable layout properties. These can be provided by a trained ML model, e.g. a regression

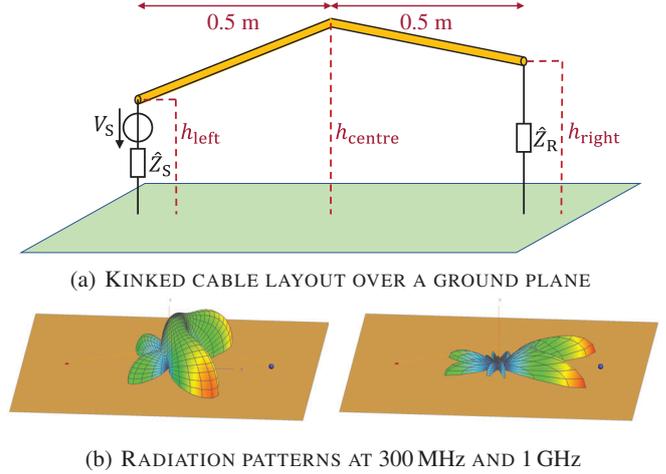


FIGURE 5: MoM SIMULATION TO EVALUATE THE EMS OF AN AUTOMOTIVE WIRE HARNESS.

neural network of moderate size, which in turn can serve to generate further data for *OpenTwin*. Instead of performing a computationally intensive CEM simulation of a complex wire harness layout, a trained ML model can then be used to predict the susceptibility of such a harness to any (I)EMI signal of interest.

## VI. ML TO DETECT (I)EMI IN VEHICLES

As previously explained, AVs will challenge EMC standards and industrial procedures to make them functionally safe. (I)EMI-related challenges for autonomous systems in an unfriendly EM environment have been examined before [1]. The interactions of efficient PE for electrical drives and advanced signal processing for autonomous operations require new concepts of EMC processes during the design of AVs. It is worthwhile to investigate methodologies that potentially protect an AV against (I)EMI by a permanent online analysis of its inherent data traffic. Such a task is ideally suited for ML that can analyse a large dataset and its underlying patterns in real-time. But even without an onboard implementation on fast hardware it is beneficial to assess the EMC of certain design variants by such a method to harden an AV against (I)EMI. In the latter case, online data are replaced by data gained via circuit simulation and *OpenTwin* provides an ideal framework for data management and processing.

In this project, autoencoders (AEs), which are usually employed for intelligent dimensional reduction of high-dimensional input data, are implemented via neural networks [16], [17]. In addition to mapping data on a low-dimensional so-called *latent space* before further data analysis, as depicted in Fig. 6, AEs are here also used as adaptive filters to detect anomalies on AV system communication signals.

Normal patterns of the internal data transfer are learned during operation and disturbed signals are quickly identified. They can also be classified after mapping on the *latent space*. Depending on the result of an online analysis, an

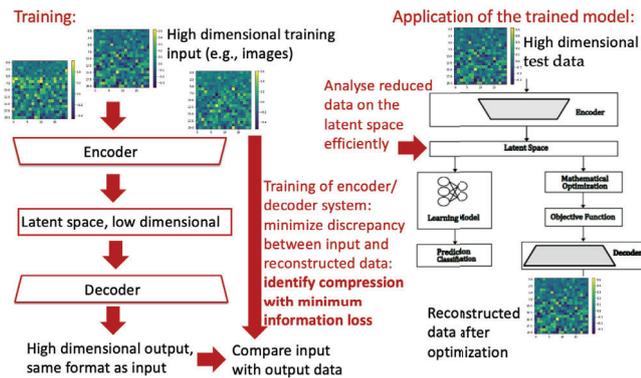


FIGURE 6: AE FOR DATA COMPRESSION.

AV's (sub-)system can be switched into a fail-safe mode or can employ built-in redundancies. Such methods can also be applied to detect changes of the system due to sudden failure of components, caused by aging or by (I)EMI.

An AE consists of an encoder with the shape of the incoming data, e.g. time series of simulated or measured currents, as the input format and a low-dimensional output shape, configuring the *latent space* [18]. Its second part is a downstream decoder with interchanged input and output shapes. The goal of training AEs is to obtain a minimum information loss when compressing data. This is achieved by processing the data through the encoder and decoder and comparing the resulting reconstructed data to the original data with respect to an adequate metric, as shown in Fig. 6. The AE adapts its internal weights to the incoming data and *learns* to compress them with a minimum information loss. Hence, the resulting *latent space* is an ideal feature space for high-dimensional data. For anomaly detection, the sequence consisting of an encoder and a downstream decoder is only trained by *normal* data and learns to reproduce the underlying pattern in high accuracy. Data from specimens deviating from the pattern will typically be mapped onto an output that differs significantly from the input and can thus be identified by a high information loss.

## VII. CONCLUSIONS

This paper presents the cross-domain platform *OpenTwin* supporting digital twin development for AVs with a focus on EMC. The data stored, labelled and interconnected on the platform stem from (I)EMI measurements on automotive PE and sensor components, as well as from simulations on automotive wire harness layouts presented in this work. Furthermore, an AE-based data analysis is detailed, which enriches the *OpenTwin* data. Next steps involve the AI-supported extraction of hidden patterns in the available data and predictions on the EMS and EMI of automotive DUTs.

## ACKNOWLEDGMENT

This paper is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which we

gratefully acknowledge [project ESAS – Elektromagnetische Störfestigkeit autonomer Systeme]. dtec.bw is funded by the European Union – NextGenerationEU.

## REFERENCES

- [1] U. Aizpurua, T. Brandt, M. Hagel, E. Kampert, M. Wendt, S. Dickmann, K. F. Hoffmann, and M. Stierner, "Avoiding electromagnetic interference induced risks for autonomous driving," in *dtec.bw-Beiträge der Helmut-Schmidt-Universität*, Band 1, 2022, pp. 174–180.
- [2] AIAA, "Digital twin: Definition & value—an AIAA and AIA position paper," AIAA Digital Engineering Integration Committee and others, Reston, VA, USA, White paper, Dec. 2020.
- [3] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108 952–108 971, 2020.
- [4] OpenTwin Organization, "OT-OpenTwin," [Online], Last accessed: 2024-10-25. [Online]. Available: <https://github.com/OT-OpenTwin>
- [5] J. Wagner and P. Thoma, "Secure cloud-based simulation," in *2022 Kleinheubach Conference*, 2022, pp. 1–4.
- [6] S. S. Kulkarni, M. Wendt, S. Brandt, K. F. Hoffmann, E. Kampert, and S. Dickmann, "Conducted emissions of an upto 23kW SiC interleaved buck converter under varied operative conditions," in *2024 International Symposium on Electromagnetic Compatibility – EMC Europe*, 2024, pp. 1–6.
- [7] M. Wendt, E. Kampert, J. Wendt, K. F. Hoffmann, and S. Dickmann, "Conducted EMI from GaN-based 48V to 12V DC-DC converters for automotive applications," in *PCIM Europe 2024; International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management*, 2024, pp. 2337–2344.
- [8] C. Keller and K. Feser, "Fast emission measurement in time domain," *IEEE Transactions on Electromagnetic Compatibility*, vol. 49, no. 4, pp. 816–824, 2007.
- [9] U. Aizpurua, E. Kampert, and S. Dickmann, "Improved determination of the electric field in radiated IEMI tests with EMP and UWB signals," in *2024 International Symposium on Electromagnetic Compatibility – EMC Europe*, 2024, pp. 1–6.
- [10] T. Brandt, S. Fisahn, M. Schaarschmidt, U. Aizpurua, E. Kampert, and S. Dickmann, "Investigating electromagnetic interference induced risks for autonomous driving," *Advances in Radio Science*, vol. 22, pp. 1–8, 2024, Last accessed: 2024-10-25. [Online]. Available: <https://ars.copernicus.org/articles/22/index.html>
- [11] Analog Devices, Inc., "LTspice," Last accessed: 2024-10-25. [Online]. Available: <https://www.analog.com/en/resources/design-tools-and-calculators/ltspice-simulator.html>
- [12] Altair Engineering Inc., "FEKO," Last accessed: 2024-10-25. [Online]. Available: <https://www.altair.de/feko/>
- [13] A. R. Ruddle and A. J. M. Martin, "Adapting automotive EMC to meet the needs of the 21st century," *IEEE Electromagnetic Compatibility Magazine*, vol. 8, no. 3, pp. 75–85, 2019.
- [14] U. Aizpurua, E. Kampert, and S. Dickmann, "EMC study of automotive wire harness configurations in a GTEM cell," in *2023 International Symposium on Electromagnetic Compatibility – EMC Europe*, 2023, pp. 1–6.
- [15] T. Liang, G. Spadacini, F. Grassi, and S. A. Pignari, "Worst-case wideband radiated IEMI for unshielded and shielded cables: A statistical analysis of the main influencing parameters," *IEEE Transactions on Electromagnetic Compatibility*, vol. 61, no. 4, pp. 1244–1252, 2019.
- [16] M. Stierner, M. Hagel, U. Aizpurua, M. ElSayed, and I. Cahani, "Machine learning based data analysis for electromagnetic reverberation chambers\*," in *2023 Kleinheubach Conference*, 2023, pp. 1–4.
- [17] M. Stierner, "Machine learning based classification and prediction of electromagnetic absorption in electrical reverberation chambers," in *2023 XXXVth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, 2023, pp. 1–4.
- [18] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, 2014, pp. 4–11.

# Synergistic Integration of Physical and Digital Simulation of a Shoulder Exoskeleton

Rajal Nagwekar\*, Seyed Milad Mir Latifi, Ramazan Gökay, Robert Weidner  
*Laboratory of Production Technology  
 Helmut-Schmidt-University/University  
 of the Armed Forces Hamburg  
 Hamburg, Germany*  
 \*rajal.nagwekar@hsu-hh.de

Alexander Stark, Julian Öltjen  
*voraus robotik GmbH  
 Hannover, Germany*

Marc Kilian Klankers, Adrian Rudloff,  
 Jochen J. Steil  
*Institute of Robotics and Process  
 Control  
 Technische Universität Braunschweig  
 Braunschweig, Germany*

**Abstract** – Support systems for humans are becoming an important aspect in today's workplace due to the increasing workload and need for productive tasks. For this purpose, industrial exoskeletons are a promising approach, especially for physically demanding tasks. To develop, evaluate, and optimize assistive systems, it is necessary to build a comprehensive understanding of these systems by investigating their dynamics and kinematics. This can be achieved by simulating all interacting components, including the task, the human, and the robotic system. This paper describes a synergy of various elements to overcome the challenges while researching and developing exoskeletons.

**Keywords** – Exoskeleton, motion capture system, robot control, virtual reality

## I. INTRODUCTION

Exoskeletons, also known as wearable robots, represent a cutting-edge advancement in the field of assistive technology and robotics [1]. These devices are designed to augment, reinforce, or restore human performance by providing external mechanical support. Initially conceptualized for military applications, exoskeletons have since found significant applications in medical rehabilitation, industrial ergonomics, and personal mobility enhancement. [1]

The use of simulation techniques is crucial in the development and implementation of exoskeletons. Without the requirement for large physical prototypes, simulation offers a virtual environment for designing, evaluating, testing, and refining exoskeleton ideas under considered settings. This method finds possible challenges early in the design phase of such systems, which speeds up development, lowers costs, and improves safety. Additionally, a variety of real-world conditions can be replicated using simulations, which enables researchers to maximize exoskeleton performance and ergonomics. Exoskeleton development is dependent on the identification and study of particular use cases. Use cases offer realistic circumstances that direct exoskeleton technology development, testing, and improvement.

However, available data indicates that a variety of industrial tasks and user attributes have a major impact on how effective exoskeletons are [2]. A training framework that can allow employees to experience various exoskeletons in their workplace while gaining the motor skills required to operate an exoskeleton safely and effectively is needed, as exoskeleton operation needs to be meaningfully and effectively integrated

into the job or work context and there is a great deal of diversity in user, work, and exoskeleton characteristics. Virtual reality, another Industry 4.0 technology, has demonstrated an excessive amount of interest in taking human factors into account throughout the lifetime of various industrial operations and products [3], [4], particularly when conventional technologies and decision support tools are insufficient to address novel difficulties [5]. These days, a broad spectrum of enterprises employ virtual reality for a variety of purposes, including education, decision support, product review, maintenance assistance, reproducing industrial activities, enhancing resource planning, and more. Virtual reality has demonstrated its usefulness in practical production. Lastly, evaluation techniques are crucial for producing insightful data and carrying out an unbiased investigation of the impacts of exoskeletons in a real-world workplace [6], [7]. The majority of recent research lacks standardized parameters or higher sample sizes [8], [9].

In order to benchmark exoskeletons uniformly, the EVO-MTI project has taken the lead in this regard and created a framework and methodology for harmonizing the testing of assistive robotic systems performance. Improving robotic systems' comparability and reproducibility as well as developing a new system is the aim. This paper aims to create various physical and digital simulations, integrate the resulting data into a comprehensive database, and introduce a methodology for using this data in exoskeleton development.

## II. METHODOLOGY

Relevant hardware-level infrastructure components are further introduced in the sections that follow. Additionally, relevant analytical techniques that can offer valuable data for simulation models are described. A variety of physical prototypes that can reproduce different scenarios are used, as seen in Figure 1. These prototypes simulate scenarios with and without (S1) assistance. Then, a simulation of the exoskeleton (S2A) is performed, and then a simulation of the human model (S2B). Ultimately, these prototypes are combined into a single, comprehensive simulation (S3) [10]. These parts are implemented individually or together depending on the iterative stage of the transdisciplinary development process.

### A. Simulated Use Case

In this scenario, different tasks are replicated and can be performed without or with exoskeletons in a standardized laboratory environment. A test parkour is used to summarize different physically demanding tasks like baggage lifting, nursing, construction or assembly line. Users can use different exoskeletons while performing the same task.

For this, use cases play a vital role in the development and evaluation of exoskeletons, providing real-world scenarios that guide the design and evaluation of these devices. By analysing specific use cases, developers can identify common repetitive motions, errors associated with it and stress induced in muscles during the task allowing for targeted improvements in exoskeleton design. The importance of simulating use cases is listed below:

- Design Optimization: Use cases help in understanding the specific demands of different tasks, allowing engi-

neers to design exoskeletons that can support and enhance repetitive motions effectively. This includes optimizing the range of motion, joint alignment, and load distribution to minimize stress and strain.

- Performance Evaluation: Evaluating exoskeleton performance in real-world use cases is essential for identifying potential errors and inefficiencies. This involves assessing the device's ability to reduce fatigue, enhance performance, and maintain consistency in repetitive tasks.
- User-Centred Design: Engaging users in specific use cases provides valuable feedback on the usability and comfort of exoskeletons. This feedback is critical for making iterative improvements that ensure the device meets the needs and expectations of users in performing repetitive motions.

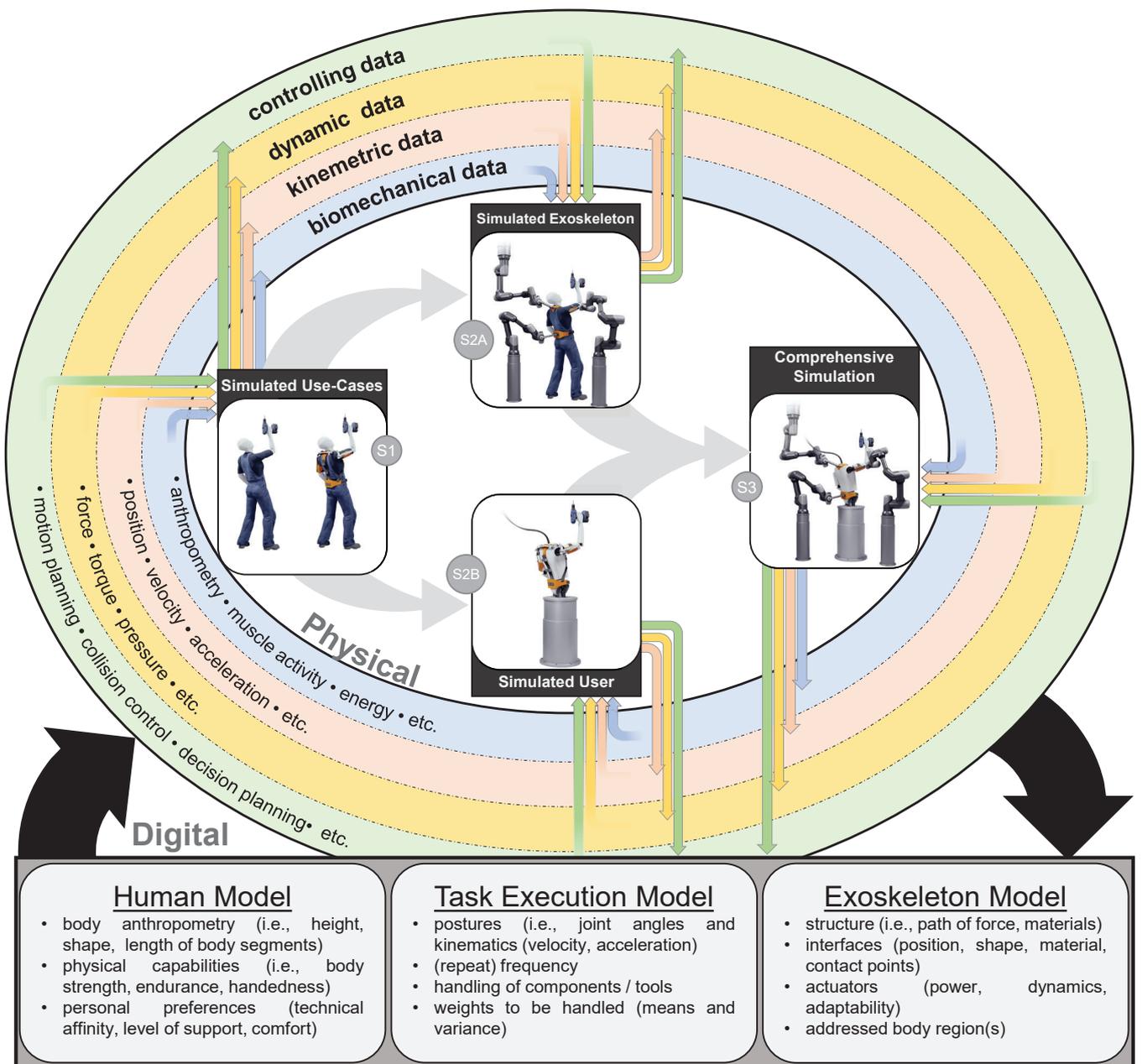


FIGURE 1: DATASTREAM AND ELEMENT CONNECTION WITHIN THE DEVELOPMENT ENVIRONMENT

### B. Simulated Exoskeleton

Besides using real exoskeletons, the simulation of exoskeletal characteristics in a robot-based environment by using cobots is implemented in this scenario. Relevant exoskeletal properties are modelled and simulated by multiple cobots where interaction occurs at the coupled elements that is the exoskeletal interfaces. Parameterizable properties of the exoskeleton such as joint states, and control parameters acting at the same physical interfaces are needed to simulate the exoskeleton and it is carried out in an environment that tracks motion trajectories and provides defined forces and torque.

### C. Simulated User

Repetitive human motion is integral to many use cases but is often accompanied by challenges such as stress, strain, and errors. The development of exoskeletons that can effectively support these motions requires a deep understanding of biomechanics and the specific demands of various use cases. By leveraging detailed use case analysis, developers can design exoskeletons that enhance performance, reduce physical strain, and minimize errors in repetitive tasks. This approach not only improves the functionality and safety of exoskeletons but also ensures they are tailored to meet the real-world needs of the users. Challenges in repetitive motion are as follows:

- **Stress and Strain:** Continuous repetition of the same motions can lead to muscle stress and strain, increasing the risk of injuries. This is particularly problematic in industrial settings where workers are required to perform repetitive tasks for long hours.
- **Biomechanical Variability:** Individual differences in anatomy and biomechanics mean that the same task can be performed differently by different people. This variability can lead to inconsistent performance and potential inaccuracies.
- **Motor Learning and Control:** Learning and controlling repetitive motions involve complex motor processes that can be prone to errors. These errors can result from incorrect movement patterns, lack of coordination, or inadequate feedback mechanisms.

In order to reduce the tiredness of the user while performing a repetitive task, achieving standardized movements to improve the possibilities of comparing different systems as well as to measure the additional aspects like joint torque, humanoid-like robot simulate human movements.

### D. Comprehensive Simulation

Humans, tasks and exoskeletons are physically simulated simultaneously in this scenario. Simulation is done using cobot coupling and testing platforms. This stage highly depends on the S2A and S2B's technical realization.

## III. IMPLEMENTATION

### A. Simulation of exoskeleton properties

In order to simulate the exoskeleton properties, a Robot Operating System (ROS) based platform as shown in Figure 2 is developed for simulating the exoskeleton properties and controlling the cobot accordingly. The Vicon node streams the real-time data of the markers detected in the motion capture environment and calculates the shoulder arm angle accordingly. This data is then used to calculate the simulated support torque in the Exo Node. The necessary support characteristics in the form of different support curves can be imported in this

node. The human pose and support torque are used as an input for generating the trajectory and torque of the cobot. A force torque sensor attached between the end effector of the cobot and the interface helps formulate a feedback loop for the control of the cobot.

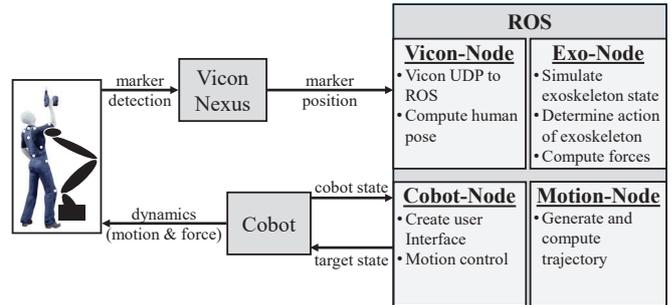


FIGURE 2: ROS ARCHITECTURE FOR SIMULATING EXOSKELETAL PROPERTIES

The Exo node can simulate multiple exoskeletons that provide different support characteristics at different interfaces. In the implementation below, the shoulder exoskeleton Lucy [11] is used for the simulating support torque based on the angle of the arm. The use case used here is the above-head drilling motion. The user is attached with markers at the arm, shoulder and hip as shown in Figure 3 in the form of red spheres. This is done in order to gain the shoulder angle of the user while the user is performing the task in the Vicon motion capture environment.

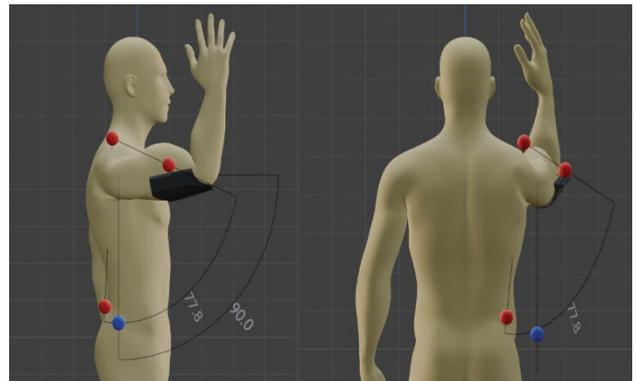


FIGURE 3: SIMULATION OF ARM ANGLE

In Figure 3, a visualization of the human model in Blender is shown. The shoulder arm angle calculated in the Vicon node in ROS is realized with respect to the actual arm angle. Here, the actual shoulder angle (90°) and calculated shoulder angle (77.8°) are illustrated. This error occurs since the shoulder joint of a human being, having seven Degrees of freedom (DOF), is a complex joint to simulate, the human error caused during the placing of markers that are not aligned as per the actual joints and the inaccuracy caused by motion capture due to blind spots.

Figure 4 illustrates the graph between the actual arm angle vs. the arm angle calculated by the Vicon node in ROS based on the three-marker set. The red and blue plots are data acquired by two test persons. In order to fit the curve, a regression is performed which is then exhibited by the grey line.

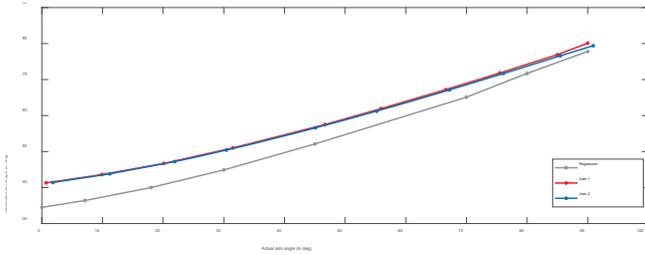


FIGURE 4: REGRESSION FOR COMPENSATION OF ANGLE DIFFERENCE

In order to calibrate this error in the tracking arm angle, a localization algorithm in addition to a regression model is implemented. For this, every test person has to initially hold the arm position at standard angles ( $0^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ ,  $90^\circ$ ). The standard angles can then be manually entered in respect to the tracked arm angle. The calibration algorithm is then implemented on this data to give a localization output to correct the error of the tracking system.

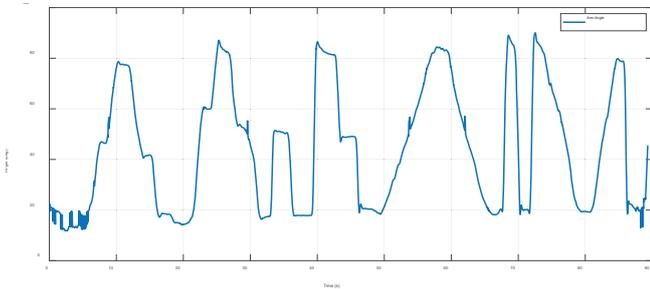


FIGURE 5: CALCULATED ARM ANGLE IN A CYCLIC MOTION

Since the use cases have to be realized by a repetitive motion, it is then observed in Figure 5 that while the test person performs a cyclic drilling motion eight times, the reach of the shoulder arm angle as well as the acceleration of the arm keeps differing. The test person may not be familiarized with the task and is in a new environment, which may cause this human error.

For the purpose of reducing the error in angle calculation and obtaining a standardized use case data for simulating and realizing different scenarios, an approach to simulate a human motion, is discussed.

### B. Humanoid

The implementation of robotic-based systems can facilitate the achievement of high reproducibility of human movements. Collaborative robots (cobots) are utilized to model human-like movement. In addition to their compact dimensions, cobots are distinguished by their versatile parameterization, mobility, and integrability for this purpose. In a methodology outlined in [12] a physical arm, constructed as a 3D-printed part, is attached to the end effector of a seven-axis Franka Panda cobot. As shown in Figure 6, the exoskeleton is positioned on a stand, and the 3D-printed arm is coupled to the arm shells of the exoskeleton. Previously recorded motion data from the motion lab were utilized and implemented to the cobot allowing for the execution of the movement by the 3D printed arm and the collection of data for further investigation.

### C. Virtual Reality

In this section, we shall discuss the implementation of virtual reality (VR) in the development and evaluation process of this project. The virtual reality approach can help reduce

human error. Data analysis to control parameterizable properties, realization of the influence of various sensors that are integrated will get easier. It will also play a crucial role in getting the test person familiarised with the task and guiding them to perform the task and not get distracted due to a wearable robot attached to the body of the test person. Evaluation of exoskeletons requires testing different exoskeletons in different use cases. Building a test station to realize multiple use cases can be difficult due to lack of space. The integration of VR will also benefit with respect to building multiple test stations to realize different use cases. A VR environment will give the option to customize use cases and tools as needed. The position of the controller in the VR environment will also aid as an efficient feedback loop for the arm angle calibration. A safety approach can also be induced in this environment where the user can communicate via a controller when some technical issues are occurring. An alert when a malfunction occurs in the robotic system can be visualized in the VR environment.

The future work in this project is to use the Quest 3 VR headset in combination with a Unity 3D software platform. In Figure 7, a flowchart of the workflow of the VR environment is presented. The idea is to take in user inputs via the VR controller as well as use the motion capture for calibration of the position of the test person in the VR environment. In the next step, the test person will be given the choice to choose which exoskeleton they are currently wearing in the VR environment. All data such as the accuracy of the task performed, acceleration, and position of the controller will be logged.

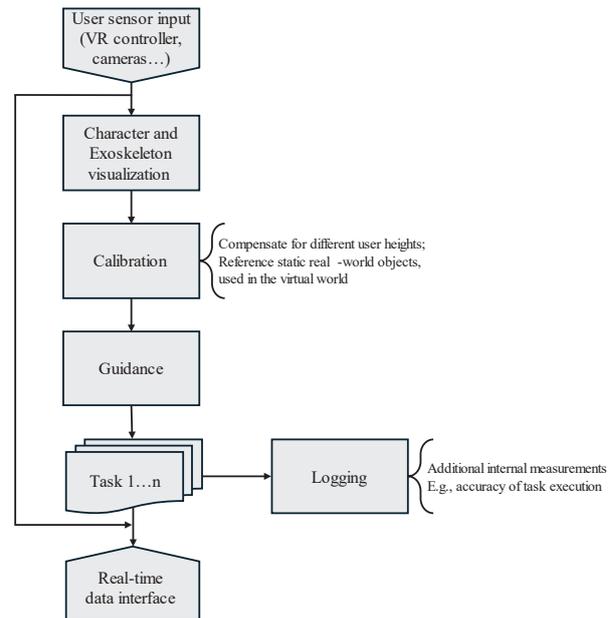


FIGURE 7: VIRTUAL REALITY FLOWCHART

## IV. CONCLUSION

Despite the number of studies carried out to date, there are still a number of issues with research methodology and exoskeleton evaluation. An assessment of the systems in actual working environments is required to comprehend their consequences. Standardized test protocols are also necessary to obtain a sizable and comparable data set.

Towards this goal, this paper presented a plan to collaboratively progress the emerging development and evaluation of

exoskeletons by strengthening the exchange between developers, researchers, and end users and collecting data in the future. In a continuous, multicentric prospective study strategy, this should be helpful to add more data sets.

It will be possible to test and assess exoskeletons in specially designed Parkours, where working conditions are as close to the real-world as possible. This allows research data to be generated specific to an exoskeleton and allows for in-trainividual comparisons of exoskeleton users.

The synergy of various core elements provides a foundation for standardizing use cases, which enables a systematic investigation of exoskeletons. This standardization then allows for the variation of parameters, which in turn facilitates the optimization of support systems at both the hardware and software levels. Furthermore, the integration of these core elements into a virtual reality environment allows for the investigation of human perception and the acceptance of support systems through immersion.

#### ACKNOWLEDGEMENT

This research paper [project EVO-MTI] is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] Pons, J. L. (2008). *Wearable robots: biomechatronic exoskeletons*. John Wiley & Sons.
- [2] Kim, S., Moore, A., Srinivasan, D., Akanmu, A., Barr, A., Harris-Adamson, C., Rempel, D. & Nussbaum, M. A. (2019). Potential of exoskeleton technologies to enhance safety, health, and performance in construction: Industry perspectives and future research directions. *IIEE Transactions on Occupational Ergonomics and Human Factors*, 7(3-4), 185-191.
- [3] Chryssolouris, George, et al. "A hybrid approach to the verification and analysis of assembly and maintenance processes using Virtual Reality and Digital Mannequin technologies." *Virtual and augmented reality applications in manufacturing* (2004): 97-110.
- [4] Chryssolouris, G., Mavrikios, D., Fragos, D., Karabatsou, V., & Alexopoulos, K. (2004). A hybrid approach to the verification and analysis of assembly and maintenance processes using Virtual Reality and Digital Mannequin technologies. *Virtual and augmented reality applications in manufacturing*, 97-110.
- [5] Tolio, T., Ceglarek, D., ElMaraghy, H. A., Fischer, A., Hu, S. J., Laperrière, L., Newmen, S.T. & Váncza, J. (2010). SPECIES—Co-evolution of products, processes and production systems. *CIRP annals*, 59(2), 672-693.
- [6] Masood, J., Dacal-Nieto, A., Alonso-Ramos, V., Fontano, M. I., Voilqué, A., & Bou, J. (2019). Industrial wearable exoskeletons and exosuits assessment process. In *Wearable Robotics: Challenges and Trends: Proceedings of the 4th International Symposium on Wearable Robotics, WeRob2018, October 16-20, 2018, Pisa, Italy 3* (pp. 234-238). Springer International Publishing.
- [7] Planas-Lara, A. E., Ducun-Lecumberri, M., Tomás-Royo, J. A., Marín, J., & Marín, J. J. (2022). Objective techniques to measure the effect of an exoskeleton. In *Wearable Robotics: Challenges and Trends: Proceedings of the 5th International Symposium on Wearable Robotics, WeRob2020, and of WearRAcon Europe 2020, October 13–16, 2020* (pp. 577-581). Springer International Publishing.
- [8] Grazi, L., Chen, B., Lanotte, F., Vitiello, N., & Crea, S. (2019, June). Towards methodology and metrics for assessing lumbar exoskeletons in industrial applications. In *2019 II workshop on metrology for industry 4.0 and IoT (MetroInd4. 0&IoT)* (pp. 400-404). IEEE.
- [9] Hoffmann, N., Prokop, G., & Weidner, R. (2022). Methodologies for evaluating exoskeletons with industrial applications. *Ergonomics*, 65(2), 276-295.
- [10] Hoffmann, N., Latifi, S. M. M., Nagwekar, R., & Weidner, R. (2023, September). Towards a Robotic-Based Development Environment for Designing and Evaluating Exoskeletons. In *ISR Europe 2023; 56th International Symposium on Robotics* (pp. 354-360). VDE.
- [11] Otten, B. M., Weidner, R., & Argubi-Wollesen, A. (2018). Evaluation of a novel active exoskeleton for tasks at or above head level. *IEEE Robotics and Automation Letters*, 3(3), 2408-2415.
- [12] Klankers, M. K., Rudloff, A., Mohammadi, P., Hoffmann, N., Mir Latifi, S. M., Gökay, R., Nagwekar, R., Weidner, R. & Steil, J. J. (2024). Lessons Learned from Investigating Robotics-Based, Human-like Testing of an Upper-Body Exoskeleton. *Applied Sciences*, 14(6), 2481.

# Erfolg braucht Taktik – nicht bloß Masse. Wie die Bundeswehr mit GhostPlay das Defense Metaverse zur Fähigkeitsentwicklung nutzt

Heiko Borchert

*Ko-Direktor Defense AI Observatory  
Helmut-Schmidt-Universität/Universität  
der Bundeswehr Hamburg  
Hamburg, Deutschland  
hb@defenseai.eu*

Christian Brandlhuber

*CTO 21strategies  
München, Germany  
christian.brandlhuber@21strategies.com*

Gary S. Schaal

*Ko-Direktor Defense AI Observatory  
Helmut-Schmidt-Universität/Universität  
der Bundeswehr Hamburg  
Hamburg, Deutschland  
gschaal@hsu-hh.de*

**Kurzfassung** – GhostPlay hat die Taktikentwicklung für rote und blaue Kräfte mit künstlicher Intelligenz (KI) als souveräne deutsche Fähigkeits- und Technologieinnovation etabliert. Im hochauflösenden und hyperrealistischen digitalen Zwilling des Gefechtsfeldes – dem Defense Metaverse – werden KI-basierte Entscheidungsverfahren entwickelt, die es erlauben, operative Fähigkeiten im Hinblick auf ihren Mehrwert für den Missionserfolg zu erproben. Die Interaktion mit einem nicht-kooperativen Gegner, dessen Fähigkeiten sich z. B. aus dem Wissen über gegnerische Kapazitäten, die Auswertung aktueller Konflikte und der Exploration künftig verfügbarer technischer Optionen ergeben, stellt den Kern des Ansatzes dar. Der Beitrag beleuchtet anhand der gemeinsam mit dem Amt für Heeresentwicklung durchgeführten Untersuchung, wie im Defense Metaverse Konzepte und Fähigkeiten entwickelt werden, um die Leistungsfähigkeit von Hubschraubern durch den Einsatz von Pseudomunition (Air-Launched Effects, ALE) gegen rote bodengebundene Flugabwehr (Ground-Based Air Defense, GBAD) zu steigern. Von entscheidender Bedeutung ist dabei nicht bloß der Einsatz unbemannter Systeme in Massen, sondern vielmehr die Taktik zum Anflug auf das gegnerische Ziel, zur Verwirrung bzw. Überwindung der gegnerischen Sensorik und Effektorik sowie zur emergenten Koordination der Schwarmmitglieder.

**Stichworte** – Defense Metaverse, Fähigkeitsentwicklung, Künstliche Intelligenz, Red Teaming, Selbstkoordination.

## NOMENKLATUR

AHEntwg	Amt für Heeresentwicklung
ALE	Air-Launched Effects
C-UAV	Counter-Unmanned Aerial Vehicle
DTED	Digital Terrain Elevation Data
GBAD	Ground-Based Air Defense
KI	Künstliche Intelligenz
LIDAR	Light Detection and Ranging
MDO	Multi-Domain Operations
NLOS	Non-Line of Sight
SAR	Synthetic Aperture Radar
SDD	Software Defined Defense
UAV	Unmanned Aerial Vehicle

## I. QUANTITÄT ALS QUALITÄT – VORSICHT VOR VOREILIGEN SCHLUSSFOLGERUNGEN!

2014 gehörten Robert O. Work, damaliger stellvertretender US-Verteidigungsminister, und Shawn Brimley mit ihrer Arbeit „20YY: Preparing for War in the Robotic Age“ [1] zu den ersten Autoren, die in Schwärmen autonomer Systeme ein potenziell revolutionäres neues Element der Kriegführung erkannten. Masse, so ihr Argument, zeichne sich als entscheidendes Kriterium für den Erfolg im Kampf gegen einen ebenbürtigen Gegner ab. „Quantität als Qualität“ hat sich seither zu einem Leitmotiv der Streitkräfteentwicklung der EU- und NATO-Staaten sowie ihrer Gegner entwickelt. Unlängst präsentierte Kathleen Hicks, gegenwärtig stellvertretende US-Verteidigungsministerin, mit der Replicator-Initiative einen neuen Vorstoß, der umfassend auf Masse setzt [2]. Auch die Angriffe Russlands gegen die Ukraine seit 2022 und der im April 2024 erfolgte Einsatz von 300 Drohnen und Marschflugkörpern Irans gegen Israel scheinen zu belegen, dass sich diese Überlegung auch in aktuellen Konflikten widerspiegelt.

Ohne Zweifel spielt Masse eine wichtige Rolle, um z. B. gegen einen leistungsfähigen Gegner zu bestehen, etwa indem militärische Fähigkeiten verteilt und über verschiedene Dimensionen koordiniert zum Einsatz kommen. Aber, und diese einschränkende Frage ist wichtig, ist es tatsächlich allein die Masse, die den Ausschlag geben kann?

NEIN – lautet die Antwort, die der vorliegende Beitrag darauf gibt. Umfassende Arbeiten, die gemeinsam mit dem Amt für Heeresentwicklung (AHEntwg) im Rahmen des Vorhabens Wild Hornets durchgeführt wurden, deuten – mit Blick auf die spezifische Interaktion zwischen Drehflüglern und Pseudomunition (Air-Launched Effects, ALE) gegen bodengebundene Flugabwehr (Ground-Based Air Defense, GBAD) – klar darauf hin, dass dies nicht der Fall ist. Entscheidend ist vielmehr, dass diese Masse mit kluger Taktik eingesetzt wird. Dieser Hinweis ist umso relevanter, als er verdeutlicht, wie das Defense Metaverse, der digitale Zwilling des Gefechtsfeldes, der Elemente der analytischen und konstruktiven Simulation mit leistungsstarkem, KI-basiertem Red Teaming kombiniert, Planer und Konzeptentwickler vor Trugschlüssen bewahren kann, die weitreichende Folgen für die Streitkräfteentwicklung haben. Wild Hornets und die Fähigkeitsentwicklung im Defense Metaverse stehen damit stellvertretend für neue

Möglichkeiten, die Software Defined Defense (SDD) den Streitkräften bietet, um sich – ausgehend von analytische Konzepten, Designparametern über neue Systementwicklungen bis in den Einsatz und zur Ausbildung – auf eine grundlegende neue strategische Ausgangslage in Europa und die damit verbundenen Folgen für den Einsatz von Streitkräften im Rahmen der Landes- und Bündnisverteidigung [3] vorzubereiten.

## II. DEFENSE METAVERSE: WORUM GEHT ES?

1992 prägte Neal Stephenson [4] den Begriff des Metaverse. Er beschrieb damit einen dreidimensionalen virtuellen Raum, in dem Menschen mit Softwareagenten interagieren. Jahrzehnte später ist der Begriff in Form des digitalen Zwilling als digitale Kopie physischer Objekte und Prozesse Realität. Im Kontext der Industrie 4.0 [5] dienen digitale Zwillinge von Produktionsanlagen und -abläufen oder von unternehmerischen Lieferketten beispielsweise dazu, Entwicklungs- und Fertigungsprozesse effizienter zu gestalten oder Lieferengpässe in der Versorgung in der digitalen Kopie frühzeitig zu erkennen und damit im realen Umfeld zu vermeiden.

Das Defense Metaverse setzt auf dieser Idee auf. Mit dem digitalen Zwilling des Gefechtsfeldes wird ein Abbild des Einsatzumfeldes geschaffen, das dieses möglichst realitätsrecht wiedergibt [6]. Dabei geht es um Aspekte wie z. B. ländliche oder städtische Infrastruktur, Terrain, Vegetation, Wetterbedingungen oder realitätsnahe Sensor- und Effekteigenschaften. Von zentraler Bedeutung für das mit GhostPlay aufgesetzte Defense Metaverse ist, Angreifer (rote Kräfte) und Verteidiger (blaue Kräfte) in simulierten Auseinandersetzungen (Multi-Agent-Learning & Deep-Reinforcement Learning) gegeneinander antreten zu lassen. Dabei ist das auf Selbstlernen basierende, nicht-kooperative Verhalten der roten Kräfte entscheidend, um bestehende Ansätze hinsichtlich der Konzeption, der Eigenschaften und der Fähigkeiten der blauen Kräfte kritisch zu hinterfragen und zu überprüfen.

Der im Rahmen von GhostPlay entwickelte digitale Zwilling der Einsatzumgebung unterscheidet sich aufgrund seiner vertikalen und horizontalen Durchlässigkeit von bislang verfügbaren Lösungen der analytischen bzw. konstruktiven Simulation im Verteidigungsumfeld:

- *Vertikale Durchlässigkeit:* Das mit GhostPlay konstruierte Defense Metaverse verbindet analytische und konstruktive Aspekte in einem neuen Aggregations- bzw. Disaggregationsansatz. Das erlaubt es, taktische Verhaltensweisen bei hoher Dichte an exemplarischen Szenarien auf ihre statistischen Erfolgsaussichten zu untersuchen. Gleichzeitig bleibt die Analyse von Einzelszenarien weiterhin möglich. Damit behebt das Defense Metaverse ein Problem von Wargaming und Ansätzen der analytischen Simulation: Im Unterschied zu diesen hängen die Ergebnisse nicht mehr von Analyteneinschätzungen ab, die verzerrt sein können (Bias), sondern werden in Einzelszenarien berechnet und aggregiert. Die sehr hohe Zahl an durchlaufenen Szenarien ermöglicht eine sehr hohe Abdeckung missionsrelevanter Anforderungen, ohne dass deren Aussagekraft aufgrund analytischer Verzerrungen eingeschränkt würde.
- *Horizontale Durchlässigkeit:* Parallel ermöglicht das in dieser Form konzipierte Defense Metaverse auch eine neue Form der cyber-physischen Durchlässigkeit. So können, abhängig von den zu untersuchenden

Fragestellungen, z. B. probabilistische Modelle durch genauere signalbasierte Modelle simuliert werden. Ebenso ist es möglich, z. B. leistungsfähige generative KI-Methoden einzusetzen, um realistische Sensorcharakteristiken im Detail zu modellieren. Die horizontale Durchlässigkeit erlaubt es zudem, wie in Abschnitt V diskutiert wird, die im Defense Metaverse entwickelten KI-Taktiken physisch auf unbemannte Systeme im Einsatz zu übertragen und deren Leistungsfähigkeit unter Echtbedingungen zu bewerten.

Das Defense Metaverse ermöglicht es, Millionen von Szenarien zu durchlaufen, um die analytisch erforderlichen Antworten zu generieren. In konstruktiven Simulationen klassischen Zuschnitts ist dies ein Problem. Der Umstand, dass wesentliche taktische Bewegungen des Gegners in der Regel gescriptet werden müssen, gestaltet die Szenariobildung aufwändig. Dieser Umstand reduziert die Szenariovielfalt drastisch. Das schränkt wiederum die Aussagekraft der Analyse ein, denn diese ist nur für den untersuchten Szenariosatz repräsentativ. Damit werden z. B. überraschende Aktionen eines Gegners oder seine Flexibilität, die bekannt, aber nicht in die Szenarien integriert werden, kaum betrachtet.

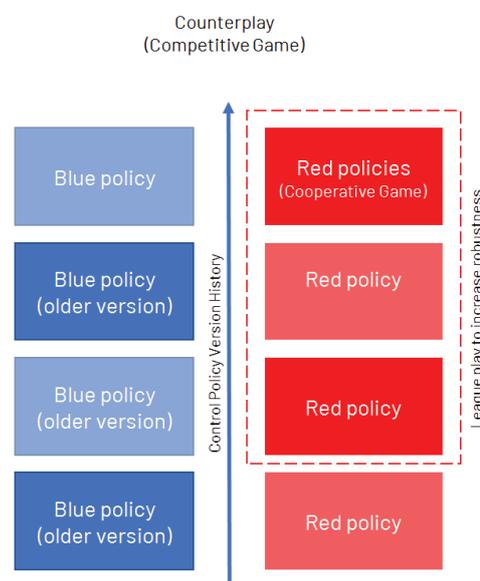


ABBILDUNG 1: COUNTER-PLAY-ANSATZ IN GHOSTPLAY

Um dieses Problem zu beheben, hat GhostPlay die Möglichkeit geschaffen, gegnerische Taktiken von Einzelobjekten, Gruppen und Verbänden im Verfahren des Counter-Play automatisch zu entwickeln (Abbildung 1). Dieser Ansatz erlaubt eine deutlich größere Szenarioabdeckung als klassische Simulationen. Außerdem eröffnet dieser Ansatz die Möglichkeit, Szenarien mit einer sehr hohen Anzahl koordinierter Objekte auf Ebene der jeweiligen Objektdetails zu simulieren und dadurch z. B. Kampfeinsätze hoher Intensität abzubilden.

## III. WILD HORNETS: DIE ANFORDERUNGEN DES AMTES FÜR HEERESWICKLUNG

Konzeptioneller Bezugspunkt für Wild Hornets (wilde Hornissen) sind die laufenden US-amerikanischen Bemühungen zur Entwicklung eines künftigen Ökosystems der Drehflügler. Im Kern geht es dabei um das Zusammenwirken von bemannten und unbemannten Systemen, sogenannten Air-Launched Effects (ALE) [7], [8]. Ausgehend davon entwickelt

das Amt für Heeresentwicklung (AHEntwg) seit 2022 eigene Konzeptvorstellungen zum taktischen Einsatz von ALE. Diese Arbeit sollte GhostPlay mit der Bereitstellung seines Simulationsumfeldes und der Entwicklung geeigneter KI-Taktiken unterstützen. In der Zusammenarbeit sollte aufgezeigt und nachgewiesen werden, welchen fähigkeitssteigernden Mehrwert KI für die Konfiguration und die Steuerung bzw. Führung von ALE im Einsatz gegen die leistungsfähige rote GBAD schafft. Daraus leiteten sich zwei zentrale Aufgabenstellungen ab:

- Evaluieren geplanter Konzepte für den taktischen ALE-Einsatz im Zusammenwirken mit bemannten Systemen gegen eine leistungsfähige rote GBAD.
- Szenariobezogenes Beschreiben und Simulieren der Schlüsselparameter (z. B. Sensorik, Effektorik), des Umfangs, der Konfiguration sowie des taktischen Verhaltens verschiedener ALE-Optionen.

Um die mit GhostPlay entwickelten Entscheidungsverfahren für Wild Hornets zu nutzen und die analytische Simulation durchzuführen, hat das AHEntwg spezifische Vignetten entwickelt und bereitgestellt sowie hierfür unterschiedliche Forschungsfragen definiert.

Die Basisvignette definiert dazu die gegnerische GBAD. Zu Beginn wurde der für GhostPlay umfassend modellierte Gepard Flugabwehrkanonenpanzer eingesetzt. Dieser GBAD-Typ wurde während der Projektentwicklung mit weiteren Komponenten ergänzt, die im nächsten Abschnitt diskutiert werden. Ferner umschreibt die Basisvignette in allgemeiner Weise die einzusetzenden ALE. Für die GBAD und die ALE wurden jeweils Reichweite und Sensorik spezifiziert, für die ALE zusätzlich deren Geschwindigkeit. Zudem zeigt eine taktische Geländekarte auf, wo die Stellungen der gegnerischen GBAD vermutet werden. In der Ausgangslage schützten die GBAD ein gegnerisches Hochwertziel. Die angreifenden ALE hatten die Aufgabe, die vorgelagerte GBAD zu neutralisieren bzw. zu zerstören, um auch das Hochwertziel ins Visier zu nehmen.

Die unterschiedlichen Vignetten beschreiben die Art und Weise, wie die Ziele angegriffen werden sollten, in

zunehmend anspruchsvolleren Missionen. Dabei ging es darum, die Leistungsfähigkeit der KI-Taktik umfassend beurteilen zu können, nicht darum, eine bestimmte Angriffstaktik nachzumodellieren. Wie in [9] beschrieben, sollte die KI vielmehr unter Berücksichtigung der generellen Vorgaben für den Missionserfolg frei explorieren, wie sie ihr Missionsziel am besten erfüllt. Die mit Wild Hornets entwickelten ALE-Angriffstaktiken sind damit selbst erlernt und nicht vorgegeben.

#### IV. DEFENSE METAVERSE FÜR WILD HORNETS

Den Überlegungen des AHEntwg liegt ein spezifisches taktisches Anforderungsprofil zugrunde. Das Defense Metaverse musste für diese Anforderungen (z. B. Flüge in Höhe von Baumwipfeln, Vegetationsunterschiede) spezifiziert werden, denn nur damit war es möglich, die taktische Raffinesse zu erforschen, die es einer KI-Taktik ermöglicht, die Einsatzbedingungen optimal zum eigenen Vorteil zu nutzen.

Neben der Modellierung der roten und blauen Kräfte machte dies umfassende Arbeiten an der Modellierung der Einsatzumgebung erforderlich (Abbildung 2). Die vom AHEntwg bereitgestellten Vignetten gehen von einem durchschnittlichen Gelände aus, das für Ost- und West-Europa repräsentativ ist. Konkret wurden Geländeabschnitte in der Region Roth, Allersberg, Neumarkt in der Oberpfalz in Franken/Bayern ausgewählt. Die realitätsnahe Simulation erforderte Geländedaten mit entsprechend hoher Auflösung. Die typischerweise für Simulationen verwendeten Digital Terrain Elevation-Daten (DTED) waren für Wild Hornets nicht ausreichend. Stattdessen wurden Daten benötigt, die idealerweise durch SAR-Satelliten gewonnen und mit LiDAR-Vermessungsflügen augmentiert werden. Konkret griff das Entwicklerteam für Wild Hornets auf Daten des Landesamtes für Digitalisierung, Breitband und Vermessung Bayern zurück, die dieses offen und kostenlos anbietet. In mehreren Entwicklungsschritten wurde die eigene Methode der Terrainerstellung im Defense Metaverse in einen produktionsreifen Prozess überführt, der für den Aufbau einer beliebigen zu simulierenden Einsatzumgebung genutzt werden kann. Mit diesem Prozess kann grundsätzlich innerhalb von ein bis zwei Wochen eine neue Einsatzumgebung mit der hohen Auflösung des



ABBILDUNG 2: MODELLIERUNG DER EINSATZUMGEBUNG FÜR WILD HORNETS

GhostPlay-Ansatzes erstellt werden. Dazu wurden eigenständige Softwareprozesse entwickelt, die die Vorverarbeitung und Aufbereitung aller Eingangsdaten und deren Umsetzung auf ein einheitliches Weltmodell weitgehend automatisch und parallel durchführen. Die so erstellten und transformierten Daten werden mit korrekter Geo-Referenzierung automatisch fusioniert und für die Simulation aufbereitet.

In analoger Weise wurden neben dem Terrain, den Gebäuden und spezifischen Wettereffekten – das GhostPlay Defense Metaverse simuliert Einsätze bei unterschiedlichen Hell-Dunkel-Graden, Schnee, Eisregen, Nebel etc. – auch die roten und blauen Waffensysteme modelliert. Gegenwärtig umfasst die Modellbibliothek dazu folgende Systeme:

- ALE und Lenkflugkörper: Switchblade 600 (Aero-Environment), Altius 600 (Anduril), Hero (Uvision/Rheinmetall), Spike NLOS (Rafael)
- Drehflügler: Tiger und EC145M (beide Airbus Helicopters)
- GBAD: Gepard Flugabwehrkanonenpanzer (KMW), Patriot (RTX Corp und Lockheed Martin), S-400 Triumpf (Almas-Anteil), 96K6 Pantsir S1 und S2 (Rostec)
- Radarsysteme: 76N6 Tiefflugradar (Clam Shell, LEMZ)

Zusätzlich zu diesen bekannten Systemen fließen in Wild Hornets auch Analysen des Einsatzes sogenannter First Person View (FPV)-Drohnen ein. Diese haben sich, wie die Entwicklungen in der Ukraine zeigen, zu einer veritablen Bedrohung für Bodentruppen entwickelt. Sie stellen auch für die GBAD eine erhebliche Bedrohung dar, weil sie noch kleiner als die beschriebenen ALE-Klassen und damit noch schwerer zu entdecken bzw. zu bekämpfen sind. Speziell die aggressiven Flugmanöver dieser Systeme, sofern sie von einem geübten Piloten gesteuert werden, stellen erhebliche Anforderungen an das Target Tracking der GBAD.

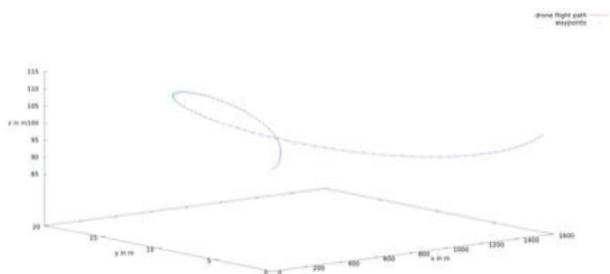


ABBILDUNG 3: TRAJEKTORIE EINER ANN-GESTEUERTEN FPV-DROHNE IN GHOSTPLAY

Gemeinsam mit dem kroatischen FPV-Hersteller Orqa wurden daher die bestehenden Simulationsmodelle von GhostPlay validiert, um die Entwicklung von Anforderungen und Modellen für diese neue Bedrohung zu analysieren. Dazu flogen professionelle FPV-Piloten dieses Herstellers auch im GhostPlay Defense Metaverse. Gestützt darauf wurde eine Entwicklungslinie implementiert, bei der eine Neuronal-Netz-Architektur (ANN) trainiert wird, um die optimalen Flugwege zu einem Ziel zu berechnen und direkt in den FPV-Aktuator-Inputs umzusetzen. Dieser Schritt verfolgte das Ziel, die gegenwärtige GBAD in möglichst kurzer Zeit zu erreichen. Dabei sollte das Netz lernen, auch am aero-dynamischen Limit zu fliegen, falls dies der Zielerreichung dient (Abbildung 3).

## V. ERGEBNISSE: TAKTIK, TAKTIK UND TAKTIK

Die Wild Hornets-Vignetten des AHEntwg waren mit spezifischen Untersuchungsfragen verbunden, die die Konzeptentwicklung des Amtes unterstützen. Dabei geht es um Aspekte wie z. B. die schrittweise oder gleichzeitige Zielbekämpfung, die Anflugrichtung und -höhe, den Formationsflug oder die Zusammensetzung eines ALE-Schwarms (homogen oder heterogen mit ALE unterschiedlichen Typs). Die Hauptergebnisse von mehr als 1,2 Millionen Szenarioläufen, die wir an dieser Stelle nur cursorisch zusammenfassen, sind ebenso eindeutig wie überraschend:

- *Erstens ist Saturierung mit Masse nicht das Allheilmittel:* Die Idee, einen Gegner mit bloßer Masse überwinden zu können, ist nicht zielführend – vor allem dann nicht, wenn der Gegner über eine intelligente Flugabwehr verfügt. Die Szenarien machen deutlich, dass eine qualifizierte GBAD insbesondere für den Nahbereich große ALE-Schwärme, die unvorteilhaft eingesetzt werden, verhältnismäßig leicht zerstört. Langsam und tieffliegende ALE können dagegen unter realistischen Bedingungen sehr effektiv gegen eine solche GBAD eingesetzt werden – insbesondere dann, wenn sie auch Anflughöhe und -geschwindigkeit variieren. Der Schlüssel zum Erfolg liegt dabei in einer Angriffstaktik, die es ermöglicht, das Ziel weitgehend unentdeckt anzufliegen und dabei auch Überraschungsmoment (z. B. Anflug im Sinkflug aus der Sonne) zu nutzen. Das lässt die Schlussfolgerung zu, dass die meisten Demo-Szenarien für die UAV-Abwehr, die zeigen, wie ein C-UAV-System stationär und meist in Formation neben- oder hintereinander fliegende Quadrocop-ter zerstört, die Realität nur ansatzweise korrekt wiedergeben.
- *Zweitens können ALE als „Türöffner“ genutzt werden:* ALE können für die Zieleinweisung anderer Lenkflugkörper erfolgreich genutzt werden. So wurde mit Wild Hornets eine Taktik entwickelt, mit der ALE ein Ziel aufklären und die entsprechenden Positionsdaten an einen Lenkflugkörper übermitteln, der das Ziel zerstört. Die ALE profitieren dabei von der Möglichkeit, sich dem Ziel weitgehend unerkannt nähern zu können. Entscheidend ist aber auch hierfür die bereits erwähnte Möglichkeit des Tief- und Tiefstfluges der ALE, die – nach öffentlich frei verfügbaren Informationen – gegenwärtig noch nicht gegeben ist.
- *Drittens ist ALE-Kooperation entscheidend:* Das mit deutlichem Abstand am klarsten herausgearbeitete Ergebnis ergibt sich aus dem Umstand, dass Kooperation für den Erfolg des ALE-Schwarms gegen eine (auch intelligente) GBAD entscheidend ist. Verstehen es die Schwarmmitglieder, ihre Aufgaben und Sub-Aufgaben, die im Rahmen der Mission zu erfüllen sind, kooperativ selbst zuzuteilen und damit verteilt und koordiniert zu agieren, sind sie am effektivsten, weil Rollen und Timing einzelner ALE situativ abgestimmt und neu zugewiesen werden können. Dabei verdeutlichen die Simulationsläufe, dass – nach heutigem Stand der Technik – ein sich selbst koordinierender und kooperativer ALE-Schwarm mit 12–16 Mitgliedern das Missionsziel mit höherer Wahrscheinlichkeit erreicht als ein Schwarm mit 100 Mitgliedern, der nicht über diese Fähigkeit verfügt.

Aus der Gesamtbetrachtung dieser drei Ergebnisse resultiert die Einsicht, dass ein selbstlernender, kooperativer und sich koordinierender Aufklärungs- und Zielbekämpfungsschwarm (Recce-Strike Swarm) sehr gute Aussichten auf hohe Missionserfolge aufweist. Daher wurde mit Wild Hornets auf Wunsch des AHEntwg zusätzlich eine sophisticatede Aufklärungstaktik entwickelt, die wesentlich zu diesem Ergebnis beiträgt. Diese basiert nicht auf dem gegenwärtig bekannten Sweep-Ansatz, wonach UAV für die Aufklärung ein Gelände in regelmässigen parallelen Trajektorien abfliegen (Abbildung 4, links). Dieser Ansatz macht die eigenen Kräfte für den Gegner in hohem Masse berechenbar und damit verwundbar.

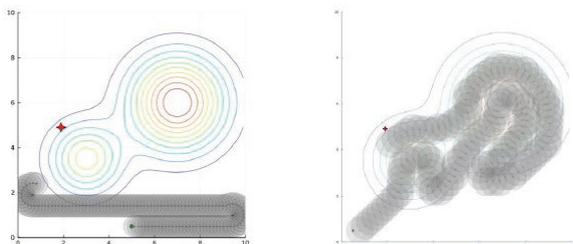


ABBILDUNG 4: AUFKLÄRUNGSSUCHSTRATEGIE SWEEP (LINKS) IM VERGLEICH MIT DEM WILD HORNETS-ANSATZ (RECHTS)

Im Unterschied dazu nutzt die neue Strategie Vorwissen über die möglichen Positionen gegnerischer Stellungen zur optimalen Steuerung der Aufklärungsdrohnen. Diese neue Suchstrategie (Abbildung 4, rechts) plant auf dieser Basis einen optimalen Suchpfad. Für überflogene bzw. beobachtete Positionen, an denen kein Ziel entdeckt wurde, wird die Wahrscheinlichkeit reduziert, dass sich das Ziel dort aufhält. Dabei wird auch die Möglichkeit berücksichtigt, dass sich das Ziel an der Position befindet, aber nicht entdeckt wurde. Im Verlauf der Suchmission kristallisieren sich die potenziellen Zielstandorte immer deutlicher heraus. In der Summe ist dieser Ansatz für den Gegner deutlich schwieriger als Suchmission erkennbar, denn das Verhalten der Aufklärungsdrohne erscheint randomisiert. Mit Wild Hornets wurde insbesondere eine kooperative Suchstrategie entwickelt, durch die mehrere Drohnen mit unterschiedlichen Sensoren ihre Anflugpfade so koordinieren, dass ein vorgegebenes Suchgebiet in kürzester Zeit aufgeklärt und eine optimale Sensor-Effektor-Zu- und Einweisung mit minimaler Latenz erreicht werden kann. Abbildung 5 zeigt die resultierenden Angriffstrajektorien, die selbst von einer leistungsfähigen SHORAD kaum abgewehrt werden können

Im Ergebnis liegt dank Wild Hornets inzwischen eine kombinierte Aufklärungs- und Zielbekämpfungstaktik vor, die für den Einsatz von ALE und gemischten Schwärmen in großen Stellungsräumen eingesetzt werden kann. Am 2. Mai 2024 wurden erste Auskoppelungen dieser Taktik beim Innovationsstag des Wehrwissenschaftlichen Instituts für Werk- und Betriebsstoffe (WIWeB) erfolgreich auf einen echten Drohenschwarm übertragen (cyber-physische Durchlässigkeit) und in einer Live-Demonstration vorgeführt.

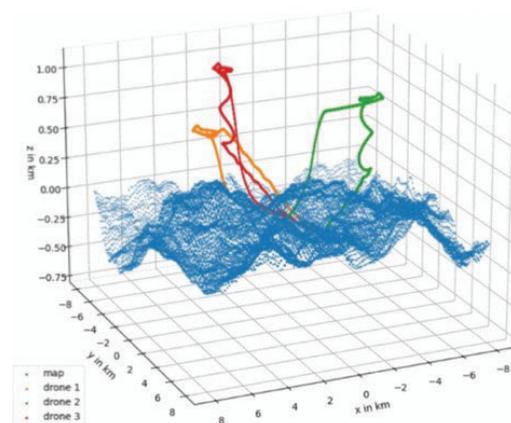


ABBILDUNG 5: SELBSTKOORDINIERENDER SCHWARM MIT AUFKLÄRUNGS- UND ZIELBEKÄMPFUNGSDROHNEN

## VI. AUSBLICK

Das hier am spezifischen Anwendungsbeispiel des Einsatzes von ALE in Kombination mit Drehflüglern vorgestellte Defense Metaverse kann zur Unterstützung vielfältiger Aufgabenstellungen der Bundeswehr genutzt werden. Neben der in diesem Aufsatz diskutierten Konzept- und Fähigkeitsentwicklung ist es möglich, das Defense Metaverse für die kontinuierliche Validierung und Verbesserung von Konzepten, Designs und technischer Parameter während der Entwicklung bzw. des Einsatzes neuer Waffensysteme zu nutzen. Parallel zur Entwicklung von GhostPlay und Wild Hornets ist es dem Projektteam auch gelungen, die Darstellung einer leistungsfähigen roten Flugabwehr anhand des beschriebenen Ansatzes in die laufenden Arbeiten für das Future Combat Air System (FCAS) und das Multi Ground Combat System (MGCS) zu integrieren.

Zusätzlich ist in der beantragten Verlängerung von GhostPlay für die Jahre 2025-26, die noch bewilligt werden muss, eine wesentliche konzeptionelle Weiterentwicklung angedacht. In dieser Phase liegt der Schwerpunkt in der Dimension See insbesondere mit dem Ziel

- des Dimensionentransfers der selbsterlernten Taktiken in das Über- und Unterwasserumfeld und der Steigerung der Leistungsfähigkeit dieser Technik für dimensionsübergreifende Operationen (Multi-Domain Operations, MDO);
- die selbsterlernten kooperativen Taktiken zu erweitern (1) von der Steuerung dimensionsspezifischer Schwärme zu dimensionsübergreifenden, sich (vollständig) selbst steuernden Schwarmverbänden sowie (2) zum Schutz vor dimensionsübergreifend agierenden Schwärmen des Gegners (Multi-Domain Counter Swarming).

## VII. FAZIT: GREIFBARER FÄHIGKEITSGEWINN FÜR DIE BUNDESWEHR

GhostPlay hat die Entwicklung von KI-Taktiken für rote und blaue Kräfte als souveräne deutsche Fähigkeits-Technologieinnovation etabliert. GhostPlay zeigt, dass KI lernen kann, komplexes taktisches Verhalten in Interaktion mit einem intelligenten Gegner unter Bedingungen (1) eines nur teilweise beobachtbaren Zustandsraums und (2) unvollständiger Information zu erzeugen. Militärische Experten bewerten

diese Taktiken als relevant, realistisch und Erfolg versprechend. Mit dem Einsatz eines KI-gesteuerter Red Teaming u. a. für die Konzept- und Fähigkeitsentwicklung oder der Analyse des taktischen Einsatzwertes geplanter Verteidigungssysteme hat sich GhostPlay eine Alleinstellung erarbeitet, die sich auch in der Überführung der Technologie in laufende Entwicklungsvorhaben der Bundeswehr und dem wachsenden Interesse von Streitkräften der EU-/NATO-Staaten niederschlägt. Aus GhostPlay resultiert – als Ergebnis der Zusammenarbeit mit dem AHEntwg – eine gefechtsaugliche Aufklärungstaktik für ALE. Diese plant ausgehend von Vorwissen einen optimalen Suchpfad, um feindliches Gelände nach möglichen GBAD-Elementen zu untersuchen. Parallel wurde diese Aufklärungstaktik mit einer Zielbekämpfungstaktik erweitert, so dass ein Schwarm aus unterschiedlichen unbemannten Systemen (autonom) gegnerische Ziele erfolgreich aufklären und bekämpfen kann. Erste Auskoppelungen dieser Taktiken wurden am 2. Mai 2024 beim Innovationstag des WIWeB im Fliegerhorst Erding mit UAV erfolgreich live geflogen.

#### DANKSAGUNG

Die Autoren danken General Klaus Frauenhoff, Oberst Guido Krahl und Oberstleutnant Marcus Preuß für die ausgezeichnete Zusammenarbeit und die Zustimmung zur Veröffentlichung ausgewählter Projektergebnisse. Diese Forschungsarbeit wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

#### LITERATUR

- [1] Work, R. O.; Brimley, S., 20YY. *Preparing for War in the Robotic Age*, Washington, DC, Center for New American Security, 2014.
- [2] Hicks, K., „Unpacking the Replicator Initiative“, Arlington, Virginia, 6. September 2023, <https://www.defense.gov/News/Speeches/Speech/Article/3517213/deputy-secretary-of-defense-kathleen-hicks-remarks-unpacking-the-replicator-initi/> (Zugriff am 16. April 2024).
- [3] Integrierte Sicherheit für Deutschland. *Nationale Sicherheitsstrategie*, Berlin, Bundesregierung, 2023.
- [4] Stephenson, N., *Snow Crash*, New York, Bantam Books, 1992.
- [5] Garrels, K., Der Digitale Zwilling in Industrie 4.0. Eine Kurz-Einführung zu Merkmalen, Teilmodellen und Verwaltungsschalen, Präsentation, *Plattform Industrie 4.0*, 22. Oktober 2020, [https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/presentation\\_der-digitale-Zwilling.html](https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/presentation_der-digitale-Zwilling.html) (Zugriff am 16. April 2024).
- [6] Henckel, O., „KI-Gefechtssimulation für die Bundeswehr: GhostPlay“, *Europäische Sicherheit und Technik*, Juli 2023, S. 39-40.
- [7] Bassler, Ch; McCarton, M.; Sharp, T.; Chang, J., *Living within one's means: revisiting defense acquisition and affordability A case study of the Army's future vertical lift program*, Washington, DC., CSBA, 2021.
- [8] Judson, J., „US Army leaning into launched effects for modernized battlefield“, *Defense News*, 27 März 2024, <https://www.defense-news.com/digital-show-dailies/global-force-symposium/2024/03/26/us-army-leaning-into-launched-effects-for-modernized-battlefield/> (Zugriff am 16. April 2024).
- [9] Borchert, H.; Brandlhuber, Ch.; Brandstetter, A.; Schaal, G. S., Free Jazz on the Battlefield. How GhostPlay's AI Approach Enhances Air Defense, Hamburg, *Defense AI Observatory*, 2022, [https://defenseai.eu/daio\\_study2203](https://defenseai.eu/daio_study2203) (Zugriff am 16. April 2024).

# Systematische Entwicklung einer Architektur für die virtuelle Inbetriebnahme eines automatisierten Produktionssystems

Omar Ismail, Lasse Beers, Felix Gehlhoff

*Institut für Automatisierungstechnik*

*Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg*

Hamburg, Deutschland

omar.ismail@hsu-hh.de

Nihar H. Shah

*Airbus Operations GmbH*

Hamburg, Deutschland

nihar.shah@airbus.com

Alexander Fay

*Lehrstuhl für Automatisierungstechnik*

*Ruhr-Universität Bochum*

Bochum, Deutschland

alexander.fay@rub.de

**Zusammenfassung**—Die Unternehmen der Fertigungsindustrie stehen heute unter erheblichem Druck, einerseits hohe Qualität bei komplexen Produktionsprozessen sicherzustellen und andererseits einen schnellen Ratenhochlauf zu gewährleisten. Um diese Herausforderungen effizient und fehlerfrei zu bewältigen, gewinnt die Virtuelle Inbetriebnahme (VIBN) immer mehr an Bedeutung. Eine VIBN-Architektur ermöglicht eine tiefgehende Analyse des SPS-Codes und des Ressourcenverhaltens mithilfe von 3D-Modellen und Steuerungsimulationen vor der tatsächlichen Implementierung, wodurch Fehler während der Inbetriebnahme signifikant reduziert werden können. Das Fehlen von standardisierten Vorgehensweisen für die Entwicklung solcher VIBN-Architekturen stellt dennoch eine Herausforderung dar. Ziel dieses Beitrags ist die Entwicklung einer anpassungsfähigen VIBN-Architektur, die speziell auf die Anforderungen der Flugzeugfertigung zugeschnitten ist. Dabei zeigt der Beitrag die Wirksamkeit der Anwendung der SPES-Methodik (Software Platform Embedded Systems) zur Entwicklung einer solchen Architektur auf und identifiziert gleichzeitig Bereiche zur Verbesserung, insbesondere bei der Entscheidungsunterstützung. Der Beitrag bietet damit einen Leitfaden für Forscher und Praktiker im Bereich der VIBN von automatisierten Produktionssystemen, indem er eine strukturierte Vorgehensweise zur Entwicklung komplexer Systeme aufzeigt.

**Stichworte**—Virtuelle Inbetriebnahme, Methodik der Software Platform Embedded Systems, modellbasierte Systementwicklung, automatisierte Produktionssysteme, Flugzeugfertigung

## I. EINLEITUNG

Die Fertigungsindustrie und insbesondere die Flugzeugfertigung zeichnen sich durch ein hohes Maß an Präzision bei gleichzeitiger Forderung nach hoher Qualität, Effizienz und Anpassungsfähigkeit bei der Montage komplexer Bauteile aus. Gleichzeitig besteht bei der Einführung neuer Produktionssysteme ein hoher Druck, die geforderten Produktionsraten zu erreichen und eine möglichst komplikationslose Hochlaufphase zu gewährleisten. [1]. Insbesondere in der Strukturmontage von Rumpfsektionen mit zahlreichen Löchern, die präzises

Bohren und Nieten erfordern, ist der Bedarf an sorgfältiger Planung und fortschrittlichen Automatisierungssystemen hoch [2], [3]. Speziell diese Automatisierungssysteme weisen oft mehrstufige Hierarchien auf, was die Komplexität und damit auch die Wahrscheinlichkeit von Softwareproblemen erhöht [4]. Inmitten dieses Umfeldes wird die Bedeutung der VIBN deutlich. VIBN ist eine Technologie zur Reduzierung von Hochlaufzeiten durch Parallelisierung von Entwicklungs- und Inbetriebnahmeprozessen [1]. Die speziell für bestimmte Anwendungen entwickelte VIBN kann Fehler während der Inbetriebnahme signifikant reduzieren, indem eine gründliche Analyse des SPS-Codes und des Ressourcenverhaltens mithilfe von 3D-Modellen und Steuerungsimulationen vor der eigentlichen Implementierung ermöglicht wird [1]. Obwohl Methoden zur Entwicklung von VIBN-Architekturen existieren, sind sie in der Regel domänen- oder anwendungsspezifisch und berücksichtigen nicht die vielfältigen Anforderungen der Flugzeugfertigung [5]. Die Modellbasierte Systementwicklung (Model-based Systems Engineering, MBSE) bietet wertvolle Methoden und Ansätze zur Entwicklung komplexer Systeme und legt damit den Grundstein für die systematische Entwicklung von VIBN-Architekturen. Ein bekannter Ansatz aus dem Bereich der MBSE ist die SPES-Methodik, die einen systematischen, Hardware- und Softwarekomponenten umfassenden Lösungsansatz für eingebettete Systeme bietet [6]. Aus diesen Beobachtungen ergibt sich das Forschungsziel, zu untersuchen, wie die SPES-Methodik die Entwicklung einer VIBN-Architektur für komplexe Automatisierungssysteme effektiv unterstützen kann. Dieser Beitrag ist wie folgt gegliedert. Im Abschnitt II werden die Grundlagen von VIBN und SPES vorgestellt. Im Abschnitt III wird die VIBN-Architektur systematisch mit Hilfe der SPES-Methodik entwickelt, die die Problembeschreibung des Anwendungsfalls, die Definition der funktionalen und nicht-funktionalen Anforderungen sowie

eine Beschreibung der abgeleiteten Architektur umfasst. Die entwickelte VIBN-Architektur und die Anwendbarkeit von SPES werden im Abschnitt IV diskutiert, während im Abschnitt V die Ergebnisse zusammengefasst und mögliche neue Forschungsfragen aufgezeigt werden.

## II. GRUNDLAGEN

### A. Grundlagen der virtuellen Inbetriebnahme

Der Inbetriebnahmeprozess stellt sicher, dass ein System, eine Komponente oder eine Anlage entsprechend den betrieblichen Anforderungen geplant, installiert, getestet, betrieben und gewartet wird. Dieser Prozess beginnt in der Regel, sobald das System vor Ort installiert ist. Abbildung 1 vergleicht die reale Inbetriebnahme mit der VIBN für Systemlieferanten und Betreiber. Der Prozess, der VIBN ist ein fortschrittlicher simulationsbasierter Ansatz, der darauf abzielt, den Inbetriebnahmeprozess zu verbessern und bereits frühzeitig im virtuellen Raum gestartet werden kann. Bei der VIBN kommt eine virtuelle Testumgebung hinzu, die Simulationsmodelle und virtuelle Tests beinhaltet. Durch die Verknüpfung von Steuerungs-, Bauteil- und Maschinenmodellen ermöglicht die VIBN eine frühzeitige Überprüfung und Optimierung des Anlagenverhaltens. Die VIBN kann die eigentliche Inbetriebnahme zwar nicht vollständig ersetzen, hilft aber bei der frühzeitigen Erkennung und Behebung von Fehlern, verkürzt die Inbetriebnahmezeit, erhöht die Systemqualität und minimiert die mit der Inbetriebnahme vor Ort verbundenen Kosten. Darüber hinaus kann die VIBN die Kommunikation und Koordination zwischen Lieferanten und Betreibern verbessern, was eine schnellere Umsetzung von Änderungen und fortgeschrittenen Schulungsszenarien ermöglicht [7].

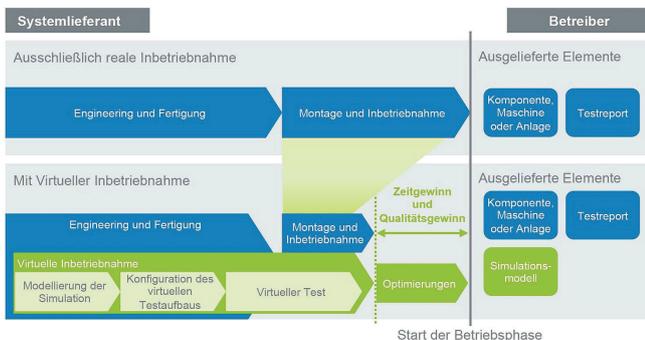


Abbildung 1: Vergleich von realer Inbetriebnahme und VIBN [7]

Der Einsatz von VIBN spielt eine entscheidende Rolle bei der Bewältigung der hohen Komplexität moderner Produktionsanlagen. Diese müssen unter sorgfältiger Berücksichtigung der vorherrschenden Anforderungen entwickelt werden, was die Definition und Charakterisierung wesentlicher Komponenten für die Architektur erfordert.

Die Integration von Verhaltensmodellen, z.B. durch eine Co-Simulationsumgebung, in den VIBN-Workflow ist entscheidend für die Validierung von Steuerungsprogrammen,

durch die umfassende Tests und Fehlerbehebung ermöglicht werden. Auf diese Weise werden die Kommunikationsstruktur und Interoperabilität vor der eigentlichen Inbetriebnahme getestet [8]. Technologien wie ein Functional Mock-Up Interface (FMI) erleichtern eine nahtlose Co-Simulation und AutomationML verbessert den Austausch von Engineering-Daten [9]. Beides ist wichtig im Validierungsprozess. Dazu gehören auch die Integration von Sicherheitsfunktionen, um Systemausfallzeiten zu reduzieren und umfassende Mitarbeiterschulungen [10]. Standard-Technologien, die für Konnektivität und Kommunikation innerhalb von VIBN-Architekturen eingesetzt werden, umfassen Technologien wie OPC UA, TCP/IP oder FMI.

Die Entwicklung einer VIBN-Architektur umfasst mehrere Komponenten. Komponenten wie das Verhaltensmodell, das erweiterte Geometriemodell und die Bus-Emulation ermöglichen die Validierung wichtiger Funktionalitäten, einschließlich der Simulation von Schnittstellen, der internen Verhaltenslogik, der Visualisierung des Systemmodells und der Kommunikation mit realen SPSEN [11]. Virtuelle Darstellungen physischer Objekte, die für VIBN und Digital Twins [12], [13] unerlässlich sind, spielen eine wichtige Rolle bei Simulationen und Tests und verringern die Abhängigkeit von physischen Tests.

Die Entwicklung einer VIBN-Architektur erfordert jedes Mal einen individuellen Ansatz, der auf den spezifischen Anwendungsfall zugeschnitten ist. Faktoren wie die Verfügbarkeit der Emulation des Anlagenverhaltens, die Auswahl der Software, Ressourcenauswahl und Budgetbeschränkungen beeinflussen das Design der VIBN-Architektur. Abhängig davon, ob die Software vom gleichen Hersteller oder von verschiedenen Herstellern stammt, besteht die Notwendigkeit von Co-Simulationsmethoden wie FMIs und Datenschnittstellen [8]. Die Integration des Steuerungssystems in ein Simulationsmodell umfasst im VIBN-Kontext verschiedene Konfigurationen und Methoden [14]. Diese Konfigurationen unterscheiden sich je nach Faktoren wie Typ und Umfang der Modelle, Varianten der Simulation und den Automatisierungssystemen sowie Typ und Umfang der simulierten Automatisierungssystemstrukturen. Folglich ist die Wahl der VIBN-Setup-Konfiguration, wie Model-in-the-Loop (MiL), Hardware-in-the-Loop (HiL), Software-in-the-Loop (SiL) oder Reality-in-the-Loop (RiL), eng mit den Anforderungen, der Notwendigkeit und der Verfügbarkeit der Elemente des Problems verbunden.

### B. Grundlagen der Modellierung nach SPES

Das SPES-Modellierungsframework zielt darauf ab, sprach- und werkzeugunabhängige Modellierungsmethoden für den Entwurf eingebetteter Systeme bereitzustellen [15]. Dieses generische MBSE-Framework findet Anwendung in verschiedenen Domänen. Das Kernkonzept definiert zwei Hauptansätze: Abstraktionsebenen und Sichten, die zusammen einen zweidimensionalen Engineering-Raum bilden [15].

Auf der horizontalen Achse werden die Sichten in vier Abschnitte unterteilt: *Anforderungs-, Funktionale-, Logische-*

und *Technische* Sicht. Die Anforderungssicht ermöglicht es, den Prozess der Erfassung, Dokumentation und Validierung von Anforderungen zu gestalten. Die Anforderungssicht umfasst die Definition des Kontextmodells des Systems, des Zielmodells, der Szenariomodelle (Use-Cases) und des lösungsorientierten Anforderungsmodells. Die *Funktionale Sicht* umfasst Blackbox- und Whitebox-Modelle und konzentriert sich auf die Funktionalitäten des Systems. In der Blackbox wird die funktionale Struktur auf hoher Ebene beschrieben und iterativ in feinere Funktionen zerlegt. Die Whitebox untersucht das Verhalten der Funktionen und berücksichtigt das Ablaufverhalten sowie Eingaben, Ausgaben und Abhängigkeiten zwischen den Funktionen. Die *Logische Sicht* beschreibt die strukturelle Zusammensetzung des Systems, indem das System of Interest (SoI) in logische Komponenten zerlegt wird, die in der Lage sind, die definierten Funktionalitäten auszuführen. Sie ordnet Funktionen den logischen Komponenten zu, definiert Schnittstellen zwischen den Komponenten und beschreibt die Interaktion und das Verhalten von Teilsystemen während der Ausführung von Funktionen. Schließlich konzentriert sich die *Technische Sicht* auf technische Spezifikationen und verfeinert die in der Logischen Sicht modellierten Elemente. Sie definiert Systeme, Teilsysteme und analysiert Systemvarianten hinsichtlich ihrer Funktionalität und Eignung. Die vertikale Achse repräsentiert verschiedene Abstraktionsebenen, die den spezifischen Systemebenen entsprechen. Sie ermöglicht die Modellierung eines vollständigen Systems oder nur eines seiner Designelemente [15].

### III. ANGEPASSTER SPES-ABLAUF FÜR DIE ENTWICKLUNG VON VIBN-ARCHITEKTUREN

Der verwendete Modellierungs-Workflow, der in [16] veröffentlicht wurde, folgt der SPES-Methodik, die im vorherigen Abschnitt beschrieben wurde. Der Workflow umfasst die Entwicklung der vier Sichten und wurde mit Hilfe des Softwaretools Magic Systems of Systems Architect (MSoSA) in der Systems Modeling Language (SysML) implementiert.

#### A. Anwendungsbeispiel

Eines der Hauptziele bei der Anwendung des angepassten SPES-Workflows besteht darin, die Effizienz bei der Entwicklung von VIBN-Architekturen zu steigern und ihn für verschiedene Anwendungsfälle komplexer automatisierter Systeme zu nutzen. Einer dieser Anwendungsfälle ist das iMOD-Projekt, dessen Ziel es ist, die Effizienz, Qualität und Anpassungsfähigkeit der Flugzeugfertigung zu verbessern. Erreicht werden soll dies durch die Einführung eines speziellen Automatisierungssystems für das Bohren und Nieten in der Rumpfmontage sowie durch neue Engineeringansätze [2], [16].

Im Rahmen des iMOD-Projekts wird eine prototypische Montagesstation für die Serienproduktion zukünftiger Flugzeugmodelle aufgebaut. Diese dient als Validierungsumgebung für eine hochautomatisierte Lösung, in der Bohr- und Nietprozesse für die Montage von Flugzeugrümpfen getestet werden. Das automatisierte Montagesystem ist in der Lage,

sowohl die Längs- als auch die Quernaht herzustellen. Die komplexen Prozesse werden von mehreren individuellen Robotern mit multifunktionalen Endeffektoren ausgeführt, die über eine dezentrale Steuerung verfügen und eigenständig Entscheidungen treffen können. Das iMOD-Projekt umfasst auch die Integration wesentlicher IT/OT-Ausrüstung zur Steuerung der automatisierten Komponenten sowie die Implementierung erforderlicher Sicherheitsmaßnahmen für die Plattform.

Die Komplexität des beschriebenen Automatisierungssystems erfordert eine gründliche Validierung vor der eigentlichen Inbetriebnahme. Aufgrund der anspruchsvollen, eng getakteten Inbetriebnahmepläne ist der Einsatz von VIBN daher unumgänglich. Die Entwicklung einer geeigneten VIBN-Architektur ist jedoch eine aufwändige und komplexe Aufgabe, da es keine allgemeingültige Lösung gibt. Sie muss individuell an die Bedürfnisse und Anforderungen der Anlage angepasst werden. Im folgenden Abschnitt wird daher zunächst die Anforderungssicht modelliert, um die notwendigen Anforderungen an die entsprechende VIBN-Architektur zu identifizieren.

#### B. Anforderungssicht

Um alle funktionalen Anforderungen (FA) und nicht-funktionalen Anforderungen (NFA) für die VIBN-Architektur zu identifizieren, schlägt die SPES-Methodik eine Reihe von Schritten vor, einschließlich der Definition des Systemkontexts und der Definition von Anwendungsfällen.

Der Systemkontext wird mit Hilfe von SysML Internal Block Diagrams (ibd) modelliert, siehe Abbildung 2. Dieses Diagramm bietet einen umfassenden Überblick über die zu berücksichtigenden Kontextelemente und ihre Interaktion mit dem zu entwickelnden VIBN-System. Es umfasst externe Systemkomponenten, wie das *Industrial Network*, und Stakeholder, die über eine definierbare Schnittstelle mit dem System interagieren. Der Systemkontext dient als Grundlage für die Ableitung von NFA. Das VIBN muss in das *Industrial Network*, genauer gesagt in eine Product Lifecycle Management (PLM)-Umgebung, integriert werden, um die Kompatibilität mit der etablierten industriellen IT-Infrastruktur zu gewährleisten. Die *Station* verfügt über eine bidirektionale Echtzeitkommunikation mit dem SoI, die die Interoperabilität mit den Steuerungssystemen und den realen Geräten ermöglicht. Aus diesem Modell lassen sich die Anforderungen *NFA4* und *NFA5* ableiten.

Die Stakeholder *Expert User*, *PLC Programmer* und *Offline Programmer* müssen ebenfalls betrachtet werden. Um eine übersichtliche Darstellung zu gewährleisten, wird darauf verzichtet weitere Stakeholder wie den *Bediener* und *Wartungs-Mitarbeiter* in der Abbildung darzustellen. Der *Offline Programmer* muss in der Lage sein, Robotermodelle hinzuzufügen, zu ändern oder zu entfernen. Darüber hinaus müssen sowohl der *Expert User* als auch der *PLC Programmer* die Möglichkeit besitzen Hardware- und Softwaremodule hinzuzufügen, zu ändern oder zu entfernen. Daraus ergibt sich die Anforderung *NFA1* eines modularen Designs und der Skalierbarkeit der Architektur. Darüber hinaus benötigen

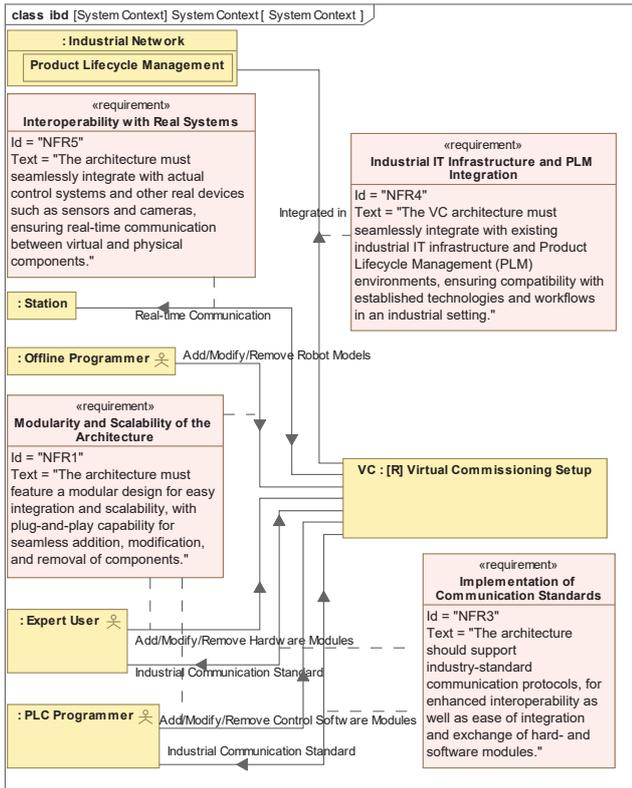


Abbildung 2: Anforderungssicht - Systemkontext [17]

sowohl der *Expert User* als auch der *PLC Programmer* bidirektionale Kommunikationsschnittstellen, welche auf industriellen Kommunikationsstandards basieren. Die Verwendung von industrieeüblichen Kommunikationsprotokollen verbessert die Interoperabilität und ermöglicht die Integration und den Austausch von Hard- und Softwaremodulen. Dies ist in *NFA3* festgehalten.

Die Anforderung *NFA2* (Adaptive und dezentrale Steuerungsarchitektur) ist im System Context Diagramm nicht explizit dargestellt, ergibt sich aber aus der Notwendigkeit einer Steuerungsarchitektur, die eine flexible Anpassung und Integration von Modulen und Komponenten zur Laufzeit ermöglicht. Dies ist besonders wichtig, da einige der hinzugefügten Komponenten teilweise ihre eigene Steuerungszelligenz haben, aber nahtlos in das System integriert werden müssen. Das Ziel ist eine dezentrale Architektur mit hierarchischen und dezentralen Funktionen, um ein effizientes Systemmanagement zu gewährleisten.

Abbildung 3 stellt die Anwendungsfälle des Systems in einem SysML Use Case Diagramm dar. Der Hauptanwendungsfall, *Validate Complete System*, umfasst verschiedene Unteranwendungsfälle. Der Anwendungsfall *Computationally Efficient VC System* zielt darauf ab, die Effizienz und Leistung des VIBN-Setups zu optimieren. Von diesem wird die Anforderung *FA1* abgeleitet. Der Anwendungsfall *Exchange of Simulation Models Across Platforms* bezieht sich

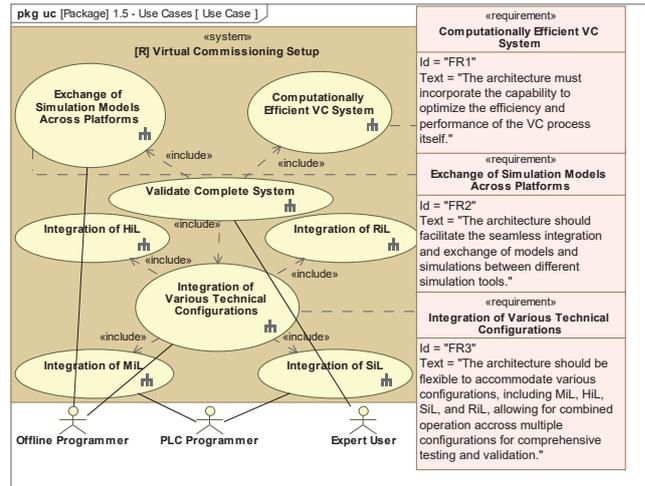


Abbildung 3: Anforderungssicht - Anwendungsfalldiagramm [17]

auf die Notwendigkeit einer nahtlosen Integration und des Austauschs von Modellen zwischen verschiedenen Simulationswerkzeugen und resultiert in *FA2*. Der Anwendungsfall *Integration of Various Technical Configurations* zielt darauf ab, die Flexibilität bei der Aufnahme verschiedener technischer Konfigurationen wie MiL, HiL, SiL und RiL zu gewährleisten und führt zu *FA3*. Das Diagramm hebt auch die Stakeholder hervor, die an diesen Anwendungsfällen beteiligt sind.

*Funktionale und nicht-funktionale Anforderungen* - Im Zuge der Entwicklung einer VIBN-Architektur für den Anwendungsfall iMOD wurden der System Context (vgl. Abbildung 2) und die Use Cases (vgl. Abbildung 3) modelliert und daraus die funktionalen und nicht-funktionalen Anforderungen abgeleitet. Auch wenn diese Anforderungen nur einen Teil der gesamten Anforderungen umfassen, dienen speziell diese als Referenzpunkt für die Bewertung und dem Vergleich mit bestehenden vorhandenen Architekturen in der Literatur. Diese Anforderungen sind so vielseitig, sodass die entsprechende VIBN-Architektur auch auf anderen Anwendungsfälle anwendbar ist, die über die iMOD Station hinausgeht.

- NFA1 Modularität und Skalierbarkeit der Architektur
- NFA2 Anpassungsfähige und dezentralisierte Steuerungsarchitektur
- NFA3 Implementierung von Kommunikationsstandards
- NFA4 Industrielle IT-Infrastruktur und PLM-Integration
- NFA5 Interoperabilität mit realen Systemen
- FA1 Recheneffizientes VIBN-System
- FA2 Plattformübergreifender Austausch von Simulationsmodellen
- FA3 Integration von verschiedenen technischen Konfigurationen

*Analyse bestehender VIBN-Setups* - Bevor die funktionale Sicht definiert wurde, wurden bestehende VIBN-Setups aus der Literatur hinsichtlich ihrer Eignung analysiert. Dazu wurde ein Abgleich der definierten Anforderungen mit den

bestehenden Ansätzen durchgeführt und die Ergebnisse in Tabelle I zusammengefasst. Jeder Ansatz präsentiert seine eigenen Zielsetzungen und spezifischen Anwendungsbereiche, die das Design und den Fokus der jeweiligen Architekturen beeinflussen. Der Vergleich ergab, dass zwar einige Architekturen bestimmte Anforderungen effektiv adressierten, jedoch keiner der Ansätze alle definierten Anforderungen erfüllt. Daraus ergibt sich die Notwendigkeit, einen neuen VIBN-Architekturansatz zu entwickeln, der diese generischen Anforderungen erfüllt und damit sowohl im Kontext des iMOD-Anwendungsfalles als auch in anderen Anwendungsfällen eingesetzt werden kann.

Tabelle I: Vergleich von VIBN-Architektur Ansätzen [17]

Referenz	NFA1	NFA2	NFA3	NFA4	NFA5	FA1	FA2	FA3
Makris et al. [18]	1	1	1	0	0	0	0	0
Ko et al. [19]	1	0	1	0	1	0	0	1
Stiß et al. [9]	1	1	1	0	0	0	1	1
Vermaak and Niemann [20]	1	0	1	0	1	0	0	1
Hofmann et al. [21]	1	0	1	0	1	0	0	1
Illmer et al. [22]	1	0	1	0	0	0	0	0
Barbieri et al. [13]	1	0	1	0	0	0	0	0
Abassi et al. [10]	1	0	1	0	0	0	0	0
Bendjelloul et al. [23]	1	1	1	0	1	0	0	1
Raffaelli et al. [24]	1	0	1	0	1	0	0	1
Ismail and Fay [3]	0	0	1	0	0	1	0	0

Legende: 1 = erfüllt Anforderung, 0 = berücksichtigt Anforderung nicht

### C. Funktionssicht

Die funktionale Sicht vertieft die Umsetzung der identifizierten Anwendungsfälle. In der Blackbox-Perspektive werden Funktionen zunächst in einem Blockdefinitionsdiagramm definiert. Übergeordnete Systemfunktionen werden zunächst zu Funktionen, die dann iterativ in granulare Funktionen, sogenannte atomare Blackbox-Funktionen, zerlegt werden.

In einem zweiten Schritt wird innerhalb der Whitebox-Perspektive der zeitliche und logische Ablauf der Systemfunktionen beschrieben (vgl. Abbildung 4). Der Anwendungsfall *Austausch von Simulationmodellen zwischen Plattformen* aus Abbildung 3 wird hier in vier atomare Blackbox-Funktionen zerlegt. Die erste Funktion generiert das zu untersuchende Modell, die zweite Funktion testet das Steuerungsverhalten des Modells, die dritte realisiert den Austausch und die Integration von Modellen zwischen verschiedenen Simulationswerkzeugen, und die vierte testet das Modell in der 3D-Simulationssoftware.

Die Modellierung der funktionalen Sicht ist ein iterativer Prozess. So ist es möglich, dass innerhalb der Whitebox-Perspektive übersehene Funktionen identifiziert werden. Diese müssen nachträglich modelliert werden. Außerdem hat dies direkten Einfluss auf die Modellierung der Blackbox-Perspektive da dort ebenfalls vergessene Funktionen ergänzt werden müssen. Dieser dualistische Ansatz, der sowohl die Blackbox- als auch die Whitebox-Perspektiven umfasst, trägt dazu bei, sowohl die vollständigen Funktionen als auch die

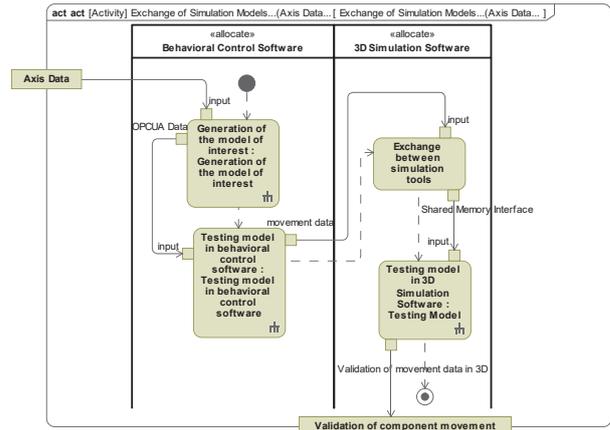


Abbildung 4: Funktionssicht - Whitebox-Perspektive [17]

komplexen Details des Verhaltens des Systems zu definieren und somit zur umfassenden Entwicklung der VIBN-Architektur beizutragen.

### D. Logische Sicht

Die logische Sicht beginnt mit der Definition der logischen Subsysteme in einem Blockdiagramm. Ausgehend von dem zu entwickelnden SoI werden die Teilkomponenten definiert, die zur Erfüllung der Systemfunktionen benötigt werden. In einem zweiten Schritt werden mit Hilfe von Interfaceblöcken, die Interaktionen zwischen den Komponenten definiert vgl. Abbildung 5. Diese werden unter anderem auch aus den Schnittstellen zwischen dem System und dem System Context sowie den funktionalen Schnittstellen abgeleitet.

Nach der Modellierung der funktionalen und der logischen Sicht werden diese miteinander verknüpft, indem jede Funktion einer logischen Komponente zugeordnet wird. Im Beispiel des *Exchange of Simulation Models Across Platforms* und seiner vier Funktionen, dargestellt in Abbildung 4, sind für jede Funktion spezifische Subsysteme vorgesehen, z.B. wird für die Generierung und den Test des Modells eine Verhaltenssteuerungssoftware benötigt und für den Test in einer 3D-Umgebung zusätzlich eine 3D-Simulationssoftware. Damit Daten zwischen diesen zwei Software-Werkzeugen ausgetauscht werden können, bedarf es weiterhin einer logischen Schnittstelle zwischen den entsprechenden logischen Komponenten. Auf die gleiche Weise wird auch jede andere Funktion einer logischen Komponente zugewiesen. Die Definition der benötigten Teilsysteme sowie ihrer Schnittstellen ist dabei stark von der Expertise und Erfahrung des Entwicklers abhängig. Das Übersehen von Komponenten oder Schnittstellen in der logischen Sicht führt zu einem iterativen Korrekturprozess, der bis in die Anforderungssicht zurückreichen kann.

### E. Technische Sicht

Schließlich wurde in der technischen Sicht die eigentliche VIBN-Architektur definiert. Basierend auf den logischen

Komponenten werden nun geeignete Software- und Hardwaremodule identifiziert. Dabei müssen auch die notwendigen logischen Schnittstellen berücksichtigt werden. Während in der logischen Sicht nur eine abstrakte Kommunikationsschnittstelle definiert wurde, werden diese Schnittstellen nun vollständig spezifiziert. Abbildung 5 veranschaulicht exemplarisch das komplexe Zusammenspiel von Hardware- und Softwarekomponenten sowie der dafür benötigten Schnittstellen zum Datenaustausch. Dabei werden auch die Schnittstellen zu den Stakeholdern berücksichtigt bspw. für das Programmieren der Steuerung.

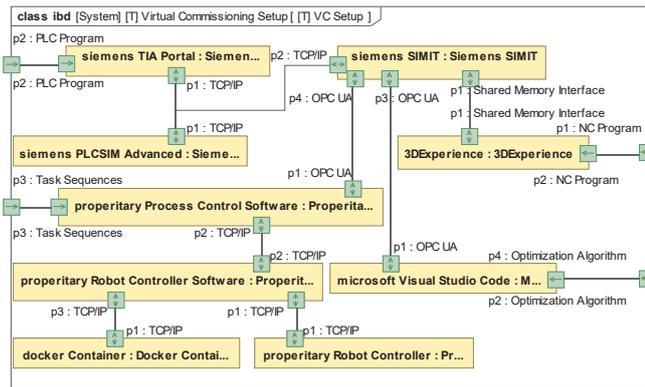


Abbildung 5: Technische Sicht -Komponenten und Kommunikationsschnittstellen [17]

Die 3DEXPERIENCE-Software wurde als geeignete Software für die 3D-Simulationssoftware und das TIA-Portal für die SPS-Programmierung identifiziert. Die Auswahl der Softwarekomponenten ist ein komplexer Prozess, der von Expertenwissen geleitet wird, um sicherzustellen, dass die ausgewählte Software die spezifischen Anforderungen erfüllt. Die Kommunikationsprotokolle OPC-UA, TCP/IP sowie die Shared-Memory-Schnittstelle wurden als geeignete Mechanismen identifiziert, um den Informationsaustausch zu gewährleisten. Die in der Anforderungssicht gesammelten NFAs, wie *NFA3 Implementation of Communication Standards*, können nun validiert werden. Nachdem alle Sichten durchlaufen wurden, kann eine nicht ausreichend spezifizierte technische Komponente ausgewählt werden und für die Komponente der gesamte Prozess erneut iterativ durchlaufen werden.

#### IV. DISKUSSION

Die resultierende Architektur, die durch den hier vorstellten Ansatz entwickelt wurde, erfüllt die identifizierten Anforderungen. Dennoch hat die an SPES angelehnte Methode Einschränkungen: Es fehlt an einer Unterstützung bei Entscheidungen zur Softwareauswahl, zur Auswahl technischer Konfigurationen, zur Auswahl von Kommunikationsprotokollen und zur Entscheidung über die Verwendung vorhandener Simulationsmodelle. Es fehlt auch an Klarheit bei der Unterscheidung zwischen verschiedenen Steuerungsebenen (z.B. Station Controller, Ressourcencontroller, Prozesssteuerung). Zudem werden bestehende Softwarekompetenzen innerhalb

eines bestimmten Unternehmens während des Entwicklungsprozesses nicht berücksichtigt. Eine große Herausforderung besteht im potenziell iterativen Prozess für Updates, insbesondere bei der Identifizierung fehlender Funktionen während der Blackbox- und Whitebox-Implementierung in der funktionalen Sicht, fehlender Softwareschnittstellen in der logischen Sicht oder der Auswahl einer bestimmten Komponente (Software oder Hardware) in der technischen Sicht.

Zusammenfassend dient die SPES-Methodik als guter Leitfaden für die Entwicklung einer robusten VIBN-Architektur, jedoch muss die Methodik an einigen Stellen angepasst werden. Es ist erforderlich, dass Entwickler aufwandsarm kontinuierlich verschiedene Aspekte der Architektur basierend auf neuen Erkenntnissen, sich ändernden Anforderungen oder technologischen Fortschritten überprüfen und verbessern können sollten. Diese Verbesserung könnte in bestimmten Aspekten bei der Entscheidungsunterstützung helfen und die Anwendbarkeit für industrielle Szenarien verbessern.

#### V. FAZIT UND AUSBLICK

Dieser Beitrag stellt die systematische Entwicklung einer VIBN-Architektur für ein automatisiertes Produktionssystem in einem Industrieunternehmen im Bereich der Flugzeugfertigung vor.

Die Anwendung der SPES-basierten Methodik wurde als systematische Lösung für die Entwicklung einer VIBN-Architektur verwendet. Basierend auf einer Untersuchung grundlegender Aspekte der VIBN und SPES sowie einer detaillierten Problemanalyse wurde ein angepasster Arbeitsablauf für die Entwicklung von VIBN-Architekturen skizziert, der die Anforderungs-, Funktions-, Logik- und technische Sicht umfasst. Die Anforderungssicht wurde detailliert beschrieben und wichtige funktionale und nicht-funktionale Anforderungen identifiziert, die für die Entwicklung der VIBN-Architektur notwendig sind. Die Modellierung verschiedener, aufeinander aufbauender Sichten ermöglicht ein strukturiertes Vorgehen, welches in der Lage ist, bei der Definition einer VIBN-Architektur zu unterstützen. Zusammenfassend bietet der systematische Entwicklungsprozess dieses Beitrags wertvolle Einblicke für Forscher und Praktiker auf dem Gebiet der VIBN für automatisierte Produktionssysteme, indem er einen Rahmen bietet, der potenziell als Blaupause für ähnliche Anwendungen in verschiedenen industriellen Bereichen dienen kann. Obwohl die entwickelte VIBN-Architektur die identifizierten Anforderungen erfüllen kann, besteht Potenzial zur Verbesserung und Anpassung bei der Anwendung der Methodik. Eine Schwachstelle, die identifiziert wurde, ist die fehlende Entscheidungsunterstützung bei der Auswahl von Hard- und Softwarekomponenten sowie deren Schnittstellen. Insbesondere bei der dynamischen Anpassung von Funktionen und technischen Komponenten in industriellen Szenarien besteht ein hoher Modellierungsaufwand. Zukünftige Forschungsarbeiten könnten sich darauf konzentrieren, die Entscheidungsunterstützung der Methodik zu verbessern und branchenspezifische Herausforderungen zu adressieren, um die Entwicklung von VIBN-Architekturen weiter zu optimieren.

## DANKSAGUNG

Diese Forschungsarbeit aus dem Projekt iMOD wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] A. Jain, D. A. Vera, and R. Harrison, "Virtual Commissioning of Modular Automation Systems," *IFAC Proceedings Volumes*, vol. 43, no. 4, pp. 72–77, 2010.
- [2] F. Gehlhoff, H. Nabizada, M. Weigand, L. Beers, O. Ismail, A. Wenzel, A. Fay, P. Nyhuis, W. Lagutin, and M. Röhrig, "Challenges in Automated Commercial Aircraft Production," *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 354–359, 2022.
- [3] O. Ismail and A. Fay, "Optimizing virtual commissioning of a robotic system using process mining and footprints conformance checking," in *2023 IEEE 21st INDIN*, 2023, pp. 1–6.
- [4] B. Vogel-Heuser, G. Kegel, K. Bender, and K. Wucherer, "Global information architecture for industrial automation," *Automatisierungstechnische Praxis (atp)*, vol. 51, no. 1, pp. 108–115, 2009.
- [5] N. H. Shah, P. Le Henaff, C. Schiffer, M. Krammer, and M. Benedikt, "Accurate robot simulation for industrial manufacturing processes using fmi and dcp standards," in *Proceedings of 14th Modelica Conferences*, 2021.
- [6] K. Pohl, M. Broy, H. Daembkes, and H. Hönninger, *Advanced Model-Based Engineering of Embedded Systems*. Springer, 2016.
- [7] H. Rauhen, C. Mosch, and E. Axmann, "VDMA Virtual Commissioning Guide: Recommendations for action for economic entry," 2020.
- [8] S. Süß, A. Strahilov, and C. Diedrich, "Behaviour simulation for virtual commissioning using co-simulation," in *2015 IEEE 20th ETFA*, 2015, pp. 1–8.
- [9] S. Süß, S. Magnus, M. Thron, H. Zipper, U. Odefey, V. Fäßler, A. Strahilov, A. Klodowski, T. Bär, and C. Diedrich, "Test methodology for virtual commissioning based on behaviour simulation of production systems," in *2016 IEEE 21st ETFA*, 2016, pp. 1–9.
- [10] A. Abassi, J. H. Lugo-Calles, R. Balderas-Hill, and N. Lassabe, "Virtual Commissioning Implementation of Industrial Painting Process," in *2nd International Conference CompAuto*. IEEE, 2022, pp. 65–68.
- [11] Z. Liu and C. Diedrich, "Component modeling for different application scenarios in the virtual commissioning," in *2020 IEEE 7th ICISCE*, 2020, pp. 1774–1778.
- [12] D. Rovere, M. Silvestri, G. D. Maso, H. Dzafic, and P. Pedrazzoli, "Streamline 3D simulation model development for virtual commissioning with IEC61499," *IFAC-PapersOnLine*, vol. 55, no. 10, pp. 773–778, 2022.
- [13] G. Barbieri, A. Bertuzzi, A. Capriotti, L. Ragazzini, D. Gutierrez, E. Negri, and L. Fumagalli, "A virtual commissioning based methodology to integrate digital twins into manufacturing systems," *Production Engineering*, vol. 15, no. 3-4, pp. 397–412, 2021.
- [14] N. Striffler and T. Voigt, "Concepts and trends of virtual commissioning – A comprehensive review," *Journal of Manufacturing Systems*, vol. 71, pp. 664–680, 2023.
- [15] K. Pohl, H. Hönninger, R. Achatz, and M. Broy, *Model-Based Engineering of Embedded Systems*. Springer, 2012.
- [16] L. Beers, M. Weigand, H. Nabizada, and A. Fay, "MBSE Modeling Workflow for the Development of Automated Aircraft Production Systems," in *2023 IEEE 28th ETFA*, 2023, pp. 1–8.
- [17] O. Ismail, L. Beers, F. Gehlhoff, N. H. Shah, and A. Fay, "Systematic development of a virtual commissioning architecture for an automated production system," in *2024 IFAC 18th INCOM*, 2024, accepted for publication.
- [18] S. Makris, G. Michalos, and G. Chryssolouris, "Virtual commissioning of an assembly cell with cooperating robots," *Advances in Decision Sciences*, vol. 2012, 2012.
- [19] M. Ko, E. Ahn, and S. C. Park, "A concurrent design methodology of a production system for virtual commissioning," *Concurrent Engineering*, vol. 21, no. 2, pp. 129–140, 2013.
- [20] H. Vermaak and J. Niemann, "Virtual commissioning: A tool to ensure effective system integration," in *2017 IEEE ECMSM*, 2017, pp. 1–6.
- [21] W. Hofmann, J. H. Ulrich, S. Lang, T. Reggelin, and J. Tolujew, "Simulation and Virtual Commissioning of Modules for a Plug-and-Play Conveying System," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 649–654, 2018.
- [22] B. Illmer, M. Karkowski, and M. Vielhaber, "Petri net controlled virtual commissioning – A virtual design-loop approach," *Procedia CIRP*, vol. 91, pp. 152–157, 2020.
- [23] A. Bendjelloul, B. Mihoubi, M. Gaham, M. Moufid, and B. Bouzouia, "A framework for an effective virtual commissioning of agent-based cyber-physical production systems integrated into manufacturing facilities," *Concurrent Engineering*, vol. 30, no. 4, pp. 399–410, 2022.
- [24] R. Raffaelli, P. Bilancia, F. Neri, M. Peruzzini, and M. Pellicciari, "Engineering Method and Tool for the Complete Virtual Commissioning of Robotic Cells," *Applied Sciences*, vol. 12, no. 6, p. 3164, 2022.

# Production-Related Testing and Optimisation Environment for Hardware and Software in Aircraft Production of Today and Tomorrow

Sara Abshagen, Alain Chahine, Nihar H. Shah, André Ullmann,  
Martin Röhrig  
*Airbus Operations GmbH*  
Hamburg, Deutschland

Hamied Nabizada, Lasse Beers, Omar Ismail, Maximilian  
Weigand, Felix Gehlhoff  
*Institute for Automation Technology*  
Helmut-Schmidt-Universität/Universität der Bundeswehr  
Hamburg, Deutschland

Alexander Wenzel, Peter Nyhuis  
*Institute of Production Systems and Logistics,*  
*Leibniz-University Hannover*  
Hannover, Deutschland

Alexander Fay  
*Chair of Automation*  
*Ruhr-Universität Bochum*  
Bochum, Deutschland  
alexander.fay@rub.de

**Abstract** – In response to the predicted growth trends in the aerospace industry, a gradual implementation of future industrial production systems with a strong focus on robotics and automation is inevitable. These technologies enhance production performance by automating repetitive tasks, allowing workers to concentrate on more complex activities. This paper introduces a new automation concept and validation system for aircraft structural assembly, set up at the Center for Applied Aviation Research in Hamburg. This system enables the validation of the research topics in the dtec.bw-funded project iMOD and includes recent key innovations developed by Airbus like the Multi-Function End Effector, which can perform multiple processes.

**Keywords** – *Aircraft production, structural assembly production optimisation, robotics, automation*

## I. INTRODUCTION

The aircraft industry deals with the challenge of achieving high accuracy in the assembly of extremely large products. Efforts to address this challenge using robotics designed for the automotive industry have proven unsuccessful due to fundamental differences in the requirements and nature of the components involved. Specifically, the components in the Major Component Assembly (MCA) and Final Assembly Line (FAL) are unique due to their size and exhibit production-related geometric deviations. Despite these challenges, the assembly process requires a high level of precision and quality [1].

The importance of precision in the aircraft industry cannot be overstated. Even minor errors can lead to significant safety risks and increased production costs, necessitating the development of specialized automation systems tailored to the unique challenges of aircraft assembly.

One of the primary issues for the specialized automation system are the geometry deviations that occur during the production of large aircraft components. These deviations can significantly impact the assembly process, making it difficult for

standard robotic systems to achieve the required precision. Therefore, there is a need for adaptive, modular, and flexible automation solutions that can accommodate these deviations and ensure accurate assembly.

In the dtec.bw-funded project “Intelligent modular robotics and integrated production design in aircraft construction” (iMOD) research is being conducted on how to optimise the automated structural assembly in aircraft manufacturing through the use of advanced digital engineering methodologies, including Model-based System Engineering (MBSE), AI Planning, and simulation models.

A robotic system developed by Airbus is used as a validation platform for the research results of iMOD. This platform is capable of performing both longitudinal and orbital joining processes of the aircraft fuselage within a single station. Currently, this validation platform, including the new robotic system and a fuselage, is being set up as a research demonstrator at the Center for Applied Aviation Research (ZAL).

The remainder of the paper is organized as follows: Section II provides a state-of-the-art overview of production systems for structural assembly at Airbus. Section III presents detailed insights into the iMOD validation platform. Section IV outlines the research topics that are being investigated as part of the project and evaluated on the validation platform. Section V summarizes the results and outlines future research directions.

## II. STATE OF THE ART AT AIRBUS

In the area of part and sub-assembly production, automation is already state of the art. For example, shells for the aircraft fuselage with stiffening elements are already assembled fully automated. These automatic drilling and riveting machines are very large and usually the workpiece, in this case the shell, is moved to the tool, which is the drilling and riveting end effector.

For the future of production, automation solely at the component level is not sufficient. A consistent portfolio of automation along the whole value chain is required. This contribution deals with the automation for connecting the longitudinal and orbital joints in the MCA and/or the FAL. The target is to perform most operations in one station with one single system, where today several automation solutions and manual work are required. The main advantage of this station is that it combines different automated drilling and riveting processes to minimize human-robot interaction and achieve efficient system utilization.

One of the systems used at Airbus is the Flextrack (see Figure 1), which is primarily employed for longitudinal or orbital joints on A320 family fuselage pre-assembly lines. Flextrack robotic systems consist of flexible rails that can be easily disassembled and reassembled around the aircraft, allowing lightweight and modular drilling robots to move alongside the fuselage, without the need for separate holding devices [2].

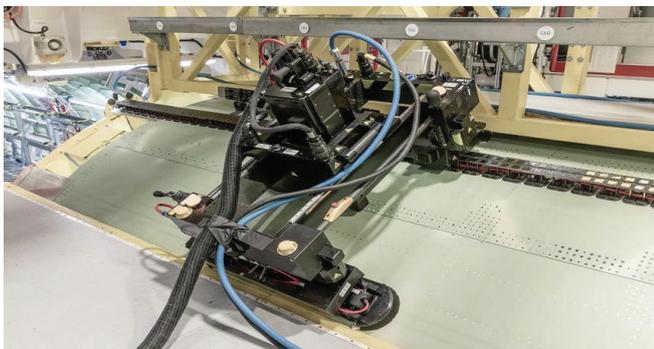


FIGURE 1: FLEXTRACK SYSTEM ON LONGITUDINAL JOINT

Other systems mainly used for orbital joints are 6-axis-robots based on floor-mounted linear axis (see Figure 2). Due to the combination of reach and load capacity, the working area is limited. Furthermore, the size of the components requires several systems to carry out the work in a coordinated manner.



FIGURE 2: NEW AIRBUS STRUCTURE LINE A320 FOR ORBITAL JOINT [2]

End effectors that can reference, drill, and rivet are required to carry out the work efficiently. Given the weight of these multi-functional end effectors (MFEE), a strong support structure and handling technology is needed to enable the required accuracies and forces. Due to the weight of the end effector, the majority of services or repairs are performed directly on the robot in the event of malfunctions. This leads to reduced equipment effectiveness due to planned stops and delays in lead time in case of breakdowns.

Against the background of the production ramp-up, it is imperative to establish an efficient automation system. For example, the end effector should be easily replaceable with a spare part, allowing maintenance, repair, and overhaul (MRO) to be carried out outside the station, thus minimizing any impact on ongoing production.

### III. IMOD VALIDATION PLATFORM

The dtec.bw-funded research project iMOD addresses the need to increase efficiency against the backdrop of an increasing number of variants. The focus here is on the use of advanced Industry 4.0 technologies and the digitalization and automation of manufacturing processes and production workflows.

The project has the following objectives:

- to digitally map resources, products, and processes via the MBSE approach which can be used in future early engineering phases,
- to support the optimisation of assembly and inspection activities and
- to develop corresponding algorithms for collaborative order execution and validation.

These objectives are to be achieved through a collaboration between the Helmut Schmidt University/University of the Federal Armed Forces (HSU/UniBw H) and Airbus. To facilitate this, a validation platform is currently being set up in Hall A of the ZAL. This platform will serve as a research demonstrator to validate the research results for achieving the project objectives. A virtual model of the validation platform is presented in Figure 3.

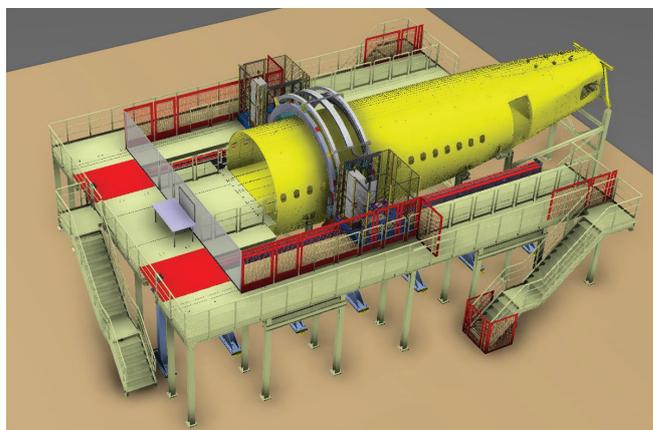


FIGURE 3: 3D MODEL OF THE IMOD VALIDATION PLATFORM

The iMOD validation platform includes a jig to provide access and carry the fuselage section 17-19 of an A320 and the robotic system. The complete robotic system, incl. actuators, MFEEs (co-developed with Setitec), and controls have been developed by Airbus and will be provided to the iMOD project.

The robotic system consists of an orbital gantry that moves on an auxiliary axis. The two MFEEs are installed on the gantry to enable process parallelization. Auxiliary devices required for the process, along with the control system, move with the gantry.

The scope of the automated tasks consists of performing drilling and fastening of the joints and will be split into sequences based on the provided NC-part program. The flow of each sequence, regardless of whether it involves a longitudinal or orbital joint, is fundamentally based on the repetition of the following tasks:

1. Load process parameter for sequence
2. Local referencing
3. Move to drilling point, read process parameters for current hole, and rotate carousel for correct cutter
4. Clamping on fuselage
5. Measure normality
6. Drilling and countersinking, including monitoring the process and collecting process data
7. Insert the fastener with applied sealant while receiving the fastener for the next hole
8. Rotate carousel and control squeeze sealant ring with camera of MFEE
9. Unclamping
10. Repeat beginning from Point 3 until sequence is finished

This process flow will be repeated until all sequences are done and the joint is finished. Each sequence will be performed by one MFEE.

#### A. Virtual Commissioning

Over the last decade, advancements in digitalization technology have enabled comprehensive testing, also known as validation and verification (V&V), of individual devices to entire plants within a virtual environment before actual implementation. Virtual Commissioning (VC) is one such technological advancement, designed to validate production system control programs through virtual simulation models. It is particularly useful for simulating complete automated manufacturing processes in aircraft production [1], [4]. Manufacturing systems in the aerospace domain are highly complex due to component size, tight tolerance requirements, process variation, and multi-tier manufacturing processes. Accurate simulations of robots and other programmable assets are needed to lower the risk of collisions and manufacturing downtimes [5]. Moreover, the VC environment allows re-evaluation of the control logic in case of design change, software updates, or new feature deployment without disrupting the actual production schedule. By testing and validating production systems virtually before physical construction, VC aims to reduce commissioning time and costs while improving system accuracy and reliability [6].

As mentioned earlier in the section, the iMOD robotic system carries two MFEEs on a gantry system. Robots are controlled by a custom controller chip with a specialized real-time Linux-based operating system. This controller is selected for its small, flexible, embedded computer for industrial applications that require high performance and reliability. It combines an ARM processor, an FPGA, and a high-density connector to interface with application-specific I/O. The gantry systems are operated by software PLCs and drive control systems with safety encoders. The overall manufacturing process and coordination are enabled through process orchestration and the

control software. Since there are multiple robot controller systems and the co-simulation between control software and robotic simulations, a modular architecture using modelling and communication standards is proposed [7]. The VC system for this automation system is created as part of the project to support product development, commissioning, and operational scenarios. The VC testbed is the core of the validation platform allowing the utilization of Test-Driven Development (TDD) methodology for the project. It allows the implementation of a V-cycle for development and better testing of the concept. Furthermore, criteria such as modularity, ease of use, accuracy of simulation, as well as the cost of the simulation system are considered for the selection of the architecture.

Due to the lack of availability of accurate simulation models, a Hardware-in-the-loop (HiL) approach is used for the robot controller and the safety-relevant components due to their criticality (see Fig. 4). On the other hand, software-in-the-loop solutions are utilized for PLC control logic, motor simulation, as well as process orchestration software. The iMOD VC solution co-simulates these models to simulate accurate behaviour of the automation and robotic systems. Several standard interface and modelling technologies such as FMI/FMU, OPC UA, and Docker are utilized for successful co-simulation.



FIGURE 4: iMOD - VIRTUAL COMMISSIONING WORKBENCH

Standalone simulation models and software are run on the same simulation PC with standard containerization technology to reduce reliance on HiL. For both PLC and robotic system simulation, it is crucial to work closely with the solution supplier and the controller supplier to make overall solutions and individual components compatible with the simulation workbench. Due to the robotic systems, moving components, and the danger of collision, the simulation workbench is connected with a visualization platform to visualize robot motions and check for possible collisions with the environment. Here as well, a standard communication protocol such as OPC UA is used for simplicity of the solution.

A derivative of TDD is used for the validation of the system. The VC solution will first be used to validate elementary models such as motors, robot controllers, gantry drives, etc. Afterwards, integrated testing of one robotic system will be performed. Once the tests are successful and suggested changes are implemented, both robotic systems will be co-simulated for behavioural tests with NC programs before the system is made operational.

#### IV. RELATED RESEARCH WORK

The iMOD validation platform supports extensive academic research, including five Ph.D. theses and various bachelor and master theses, advancing future automation and digitalization in aircraft production. This section introduces the current research efforts within the iMOD project.

##### A. MBSE Modeling

In the initial phases of engineering, the application of MBSE methods proves to be highly advantageous. These methods facilitate the analysis of various production system configurations to evaluate their effectiveness in accommodating the significant variability stemming from individual customer requirements.

Typically, system engineers encounter difficulties in responding to changes within the system. Moreover, they often face challenges in identifying and assessing implementation alternatives, a process that can be time-intensive and requires considerable manual effort [8]. Consequently, the modelling and management of production resources have gained increasing importance in recent years [9]. This is particularly relevant in the aircraft manufacturing sector, where the future use of reconfigurable resources is anticipated to rise [10].

To address these challenges, a modelling workflow was developed to enable consistent modelling of production systems during the early engineering stages of aircraft manufacturing. This workflow aims to offer a comprehensive solution to the difficulties associated with modelling complex production systems, thereby supporting effective decision-making throughout the design process. This solution includes detailed modelling of individual resource components, their interfaces, and capabilities to identify suitable components [11].

To support and guide the modelling process, Domain Specific Modelling Languages (DSML) have been designed based on established standards. These languages address the existing need for standardization and the lack of terminology and reusability in the practical application of MBSE methods [12]. Two notable examples are the SysML profile, which facilitates process modelling in accordance with the VDI3682 standard [13], and the standardized description of robotic resources [14].

The iMOD validation platform model aids not only in system development but also in simulation purposes. Based on these simulation results, it becomes possible to determine the optimal tool configuration for the MFEE to minimize the need for reconfiguration. The tools required for executing the processes are derived from the product model, particularly from rivet information such as rivet type, diameter, and the number and position of the rivets [15].

##### B. AI-based Process Planning

In early engineering phases, it is crucial to compare different feasible system configurations to identify the most effective solutions [16]. For instance, while adding an additional MFEE to the gantry of the iMOD validation platform may reduce the lead time, it could also lead to increased idle times for other MFEEs. To make informed decisions and enable such comparisons, a preliminary process plan is needed to consider specific values for cycle times and resource utilization. This plan must define a sequence of processes necessary to manufacture the given product and allocate an appropriate resource to each individual process.

The description and algorithmic solution of such problems underlies the research area of AI Planning [17]. Within this area, the Planning Domain Definition Language (PDDL) has become the de facto standard for describing planning problems [18]. PDDL is a formal language used to represent the tasks, actions, and resources involved in a planning problem. However, modelling PDDL descriptions is often particularly challenging, time-consuming, error-prone [19], and can lead to data inconsistency issues [20].

To address these challenges, the iMOD research project focuses on developing methods to automate the generation of PDDL descriptions by leveraging existing engineering artifacts such as system models and product models. A workflow was developed in [21] that specifies the necessary steps to process existing knowledge and transform it into PDDL descriptions. System and product models are systematically enriched through a SysML profile for PDDL [22], enabling algorithms to accurately retrieve the required information. Consequently, PDDL descriptions for various system configurations can be created during the engineering phase of the robotic system. These descriptions are used to generate process plans, allowing for performance assessments of the system.

##### C. Continuous Virtual Commissioning in Operation

VC is traditionally used in the early development phases up to the commissioning phase of automated systems, enabling thorough analysis and testing of control programs and resource behaviour using 3D models and control emulations [6]. In the iMOD validation platform, VC will be used to virtually simulate and validate control program updates for the robotic system and its lightweight MFEE. This process ensures that the system's drilling and fastening operations on both orbital and longitudinal joints are thoroughly tested in a virtual environment. By doing so, potential issues can be identified and resolved before implementation, minimizing disruption and ensuring seamless integration during the manufacturing process of aircraft fuselage sections.

Since VC is traditionally used only during the development and commissioning phases, disruptions that occur during the operational phase remain a significant source of downtime for automation systems, such as in the iMOD validation platform. Therefore, VC is employed to optimise the testing and validation procedures of the changes in the control program in the operation of automated manufacturing processes. These changes in the control program are an attempt to mitigate disruptions like delays or collisions, which occur in the automation system during operation. Hence, the aim is to develop a method that focuses on testing and validating only the changes in the updated control programs, thus enabling the continuous operation of the automation system with minimal interruption [23].

This approach extends the traditional use of VC by integrating it into both the preparation and operational phases through Continuous Virtual Commissioning in Operation (CVCO) [23]. This method builds upon the concept of Integrated Virtual Preparation and Commissioning (IVPC) [24], a holistic approach that utilizes the same virtual models of the automation system throughout the preparation, control system implementation, and commissioning phases. By incorporating VC into the operational phase, CVCO ensures that any changes in the control programs are identified and validated in real-time, allowing for immediate adjustments and minimizing downtime.

#### D. Logistics-oriented analysis of highly automated production systems

The first study within iMOD's logistics division looked at hybridizing different production structures to increase throughput time potential while increasing variant complexity in aircraft manufacturing. In previous articles, the main potentials, challenges, and basic approaches to realizing hybrid production principles were presented. Based on this, a concept was developed for the ideal combination of automated and manual processes in a hybrid production structure. [25, 26]

Based on the structural analysis, research was conducted to derive the best possible supply strategy for highly automated production systems. Inventory management in the production planning and control (PPC) process was identified as a key lever for a reliable order fulfilment process. The goal was to extend a logic for inventory dimensioning in order to supply the relevant production areas with the right quantity at the right time. Low standardization and lack of supplier diversity posed a particular challenge. Therefore, an approach for predictive inventory management based on product portfolio characteristics was developed. In addition, a new mathematical model for dimensioning safety stocks under the increasing influence of volatility is developed [27].

In order to analyse the reliability of an automated system and the associated impact on downstream processes in the early engineering phase, it is not enough to focus solely on material availability and warehouse management. The integration of resilience analyses into risk management for the systematic identification of vulnerabilities in systems and their linkage with upstream or downstream systems was identified as a key factor for logistics-oriented risk assessment. This is necessary to provide support to engineers in the early development phases with regard to the cost-benefit dilemma and to assist in the identification and evaluation of risk strategies through static or dynamic resilience. A framework is presented that shows how risk assessment and resilience analysis can be combined to support decision makers in the above-mentioned dilemma. Furthermore, a concept for utilizing different information from different MBSE models has been developed [28].

#### V. SUMMARY & OUTLOOK

This paper presents optimisation approaches for automation concepts in aircraft production. To thoroughly test these concepts, a validation platform is being set up at ZAL in Hamburg, in which the various research approaches are being validated as part of the dtec.bw-funded iMOD research project. This will lay the foundation for the automated aircraft production of the future.

Furthermore, this validation platform will pave the way for future research activities, including cooperative task execution of multiple robotic systems, algorithm development for production-specific applications, data-driven process optimisation, automated inline monitoring of assembly processes, and the integration of AI technologies under consideration of compliance and certification guidelines like EASA or ISO.

#### ACKNOWLEDGEMENT

This research project iMOD is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] F. Gehlhoff et al., "Challenges in Automated Commercial Aircraft Production," *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 354–359, 2022.
- [2] Airbus SAS, "Developing the automated aerospace production systems of the future," 2024, [Online]. Available: <https://www.airbus.com/en/innovation/industry-4-0/robotics>. [Accessed: 27-Jun-2024].
- [3] M. Richardson, "Airbus inaugurates Hamburg A320 structure assembly line," *Aerospace Manufacturing*, 2019, [Online]. Available: <https://www.aero-mag.com/airbus-inaugurates-hamburg-a320-structure-assembly-line/category>. [Accessed: 27-Jun-2024].
- [4] O. Ismail and A. Fay, "Optimizing Virtual Commissioning of a Robotic System using Process Mining and Footprints Conformance Checking," 2023 IEEE 21st International Conference on Industrial Informatics (INDIN), 2023.
- [5] N. Shah et al., "Accurate Robot Simulation for Industrial Manufacturing Processes using FMI and DCP Standards," 14th International Modelica Conference, 2021.
- [6] H. Rauen et al., "Virtual Commissioning Guide: Recommendations for action for economic entry," Verband Deutscher Maschinen- und Anlagenbau e.V. (VDMA), 2020.
- [7] O. Ismail et al., "Systematic Development of a Virtual Commissioning Architecture for an Automated Production System," 18th IFAC Symposium on Information Control Problems in Manufacturing (INCOM), 2024, accepted for publication.
- [8] P. Z. Chadzynski et al., "Enhancing Automated Trade Studies using MBSE, SysML and PLM," in 28th Annual INCOSE International Symposium, 2018.
- [9] M. Sanfilippo et al., "Modeling Manufacturing Resources: An Ontological Approach," *Product Lifecycle Management to Support Industry 4.0* (PLM 2018), 2018.
- [10] R. Arista et al., "Industrial Resources in the design of Reconfigurable Manufacturing Systems for aerospace: A systematic literature review," *Computers in Industry*, vol. 142, p. 103719, 2022.
- [11] L. Beers et al., "MBSE Modeling Workflow for the Development of Automated Aircraft Production Systems," 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), 2023.
- [12] T. A. McDermott et al., "Benchmarking the Benefits and Current Maturity of Model-Based Systems Engineering across the Enterprise," Technical Report, Systems Engineering Research Center (SERC), 2020.
- [13] L. Beers et al., "A SysML Profile for the Standardized Description of Processes during System Development," 2024 IEEE International Systems Conference (SysCon), 2024.
- [14] L. Beers et al., "Entwicklung eines generischen Modells für die standardisierte Beschreibung von Ressourcen in der Luftfahrtproduktion," *dtec.bw-Beiträge der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw - Band 1*, 2022.
- [15] L. Beers et al., "Towards an MBSE Approach for Modeling Complex Production Systems based on Industrial Standards," 29th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2024, accepted for publication.
- [16] M. Törmänen et al., "Integrating Multi-Disciplinary Optimization into the Product Development Process using Model-Based Systems Engineering (MBSE)," NAFEMS World Congress, 2017.
- [17] M. Ghallab et al., *Automated Planning and Acting*, Cambridge University Press, 2016.
- [18] C. Mayr-Dorn et al., "Evaluating PDDL for programming production cells: a case study," 2022 IEEE/ACM 4th International Workshop on Robotics Software Engineering (RoSE), 2022.
- [19] A. Lindsay, "On Using Action Inheritance and Modularity in PDDL Domain Modelling," *Proceedings of the International Conference on Automated Planning and Scheduling*, 2023.
- [20] S. Bhatnagar et al., "On the Challenges of on-the-fly Knowledge Acquisition for Automated Planning Applications," *Proceedings of the 14th International Conference on Agents and Artificial Intelligence*, 2022.

- [21] H. Nabizada et al., “Model-based Workflow for the Automated Generation of PDDL Descriptions,” 29th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2024, in review.
- [22] H. Nabizada et al., “SysML-Profil zur automatisierten Generierung von PDDL-Beschreibungen aus Systemmodellen,” in Tagung Entwurf komplexer Automatisierungssysteme (EKA), 2024.
- [23] O. Ismail et al., “Efficient Identification and Validation of Control Program Changes in an Automated Production System using Virtual Commissioning,” 29th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2024, accepted for publication.
- [24] M. Dahl et al., “Automatic Modeling and Simulation of Robot Program Behavior in Integrated Virtual Preparation and Commissioning,” *Procedia Manufacturing*, vol. 11, pp. 284–291, 2017.
- [25] A. Wenzel et al., “Hybrid Production Principles: A Framework for the Integration in Aircraft Manufacturing,” in *Smart, Sustainable Manufacturing in an Ever-Changing World*, 2023.
- [26] A. Wenzel et al., “Neue Produktionsstrukturen für die Flugzeugfertigung der Zukunft,” *dtec.bw-Beiträge der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw - Band 1*, 2022.
- [27] A. Wenzel et al., “Anticipatory Inventory Management For Realizing Robust Production Processes In Engineer-To-Order Manufacturing: A Modeling Approach”, Conference on Production Systems and Logistics: CPSL 2023, 2023.
- [28] A. Wenzel et al., “Framework for Logistics-oriented Risk Assessment by Integrating Resilience Analysis”. *Procedia CIRP*, 2024, accepted for publishing.

# Knowledge Graphs for the Enhancement of Process Planning in Manufacturing

Jonathan Reif, Tom Jeleniewski, Felix Gehlhoff

*Institute for Automation Technology  
Helmut Schmidt University/*

*University of the Federal Armed Forces Hamburg  
Hamburg, Deutschland*

{jonathan.reif, tom.jeleniewski, felix.gehlhoff}@hsu-hh.de

Alexander Fay

*Chair of Automation  
Ruhr University Bochum  
Bochum, Deutschland  
alexander.fay@rub.de*

Constantin Hildebrandt

*CTC GmbH (An Airbus Company)  
Stade, Deutschland  
constantin.hildebrandt@airbus.com*

**Abstract**—The increasing complexity of modern manufacturing systems, driven by demands for sustainability, shorter product life cycles, and greater customization, necessitates frequent reconfiguration and redesign of production processes. In this context, knowledge about parameter interdependencies is crucial. Additionally, simulations play a vital role by enabling virtual testing of various parameter configurations to optimize key performance indicators such as energy consumption, emissions, processing time, and costs.

This project report presents an integrated approach to enhance process planning by incorporating semantic models that describe process parameter interdependencies. Furthermore, we describe a concept for automating the generation of simulation sequences, thereby reducing the complexity and effort involved in manual planning. The combination of these approaches is demonstrated using a web-based application, showcasing the potential to support efficient and sustainable manufacturing practices.

**Index Terms**—Semantic Web, Simulation, Parameter Interdependencies

## I. INTRODUCTION

Due to a rising demand for customized products and a consequent reduction in lot sizes, contemporary production systems are becoming increasingly complex [1]. Additionally, manufacturing companies encounter significant challenges due to the wide range of possible configurations in production systems [2]. Current efforts to enhance sustainability and reduce emissions, energy consumption, and waste require considering various parameters and indicators in process design. However, only a limited number of experts possess in-depth knowledge about a production system or specific subareas, as such expertise is often confined to particular process steps.

A thorough and comprehensible documentation of the production system is essential for disseminating this knowledge to other stakeholders within a manufacturing company [3]. One critical area where this understanding is crucial is in the (re)design of production processes. Knowledge about the fundamental behavior of processes, the impact of parameter variations, and the effects of process changes (e.g., different raw materials) is essential. Complex production processes, particularly those with high unit costs, are often prohibitively expensive to conduct iteratively during the design phase [4]. Simulation provides a method for enhancing process knowledge virtually and allows for testing various parameter configurations

in a cost-effective and timely manner [5]. However, selecting and planning these simulations is not trivial. Therefore, it is necessary to also store knowledge about available simulations in a semantic model to support their use.

An increasingly important approach for storing knowledge in a machine-readable form is the use of *Semantic Web* technologies, particularly ontologies. An ontology is defined as a "formal, explicit specification of a shared conceptualization" [6]. The *SPARQL Protocol and RDF Query Language* (SPARQL) can be utilized to retrieve knowledge stored in an ontology. Semantic Web technologies have proven useful and are employed, for instance, as information models in manufacturing [3]. This project report aims to describe the application potentials of Semantic Web technologies in the context of process (re)design and to demonstrate practical suitability through a web-based demonstrator that implements various concepts for describing known parameter influences and supporting the simulation of such process information.

The background is introduced in Section II, followed by an explanation of the developed concept in Section III. Section IV describes the implementation of the concept in a prototypical demonstrator. Finally, Section V provides the conclusion and outlines future work.

## II. BACKGROUND

As outlined by Gill et al. [7], semantic web technologies present extensive opportunities for application in the manufacturing domain. These technologies enable the seamless integration of diverse models — structural, functional, and behavioral — by providing clear semantics, which ensure that design data can be effectively reused in subsequent lifecycle phases such as manufacturing and maintenance. This possibility enhances overall efficiency. [7]

Semantic technologies facilitate the formalization and integration of expert knowledge and decentralized data, which is crucial for understanding and managing process parameter interdependencies. By using ontologies, these technologies allow for the precise specification of process parameters and their interrelations, providing a comprehensive framework for documenting and retrieving process knowledge.

In the context of simulation support, semantic technologies can support the automated selection and planning of simulation sequences by enabling a comprehensive description of simulation resources, including input and output parameters, as well as other important properties. This supports virtual testing and optimization of manufacturing processes without the need for physical prototypes, thereby reducing costs and development time while maintaining high-quality production outcomes. The integration of simulations with semantic models allows for the automated planning of simulation sequences, ensuring that the most appropriate simulations are selected and executed based on defined criteria and interdependencies.

In summary, semantic technologies offer a robust framework for integrating and managing digital artifacts in manufacturing. They address data interoperability challenges and enable more efficient and effective lifecycle management, particularly in the areas of process (re)design, parameter interdependency analysis, and simulation support.

### III. CONCEPT

Based on the identification of individual application potentials as described by Gill et al. [7], a comprehensive concept to leverage these potentials was developed. The primary objective of this concept is to address user queries. These queries relate to various process variables that arise from executing the process with specific parameters.

The concept aims to support users in the context of process (re)design by providing information about process parameter interdependencies or, when necessary, facilitating the virtual generation of this information through simulations. To address user queries, different methods complementing each other, each accompanied by its own developed concept, were identified.

If the process has already been executed with the parameters for which the user seeks process variables, the query can be answered using historical process data. Otherwise, the required knowledge can be derived from functional interdependencies or generated via simulations.

The core of the concept is a knowledge model in the form of an ontology. This approach ensures that the concept remains as generic as possible. By exchanging the application-specific ABox in the knowledge model, the concept can be applied to other use cases or processes while the structure of the knowledge model (TBox) remains. The knowledge model integrates both, machine data and expert knowledge.

In this section, we present the sub-concepts *Process Knowledge Retrieval*, *Historical Process Data*, *Parameter Interdependency Description* and *Simulation Support* before demonstrating their combined use in a web application in Section IV. The sub-concepts and their interaction are illustrated in Figure 1.

#### A. Process Knowledge Retrieval

An ontology includes not only a conceptual vocabulary but also precise definitions for each term, delineating the relation among concepts. The methodological framework detailed by

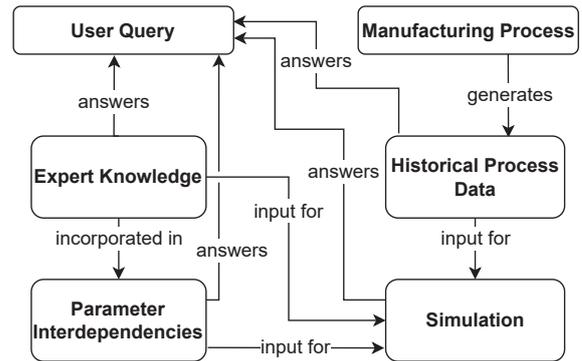


Figure 1: Presented sub-concepts and their interactions

Hildebrandt et al. [3] was used for developing the ontology. This framework is designed for the structured development of ontologies within manufacturing environments and was applied during workshops with project partners.

To enhance generalizability and standardization while reducing modeling effort, *Ontology Design Patterns* (ODPs) are utilized. These patterns provide modeling solutions for recurring problems in ontology design [8]. Unlike an ontology developed from scratch, this approach allows for high reusability of previous models, ensuring that developments in this research area can be reused in other applications. These ODPs are maintained separately and imported into an alignment ontology, facilitating the integration of additional standards and ODPs as needed [3].

Following the requirement analysis, ODPs based on industrial standards and containing the necessary terminologies for the application were identified. For instance, the ODP based on the VDI/VDE 3682 guideline for Formalized Process Description was implemented in the concept to address recurring issues in describing discrete or continuous production processes [9]. To ensure compliance with these guidelines during modeling, a web-based modeling tool, introduced by Nabizada et al. [10] following VDI/VDE 3682, was used. A mapping feature enables automated conversion of process descriptions from the modeling tool's export format into an ontology.

To specify additional information or data, such as process parameters, the DIN EN 61360 standard [11] and its corresponding ODP are used, which allow detailed descriptions of data elements with associated types and instances. For data element type descriptions, unit specification is limited to freely selectable strings. However, it is crucial to define terminologies for standardized units, so an additional ODP containing the UNECE-recommended classification of measurement units is included. The proposed model also uses this ODP to assign standardized units to type definitions. A particular focus is placed on the relations between equipment, processes, and products, as well as on the specification of related process parameters. However, the process description alone cannot

adequately represent the extent to which input parameters influence specific output variables or the impact of changes in input variables. This underscores the necessity to expand the ontology with additional concepts and terminologies.

### B. Historical Process Data

In the context of manufacturing processes, systematic collection and storage of various machine data are performed. The recording and documentation of all events, state changes, and process steps enable the creation of a comprehensive process history. To effectively utilize these data, it is essential to embed them within a contextual framework that allows user-oriented insights and interpretations.

A key element in this process and the overall concept is linking the machine data to an ontology. This linkage enables the interpretation of data concerning their semantic meaning and context, focusing on the virtualization approach.

By making historical data available in the ontology and enriching it with a semantic meaning there is the possibility to utilize this data in the concept to answer user queries.

### C. Parameter Interdependency Description

In the proposed concept, it is essential to clearly articulate the known interdependencies between parameters, especially when assessing the impacts of process changes. The process ontology described in Section III-A has been extended by Jeleniewski et al. [12] to facilitate the representation of process parameters and their interdependencies. A class diagram of the developed ontology is shown in Figure 2.

In this model, concepts from the *Formalized Process Description* (FPD) as outlined in the guideline *VDI/VDE 3682* [13] are incorporated. Furthermore, the ODP from the *DINEN 61360* standard [14], which includes a framework for specifying data elements with types and instance descriptions, is used for modeling process parameters. Restrictions can also be applied to these parameters, as demonstrated by Jeleniewski et al. [15]. To add unit of measurement information to the data elements, type descriptions are additionally classified using a *UNECE* unit of measurement class [16].

Additionally, the model integrates the OpenMath ontology, as presented by Wenzel [17], which enables the representation of mathematical expressions and functions. These mathematical expressions are correlated in the model with the process description elements of *VDI/VDE 3682* and their corresponding data elements to formally describe interdependencies among process parameters.

One advantage of the FPD is the ability to model at various levels of abstraction, allowing for both detailed and generalized process representations as needed. Jeleniewski et al. [15] propose a method to unify the concepts of composition and decomposition from both mathematical and process description perspectives, linking decomposable process operators with descriptive functions. Since interdependency descriptions are directly associated with a process operator, this methodology supports the decomposition and composition of process operators and their corresponding mathematical expressions

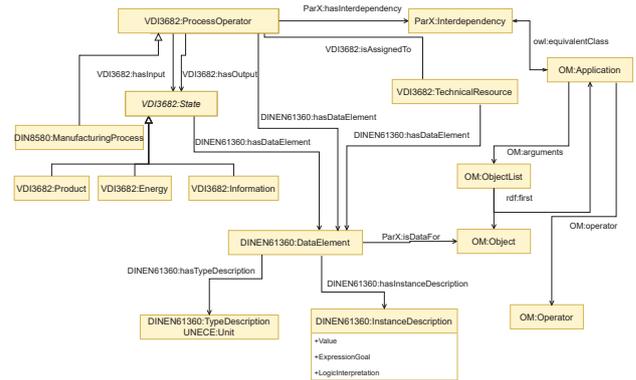


Figure 2: Class diagram of the Parameter Interdependency Ontology (ParX) based on [19]

for describing interdependencies. However, this flexibility can complicate the integration of mathematical expressions for interdependency descriptions.

To address this challenge, Jeleniewski et al. [18] have developed a structured method for integrating interdependencies. Additionally, Jeleniewski et al. [12] provide an extensive modeling and querying example using SPARQL. A method to enhance the SPARQL queries with a data consistency check, ensuring the return of correct process-related data along with the corresponding mathematical expressions, is detailed in Jeleniewski et al. [15].

### D. Simulation Support

As outlined in I, simulations play a crucial role in analyzing manufacturing processes. They can depict various elements of the production process and related systems, facilitating the virtual generation of process information and enabling cost-effective testing of different process parameter configurations within reasonable time frames [5].

However, a single simulation often is not able to provide a complete understanding of specific process settings. Instead, multiple simulations are needed, where the outputs of one serve as inputs for another, forming a sequence of simulations. Moreover, simulations of the same object of consideration, e.g. the same process step, are often available at different complexity levels, necessitating a selection based on time and resource constraints. The general approach is to simulate most of the system at the lowest complexity and only simulate specific parts in more detail when necessary [20].

Due to these factors, manually creating simulation sequences is complex, time-consuming, and prone to errors. It requires in-depth knowledge of simulations, their parameters, and the entire production process. To automate the creation of simulation sequences, a machine-interpretable model is essential. This model must describe simulations, their functionalities, and relevant properties.

This model forms the basis for the method introduced by Reif et al. [21] for planning simulation sequences, enabling the

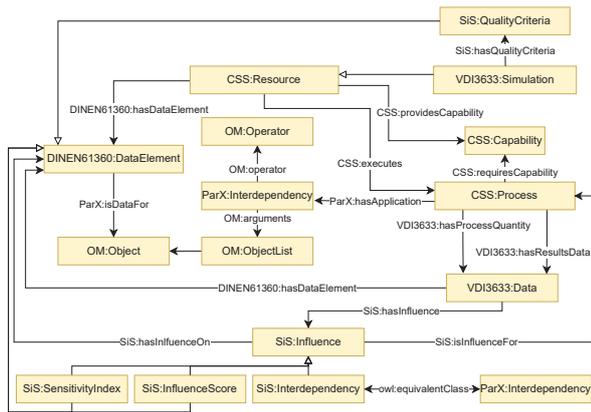


Figure 3: Components of the *Simulation Support* (SiS) ontology [24]

selection and planning of simulations. A class diagram of the model, which was implemented as an ontology is illustrated in Figure 3.

The model integrates the industrial standard VDI 3633 [22] to outline various simulations along with their input and output parameters. Additionally, it employs the *Capability-Skill-Service* (CSS) model [23] to describe the capabilities of these simulations. In the CSS context, capabilities are abstract descriptions of functions required by processes and provided by resources [23]. This capability description within the model enables the selection of simulations based on their ability to generate specific process information. Furthermore, the model also utilizes the *DINEN 61360* standard [11] to model characteristics of parameters.

Moreover, the model incorporates quality criteria to assist in choosing simulations that are more suitable for particular use cases. The description of the influence of parameters is critical in this selection process. If an input parameter significantly influences the output of a simulation, it should receive more consideration during the selection process, or in the context of simulation sequences, the simulation generating this input parameter should be prioritized over others with negligible influence on the output. The model provides three methods to describe these parameter influences, one of which utilizes the *Parameter Interdependency* (ParX) ontology [19] to depict parameter influences in the form of mathematical functions. Other types of influences are expressed as an index resulting from sensitivity analysis or as a score resulting from expert knowledge.

#### IV. PRODUCTION ONTOLOGY DEMONSTRATOR

The combination of the concepts described in Section III was implemented in a demonstrator to showcase their potential and usability in an industrial setting. This demonstrator enables experts from various domains to explore approaches for knowledge description, discovery, retrieval, and generation. The primary goals are to demonstrate the potential of these

technologies, encourage professional discussions, and align research approaches with industrial needs.

During the development of the demonstrator, significant focus was placed on understanding the interactions between various parameters and their effects on both the process and its output variables. Given the high costs and time associated with real production processes, there is a strong interest in obtaining process knowledge, particularly about the influence of different input parameters on outputs, without physical execution. This approach allows for testing different process variants with varying parameters.

Initial solutions to these challenges were implemented in the presented demonstrator. The various functions implemented in the demonstrator are shown on *YouTube*<sup>1</sup>. It can query various aspects of the knowledge model, including process descriptions based on the product-process-resource approach according to the FPD and associated information or process parameters. These parameters are enriched with information about attributes, units of measure, and additional information such as influences and restrictions (upper and lower limits). Simple physical relations were also modeled in graph form and made available for querying. Additionally, the demonstrator supports users through a graphical interface to create, visualize, and delete knowledge.

Furthermore, process information can be generated through prototypical simulations connected to the demonstrator. Users can input various process parameters via a flexible user interface to query whether historical process information exists for the respective parameter combinations and select simulations to generate missing process information.

It is also possible to execute selected simulations with the input parameters, thereby generating new process information. Due to the combination of process models including the description of interdependencies and the simulation models, process information can also support the execution of simulations. For example, production restrictions contained in the process model can simultaneously restrict the input of parameters in the simulation. Other benefits of linking the information models include connecting real process parameters with simulated process parameters. This linkage allows for diverse input sources for simulation input parameters, such as historical data or process interdependencies. Additionally, it facilitates the simulation of missing process information or the derivation of information through functional parameter interdependencies.

Similar to the knowledge model, the development of the demonstrator emphasized generalizability and transferability to other processes through ontology-based approaches.

The demonstrator, described in this section, was developed as a web application using *Angular*<sup>2</sup> for the frontend and *NestJS*<sup>3</sup> for the backend. As shown in Figure 4, the user interacts with the frontend interface, which triggers HTTP requests

<sup>1</sup><https://www.youtube.com/watch?v=Jt39YbPkWVk>

<sup>2</sup><https://angular.dev/>

<sup>3</sup><https://nestjs.com/>

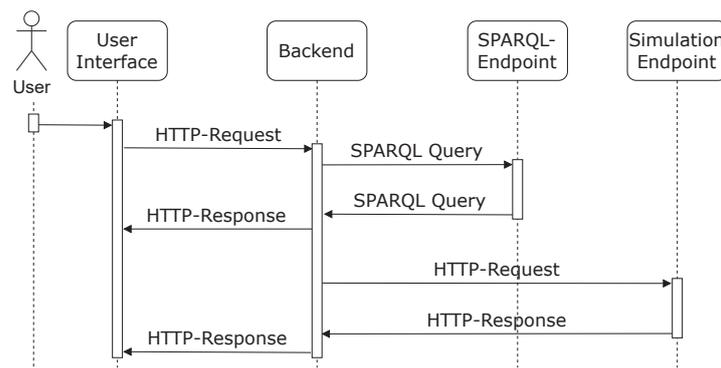


Figure 4: Sequence diagram of the production ontology demonstrator

to the backend. The backend communicates with a SPARQL endpoint, implemented in a graph database via SPARQL queries. Similarly, users can interact with the Simulation-Endpoint by triggering HTTP requests to the backend through the frontend, executing simulations in the functional mockup unit format connected to the demonstrator via a web interface. This allows users to generate required process information.

The demonstrator is designed so that users do not have to interact directly with the ontology via SPARQL, but are supported by an intuitive user interface for every action. This ensures that the created information model is accessible not only to ontology experts but also to other stakeholders in a manufacturing company.

## V. CONCLUSION AND FUTURE WORK

This project report describes an approach to enhancing process planning in manufacturing by integrating semantic models and simulations. By leveraging knowledge about process parameter interdependencies and supporting the selection and planning of simulation sequences, the proposed concept offers ways to improve efficiency and sustainability in manufacturing systems.

The developed ontology-based knowledge model facilitates the representation of complex manufacturing processes and their interdependencies, allowing for better documentation, retrieval, and application of process knowledge. This is particularly valuable given the increasing demands for customization and sustainability in manufacturing. The concept of linking historical process data with semantic models further enriches the knowledge base, enabling more informed decision-making.

The demonstrator highlights the practical potential of these concepts, offering a user-friendly interface for querying and generating process information. This tool supports users in both retrieving existing knowledge and generating new knowledge through simulations.

Future work will focus on expanding the scope and functionality of the demonstrator. This includes incorporating more advanced simulation capabilities, improving the integration of real-time data, and enhancing the user interface to support a wider range of industrial applications. Additionally, efforts

will be made to further standardize the knowledge models, promoting broader adoption and interoperability across different manufacturing systems.

Moreover, research will continue to refine the automated planning of simulation sequences, exploring more sophisticated algorithms and techniques to optimize the selection and execution of simulations. These advancements aim to further reduce the complexity and effort required in manual process planning, ultimately driving greater efficiency and sustainability in manufacturing practices.

In conclusion, the integration of Semantic Web Technologies and simulations represents a promising direction for the future of process planning in manufacturing. By harnessing the power of knowledge graphs and simulation, this approach provides a robust framework for addressing the challenges of modern manufacturing, paving the way for more efficient, sustainable, and customizable production processes.

## ACKNOWLEDGMENT

This project report [project LaiLa] is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

## REFERENCES

- [1] P. Leitão, “Agent-based distributed manufacturing control: A state-of-the-art survey,” *Engineering Applications of Artificial Intelligence*, vol. 22, no. 7, pp. 979–991, 2009.
- [2] M. Garetti, L. Fumagalli, and E. Negri, “Role of Ontologies for CPS Implementation in Manufacturing,” *Management and Production Engineering Review*, vol. 6, no. 4, pp. 26–32, 2015.
- [3] C. Hildebrandt, A. Köcher, C. Küstner, C.-M. Lopez-Enriquez, A. W. Müller, B. Caesar, C. S. Gundlach, and A. Fay, “Ontology Building for Cyber-Physical Systems: Application in the Manufacturing Domain,” *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 3, pp. 1266–1282, 2020.
- [4] S. N. Grigoriev, A. A. Kutin, and M. V. Turkin, “Modelling Complex Production Processes in Aerospace Industry based on Dimensional Analysis,” *Procedia CIRP*, vol. 7, pp. 473–478, 2013.
- [5] D. Mourtzis, “Simulation in the design and operation of manufacturing systems: state of the art and new trends,” *International Journal of Production Research*, vol. 58, no. 7, pp. 1927–1949, 2020.
- [6] R. Studer, V. Benjamins, and D. Fensel, “Knowledge engineering: Principles and methods,” *Data & Knowledge Engineering*, vol. 25, no. 1-2, pp. 161–197, 1998.

- [7] M. S. Gill, J. Reif, T. Jeleniewski, M. Weigand, and A. Fay, "Application potentials of Semantic Technologies for Digital Twins in Aircraft Design, Manufacturing and Maintenance."
- [8] A. Gangemi and V. Presutti, "Ontology Design Patterns," in *Handbook on Ontologies*, S. Staab and R. Studer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 221–243.
- [9] VDI, "Richtlinie VDI/VDE 3682: Formalisierte Prozessbeschreibungen," Mai 2015.
- [10] H. Nabizada, A. Köcher, C. Hildebrandt, and A. Fay, "Offenes, web-basiertes Werkzeug zur Informationsmodellierung mit Formalisierter Prozessbeschreibung," in *Automation 2020*. VDI Verlag, 2020, pp. 443–454.
- [11] DIN EN 61360-1, "Standard data element types with associated classification scheme - Part 1: Definitions - Principles and methods (IEC 61360-1:2017)," 07.2018.
- [12] T. Jeleniewski, H. Nabizada, J. Reif, A. Köcher, and A. Fay, "A Semantic Model to Express Process Parameters and their Interdependencies in Manufacturing," in *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*. IEEE, 2023, pp. 1–6.
- [13] VDI/VDE 3682:1, "Formalised Process Descriptions - Concept and Graphic Representation," 2015.
- [14] DIN EN 61360-1, "Standard data element types with associated classification scheme - Part 1: Definitions - Principles and methods (IEC 61360-1:2017)," 2018.
- [15] T. Jeleniewski, J. Reif, F. Gehlhoff, and A. Fay, "Parameter Interdependencies in Knowledge Graphs for Manufacturing Processes," in *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2024, pp. 1–8.
- [16] UNECE, "Recommendation No. 20: Codes for Units of Measure Used in International Trade," 2010.
- [17] K. Wenzel, "OpenMath-RDF: RDF encodings for OpenMath objects and Content Dictionaries," in *31st OpenMath Workshop*, 2021.
- [18] T. Jeleniewski, J. Reif, and A. Fay, "Integrating Interdependencies in Semantic Manufacturing Process Description Models," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2023, pp. 1–4.
- [19] T. Jeleniewski, H. Nabizada, J. Reif, A. Köcher, and A. Fay, "A Semantic Model to Express Process Parameters and their Interdependencies in Manufacturing," in *2023 IEEE 32nd ISIE*. IEEE, 2023, pp. 1–6.
- [20] P. Puntel-Schmidt and A. Fay, "Levels of Detail and Appropriate Model Types for Virtual Commissioning in Manufacturing Engineering," *IFAC-PapersOnLine*, vol. 48, no. 1, pp. 922–927, 2015.
- [21] J. Reif, T. Jeleniewski, and A. Fay, "An Approach to Automating the Generation of Process Simulation Sequences," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2023, pp. 1–4.
- [22] VDI 3633:1, "Simulation of systems in materials handling, logistics and production," 12.2014.
- [23] A. Köcher, A. Belyaev, J. Hermann, J. Bock, K. Meixner, M. Volkman, M. Winter, P. Zimmermann, S. Grimm, and C. Diedrich, "A reference model for common understanding of capabilities and skills in manufacturing," *at - Automatisierungstechnik*, vol. 71, no. 2, pp. 94–104, 2023.
- [24] J. Reif, T. Jeleniewski, A. Köcher, T. Frerich, F. Gehlhoff, and A. Fay, "Semantic Capability Model for the Simulation of Manufacturing Processes," in *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. SCITEPRESS - Science and Technology Publications, 2024.

# Fortschritte der interdisziplinären Forschung zur Digitalisierung in der Leichtbauproduktion: Anwendungen und Optimierungen in der LaiLa Modellfabrik

Vincent Adomat\*, Maryam Ahanpanjeh, Christian Kober, Marc Fette, Jens P. Wulfsberg

*Institute of Production Engineering*

*Helmut Schmidt University/University of the Federal Armed Forces Hamburg*

Hamburg, Germany

\*vincent.adomat@hsu-hh.de

**Zusammenfassung** — Die Modellfabrik des dtec Labors für intelligente Leichtbauproduktion (LaiLa) hat sich in den letzten drei Jahren als Digitalisierungslabor für intelligente Leichtbauproduktion etabliert und den Wissenstransfer zwischen Grundlagenforschung, anwendungsnaher Forschung und industrieller Anwendung in der Luftfahrt vorangetrieben. Durch die Implementierung und Optimierung von Technologien wie energieoptimierter Produktionsplanung und -steuerung, robotergestütztem kontinuierlichem Ultraschallschweißen und digitalen Zwillingen konnten signifikante Fortschritte in Bezug auf Qualitätssteigerung, individualisierte Fertigung, Ressourceneffizienz und Transparenz erzielt werden. Der vorliegende Beitrag präsentiert die gewonnenen Erkenntnisse und zeigt auf, wie die dargestellte Digitalisierung der Modellfabrik dazu beiträgt, Nachhaltigkeit und Effizienz in der Leichtbauproduktion durch intelligente Vernetzung, automatisierte Produktionssysteme, Echtzeit-Abbildung, Überwachung und Steuerung sowie flexibles Produktionsdesign zu erreichen. Darüber hinaus werden die Potenziale für die Weiterentwicklung und Skalierung der erprobten Lösungen in der industriellen Anwendung diskutiert, um den Weg für eine zukunftsorientierte und wettbewerbsfähige Leichtbauproduktion in der Luftfahrtindustrie und darüber hinaus zu ebnen. Im Einklang mit den übergeordneten Zielen von LaiLa und dtec trägt die Modellfabrik dazu bei, die Digitalisierung in der Leichtbauproduktion voranzutreiben und die Wettbewerbsfähigkeit der deutschen Luftfahrtindustrie zu stärken. Durch die enge Zusammenarbeit von Partnern aus Industrie und Wissenschaft fördert LaiLa den Technologietransfer und die Entwicklung innovativer Lösungen.

**Keywords** — Digitalisierung, Leichtbauproduktion, Produktionsplanung, Ultraschallschweißen, Digitale Zwillinge, Nachhaltigkeit

## I. EINLEITUNG

Unternehmen müssen eine Vielzahl von Strategien verfolgen, die sowohl technologische als auch organisatorische Aspekte berücksichtigen, um wettbewerbsfähig zu bleiben. Angesichts der zunehmenden Herausforderungen des Klimawandels und der Notwendigkeit einer nachhaltigen klimaneutralen Produktion gewinnen energieoptimierte Prozesse zunehmend an Bedeutung. Die Implementierung solcher Prozesse und die Nutzung umweltfreundlicher Materialien

helfen, ökologische Anforderungen zu erfüllen und gleichzeitig Kosten zu senken. Die Verwendung von kohlenstofffaserverstärkten Kunststoffen (CFK) ist ein zentraler Aspekt für eine nachhaltigere Produktion und Produktentwicklung, da sie zur Erreichung von Umweltzielen wie der Reduzierung von Treibhausgasemissionen und der Ressourceneffizienz beitragen können. Trotz deren geringen Gewichts bieten CFK-Bauteile eine hervorragende Festigkeit und Steifigkeit. Diese Materialien leisten einen wesentlichen Beitrag zur Modernisierung und Nachhaltigkeit der Luftfahrtindustrie, indem sie die Effizienz, Leistung und Umweltfreundlichkeit von Flugzeugen verbessern.

Unternehmen müssen nicht nur effizienter werden, sondern auch flexibler, um den steigenden Anforderungen gerecht zu werden, ohne die Produktqualität zu beeinträchtigen. Ein zentraler Schritt ist die Förderung von Innovationen. Durch Investitionen in Forschung und Entwicklung können Unternehmen kontinuierlich neue Produkte und Technologien entwickeln. Der Einsatz fortschrittlicher Technologien zur Optimierung von Produktionsprozessen und die Umsetzung von Industrie 4.0-Prinzipien unterstützen die Digitalisierung und Vernetzung der Produktionsabläufe, was zu höherer Effizienz und besserer Anpassungsfähigkeit führt.

Dabei sind umfassende Qualitätskontrollen und kontinuierliche Verbesserungsprozesse von großer Bedeutung. Schlüsseltechnologien wie Künstliche Intelligenz (KI), das Internet der Dinge (IoT) und Digitale Zwillinge (DTs) spielen hierbei eine entscheidende Rolle. Der Einsatz digitaler Werkzeuge zur Überwachung und Verbesserung der Qualität bietet vielversprechende Ansätze zur Steigerung der Wettbewerbsfähigkeit. Am Composite Technology Center (CTC) GmbH, dem Leichtbau-Technologie- und Innovationszentrum von Airbus in Deutschland, wird eine zentrale Laborplattform eingerichtet. Diese Plattform dient als praxisnahe Entwicklungs- und Validierungsumgebung für alle Arbeitspakete und die im Projekt definierten Anwendungsbereiche. Drei dieser Anwendungsbereiche werden im Folgenden bezüglich der wissenschaftlichen sowie industriellen Relevanz vorgestellt

und das damit einhergehende systematische Vorgehen im Projekt erläutert.

## II. WEITERENTWICKLUNG ENERGIEOPTIMierter PPS IN DER KOMPOSITAVERARBEITUNG: ERKENNTNISSE UND FORTSCHRITTE DER LAiLa MODELLFABRIK

Die Herausforderungen des Klimawandels und die daraus resultierenden Anforderungen an eine nachhaltige und klimaneutrale Produktion haben in den letzten Jahren auch in Unternehmen aus dem Bereich der Kompositverarbeitung nochmals an Bedeutung zugelegt. Kompositwerkstoffe, insbesondere kohlenstoffbasierte faserverstärkte Kunststoffe (CFK, FVKs), bieten aufgrund ihrer hervorragenden Eigenschaften und ihres Potenzials zur Gewichtsreduktion bereits heute zahlreiche Vorteile für eine ressourcenschonende und energieeffiziente Fertigung [1]. Um jedoch auch in Zukunft wettbewerbsfähig zu bleiben und den steigenden Anforderungen an Nachhaltigkeit gerecht zu werden, ist eine kontinuierliche Optimierung der Produktionsplanung und -steuerung (PPS) in der Kompositverarbeitung unerlässlich. In diesem Kapitel werden die Fortschritte und Erkenntnisse der letzten zwei Jahre in der LaiLa Modellfabrik im Bereich der energieoptimierten PPS vorgestellt und diskutiert, um aufzuzeigen, wie Digitalisierung dazu beitragen kann, die Kompositverarbeitung langfristig international wettbewerbsfähig und klimaneutral zu gestalten.

In den letzten zwei Jahren hat die Forschung im Rahmen des dtec Labors für intelligente Leichtbauproduktion (LaiLa) deutliche Fortschritte bei der Entwicklung von Methoden und Werkzeugen zur Erweiterung der PPS in der Kompositverarbeitung um die Perspektive „Energiebedarf und Energiekosten“ erzielt. Durch die Arbeit an angepassten Algorithmen und Optimierungsmodellen [2] zur dynamischen Anpassung der Produktionsplanung konnten erste Erkenntnisse in realen Produktionsumgebungen gewonnen werden, die Unternehmen in die Lage versetzen, energieintensive Prozesse kostenoptimal zu planen und durchzuführen. Die entwickelten Ansätze zielen darauf ab, Kosten und CO<sub>2</sub>-Emissionen zu reduzieren, indem energie- und arbeitsintensive Prozesse identifiziert, entkoppelt und zeitlich verlagert werden. Diese ersten Tests und Validierungen der Methoden und Werkzeuge deuten einen potenziellen Nutzen für Unternehmen, nicht nur in der Kompositverarbeitung, an. Durch die Integration der Perspektive „Energiebedarf und Energiekosten“ in das eigene PPS-System könnte das dtec Projekt LaiLa zukünftig auch einen dauerhaften Beitrag zur Steigerung der Wettbewerbsfähigkeit und Nachhaltigkeit der Branche in Zeiten zunehmend volatiler Energiemärkte und wachsender Anforderungen an den Klimaschutz leisten. Im Folgenden werden die bisherigen Ergebnisse und Erkenntnisse näher beleuchtet und diskutiert, um das Potenzial der entwickelten Lösungen aufzuzeigen und weitere Schritte für deren Umsetzung in der Praxis abzuleiten.

Der Digitalisierung kommt eine zentrale Rolle bei der Optimierung des Energieverbrauchs in der Verarbeitung von kohlenstofffaserverstärkten Kunststoffen (CFK) zu. Durch den Einsatz von Sensoren und dezentralen Messköpfen konnten

in LaiLa detaillierte Daten über den Energieverbrauch von Prozessen und einzelnen Komponenten gesammelt werden. Diese Daten ermöglichen eine tiefgehende Analyse und Optimierung der Produktionsprozesse. Temporäre Messungen helfen dabei, die Energieverbräuche einzelner, wiederkehrender Prozessschritte zu identifizieren. So wird eine Optimierung von Heiz- und Kühlsystemen sowie die Nutzung von Abwärme für Parallelprozesse ermöglicht [3].

Eckhoff et al. [3] zeigen zudem, dass durch die gezielte Messung und Analyse der Energieverbräuche konkrete Energietreiber identifiziert und gezielte Verbesserungen vorgenommen werden können. Ein Beispiel für die Anwendung temporärer Messungen ist der Fertigungsprozess im Resin-Transfer-Moulding Verfahren (RTM) für strukturelle Großbauteile aus CFK. So wurden Leistungsmessungen der Harz-Injektionsanlage vorgenommen. Die einzelnen Prozessschritte sind in den gemessenen Kurven anhand ihrer unterschiedlichen Energieverbräuche zu erkennen. Durch die Vermessung aller einzelnen Systeme wie Heizungssysteme, Kühlsysteme sowie der Injektionsanlage konnten Wechselwirkungen und Zusammenhänge analysiert werden. Dies ermöglichte es, bei der Messung der Verbraucher konkrete Energietreiber zu identifizieren. Eckhoff et al. [3] haben gezeigt, dass bei bestehenden Anlagen durch eine solche Analyse gezielte Verbesserungen erreicht werden können, beispielsweise durch optimierte Regler, effizientere Aktuatoren oder eine verbesserte Steuerung der Anlage. Die Energiedatenerfassung über modernes Messequipment ist für alle Prozessschritte essenziell, um eine energieoptimierte Produktionsplanung und Steuerung zu gewährleisten.

In [4] stellen die Autoren ein Mixed-Integer-Linear-Programming (MILP) Modell vor, das darauf abzielt, die Produktionsplanung in der diskreten Fertigung unter Berücksichtigung variabler Energiepreise zu optimieren. Dieses Modell ist ein zentraler Baustein der energiebewussten PPS-Forschung in LaiLa. Es unterscheidet zwischen energieintensiven und arbeitsintensiven Prozessen und ermöglicht es, energieintensive Prozesse in Zeiten niedrigerer Energiepreise zu verlagern, um Kosten zu senken und die Wettbewerbsfähigkeit zu steigern. Das Modell berücksichtigt wesentliche Parameter wie den Energiebedarf und die Energiekosten, die Verfügbarkeit und Kosten von Maschinen sowie die Verfügbarkeit und Kosten von qualifizierten Arbeitskräften. Die Zielfunktion maximiert die Gesamteinnahmen aus den Produktionsaktivitäten, indem alle relevanten Kostenfaktoren von den jeweiligen Einnahmen abgezogen werden. Nebenbedingungen stellen sicher, dass jede Operation höchstens einmal geplant wird, keine Maschine in einer Periode mehrfach belegt wird, die Maschinenbelegung die verfügbare Kapazität nicht überschreitet, die Mitarbeiterkapazität pro Periode und Kompetenz nicht überschritten wird und die Energiekapazität pro Periode nicht überschritten wird. Als Ergebnis gibt das Modell eine optimale zeitliche Verteilung der einzelnen zur Bearbeitung vorgesehenen Aufträge auf die verfügbaren Maschinen gemäß Abbildung 1 aus

Zur Validierung des Modells wurden numerische Experimente mit verschiedenen Datensätzen durchgeführt, die auf der

period	57	58	59	60	61	62	63	64	65
date	03.01.2022	03.01.2022	03.01.2022	03.01.2022	03.01.2022	03.01.2022	03.01.2022	03.01.2022	03.01.2022
time	09:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00
energy cost	122.93	110.17	93.38	96.73	82.37	80.9	80.59	98.00	123.51
machine 1			operation 21		operation 41	operation 26		operation 7	
machine 2					operation 4	operation 13			
machine 3							operation 38		
machine 4	operation 36	operation 40	operation 12						
machine 5	operation 42	operation 29	operation 44						
machine 6			operation 25		operation 22	operation 27			
machine 7			operation 1						
machine 8									
machine 9									
machine 10		operation 40	operation 35			operation 24	operation 47		
machine 11		operation 29		operation 50	operation 18	operation 46		operation 45	
machine 12									
machine 13					operation 43	operation 28	operation 6		
machine 14		operation 5	operation 9	operation 37	operation 25	operation 33		operation 10	
machine 15									
machine 16								operation 3	
machine 17									
machine 18					operation 11	operation 34	operation 14	operation 2	
machine 19	operation 30	operation 20				operation 14	operation 17	operation 15	
machine 20				operation 19			operation 16	operation 49	

ABILDUNG 1: EXEMPLARISCHE PLANUNGSTABELLE, WIE SIE DAS OPTIMIERUNGSMODELL AUS [4] AUSGIBT.

Modellfabrik des Projektes basieren. Diese Experimente zeigen, dass das Modell in der Lage ist, realistische Produktionszenarien zu optimieren und die Gesamteinnahmen zu steigern. Beispielsweise wurde für einen mittelgroßen Datensatz eine Einnahmesteigerung von 4,66 % durch die Berücksichtigung variabler Energiepreise erzielt. Die Ergebnisse verdeutlichen, dass eine flexible Planung energieintensiver Prozesse in Zeiten niedrigerer Energiepreise zu einer Erhöhung der Gesamteinnahmen führen kann. Dies unterstreicht die Bedeutung der Integration variabler Energiepreise in die Produktionsplanung, um sowohl wirtschaftliche als auch ökologische Ziele zu erreichen.

[2] erweitert das bestehende Mixed-Integer-Linear-Programming (MILP) Modell zur energiepreisoptimierten Produktionsplanung um spezifische Aspekte der Herstellung von kohlenstofffaserverstärkten Kunststoffen (CFRP). Diese Erweiterung berücksichtigt die besonderen Anforderungen und Einschränkungen der CFRP-Fertigung, wie die begrenzte Haltbarkeit von Materialien, die Notwendigkeit der Materialvorbereitung, das Öffnen von Formen während des RTM-Zyklus, wiederholte Autoklav-Aushärtungszyklen, das hohe individuelle Investitionsrisiko und die hohen Kosten für Verzögerungen. Das Modell integriert diese spezifischen Produktionsbedingungen und optimiert die Planung, indem es energieintensive Prozesse in Zeiten niedrigerer Energiepreise verlagert. Die experimentellen Ergebnisse zeigen, dass das Modell nicht nur für kleinere, sondern auch für größere industrielle Szenarien geeignet ist und eine signifikante Reduzierung der Energiekosten bei gleichzeitiger Maximierung der Einnahmen ermöglicht. Die Erweiterung hebt hervor, dass die Flexibilisierung der Produktionsplanung durch die Berücksichtigung variabler Energiepreise nicht nur zur Kostensenkung beiträgt, sondern auch die Wettbewerbsfähigkeit in einem globalisierten Markt sichert und die Umstellung auf erneuerbare Energien fördert. Zukünftige Arbeiten sollen das Modell um Techniken zur Vorhersage von Energiepreisen erweitern und betriebliche Abhängigkeiten detaillierter berücksichtigen.

Die Erprobung der belastungsorientierten Auftragsfreigabe (BoA) als Methode der Produktionsplanung und -steuerung (PPS) unter Berücksichtigung variabler Energiepreise stellt einen weiteren, jedoch heuristischen Ansatz der PPS-Optimierung dar. Die BoA, die ursprünglich darauf abzielt, die Produktionskapazitäten durch gezielte Freigabe von Aufträgen zu optimieren und Engpässe zu vermeiden, muss hierfür um zusätzliche Parameter und Entscheidungsvariablen erweitert werden. Insbesondere müssen die Energiepreise als

dynamische Größe in die Freigabeentscheidungen integriert werden. Dies bedeutet, dass neben der Berücksichtigung der Maschinen- und Personalkapazitäten auch die aktuellen und prognostizierten Energiepreise in die Entscheidungsfindung einfließen [5]. Die Anpassung der BoA erfordert die Entwicklung eines erweiterten Modells, das die zeitliche Verschiebung energieintensiver Prozesse in Perioden niedrigerer Energiepreise ermöglicht. Hierzu müssen die Energiebedarfe der einzelnen Aufträge ebenfalls bekannt sein und in Relation zu den variablen Energiepreisen gesetzt werden. Die Erweiterung der BoA um diese energiepreisabhängigen Entscheidungsvariablen ermöglicht es, die Produktionsplanung selber flexibler und kosteneffizienter zu gestalten, da auch Optimierungen mehrdimensionaler Systeme mit verhältnismäßigem Rechenaufwand näherungsweise gelöst werden können. Experimentellen Ergebnisse zeigen, dass durch die Integration variabler Energiepreise in die BoA eine relevante Reduzierung der Energiekosten bei gleichzeitiger Maximierung der Produktionsleistung erreicht werden kann. Dieses Thema ist ein aktueller Arbeitsschwerpunkt und wird in zukünftigen Forschungsarbeiten weiter vertieft, um die praktische Anwendbarkeit und die wirtschaftlichen Vorteile dieser Methode zu maximieren. Die belastungsorientierte Auftragsfreigabe (BoA) bleibt auch heute noch in der Produktionsplanung und -steuerung (PPS) relevant. Ihre weitläufige industrielle Verwendung beruht auf ihrer Fähigkeit, Produktionskapazitäten durch gezielte Freigabe von Aufträgen zu optimieren und Engpässe zu vermeiden. In den letzten zwei Jahren wurden an der LaiLa Modellfabrik umfangreiche Energiemessungen durchgeführt, um den Energieverbrauch in verschiedenen Produktionsprozessen detailliert zu erfassen. Basierend auf diesen Daten wurde ein Mixed-Integer-Linear-Programming (MILP) Optimierungsmodell entwickelt, das eine energiebewusste Produktionsplanung und -steuerung (PPS) ermöglicht. Dieses Modell optimiert die zeitliche Verlagerung energieintensiver Prozesse unter Berücksichtigung variabler Energiepreise, um Kosten zu senken und die Wettbewerbsfähigkeit zu steigern. Parallel dazu wurde die belastungsorientierte Auftragsfreigabe (BoA) untersucht und um energiepreisabhängige Entscheidungsvariablen erweitert. Dies ermöglicht eine flexible und kosteneffiziente Produktionsplanung, die auf aktuelle und prognostizierte Energiepreise reagiert. Ein weiterer Forschungsschwerpunkt lag auf der Analyse der Prozessketten für kohlenstofffaserverstärkte Kunststoffe (CFK), wobei energieintensive Prozessbestandteile identifiziert und das Potenzial für eine energieoptimierte PPS-Planung untersucht wurden. [6] Zusätzlich führten die Analysen der Messdaten zu direkten Verbesserungen der Energieeffizienz, was weitere positive Ergebnisse über die Optimierung der PPS hinaus ermöglichte. Insgesamt zeigen die Forschungsergebnisse, dass eine integrierte Betrachtung von Energieverbrauch und Produktionsplanung erhebliche wirtschaftliche und ökologische Vorteile bietet.

### III. ROBOTERGEFÜHRTES KONTINUIERLICHES ULTRASCHALLSCHWEISSEN

Der Einsatz von carbonfaserverstärktem Kunststoff (CFK) in der Luftfahrt steigt seit den 1960er-Jahren aus ökonomischen und ökologischen Gründen. Faserverstärkte Thermoplaste bieten im Vergleich zu duroplastischen Verbundwerkstoffen mehrere Vorteile, die sie für verschiedene Anwendungen in der Luftfahrtindustrie attraktiv machen, wie z. B. hohe Schlagzähigkeit, Bruchzähigkeit, Recyclingfähigkeit sowie kurze Verarbeitungszeit [7], [8]. Ihre Schweißbarkeit ist einer ihrer größten Vorteile. Das Ultraschallschweißen (USW), eine der in den letzten Jahren am weitesten verbreiteten Füge-technologien für thermoplastische Verbundwerkstoffe, zeichnet sich durch seine hervorragende Festigkeit, Energieeffizienz, kurze Verarbeitungszeit und seine Fähigkeit zur vollständigen Automatisierung aus [9], [10], [11]. Beim kontinuierlichen Ultraschallschweißen gibt es zwei wichtige Phasen: Vibration und Konsolidierung. Während jeder Phase bestimmen einige Schlüsselparameter und Kriterien die Qualität der endgültigen Schweißnaht. In der Vibrationsphase erfolgt die Erwärmung durch mechanische Schwingungen mit hoher Frequenz (20–40 kHz) und niedriger Amplitude (50–100  $\mu\text{m}$ ). Ein Wandler wandelt elektrische Schwingungen in mechanische Schwingungen um, die dann über eine Sonotrode auf die Schweißteile übertragen werden. Diese Schwingungen versetzen Molekülketten in Schwingung, wenn sie auf ein Material treffen. Wenn sich die Moleküle bewegen, reiben sie aneinander und erzeugen Reibungswärme. Dadurch steigt die Temperatur an der Kontaktfläche an, und das thermoplastische Harz erreicht seine Glasübergangstemperatur ( $T_g$ ). Bei fortgesetzter Erwärmung findet die viskoelastische Wärme statt und die Fügezone beginnt zu schmelzen [11], [12]. In der nächsten Phase wird die Schweißnaht unter einer Konsolidierungseinheit unter die  $T_g$  abgekühlt.

Der Einsatz von Industrierobotern wird zur Notwendigkeit, wenn das Ultraschallschweißen zum Fügen komplexer und großer Bauteile eingesetzt wird. Industrieroboter sind flexibler als Schweißmaschinen und besitzen einen großen Arbeitsbereich, allerdings bringen sie trotz ihrer Vorteile auch Herausforderungen mit sich. Industrieroboter verfügen über geringe absolute Genauigkeit aufgrund der mangelnden Steifigkeit und deren seriellen Kinematik [13]. Zudem sind sie durch eine enge Wechselwirkung zwischen dem Prozess und der mechanischen Roboterstruktur gekennzeichnet. Die am Prozess beteiligten Kräfte führen zu Auslenkungen des Roboters und des Endeffektors, welche Abweichungen von der gewünschten Bahn verursacht und den Ablauf vom Prozess beeinträchtigt [14]. Daher ist es eine große Herausforderung die Roboter in manchen Applikationen einzusetzen, insbesondere bei Prozessen mit hohen statischen und dynamischen Kräften. Die Fähigkeit, jederzeit gleichbleibend gute Schweißnähte zu liefern und gleichzeitig robust genug zu sein, um mit Fertigungstoleranzen umzugehen, ist von einer großen Bedeutung, um den Prozess auf ein höheres Technologieniveau zu bringen.

Dieses erfordert jedoch ein gutes Verständnis des Prozesses und des Verhaltens des Roboters sowie der Wechselwirkungen zwischen beiden.

In der Vibrationsphase ist es von entscheidender Bedeutung, die Prozessparameter richtig zu wählen, um die notwendige Energie für die Herstellung einer Schweißverbindung zu erzeugen. Diese Prozessparameter sind Schwingungsfrequenz, Amplitude und statische Kraft der Sonotrode. Außerdem muss sichergestellt werden, dass die erzeugte Energie in angemessener Weise in die Fügezone übertragen wird. Die Energie der Schwingungsamplitude, die sich in der Schweißzone verteilt, hängt von der richtigen Ausrichtung der Sonotrode in Bezug auf das obere Bauteil ab [15], [16]. Die maximale Wärmeabgabe in einer Schweißzone kann nur bei vertikaler Positionierung der Sonotrode erreicht werden. Während der Konsolidierungsphase muss genügend Wärme von der Fügezone abgeführt werden, um sie unter die Kristallinitätstemperatur zu bringen. Dies hängt von der Konsolidierungskraft und der Konsolidierungszeit ab. Die richtige Auswahl der Prozessparameter und die richtige Einstellung der wichtigen Kriterien reichen nicht aus, um Schweißnähte von wiederholbarer Qualität zu erzielen, da es während des Schweißprozesses zu Abweichungen kommen kann. Um wiederholbar hochwertige Schweißnähte zu erzielen, müssen diese Parameter und Kriterien überwacht und schließlich kontrolliert werden.

Das Ziel dieser Forschungsarbeit im Rahmen des LaiLa-Projektes ist die Erzeugung reproduzierbarer Schweißqualität durch Erstellung einer automatisierten Schweißbahn beim robotergeführten kontinuierlichen Ultraschallschweißen mittels sensorgestützter Bewegungsausführung und Bewegungsbeobachtung, wo auch die wichtigsten Prozessparameter überwacht und eventuell geregelt werden. Abbildung 2 zeigt die zu entwickelnde automatisierte Bahnplanung, bei der die wichtigen Prozessparameter überwacht und gegebenenfalls geregelt werden.

Wie bereits erwähnt sind die Schweißkraft und die Konsolidierungskraft zwei Schlüsselparameter, die die Qualität von der Schweißnaht beeinflussen. Diese Kräfte werden unabhängig voneinander durch pneumatische Zylinder aufgebracht. Allerdings bleiben die eingestellten Kräfte durch die Interaktionen zwischen den Komponenten am Endeffektor während des Prozesses nicht konstant. Um diese Kräfte zu überwachen und bei Bedarf zu regeln, wurden unter der Kolbenstange der Pneumatikzylinder 1-D-Kraftsensoren integriert und die Kräfte (Z-Bewegungen von Komponenten am Endeffektor) geregelt. Ein zusätzlicher Kraftsensor wurde an der Rolle eingebaut, um die Vorspannkraft zu überwachen und zu regeln. Abbildung 3 zeigt die integrierten Kraftsensoren an der Sonotrode, Konsolidierungseinheit und der Rolle.

Die Ausrichtung der Sonotrode ist eines der wichtigsten zu überwachenden Kriterien, die die Schwingungsübertragung und die Gleichmäßigkeit der Druckverteilung während der Vibrationsphase beeinflusst. Zu diesem Zweck wurde eine Inertialmesseinheit (IMU) in der Nähe der Sonotrode am Endeffektor angebracht und ein auf Quaternionen basierender Kalman-

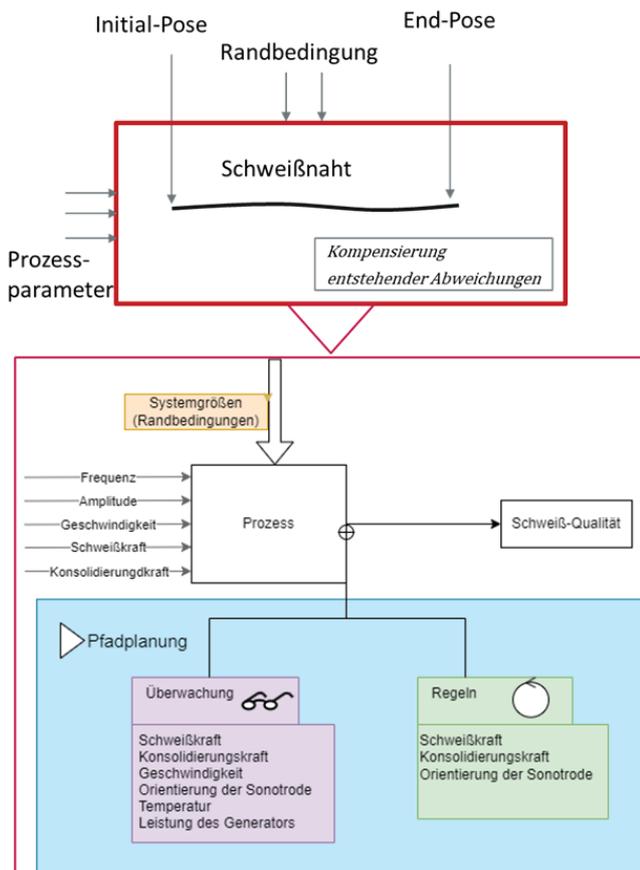


ABBILDUNG 2: ENTWICKELTE SENSORBASIERTE STRATEGIE FÜR AUTOMATISIERTE PFADPLANUNG

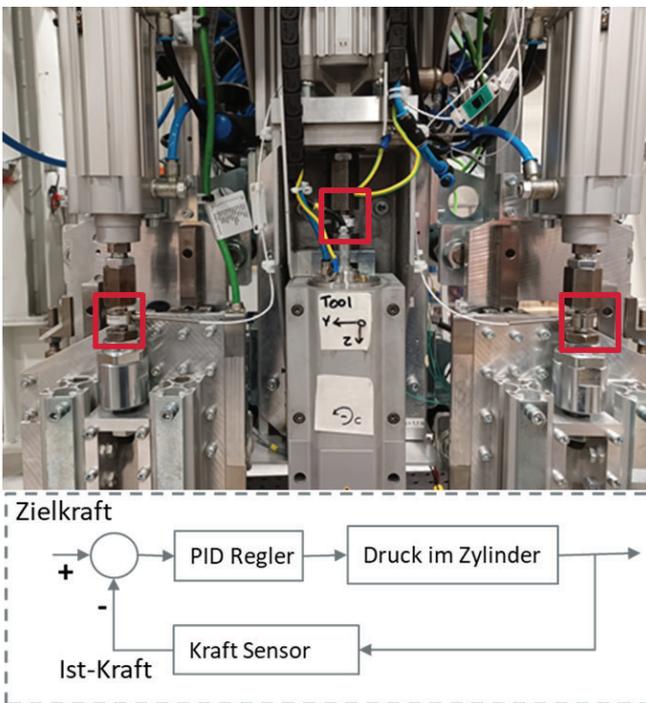


ABBILDUNG 3: INTEGRIERTE KRAFTSENSOREN UND DER REGELKREIS

Filter (EKF) integriert, um die Ausrichtung der Sonotrode in Echtzeit zu messen. Eine IMU, die in dieser Arbeit verwendet wird, besteht aus einem Drei-Achsen-Gyroskop und einem Drei-Achsen-Beschleunigungsmesser. Das Gyroskop misst die Winkelgeschwindigkeit des Objekts und der Beschleunigungssensor misst die externen Kräfte in Bezug auf den IMU-Rahmen (Körperrahmen). Durch die Integration der Winkelgeschwindigkeitsdaten von Gyroskop über die Zeit kann die IMU Änderungen in der Ausrichtung abschätzen, allerdings neigen Gyroskope dazu, mit der Zeit abzudriften, welches mit Hilfe von Beschleunigungsmesserdaten korrigiert werden kann. Da der Beschleunigungssensor bestimmte Kräfte, einschließlich der Schwerkraft, misst, kann er die Ausrichtung des Geräts in Bezug auf die vertikale Drehachse bestimmen. EKF wurde verwendet, um Daten von Beschleunigungsmessern und Gyroskopen zusammenzuführen und die Ausrichtung der Sonotrode um die X- (Schweißrichtung) und Y-Achse zu bestimmen. Mithilfe von diesem Sensor kann die vertikale Ausrichtung der Sonotrode auf den oberen Fügepartner sichergestellt werden.

Zusammenfassend gibt es bei einem Roboter-Ultraschallschweißprozess eine signifikante Wechselwirkung zwischen Prozess und Roboterdynamik. Das Schweißen komplexer Teile mit Doppelkrümmungen und Rampen erfordert eine aktive Bahnüberwachung und -steuerung, um jederzeit die richtige Positionierung und Ausrichtung der Komponenten zu erreichen. Ziel dieser Forschung ist es, die abgebildete Kompensationsstrategie mit Hilfe geeigneter Sensoren weiterzuentwickeln, um das wiederholbare Schweißen komplexer und gekrümmter Strukturen zu ermöglichen.

#### IV. DIE BEDEUTUNG VON FIDELITY IM KONTEXT DES NUTZENS VON DIGITALEN ZWILLINGEN

Eine Modellfabrik dient in erster Linie der Demonstration der Umsetzbarkeit und Machbarkeit neuer Technologien. In LaiLa wird jedoch auch der Transfer und die Implementierung dieser Technologien in die Industrie erforscht und durch die Entwicklung entsprechender Frameworks und Methoden unterstützt. Es wurde in der Praxis häufig beobachtet, dass im Rahmen der Digitalisierung der Fertigung, beispielsweise durch den Einsatz von Digitalen Zwillingen (DT), eine Diskrepanz zwischen dem erzielten Mehrwert und den entstehenden Mehrkosten besteht. Dieser Abschnitt fasst daher die Ergebnisse einer früheren Veröffentlichung zusammen, die diesen Aspekt vertieft analysiert [17].

Im Zuge der Umstellung auf Industrie 4.0 steigt die Nachfrage nach digitalen Werkzeugen und fortschrittlichen Technologien kontinuierlich. Gleichzeitig nimmt die Komplexität in der Fertigungsindustrie zu, da die Anforderungen an Individualisierung und Produktionsflexibilität zunehmen [18]. Um diesen Anforderungen über den gesamten Produktlebenszyklus hinweg gerecht zu werden, müssen Unternehmen ihre Effizienz kontinuierlich steigern und agiler werden, ohne dabei die Produktqualität zu beeinträchtigen [19]. Die Digitalisierung und ihre zugrundeliegenden Schlüsseltechnologien wie Künstliche Intelligenz (KI), Internet der Dinge (IoT) und DTs

bieten hierbei einen vielversprechenden Ansatz zur Steigerung der Wettbewerbsfähigkeit [20]. DTs sind ein technologisches Konzept, das zunehmend Beachtung in der Industrie findet [21]. Allerdings gibt es immer noch Uneinigkeit darüber, was einen DT genau ausmacht. Seit der ersten Definition im Jahr 2002 sind zahlreiche weitere entstanden [22], [23], [24]. Ein DT umfasst in der Regel eine physische Realität, ihre virtuelle Repräsentation und den Datenaustausch zwischen beiden. Trotz dieser grundlegenden Elemente besteht weiterhin kein allgemeiner Konsens über die Definition eines DTs. Zudem ist das Bewusstsein für die Beziehungen zwischen den variablen Elementen eines DTs gering, was zu einem unzureichenden Verständnis des Konzepts und seiner Vorteile führt [25]. Unternehmen investieren oft in DTs, ohne konkrete Ziele zu definieren, was dazu führt, dass praktische Implementierungen in der Fertigungsindustrie selten sind [26], [27], [28]. Zudem fehlt es an einer Bewertung der Kosten für die Erstellung und Nutzung von DTs, was zu Akzeptanzproblemen führt [17].

Eine wesentliche Variable, die den Nutzen von DTs beeinflusst, ist die DT-Fidelity. Fidelity beschreibt die "Anzahl der Parameter, ihre Genauigkeit und ihr Abstraktionsniveau" [29]. Um durch die Implementierung von DTs Vorteile zu erzielen, muss die Fidelity bereits in der Entwurfs- und Konzeptentwicklungsphase berücksichtigt werden [30]. Akademische Definitionen legen nahe, dass ein DT mit hoher Fidelity modelliert werden sollte [23], [31]. Jedoch ist eine hohe Fidelity für den praktischen Einsatz nicht immer erforderlich [30], da sie zu höheren Kosten führt [32], [33]. Daher ist es entscheidend, ein geeignetes Fidelityniveau festzulegen. Bestehende Verfahren zur Auswahl geeigneter Fidelityniveaus sind jedoch oft subjektiv und erfordern Erfahrung, die häufig nicht vorhanden ist. Trotz der Relevanz von DT-Fidelity mangelt es weiterhin an Methoden, die potenzielle Implementierungen anleiten [30]. Ein Ansatz zur Ermittlung geeigneter Fidelitystufen in Produktionsumgebungen ist das Digital Twin Fidelity Requirements Model (DT-FRM) [30]. Dieses Kapitel trägt zum Verständnis des DT-Konzepts und der Beziehungen zwischen den relevanten Variablen bei und untersucht die Auswirkungen der DT-Fidelity auf den Nutzen, definiert als Differenz zwischen Einsparpotenzial und Kosten. Die in diesem Kapitel vorgestellten Forschungsergebnisse schließen die Lücke im Verständnis des Einflusses von DT-Fidelity. Zudem wird untersucht, ob die analysierten Zusammenhänge grafisch dargestellt werden können, um als Referenz im DT-Konzeptualisierungs- und Designprozess zu dienen. Ziel ist es, realistischere Erwartungen bei potenziellen Nutzern zu erzeugen und die Akzeptanz in Organisationen zu erhöhen.

Design Science Research (DSR) wurde verwendet, um die DT Benefit Curves (DT-BC) zu entwickeln und zu validieren, die in mehreren Designzyklen und basierend auf einer umfangreichen Literaturanalyse erstellt wurden [34]. Die Entwicklung der DT-BCs begann mit einer Analyse relevanter Literatur zu ähnlichen Konzepten in der Fertigungsindustrie, gefolgt von einer schrittweisen Anpassung der Suchbegriffe und der Anwendung der Analogiemethode. Theoretisches und empirisches Wissen wurden in den Entwürfen verarbeitet. Vorläufige

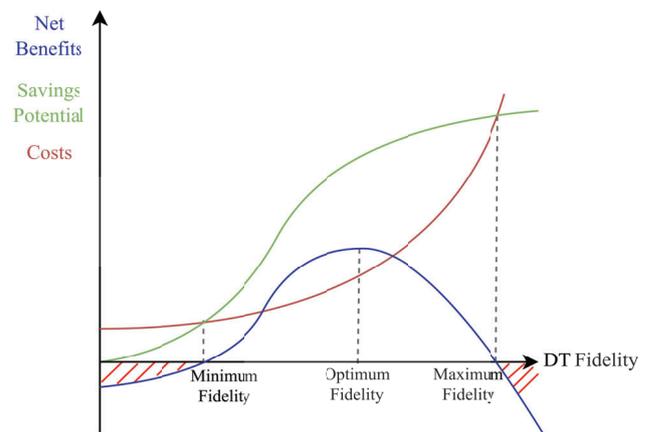


ABBILDUNG 4: DIGITAL TWIN BENEFIT-CURVES AUS [17]

Entwürfe wurden in Diskussionen mit Industrieexperten und Wissenschaftlern überprüft und angepasst. Schließlich wurden die Entwürfe in einer Online-Fokusgruppe bewertet und weiterentwickelt, bis die finalen DT-BCs festgelegt wurden [35]–[37], siehe Abbildung 4.

In diesem Kapitel wird der Verlauf der Kurven für das Einsparpotenzial und die Kosten unter Bezugnahme auf die einschlägige Literatur erläutert. Abbildung 4 veranschaulicht die typische Entwicklung der Kosten und des Einsparpotenzials in Abhängigkeit von der DT-Fidelity, wobei der Schwerpunkt auf relativen Relationen und nicht auf absoluten Werten liegt. Die Kostenkurve schneidet die Ordinate oberhalb des Ursprungs, da zur Realisierung des DT zunächst Investitionskosten anfallen, beispielsweise für Hardware, Software und Personal. Diese Kosten sind unabhängig von der tatsächlichen DT-Fidelity und hängen stark vom Digitalisierungsreifeegrad der Organisation ab. Der exponentielle Verlauf der Kostenkurve basiert auf der Analogie zu den Komplexitätskosten: Je höher die Komplexität, desto höher sind die Kosten für Modellierung, Berechnung, Modifikation und Interpretation der Ergebnisse. Tools und Knowhow sind meist schon auf einem niedrigen Fidelityniveau vorhanden, während mit zunehmender Fidelity die Lücke zwischen vorhandener und benötigter Erfahrung wächst. Infolgedessen steigt der Aufwand zur Erreichung eines höheren Fidelity-Levels exponentiell. Dieser Effekt ist das Gegenteil des Lernkurveneffekts, bei dem die Kosten mit zunehmender Erfahrung sinken. Zudem erfordert die Realisierung von DTs mit sehr hoher Fidelity oft individuelle Spezialsoftware oder kostspielige Kombinationen mehrerer Softwaretools, was weitere Forschung erforderlich machen könnte. Eine weitere Analogie ist die Progression der Zehnerregel: Je später im Entwicklungsprozess eine Änderung notwendig wird, desto teurer ist sie, da die Komplexität und Abhängigkeiten des Systems zunehmen. Dieser Zusammenhang lässt sich auch auf die DT-Fidelity übertragen.

Hinsichtlich des Einsparpotenzials folgt der typische Trend einer S-Kurve. Zunächst ist der Wissenszuwachs durch ein Low-Fidelity-Modell relativ gering. In diesem Stadium können die meisten Erkenntnisse wahrscheinlich mit anderen Methoden als der Verwendung des DTs gewonnen werden. Ab einer

bestimmten Mindestfidelity übersteigt das Einsparpotenzial jedoch die Kosten, sodass der Nettonutzen positiv wird. Dieser Übergang erfolgt, da mit zunehmender Fidelity neues Wissen aufgedeckt und durch den DT adressiert wird. Die Steigung der Einsparpotenzialkurve erreicht an ihrem Drehpunkt ein Maximum, abhängig von der zielgerichteten Nutzung des DTs. Ab einem bestimmten Fidelityniveau steigt das Einsparpotenzial jedoch nicht mehr weiter an und erreicht einen asymptotischen Wert. An diesem Punkt dominieren die Kosten durch ihren exponentiellen Verlauf, sodass der Nettonutzen negativ wird. Das Optimum des Nettonutzens liegt somit zwischen der minimalen und der maximalen Fidelity, abhängig vom spezifischen Anwendungsfall. Abbildung 4 zeigt daher eine typische, beispielhafte Kurve, die für DTs charakteristisch ist.

Zusammenfassend schärfen die Diagramme das Bewusstsein und das Verständnis dafür, wie die Fidelity den Nettonutzen beeinflusst, der durch die Anwendung von DTs in Fertigungsunternehmen erzielt wird. Dieses gesteigerte Bewusstsein führt zu einer besseren Einschätzung, ob sich die hohen Anfangskosten für die Realisierung von DTs lohnen. Außerdem kann die Akzeptanz von DTs bei Stakeholdern durch die Visualisierung der Beziehungen anhand der vorgestellten Grafiken verbessert werden. Empirische Validierungen haben gezeigt, dass die Kennlinien als relevant wahrgenommen werden und ein besseres Verständnis ermöglichen.

#### V. ZUSAMMENFASSUNG UND AUSBLICK

Der vorangegangene Text stellt die bisherige Optimierung moderner Produktionstechnologien im Rahmen des Forschungsprojektes „LaiLa“ dar. Ein zentraler Schwerpunkt des Projektes liegt auf der Digitalisierung und Nachhaltigkeit. Unternehmen müssen technologische und organisatorische Strategien verfolgen, um wettbewerbsfähig und nachhaltig zu bleiben. Kohlenstofffaserverstärkte Kunststoffe (CFK) spielen hierbei eine zentrale Rolle, insbesondere in der Luftfahrtindustrie. Technologien wie Industrie 4.0, Künstliche Intelligenz (KI), Internet der Dinge (IoT) und Digitale Zwillinge (DTs) sind entscheidend für die Steigerung der Effizienz und Nachhaltigkeit. Die LaiLa Modellfabrik hat durch die Implementierung und Optimierung dieser Technologien signifikante Fortschritte in Bezug auf Qualitätssteigerung, individualisierte Fertigung, Ressourceneffizienz und Transparenz erzielt. Ein weiterer wichtiger Aspekt ist das Ultraschallschweißen und der Einsatz von Industrierobotern. Der Einsatz von Industrierobotern ist aufgrund ihrer Flexibilität notwendig, jedoch herausfordernd, da sie eine geringe absolute Genauigkeit aufweisen und Wechselwirkungen zwischen dem Prozess und der Roboterstruktur bestehen. Im Rahmen des LaiLa-Projektes wird daran gearbeitet, diese Herausforderungen durch die Entwicklung einer automatisierten Schweißbahn mittels sensorgestützter Bewegungsausführung zu bewältigen, bei der die wichtigsten Prozessparameter und -faktoren überwacht und gegebenenfalls geregelt werden. Zur Optimierung der Produktionsplanung wurde ein Mixed-Integer-Linear-Programming (MILP) Modell entwickelt, das die Produktionsplanung unter Berücksichtigung variabler Energiepreise optimiert. Die Er-

gebnisse zeigen, dass eine flexible Planung energieintensiver Prozesse in Zeiten niedrigerer Energiepreise die Gesamteinnahmen steigern und die Energiekosten senken kann. Dieses Modell ist ein zentraler Baustein der energiebewussten Produktionsplanung und -steuerung (PPS) im LaiLa-Projekt und ermöglicht es, energieintensive Prozesse in Zeiten niedrigerer Energiepreise zu verlagern. Ein weiterer Aspekt der Forschung ist die Untersuchung von Digitalen Zwillingen. Ein Digitaler Zwilling umfasst eine physische Realität, ihre virtuelle Repräsentation und den Datenaustausch zwischen beiden. Die Genauigkeit und Detailliertheit eines Digitalen Zwillings (Fidelity) beeinflusst den Nutzen und die Kosten. Es ist entscheidend, ein geeignetes Fidelityniveau festzulegen, um den maximalen Nettonutzen zu erzielen. Im Rahmen des LaiLa-Projektes wird daran gearbeitet, geeignete Fidelityniveaus zu bestimmen und die Akzeptanz von Digitalen Zwillingen in der Industrie zu erhöhen.

Im weiteren Verlauf des Projektes „LaiLa“ wird die Weiterentwicklung der sensorbasierten Technologie angestrebt, um die Überwachung und Regelung der Schweißparameter zu verbessern. Dies umfasst die Integration fortschrittlicher Sensoren, die eine präzisere Erfassung und Steuerung der Schweißprozesse ermöglichen. Zudem soll das bestehende Mixed-Integer-Linear-Programming (MILP) Modell erweitert werden, indem Techniken zur Vorhersage von Energiepreisen und eine detailliertere Berücksichtigung betrieblicher Abhängigkeiten integriert werden. Dies wird die Flexibilität und Effizienz der energieoptimierten Produktionsplanung weiter erhöhen. Ein weiterer Schwerpunkt liegt auf der Optimierung der Digitalen Zwillinge, wobei Methoden zur Bestimmung des optimalen Fidelityniveaus für verschiedene Anwendungsfälle entwickelt werden sollen. Schließlich wird die Erforschung und Implementierung nachhaltiger Materialien und Prozesse in der Produktion fortgesetzt, um die ökologischen und ökonomischen Vorteile weiter zu maximieren.

Der Forschungsausblick des Projektes „LaiLa“ sieht eine Erweiterung der Anwendung von Digitalen Zwillingen vor, um deren praktische Implementierung und Akzeptanz in verschiedenen Industrien zu untersuchen. Zudem wird der Einsatz von Künstlicher Intelligenz (KI) zur weiteren Optimierung der Produktionsprozesse und zur Vorhersage von Prozessabweichungen erforscht. Ein weiterer Schwerpunkt liegt auf der Entwicklung neuer Modelle und Algorithmen zur Reduzierung des Energieverbrauchs und der CO<sub>2</sub>-Emissionen, um energieeffiziente Produktionssysteme zu fördern. Schließlich wird die interdisziplinäre Forschung vorangetrieben, um die Zusammenarbeit zwischen verschiedenen Disziplinen zu fördern und innovative Lösungen für komplexe Produktionsherausforderungen zu entwickeln.

#### DANKSAGUNG

Die Autoren danken dtec.bw – dem Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr und CTC GmbH – dem Composite Technology Center (An Airbus Company) für die Unterstützung dieser Arbeit im

Rahmen des Projekts "LaiLa – Labor für intelligente Leichtbauproduktion". Dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] V. Adomat, M. Ahanpanjeh, C. Kober, M. Fette, and J. P. Wulfsberg, "Interdisziplinäre forschungsperspektiven auf die digitalisierung in der leichtbauproduktion und anwendungsmöglichkeiten in der laila modellfabrik." [Online]. Available: <https://openhsu.ub.hsu-hh.de/handle/10.24405/14557>
- [2] V. Adomat, D. Trautwein, J. Eckhoff, and J. P. Wulfsberg, "Redesigning industrial composite production planning and control for energy awareness," *Procedia CIRP*, vol. 119, p. 582–588, 2023. [Online]. Available: <http://dx.doi.org/10.1016/j.procir.2023.02.152>
- [3] J. Eckhoff, V. Adomat, and M. Fette, "Energieoptimierte produktionsplanung: Am beispiel der faserverbundkunststoffverarbeitenden luftfahrtindustrie," *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, vol. 118, no. 3, p. 133–137, Mar. 2023. [Online]. Available: <http://dx.doi.org/10.1515/zwf-2023-1032>
- [4] V. Adomat, C. Kober, M. Ahanpanjeh, and J. P. Wulfsberg, "A linear programming model for renewable energy aware discrete production planning and control," 2023. [Online]. Available: <https://www.repo.uni-hannover.de/handle/123456789/13602>
- [5] S. Willeke, S. Wesebaum, and G. Ullmann, "Integration zeitvariabler energiekosten in die belastungsorientierte auftragsfreigabe," p. Issue 07, 2015. [Online]. Available: <http://www.logistics-journal.de/not-reviewed/2015/07/4200>
- [6] J. Eckhoff, V. Adomat, M. Fette, and J. P. Wulfsberg, "Energy optimized process design and scheduling in the field of large scale cfrp parts," in *SAMPE Europe Conference 2023*. Madrid, Spain: Helmut Schmidt University, 2023.
- [7] "Review of thermoplastic composites in AerospaceIndustry," vol. 3, no. 1. [Online]. Available: <https://skeenapublishers.com/journal/ijeti/IJETI-03-00031.pdf>
- [8] S. L. Omairey, S. Sampethai, L. Hans, C. Worrall, S. Lewis, D. Negro, T. Sattar, E. Ferrera, E. Blanco, J. Wighton, L. Muijs, S. L. Veldman, M. Doldersum, R. Tonnaer, N. Jayasree, and M. Kazilas, "Development of innovative automated solutions for the assembly of multifunctional thermoplastic composite fuselage," vol. 117, no. 5, pp. 1721–1738. [Online]. Available: <https://link.springer.com/10.1007/s00170-021-07829-2>
- [9] B. Jongbloed, J. Teuwen, R. Benedictus, and I. F. Villegas, "On differences and similarities between static and continuous ultrasonic welding of thermoplastic composites," vol. 203, p. 108466. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1359836820335149>
- [10] I. F. Villegas, "Strength development versus process data in ultrasonic welding of thermoplastic composites with flat energy directors and its application to the definition of optimum processing parameters," vol. 65, pp. 27–37. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S13598335X14001614>
- [11] Zongbo Zhang, Xiaodong Wang, Yi Luo, Zhenqiang Zhang, and Liding Wang, "Study on heating process of ultrasonic welding for thermoplastics," vol. 23, no. 5, pp. 647–664. [Online]. Available: <http://journals.sagepub.com/doi/10.1177/0892705709356493>
- [12] S. Tutunjian, O. Eroglu, M. Dannemann, N. Modler, and F. Fischer, "A numerical analysis of an energy directing method through friction heating during the ultrasonic welding of thermoplastic composites," vol. 33, no. 11, pp. 1569–1587. [Online]. Available: <http://journals.sagepub.com/doi/10.1177/0892705719833108>
- [13] K. Wu, J. Li, H. Zhao, and Y. Zhong, "Review of industrial robot stiffness identification and modelling," vol. 12, no. 17, p. 8719. [Online]. Available: <https://www.mdpi.com/2076-3417/12/17/8719>
- [14] M. Ahanpanjeh, M. Fette, and J. P. Wulfsberg, "Contribution of process parameters to geometric deviations of the robot for continuous ultrasonic welding process of thermoplastic composites," in *2023 28th International Conference on Automation and Computing (ICAC)*, 2023, pp. 1–7.
- [15] Q. Zhi, L. Lu, Z.-X. Liu, and P.-C. Wang, "Influence of horn misalignment on weld quality in ultrasonic welding of carbon fiber/polyamide 66 composite," vol. 97, no. 5, pp. 133–143.
- [16] C. B. G. Brito, J. J. Teuwen, C. A. Dransfeld, and I. F. Villegas, "The effects of misaligned adherends on static ultrasonic welding of thermoplastic composites," *Composites Part A: Applied Science and Manufacturing*, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:246233489>
- [17] C. Kober, B. N. Algan, M. Fette, and J. P. Wulfsberg, "Relations of Digital Twin Fidelity and Benefits: A Design-to-Value Approach," *Procedia CIRP*, vol. 119, pp. 809–815, 2023. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2212827123005759>
- [18] W. ElMaraghy, H. ElMaraghy, T. Tomiyama, and L. Monostori, "Complexity in engineering design and manufacturing," *CIRP Annals*, vol. 61, no. 2, pp. 793–814, 2012. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0007850612002004>
- [19] A. Sanders, C. Elangeswaran, and J. Wulfsberg, "Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing," *Journal of Industrial Engineering and Management*, vol. 9, no. 3, pp. 811–833, 2016.
- [20] E. Oztemel and S. Gursev, "Literature review of Industry 4.0 and related technologies," *Journal of Intelligent Manufacturing*, vol. 31, no. 1, pp. 127–182, Jan. 2020. [Online]. Available: <https://doi.org/10.1007/s10845-018-1433-8>
- [21] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020, conference Name: IEEE Access. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9103025>
- [22] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihm, "Digital Twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896318316021>
- [23] E. H. Glaessgen and D. S. Stargel, "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles," Honolulu, HI, 2012, nTRS Author Affiliations: NASA Langley Research Center, Air Force Office of Scientific Research NTRS Report/Patent Number: NF1676L-13293 NTRS Document ID: 20120008178 NTRS Research Center: Langley Research Center (LaRC). [Online]. Available: <https://ntrs.nasa.gov/citations/20120008178>
- [24] E. VanDerHorn and S. Mahadevan, "Digital Twin: Generalization, characterization and implementation," *Decision Support Systems*, vol. 145, p. 113524, Jun. 2021. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0167923621000348>
- [25] C. Kober, F. Gomez Medina, M. Benfer, J. P. Wulfsberg, V. Martinez, and G. Lanza, "Digital Twin Stakeholder Communication: Characteristics, Challenges, and Best Practices," *Computers in Industry*, vol. in print, 2024. [Online]. Available: <https://www.ssrn.com/abstract=4692283>
- [26] R. Joppen, A. Lipsmeier, C. Tewes, A. Kühn, and R. Dumitrescu, "Evaluation of investments in the digitalization of a production," *Procedia CIRP*, vol. 81, pp. 411–416, Jan. 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2212827119303762>
- [27] C. Kober, M. Fette, and J. Wulfsberg, "Challenges of Digital Twin Application in Manufacturing," in *2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Dec. 2022, pp. 0162–0168.
- [28] T. Riedelsheimer, P. Lünemann, S. Wehking, and L. Dorfhuber, "Digital Twin Readiness Assessment," 2020.
- [29] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the Digital Twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, May 2020. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1755581720300110>
- [30] C. Kober, V. Adomat, M. Ahanpanjeh, M. Fette, and J. Wulfsberg, "Digital Twin Fidelity Requirements Model For Manufacturing," in *Proceedings of the 3rd Conference on Production Systems and Logistics (CPSL)*, Vancouver, 2022, pp. 595–611.
- [31] M. Grieves, "Digital Twin: Manufacturing Excellence through Virtual Factory Replication," 2014. [Online]. Available: <https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf>
- [32] L. Zhang, L. Zhou, and B. K. Horn, "Building a right digital twin with model engineering," *Journal of Manufacturing Systems*, vol. 59, pp. 151–164, Apr. 2021. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0278612521000455>

- [33] C. Kober, M. Fette, and J. P. Wulfsberg, "A Method for Calculating Optimum Digital Twin Fidelity," *Procedia CIRP*, vol. 120, pp. 1155–1160, 2023. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2212827123008739>
- [34] J. vom Brocke, A. Hevner, and A. Maedche, Eds., *Design Science Research. Cases*, ser. Progress in IS. Cham: Springer International Publishing, 2020. [Online]. Available: <http://link.springer.com/10.1007/978-3-030-46781-4>
- [35] M. C. Tremblay, A. R. Hevner, and D. J. Berndt, "Focus Groups for Artifact Refinement and Evaluation in Design Research," *Communications of the Association for Information Systems*, vol. 26, pp. 599–618, Jan. 2010, publisher: Association for Information Systems. [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=plh&AN=52686571&site=ehost-live&custid=ns065117>
- [36] T. Dresing and T. Pehl, *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende*, 8th ed. Marburg: dr. dresing & pehl GmbH, 2018.
- [37] P. Mayring, *Qualitative Content Analysis: A Step-by-Step Guide*, 1st ed. Thousand Oaks: SAGE Publications Ltd, Dec. 2021.

# UAV-Based Hyperspectral Imaging of Iron(III) on Steel Infrastructure

Aaron Wickers

*Chair of control engineering*

HELMUT SCHMIDT UNIVERSITY/  
UNIVERSITY OF THE FEDERAL ARMED FORCES

Hamburg, Germany

wickersa@hsu-hh.de

Dominik Thomas

*Chair of steel structures*

HELMUT SCHMIDT UNIVERSITY/  
UNIVERSITY OF THE FEDERAL ARMED FORCES

Hamburg, Germany

dominik.thomas@hsu-hh.de

Dr.-Ing. Mirco Alpen

*Chair of control engineering*

HELMUT SCHMIDT UNIVERSITY/  
UNIVERSITY OF THE FEDERAL ARMED FORCES

Hamburg, Germany

mirco.alpen@hsu-hh.de

Univ.-Prof. Dr.-Ing. Max Gündel

*Chair of steel structures*

HELMUT SCHMIDT UNIVERSITY/  
UNIVERSITY OF THE FEDERAL ARMED FORCES

Hamburg, Germany

max.guendel@hsu-hh.de

Univ.-Prof. Dr.-Ing. Joachim Horn

*Chair of control engineering*

HELMUT SCHMIDT UNIVERSITY/  
UNIVERSITY OF THE FEDERAL ARMED FORCES

Hamburg, Germany

joachim.horn@hsu-hh.de

**Abstract**— The inspection of infrastructure, such as bridges or floodgates, requires a significant human workload. The authors propose a general workflow for Unmanned Aerial System-based inspection of steel infrastructure using a hyperspectral camera to detect the spectral signature of iron(III) in rust, reducing the human workload by assigning visual inspection tasks to a UAS. A general scheme for corrosion damage derived from templates of different infrastructure operators and an algorithm for Iron(III) detection are presented. Furthermore, the paper presents an exemplary approach for UAS path planning and validates the concepts on point cloud data of the Travehafenbrücke in Hamburg, Germany.

**Keywords** - Inspection, UAS, Path Planning, Corrosion

## I. MOTIVATION

The annual cost of corrosion is estimated to be 3 to 4% of the global gross domestic product while early and accurate corrosion detection could decrease this cost by 15 to 35% [1]. The workflow for detection can be improved by going from manual inspection labour to applying drones with automatic path planning and hyperspectral imaging, as shown in this paper. Automatic path planning takes the workload from manual flight planning from the drone operator and optimizes the inspected area per time, while the hyperspectral camera allows the detection of the specific spectral signature

of iron(III), which is a common ion occurring in atmospheric corrosion of steel.

## II. INFRASTRUCTURE DIAGNOSTICS

### A. German standards for corrosion and coating evaluation

As required by law, infrastructure in Germany has to be inspected completely at least every six years, including every single component as well as closed and hardly reachable areas [2].

For steel constructions the intact state of the coating protection is of special interest. Therefore, the spotted corrosion areas are documented in size and location. The evaluation of the infrastructure agencies is influenced by the standards DIN EN ISO 4628-2 bis -5. The following defects are visually inspected at a close distance of around 30 cm, according to ISO 13076 and rated into different condition classes [3]:

- Blistering according to ISO 4628-2,
- Rusting according to ISO 4628-3,
- Cracking according to ISO 4628-4,
- Delamination according to ISO 4628-5,

The aim of the MISDRO project is to automatize and objectify this currently manual task by drone inspection and image segmentation algorithms for hyperspectral images [4].

### B. Defect detection from hyperspectral images

For the detection of coating defects, images are taken with a hyperspectral snapshot camera. The system is turned on and after a warm-up period of 5 minutes, a white calibration is done by flying at a height of around 1,50 cm over a Zenith Lite® Diffuse Target, 1000x1000x12 mm with an average reflectivity of 95%. Dark current calibration is done afterwards with a closed lid.

Different illumination conditions at the calibration point and inspection area lead to a different excitation spectrum, however these changes are neglected, as there are no calibration targets at the inspection side, and the illumination is expected to change the absolute values, but not the shape of the reflectance spectrum. In addition, atmospheric vapour absorption is neglected due to the close distance (< 20 m) between the image area and the sensor. The hyperspectral images are saved in a hyperspectral data cube  $(x, y, \lambda)$  with the spatial dimensions  $x, y$  and the corresponding wavelength  $\lambda$ , where one voxel contains the measured light intensity at the given spatial position and wavelength.

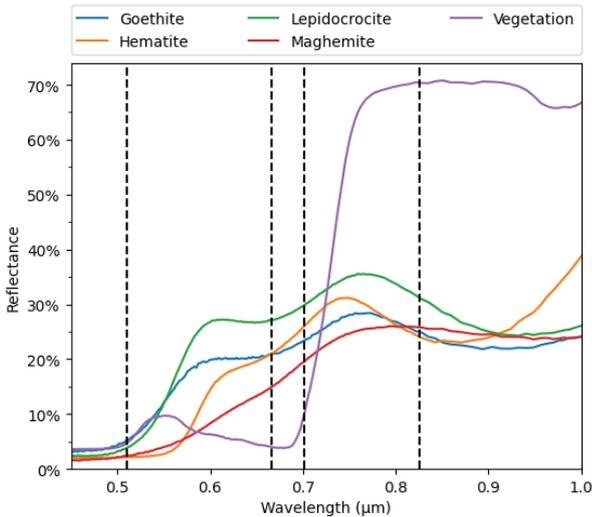


FIGURE 1. Reflectance spectra of some Iron(III)-containing minerals compared to the red edge of vegetation and selected wavelengths for hyperspectral Iron(III)-detection. Values at 510, 666, 702 and 826 nm are taken for Iron(III) detection. Reflectance spectra from USGS Spectral Library 7 [5]

The atmospheric steel corrosion leads to the build-up of Iron(III)-oxides, -hydroxides and -oxide-hydroxides, where the iron(III) has specific allowed electron transitions in the optical range [6]. These electron transitions happen in the atomic shell of the iron and thus do not depend on the chemical composition and crystal structure. Exemplary spectra of iron(III) containing minerals and vegetation is shown in Figure 1. The reflectance increase in the range from 500 nm to 750 nm overlap with the reflectance increase of vegetation from 680 nm to 730 nm and has to be accounted in the analysis.

### C. Spectral analysis

For the spectral analysis, several algorithms are widely applied, e.g. the Normalized Difference Vegetation Index (NDVI) or spectral angle mapper to calculate the similarity of a reflectance spectrum to spectra from a database. In this paper, we propose a simple spectral algorithm for corrosion detection in Algorithm 1.

---

#### Algorithm 1: Algorithm for iron(III) corrosion detection

---

**input :** Reflectance signal  $s_\lambda = R(\lambda)$  between 0 % and 100 % of a HSI pixel at wavelength  $\lambda = 510, 666, 702$  and  $826$  nm

**output:** Fe(III) index value between -1 and 1

```

1  $s_{hNDVI} = (s_{826nm} - s_{666nm}) / (s_{826nm} + s_{666nm})$ 
2 if  $s_{hNDVI} > 0.4$  then
3   | return -1 /* vegetation */
4  $s_{hNDFe3I} = (s_{702nm} - s_{510nm}) / (s_{702nm} + s_{510nm})$ 
5 return  $s_{hNDFe3I}$ 

```

---

The algorithms gives each spatial pixel a value between -1 and 1, with higher values indicating higher possible iron(III) content. Vegetation is filtered out by using a hyperspectral NDVI. In this case, spatial pixels with NDVI values greater 0.4 indicate a healthy vegetation and are therefore ignored [7]. The used wavelengths can also be found in Figure 1.

Spatial pixels given higher Iron(III) indicating values than 0.4 are taken as corrosion. An example image is given in Figure 2. The image shows two disadvantages of the simple algorithm: First, the top left corner contains highlighted spatial pixels, with a wrong calibration, because the white reference target did not cover the whole image during calibration. Second, the upper half of the image detected some spatial pixels of the sky as possible corrosion areas, due to a backlight increase, leading to wrong reflectance values above 100%. Both issues can be addressed by filtering out spatial pixels above this threshold.

By using the field of view of the camera and distance to the area of interest, the physical size of the corroded area can be determined and used for evaluation according to the rust grade in ISO 4628-3.

## III. UAS & SENSORS

The following section presents the sensor and UAS setup deployed, enabling a fast and inexpensive inspection of the structure without the need for long closure intervals with regard to the specific inspection requirements.

### A. Hyperspectral camera

The MISDRO project employs a Cubert Ultris X20 Plus as a hyperspectral snapshot camera with a weight of 630 g, dimensions of 86x121x105 mm, 640x640 spatial pixels, 164 spectral bands from 350 to 1000 nm with 4 nm spectral sampling and a FWHM of 10 nm each. Each pixel has a dynamic range of 12 bit.

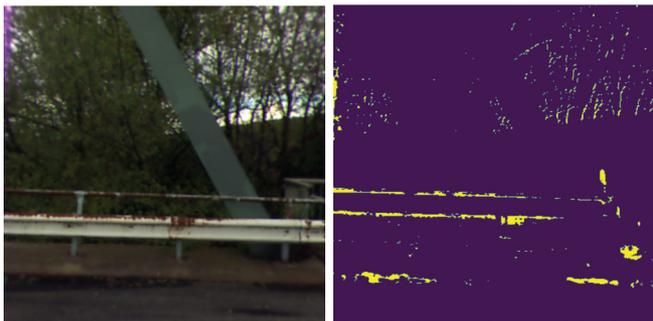


FIGURE 2. RGB-representation of an hyperspectral image of corroded areas on a bridge (left) and potential iron(III)-containing pixels marked by algorithm 1 (right)

The spatial resolution is increased via pansharpening, by using a 1886x1886 pixel greyscale sensor from a second camera mounted on top of the first.

The camera employs light field technology resulting in a minimum distance of 6 m to the AOI for matching pixel signals. The minimum distance can be adjusted by recalibration with a specific calibration target. The field of view is 35°. Data readout is done by global shutter with a maximum frame rate of 4 Hz and allowed integration times between 0.1 and 1000 ms. The maximum power consumption is 8 Watt.

### B. UAS

For an inspection mission with the hyperspectral camera introduced in chapter III-A a sufficient UAS has to be developed. Therefore, the requirements and the solution are addressed in the following section. Different aspects have to be considered for the concept of a UAS. Favourably, the UAS for an inspection mission should be as compact as possible to reduce the dimensions and therefore create a better accessibility regarding the infrastructure. Further, a long flight time is beneficial to reduce maintenance time and a stable, yet agile flight behaviour is necessary to ensure the reliable recording of hyperspectral images. However, the UAS has to be able to lift the inspection sensor with the associated gimbal and has to hold other sensors for navigation purposes as well. Finally, the UAS should be capable to record images at any angle such that the UAS can access a desired inspection point from above or below. Balancing this variety of different interests, we developed a UAS with the dimensions of 860x860x700 mm. The UAS is capable to lift 3 kg additional payload corresponding to the approximated 1.5 kg for an inspection sensor mounted on a gimbal and another 1.5 kg for a LIDAR, which is used for navigation purposes. The total weight adds up to about 12.5 kg, including a battery pack of 10 Ah ensuring a total flight time of at least 15 min. The UAS is designed with four arms and eight coaxial rotors in X-configuration as depicted in Figure 3. Also, the c-shaped construction around the centre of the UAS is shown, which holds the LIDAR and the camera gimbal on opposite sides. This unique part solves the mentioned issue of ensuring images to be taken from above and below. During an exchange

of the battery pack the construction can be rotated by 180° for taking images in the other direction. This construction ensures an overall compact design granting the benefits of the stable flight behaviour of a well-known multicopter. Further to mention, the UAS is equipped with a *Pixhawk 6X FCU*, a *Gigabyte BRIX* onboard pc and common telemetry- and GPS-modules. The onboard pc is used for sensor fusion and the control of the hyperspectral camera. The communication is based on ethernet and ROS2. The UAS's dimensions are used to create a collision model, which also takes the different mounting options into account. Therefore, the models' height differs with  $\pm 30$ cm. The models will be considered for path planning and collision avoidance purposes.



FIGURE 3. Rendering of the developed UAS with the C-shaped sensor-mounting

## IV. PATH PLANNING

The proposed UAS design has major influence on the mission planning that will be briefly discussed in the following. For more information about the used algorithms, the reader is encouraged to follow the upcoming publications of the authors. Note that the rotation of the c-shaped construction has to be considered during the process. The mission planning therefore creates two separate missions for points approached from above and below the UAS. Only one of them will be visualized in the results.

The general concept of the path planning is depicted in fig. 4. Starting from a point cloud recorded at the infrastructure or a BIM model, an algorithm performs an extraction of the most significant surfaces of the model. The precision can be customized and is also depending on the point cloud density or the overall quality of the underlying data. Considering the derived sensor model of the hyperspectral camera relevant inspection points are placed among the extracted surfaces with respect to a certain overlap between the individual images. With the help of the collision model of the UAS and the infrastructure model waypoints in the  $\mathbb{R}^3$  are created. An optimization algorithm takes into account hard constraints such as an unblocked line of sight and a collision-free spacing towards the structure but also factors like the viewing angle, the distance to the desired inspection point and lighting conditions. After the best fitting waypoints have been evaluated, they are sorted for separate missions depending on the mounting of

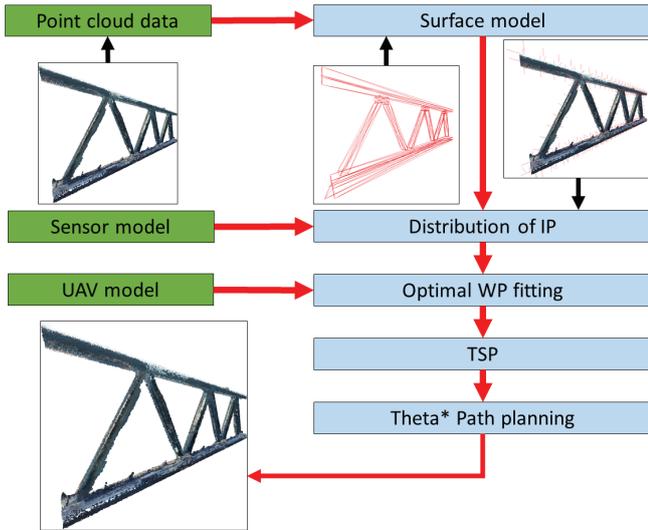


FIGURE 4. Visualization of the general workflow of the mission planning

the camera. The missions further are processed by a Travelling Salesman Problem (TSP) with an approximated Euclidean distance to optimize for path length and therefore required inspection time. Finally, a *Theta\** algorithm calculates an exact path for the sorted mission, which then is handed to the UAS. An exemplary 3D visualization of an inspection of three bridge elements extracted from the Travehafenbrücke located at the port of Hamburg, Germany, is shown in Figure 6. A 2D representation with respect to the  $x-y$ - and the  $x-z$ -plane is depicted in Figure 5. The mission is calculated for the introduced UAS with the hyperspectral camera mounted on the bottom. The red lines indicate the path and the dashed black lines roughly show the outlines of the bridge. The graph shows a meandering path in the  $x-y$ -plane to inspect the upper side of the bridge. Further, with respect to a certain safety distance, waypoints along the sides of the bridge are taken into account. The groups of points in the middle of the left tile are placed in the small areas inside the triangular structures, that allow the UAS to maintain the required safety distance. Those points can also be recognized in the  $x-z$ -plane depicted in the right tile of the figure. This tile however mainly shows the full vertical coverage of the bridge. Since the algorithm presented is restricted to movements along a grid, small adjustments of the path are necessary to move between the navigation grid and the exact waypoints for the inspection. Our current work focuses on improvements for a smoother path by implementing a cluster algorithm, that combines nearby waypoints into a single waypoint. Also, the current direction of movement will be taken into account to eliminate unnecessary changes of the paths' direction. However, the result shows an appropriate path for a UAS with respect to a predefined safety distance derived from the UAS model. The presented work can easily be adapted to other sensors such as RGB- or thermal cameras and other UAS configurations or collision models.

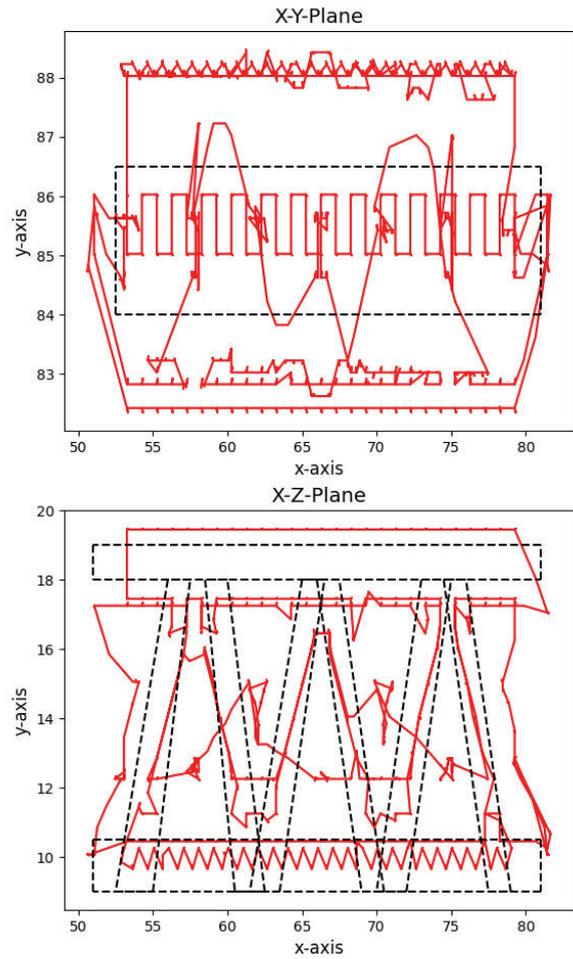


FIGURE 5. 2D Visualization of an inspection mission with the HSI camera mounted below in the X-Y-Plane (top) and X-Z-Plane (bottom)

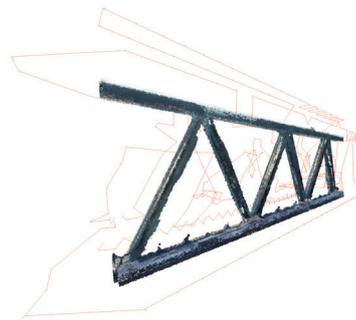


FIGURE 6. 3D Visualization of the inspection mission from fig. 5

## V. CONCLUSION

In this paper the authors present algorithms and the rudimentary concept of a UAS, that are capable to detect surfaces on a building and inspect those automatically with a mounted hyperspectral camera. Also an algorithm for the evaluation of the gathered data is presented, such that the detection of iron(III) from atmospheric steel corrosion according to ISO 4628-3 is possible. The proposed algorithm takes the increase of iron(III) and vegetation between 650 nm and 850 nm into consideration, while different illumination conditions between the calibration target and inspection area are neglected.

The current path planning is not optimized yet and will be improved in the future. Cluster algorithms and other optimizations will therefore be taken into account to achieve a much smoother and thus more reasonable path. However, the deployed system is already capable of reducing the closure times for inspection missions and reduce the costs and the effort of the necessary maintenance procedures. The authors work will focus on improvements and evaluation for the given algorithms. Test scenarios on real infrastructures with the presented system are planned for the near future.

## ACKNOWLEDGMENT

This research paper and the project MISDRO are funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

## REFERENCES

- [1] NACE IMPACT. “International Measures of Prevention, Application, and Economics of Corrosion Technology Study”. In: *NACE Int* (2016).
- [2] Iris Hindersmann and Stefan Staub. “Herausforderungen für ein Lebenszyklusmanagement für Bauwerke der Verkehrsinfrastruktur – Ergebnisse von leitfadengeführten Experteninterviews”. In: *Bautechnik* 99.11 (2022), pp. 817–824.
- [3] *Beschichtungsstoffe – Beleuchtung und Durchführung für visuelle Abmusterungen von Beschichtungen*. Norm. Feb. 2020.
- [4] Dominik Thomas et al. “Drohngestützte, multivariate Inspektionssysteme zur Zustandserfassung von Stahlbrücken und Stahlwasserbauten”. In: (2022).
- [5] RF Kokaly et al. “Usgs spectral library version 7 data: Us geological survey data release”. In: *United States Geological Survey (USGS): Reston, VA, USA* 61 (2017).
- [6] Dominik Thomas and Max Gündel. “Hyperspectral imaging systems for corrosion detection from remotely operated vehicles”. In: *ce/papers* 6.5 (2023), pp. 934–938.
- [7] Andrés Viña, Geoffrey M. Henebry, and Anatoly A. Gitelson. “Satellite monitoring of vegetation dynamics: Sensitivity enhancement by the wide dynamic range vegetation index”. In: *Geophysical Research Letters* 31.4 (2004).

# Echtzeit-Optimierung für die Steuerung von energetisch flexiblen Anlagenverbänden

Vincent Henkel, Lukas P. Wagner, Lasse M. Reinpold und Felix Gehlhoff

*Institut für Automatisierungstechnik*

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Hamburg, Deutschland

vincent.henkel@hsu-hh.de

Alexander Fay

*Lehrstuhl für Automatisierungstechnik*

Ruhr Universität Bochum

Bochum, Germany

alexander.fay@rub.de

Julian Jepsen

*Institut für angewandte Werkstofftechnik*

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Hamburg, Deutschland

jepsen@hsu-hh.de

**Zusammenfassung**—Die rasche Expansion erneuerbarer Energiequellen führt zu erheblicher Volatilität und Unvorhersehbarkeit in der Energieversorgungskette, was die Stabilität und Zuverlässigkeit des Stromnetzes herausfordert. Diese Arbeit präsentiert daher eine Methode, die bestehende statische Optimierungsmodele flexibler Energiequellen echtzeitfähig macht und damit ermöglicht, dass diese sowohl für die mittelfristige Planung als auch für Echtzeitanwendungen eingesetzt werden können. Durch die Anpassung statischer Modelle für die Echtzeitanwendung ermöglicht eine in diesem Beitrag vorgeschlagene zweistufige Optimierungsstrategie flexible Anpassungen von Betriebsplänen, wodurch erneuerbare Energien auch trotz unvorhergesehener Schwankungen bestmöglich genutzt werden können. Dieser Ansatz gewährleistet nicht nur die Zuverlässigkeit des Netzes, sondern verbessert auch die wirtschaftliche Effizienz durch die Optimierung der Ressourcennutzung. Die Wirksamkeit dieser Methode wird durch eine Fallstudie mit einem System von Elektrolyseuren demonstriert, die deutliche Vorteile gegenüber traditionellen statischen Optimierungsmethoden bei der Ausrichtung des Energieverbrauchs auf die Erzeugung erneuerbarer Energie aufzeigt.

**Stichwörter**—Energieflexibilität, Unsicherheiten, Zweistufige Optimierung

## I. EINLEITUNG

Die rasche Ausweitung von erneuerbaren Energien (EE) erhöht die Volatilität in der Energieversorgungskette erheblich, was fortschrittliche Regelungsstrategien zur Sicherstellung der Stabilität und Zuverlässigkeit des Netzes erforderlich macht [1]. Dadurch steigt die Bedeutung der Flexibilität von Energieressourcen, die die Fähigkeit einer Ressource darstellt, ihre Energieerzeugung oder ihren -verbrauch zu modifizieren [2].

Um flexible Energieressourcen möglichst optimal zu betreiben werden häufig numerische Optimierungsansätze verwendet. Bestehende Optimierungsansätze berücksichtigen dabei häufig keine Unsicherheiten [3][4], wie die von Prognosen der EE-Erzeugung, und fokussieren sich stattdessen auf eine wirtschaftliche Optimierung, bei der grobe zeitliche Auflösungen ausreichen [5]. Diese strategische Optimierung wird jedoch

meist unter erheblichen Unsicherheiten durchgeführt [6], was die Notwendigkeit feinerer zeitlicher Auflösungen des Optimierungsmodells zur Bewältigung sowohl der EE-Integration als auch des Unsicherheitsmanagements unterstreicht [7]. Hohe zeitliche Auflösungen gehen allerdings fast immer mit einer hohen numerischen Komplexität von Optimierungsmodellen einher, weshalb der Rechenaufwand für deren Lösung schnell unpraktikabel hoch wird [5].

Daher kann ein mehrschichtiger Optimierungsansatz, bekannt für seine Skalierbarkeit und Anpassungsfähigkeit, wie in Abb. 1 dargestellt, eingesetzt werden. Jede Schicht innerhalb der hierarchischen Struktur des Ansatzes verwaltet spezifische Entscheidungsvariablen und überwacht unterschiedliche Parameter und Variablen anderer Schichten. Diese hierarchische Organisation erleichtert die Abstraktion der Komplexität und ermöglicht die Bereitstellung von Diensten für höhere Schichten. Für die betriebliche Optimierung können diese Schichten genutzt werden, um verschiedene Optimierungsziele zu erreichen [8].

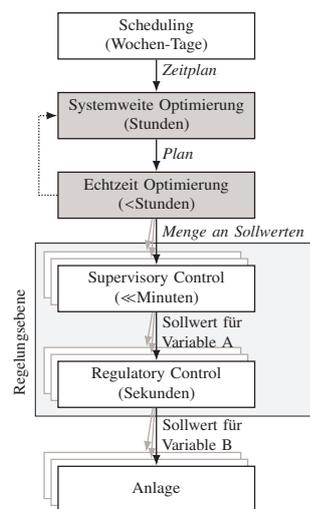


Abbildung 1: Hierarchische Optimierungsstruktur [8]

Ein zentrales Merkmal der hierarchischen Optimierungsstruktur ist die unterschiedliche zeitliche Auflösung jeder Schicht, die eine effektive Reaktion sowohl auf dynamische Veränderungen als auch auf langfristige Trends ermöglicht. Die *Scheduling*-Schicht, die mit der längsten zeitlichen Auflösung (von Wochen bis Tagen) arbeitet [8], erzeugt einen *Zeitplan*.

Auf Grundlage dieses *Zeitplans* ermittelt die *systemweite Optimierung* (SWO)-Schicht einen *Plan*. Die *Echtzeit-Optimierung* (RTO)-Schicht arbeitet mit einer mittleren Auflösung, typischerweise Stunden bis Minuten<sup>1</sup>, und stimmt die operative Umsetzung mit strategischen Vorgaben ab [8]. Sie definiert eine strukturierte *Menge an Sollwerten* (MaS), wobei der aktuelle Sollwert über die Regelungsebene zur Anlage übertragen wird.

Darüber hinaus bietet RTO, basierend auf dem Konzept der rollierenden Planung, einen vielversprechenden Ansatz zum Umgang mit Unsicherheiten. Während traditionelle strategische Optimierungsmethoden nicht darauf ausgelegt sind, auf kurzfristige Änderungen zu reagieren, erhöht RTO die Anpassungsfähigkeit bei Unsicherheiten und macht es so zu einem leistungsstarken Werkzeug für wirtschaftliche Optimierung und verbesserte Betriebsplanung [9].

Die meisten Optimierungsmodelle flexibler Energieressourcen benötigen Prognosen, etwa der EE-Erzeugung. Dadurch werden Prognosen in allen Schichten hierarchischer Optimierungsansätze verwendet. Die Unsicherheit und somit die Genauigkeit dieser Prognosen korrelieren mit dem Prognosehorizont [10]. Folglich hat der *Zeitplan* die höchste Unsicherheit, mit abnehmender Unsicherheit bis hin zur RTO-Schicht.

Die unterste Schicht agiert mit der kürzesten zeitlichen Auflösung und reagiert innerhalb von Minuten bis Sekunden auf unmittelbare Veränderungen [8]. In dieser Schicht kann der Anlagenverbund in einzelne Anlagen unterteilt werden, wobei jede Anlage ihre eigene Regelung hat.

Der vorherrschende Fokus auf statische Optimierung [7] führt oft dazu, dass die dargestellten Schichten isoliert betrachtet werden. Infolgedessen ignorieren viele Publikationen die Wechselwirkungen zwischen RTO und SWO (grau dargestellte Schichten in Abb. 1). Trotz seines Potenzials für erhebliche (energie-)wirtschaftliche Vorteile hat die Einführung von RTO in realen Anwendungen die Erwartungen nicht erfüllt, wodurch deren Potenzial weitgehend ungenutzt bleibt [6][9]. Ein möglicher Grund hierfür könnte den hohen Kosten zugeschrieben werden, die mit der Entwicklung und Implementierung von RTO-Lösungen verbunden sind [11].

Um das ungenutzte Potenzial auszuschöpfen und einen effizienteren sowie wirtschaftlicheren Anlagenbetrieb zu gewährleisten, untersucht diese Arbeit die Überführung bestehender statischer Optimierungsmodelle (z. B. aus SWO) in ein Modell zur gleichzeitigen Anwendung von SWO und RTO. Als Ergebnis entsteht ein zweistufiger Optimierungsansatz, der das gleiche Optimierungsmodell sowohl für die Langzeit- als auch für die Kurzzeitoptimierung nutzt, wobei

die SWO die Zielsetzung der RTO vorgibt (siehe Abb. 1). Die Anpassung der Automatisierungspyramide nach Skogestad [8] (wie in Abb. 1 gezeigt) wird in dieser Arbeit angewendet. Folglich ist der Ansatz in der Lage, auf Unsicherheiten und kurzfristige Abweichungen zu reagieren, und begegnet dem Problem, dass unterschiedliche Optimierungsmodelle über die Schichten hinweg zu Inkonsistenzen führen können [9]. Die Optimierungsziele auf der SWO- und der RTO-Ebene sind ähnlich, da beide Schichten darauf abzielen, den Nutzen zu maximieren, z. B. die Kosten zu minimieren [9]. Daher besteht der Hauptunterschied zwischen den beiden Schichten in ihrer zeitlichen Auflösung [9]. Kürzlich haben Reinpold et al. [4] gezeigt, dass geeignete Optimierungsmodelle für die SWO auch das reale Echtzeitverhalten einer Ressource genau darstellen können, wobei jedoch der Optimierungshorizont aufgrund des hohen Rechenaufwandes limitiert ist. Dadurch eignen sich solche Modelle für die doppelte Nutzung sowohl in der Langzeitplanung als auch in der Kurzzeitoptimierung.

Zusammenfassend sind die Beiträge dieser Arbeit die folgenden:

- eine Methode zur Anpassung bestehender statischer Optimierungsmodelle für ihre doppelte Verwendung in SWO und RTO,
- ein Algorithmus zur kontinuierlichen Lösung des zweistufigen Optimierungsproblems unter Einbeziehung aktualisierter Prognoseinformationen,
- eine Bewertung und Validierung der vorgeschlagenen Methode und Algorithmen durch eine Fallstudie.

Diese Arbeit ist wie folgt strukturiert: Abs. II bietet eine Analyse verwandter Arbeiten und beschreibt die Forschungslücke. Abs. III beschreibt die Methode zur kontinuierlichen Lösung des zweistufigen Optimierungsproblems einschließlich des Ansatzes zur Anpassung bestehender statischer Optimierungsmodelle. Abs. IV demonstriert die Validierung durch eine Fallstudie. Abs. V fasst diese Arbeit zusammen und beschreibt weiteren Forschungsbedarf.

## II. STAND DER WISSENSCHAFT

Dieser Abschnitt analysiert verwandte Arbeiten zur Optimierung flexibler Energieressourcen, wobei der Schwerpunkt auf deren Steuerung liegt, insbesondere im Umgang mit Unsicherheiten.

Sun und Leto [12] entwickelten ein Modell zur Optimierung von Geboten auf Energiemärkten für erneuerbare Quellen und flexible Anlagen, das darauf abzielt, Gewinne zu maximieren und gleichzeitig Risiken durch EE-Unsicherheiten zu mindern. Ihr Modell, das sich auf die Planung und SWO konzentriert, vernachlässigt RTO, das entscheidend für die Reaktion auf kurzfristige Schwankungen ist.

Vedullapalli et al. [13] stellten ein Planungsmodell für Batterien, Heizung, Lüftung und Klimatisierung vor, das RTO verwendet, um Energiekosten zu minimieren und sich auf die kurzfristige Planung konzentriert, ohne höhere Planungsfunktionen wie SWO. Dieser Ansatz unterstreicht die Notwendigkeit der Echtzeitsteuerung, jedoch fehlt die Integration langfristiger strategischer Planung zur Vermarktung von Flexibilität.

<sup>1</sup>Der Zeitmaßstab von RTO unterscheidet sich von kürzeren Zeitmaßstäben in anderen Echtzeit-Kontexten, wie etwa in der Regelungstechnik.

Tsay et al. [14] schlugen ein Rahmenwerk für den energieflexiblen Betrieb von industriellen Luftzerlegungsanlagen unter Verwendung von RTO vor. Ihre Arbeit betont die Notwendigkeit, Prozessdynamiken in der Produktionsoptimierung aufgrund der variablen Natur der Strompreise zu berücksichtigen, und behandelt dies mit einem speziell für Luftzerlegungsanlagen entwickelten dynamischen Optimierungsansatz. Obwohl in der Studie die Echtzeit-Strompreise erwähnt werden, werden sie nur mit einer stündlichen Auflösung berücksichtigt, was keine Reaktion auf kurzfristige Abweichungen zulässt.

Die Studie von Flamm et al. [15] präsentiert ein angepasstes Optimierungsmodell für einen Elektrolyseur. Dieses Modell speist einen RTO-Regler, der darauf abzielt, die Produktionskosten durch Anpassung an schwankende Strompreise und PV-Erzeugung zu minimieren [15]. Die Studie nutzt jedoch deterministische Prognosen, was Fragen zur Angemessenheit der Methode im Umgang mit Prognoseunsicherheiten aufwirft [15].

Die von Alabi et al. [16] und Alirezazadeh et al. [17] vorgeschlagenen Methoden behandeln Optimierung von Systemen flexibler Energieanlagen, konzentrieren sich jedoch hauptsächlich auf langfristige Planungsstrategien ohne Einbeziehung von RTO. Alabi et al. [16] präsentieren einen Ansatz, der die Markowitz-Portfolio-Risikothorie nutzt, um Unsicherheiten zu verwalten. Ähnlich ist die von Alirezazadeh et al. [17] vorgestellte Methode zur Optimierung flexibler Erzeugung auf lineare und nicht-lineare Programmierung für die Einsatzplanung von Anlagen beschränkt, bleibt jedoch auf statische SWO beschränkt, ohne die für kurzfristige Variationen erforderliche Anpassungsfähigkeit zu berücksichtigen.

Die Studie von Yang et al. [18] nutzt einen Ansatz zur Optimierung von Elektrolyseuren, der darauf abzielt, die Rentabilität der Wasserstoffproduktion unter Nutzung von Solar- und Netzstrom zu steigern. Sie berücksichtigt Faktoren wie die Verfügbarkeit von Solarenergie und Strompreise in ihrer langfristigen SWO-Strategie. Ähnlich wie die Arbeit von Alabi et al. [16] vernachlässigt die Arbeit von Yang et al. [18] die Notwendigkeit von RTO, um auf kurzfristige Schwankungen bei EE-Quellen zu reagieren.

Ireshika und Kepplinger [19] untersuchten das Management der Ladung von Elektrofahrzeugen unter Unsicherheit mit einem zweistufigen Optimierungsansatz, wobei sie sich auf Unsicherheiten wie nicht-elastische Nachfrage und das Nutzungsverhalten von Elektrofahrzeugen konzentrierten.

Dumas et al. [20] führten einen hierarchischen Optimierungsansatz für Netze ein, der sich auf die Koordination zwischen operativer Planung und RTO konzentriert. Auf der Ebene der operativen Planung werden Entscheidungen basierend auf Prognosen getroffen, um die Energiekosten über einen oder mehrere Tage zu optimieren. Die RTO-Ebene passt die Operationen basierend auf den tatsächlichen Bedingungen und Prognosefehlern innerhalb des aktuellen Marktzeitraums an.

Während Ireshika und Kepplinger [19] und Dumas et al. [20] jeweils die Vorteile eines zweistufigen Optimierungsansatzes zur Steuerung flexibler Energieanlagen unter Unsicherheit betonten. Anstatt bestehende Ansätze wiederzuverwenden, spe-

zialisieren sich beide Arbeiten auf die Anwendung spezifischer Optimierungsmodelle für bestimmte Probleme innerhalb ihrer Domänen und verwenden unterschiedliche Modelle für die verschiedenen Optimierungsebenen.

Die Analyse verwandter Arbeiten zeigt eine Lücke in der Integration von SWO und RTO in einen einheitlichen Rahmen, was auf einen Bedarf an generalisierten Methoden zur Weiterentwicklung hinweist. Bestehende Ansätze konzentrieren sich entweder auf SWO oder RTO, wobei Strategien zur Adressierung kurzfristiger Abweichungen selten diskutiert werden [7]. Dies unterstreicht die Notwendigkeit, über die Verwendung spezifischer Modelle für einzelne Anwendungsfälle hinaus Methoden zu entwickeln, die eine umfassende und wiederverwendbare Lösung für die dynamische Anpassung von Optimierungsstrategien bieten.

### III. METHODE ZUR LÖSUNG DES ZWEISTUFIGEN OPTIMIERUNGSMODELLS

Dieser Abschnitt beschreibt die Methode zur kontinuierlichen Lösung von SWO und RTO Modellen für die anschließende Steuerung von Systemen energetisch flexibler Anlagen. Das Konzept für einen zweistufigen Optimierungsansatz unter Verwendung bestehender Optimierungsmodelle wird in Abs. III-A umrissen. Die notwendigen Anpassungen, um ein bestehendes Optimierungsmodell in ein Modell zu verwandeln, das sowohl mit RTO als auch mit SWO kompatibel ist, werden in Abs. III-B erläutert. Der Ansatz zur kontinuierlichen Lösung des zweistufigen Optimierungsproblems wird in Abs. III-C dargestellt.

#### A. Konzept für die zweistufige Optimierung

Das Konzept, dargestellt in Abb. 2, nutzt SWO, um einen *Plan* für den gesamten Optimierungshorizont  $\mathcal{T}$ , z. B. einen ganzen Tag, aufgeteilt in Intervalle  $\Delta\tau$ , zu entwickeln. Individuelle Zeitpunkte werden als  $\tau$  bezeichnet. Um dies zu erreichen, werden Langzeitprognosedaten einbezogen, um sicherzustellen, dass der *Plan* die erwarteten zukünftigen Bedingungen oder Anforderungen widerspiegelt. Die Ergebnisse dieser Planungsphase setzen Startpunkte für RTO fest, d. h., sie definieren Startbedingungen wie die Zustände des Systems für jeden Zeitpunkt  $t_\tau$ . Dieser Ansatz wird zyklisch durchgeführt, wobei bei jeder Optimierung neue, aktualisierte Prognosen berücksichtigt werden.

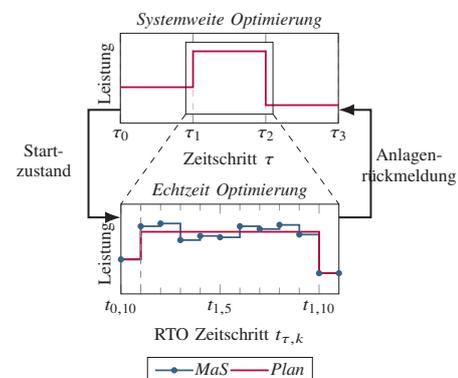


Abbildung 2: Schematische Darstellung des Konzepts

Nach dieser Anfangsphase verfeinert RTO die Optimierung mit einer höheren Auflösung  $\Delta t$ , etwa auf Minutenbasis, bezeichnet als Zeitschritte  $t_{\tau,k}$  über einen Optimierungshorizont  $T$ . Die Länge des RTO-Horizonts  $|T|$  entspricht einem Zeitschritt der SWO  $\Delta\tau$ . Diese feinere Granularität ermöglicht die Berücksichtigung unmittelbarer Schwankungen und gewährleistet, dass sich die Betriebsabläufe schnell an sich entwickelnde Szenarien anpassen können. Jegliche während der RTO-Phase festgestellten Änderungen, wie etwa der Ausfall von Anlagen, werden dann in die SWO zurückgespeist, um sicherzustellen, dass das Verhalten des realen Systems in den nachfolgenden Planungsperioden  $t_\tau$  der SWO widergespiegelt wird. Dieser zyklische Prozess schafft einen dynamischen und reaktionsfähigen Optimierungsrahmen, der langfristige strategische Planung nahtlos mit sofortigen betrieblichen Anpassungen integriert und dadurch sowohl die Effizienz als auch die Anpassungsfähigkeit erheblich verbessert.

### B. Zweifache Nutzung von Optimierungsmodellen

Um die Umwandlung statischer Optimierungsmodelle in dynamische RTO-Modelle zu erleichtern, müssen einige vorbereitende Schritte durchgeführt werden. Eine grundlegende Voraussetzung ist jedoch, dass ein statisches, lösbares Optimierungsmodell eines Systems flexibler Anlagen vorhanden ist. Das statische Optimierungsmodell erfordert Modifikationen, um eine Unterscheidung zwischen seiner Anwendung in SWO und RTO zu ermöglichen. Speziell für den Einsatz in RTO ist es notwendig, historische Werte von Variablen festzulegen, einschließlich der Zustände von Anlagen oder Startwerte für die RTO-Zeitschritte. Dies wird erreicht, indem die jeweilige Entscheidungsvariable dem entsprechenden Wert gleichgesetzt wird. Diese Werte werden aus den Ergebnissen der SWO oder aus den Ergebnissen vorheriger RTO-Zeitschritte abgeleitet. Weiterhin muss die für jeden Zeitschritt  $\tau$  beschaffte Energiemenge als Ziel für den entsprechenden RTO-Horizont ab  $t_{\tau,0}$  festgelegt werden. Dies wird durch die in Gl. 1 gezeigte Nebenbedingung sichergestellt, in der die Energie in einem SWO-Zeitschritt  $\tau$  gleich der Summe der Energie sein muss, die aus dem Netz in allen entsprechenden RTO-Zeitschritten  $t$  bezogen wird.

$$\sum_{t_{\tau,k} \in |T|} P_{\text{el, Netz}, t_{\tau,k}} \cdot \Delta t = P_{\text{el, Netz}, \tau} \cdot \Delta\tau \quad \forall \tau \quad (1)$$

Dies ist besonders bedeutend bei der Abrechnung von Regelarbeit, da die Kosten für diese Dienstleistungen den Parteien zugeordnet werden, die für Abweichungen vom *Plan* verantwortlich sind.

### C. Lösung des zweistufigen Optimierungsmodells

Der Prozess zur Lösung des zweistufigen Optimierungsmodells ist in Abb. 3 dargestellt. Zunächst werden die anfänglichen Parametereinstellungen definiert, wie der Systemzustand zu Beginn der statischen Optimierung (Schritt 1). Anschließend werden Langzeitprognoseinformationen, wie Stromdaten aus Spotmärkten oder EE-Erzeugung, importiert (Schritt 2) und verwendet, um das statische Optimierungsmodell für die

Zeitschritte  $[\tau_0, \tau_{|T|}]$  zu lösen (Schritt 3). Die Ergebnisse dieser betrieblichen Planung für den Zeitschritt  $\tau$  werden gespeichert (Schritt 4) und anschließend als Anfangswerte für das RTO-Modell übernommen ( $\tau_j = t_{\tau_j,0}$ , Schritt 5).

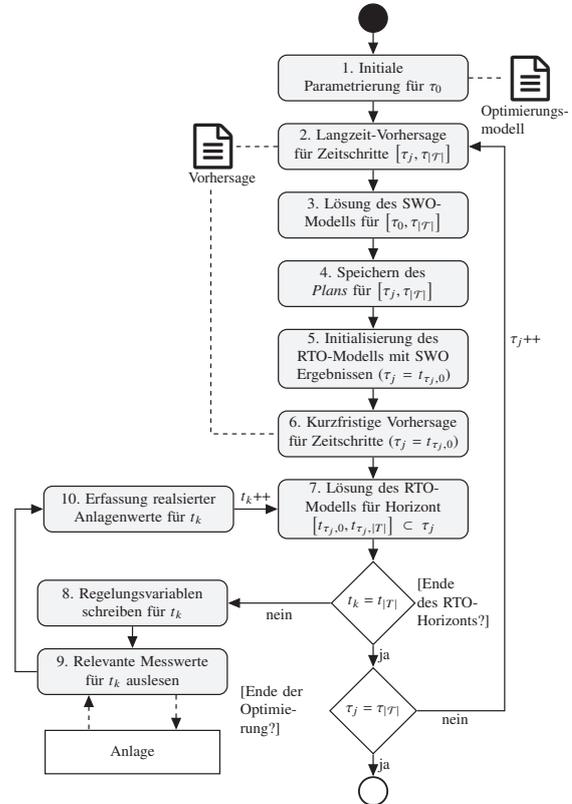


Abbildung 3: Flussdiagramm für die Verwendung eines bestehenden Optimierungsmodells in RTO

Um angemessene Reaktionen zu ermöglichen, werden kurzfristige Prognosen mit hoher Auflösung, wie die von EE-Erzeugung, eingebunden (Schritt 6). Anschließend wird das RTO-Modell basierend auf dem festgelegten RTO-Horizont  $T$  gelöst, was geeignete Anpassungen an kurzfristige Schwankungen ermöglicht (Schritt 7). Danach prüft das System, ob das Ende des RTO-Horizonts  $t_k = t_{|T|}$  erreicht wurde. Wenn nicht, werden die optimierten Sollwerte an den Prozess übermittelt (Schritt 8), und relevante Messwerte werden empfangen (Schritt 9). Diese gemessenen Werte können dann während RTO berücksichtigt werden (Schritt 10). Die aktuellen Sollwerte dienen dann als neue Startwerte für die nächste RTO-Iteration (Schritt 10).

Beim Erreichen des Endes des RTO-Horizonts mit  $t_k = t_{|T|}$  prüft das System auch, ob das Ende des Planungsintervalls  $\mathcal{T}$  erreicht wurde. Wenn ja, endet der Prozess. Dieser Prozess wird jedoch typischerweise kontinuierlich ausgeführt, was bedeutet, dass in den meisten Fällen kein Ende erreicht wird. Stattdessen werden neue Startwerte für den nachfolgenden Planungszeitschritt  $\tau$  aus der statischen Optimierung gewonnen, und der RTO-Prozess beginnt erneut (Schritt 5). Beide Stufen werden immer für die gesamte Länge des jeweiligen

Optimierungshorizonts  $\mathcal{T}/T$  gelöst, wobei neu verfügbare Informationen, wie Prognosen für zukünftige Zeitschritte und historische Werte, einbezogen werden. Daher wird jedes Modell  $x$ -Mal gelöst, wobei  $x$  entsprechend  $\frac{\mathcal{T}}{\Delta\tau}$  oder  $\frac{T}{\Delta\tau}$  für SWO bzw. RTO steht.

Die Anwendung der in diesem Abschnitt beschriebenen Methode erzeugt *Pläne* für die Steuerung eines Systems von Ressourcen. Bei Verfügbarkeit neuer Daten, wie Prognosen, wird eine aktualisierte MaS oder ein *Plan* für die verbleibenden Zeitschritte des jeweiligen Optimierungshorizonts generiert.

#### IV. VALIDIERUNG

Die Fallstudie demonstriert die Anwendbarkeit der vorgeschlagenen Methode, indem sie auf ein System von Elektrolyseuren angewendet wird, die sowohl Strom aus dem Netz als auch von einem Windpark beziehen. Zu diesem Zweck beschreibt dieser Abschnitt die Einrichtung der Validierung (Abs. IV-A) sowie die Ergebnisse (Abs. IV-B & Abs. IV-C).

##### A. Setup der Validierung

Das zur Validierung der Methode verwendete Optimierungsmodell basiert auf der von Wagner et al. [21] entwickelten und validierten Optimierungsmodellstruktur (siehe Optimierungsmodell 1). Dabei wurde eine Menge von Nebenbedingungen zur Darstellung von Flexibilitätsmerkmalen entwickelt [21]. Dies umfasst Nebenbedingungen für Betriebsgrenzen (min./max. Werte für Flüsse), Input-Output-Beziehungen sowie Systemzustände und zugehörige Nebenbedingungen [21]. Die Parameter für jeden Elektrolyseur wurden unter Verwendung der von Wagner und Fay [22] entwickelten Methodik zur Parametrisierung von Optimierungsmodellen bestimmt. Dies generiert ein Datenmodell für die Parametrisierung des Optimierungsmodells, basierend auf Betriebsdaten [22].

**min** Energiekosten (Intraday-Markt)  
**subject to**

- Energiebilanz unter Einbeziehung von Netzstrom und Strom aus erneuerbaren Energien und Anschluss von Elektrolyseuren
- Betriebsgrenzen des Systems
- Betriebsgrenzen jeder Ressource
- Input-Output-Beziehungen
- Ziel für Energie-Input
- Systemzustände
- Folgezustände
- Haltedauern
- Ramp

Optimierungsmodell 1: Modell für Fallstudie, basierend auf von Wagner et al. [21] entwickelter Modellstruktur

Optimierungsmodell 2 zeigt, wie das bestehende, durchführbare Optimierungsmodell 1 für den Einsatz in RTO erweitert wird, wie in Abs. III-B beschrieben.

Für Prognosen zukünftiger EE-Erzeugung verwendet diese Arbeit Daten eines echten Windparks, die von Anvari et al. [10] mit einer Auflösung von 1 Hz veröffentlicht wurden. Aus diesem Datensatz wurden sowohl  $|\mathcal{T}|$  Langzeitprognosen (siehe Abb. 4a) als auch  $|T|$  entsprechende Kurzzeitprognosen für jeden Zeitschritt  $\tau$  (siehe Abb. 4b) generiert.

**max** Wasserstoff-Energiemenge **subject to**

- Nebenbedingungen von Optimierungsmodell 1
- Ziel zur Vermeidung von Strafkosten (Gl. 1)
- Fixierte Werte vergangener Perioden

Optimierungsmodell 2: Weitere Nebenbedingungen für RTO

Diese Prognosen wurden iterativ zu jedem Zeitschritt erstellt, wobei Datenpunkte für vergangene Zeitschritte beibehalten wurden. Um das Problem der Vorhersageunsicherheiten anzugehen, wurde eine größere Unsicherheit für weiter in der Zukunft liegende Zeitschritte eingeführt. Dies lässt sich beispielhaft an den Abweichungen der Werte zum Zeitschritt  $\tau_1$  der Prognosen, die jeweils bei  $\tau_0$  und  $\tau_1$  erstellt wurden, in Abb. 4a (grauer Kreis) erkennen. Dort hat die bei Zeitschritt  $\tau_0$  (blau) erstellte Prognose die Erzeugung erneuerbarer Energien bei  $\tau_1$  überschätzt, was eine Korrektur der Prognose für  $\tau_1$  (rot) erforderlich machte. Eine ähnliche Interpretation kann auf die Kurzzeitprognosen in Abb. 4b angewendet werden.

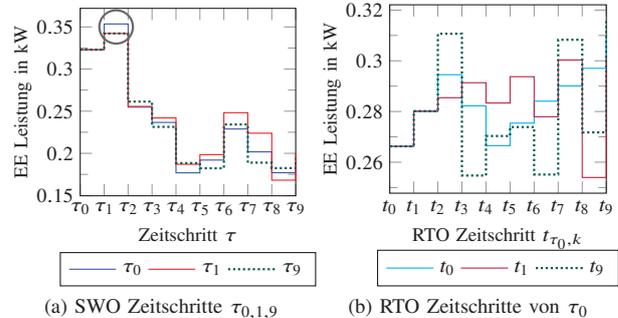


Abbildung 4: Vorhersagen und deren Unsicherheiten (Daten von Anvari et al. [10])

Die Methode wurde in Java unter Verwendung eines auf IBM ILOG CPLEX basierenden Optimierungsmodells implementiert<sup>2</sup>.

##### B. Ergebnisse der Validierung

Die in Abs. III beschriebene Methode wird unter Verwendung der Optimierungsmodelle 1 und 2 angewendet, wobei alle Modelle wie in Abb. 3 dargestellt gelöst werden. Dabei werden Preise verwendet, die vom europäischen Intraday-Markt bezogen werden [23], wie in Abb. 5 illustriert.

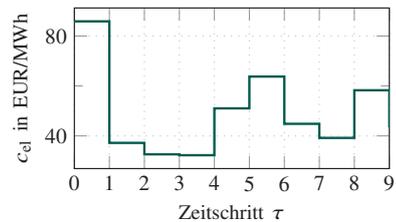


Abbildung 5: Preis für elektrische Energie [23]

<sup>2</sup>Implementation: <https://github.com/lukas-wagner/TwoStageOpt>

Die Berechnungen wurden auf einem Windows 10-System mit einem Intel Core i7-11700 Prozessor und 16 GB RAM mit einer Optimalitätslücke von  $10^{-3}$  durchgeführt. Die Methode für den zweistufigen Optimierungsansatz wurde für  $|\mathcal{T}| = 10$  SWO-Zeitschritte und jeweils  $|T| = 10$  RTO-Zeitschritte angewendet. Die Berechnungszeit für jeden Zeitschritt  $\tau$  einschließlich der entsprechenden RTO-Zeitschritte  $t$  betrug insgesamt etwa 10 Sekunden, wobei zeitliche Auflösungen von 0,25 Stunden für SWO und 0,025 Stunden für RTO verwendet wurden.

Darüber hinaus wird die MaS, die durch die RTO generiert wurde, an ein Simulationsmodell von Elektrolyseuren übertragen, wie in [24] beschrieben, unter Verwendung von OPC-UA (Schritte 9 & 10 in Abb. 3). Zu diesem Zweck wird die von Reinbold et al. [4] entwickelte Methode verwendet. Dieser Prozess wurde erfolgreich durchgeführt, und die aufgezeichneten Werte werden nachfolgend analysiert.

1) *Ergebnisse der systemweiten Optimierung:*

Abb. 6 zeigt den durch das SWO erstellten *Plan*, der die in Abb. 4a dargestellten Prognosen berücksichtigt. Zur Verdeutlichung werden nur der im ersten Zeitschritt  $\tau_0$  optimierte *Plan* und der endgültige, realisierte MaS, der unter Berücksichtigung aller vergangenen Entscheidungen bei  $\tau_9$  generiert wurde, dargestellt.

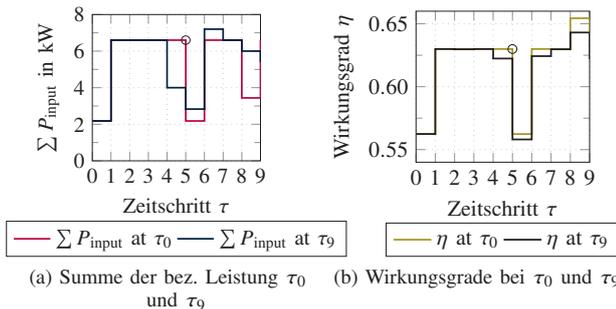


Abbildung 6: Ergebnisse der SWO

Abb. 6a zeigt, dass der bei  $\tau_0$  erstellte *Plan* mit dem realisierten *Plan* ( $\tau_9$ ) bis einschließlich Zeitschritt  $\tau_4$  übereinstimmt. Dieser anfängliche *Plan* wird unter der Annahme einer perfekten Voraussicht erstellt und profitiert vom größtmöglichen Flexibilitätspotenzial.

Ab  $\tau_5$  ist bereits ein erheblicher Anteil der Zeitschritte ausgeführt worden, was die Fähigkeit einschränkt, auf aktualisierte Prognosedaten während der verbleibenden Zeitschritte von SWO zu reagieren, aufgrund der auferlegten Einschränkungen. Folglich führt dies zu Abweichungen zwischen dem bei  $\tau_0$  optimierten *Plan* und dem realisierten *Plan* ab diesem Zeitpunkt, da das Flexibilitätspotenzial abnimmt. Diese Einschränkung spiegelt sich in einem Rückgang der Wasserstoffproduktion um 0,5 % zwischen dem anfänglichen Stadium bei  $\tau_0$  und dem letzten Schritt  $\tau_9$  wider. Der Hauptgrund dafür ist die Unsicherheit der Prognosen, wobei 1,5 % weniger EE-Erzeugung realisiert wurden als ursprünglich vorhergesagt. Dies führt auch zu einer geringeren Gesamteffizienz von

62,3 % im Vergleich zum anfänglichen *Plan* von 62,6 % bei  $\tau_0$ . Der Vergleich der Effizienzen bei  $\tau_0$  und  $\tau_9$  wird in Abb. 6b dargestellt und veranschaulicht die Auswirkungen der genannten Unsicherheiten auf die Effizienz des Systems.

2) *Ergebnisse der Echtzeit-Optimierung:* Abb. 7 zeigt die Ergebnisse von RTO für einen exemplarischen Zeitschritt  $\tau_4$ . In Abb. 7a werden die Prognose für die EE-Erzeugung unter dem SWO für den Zeitschritt  $\tau_4$  sowie die abweichende Prognose bei  $t_{\tau_4,9}$  dargestellt. Zu erwähnen ist, dass die Erzeugung von EE die langfristige Vorhersage des SWO konstant übertrifft. Dies führt folglich zu einer Erhöhung der Wasserstoffproduktion. Das realisierte MaS bei  $t_{\tau_4,9}$  ergibt 0,66 kWh Wasserstoffenergie anstelle der geplanten Menge von 0,62 kWh bei  $\tau_4$ , was die erhöhte Verfügbarkeit von EE nutzt (siehe Abb. 7b).

Weiterhin, wie in Abb. 7b gezeigt, kann ein signifikanter Rückgang der Leistung bei  $t_7$  beobachtet werden. Dies wird auf den Anstieg der Prognose für EE von  $t_8$  bis  $t_9$  im Vergleich zu früheren Prognosen zurückgeführt (siehe  $t_{\tau_4,7}$ ). Infolgedessen wurde die Wasserstoffproduktion auf spätere Zeitschritte  $t_8$  und  $t_9$  verschoben, was zu einer Verringerung der Leistungsaufnahme bei  $t_7$  führte.

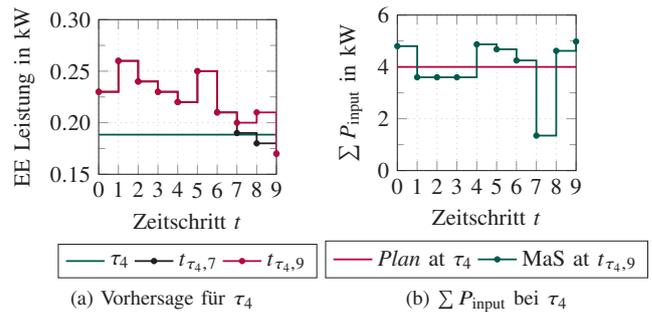


Abbildung 7: RTO MaS bei  $\tau_4$

C. *Validierung der Methode zur zweistufigen Optimierung*

Abb. 8 zeigt die realisierte Wasserstoffproduktion in jedem Zeitschritt für das SWO sowie die Summe aller Wasserstoffmengen im letzten RTO-Zeitschritt jedes SWO-Zeitschritts  $t_{\tau_j,9}$ . Dabei zeigt sich, dass durch die SWO bereits robuste Ergebnisse erzielt wurden, mit einer durchschnittlichen Abweichung von +2 % zwischen dem optimierten *Plan* und dem realisierten MaS über alle Zeitschritte  $\tau$ , die einen Bereich von -3,5 % bis +15 % umfassen. Obwohl die SWO im Allgemeinen günstige Ergebnisse liefert, unterstreichen Abweichungen, die +15 % in Perioden hoher EE-Unsicherheit überschreiten, die Bedeutung der RTO. Diese Abweichungen verdeutlichen, dass die RTO entscheidend ist, um auf kurzfristige Schwankungen und Unsicherheiten in der Verfügbarkeit von erneuerbaren Energien zu reagieren, wodurch die Effizienz und Zuverlässigkeit des Gesamtsystems verbessert wird.

Die inhärenten Unsicherheiten bei EE, die sich weiterer Quantifizierung widersetzen und von der Qualität der Prognose abhängen, können effektiv durch RTO adressiert werden, was ein wertvolles Instrument für angemessene Reaktionen bietet.

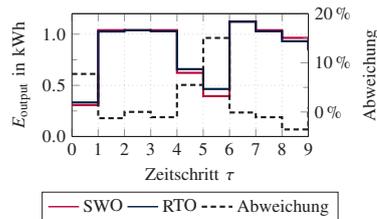


Abbildung 8: Optimierungsergebnisse für den ganzen Horizont

Der in Abs. III dargestellte Ansatz zur Implementierung einer zweistufigen Optimierungsstrategie ist besonders vorteilhaft für hybride Energiequellen, die sowohl EE-Quellen als auch das Stromnetz integrieren. Zunächst bestimmt die SWO die erforderliche Energiebeschaffung vom Energiemarkt mit überschaubarem Rechenaufwand und einem langfristigen Planungshorizont. Die zweite Optimierungsstufe, RTO, nutzt dann die vorher festgelegte Energiebeschaffung aus dem Energiemarkt und die durch SWO ermittelten Ressourcenzustände, um auf kurzfristige Schwankungen oder Abweichungen innerhalb optimierter, definierter Rahmenbedingungen zu reagieren.

Die Nutzung eines einzigen Optimierungsmodells hätte nicht für kontinuierlich aktualisierte Prognosen gesorgt, da dies die kontinuierliche Lösung des Modells für den gesamten Horizont gleich  $\tau$  bei Verfügbarkeit neuer Prognosen erfordert hätte. Die Rechenzeiten eines Modells bei hoher zeitlicher Auflösung, z. B. 1,5 Minuten, wären zu groß gewesen, um nützlich zu sein (zeitgerechte Verfügbarkeit des Plans) [5]. Ein solcher Ansatz ist auch angesichts der hohen Unsicherheit von Prognosen für die ferne Zukunft ungeeignet, was die Auflösung von 1,5 Min. ungerechtfertigt erscheinen lässt und die Anwendung des vorgestellten zweistufigen Optimierungsansatzes unterstreicht.

Zusammenfassend zeigen die Ergebnisse die Wirksamkeit der in Abs. III vorgestellten Methode und bestätigen, dass der zweistufige Optimierungsansatz erfolgreich auf der Basis eines bestehenden statischen Optimierungsmodells angewendet werden kann. Dieser Ansatz optimiert die Beschaffung von Energie aus einem Spotmarkt parallel und erleichtert gleichzeitig die Integration von variablen EE-Quellen. Dies ist besonders wichtig angesichts des zunehmenden Ausbaus von EE.

Darüber hinaus wird deutlich, dass bestehende Optimierungsmodelle (SWO) für diesen zweistufigen Optimierungsansatz angepasst werden können, ohne signifikante Änderungen zu erfordern (siehe Abs. III-B). Dies erleichtert den Übergang von reiner Planung zum Echtzeitbetrieb, in dem kurzfristige Schwankungen berücksichtigt werden.

## V. FAZIT

In dieser Arbeit wurde eine neuartige zweistufige Optimierungsmethode entwickelt, die die Lücke zwischen SWO und RTO überbrückt und dynamische Anpassungen an die Volatilität von EE-Quellen ermöglicht. Durch die Nutzung bestehender statischer Optimierungsmodelle und deren Anpassung für den gleichzeitigen Einsatz als RTO-Modelle wird die Integration von EE erleichtert, was sowohl die Stabilität des Stromnetzes als auch die betriebliche Effizienz des Stromnetzes und einzelner flexibler Energieressourcen verbessert. Diese

Methode zeichnet sich durch ihre Reaktionsfähigkeit auf unvorhergesehene Ereignisse aus, indem sie eine optimierte Ressourcennutzung in nahezu Echtzeit gewährleistet. In der durchgeführten Fallstudie konnte so gezeigt werden, dass die Strombeschaffungskosten an der Intraday-Strombörse verringert werden können.

Die Fallstudie an einem System von Elektrolyseuren, das eine hybride Energieversorgung sowohl aus dem Netz als auch von einem Windpark nutzt, validierte die Wirksamkeit der Methode. Die Ergebnisse zeigten nicht nur einen nachhaltigeren Betrieb durch die optimierte Nutzung erneuerbarer Ressourcen, sondern hoben auch das Potenzial für wirtschaftliche Vorteile hervor, da aktualisierte Prognosen in die betriebliche Planung einbezogen werden können. Durch die Feinabstimmung des Energieverbrauchs an die Verfügbarkeit von Windenergie wurden eine signifikante Verbesserung der Systemeffizienz und eine Reduzierung der Abhängigkeit von nicht-EE-Quellen beobachtet.

Zusammenfassend bietet die entwickelte zweistufige Optimierungsstrategie ein robustes und flexibles Rahmenwerk für das Energiemanagement und ebnet den Weg für eine nachhaltigere und wirtschaftlich rentable Regelung flexibler Energieanlagen angesichts des wachsenden EE-Einflusses.

Zukünftige Arbeiten könnten sich auf die Anwendung des Ansatzes auf reale Anlagen fokussieren. Darüber hinaus scheint die Anwendung des Ansatzes im Rahmen eines Multi-Agenten-Systems als zielführend, da diese von Natur aus geeignet sind, Ressourcen unter Unsicherheit darzustellen und zu steuern [25].

## DANKSAGUNG

Diese Forschungsarbeit aus dem Projekt OptiFlex wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

Zudem wurde diese Forschung im Projekt eModule (Förderkennzeichen: 03HY116) vom Bundesministerium für Bildung und Forschung auf Beschluss des Deutschen Bundestages gefördert.

## LITERATUR

- [1] O. Smith, O. Cattell, E. Farcot, R. D. O’Dea, und K. I. Hopcraft, “The effect of renewable energy incorporation on power grid stability and resilience,” *Science advances*, 2022.
- [2] A. Ulbig und G. Andersson, “On operational flexibility in power systems,” in *2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012, pp. 1–8.
- [3] M. Lindner, B. Grosch, G. Elserafi, B. Dietrich, und M. Weigold, “Holistic Approach for an Energy-Flexible Operation of a Machine Tool with Cooling Supply,” *Energies*, vol. 16, no. 9, p. 3943, 2023.
- [4] L. M. Reinhold, L. P. Wagner, L.-T. Reiche, und A. Fay, “Experimental Setup for the Evaluation of Optimization Strategies for Flexible Energy Resources,” in *2nd IEEE Industrial Electronics Society Annual Online Conference (ONCON)*, 2023.
- [5] L. P. Wagner, M. Kiltath, L. M. Reinhold, und A. Fay, “Required Level of Detail of Optimization Models for the Control of Flexible Energy Resources,” in *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids*, Glasgow, UK, 2023.
- [6] S. Engell und I. Harjunkoski, “Optimal operation: Scheduling, advanced control and their integration,” *Computers & Chemical Engineering*, vol. 47, pp. 121–133, 2012.

- [7] L. P. Wagner, L. M. Reinpold, M. Kiltthau, und A. Fay, "A Systematic Review of Modeling Approaches for Flexible Energy Resources," *Renewable and Sustainable Energy Reviews*, vol. 184, p. 113541, 2023.
- [8] S. Skogestad, "Advanced control using decomposition and simple elements," *Annual Reviews in Control*, vol. 56, p. 100903, 2023.
- [9] D. Krishnamoorthy und S. Skogestad, "Real-time optimization as a feedback control problem – a review," *Computers & Chemical Engineering*, vol. 161, p. 107723, 2022.
- [10] M. Anvari, G. Lohmann, M. Wächter, P. Milan, E. Lorenz, D. Heinemann, M. R. R. Tabar, und J. Peinke, "Short term fluctuations of wind and solar power systems," *New Journal of Physics*, 2016.
- [11] M. Bauer und I. K. Craig, "Economic assessment of advanced process control – a survey and framework," *Journal of Process Control*, vol. 18, no. 1, pp. 2–18, 2008.
- [12] C. Sun und S. Leto, "A novel joint bidding technique for fuel cell wind turbine photovoltaic storage unit and demand response considering prediction models analysis Effect's," *International Journal of Hydrogen Energy*, vol. 45, no. 11, pp. 6823–6837, 2020.
- [13] D. T. Vedullapalli, R. Hadidi, und B. Schroeder, "Optimal Demand Response in a building by Battery and HVAC scheduling using Model Predictive Control," in *2019 IEEE/IAS 55th I&CPS*, Calgary, 2019.
- [14] C. Tsay, A. Kumar, J. Flores-Cerrillo, und M. Baldea, "Optimal demand response scheduling of an industrial air separation unit using data-driven dynamic models," *Computers & Chemical Engineering*, pp. 22–34, 2019.
- [15] B. Flamm, C. Peter, F. N. Büchi, und J. Lygeros, "Electrolyzer modeling and real-time control for optimized production of hydrogen gas," *Applied Energy*, vol. 281, p. 116031, 2021.
- [16] T. M. Alabi, L. Lu, und Z. Yang, "A novel multi-objective stochastic risk co-optimization model of a zero-carbon multi-energy system (ZCMES) incorporating energy storage aging model and integrated demand response," *Energy*, vol. 226, p. 120258, 2021.
- [17] A. Alirezazadeh, M. Rashidinejad, A. Abdollahi, P. Afzali, und A. Bakhsai, "A new flexible model for generation scheduling in a smart grid," *Energy*, 2020.
- [18] Y. Yang, B. de La Torre, K. Stewart, L. Lair, N. L. Phan, R. Das, D. Gonzalez, und R. C. Lo, "The scheduling of alkaline water electrolysis for hydrogen production using hybrid energy sources," *Energy Conversion and Management*, vol. 257, p. 115408, 2022.
- [19] M. A. S. T. Ireshika und P. Kepplinger, "Uncertainties in model predictive control for decentralized autonomous demand side management of electric vehicles," *Journal of Energy Storage*, vol. 83, p. 110194, 2024.
- [20] J. Dumas, S. Dakir, C. Liu, und B. Cornélusse, "Coordination of operational planning and real-time optimization in microgrids," *Electric Power Systems Research*, vol. 190, p. 106634, 2021.
- [21] L. P. Wagner, L. M. Reinpold, und A. Fay, "Design Patterns for Optimization Models of Flexible Energy Resources," in *2nd IEEE Industrial Electronics Society Annual Online Conference (ONCON)*, online, 2023.
- [22] L. P. Wagner und A. Fay, "Methodology for Deriving Parameters for Optimization Models of Systems of Flexible Energy Resources," *IEEE Open Journal of the Industrial Electronics Society*, 2024, submitted.
- [23] "Continuous intra-day prices for Germany." [Online]. Available: [www.epexspot.com](http://www.epexspot.com)
- [24] V. Henkel, M. Kiltthau, F. Gehlhoff, L. Wagner, und A. Fay, "Cost optimized scheduling in modular electrolysis plants," in *2024 International Conference on Information Technology (ICIT)*. Bristol, UK: IEEE, 2024.
- [25] L. M. Reinpold, L. P. Wagner, F. Gehlhoff, M. Ramonat, M. Kiltthau, M. S. Gill, J. T. Reif, V. Henkel, L. Scholz, und A. Fay, "Systematic comparison of software agents and Digital Twins: differences, similarities, and synergies in industrial production," *Journal of Intelligent Manufacturing*, 2024.

# Chatbot-basierte Interaktion mit Ontologien: Nutzung von Large Language Models und domänenspezifischen Standards

Milapji Singh Gill, Jonathan Reif, Tom Jeleniewski, Felix Gehlhoff  
*Institut für Automatisierungstechnik*  
*Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg*  
Hamburg, Deutschland  
{milapji.gill, jonathan.reif, tom.jeleniewski, felix.gehlhoff}@hsu-hh.de

Alexander Fay  
*Lehrstuhl für Automatisierungstechnik*  
*Ruhr-Universität Bochum*  
Bochum, Deutschland  
alexander.fay@rub.de

**Zusammenfassung**—Im vorliegenden Beitrag wird ein Konzept vorgestellt, das Large Language Models (LLMs) und eine Chatbot-Benutzeroberfläche nutzt, um eine intuitive Interaktion zwischen Anwendern und Ontologien zu ermöglichen. Eingaben in natürlicher Sprache werden dabei in SPARQL-Abfragen umgewandelt, um das Faktenwissen der Ontologie abzufragen und das Risiko von Fehlinformationen zu minimieren. Zur Steigerung der Ergebnisqualität wird die Ontologie mithilfe von etablierten, domänenspezifischen Standards um zusätzliche Kontextinformationen zu modellierten Klassen und Relationen erweitert. Eine experimentelle Untersuchung wurde durchgeführt, um die Genauigkeit der mittels LLMs generierten SPARQL-Abfragen zu ermitteln. Die vorläufigen Ergebnisse zeigen den Mehrwert des Konzepts für die Ontologieabfrage.

**Index Terms**—Semantic Web, Ontologien, Large Language Model, Cyber-physische Systeme, Industrie 4.0

## I. EINFÜHRUNG

Ontologien können im Kontext der Industrie 4.0 und insbesondere bei dynamisch vernetzten *Cyber-physischen Systemen* (CPS) helfen, die stetig steigende Komplexität zu bewältigen. Einerseits sind sie geeignete formale Beschreibungsmittel, um die Datenintegration zu erleichtern. Andererseits wird durch die Etablierung einer eindeutigen Semantik die Interoperabilität zwischen CPS ermöglicht [1]. Diese Eigenschaften sind insbesondere mit Blick auf industrielle Anwendungen von großem Nutzen, da notwendige Daten meist aus diversen heterogenen Quellen stammen, jedoch für verschiedenste Anwendungsgebiete der Industrie 4.0 miteinander verknüpft werden müssen. Darüber hinaus sind Ontologien auch bei der Entwicklung, dem Betrieb und der Instandhaltung von CPS von großer Relevanz. Ihr Einsatz in Assistenzsystemen als Mittel zur Formalisierung von Domänenwissen bildet die Grundlage für eine verbesserte Entscheidungsfindung [2, 3, 4].

Diesen großen Vorteilen stehen jedoch ebenso Herausforderungen gegenüber. Aufgrund ihrer inhärenten Komplexität sind Ontologien schwer verständlich und für den Anwender nicht einfach zu handhaben [5]. Traditionell erfolgt die Wissensabfrage auf Basis vordefinierter *Competency Questions* (CQs) und statischer *SPARQL Protocol and RDF Query Language* (SPARQL)-Abfragen [1], die für Anwender ohne Se-

mantic Web-Expertise weder intuitiv noch benutzerfreundlich sind. Diese Problematik wird durch die geringe Anzahl an Ontologie-Experten im industriellen Umfeld verstärkt [1]. Diese Umstände führen dazu, dass die Flexibilität und Effizienz der Wissensabfrage erheblich eingeschränkt werden.

In diesem Zusammenhang kann die Interaktion zwischen einer domänenspezifischen Ontologie und einem Anwender durch den Einsatz von Chatbots und *Large Language Models* (LLMs) deutlich verbessert werden [5, 6]. Derzeit gewinnen auf LLMs basierende Assistenzsysteme, angetrieben durch die neuesten Errungenschaften in der aktuellen Forschung, in verschiedensten Domänen an Bedeutung. Die ausschließliche Abhängigkeit von auf LLMs basierenden Ansätzen birgt allerdings erhebliche Risiken. Aufgrund der fehlenden Nachvollziehbarkeit von generierten Antworten und der kreativen Interpretationsfähigkeiten sowie Halluzinationen dieser Modelle kann die Glaubwürdigkeit der wahrscheinkeitsbasierten Antworten nicht sichergestellt werden. Folglich könnten falsche Informationen an den Nutzer übermittelt werden. Dies stellt insbesondere in industriellen Anwendungen, bei denen Falschinformationen bzw. halluzinierte Informationen erhebliche wirtschaftliche und sicherheitstechnische Folgen haben könnten, eine außerordentliche Bedrohung dar.

Insofern bietet es sich an, die Vorteile von Ontologien, insbesondere die formale, strukturierte Bereitstellung von Faktenwissen, mit denen von LLMs, die eine intuitive Anwenderschnittstelle ermöglichen, zu kombinieren. Im Folgenden wird ein Konzept vorgeschlagen, das den Prozess der automatisierten SPARQL-Abfragegenerierung durch die Nutzung von LLMs und Informationen aus domänenspezifischen Standards verbessern soll. Dieses Konzept zielt darauf ab, die Benutzerfreundlichkeit durch Bereitstellung einer intuitiven Benutzeroberfläche für die Interaktion mit komplexen Ontologien zu erhöhen.

Der Aufbau des Beitrags gestaltet sich wie folgt: In Abschnitt II werden die Anforderungen an das vorgeschlagene Konzept eingeführt und anschließend der Stand der Wissenschaft analysiert. Abschnitt III erläutert das Konzept für die Chatbot-basierte Interaktion mit Ontologien. Vorläufige

Ergebnisse aus einer ersten experimentellen Studie werden in Abschnitt IV präsentiert. Abschließend fasst Abschnitt V den vorliegenden Beitrag zusammen und skizziert die Handlungsfelder für zukünftige Forschung.

## II. ANFORDERUNGEN UND STAND DER WISSENSCHAFT

### A. Anforderungen

#### **A1: Intuitive Interaktion zwischen dem Anwender und der individuell erstellten Ontologie**

Die Nutzung von Ontologien stellt für Nicht-Experten eine erhebliche Herausforderung dar. Grundsätzlich sollten Anwender in der Lage sein, Abfragen in ihren eigenen Worten zu formulieren, anstatt spezifische Fachbegriffe oder Codes verwenden zu müssen [7]. Eine wichtige Anforderung ist daher, die Hürden zur Nutzung zu senken, indem Anwender bei der Kommunikation mit der Ontologie unterstützt werden. Dies beinhaltet die Entwicklung einer benutzerfreundlichen und intuitiven Oberfläche, die es ermöglicht, auf natürliche und unkomplizierte Weise mit der Ontologie zu interagieren.

#### **A2: Flexible Abfragen an die Ontologie basierend auf den aktuellen Informationsbedürfnissen des Anwenders**

Die Möglichkeit, flexibel Abfragen an die Wissensbasis zu stellen, trägt ebenfalls zur Steigerung der Benutzerfreundlichkeit des Assistenzsystems bei. Diese Abfragen sollten an die aktuellen Informationsbedürfnisse des Anwenders angepasst werden können. Diese Flexibilität sorgt dafür, dass der Anwender die relevantesten und nützlichsten Informationen aus der Ontologie abfragen kann [8].

#### **A3: Genauigkeit und Nachvollziehbarkeit von Antworten**

Die Genauigkeit der bereitgestellten Informationen ist die notwendige Voraussetzung, damit ein Assistenzsystem vom Anwender als hilfreich empfunden wird [9]. Darüber hinaus muss der Anwender die bereitgestellten Informationen auch nachvollziehen können, damit er diesen traut. Daher ist die dritte Anforderung, die Genauigkeit und Nachvollziehbarkeit der vom Assistenzsystem bereitgestellten Antworten zu gewährleisten. Diese Anforderung gewinnt besonders im industriellen Umfeld an Relevanz, da ungenaue Informationen kostspielige Fehler, Ineffizienzen und sogar Sicherheitsrisiken zur Folge haben können.

### B. Stand der Wissenschaft

Chen et al. [10] präsentieren ein Framework für das semantische Embedding von *Web Ontology Language* (OWL)-Ontologien. Dieses Framework nutzt eine Kombination aus Random Walks und Word-Embedding-Techniken, um die Semantik von Ontologien zu kodieren, indem deren Graphenstruktur sowie die lexikalischen Informationen und logischen Konstrukte berücksichtigt werden. Die Ergebnisse deuten auf eine hohe Genauigkeit der Antworten hin. Allerdings ist aufgrund der alleinigen Generierung von Antworten durch ein LLM die Nachvollziehbarkeit der generierten Antworten nicht sichergestellt.

Chen et al. [5] stellen ein System vor, das darauf ausgelegt ist, SPARQL-Abfragen für sogenannte Frage-Antwort-Systeme effizient zu generieren. Das Hauptziel des Systems ist

es, die Kosten für Abfragen zu reduzieren. Gleichzeitig wird eine hohe Genauigkeit bei der Generierung von SPARQL-Abfragen beibehalten, die zum Abrufen von Antworten aus Datenbanken verwendet werden. Der Ansatz verwendet ein *Recurrent Neural Network* (RNN), um SPARQL-Abfragen aus gelernten und gelabelten Schlüsselwörtern zu generieren. Darauf aufbauend beschreiben Chen et al. [11] die Verbesserung von Frage-Antwort-Systemen durch fortgeschrittene *Natural Language Processing* (NLP)-Techniken und Multi-Label-Klassifikation, ebenfalls unter Verwendung von RNN. Sie heben den Einsatz von NLP hervor, um Anwenderanfragen in natürlicher Sprache zu interpretieren und zu verarbeiten. Dabei werden die Anfragen in ein Format umgewandelt, das effektiv zur Generierung von SPARQL-Abfragen genutzt werden kann. Dies umfasst den Einsatz von Technologien wie Tokenization, Lemmatization und Part-of-Speech-Tagging, um die semantische Struktur der Abfragen zu verstehen. Obwohl beide Arbeiten ebenfalls die Idee verfolgen, Anwenderfragen in SPARQL-Abfragen zu übersetzen, verzichten sie auf den Einsatz von LLMs [5, 11].

Avila et al. [6] führten Experimente durch, um die Fähigkeit von *ChatGPT* (GPT-3.5) zu bewerten, Fragen in natürlicher Sprache in Bereichen wie Familie und Beruf mithilfe von Wissensgraphen zu beantworten. Verschiedene Setups wurden getestet, darunter die direkte Beantwortung von Fragen sowie die Text-zu-SPARQL-Übersetzung unter Verwendung der terminologischen Box (T-Box), der assertionalen Box (A-Box) oder beider Komponenten. Die Ergebnisse zeigten, dass der Text-zu-SPARQL-Ansatz unter Nutzung sowohl der T-Box als auch der A-Box die beste Leistung erbrachte. Weiterhin präsentieren Avila et al. [6] ein Framework, das darauf ausgelegt ist, die Übersetzung von Fragen in natürlicher Sprache in SPARQL-Abfragen zu optimieren. Dieses Framework setzt sich aus zwei Phasen zusammen: In der Offline-Phase werden Indizes generiert, welche Begriffe aus der T- und A-Box ihren *Uniform Resource Identifiers* (URIs) zuordnen. In der Online-Phase werden diese Indizes genutzt, um die Fragen in natürlicher Sprache in SPARQL-Abfragen zu übersetzen und Antworten zu generieren. Durch die Reduzierung der Anzahl der verarbeiteten Tokens verringert das Framework die Wahrscheinlichkeit von Halluzinationen und verbessert die Unterstützung für große Wissensgraphen. Allerdings wurden die Auswirkungen der Bereitstellung von Graphenerklärungen für das LLM sowie das Komplexitätsniveau, auf dem das LLM zuverlässig SPARQL-Abfragen erzeugen kann, nicht untersucht.

## III. KONZEPT FÜR DIE CHATBOT-BASIERTE INTERAKTION MIT ONTOLOGIEN

Im Folgenden wird ein Konzept vorgestellt, das eine auf Chatbots basierende Benutzeroberfläche für die Interaktion mit Ontologien nutzt und flexible Abfragemöglichkeiten realisiert. Wie in Abschnitt II erwähnt, verwenden einige der verwandten Arbeiten LLMs direkt für die Wissensabfrage. Dies ist jedoch, wie in Abschnitt I erläutert, mit erheblichen Risiken verbunden, insbesondere im industriellen Umfeld. Daher dient

in unserem Konzept die Ontologie mitsamt SPARQL-basierter Abfragen als primäre Wissensquelle. Um Fachexperten mit begrenzten Kenntnissen beim Umgang mit Semantic-Web-Technologien zu unterstützen, werden in diesem Ansatz LLMs verwendet, um SPARQL-Abfragen aus Fragen, die in natürlicher Sprache formuliert worden sind, zu generieren. Abbildung 1 veranschaulicht das Konzept zur Abfrage der Ontologie.

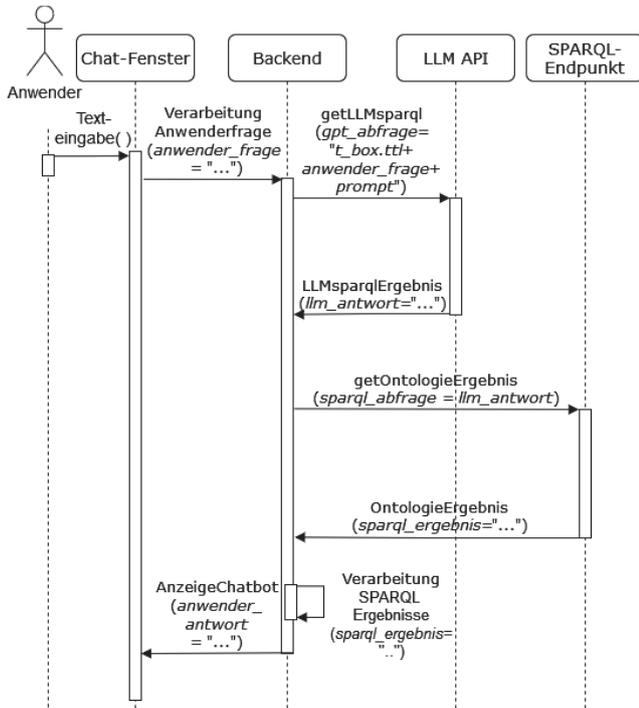


Abbildung 1: Konzept für die Chatbot-basierte Interaktion mit Ontologien [12]

Der Ablauf stellt sich wie folgt dar: Anwender interagieren über ein Chat-Fenster mit der Ontologie, indem sie Fragen in natürlicher Sprache stellen, die im Backend verarbeitet werden. Diese Abfragen werden über ein *Application Programming Interface* (API) an das LLM gesendet, wobei vordefinierte Prompts um das Schema der T-Box ergänzt werden. Durch die Einbeziehung der T-Box in die Prompts wird das explizite A-Box-Wissen von dem LLM separiert und ist nur über die SPARQL-Abfragen zugänglich. Dieser Ansatz stellt sicher, dass sensible Informationen des Unternehmens geschützt bleiben, da die Anwenderfragen nicht unmittelbar vom LLM beantwortet werden. Stattdessen verwendet das Modell die vordefinierten Prompts, um Fragen in SPARQL-Abfragen umzuwandeln. Diese Abfragen werden dann an das Backend zurückgegeben und schließlich dem SPARQL-Endpunkt übergeben, um die Instanzdaten der A-Box abzufragen. Die Ergebnisse werden im Backend verarbeitet und über die Benutzeroberfläche angezeigt. Dieser Ansatz gewährleistet, dass das LLM keine inkorrekten oder erfundenen Antworten produziert, da nur Faktenwissen abgerufen wird, welches in

der Ontologie hinterlegt ist. Allerdings sind die verifizierbaren Antworten nicht vollständig validiert. Generierte Abfragen können weiterhin fehlerhaft sein, was zu Antworten führt, die nicht der ursprünglichen Frage entsprechen. In manchen Fällen werden überhaupt keine Antworten zurückgegeben.

Da potenzielle Anwender oftmals nicht mit der Terminologie der Ontologie vertraut sind, ist es wichtig, sie bei der Formulierung der Fragen über das Chat-Fenster zu unterstützen. Besondere Aufmerksamkeit sollte den folgenden Aspekten gewidmet werden, um eine effektive Anwenderinteraktion zu gewährleisten:

1) *Prompts*: Da weder der Anwender noch das LLM über die T-Box der Ontologie informiert sind, gilt es zusätzliches Kontextwissen zu den chatbasierten Fragen in Form von Prompts zu ergänzen. Diese Prompts müssen die T-Box einschließlich ihrer Klassen, Eigenschaften und Beziehungen beinhalten. Durch die Integration ausführlicher Beschreibungen der Ontologie in die Prompts kann das LLM die Anwenderfragen besser interpretieren und sie in präzise SPARQL-Abfragen umwandeln. Diese Prompts, einschließlich der T-Box, dienen als Leitfaden für die Übersetzung domänenspezifischen Wissens in ausführbare SPARQL-Abfragen. Sie stellen sicher, dass das LLM den notwendigen Kontext und die Spezifikationen für eine genaue Formulierung der SPARQL-Abfragen erfasst.

2) *Erstellen von Ontology Design Patterns*: Die Erstellung von Ontologien ist mit großem Modellierungsaufwand verbunden. Eine modulare und auf Standards basierende Ontologie kann diesen Aufwand langfristig durch die Möglichkeit der Wiederverwendung entwickelter ontologischer Artefakte reduzieren. Vor diesem Hintergrund beschreiben Hildebrandt et al. [1] einen systematischen methodischen Ansatz zur Entwicklung von Ontologien, aufbauend auf modularen *Ontology Design Patterns* (ODPs), die auf etablierte Terminologien aus Standards zurückgreifen. Wenn diese angepassten ODPs, die einer T-Box für den Problemkontext ähneln, zu den chatbasierten Fragen ergänzt werden, kann das LLM die Strukturen und Terminologien untersuchen, um die Anfrage präzise in SPARQL gemäß den ODPs zu übersetzen.

Die Erweiterung von ODPs mit `rdfs:comments` Annotationen ist für das beschriebene Konzept entscheidend, da hierdurch zusätzlicher Kontext über die Klassen, Objekt- und Dateneigenschaften mitgeliefert werden kann. Diese zusätzliche Kontextebene hilft dem LLM, Begriffe, die mehrdeutig sein könnten oder aufgrund ihrer wörtlichen Bedeutung mehrere Interpretationen zulassen, einzuordnen. Durch die Nutzung von `rdfs:comment` gewinnt das LLM tiefere Einblicke in die Semantik der modellierten Konzepte, wodurch seine Fähigkeit verbessert wird, Anwenderfragen präzise in SPARQL-Abfragen umzuwandeln. Diese Strategie stellt sicher, dass die generierten Abfragen enger mit der zugrundeliegenden Ontologie übereinstimmen, Fehlinterpretationen minimiert werden und die Zuverlässigkeit der SPARQL-Abfragen erhöht wird.

#### IV. VORLÄUFIGE ERGEBNISSE

Zur Untersuchung des vorgestellten Konzepts wurde eine experimentelle Studie durchgeführt, um die Fähigkeit von

LLMs zur SPARQL-Abfragengenerierung zu überprüfen. Es wurde *ChatGPT-4o* verwendet, um SPARQL-Abfragen für verschiedene ODPs zu erstellen. Dabei kamen Prompts zum Einsatz, die spezifische ODP-Informationen in Klartext und eine entsprechende Frage enthielten, die die Abfrage beantworten sollten.

Die in der Studie verwendeten ODPs waren die VDI 3682 (*Formalisierte Prozessbeschreibung*) [13], die DIN EN 61360 (*Datenbank für elektrische Bauteile*) [14] und die VDI 2206 (*Entwicklung mechatronischer und cyber-physischer Systeme*) [15]. Die Fragen wurden in zwei unterschiedlichen Stilen formuliert, um den Einfluss der Formulierung auf die Qualität der Abfrage zu bewerten. Zum einen wurden *standardkonforme Fragen* (SKF) gestellt, die sicherstellten, dass die Terminologie den etablierten Standards entsprach. Zum anderen wurden *nicht-standardkonforme Fragen* (NSKF) formuliert, die eine allgemeinere Terminologie einbezogen, wie sie typischerweise von Nicht-Experten verwendet wird.

Zusätzlich wurden diese Fragen unter Verwendung der ODPs sowohl mit als auch ohne Annotationen mithilfe von `rdfs:comment` gestellt, um zu ermitteln, ob solche Kommentare die Qualität der Abfragen verbessern.

Tabelle I: Untersuchte Fragekategorien gemäß [16]<sup>1</sup>

Category	SKF Beispiel	ODP
Boolean	<i>Ist der Sensor Teil eines Moduls im System?</i>	VDI 2206
Count	<i>Wie viele technische Ressourcen sind im System enthalten?</i>	VDI 3682
Rank	<i>Können die im Modell enthaltenen Werte in aufsteigender Reihenfolge angegeben werden?</i>	DIN EN 61360
Simple	<i>Welche Prozessoperatoren werden in Prozess X verwendet?</i>	VDI 3682
String	<i>Gibt es ein DataElement mit dem Namen „ResultAccuracy“?</i>	DIN EN 61360
Two Hop	<i>Welche Komponenten sind Teil eines Moduls und welchem System gehört dieses Modul an?</i>	VDI 2206
Two Intent	<i>Aus welchen Prozessoperatoren besteht Prozess X? Welchen technischen Ressourcen sind diese Prozessoperatoren zugeordnet?</i>	VDI 3682

Komplexität

Die Komplexität der Fragen wurde gemäß dem von Rony et al. [16] vorgeschlagenen Schema in sieben Kategorien eingeteilt, wie in Tabelle I dargestellt. *Boolean*, *Count* und *Rank* repräsentieren einfachere Abfragen an die Ontologie. Diese zielen darauf ab, ein Wahr/Falsch-Ergebnis, eine numerische Anzahl oder eine Rangfolge zu liefern. *Simple*, *String* und *Two Hop* erfordern das Abfragen komplexerer Graphbeziehungen oder spezifischerer Wörter. Dies setzt ein größeres semantisches Verständnis voraus. *Two Intent* ist die komplexeste

<sup>1</sup>Zum besseren Verständnis wurden die gestellten Fragen ins Deutsche übersetzt. Während der Untersuchung wurde ausschließlich englische Terminologie verwendet.

Kategorie, da sie im Wesentlichen zwei Antworten und die Zusammenführung mehrerer Tripel erfordert. Für jede Kategorie wird in Tabelle I ein Beispiel mit einer SKF zusammen mit dem entsprechenden ODP aufgelistet. Insgesamt umfasst die experimentelle Studie 84 Fragen.

Vorläufige Ergebnisse deuten auf eine allgemein gute Genauigkeit des getesteten LLM hinsichtlich der Generierung von SPARQL-Abfragen hin. In Tabelle II wurden die Fragekategorien, die ähnliche Muster in den Ergebnissen aufwiesen, in drei Cluster eingeteilt. Die Ergebnisse zeigen, dass einfachere Fragen (*Boolean*, *Count*, *Rank*) im Allgemeinen genauere SPARQL-Abfragen lieferten, unabhängig von der Formulierung (SKF vs. NSKF) und dem Hinzufügen von `rdfs:comment` in den getesteten ODPs. Für komplexere Kategorien (*Simple*, *String*, *Two Hop*, *Two Intent*) wurde jedoch festgestellt, dass eine präzise Formulierung von SKF sowie die Hinzunahme von Kommentaren die Qualität der Abfragen erheblich beeinflusste. Bei den komplexesten Fragen (*Two Intent*) erzeugte *ChatGPT-4o* nur bei Vorliegen von SKF und Kommentaren korrekte Abfragen. Ohne Kommentare und bei Verwendung von NSKF wurden oft ungenaue oder unpräzise Abfragen erzeugt, wobei es typischerweise daran scheiterte, die korrekte Instanz zu identifizieren. Die Ergebnisse legen nahe, dass ODPs, die mit `rdfs:comment` erweitert wurden, präzisere Abfragen generierten. Dies stützt die Hypothese, dass detaillierte Kommentare in Ontologien einen positiven Einfluss auf die automatisierte SPARQL-Abfragegenerierung haben. Folglich kann geschlossen werden, dass Annotationen nicht nur das menschliche Verständnis von Ontologien verbessern, sondern auch signifikante Vorteile für LLMs hinsichtlich der Genauigkeit und Effektivität der Abfragegenerierung bieten. Insgesamt konnte die Eignung von LLMs für die (automatisierte) SPARQL-Abfragegenerierung gezeigt werden.

Tabelle II: Vorläufige Ergebnisse: Prozentsatz der korrekt generierten SPARQL-Abfragen

Kategorien	ohne Kommentar		mit Kommentar	
	SKF	NSKF	SKF	NSKF
Boolean, Count, Rank	100%	100%	100%	100%
Simple, String, Two Hop	89%	44%	100%	78%
Two Intent	67%	0%	100%	67%

## V. ZUSAMMENFASSUNG UND ZUKÜNFTIGE ARBEITEN

In diesem Beitrag wurde ein Konzept beschrieben, das LLMs und domänenspezifische Standards zur Generierung von SPARQL-Abfragen nutzt. Das übergeordnete Ziel dieses Konzepts ist es, die Interaktion von Benutzern mit Ontologien zu vereinfachen. Dieser Ansatz kann insbesondere im industriellen Kontext von Vorteil sein, da die intuitive Benutzung von LLM-basierten Chat-Anwendungen mit der formalen, strukturierten Wissensbereitstellung durch Ontologien kombiniert wird. Eine experimentelle Studie mit *ChatGPT-4o* wurde durchgeführt, um die Genauigkeit der generierten SPARQL-Abfragen unter verschiedenen Randbedingungen zu bewerten.

Die Ergebnisse heben den Mehrwert der Einbeziehung weiterer Kontextinformationen mittels `rdfs:comment` hervor. Zukünftig sollte natürlich auch die Anwendung weiterer LLMs für das beschriebene Konzept in Betracht gezogen werden.

Die Entwicklung von Strategien zur Reduzierung von Fehlern bei der Generierung von SPARQL-Abfragen ist ein wichtiges zukünftiges Forschungsthema. Dies gilt insbesondere für komplexe Abfragen und zielt darauf ab, die Genauigkeit und Zuverlässigkeit zu verbessern. Ein weiterer Schwerpunkt liegt auf der Verbesserung der Benutzerinteraktion mit auf Ontologien basierenden Systemen. Dies kann durch Verfeinerung der Prompts und durch Bereitstellung weiterer Kontextinformationen zur Ontologie für das LLM erreicht werden. Darüber hinaus ist eine weitergehende Untersuchung der Auswirkungen detaillierter `rdfs:comments` auf die Qualität der generierten SPARQL-Abfragen erforderlich. Dies schließt Tests mit komplexeren Ontologien und Variationen im Detailgrad der `rdfs:comments` ein. Robuste Validierungsmechanismen müssen ebenfalls implementiert werden. Sie sollen die Genauigkeit und Nachvollziehbarkeit der generierten Abfragen sicherstellen, insbesondere für industrielle Anwendungsfälle.

#### DANKSAGUNG

Diese Forschungsarbeit aus den Projekten ProMoDi und LaiLa wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

#### LITERATUR

- [1] C. Hildebrandt, A. Köcher, C. Kustner, C.-M. Lopez-Enriquez, A. W. Muller, B. Caesar, C. S. Gundlach, and A. Fay, "Ontology Building for Cyber-Physical Systems: Application in the Manufacturing Domain," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 3, pp. 1266–1282, 2020.
- [2] T. Jeleniewski, H. Nabizada, J. Reif, A. Köcher, and A. Fay, "A Semantic Model to Express Process Parameters and their Interdependencies in Manufacturing," in *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*. IEEE, 2023, pp. 1–6.
- [3] J. Reif, T. Jeleniewski, and A. Fay, "An Approach to Automating the Generation of Process Simulation Sequences," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2023, pp. 1–4.
- [4] M. S. Gill and A. Fay, "Utilisation of semantic technologies for the realisation of data-driven process improvements in the maintenance, repair and overhaul of aircraft components," *CEAS Aeronautical Journal*, vol. 15, no. 2, pp. 459–480, 2023.
- [5] Y.-H. Chen, E. J.-L. Lu, and Y.-Y. Lin, "Efficient SPARQL Queries Generator for Question Answering Systems," *IEEE Access*, vol. 10, pp. 99 850–99 860, 2022.
- [6] C. V. S. Avila, V. M. Vidal, W. Franco, and M. A. Casanova, "Experiments with text-to-SPARQL based on ChatGPT," in *2024 IEEE 18th International Conference on Semantic Computing (ICSC)*. IEEE, 2024, pp. 277–284.
- [7] M. Yani and A. A. Krisnadi, "Challenges, Techniques, and Trends of Simple Knowledge Graph Question Answering: A Survey," *Information*, vol. 12, no. 7, p. 271, 2021.
- [8] J. Sai Sharath and R. Banafsheh, "Conversational Question Answering Over Knowledge Base using Chat-Bot Framework," in *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*. IEEE, 2021, pp. 84–85.
- [9] J. Martinez-Gil, S. Yin, J. Küng, and F. Morvan, "Knowledge Graph Augmentation for Increased Question Answering Accuracy," in *Transactions on large-scale data- and knowledge-centered systems LII*, A. Hameurlain and A. M. Tjoa, Eds. Berlin: Springer, 2022, vol. 13470, pp. 70–85.
- [10] J. Chen, P. Hu, E. Jimenez-Ruiz, O. M. Holter, D. Antonyrajah, and I. Horrocks, "OWL2Vec\*: embedding of OWL ontologies," *Machine Learning*, vol. 110, no. 7, pp. 1813–1845, 2021.
- [11] Y. Chen, E. J.-L. Lu, and Jin-De, *Boosting Question Answering Systems with Multi-Label Classification Techniques*, 2023.
- [12] J. Reif, T. Jeleniewski, M. S. Gill, F. Gehlhoff, and A. Fay, "Chatbot-Based Ontology Interaction Using Large Language Models and Domain-Specific Standards," 2024.
- [13] VDI/VDE 3682:2, "Formalised Process Descriptions - Information Model," 05.2015.
- [14] DIN EN 61360-1, "Standard data element types with associated classification scheme - Part 1: Definitions - Principles and methods (IEC 61360-1:2017)," 07.2018.
- [15] VDI/VDE 2206, "Development of mechatronic and cyber-physical systems," 11.2021.
- [16] M. R. A. H. Rony, U. Kumar, R. Teucher, L. Kovriguina, and J. Lehmann, "SGPT: A Generative Approach for SPARQL Query Generation From Natural Language Questions," *IEEE Access*, vol. 10, pp. 70 712–70 723, 2022.

# Prävention und Beseitigung von Fehlerursachen im Kontext von unbemannten Fahrzeugen

Aron Schnakenbeck\*, Christoph Sieber, Luis Miguel Vieira da Silva,  
Felix Gehlhoff  
Institut für Automatisierungstechnik  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
aron.schnakenbeck@hsu-hh.de

Alexander Fay  
Lehrstuhl für Automatisierungstechnik  
Ruhr-Universität Bochum  
Bochum, Deutschland

**Zusammenfassung**—Unbemannte Fahrzeuge sind durch zunehmende Autonomie in der Lage in unterschiedlichen unbekanntem Umgebungen zu operieren. Diese Flexibilität ermöglicht es ihnen Ziele eigenständig zu erfüllen und ihre Handlungen dynamisch anzupassen ohne starr vorgegebenen Steuerungscode. Allerdings erschwert ihr autonomes Verhalten die Gewährleistung von Sicherheit und Zuverlässigkeit bzw. der Verlässlichkeit, da der Einfluss eines menschlichen Bedieners zur genauen Überwachung und Verifizierung der Aktionen jedes Roboters begrenzt ist. Daher werden Methoden sowohl in der Planung als auch in der Ausführung von Missionen für unbemannte Fahrzeuge benötigt, um die Sicherheit und Zuverlässigkeit dieser Fahrzeuge zu gewährleisten. In diesem Artikel wird ein zweistufiger Ansatz vorgestellt, der eine Fehlerbeseitigung während der Missionsplanung und eine Fehlerprävention während der Missionsausführung für unbemannte Fahrzeuge sicherstellt. Die Fehlerbeseitigung basiert auf formaler Verifikation, die während der Planungsphase der Missionen angewendet wird. Die Fehlerprävention basiert auf einem regelbasierten Konzept, das während der Missionsausführung angewendet wird. Der Ansatz wird an einem Beispiel angewendet und es wird diskutiert, wie die beiden Konzepte sich ergänzen und welchen Beitrag sie zu verschiedenen Aspekten der Verlässlichkeit leisten.

**Index Terms**—Unbemannte Fahrzeuge, GRAFCET, Model Checking, Laufzeitverifikation

## I. EINLEITUNG

Unbemannte Fahrzeuge (UxV) werden in den unterschiedlichsten Umgebungen eingesetzt, etwa im Wasser, an Land, in der Luft oder auch im Weltraum. Ihre Vielseitigkeit in Bezug auf Fähigkeiten und Erscheinungsformen eröffnet immer weitere Anwendungsszenarien [1]. Die kontinuierliche technologische Entwicklung fördert den Einsatz von autonomen UxV, die ein zunehmendes Maß an autonomem Verhalten aufweisen. Autonome UxVs zeichnen sich dadurch aus, dass sie in teilweise oder vollständig unbekanntem Umgebungen operieren können. Dies ermöglicht den Einsatz von unbemannten Bodenfahrzeugen (UGV) oder unbemannten Luftfahrzeugen (UAV). Diese führen keinen streng vordefinierten Steuerungscode aus, sondern haben die Fähigkeit selbstbestimmt zu erkennen, zu handeln und zu reagieren, um vorgegebene Ziele zu erreichen. Besonders vielversprechend ist die Kombination mehrerer heterogener autonomer UxVs zu einem Verbund. Einzelne Fahrzeuge können sich gegenseitig ergänzen und Schwächen kompensieren, sodass komplexere

Missionen ausgeführt werden können [2]. Selbst für größere Verbünde ermöglicht diese Autonomie die Unabhängigkeit von einem menschlichen Bediener und reduziert den Bedarf an Kontrolle und Überwachung auf ein Minimum.

Ein wichtiger Aspekt zur Sicherstellung der Autonomie eines Systems (hier autonome UxVs) ist dessen Verlässlichkeit (engl. *dependability*), sodass Autonomie nur mit verlässlichen Robotern erreicht werden kann. Die Verlässlichkeit eines Systems wird von Avizienis et al. [3] als Fähigkeit definiert, einen Serviceausfall zu vermeiden, der häufiger und schwerwiegender auftritt, als hinzunehmen ist. Da die Autonomie autonomer UxVs möglichst wenig menschliche Eingriffe und Überwachung vorsieht, können nicht alle auftretenden Fehler durch menschliches Eingreifen verhindert werden, und es kann nicht auf alle auftretenden Fehler durch einen menschlichen Bediener reagiert werden. Daher muss die akzeptable Fehlerhäufigkeit für autonome UxVs als extrem niedrig eingestuft werden, was die Bedeutung der Verlässlichkeit erhöht.

Avizienis et al. [3] definieren die Attribute Verfügbarkeit, Zuverlässigkeit (im Sinne von engl. *reliability*), Sicherheit (im Sinne von engl. *safety*), Wartbarkeit und Integrität zur Erreichung der Verlässlichkeit. Die Attribute *Verfügbarkeit* und *Zuverlässigkeit* sind notwendig, damit ein UxV wirklich autonom sein kann. Wenn das UxV nicht betriebsbereit oder in der Lage ist, einen korrekten Dienst fortzusetzen, kann keine Autonomie erreicht werden. *Sicherheit* muss von autonomen UxV selbst gewährleistet werden. Das Fehlen von menschlichen Eingriffen, die im Zweifel schwerwiegende Fehler verhindern, muss von den autonomen UxV selbst kompensiert werden. Eine weitere Anforderung an autonome UxVs ist deren *Wartbarkeit*. Reagiert ein autonomes UxV auf Umweltveränderungen, bspw. mit Systemanpassungen, so muss sichergestellt werden, dass das UxV Wartungen auf autonome Weise durchführen kann. Allerdings muss auch sichergestellt werden, dass die Änderungen durch solche Wartungen nicht zu neuen Fehlerursachen führen und die *Integrität* des Systems verletzen.

Zur Sicherstellung der Vielzahl dieser Attribute, und damit der Sicherstellung der Verlässlichkeit, bedarf es mehrerer, teils sehr verschiedener Methoden. Fokus dieser Arbeit sind die Aspekte der Zuverlässigkeit und Sicherheit, für deren Sicher-

stellung zwei unabhängige Konzepte vorgeschlagen werden: eine Methode zur Fehlerbeseitigung während der Planungsphase und eine Methode zur Fehlerprävention während der Ausführungsphase einer Mission.

Nach der Beschreibung von Grundlagen zur Missionsplanung für autonome UxVs in Abschnitt II wird in Abschnitt III-A ein formaler Verifikationsansatz vorgestellt. Dieser ermöglicht eine Verifikation von Missionsplänen während der Planungsphase und gegebenenfalls die Beseitigung von gefundenen Fehlerursachen aus den Plänen. Die Missionspläne werden in die Spezifikationsprache GRAFCET transformiert, um sie mittels statischer Analysen und Model Checking untersuchen zu können. Abschnitt III-B präsentiert ein Konzept, um sicherzustellen, dass autonome UxVs während der anschließenden Missionsausführung trotz minimaler menschlicher Überwachung gegebene Sicherheitsvorgaben einhalten, um so die Fehlerprävention sicherzustellen. Dieser zweistufige Ansatz ist notwendig, da autonome UxVs in einer dynamischen Umgebung operieren und nicht alle möglichen Ereignisse in der Planungsphase abgedeckt werden können. Der Ansatz wird in Abschnitt IV auf eine exemplarische Mission angewendet, die von zwei autonomen UxVs ausgeführt wird, bevor in Abschnitt V diskutiert wird, wie die beiden vorgestellten Verifikationskonzepte zusammenhängen und welchen Beitrag die Ansätze zur Erreichung der Verlässlichkeit leisten.

## II. GRUNDLAGEN

Dieser Beitrag betrachtet autonome UxV, die sich selbstständig bewegen und in einer unbekanntem und unkontrollierbaren Umgebung geeignete Aktionen wählen müssen [4]. Der gemeinsame Einsatz von autonomen UxV in einem Verbund ermöglicht die Durchführung einer Vielzahl komplexer *Szenarien*. Im Kontext dieses Artikels beschreibt ein Szenario eine Konstellation von Bedingungen und Umständen, wie die aktuell verfügbaren UxV, sowie ein übergeordnetes Ziel, das erreicht werden soll [5]. Um das übergeordnete Ziel eines Szenarios zu erreichen, werden *Missionen* verwendet. Eine Mission besteht aus einer Abfolge von einem oder mehreren *Missionskommandos* und den erforderlichen Parametern eines Missionskommandos, die einem einzelnen UxV zugewiesen werden. Zum Beispiel kann eine Mission eines autonomen UxV nur aus einem Missionskommando *Bewegung* mit den Parametern `pos_x` und `pos_y` der Zielposition in  $x$ - und  $y$ -Koordinaten sowie einem Parameter `vel` der Geschwindigkeit, mit der das UxV das Ziel ansteuern soll, bestehen. Diese Missionen werden aus einem Szenario abgeleitet [5].

Das Ableiten oder Planen von Missionen für einen Verbund autonomer UxV für ein bestimmtes Szenario kann sehr komplex werden und zu einer Vielzahl möglicher Lösungen führen, sodass eine automatisierte Planung erstrebenswert ist. Automatisierte Planung wurde im Bereich der KI-Planung viele Jahre lang untersucht, mit dem Ziel, eine Abfolge von Aktionen (im Kontext dieser Arbeit Missionskommandos) zu finden, die von einem Anfangszustand zu einem gewünschten Zielzustand führen [6]. Die am weitesten verbreitete Sprache im Bereich der KI-Planung ist die *Planning Domain Definition Language*

(PDDL) [7]. PDDL ist eine domänenunabhängige Sprache, die zur Beschreibung von Planungsproblemen verwendet wird, indem sowohl die Domäne als auch das Problem separat beschrieben werden. Die Domäne beschreibt hauptsächlich die von jeder Ressource bereitgestellten Aktionen mit ihren Vorbedingungen und Wirkungen. Das Problem definiert den Anfangszustand sowie den Zielzustand [7]. Zur Lösung eines solchen Planungsproblems werden Planer eingesetzt, die unter anderem Satisfiability Modulo Theories (SMT) verwenden, indem sie die Planungsprobleme als Erfüllbarkeitsprobleme in SMT formulieren und dann Solver verwenden, um sie zu lösen [8]. Wenn alle Gleichungen durch Zuweisung von Werten zu den Variablen erfüllt werden können, existiert ein Plan, der vom Startzustand zum Endzustand führt [8]. Eine Herausforderung bei der Umsetzung solcher Ansätze mit PDDL in der KI-Planung ist der Aufwand, der für die Erstellung eines solchen Planungsproblems erforderlich ist. Ein weiteres Problem ist, dass der Einsatz von KI-Planung in realen Anwendungen selten ist, da die Ausdruckskraft von PDDL nicht ausreicht [9].

Darüber hinaus gibt es Ansätze, die eine automatisierte Planung auf Basis von Informationsmodellen anstreben. Ansätze wie in [10] konzentrieren sich darauf, Funktionen von autonomen UxV formal zu beschreiben, um die folgenden zwei Aspekte anzugehen: Einerseits kann die Heterogenität verschiedener UxV in einem Verbund überwunden werden und einzelne UxV in einem Verbund können aufgrund des Informationsmodells bei Bedarf ausgetauscht werden. Andererseits kann die automatisierte Planung erleichtert werden. Solche Modelle sind komplex und deren Erstellung zeitaufwendig, sodass Ansätze zur automatischen Erstellung eines solchen Modells wie in [11] vorteilhaft sind. Es mangelt jedoch noch an Ansätzen, die eine automatisierte Planung basierend auf einem Informationsmodell durchführen. Bisher lag der Fokus auf dem Informationsmodell und Methoden zur Generierung des Modells, wobei formales Schließen oder KI-Planung für die automatisierte Planung vorgeschlagen wurden. Erste Ansätze in diese Richtung werden beispielsweise in [12] vorgestellt, bei denen ein Planungsproblem als Erfüllbarkeitsproblem in SMT aus einem Fähigkeitsmodell automatisch generiert und anschließend gelöst wird.

Aufgrund der Komplexität der Erstellung von Planungsproblemen einerseits und andererseits des Mangels an geeigneten Mitteln für die automatisierte Planung basierend auf formalen Modellen werden Missionen für autonome UxV oft noch manuell geplant. Sowohl manuelle als auch automatisierte Planung führt immer zu einem *Plan*, wie in Listing 1 gezeigt. Grundsätzlich umfasst ein Plan eine Abfolge von Aktionen mit den entsprechenden Parametern und stellt eine mögliche Lösung für ein definiertes Problem dar, um das zuvor festgelegte Ziel zu erreichen. Ein Plan besteht aus verschiedenen Zeitpunkten. Zu jedem Zeitpunkt können eine oder mehrere Aktionen aufgelistet werden, die zu diesem Zeitpunkt ausgeführt werden sollen. Wenn mehrere Aktionen einem Zeitpunkt zugeordnet sind, wie in Listing 1 für den Zeitpunkt 1 gezeigt, werden diese Aktionen parallel ausgeführt. Für jede

Aktion werden die erforderlichen Parameter angegeben.

Listing 1: Allgemeine Struktur eines Plans bestehend aus verschiedenen Zeitpunkten und entsprechenden Aktionen mit den erforderlichen Parametern.

```
0: Aktion_x Par_a
1: Aktion_b Par_x Par_c Par_f
   Aktion_c Par_d
2: Aktion_a Par_b
```

Missionen werden verwendet, um die Informationen aus einem Plan auf die einzelnen autonomen UxV zu übertragen. Missionskommandos entsprechen den Aktionen und sind eindeutig autonomen UxV zugeordnet, sodass eine Mission eines einzelnen autonomen UxV aus seinen Missionskommandos in der Reihenfolge besteht, in der sie im Plan vorkommen. *Bedingungen* sind erforderlich, wenn ein autonomes UxV ein Missionskommando erst nach Ausführung eines bestimmten Missionskommandos durch ein anderes UxV ausführen soll. Um fehlerhafte Pläne und damit Fehler in der Missionsausführung zu vermeiden, müssen solche Pläne in der Planungsphase verifiziert werden, was insbesondere für manuell erstellte Pläne gilt. Jedoch müssen auch automatisch erstellte Pläne je nach gewähltem Ansatz verifiziert werden, da beispielsweise komplexe Einschränkungen nicht immer in PDDL ausgedrückt werden können. Darüber hinaus müssen selbst korrekte Pläne für autonome UxV zur Laufzeit weiter verifiziert werden. Pläne beschreiben nur einen groben Ablauf, da autonome UxV in einer unbekanntem Umgebung arbeiten. Das bedeutet, dass ihr tatsächliches Verhalten leicht vom Plan abweichen oder ihn erweitern kann, ohne den Zweck des Plans zu gefährden. Autonome UxV müssen auf ihre Umgebung reagieren und daher beispielsweise ihren Weg zu einem Wegpunkt selbstständig bestimmen, indem sie Hindernisse durch Anpassung der Koordinaten umgehen. Dementsprechend müssen zwei Dinge sichergestellt werden:

- 1) Pläne müssen formal verifiziert werden, bevor sie an die UxV übergeben werden, um Fehler während der Ausführung im Voraus zu verhindern.
- 2) In der Ausführungsphase sind die UxV vielen Dynamiken ausgesetzt, daher muss ihr tatsächliches Verhalten auch während der Ausführung verifiziert werden.

### III. VERIFIZIERUNGSMETHODEN

Dieser zweistufige Ansatz wird im Folgenden vorgestellt.

#### A. Beseitigung von Fehlerursachen vor Missionsausführung

Um sicherzustellen, dass die Pläne fehlerfrei sind, wird formale Verifikation als Methode zur Beseitigung von Fehlerursachen eingesetzt. Die Beseitigung von Fehlerursachen wird von Avizienis et al. [3] als Mittel zur Reduzierung der Anzahl und Schwere von Fehlerursachen beschreiben. Luckcuck et al. [13] präsentieren eine Übersicht, wie formale Methoden im Kontext von autonomen UxV eingesetzt werden. Unter anderem werden Ansätze untersucht, die formale Methoden während des Entwurfs von Missionsplänen nutzen, welche später von Verbänden ausgeführt werden, wie z.B. in [14]–[16]. Diese Ansätze konzentrieren sich jedoch mehr auf die

Erstellung der Spezifikation, anstatt formale Verifikation zur Sicherstellung von Anforderungen einzusetzen.

Um eine formale Verifikation zu ermöglichen, müssen die Pläne in einer formalen Sprache modelliert werden. In dieser Arbeit wird GRAFCET [17] als graphische Modellierungssprache genutzt. GRAFCET wurde ursprünglich für die Modellierung von Steuerungsverhalten im Bereich der industriellen Fertigung entwickelt. GRAFCET ist für die Modellierung von Plänen geeignet, da gleichzeitig ablaufende Missionen über nebenläufige Sequenzen und Abhängigkeiten zwischen den Missionen über interne Variablen modelliert werden können. Zudem ist GRAFCET durch die graphische Repräsentation leicht verständlich und in der Automatisierung weit verbreitet. In der Vergangenheit haben einige Autoren dieses Beitrages bereits Methoden zur Verifikation von GRAFCET untersucht [18], [19], die im Folgenden auf Missionsplänen angewendet werden. Ein erster Schritt ist die Transformation der Missionspläne in GRAFCET. Anschließend werden im GRAFCET-Modell gewünschte Verhaltenseigenschaften verifiziert. Die zu verifizierenden Eigenschaften müssten dazu formalisiert werden, bei der Anwendung von Model Checking beispielsweise in einer temporalen Logik wie Computation Tree Logic (CTL) [20]. Nach Verifikation der Missionspläne können diese von den UxVs ausgeführt werden.

1) *Transformation von Missionsplänen in GRAFCET*: Die Regeln zur Transformation eines Plans in ein entsprechendes GRAFCET-Modell sind in Abbildung 1 dargestellt. Gestrichelte Linien stellen Platzhalterelemente dar, die mittels einer anderen Transformationsregel aus Abbildung 1 erzeugt werden. Für jede Mission im Plan wird eine neue Sequenz von Schritten (grafisch dargestellt durch ein Quadrat) in dem GRAFCET-Modell generiert, beginnend mit einem Anfangsschritt (doppelt umrandetes Quadrat) und endend mit einer Schlusstransition (grafisch dargestellt durch eine horizontale Linie). Für jeden Missionsbefehl wird ein Schritt in die Sequenz eingefügt, geordnet nach dem Zeitpunkt, an dem er im Plan auftritt. Während ein Schritt des GRAFCET-Modells aktiv ist, simuliert eine sogenannte Aktion die Ausführung eines Missionsbefehls `<command>`. Wenn das UxV den Befehl beendet hat, setzt es das Signal `<commandFinished>` auf `true`, der Schritt wird deaktiviert, und der nächste Schritt wird aktiviert, in dem der nächste Missionsbefehl ausgeführt wird.

Für jeden Missionsbefehl, der eine Bedingung (`<command_m>` in Abb. 1) enthält, wird eine interne boolesche Variable (`<condVar>`) eingeführt, die anzeigt, ob die Bedingung erfüllt ist. Die vorgelagerte Transitionsbedingung des Schritts, der dem jeweiligen Missionsbefehl entspricht, wird mithilfe eines  $\wedge$ -Operators um diese Variable erweitert. Die entsprechende Variable wird auf `true` gesetzt, nachdem der Befehl `<command_n>` abgeschlossen wurde, der die Bedingung erfüllt. Dies geschieht mit Hilfe einer sogenannten speichernd wirkenden Aktion, die `<conVar>` auf `true` setzt, wenn der zugehörige Schritt deaktiviert wird (dargestellt durch den Pfeil). Die Anwendung der Transformationsregeln führt zu einem

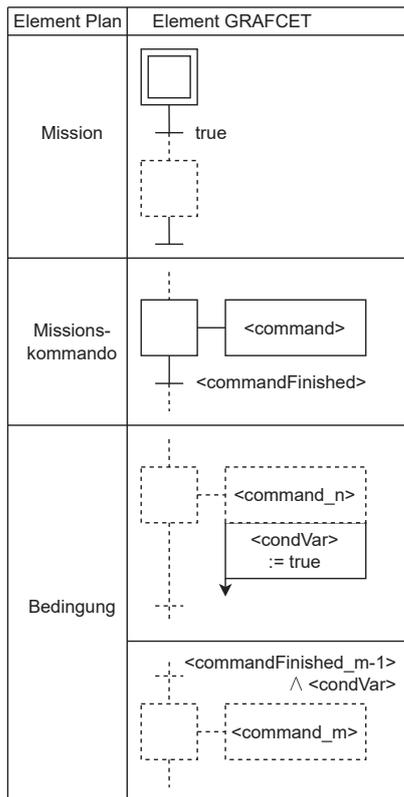


Abbildung 1: Regeln zur Transformation eines Plans (linke Seite) in ein GRAFCET-Modell (rechte Seite).

GRAFCET-Modell, das eine formale Darstellung des Plans ist und verifiziert werden kann.

2) *Eigenschaften von Missionsplänen:* Um die GRAFCET-Modelle zu verifizieren, müssen Eigenschaften definiert und vom Anwender verifiziert werden:

- **Strukturelle Anforderungen:** Es könnte verlangt werden, dass Missionen verschiedene Arten von Befehlen enthalten, wie das Greifen eines Objekts, das Fahren zu einem Ort oder das Ablegen eines Objekts. Ebenso könnte gefordert werden, dass Missionen mit einem Startbefehl beginnen und mit einem Stoppbefehl enden. Diese Eigenschaften werden mit Hilfe einer graphbasierten Analyse überprüft, die eine Tiefensuche nach Schritten durchführt, die solchen Missionsbefehlen entsprechen.
- **Chronologische Reihenfolge der Befehle:** Einige Arten von Befehlen hängen logisch von anderen ab, z. B. muss auf ein Greifen das Ablegen eines Objektes folgen. Diese Eigenschaften können mit Hilfe von temporalen Logiken formalisiert und mittels Model Checking [18] verifiziert werden. Die Formalisierung solcher Eigenschaften in temporalen Logiken erfordert ein höheres Maß an Fachwissen. Da jedoch jede Mission aus modularen und wiederverwendbaren Missionsbefehlen besteht, ist es möglich, eine bibliotheksartige Liste von Eigenschaften zu definieren, die wiederverwendet werden können. Eine solche Befehlsfolge kann in CTL als  $AG(\psi_1 \rightarrow AF\psi_2)$

formalisiert werden, was bedeutet, dass  $\psi_2$  zu einem Zeitpunkt nach  $\psi_1$  eintreten muss.

- **Freiheit von Deadlocks:** Wenn zwei UxVs gegenseitig auf die Beendigung einer bestimmten Aufgabe warten, um fortzufahren, könnte dies zu einem Deadlock führen. Mögliche Deadlocks können ebenfalls durch Model Checking erkannt werden.
- **Abwesenheit von sicherheitskritischen Situationen:** Es muss sichergestellt werden, dass bestimmte Befehle nicht gleichzeitig ausgeführt werden können, z. B. dass UxVs nicht zur gleichen Zeit den gleichen Gegenstand greifen oder dass sie nicht gleichzeitig zum selben Ort fahren. Da die UxVs die Befehle selbständig ausführen, hängt ihr genaues Verhalten von den Implementierungsdetails ab. Solche Situationen können zu einem Deadlock führen, z. B. wenn zwei UxVs versuchen, einen Zielort zu erreichen, aber gleichzeitig eine Kollision vermeiden wollen. Um diese Art von Situationen auf Missionsplanebene zu erkennen, können die entsprechenden Eigenschaften entweder mit Hilfe von Model Checking analysiert werden, oder es kann eine statische Analyse verwendet werden, wie in [19] vorgeschlagen.

Für das Model Checking können diese Situationen mithilfe von Invarianten nachgewiesen werden. Für einen Beispielplan, der zwei Missionsbefehle enthält, die ein Greifen desselben Objekts vorsehen, muss sichergestellt werden, dass die korrespondierenden Schritte im GRAFCET-Modell nicht gleichzeitig aktiv sein können:  $AG\neg(step_1 \wedge step_2)$ .

### B. Fehlerprävention während der Missionsausführung

Nach Avizienis et al. [3] umfasst Fehlerprävention zum einen die Vermeidung der Entstehung von Fehlern und zum anderen die Vermeidung der weiteren Auswirkungen bestehender Fehler. Der im vorangegangenen Kapitel erstellte und verifizierte Missionsplan räumt den einzelnen autonomen UxVs bewusst maximale Handlungsfreiheit ein. Dies hat zur Folge, dass die Missionsausführung nicht vollständig vorhersehbar und daher nur bedingt kontrollierbar ist. Erschwert wird dies zusätzlich durch unbekannte und sich ändernde Umgebungen, z. B. ist die Lage sicherheitsrelevanter Gebiete im Voraus nicht bekannt und kann sich zudem verändern. Sind solche Gebiete außerdem nicht direkt relevant für das Missionsziel, ist es möglich, dass ein autonomes UxV, trotz korrekter Wahrnehmung des Gebietes, zugehörige Informationen nicht verarbeitet. Wird z. B. ein Sperrgebiet lediglich wahrgenommen, ohne daraufhin den eigenen Pfad entsprechend anzupassen, kann dies zu unsicherem Verhalten führen. Daher ist es ratsam, den UxV parallel zum Missionsplan auch Missionsauflagen zu übermitteln. Diese Auflagen legen verbotene, hier sicherheitskritische, Verhaltensweisen fest. Eine missionsunspezifische Formulierung erleichtert die Wiederverwendung von Auflagen [21]. In [22] stellen die Autoren einen regelbasierten Ansatz zur Laufzeitverifikation vor, mit dem sie unsicheres Verhalten von unbemannten Luftfahrzeugen (UAV) erkennen konnten. Es wurden jedoch nur Sicherheitsverstöße erkannt und gemeldet,

die bereits aufgetreten waren. Eine Meldung veranlasste dann einen menschlichen Nutzer Gegenmaßnahmen einzuleiten.

Das hier vorgestellte Konzept zur missionsbegleitenden Fehlerprävention greift den Ansatz von [22] auf und erweitert ihn um eine einerseits präventiv erkennende und andererseits selbstständig reagierende Komponente. Um die Entstehung eines Fehlers mit Hilfe eines regelbasierten Systems wirksam zu verhindern, sind vier Schritte notwendig: (I) Formulierung des Fehlers, (II) Formulierung des drohenden Fehlers, (III) Fähigkeit, den Fehler aus der aktuellen Situation vorherzusagen, (IV) Gegenreaktion auf den drohenden Fehler. Diese vier Schritte werden im Folgenden anhand eines einfachen Beispiels erläutert. Zu diesem Zweck werden ein unbemanntes Bodenfahrzeug (UGV) und ein Sperrgebiet betrachtet. Das

UGV darf das Sperrgebiet nicht befahren. Der zugehörige Fehler wird wie folgt formuliert: WENN die Position des UGV im Sperrgebiet liegt, DANN liegt ein Fehler vor (I). Die aktuelle Position des UGV ist somit die Fehlerursache. Ein drohender Fehler kann nun durch die Betrachtung der bevorstehenden Position formuliert werden. WENN die bevorstehende Position des UGV im Sperrgebiet liegt, DANN liegt ein drohender Fehler vor (II). Durch die bewusste Verwendung der Variable *bevorstehende Position* innerhalb der Regel kann der zugehörige Wert auf mehrere Arten ermittelt werden. Im besten Fall sind dem UGV bereits Informationen über zukünftige Ziele und Wegpunkte bekannt. Im ungünstigsten Fall muss die bevorstehende Position anhand der aktuellen Position, des Kurses und der Geschwindigkeit für einen zu



Abbildung 2: Situation des betrachteten Anwendungsfalls mit a) UGV, b) UAV, c) Paket, d) Fabrikhalle und e) Zielposition.

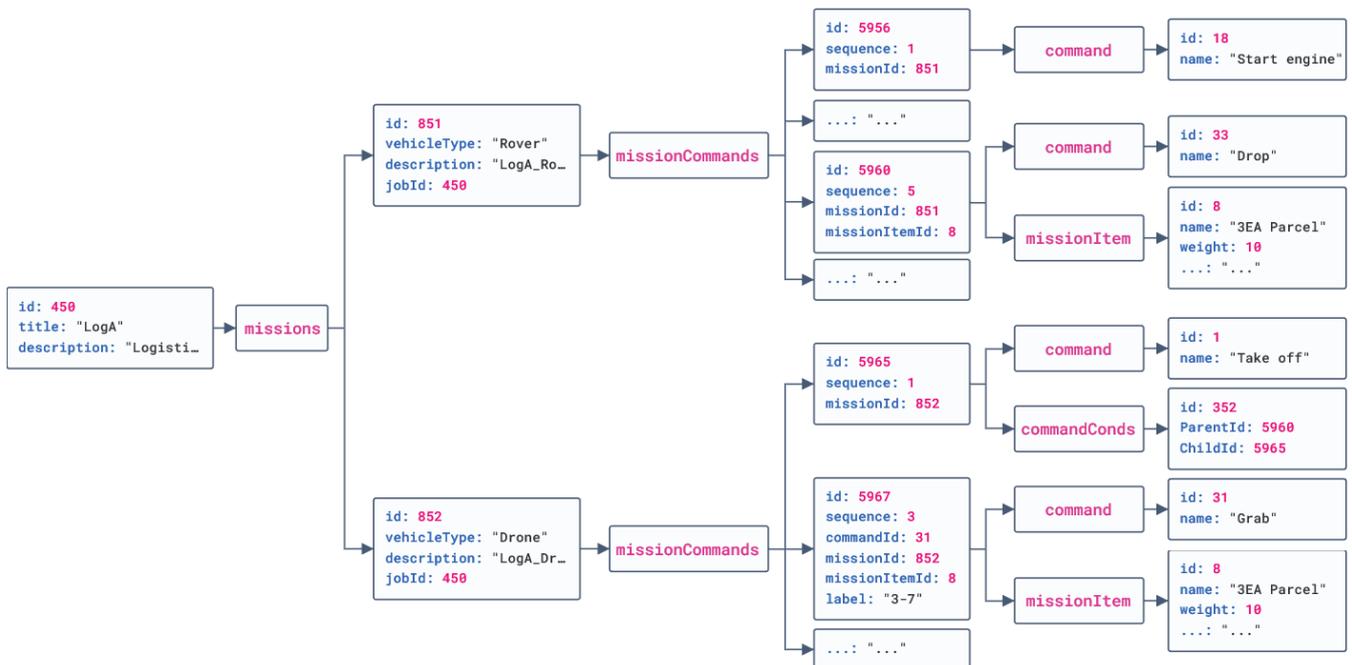


Abbildung 3: Auszug des Missionsplans im JSON-Format für das Szenario in Abbildung 2.

wählenden Zeithorizont berechnet werden, z. B. mit einer separaten Regel (III). Der Zeithorizont sollte so festgelegt werden, dass eine angemessene Reaktionszeit sowie Zeit für eine Gegenreaktion (z. B. Abbremsen) berücksichtigt wird. Allerdings kann aufgrund der relativ langen Reaktionszeit des Menschen trotz rechtzeitiger Benachrichtigung ein Fehler auftreten, bevor eine Gegenmaßnahme eingeleitet wurde. Daher kann es sinnvoll sein, am Ende der in (II) formulierten Regel bereits eine Gegenreaktion zu definieren: WENN die bevorstehende Position des UGV im Sperrgebiet liegt, DANN liegt ein drohender Fehler vor UND das UGV reduziert seine Geschwindigkeit um 50 Prozent (IV). Die hier exemplarisch vorgeschlagene Geschwindigkeitsreduzierung ermöglicht es dem UGV, einen alternativen Pfad zu bestimmen. Wenn auch dieser Pfad das Sperrgebiet kreuzt, wird die Geschwindigkeit so lange weiter reduziert, bis das UGV beim Einfahren in das Gebiet zum Stillstand kommt. Es ist zweckmäßig, dass in [22] vorgeschlagene Konzept der Benachrichtigung des menschlichen Nutzers auch bei automatischen Gegenreaktionen beizubehalten, um auch ein menschliches Eingreifen grundsätzlich zu ermöglichen.

#### IV. UMSETZUNG UND IMPLEMENTIERUNG

In diesem Abschnitt wird das Potenzial des Ansatzes und das Zusammenspiel der beiden vorgestellten Konzepte anhand eines einfachen Anwendungsfalles demonstriert. Die entsprechende Implementierung wurde mit der in [5] vorgestellten, ROS2-basierten Systemarchitektur durchgeführt. Diese ermöglicht die Trennung von Missionsplanung und -ausführung, ohne sich im Detail mit dem Aufbau, der Sensorik oder Aktorik der einzelnen UxVs auseinandersetzen zu müssen.

Im gewählten Anwendungsfall wird ein UxV-Verbund, bestehend aus einem UGV und einem UAV, verwendet, um den Transport eines Pakets durchzuführen. Dieses Paket befindet sich zunächst an einer Startposition in einer Fabrikhalle. Sein Zielort ist von einem Zaun umgeben. Da weder das UGV noch das UAV allein in der Lage sind, das Paket zu transportieren, muss das Paket außerhalb des eingezäunten Bereichs übergeben werden. Abbildung 2 zeigt die Ausgangssituation des Anwendungsfalles. Das UGV befindet sich neben dem Paket. Außerhalb befindet sich das UAV neben der Zielposition.

Das erwartete Verhalten des Verbunds besteht darin, den Transport des Pakets selbstständig durchzuführen. Der Missionsplan, der dieses Verhalten repräsentiert, wurde manuell über eine Weboberfläche geplant, die den Plan in einem JSON-Format speichert. Die resultierende JSON-Datei für das Beispielszenario ist auszugswise in Abb. 3 in einer Graphdarstellung gezeigt. Der Plan umfasst zwei Missionen: Eine für das UGV (oben) und eine für das UAV (unten). Jede Mission besteht aus einer ähnlichen Abfolge von mehreren Missionsbefehlen: *start/takeOff* → *drive/flyTo* → *grab* → *drive/flyTo* → *drop* → *drive/flyTo* → *stop/land*. Da das Paket jedoch zuerst vom UGV transportiert werden muss, wartet das UAV, bis das UGV das Paket an einer Übergabeposition abgesetzt hat. Daher wird dem *takeOff*-Befehl der UAV eine

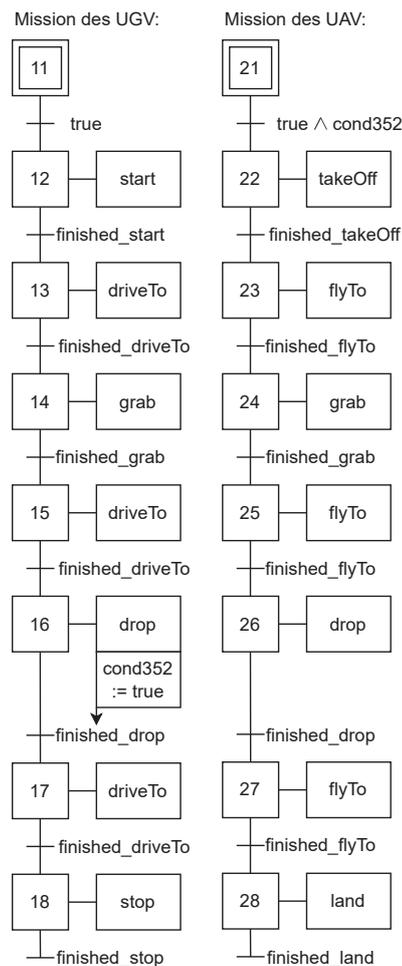


Abbildung 4: GRAFCET-Modell des Plans für den Anwendungsfall.

Befehlsbedingung hinzugefügt: Das UAV kann erst abheben, wenn das UGV das Paket abgelegt hat.

##### A. Verifikation der Missionspläne

Für die Verifikation des Plans während der Entwurfsphase wird der in Abschnitt III-A vorgestellte Ansatz auf den erstellten Plan angewendet, der in Abbildung 3 dargestellt ist. Die Anwendung der GRAFCET-Transformationsregeln resultiert in einem GRAFCET-Modell, das in Abbildung 4 dargestellt ist.

Exemplarisch werden zwei Eigenschaften verifiziert: Erstens muss sichergestellt werden, dass der Plan nicht erfordert, dass sich die UxVs näher als in einem bestimmten Sicherheitsabstand zueinander bewegen. Die Formalisierung dieser Eigenschaften kann durch Informationen aus der JSON-Datei aus Abbildung 3 unterstützt werden, in der die Zielkoordinaten z. B. für die *driveTo* und *flyTo* Befehle gespeichert sind. Wenn die Zielkoordinaten für Befehle verschiedener UxVs übereinstimmen, dürfen die entsprechenden Schritte im GRAFCET-Modell nicht gleichzeitig aktiv sein. Für das GRAFCET-Modell in Abbildung 4 bedeutet dies, dass die

Schritte 13, 14 nicht gleichzeitig mit den Schritten 23, 24 aktiv sein dürfen. In CTL kann das mit den Invarianten  $AG\neg(step_{13} \wedge step_{23})$ ,  $AG\neg(step_{14} \wedge step_{23})$ , usw. formuliert werden. Zweitens muss sichergestellt werden, dass der Plan durch die Modellierung von Bedingungen keine Verklemmung hervorruft. Für jede induzierte Bedingungsvariable (in diesem Fall nur  $cond352$ ) kann mit Hilfe der CTL-Formel  $EF(cond352 == true)$  überprüft werden, ob sie schließlich erfüllt ist. Wie in [18] vorgestellt, wurde das GRAFCET-Modell in ein Transitionssystem überführt und mit dem Model Checker ITS-Tools<sup>1</sup> verifiziert, wodurch die formalisierten Eigenschaften nachgewiesen werden konnten.

### B. Laufzeitverifikation der Missionsausführung

Die mit GRAFCET verifizierten und fehlerfreien Missionspläne können nun in ihrer ursprünglichen Form (vgl. Abb. 3) an UGV und UAV übermittelt werden. Sie gewähren den UxVs maximale Handlungsfreiheit bei der Ausführung ihrer Missionen. Wie in der Einleitung erwähnt, muss aufgrund der Autonomie der UxVs sichergestellt werden, dass während der Missionsausführung keine Fehler auftreten. Exemplarisch wird hier die Gefahr durch das Eindringen in Sperrgebiete als Fehlerursache betrachtet. Insbesondere im Zusammenhang mit UAVs werden Sperrgebiete (ugs. Geofence) zur Einhaltung von Abständen, z. B. zu Objekten und Einrichtungen, verwendet [23]. Im Kontext autonomer UxV ist es nicht zweckmäßig, feste Pfade festzulegen, um Sperrgebiete zu vermeiden. Stattdessen vermeiden autonome UxV diese Sperrgebiete effektiv, indem sie ihre Pfade selbstständig bestimmen. Abbildung 2 zeigt ein Sperrgebiet, hervorgehoben durch Pylonen. Es werden zwei einfache regelbasierte Auflagen erstellt, um sicherzustellen, dass weder das UGV noch das UAV in dieses Gebiet eindringen. Analog zu dem Ansatz in [22] werden diese Regeln innerhalb einer Ontologie mit der SWRL-Regelsprache [24] modelliert und zur Laufzeit ausgewertet. Tabelle I und Tabelle II zeigen zwei SWRL-Regeln. Regel 1 in Tabelle I erkennt vorausschauend, ob das UGV in das Sperrgebiet einfährt, und reduziert die Geschwindigkeit, um gegebenenfalls eine Neuplanung des Pfades zu ermöglichen. Wenn das UGV dennoch in das Sperrgebiet einfährt, weil es seine Geschwindigkeit nicht rechtzeitig weit genug reduziert hat oder keine Gegenreaktion erfolgte, wird es durch Regel 2 in Tabelle II gestoppt, indem seine Geschwindigkeit auf null gesetzt wird.

Für Regel 1 verlangt SWRL, dass die aktuelle Geschwindigkeit  $?velocity$  bereits in Zeile 5 in der Prämisse der Regel abgefragt wird, damit die neue Geschwindigkeit  $?newvelocity$  in der Konklusion berechnet werden kann. Da Regel 2 die Geschwindigkeit nicht berechnet, sondern fest zuweist, ist hier eine vorherige Abfrage nicht notwendig. Das UAV erhält ähnliche Regeln, die die Geschwindigkeit reduzieren oder einen Schwebeflug (engl. Loiter) bewirken. Während der Ausführung der Missionen aktualisieren beide UxVs laufend

Tabelle I: Die Regel R1 verlangsamt das UGV, wenn es in ein Sperrgebiet einfahren wird.

Line No.	SWRL-Atom
1	UGV(?myUGV)
2	$\wedge$ hasImpendingPosition(?myUGV, ?position)
3	$\wedge$ RestrictedArea(?RA)
4	$\wedge$ iswithin(?position, ?RA)
5	$\wedge$ Velocity(?myUGV, ?velocity)
6	$\rightarrow$ ImpendingFault(?myUGV, "The UGV may enter a restricted area")
7	$\wedge$ Velocity(?myUGV, ?newvelocity) $\wedge$ swrlb:multiply(0.5, ?velocity, newvelocity)

Tabelle II: Die Regel R2 bewirkt, dass das UGV anhält, sobald es sich im Sperrgebiet befindet.

Line No.	SWRL-Atom
1	UGV(?myUGV)
2	$\wedge$ hasPosition(?myUGV, ?position)
3	$\wedge$ RestrictedArea(?RA)
4	$\wedge$ iswithin(?position, ?RA)
5	$\rightarrow$ Fault(?myUGV, "The UGV is within a restricted area")
6	$\wedge$ Velocity(?myUGV, 0.0)

ihre jeweilige Wissensbasis und analysieren die Einhaltung ihrer jeweiligen Regeln.

Durch die Kombination der beiden Konzepte wird sichergestellt, dass der Packstücktransport kohärent strukturiert ist und keine sicherheitsrelevanten Störungen bei der Durchführung des Einsatzes auftreten. Das UGV holt das Paket zunächst aus der Werkshalle und transportiert es zur Übergabeposition. Die Sicherheitsauflagen verhindern einen geradlinigen Transport durch das Sperrgebiet. Sobald das Paket an der Übergabeposition abgeladen ist, startet das UAV und nähert sich, um das Paket zu übernehmen, während das UGV zurückkehrt. Abbildung 5 zeigt diese Szene im Rahmen der Mission. Nachdem das UAV das Paket am Zielort abgeladen hat, ist die Mission erfolgreich beendet.

## V. FAZIT

In diesem Artikel wurden Aspekte der Verlässlichkeit für autonome UxV untersucht. Aus der angestrebten Freiheit in der Missionsausführung und der Vielfalt von UxV ergeben sich besondere Anforderungen an den Umgang mit Fehlern. Diese Dualität von Selbstständigkeit einerseits und der Notwendigkeit der wirksamen Kontrolle andererseits erschwert die Verifikation sowohl im Rahmen der Missionsplanung als auch der Missionsausführung. Die Teilaspekte der Verlässlichkeit, insbesondere Zuverlässigkeit und Sicherheit, wurden mit Konzepten zur Fehlerbeseitigung und Fehlerprävention behandelt. Dennoch sind weder die Fehlerbeseitigung vor der Mission noch die Fehlerprävention während der Mission allein in der Lage, fehlerfreie Missionen zu gewährleisten. Die Verifikation der Missionspläne basiert nur auf unvollständigen und statischen Annahmen über die Umgebungsbedingungen, sodass trotz der funktional logischen Abfolge der Missionen deren Erfolg nicht garantiert ist. Die Verifikation der Missionsausführung berücksichtigt dagegen stärker die Umwelt und die aktuellen Bedingungen. Die damit verbundenen Regeln

<sup>1</sup><https://lip6.github.io/ITSTools-web/>



Abbildung 5: Übergabe des c) Pakets. a) UGV kehrt in die Fabrikhalle zurück und b) UAV nähert sich der Übergabeposition.

sind jedoch nicht darauf ausgelegt, den Erfolg der Mission zu gewährleisten.

Doch selbst die Kombination der beiden Konzepte erfüllt die Aspekte der Verlässlichkeit nur bedingt. Diese erfordert weitere Überlegungen zu den anderen, nicht angesprochenen Aspekten der Verlässlichkeit, wie z. B. der Wartbarkeit, wie in [3] definiert. Die Autoren bewerten eine einzelne, allumfassende Methode zur Gewährleistung der Verlässlichkeit für unzureichend und empfehlen einen modularen, kombinierbaren Ansatz zur Verifizierung verschiedener Aspekte.

In dem gezeigten Anwendungsfall ergänzen sich beide vorgestellte Konzepte, sodass die Mission erfolgreich durchgeführt werden kann. Die Autoren befürworten die Entwicklung weiterer Konzepte, insbesondere für Fehlertoleranz und Fehlervorhersage, wie sie in [3] definiert sind. Wenn solche Konzepte auf modularer Basis entwickelt werden, können sie leicht zu größeren Konstrukten kombiniert werden, was einen ganzheitlichen Ansatz für die Verlässlichkeit von autonomen UxVs erleichtert.

#### DANKSAGUNG

Diese Forschungsarbeit aus dem Projekt RIVA wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

#### LITERATUR

- [1] L. Kunze, N. Hawes, T. Duckett, M. Hanheide, and T. Krajník, "Artificial Intelligence for Long-Term Robot Autonomy: A Survey," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4023–4030, 2018.
- [2] Y. Rizk, M. Awad, and E. W. Tunstel, "Cooperative heterogeneous multi-robot systems: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1–31, 2019.
- [3] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 1, no. 1, pp. 11–33, 2004.
- [4] ISO 8373:2021(E), "Robotics – Vocabulary," 2021.
- [5] C. Sieber, L. M. Vieira da Silva, A. Fay, T. Brogt, G. Strobel, S. Berkowitz, and L. Zembrot, "A Universal Approach to Command and Control Heterogeneous Autonomous Robots," *dtec.bw-Beiträge der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw*, p. 276–280, 2022.
- [6] S. J. Russell and P. Norvig, *Artificial intelligence: A modern approach*, 4th ed., ser. Pearson Series in Artificial Intelligence. Hoboken: Pearson, 2021.
- [7] C. Aeronautiques, A. Howe, C. Knoblock, I. D. McDermott, A. Ram, M. Veloso, D. Weld, D. W. SRI, A. Barrett, D. Christianson *et al.*, "PDDL— The Planning Domain Definition Language," Technical report, Tech. Rep., 1998.
- [8] M. Cashmore, D. Magazzeni, and P. Zehtabi, "Planning for Hybrid Systems via Satisfiability Modulo Theories," *Journal of Artificial Intelligence Research*, vol. 67, pp. 235–283, 2020.
- [9] A. Rogalla, A. Fay, and O. Niggemann, "Improved Domain Modeling for Realistic Automated Planning and Scheduling in Discrete Manufacturing," in *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, vol. 1, 2018, pp. 464–471.
- [10] L. M. Vieira da Silva, A. Köcher, and A. Fay, "A capability and skill model for heterogeneous autonomous robots," *at - Automatisierungstechnik*, vol. 71, no. 2, pp. 140–150, 2023.
- [11] L. M. Vieira da Silva, A. Köcher, P. Topalis, and A. Fay, "A Python Framework for Robot Skill Development and Automated Generation of Semantic Descriptions," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2023.
- [12] A. Köcher, L. M. Vieira da Silva, and A. Fay, "Automated Process Planning Based on a Semantic Capability Model and SMT:?" [Online]. Available: <http://arxiv.org/pdf/2312.08801>
- [13] M. Luckcuck, M. Farrell, L. A. Dennis, C. Dixon, and M. Fisher, "Formal Specification and Verification of Autonomous Robotic Systems: A Survey," *ACM Comput. Surv.*, vol. 52, no. 5, 2019.
- [14] M. Kloetzer, X. C. Ding, and C. Belta, "Multi-robot deployment from LTL specifications with reduced communication," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 4867–4872.
- [15] V. Hilaire, P. Gruer, A. Koukam, and O. Simonin, "Formal specification approach of role dynamics in agent organisations: Application to the satisfaction-altruism model," *International Journal of Software Engineering and Knowledge Engineering*, vol. 17, no. 05, pp. 615–641, 2007.
- [16] K. Talamadupula, G. Briggs, T. Chakraborti, M. Scheutz, and S. Kambhampati, "Coordination in human-robot teams using mental modeling and plan recognition," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 2957–2962.
- [17] IEC International Electrotechnical Commission, "Grafcet specification language for sequential function charts," IEC 60848, 2013-02.
- [18] R. Mroß, A. Schnakenbeck, M. Völker, A. Fay, and S. Kowalewski, "Transformation of GRAFCET Into GAL for Verification Purposes Based on a Detailed Meta-Model," *IEEE Access*, vol. 10, pp. 125 665–125 665, 2022.
- [19] A. Schnakenbeck, R. Mroß, M. Völker, S. Kowalewski, and A. Fay, "A Control Flow based Static Analysis of GRAFCET using Abstract Interpretation," in *2023 IEEE 21st International Conference on Industrial Informatics (INDIN)*, 2023, pp. 1–7.
- [20] C. Baier and J.-P. Katoen, *Principles of Model Checking*, ser. The MIT Press. London, England: MIT Press, Apr. 2008.
- [21] C. Sieber, L. M. Vieira da Silva, and A. Fay, *Agilität durch Auflagen – Unterstützung der Missionsplanung für autonome Roboter*. VDI Verlag, 01 2023, pp. 691–704.
- [22] C. Sieber, L. M. Vieira da Silva, K. Grünhagen, and A. Fay, "Rule-Based Verification of Autonomous Unmanned Aerial Vehicles," *Drones*, vol. 8, no. 1, 2024.
- [23] M. N. Stevens and E. M. Atkins, "Geofencing in immediate reaches airspace for unmanned aircraft system traffic management," in *2018 AIAA Information Systems-AIAA Infotech@ Aerospace*, 2018, p. 2140.
- [24] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Groszof, M. Dean *et al.*, "SWRL: A semantic web rule language combining OWL and RuleML," *W3C Member submission*, vol. 21, no. 79, pp. 1–31, 2004.

# Ermittlung optimaler Rendezvouspunkte für heterogene Fahrzeugverbände

Anne Schweim\*, Marie Schweim, Joachim Horn

Professur für Regelungstechnik  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
\*anne.schweim@hsu-hh.de

**Kurzfassung** – In diesem Paper wird untersucht, wie Rendezvouspunkte von mindestens zwei autonomen heterogenen Fahrzeugen bestimmt werden können. Hierfür wird zunächst ein Überblick über die gängigsten Pfadplanungstechniken gegeben und dann ein Vergleich für die Anwendung bei der Lösung von Rendezvousproblemen durchgeführt. Im Anschluss werden eine Vielzahl von Rendezvouspunktproblemen, die in der Forschung betrachtet wurden, erläutert. Hierbei wird besonders auf die Lösungen eingegangen, die das Ziel haben einen optimalen Rendezvouspunkt zu finden. Danach wird das neue Rendezvousproblem, welches im Projekt RIVA gelöst werden soll, genauer beschrieben. Denn anders als bei den bekannten Problemen sollen neben multidimensionalen Kostenkarten und verschiedenen Hindernissen für die verschiedenen Fahrzeuge, die sich auch mit unterschiedlichen Geschwindigkeiten fortbewegen, nicht nur die Dauer, sondern auch der Energieverbrauch optimiert werden, wodurch die Berechnung des pareto-optimalen Rendezvouspunktes zu einem komplexen Problem wird. Die danach vorgeschlagenen Lösungsansätze zielen darauf ab, durch die Entwicklung und Anwendung geeigneter Methoden eine effektivere Koordination des heterogenen Fahrzeugverbands zu ermöglichen, wodurch neben der Logistik auch viele andere Anwendungsgebiete von einer größeren Schnelligkeit und Effizienz profitieren können.

**Stichworte** – Rendezvouspunkte, Pfadplanung, heterogene Fahrzeugverbände

## NOMENKLATUR

$s$ in m	Strecke
$t$ in s	Zeit
$v$ in m/s	Geschwindigkeit
$P$ in m	Rendezvouspunkt

## I. MOTIVATION

Autonome heterogene Fahrzeugverbände nehmen eine zunehmend wichtige Rolle in der Logistik, bei Rettungsmissionen, bei militärischen Operationen etc. ein [1]. In diesen Anwendungsgebieten können die unterschiedlichen Fähigkeiten der verschiedenen Fahrzeuge, die sich sowohl fahrend und fliegend als auch schwimmend und gehend fortbewegen können, so kombiniert werden, dass der Fahrzeugverbund Aufgaben erfüllen kann, die von einem Fahrzeug bzw. einem homogenen Fahrzeugverbund nicht erfüllt werden können. Die effi-

ziente Koordination dieser multimodalen autonomen Fahrzeuge stellt eine noch nicht abschließend gelöste Herausforderung dar, besonders wenn es keinen globalen Leitstand gibt, der die Planung übernimmt.

Eine effiziente Koordination heterogener Fahrzeugverbände ist aber von zentraler Bedeutung, denn die Anforderungen an Schnelligkeit und Präzision in logistischen Operationen nehmen immer weiter zu. Luft- und Landfahrzeuge können zum Beispiel gemeinsam genutzt werden, um die letzte Meile der Lieferung zu optimieren und somit die Lieferzeiten zu verkürzen und die Kosten zu senken [2]. In Rettungsmissionen kann die Fähigkeit, verschiedene Fahrzeuge effektiv zu koordinieren, entscheidend sein, um Leben zu retten, indem eine schnelle und präzise Zusammenarbeit zwischen Land-, Luft- und Wasserrettungsfahrzeugen ermöglicht wird [3], [4]. Auch im militärischen Kontext ist die Fähigkeit zur Koordination heterogener Einheiten ein entscheidender Faktor für den Erfolg von Operationen [5].

Die Herausforderung bei der Bestimmung optimaler Rendezvouspunkte liegt in den unterschiedlichen Bewegungseigenschaften der Fahrzeuge, wie Geschwindigkeit, Manövrierfähigkeit und Reichweite, sowie in den variierenden Umweltbedingungen, die ihre Bewegungen beeinflussen. Während Landfahrzeuge durch die Straßeninfrastruktur eingeschränkt sind, haben Luftfahrzeuge mehr Freiheitsgrade, sind aber wetteranfälliger. Wasserfahrzeuge müssen wiederum Strömungen und maritime Verkehrswege berücksichtigen.

Bisherige Ansätze verwenden eine Vielzahl von Pfadplanungstechniken, darunter Graphenbasierte-, Probabilistische- und Optimierungs-Algorithmen, sowie maschinelles Lernen und Level-Set-Methoden zur Lösung des Rendezvouspunktproblems, behandeln jedoch meist nur unimodale Fahrzeugverbände [6], [7], [8].

Zur Ermittlung optimaler Rendezvouspunkte für heterogene Fahrzeugverbände gibt dieses Paper zunächst einen umfassenden Überblick über die bestehenden Pfadplanungstechniken im Abschnitt II, sowie über bekannte Strategien zur Rendezvouspunktbestimmung im Abschnitt III. Anschließend werden in Abschnitt IV, bestehende Forschungslücken im Bereich der heterogenen Fahrzeugverbände aufgezeigt und im Abschnitt V werden mögliche Lösungsansätze vorgeschlagen. Abschließend werden im Abschnitt VI, die Ergebnisse rekapituliert und Ansätze für die zukünftige Forschung dargelegt.

## II. PFADPLANUNGSTECHNIKEN

Zur Bestimmung des optimalen Rendezvouspunktes ist die Berechnung der Pfade der einzelnen Fahrzeuge entscheidend. Es existieren diverse Algorithmen, die hierfür verwendet werden können und im Folgenden beschrieben werden.

### A. Graphenbasierte Algorithmen

Graphenbasierte Algorithmen [9] verteilen auf der Karte ein Netzwerk von Knoten und die sie verbindenden Kanten. Diese Kanten können Gewichtungen haben, die Kosten, Distanzen oder Zeiten darstellen, die benötigt werden, um von einem Knoten zum anderen zu gelangen. Die bekanntesten graphenbasierten Pfadplanungsalgorithmen sind der Dijkstra- und der A\*-Algorithmus.

#### 1) Dijkstra-Algorithmus

Der Dijkstra-Algorithmus [9] startet an einem definierten Startknoten und aktualisiert iterativ die kürzesten Wege zu allen anderen Knoten im Graphen. Bei jeder Iteration wird der Knoten mit dem aktuell kürzesten Weg ausgewählt und die Distanzen zu den benachbarten Knoten berechnet und gegebenenfalls aktualisiert. Dieser Algorithmus garantiert die Auffindung des kürzesten Pfads in einem gewichteten Graphen, ist jedoch sehr rechenintensiv. Für die Bestimmung optimaler Rendezvouspunkte heterogener Fahrzeugverbände ist der Dijkstra-Algorithmus theoretisch geeignet, jedoch aufgrund seiner hohen Rechenintensität in großen Netzwerken und dynamischen Umgebungen wenig praktikabel.

#### 2) A\*-Algorithmus

Anders als der Dijkstra-Algorithmus, berücksichtigt der A\*-Algorithmus [9] zusätzlich eine Heuristik, die die geschätzte Entfernung zum Ziel bewertet. Diese Heuristik ermöglicht eine zielgerichtete Suche, indem sie diese auf die vielversprechendsten Bereiche des Suchraums konzentriert. Die Gesamtkosten eines jeden Knoten sind die bisherigen Kosten sowie meist die euklidische Distanz zum Ziel. Durch die Kombination dieser beiden Aspekte findet der A\*-Algorithmus nicht nur den kürzesten Pfad, sondern macht dies auch effizienter als der Dijkstra-Algorithmus. Die Qualität und Effizienz des berechneten Pfads hängt jedoch von der gewählten Heuristik ab und kann somit auch schlechter bzw. langsamer als beim Dijkstra-Algorithmus sein. Der A\*-Algorithmus ist gut geeignet für die Bestimmung optimaler Rendezvouspunkte, da er effiziente und zielführende Pfadberechnungen ermöglicht. Die Wahl einer geeigneten Heuristik ist jedoch entscheidend für die Effektivität in heterogenen Fahrzeugverbänden.

### B. Potentialfeldmethode

In einem Potentialfeld [9] wird der Zielpunkt als Quelle einer attraktiven Kraft betrachtet, die mobile Einheiten anzieht, während Hindernisse abstoßende Kräfte erzeugen. Die Gesamtkraft, die auf ein Fahrzeug wirkt, ist die Summe dieser anziehenden und abstoßenden Kräfte. Durch die kontinuierliche Berechnung und Anpassung dieser Kräfte können Fahrzeuge einen kollisionsfreien Pfad zum Ziel finden. Diese Methode ist sehr schnell und effizient, kann aber in lokale Minima und bei komplexen Hinderniskonstellationen zu ineffizienten Pfaden führen. Im Falle der Rendezvouspunktbestimmung hat das jeweils andere Fahrzeug eine anziehende Kraft, aber für die Bestimmung optimaler Rendezvouspunkte ist sie aufgrund von ineffizienten Pfaden und dem Verweilen in lokalen Minima nicht geeignet.

### C. Voronoi-Diagramme

Die Kanten des Voronoi-Diagramms [9] sind maximal entfernt von den Hindernissen im Raum und bilden somit sichere und kollisionsfreie Pfade. Diese Pfade sind zwar nicht die kürzesten, aber in Umgebungen mit vielen Hindernissen die sichersten, da sie den maximalen Abstand zu diesen gewährleisten. Voronoi-Diagramme sind damit geeignet für die Bestimmung sicherer Rendezvouspunkte, finden jedoch keinen optimalen (kürzesten) Pfad. In stark hindernisreichen Umgebungen können sie dennoch von Vorteil sein.

### D. Probabilistische Algorithmen

Probabilistische Algorithmen [9] zur Pfadplanung basieren auf stochastischen Methoden, die Stichproben verwenden, um mögliche Pfade in einem Raum zu erkunden. Zu den bekanntesten probabilistischen Pfadplanungsalgorithmen gehören Rapidly-exploring Random Trees (RRT) und Probabilistic Roadmaps (PRM).

#### 1) Rapidly-exploring Random Trees (RRT)

Der RRT-Algorithmus [9] startet an einem Punkt und erweitert iterativ eine Art Baum durch zufällig ausgewählte Punkte im Raum. Neue Punkte werden so verbunden, dass sie die bestehenden Pfade im Baum erweitern, wobei Hindernisse berücksichtigt werden. Diese Methode ermöglicht schnelle Lösungen in großen und komplexen Räumen, die Pfade sind jedoch nicht optimal und müssen für die Bestimmung des optimalen Rendezvouspunktes nachträglich optimiert werden.

#### 2) Probabilistic Roadmaps (PRM)

Bei dem PRM-Algorithmus [9] wird zunächst eine zufällige Stichprobe von Punkten im freien Raum verteilt. Diese Punkte werden zu einem Graphen verbunden, indem benachbarte Punkte durch Kanten verbunden werden, die freie Pfade darstellen. Diese Methode ist sehr effizient, wenn mehrere Fahrzeuge dieselben Pfade benutzen können, jedoch ist auch hier nicht gewährleistet den optimalen Weg zu finden. Außerdem können dynamische Hindernisse nur begrenzt berücksichtigt werden.

### E. Optimierungsalgorithmen

Optimierungsalgorithmen zur Pfadplanung zielen darauf ab, den bestmöglichen Pfad zu finden, indem sie bestimmte Kriterien wie Gesamtlänge, Kosten oder Zeit minimieren. Diese Algorithmen verwenden verschiedene mathematische Techniken und Heuristiken, um die optimale Lösung zu finden oder sich ihr zu nähern. Zu den bekanntesten Optimierungsalgorithmen in der Pfadplanung gehören Genetische Algorithmen und die Partikelschwarmoptimierung (PSO).

#### 1) Genetische Algorithmen

Genetische Algorithmen [10] nutzen Prinzipien der natürlichen Selektion und Evolution zur Optimierung von Pfaden. Sie sind flexibel und anpassungsfähig, aber oft rechenaufwendig und haben Konvergenzprobleme. Damit sind sie potenziell geeignet für die Bestimmung optimaler Rendezvouspunkte, ihr Rechenaufwand und Konvergenzprobleme stellen jedoch Herausforderungen dar.

#### 2) Partikelschwarmoptimierung (PSO)

Die Partikelschwarmoptimierung [10] basiert auf dem Verhalten von Schwärmen in der Natur. Diese Algorithmen sind ebenfalls flexibel und anpassungsfähig, können jedoch ebenfalls rechenintensiv sein und Konvergenzprobleme aufweisen. Damit sind auch sie für die Bestimmung von Rendezvouspunkten geeignet, jedoch sind die Rechenintensität und

Konvergenzprobleme in großen und dynamischen Umgebungen negativ zu bewerten.

#### F. Maschinelles Lernen

Maschinelles Lernen [11] bietet die Möglichkeit, aus historischen Daten und Echtzeitinformationen zu lernen, um Pfade zu optimieren. Dies erfordert jedoch große Datenmengen für das Training der Algorithmen und kann schwer interpretierbare Ergebnisse liefern, wodurch die Validierung und das Vertrauen in die Lösung erschwert wird. Bei einer umfangreichen Datenmenge und umfangreichem Training hat das Maschinelle Lernen allerdings großes Potenzial zur Optimierung von Rendezvouspunkten und in dynamischen Umgebungen kann das fertig trainierte Modell durch Echtzeitinformationen vorteilhaft sein.

#### G. Level-Set-Methode

Die Level-Set-Methode [12] oder Niveaumengenmethode ist ein numerisches Verfahren zur Verfolgung geometrischer Objekte und deren Bewegungen. Der Pfad wird implizit als die Null-Niveaumenge einer skalaren Funktion dargestellt. Diese Funktion entwickelt sich im Laufe der Zeit, gesteuert durch partielle Differentialgleichungen (PDEs), die den Pfad zum Ziel führen und Hindernissen ausweichen. Die Methode bietet Vorteile wie eine natürliche Handhabung topologischer Änderungen, glatte Pfade und effektive Hindernisvermeidung. Bei Fahrzeugen mit vernachlässigbarer Dynamik entspricht dies einer normalen Wellenausbreitung, die vom Startpunkt ausgeht. Da jedoch dabei für das gesamte Gebiet die Iterationsschritte berechnet werden, ist sie sehr rechenintensiv und für Echtzeitanwendungen wenig geeignet, bietet dafür jedoch präzise und glatte Pfade.

#### H. Fazit

Die unterschiedlichen Pfadplanungsalgorithmen haben jeweils ihre Stärken und Schwächen, die in Tabelle 1 dargestellt sind. Optimierungsalgorithmen wie genetische Algorithmen und Partikelschwarmoptimierung sind flexibel und anpassungsfähig, aber sie sind oft rechenaufwendig und haben Konvergenzprobleme. Maschinelles Lernen bietet die Möglichkeit, aus historischen Daten und Echtzeitinformationen zu lernen, erfordert jedoch große Datenmengen und kann schwer interpretierbare Ergebnisse liefern. Graphenbasierte Ansätze, bieten eine solide mathematische Grundlage, und finden den kürzesten Weg, sind jedoch bei großen Netzwerken rechenintensiv. Der A\*-Algorithmus reduziert den Rechenaufwand im Vergleich zum Dijkstra-Algorithmus erheblich. Zudem kann der Nachteil vieler Knoten, die berechnet werden müssen, durch eine reduzierte Knotenzahl vermindert werden, indem beispielsweise nur an den Ecken von Hindernissen Knoten gelegt werden, mit denen der Algorithmus durchgeführt wird. Im Folgenden wird deshalb der A\*-Algorithmus verwendet.

Tabelle 1: Übersicht der Pfadplanungsalgorithmen

Basiert auf:	Graphen	Potentialfeldern	Geometrie	Stichproben	Optimierung		
	<b>Dijkstra, A*</b>	<b>Potentialfeld-Ansatz</b>	<b>Voronoi-Diagramm</b>	<b>PRM, RTT</b>	<b>Genetische Alg., Partikel-Schwarm</b>	<b>Maschinelles Lernen</b>	<b>Level-Set-Methode</b>
Stärke	Optimale Pfade können gefunden werden	Einfache Implementierung und Berechnung	Pfade maximieren den Sicherheitsabstand	effizient für hochdimensionale Räume	Optimale Pfade können gefunden werden	Anpassungsfähig an komplexe und dynamische Umgebungen	Optimale Pfade können gefunden werden
Schwäche	Rechenintensität hängt von der Gitterauflösung ab	Fällt leicht in lokale Minima	Hochkomplex insbesondere bei dynamischen Umgebungen	Lösungen nicht optimal und teilweise keine Lösungsfindung	Rechenintensiv, abhängig von Algorithmen variierende Effizienz	Hoher Trainingsaufwand benötigt große Datenmengen	Hoher Rechenaufwand und komplexe Implementierung
Vollständigkeit	Vollständig	Nicht vollständig	Vollständig	Wahrscheinlich vollständig			Vollständig

### III. STRATEGIEN ZUR RENDEZVOUSPUNKT-BESTIMMUNG

Besonders schwierig ist die Koordination von Fahrzeugverbänden, wenn die Fahrzeuge direkt zusammenarbeiten oder ein Rendezvouspunkt gefunden werden soll, an dem sich die Fahrzeuge treffen und beispielsweise Pakete oder Daten austauschen. Hierbei gibt es unterschiedliche Ansätze so einen Übergabepunkt zu finden. Eine Möglichkeit ist es, dass ein Fahrzeug ein anderes Fahrzeug abfängt, sodass ein Fahrzeug das Treffen herbeiführt und das andere seine Bewegung weiter fortsetzt [8], [10]. Im Unterwasserkontext gibt es im Voraus abgesprochene Rendezvouspunkte, an denen sich die Fahrzeuge treffen, da die Kommunikation sonst nicht gewährleistet werden kann [14]. In der Logistik wird die Suche nach Rendezvouspunkten betrachtet, an denen Logistikunternehmen Waren untereinander austauschen können, sodass für beide Unternehmen der Fahraufwand verringert wird [15]. Auch bei der gemeinsamen Erkundung unbekannter Gebiete durch Luft- und Landfahrzeugen gibt es Rendezvouspunkte an denen das Luftfahrzeug abgesetzt oder eingesammelt wird, um so die Vorteile beider Fahrzeuge zu nutzen, da das Landfahrzeug besser weite Strecken überbrückt, während das Luftfahrzeug an für das Landfahrzeug unerreichbare Gebiete gelangt [11].

Zur Berechnung von Rendezvouspunkten wird auch im Bereich der Spieltheorie geforscht [12]. Hier wird davon ausgegangen, dass zwei Spieler sich treffen wollen, aber keiner von beiden weiß, wo sich der andere befindet. Dabei werden Strategien vor allem auf einer Linie „on line“ und auf einem Kreis „on circle“ betrachtet und untersucht mit welchen Strategien sich die Spieler am schnellsten treffen. In einem weiteren Beispiel sind zwei Räume voller Telefone gegeben, wobei jeweils ein Telefon mit einem aus dem anderen Raum verbunden ist. Hier ist das Ziel möglichst schnell gleichzeitig ein verbundenes Telefonpaar zu benutzen.

Die bisher beschriebenen Arten Rendezvouspunkte zu bestimmen, beruhen darauf, dass sich ein Fahrzeug der Bewegungsbahn eines anderen anpasst, die Rendezvouspunkte vorgegeben werden oder nur zufällig gefunden werden können. Genauer eingegangen wird jetzt auf die Ansätze, bei denen sich beide Fahrzeuge aktiv aufeinander zu bewegen.

#### A. Physikunterricht

Ein Ansatz zur Lösung dieses Problems ist aus dem Physikunterricht bekannt. Zwei Züge fahren mit unterschiedlichen Geschwindigkeiten aufeinander zu und treffen sich an einem Rendezvouspunkt, der wie in Abbildung 1 dargestellt, berechnet werden kann, wenn die Strecke  $s$  und die Geschwindigkeiten  $v_1$  und  $v_2$  der Fahrzeuge bekannt sind.



ABBILDUNG 1: PROBLEM AUS DEM PHYSIKUNTERRICHT

Gesucht: Rendezvouspunkt P

1. Bewegungsgleichung aufstellen:

$$s = v_1 \times t + v_2 \times t$$

2. Werte einsetzen:

$$50 \text{ m} = 5 \frac{\text{m}}{\text{s}} \times t + 15 \frac{\text{m}}{\text{s}} \times t = 20 \frac{\text{m}}{\text{s}} \times t$$

3. Zeitpunkt des Treffens bestimmen:

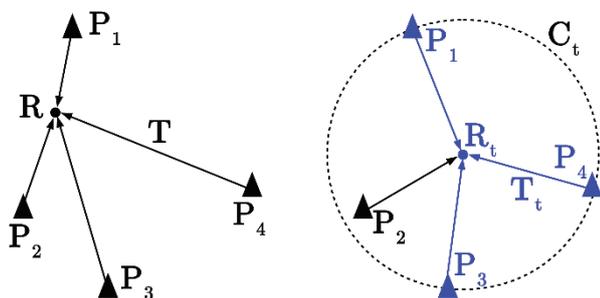
$$t = \frac{50 \text{ m}}{20 \frac{\text{m}}{\text{s}}} = 2,5 \text{ s}$$

4. Rendezvouspunkt bestimmen:

$$P = 5 \frac{\text{m}}{\text{s}} \cdot 2,5 \text{ s} = 12,5 \text{ m}$$

### B. Kreismittelpunkt

Ein weiterer Ansatz den zeitlich optimalen Rendezvouspunkt „Time Optimal Rendezvous Point“ kurz TORP zu finden wird in [6] vorgestellt. Dabei wird der TORP für eine Gruppe homogener Fahrzeuge gefunden, in dem der Mittelpunkt eines Kreises, der durch die Positionen der äußersten Fahrzeuge des Fahrzeugverbundes verläuft, berechnet wird, und dann angefahren wird. Dies ist in Abbildung 2 zu sehen. Überdies wird in dem Paper auch gezeigt, wie Hindernisse in der Berechnung integriert werden können.



(a) Rendezvous at R (b) Rendezvous at TORP,  $R_t$

ABBILDUNG 2: BESTIMMUNG DES TORP [6]

Eine Fortführung dieser Forschungsarbeit berücksichtigt unterschiedliche Geschwindigkeiten, indem ausgehend vom kürzesten Weg, der Schulphysikansatz genutzt wird. [13]

### C. Arrival-Cost-Maps

Eine weitere Möglichkeit den optimalen Rendezvouspunkt zu finden, wird in [13] und [14] anhand von Arrival-Cost-Maps beschrieben, welche auf der Level-Set-Methode beruht. Jedes Fahrzeug bestimmt ausgehend von seiner Bewegungsgeschwindigkeit und Fahrverhalten für jeden Punkt der Karte eine Ankunftszeit. Nach Kombination der Karten kann der am schnellsten erreichbare Punkte bestimmt werden.

### D. Pfadplanungsalgorithmen aus zwei Richtungen

Um die Schnelligkeit von Pfadplanungsalgorithmen zu erhöhen, kann sowohl vom Start als auch vom Ziel aus, nach dem kürzesten Weg gesucht werden. In dem Fall, dass zwei Fahrzeuge ihren Treffpunkt finden sollen, wird von beiden aus

gestartet. Dieser Ansatz wird meist genutzt, um die Pfadplanung zu beschleunigen. [15]

### E. Fazit

In der aktuellen Forschung werden beim Rendezvousproblem häufig die Fälle betrachtet, bei denen zwischen den Fahrzeugen keine Kommunikation stattfindet, bzw. keine gegenseitige Wahrnehmung der Verbundteilnehmer erfolgt. Somit wird dann auch nicht der optimale Rendezvouspunkt ermittelt. Heterogene Fahrzeuge wurden bisher selten untersucht und zum Beispiel in [8] und [13] nur insoweit betrachtet, dass nachdem der optimale Pfad gefunden wurde, der Ansatz aus dem Physikunterricht genutzt wurde, um den optimalen Treffpunkt zu bestimmen. Bei [14] wurden hingegen bereits unterschiedliche Bewegungsarten durch die Nutzung von Arrival-Cost-Maps kombiniert.

## IV. FORSCHUNGSLÜCKE

Im Projekt RIVA wird genauer untersucht, wie die Rendezvouspunkte für heterogene Fahrzeuge der Modalität Land, Luft und Wasser bestimmt werden können. Dabei haben die drei Modalitäten jeweils unterschiedliche Gebiete, in denen sie sich fortbewegen können, sowie unterschiedliche Geschwindigkeiten und Bewegungsradien. Zusätzlich soll das Pareto-Optimum hinsichtlich Energieverbrauch und Schnelligkeit gefunden werden. Einerseits soll das Paket so schnell wie möglich übergeben werden, andererseits soll der Energieverbrauch minimiert werden. Dabei muss auf die Fähigkeiten der verschiedenen Fahrzeugmodalitäten geachtet werden. Das Luftfahrzeug ist in seiner Reichweite aufgrund der Akkukapazität begrenzt und kann daher nicht lange in der Luft bleiben. Dagegen können die Land- und Wasserfahrzeuge sehr lange Strecken zurücklegen, sind dafür aber vergleichsweise langsam und in ihren Freiheitsgraden unterschiedlich stark eingeschränkt.

## V. LÖSUNGSANSÄTZE

Die ersten Ergebnisse der geplanten Rendezvouspunktbestimmung sind, wie in Abbildung 3 zu sehen, in Matlab für homogene Fahrzeuge simulativ implementiert. Oben ist die kürzeste Strecke zwischen zwei Fahrzeugen, die sich in zwei Lagerhäusern befinden mithilfe des A\*-Algorithmus berechnet worden. Der Algorithmus berücksichtigt dabei sowohl die bereits zurückgelegte Distanz als auch die euklidische Distanz als Heuristik, die die verbleibende Entfernung bis zum Ziel abschätzt. Direkt darunter ist die Anwendung des A\*-Algorithmus aus zwei Richtungen mit einem Rendezvouspunkt in der Mitte gezeigt.

Die Graustufen in den unteren beiden Bildern repräsentieren die Kosten, die berücksichtigt werden müssen, wenn die Fahrzeuge diese Gebiete durchfahren. Die Gesamtkosten für den A\*-Algorithmus setzen sich aus der Summe der Kosten diesen Punkt zu erreichen, der euklidischen Distanz und den Gebietskosten zusammen. In den weißen Bereichen liegen die Gebietskosten bei null, was bedeutet das diese Gebiete problemlos befahren werden können. In den schwarzen Bereichen hingegen sind die Gebietskosten unendlich, da die Fahrzeuge durch physische Hindernisse wie Wände nicht fahren können. Die verschiedenen Graustufen stellen unterschiedliche Gebietskosten dar, die umso stärker vermieden werden sollen, je dunkler der Farbton ist. Die dunkelgrauen Gebiete können beispielsweise hochfrequentierte Straßen sein, die aufgrund des Verkehrsaufkommens möglichst gemieden werden sollen und hohe Kosten verursachen. Dagegen sind die hellgrauen

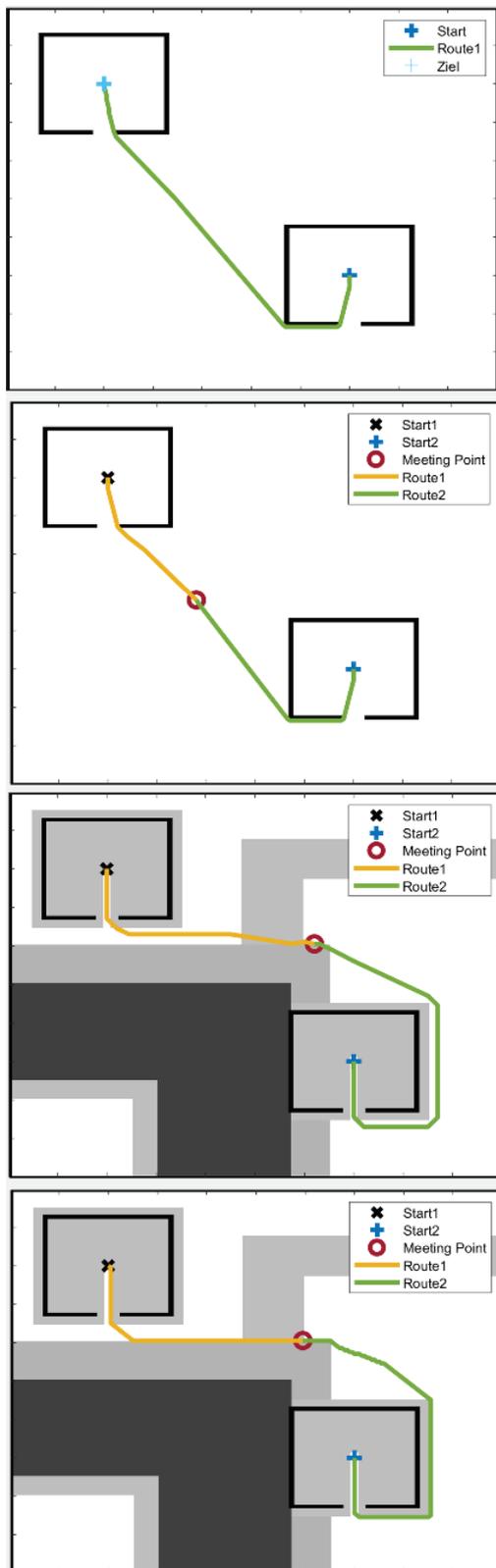


ABBILDUNG 3: RENDEZVOUSPUNKTBESTIMMUNG

Gebiete günstiger zu durchfahren und repräsentieren z. B. Fußgängerwege, die ebenfalls vermieden werden sollten, jedoch mit weniger hohen Kosten als die dunkelgrauen Gebiete durchquert werden können.

Im untersten Bild wird ein zusätzlicher Kostenfaktor eingeführt, der den Abstand zu Verbot- und Vermeidungsgebieten berücksichtigt. Hier sinken die Kosten um diese Gebiete

langsam ab um sicherzustellen, dass das Fahrzeug einen ausreichenden Sicherheitsabstand hält. Dieser zusätzliche Kostenfaktor hilft, die Sicherheit und Effizienz der Pfadplanung weiter zu verbessern, indem potenziell gefährliche oder unerwünschte Gebiete konsequenter vermieden werden.

Die Weiterentwicklung dieses Ansatzes umfasst die Bestimmung eines gemeinsamen Rendezvouspunkts für mehrere Fahrzeuge, um dann unterschiedliche Geschwindigkeiten der Fahrzeuge zu betrachten. Dabei soll berücksichtigt werden, dass die verschiedenen Fahrzeugmodalitäten unterschiedliche Bewegungsgebiete aufweisen. Diese Unterschiede sollen durch spezifische Kostenkarten für jede Modalität modelliert werden, die die unterschiedlichen Bedingungen und Einschränkungen, unter denen sich die Fahrzeuge fortbewegen können, abbilden.

Eine entsprechende Darstellung ist in Abbildung 4 dargestellt, in der verschiedene Elemente eines möglichen Szenarios gezeigt werden. Der grüne und der gelbe Kreis stellen ein Land- und ein Luftfahrzeug dar, die einen Rendezvouspunkt finden sollen. Die gestreiften Rechtecke symbolisieren die zugehörigen Fahr- und Flugverbotszonen und die schwarzen Gebiete repräsentieren feste Hindernisse, die von den Landfahrzeugen nicht durchquert werden können. Luftfahrzeuge können zwar über die Hindernisse hinwegfliegen, dies sollte jedoch vermieden werden, um potenziellen Risiken, wie Notlandungen vorzubeugen. In der Abbildung bewegt sich das grüne Landfahrzeug, ein grünes Fahrverbotsgbiet vermeidend, in Richtung des gelben Luftfahrzeugs. Das gelbe Luftfahrzeug muss jedoch eine gelbe Flugverbotszone berücksichtigen und kann sich somit entweder oberhalb dieser mit dem Landfahrzeug in den blau dargestellten Rendezvouspunkt treffen oder unterhalb, muss dann aber ein Hindernis überqueren.

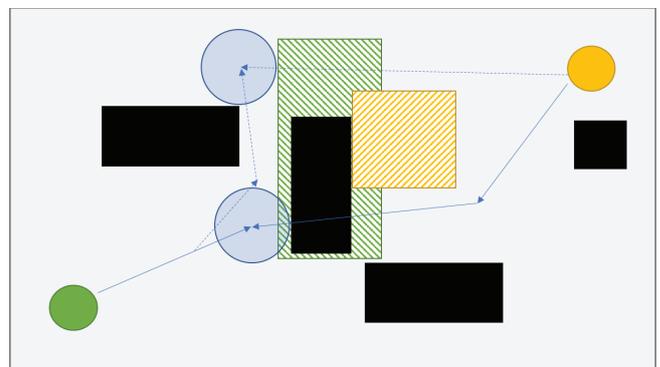


ABBILDUNG 4: MULTIDIMENSIONALE KOSTENKARTEN

Hierbei ist das Ziel das Pareto-Optimum zwischen Schnelligkeit und Energieeffizienz zu finden. Dies bedeutet, dass einerseits die schnellstmögliche Übergabe der Ladung angestrebt wird, andererseits aber auch der Energieverbrauch minimiert wird. Dies ist relevant, da das Luftfahrzeug eine begrenzte Akkulaufzeit hat, während das Landfahrzeug über eine längere Betriebszeit verfügt, jedoch langsamer ist.

Die spezifischen Kostenkarten spielen eine entscheidende Rolle bei der Optimierung. Für das Landfahrzeug umfassen diese Karten Straßen und Geländetypen, die unterschiedliche Geschwindigkeiten und Energieverbräuche verursachen. Für das Luftfahrzeug beinhalten sie Flugkorridore und Höhenprofile, die den Energieverbrauch und die Flugsicherheit beeinflussen. Durch die Verwendung dieser differenzierten Kostenkarten und einem an diese angepassten A\*-Algorithmus kön-

nen optimale Pfade für jede Fahrzeugmodalität bestimmt werden und somit der optimale Rendezvouspunkt gefunden werden.

Die Bestimmung eines optimalen Rendezvouspunkts kann zusätzlich durch die Notwendigkeit erschwert, die Kostenkarten in Echtzeit dynamisch anzupassen, um auf Veränderungen in der Umgebung oder auf Zustandsänderungen der beteiligten Fahrzeuge zu reagieren. Beispielsweise können Wetteränderungen die Flugbahnen der Luftfahrzeuge beeinflussen, während Verkehrsschwankungen die Routenplanungen der Landfahrzeuge betreffen. Solche Anpassungen erhöhen die Robustheit und Flexibilität der Pfadplanung signifikant und ermöglichen eine adaptive Reaktion auf unvorhergesehene Umstände.

Ein weiterer zu untersuchender Aspekt ist die Kommunikation und Koordination zwischen den Fahrzeugen. Durch den Einsatz von Vehicle-to-Vehicle (V2V) oder Vehicle-to-Everything (V2X) Kommunikation können die Fahrzeuge ihre Positionen, Geschwindigkeiten und geplanten Routen in Echtzeit austauschen. Dies würde die Synchronisation der Ankunftszeiten am Rendezvouspunkt verbessern und die Gesamteffizienz der Mission steigern.

Letztlich zielt diese Forschung darauf ab, die Interaktion zwischen heterogenen Fahrzeugen in komplexen Umgebungen zu optimieren. Die gewonnenen Erkenntnisse können breite Anwendungen finden, von der autonomen Logistik bis hin zu Rettungsmissionen, bei denen eine schnelle und koordinierte Bewegung entscheidend ist. Durch die Kombination von fortschrittlichen Algorithmen, spezifischen Kostenkarten und Echtzeitkommunikation können zukünftige Transportsysteme effizienter, sicherer und nachhaltiger gestaltet werden.

## VI. ZUSAMMENFASSUNG UND AUSBLICK

In diesem Paper wurden zunächst die verschiedenen Pfadplanungsalgorithmen analysiert und bewertet. Dabei setzte sich der A\*-Algorithmus aufgrund des geringen Rechenaufwandes und der Fähigkeit optimale Pfade zu finden durch, insbesondere wenn die Knoten nur in der Nähe von Eckpunkten von Hindernissen gesetzt wurden. Im Anschluss wurden die bestehenden Ansätze zur Lösung des Rendezvouspunktproblems betrachtet. Hieraus wurden aktuelle Forschungsschwerpunkte im Bereich der heterogenen Fahrzeugverbände abgeleitet und anschließend die aktuelle Forschung im Projekt RIVA vorgestellt.

In diesem Projekt wird eine präzise Berechnung pareto-optimaler Rendezvouspunkte mithilfe von multidimensionalen Kostenkarten entwickelt. Dabei hat jedes Fahrzeug seine eigenen Karten, in denen unterschiedliche Umweltbedingungen und Fahrzeugfähigkeiten mit verschiedenen Bewegungsprofilen und Energieverbräuchen berücksichtigt werden.

Zukünftige Arbeiten sollten sich auf die Erweiterung der Kostenkarten, Echtzeitkommunikation und -koordination, dynamische Anpassung der Algorithmen sowie die Simulation und Validierung in realen Szenarien konzentrieren. Dies kann die Leistungsfähigkeit und Effizienz heterogener Fahrzeugverbände weiter steigern.

## DANKSAGUNG

Diese Forschungsarbeit aus dem Projekt RIVA wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] A. Schweim, M. Zager, M. Schweim, A. Fay und J. Horn, „Unmanned vehicles on the rise: a review on projects of cooperating robot teams,“ *at - Automatisierungstechnik*, Bd. 72, Nr. 1, pp. 3-14, 2024.
- [2] P. L. Gonzalez-R, D. Sanchez-Wells und J. Andrade-Pineda, „A bi-criteria approach to the truck-multidrone routing problem,“ *Expert Systems with Applications*, 2024.
- [3] Y. Li, S. Li, W. Zhang und H. Lu, „Dynamic Route Planning for a USV-UAV Multi-Robot System in the Rendezvous Task with Obstacles,“ *J Intell Robot Syst (Journal of Intelligent & Robotic Systems)*, 2023.
- [4] M. Rodriguez, A. Al-Kaff, A. Madridano, D. Martin und A. d. La Escalera, „Wilderness Search and Rescue with Heterogeneous Multi-Robot Systems,“ *International Conference on Unmanned Aircraft Systems (ICUAS)*, 2020.
- [5] N. R. Gans und J. G. Rogers, „Cooperative Multirobot Systems for Military Applications,“ *Current Robotics Reports*, pp. 105-111, 2021.
- [6] B. Vundurthy und K. Sridharan, „Time Optimal Rendezvous for Multi-Agent Systems Amidst Obstacles - Theory and Experiments,“ *IECON*, pp. 2645-2650, 2018.
- [7] R. Parasuraman, J. Kim, S. Luo und B.-C. Min, „Multipoint Rendezvous in Multirobot Systems,“ *IEEE transactions on cybernetics*, Bd. 1, p. 310–323, 2020.
- [8] F. Kunwar, F. Wong, R. Mrad und B. Benhabib, „Time-optimal rendezvous with moving objects in dynamic cluttered environments using a guidance based technique,“ *IEEE/RSJ International Conference*, p. 283–288, 2005.
- [9] M. J. Grimble, M. A. Johnson, B. Siciliano, L. Sciavicco, L. Villani und G. Oriolo, *Robotics*, London: Springer, 2009.
- [10] N. Mathew, S. L. Smith und S. L. Waslander, „Multirobot Rendezvous Planning for Recharging in Persistent Tasks,“ *IEEE Trans. Robot. (IEEE Transactions on Robotics)*, Bd. 1, p. 128–142, 2015.
- [11] M. Zhang, H. Liang und P. Zhou, „Cooperative Route Planning for Fuel-constrained UGV-UAV Exploration,“ *2022 IEEE International Conference 2022*, p. 1047–1052.
- [12] S. Alpern und S. Gal, *The Theory of Search Games and Rendezvous*, Boston: Kluwer Academic Publishers, 2003.
- [13] O. Kulkarni, B. Vundurthy und K. Sridharan, „Rendezvous of Heterogeneous Robots in Minimum Time - Theory and Experiments,“ *Proceedings of the Advances*, 2019.
- [14] T. L. Brown, T. D. Aslam und J. P. Schmiedeler, „Determination of Minimum Time Rendezvous Points for Multiple Mobile Robots via Level Set Methods,“ *ASME 2011 International Design Engineering*, p. 787–797, 2011.
- [15] E. S. Ueland, R. Skjetne und A. R. Dahl, „Marine Autonomous Exploration Using a Lidar and SLAM,“ *ASME 2017 36th International Conference*, 2017.

# Dezentrale Informationsverteilung in Verbänden autonomer Roboter

Marvin Zager\*, Luca von Rönn, Felix Gehlhoff

Institut für Automatisierungstechnik  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
marvin.zager@hsu-hh.de

Alexander Fay

Lehrstuhl für Automatisierungstechnik  
Ruhr Universität Bochum  
Bochum, Deutschland  
alexander.fay@rub.de

**Zusammenfassung** – Der Einsatz kollaborierender Roboter in autonomen Missionen bietet signifikante Vorteile, darunter eine höhere räumliche Abdeckung und flexiblere sowie kosteneffizientere Lösungen durch spezialisierte Einheiten. Durch den Verbund autonomer Roboter kann adaptiv auf immer komplexere Umgebungen und Aufgaben reagiert werden, wovon bereits verschiedene Anwendungsbereiche profitieren. Essenziell für diese Zusammenarbeit ist der effiziente Austausch von Informationen, was jedoch aufgrund technischer Restriktionen nicht in unbegrenztem Maße möglich ist. Daher erfordern Kommunikations-einschränkungen eine sorgfältige Bewertung der Notwendigkeit jeder Nachricht, besonders in Katastrophenszenarien mit eingeschränkter Kommunikationsinfrastruktur. Dieser Beitrag untersucht bestehende Ansätze der Informationsverteilung und beschreibt eine dezentrale Methode der Informationsverteilung, welche die relevanten Aspekte zur Bewertung möglicher Kommunikation berücksichtigt. Durch selektive Kommunikation hinsichtlich des Zeitpunkts, des Inhalts und des Empfängers wird gezeigt, dass bei gleichbleibender Missionseffizienz die Netzwerklast deutlich reduziert werden kann, was besonders in Katastropheneinsätzen von Vorteil ist.

**Stichworte** – Autonomie, dezentrale Informationsverteilung, mobile Roboter, selektive Kommunikation

## I. EINLEITUNG

Der Einsatz mehrerer kollaborierender Roboter, welche in einem Verbund gemeinsam eine Mission ausführen, bringt offensichtliche Vorteile mit sich. Neben der höheren räumlichen Abdeckung erlauben kleinere und spezialisierte Roboter einen flexibleren und meist kostengünstigeren Einsatz als die Verwendung eines einzelnen Roboters, der alle Fähigkeiten allein abdecken kann. Außerdem erlaubt die Verwendung spezialisierter Roboter die Kombination unterschiedlicher Fähigkeiten, wodurch flexibler auf Herausforderungen innerhalb der Mission reagiert werden kann und der Wegfall einzelner Roboter nicht automatisch zum Abbruch der gesamten Mission führt. Eine wachsende Zahl von Anwendungen nutzt solche Roboterverbände, um Missionen autonom auszuführen. Beispiele dafür sind die Logistik [1], [2], der autonome Straßenverkehr [3], [4], [5], die Fertigung [6], [7], die Landwirtschaft [8], [9] und Katastropheneinsätze [10], [11].

Für die Kollaboration der einzelnen autonomen Roboter innerhalb des Verbunds ist dabei der Austausch von Informationen unerlässlich. Verbesserungen bei den Kommunikationstechnologien und das Aufkommen des Internet of Robotic

Things (IoRT) ermöglichen das Übertragen einer hohen Anzahl an Informationen. Dennoch sind auch diese Übertragungsmöglichkeiten begrenzt und erfordern eine Bewertung der Notwendigkeit aller möglichen Nachrichten. Einerseits ist dies generell bei einer hohen Anzahl von Robotern und dem Vorhandensein vieler Informationen (wie von zahlreichen Sensoren oder in komplexen Missionen) sinnvoll. Auch die Heterogenität von autonomen Robotern (Land-, Wasser- und Luftfahrzeuge) stellt eine große Herausforderung für Kommunikationsnetzwerke dar und erfordern ein Selektieren der möglichen Nachrichten [12]. Andererseits ist eine selektive Kommunikation insbesondere für Katastropheneinsätze unerlässlich [13], in denen die lokale Infrastruktur und Netzwerke nur eingeschränkt verwendet werden können (vgl. Abbildung 1). So waren beispielsweise nach dem Tsunami in Thailand im Jahr 2004 vielerorts sämtliche Kommunikationsverbindungen mehrere Tage unterbrochen [14]. Auch die Hilfsaktionen nach dem Erdbeben in der Türkei und Syrien im Jahr 2023 waren von Unterbrechungen im Kommunikationsnetz geprägt [15]. Dies zeigt die Notwendigkeit eines ausgereiften Entscheidungsprozesses für kommunikative Handlungen im Bereich autonomer Roboterverbände. Dabei geht es um Entscheidungen, welche Informationen wann und mit wem ausgetauscht werden sollen, um die vorgegebene Mission möglichst effizient auszuführen und gleichzeitig das Kommunikationsaufkommen zu minimieren. Obwohl der Austausch umfangreicher Informationen in autonomen Roboterverbänden unerlässlich ist, wird die Bewertung kommunikativer Handlungen in solchen Verbänden in der Literatur oft vernachlässigt [16], [17].

Die Optimierung der Kommunikation in autonomen Roboterverbänden unter bestimmten Einschränkungen erfordert

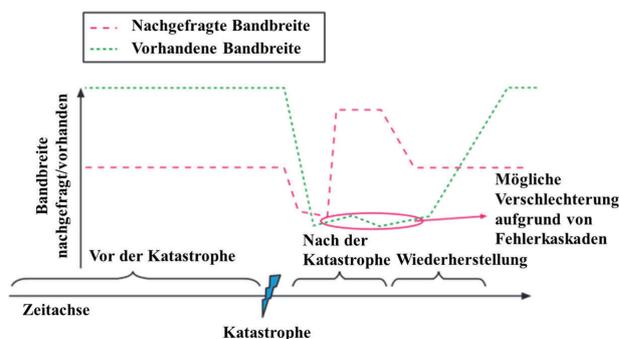


ABBILDUNG 1: BANDBREITE IN KATASTROPHENEINSÄTZEN [ADAPTIERT AUS 13]

eine Bewertung jeder möglichen kommunikativen Handlung. Die spezifischen Bewertungskriterien hängen im Allgemeinen von der Mission ab, die der autonome Roboterverbund ausführt. Direkte Bewertungsansätze [18], [19] simulieren zukünftige Missionsschritte, um die Auswirkungen der einzelnen Kommunikationsereignisse auf das Missionsziel abzuschätzen. Diese Ansätze können zwar für bestimmte Anwendungen eine optimale Leistung erbringen, erfordern aber die Fähigkeit, die Zukunft einer Mission zu simulieren, was zeit- und ressourcenintensiv sein kann. Indirekte Ansätze weisen hingegen den während der Mission verarbeiteten Informationen einen Nutzen zu und übermitteln die wertvollsten Informationen, um die Missionseffizienz positiv zu beeinflussen. Bei diesen Ansätzen wird jedoch davon ausgegangen, dass die Anwendungslogik die Informationsunsicherheit berücksichtigt, was in der Praxis oft nicht der Fall ist und die Integration in bestehende Systeme erschwert. Alternativ werden bei einer indirekten Bewertung der Nutzen einer kommunikativen Handlung durch eine spezifisch auf den Anwendungsfall zugeschnittene Nutzenfunktionen verwendet [20], [21]. Diese berücksichtigen jedoch oft nur den unmittelbaren Wert der Kommunikation, ohne die langfristigen Auswirkungen zu bewerten. Komplexere Modelle werden nur selten berücksichtigt.

Welcher Ansatz auch gewählt wird, es lässt sich feststellen, dass die Integration und Nutzung einer dezentralen Informationsverteilung in autonomen Roboterverbänden eine Herausforderung darstellen. Die wesentlichen Herausforderungen einer dezentralen Informationsverteilung lassen sich in zwei Hauptprobleme (HP) einteilen:

#### **HP1: Bewertungsprozess kommunikativer Handlungen**

In einem dezentralen System müssen Roboter selbstständig entscheiden, *wann* und *mit wem* sie *welche* Informationen austauschen. Dies erfordert eine präzise Bewertung des Nutzens und Relevanz der zu kommunizierenden Informationen. Hier fehlt es an einer einheitlichen Methode, die es erlaubt die Aspekte kommunikativer Handlungen zu bewerten und deren Auswirkungen abzuschätzen.

#### **HP2: Kollaboration innerhalb des Roboterverbunds**

Die Kollaboration in einem dezentralen Roboterverbund soll durch den Austausch von Informationen ermöglicht bzw. verbessert werden und stellt damit den eigentlichen Zweck der kommunikativen Handlungen dar. Eine selektive Informationsweitergabe ist notwendig, um die Aufgabenverteilung und effiziente Koordination der Roboter sicherzustellen und damit die Missionseffizienz zu maximieren. Insbesondere in Katastropheneinsätzen stellt dies jedoch eine Herausforderung an Kommunikationsstrategien, da hier häufig Kommunikationsbeschränkungen existieren.

Um diese Probleme zu lösen, wird im Folgenden eine Methode für eine dezentrale Informationsverteilung in Robotikanwendungen zur Förderung kollaborierenden Verhaltens vorgestellt. Diese Methode ermöglicht die selektive Verteilung sämtlicher potenzieller Informationen, die in Robotikanwendungen auftreten können, um Bandbreitenbeschränkungen zu berücksichtigen. Um diese Informationen einheitlich zu strukturieren, wird das Informationsmodell aus [22] verwendet. Weiterhin erfolgt eine Einbettung der Methode in einen Entscheidungsprozess der Robotik [23], welcher die Integration in Robotikanwendungen vereinfacht.

In Kapitel II wird mit Hilfe einer Literaturanalyse zu verschiedenen Ansätzen dezentraler Informationsverteilungen

zunächst der Forschungsbedarf herausgearbeitet, um daraus Anforderungen an eine mögliche Methode abzuleiten. Die Methode als erarbeitetes Artefakt mit den betrachteten Aspekten für kommunikative Entscheidungsprozesse wird in Kapitel III aufgezeigt. Anschließend wird in Kapitel IV die Anwendbarkeit der vorgestellten Methode in einer prototypischen Anwendung demonstriert. Zudem wird mittels einer Simulation eine Analyse des Einflusses einzelner Aspekte, die kommunikative Handlungen beeinflussen, durchgeführt. Die gewonnenen Erkenntnisse werden in Kapitel V ausgewertet und analysiert, bevor im abschließenden Kapitel VI die Ergebnisse zusammengefasst werden.

## II. LITERATURANALYSE

Um vorhandene Konzepte im Bereich der dezentralen Informationsverteilung und ihre Eignung zur Lösung der beiden Hauptprobleme zu bewerten, wurde die Fachliteratur systematisch analysiert. Zunächst wurden wissenschaftliche Fragestellungen (WF) aufgestellt, um anschließend neun verschiedene Konzepte hinsichtlich dieser Fragestellungen zu bewerten.

### A. Wissenschaftliche Fragestellungen

Die Autoren in [24] diskutieren eine Reihe von Anforderungen an den Einsatz autonomer Roboter im industriellen Anwendungsfall. Mit Fokus auf die dezentrale Informationsverteilung und der Ergänzung von Designprinzipien für die Entwicklung von verteilten Umgebungsmodellen für autonome, kollaborierende Roboter [25] lassen sich fünf relevante wissenschaftliche Fragestellungen identifizieren, die den zwei Hauptproblemen zuzuordnen sind. Diese wissenschaftlichen Fragestellungen erfüllen zwei Aufgaben: Erstens dienen sie als restriktiver Rahmen für die Entwicklung der in Kapitel III vorgestellten Methode zur selektiven Informationsverteilung. Zweitens betonen sie die spezifischen Aspekte der vorgestellten Methode und deren Alleinstellungsmerkmal im Vergleich zu vorhandenen wissenschaftlichen Konzepten. Die jeweiligen WFs gilt es zu beantworten, um die Hauptprobleme zu lösen.

#### **HP1: Bewertungsprozess kommunikativer Handlungen**

WF1: Wie wirken sich Bewertungen kommunikativer Handlungen hinsichtlich der zeitlichen Relevanz auf das kollaborative Verhalten in einem Roboterverbund aus?

WF2: Wie wirken sich Bewertungen kommunikativer Handlungen hinsichtlich des Inhalts einer Nachricht auf das kollaborative Verhalten in einem Roboterverbund aus?

WF3: Wie wirken sich Bewertungen kommunikativer Handlungen hinsichtlich der relevanten Empfänger auf das kollaborative Verhalten in einem Roboterverbund aus?

#### **HP2: Kollaboration innerhalb des Roboterverbunds**

WF4: Wie wirkt sich ein Bewertungsprozess kommunikativer Handlungen auf die Missionseffizienz eines kollaborierenden Roboterverbunds aus?

WF5: Wie wirkt sich ein Bewertungsprozess kommunikativer Handlungen auf die Netzwerklast eines kollaborierenden Roboterverbunds aus?

### B. Analyse wissenschaftlicher Arbeiten

In [26] soll durch kooperative Wahrnehmung die Sicht von autonomen Fahrzeugen verbessert und erweitert werden, um so die Sicherheit im Straßenverkehr zu erhöhen. Dafür teilt ein autonomes Fahrzeug verschiedene Informationen wie Kurs, Position, Geschwindigkeit oder Anzahl und Reichweite der Sensoren über ein lokales Netzwerk mit dem umliegenden Verkehr. Je dynamischer die detektierten Objekte sind, desto häufiger werden die darüber ermittelten Informationen auch mit dem umliegenden Verkehr kommuniziert. Die Auswirkungen unterschiedlicher Übertragungsfrequenzen auf die entstehende Netzwerklast werden ebenfalls betrachtet. In der Veröffentlichung wurde gezeigt, dass durch selektive Kommunikation die Wahrnehmung der Fahrzeuge deutlich verbessert werden kann und die dafür notwendigen Anforderungen an die kommunizierten Informationen wie Vollständigkeit und Aktualität auch bei höheren Verkehrsdichten durch lokale Netzwerke erfüllt sind. Allerdings fehlt in dieser Veröffentlichung der Vergleich mit einer Baseline, aus der deutlich wird, dass es sich um eine effiziente Art der Kommunikation handelt.

Für die Informationsverteilung in Multi-Roboter-Anwendungen wird in [27] ein mathematisches Modell entwickelt, mithilfe dessen zunächst der Informationswert (Utility) einer einzelnen Kommunikationsnachricht bestimmt wird, um anschließend die Auswirkungen auf die globale Zielfunktion zu bewerten. Mit dieser indirekten Bewertung zeigen die Autoren durch den Vergleich mit mehreren Vergleichsszenarien, dass eine selektive Kommunikation mit verschiedenen Planungshorizonten zu einer deutlichen Verringerung der Netzwerklast bei nahezu konstanter Missionseffizienz führt. Auf die Frage danach, *welche* Informationen geteilt werden, wird nicht eingegangen.

[28] untersucht ebenfalls die Informationsverteilung in Multi-Roboter-Anwendungen unter Ausnutzung des Informationswerts (hoch/gering) einer einzelnen Nachricht, wobei beispielsweise der Batteriezustand, die Position der Roboter, Bilder oder der Objektstatus übermittelt werden. Für den Fall, dass eine Nachricht mit hohem Nutzen bewertet wird, wird die entsprechende Übermittlung durchgeführt. Dabei zeigen sowohl Simulations- als auch Realexperimente die Realisierbarkeit des entwickelten Konzepts und die Auswirkungen auf Missionseffizienz und Netzwerklast im Vergleich zu einem Vergleichsszenario. Eine Bestimmung der zeitlichen Komponente, *wann* diese Nachrichten kommuniziert werden und *an wen* eine Übermittlung erfolgt, wird nicht weiter betrachtet.

Für autonome Agenten wird in [29] ein Kommunikationsmodell entwickelt, welches Ansätze der epistemischen Logik nutzt und kommunikative Handlungen indirekt bewertet. Dabei werden neue Wahrnehmungen immer dann an andere Agenten kommuniziert, wenn sie einen Mehrwert für die andere Agenten zur Erfüllung derer Aufgabe aufweisen. Die Informationen werden dabei entweder an alle oder in Abhängigkeit ihrer Aufgabe kommuniziert. Zwar zeigt sich durch den Vergleich mit mehreren Vergleichsszenarien die grundsätzliche Missionseffizienz des entwickelten Konzepts, jedoch werden die Auswirkungen auf das verwendete Kommunikationsnetzwerk nicht betrachtet.

Die Autoren in [30] beschäftigen sich mit der Informationsverteilung zwischen autonomen Drohnen, auch Unmanned Aircraft Systems (UAS) genannt. Dafür wird zunächst zwischen den grundlegenden technischen Realisierungen für die Kommunikation unterschieden, um anschließend eine Kommunikationsarchitektur im Bereich des UAS Traffic Management (UTM) vorzustellen. Es werden Informationen wie die Position, Kurs, Drohrentyp, Geschwindigkeit und Flugplan geteilt. *Wann* und *mit wem* diese Informationen ausgetauscht werden, wird nicht betrachtet.

In [31] sollen verdeckte Objekte im autonomen Straßenverkehr durch kooperative Wahrnehmung sichtbar gemacht werden. Dafür werden verschiedene Informationen wie die eigene Position, die Größe und Position der Objekte sowie der Kurs mit allen Fahrzeugen im Netzwerk geteilt. Die selbst wahrgenommenen Objekte sowie die mitgeteilten Informationen werden anschließend von jedem Fahrzeug zu einer 2D-Karte in Vogelperspektive (Birds Eye View-(BEV)-Maps) zusammengefügt. Dabei wird in dem Beitrag deutlich, dass diese kompakten BEV-Maps zu deutlich reduzierten Kommunikationsaufwand bei gleichzeitig geringerem Berechnungsaufwand führen. Nicht betrachtet werden jedoch die Auswirkungen dieses Konzepts auf die Missionseffizienz und die Netzwerklast.

Für die dezentrale Kommunikation bei autonomen Robotern werden in [32] Ansätze aus den Bereichen Multi-Agent Reinforcement Learning (MARL) und Non-Linear Model Predictive Control (NMPC) kombiniert, um den Kommunikationsbedarf in einer 3D-Umgebung zu minimieren. Dabei wird immer dann kommuniziert, wenn der Abstand unterhalb eines definierten Wertes liegt. Durch den Vergleich mit drei Vergleichsszenarien wird die Effizienz der vorgestellten Lösung deutlich. Die Untersuchung unterschiedlicher Informationstypen und Netzwerkauswirkungen werden ausgelassen.

Die Autoren in [18] beschäftigen sich mit der Optimierung der Kommunikation von autonomen Robotern bei beschränkter Bandbreite. Dafür wird ein direkter Bewertungsansatz verwendet, um zu bestimmen, *was* in einer Nachricht enthalten sein soll. Vereinfachend wird eine gitternetz-basierte Umgebung genutzt und die Belegung umliegender Zellen des autonomen Roboters als mögliche Kommunikationsinhalte definiert. Dabei konnte gezeigt werden, dass durch den entwickelten Ansatz sowohl eine gesteigerte Missionseffizienz als auch ein reduzierter Kommunikationsbedarf erfolgen kann. Eine differenzierte Betrachtung hinsichtlich des Zeitpunktes des Versendens und sowie eine Selektion bestimmter Verbundteilnehmer erfolgt nicht. Auch die Netzwerklast zur Beurteilung der Einsetzbarkeit in Katastrophensituationen wird nicht betrachtet.

Eine Selektion *welche* Informationen zwischen autonomen Agenten ausgetauscht werden sollten wird in [33] durchgeführt. Anstatt sämtliche Informationen zu senden, werden Informationen zunächst auf ihre Relevanz und ihren Einfluss für andere Agenten beurteilt. Die anschließend geteilten Informationen umfassen einerseits die wahrgenommenen Objekte und

TABELLE 1: ZUSAMMENFASSUNG DER LITERATURANALYSE

Kategorie	Quelle	Jahr	HP1			HP2	
			WF1	WF2	WF3	WF4	WF5
Autonomer Straßenverkehr	[26]	2020	x	x	x		x
	[31]	2023	x	x	x		
Autonome Roboter	[27]	2018	x		x	x	x
	[18]	2020		x		x	
	[32]	2020	x		x	x	
	[28]	2021		x		x	x
Autonome Agenten	[29]	2021	x	x	x	x	
	[33]	2022		x		x	

zum anderen die relative Position der Nachbarn. *Wann* diese Informationen ausgetauscht werden und konkret *mit welchem Agenten* wird in diesem Beitrag allerdings nicht weiter diskutiert. Die Effizienz des entwickelten Konzepts zeigt sich zwar in mehreren Vergleichsszenarien, allerdings wird auf die Netzwerklast nicht weiter eingegangen.

Tabelle 1 fasst die durchgeführte Literaturanalyse zusammen und bildet den jeweiligen Erfüllungsgrad bezogen auf die wissenschaftlichen Fragestellungen ab. Der Übersichtlichkeit halber wurden die vorgestellten Ansätze in drei Kategorien unterteilt. Die erste Kategorie umfasst Ansätze, die im Bereich des autonomen Straßenverkehrs Anwendung finden. Die zweite Kategorie ist den autonomen Robotern gewidmet, während die dritte Kategorie Ansätze enthält, die die Kommunikation innerhalb autonomer Agenten im Allgemeinen untersuchen.

### III. METHODE ZUR INFORMATIONSVERTEILUNG

Ausgehend von den in Kapitel II vorgestellten Konzepten und abgeleiteten Anforderungen, wird im folgenden Abschnitt eine Methode für eine dezentrale Informationsverteilung für autonome Fahrzeugverbände vorgestellt. Diese Methode beinhaltet Auswahlprozesse, *wann*, *was* und *mit wem* jeder Roboter kommunizieren sollte, und zielt darauf ab, die Kommunikationsaktivitäten zu minimieren, um so zur gleichmäßigen Netzwerklast beizutragen. Gleichzeitig wird so eine effiziente Ausführung der vorgegebenen Mission des Fahrzeugverbundes ermöglicht. Abbildung 2 gibt einen Überblick über die Methode deren Teilschritte im Folgenden, nach einer kurzen Darstellung grundlegender Annahmen, detailliert erläutert werden.

Es werden  $N$  kollaborierende heterogene Roboter betrachtet, denen ein übergeordnetes Verbundziel (z. B. durch den Menschen) vorgegeben ist. Ausgehend von individuellen Fähigkeiten und ihrem aktuellen Status, entscheiden die Roboter autonom über einzelne Aktionen, um bestmöglich das Verbundziel zu erreichen. Die Roboter können sich dabei frei im Raum bewegen und ihre Position (z. B. über das Global Navigation Satellite System (GNSS)) bestimmen. Die Verteilung von Informationen innerhalb des Roboterverbundes ist dabei unabhängig von der technischen Umsetzung der Kommunikation, wobei davon ausgegangen wird, dass die Roboter in der Kommunikationsreichweite nicht eingeschränkt sind. Empfangene Nachrichten werden verzögert aber im Ganzen und unverfälscht empfangen. Weiterhin sind die Roboter durch geeignete Sensorik in der Lage, ihre Umgebung wahrzunehmen

und Hindernisse zu klassifizieren. Die Roboter nutzen sowohl selbst wahrgenommene Umgebungsinformationen als auch von anderen Robotern kommunizierte Informationen, um ihr internes Weltmodell zu kreieren und zu aktualisieren. Da der Fokus der vorgestellten Methode auf dem Selektionsprozess für kommunikative Handlungen und den Auswirkungen für das kollaborierende Verhalten liegt, werden weitere Funktionalitäten, wie Kollisionsvermeidung, Pfadplanung oder Missionsmanagement als vorhanden angenommen und nicht detailliert betrachtet.

#### A. Zugrundeliegender Ablauf der Methode

Ein Auswahlprozess für kommunikative Aktivitäten kann allgemein als Entscheidungsprozess innerhalb eines Systems verstanden werden. Daher folgt das vorgestellte Konzept dem menschlichen Observe-Orient-Decide-Act [23] (Beobachten, Orientieren, Entscheiden, Handeln) Entscheidungszyklus. Der OODA-Zyklus beschreibt einen Prozess, den Instanzen autonomer mobiler Roboter durchlaufen, um sich an veränderte Situationen anzupassen und effektive Entscheidungen zu treffen.

##### 1) Observe (Beobachten)

In jedem Zeitschritt  $T$  nimmt der Roboter seine Umgebung wahr, indem er kontinuierlich Informationen über seine Umgebung und die Situation sammelt. Die Bestimmung der eigenen Position und das Detektieren von Umgebungsinformationen sind wesentliche Elemente in diesem Prozessschritt. Zusätzlich empfängt der Roboter Nachrichten von anderen Verbundteilnehmern mit einer Verzögerung  $\Delta T$ , welche sich durch die Latenz einer realen Kommunikation ergibt, und nutzt diese zur Aktualisierung seines Weltmodells.

##### 2) Orient (Orientieren)

In dieser Phase erfolgt die Aktualisierung und Verfeinerung des internen Weltmodells durch die neuen Umgebungsinformationen auf der Grundlage der Beobachtungsergebnisse. Dadurch wird eine genauere Darstellung der Umgebung und der aktuellen Situation ermöglicht. Gleichzeitig identifiziert und speichert der Roboter relevante Objekte, die als potenzielle Kommunikationsinhalte dienen können. Dieser Prozess ermöglicht es dem Roboter, die gesammelten Informationen effizient zu organisieren und für die nächste Entscheidungsphase vorzubereiten.

##### 3) Decide (Entscheiden)

Ausgehend von den gesammelten potentiellen Kommunikationsinhalten bewertet der Roboter mögliche kommunikative Handlungen, indem er die Kriterien *wann*, *was* und *an wen* analysiert. Diese Bewertung erfolgt jeweils unabhängig voneinander, wobei der einzelne Roboter die gesammelten Informationen hinsichtlich ihrer Relevanz, Dringlichkeit und der geeigneten Empfänger bewertet. Basierend auf dieser Bewertung entscheidet der Roboter, welche Nachrichten gesendet werden sollen und bereitet diese für die Übermittlung vor. Neben den kommunikativen Handlungen trifft der Roboter in

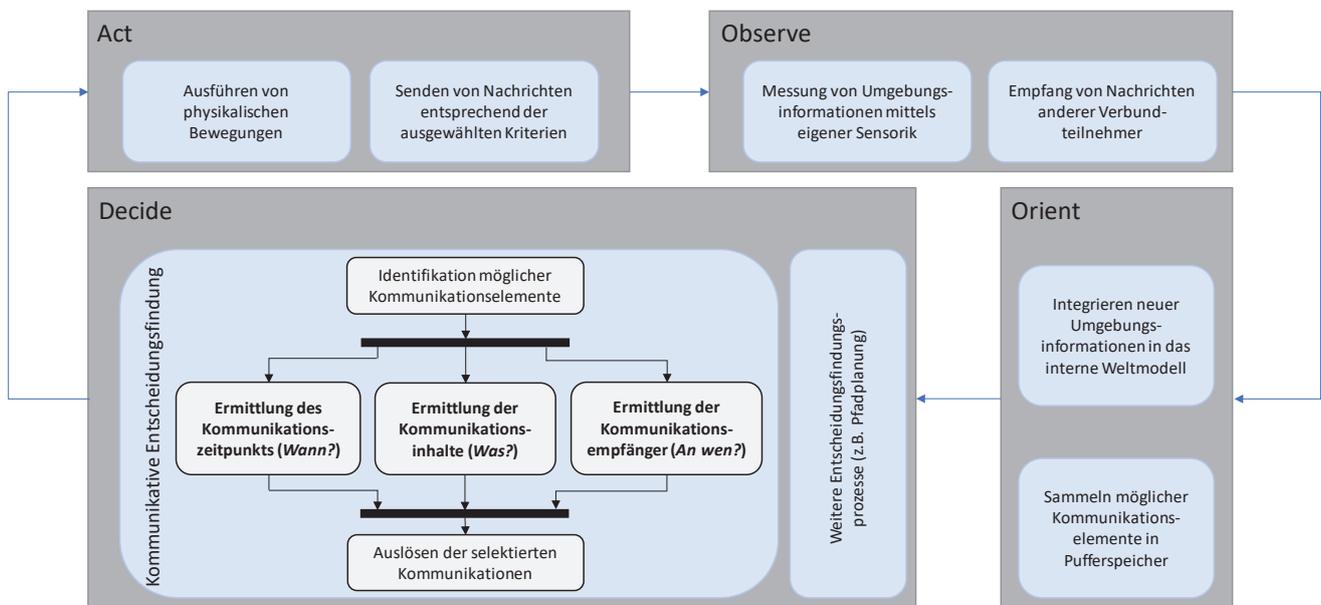


ABBILDUNG 2: PRINZIPIARDARSTELLUNG DER VERWENDETEN KOMMUNIKATIONSMETHODE

dieser Phase auch weitere Entscheidungen, die für seine Mission relevant sind, jedoch nicht im Fokus dieser Betrachtung stehen (z. B. Pfadplanung).

#### 4) Act (Handeln)

In der letzten Phase des OODA-Zyklus werden die selektierten Nachrichten übermittelt. Dies umfasst die aktive Kommunikation der ausgewählten Informationen an die entsprechenden Empfänger, um sicherzustellen, dass die relevanten Informationen zum richtigen Zeitpunkt an die richtigen Adressaten gelangen. Darüber hinaus passt der Roboter seine Position an und führt andere physische Aktionen durch, die durch vorangegangene Entscheidungen auf Basis der Informationen im Weltmodell getroffen wurden.

#### B. Ermittlung des Kommunikationszeitpunktes (*Wann?*)

Die Bestimmung des geeigneten Zeitpunkts für die Kommunikation ist entscheidend, um Netzüberlastungen zu vermeiden und die rechtzeitige Verbreitung kritischer Informationen zu gewährleisten. Jeder Roboter im Verbund verwendet eine Kombination aus lokalen und globalen Kriterien, um zu entscheiden, *wann* Informationen übertragen werden sollen.

Lokal bewertet ein Roboter seine unmittelbare Umgebung und seinen derzeitigen Status. Dabei werden kommunikative Handlungen bei ereignisgesteuerten Auslösern, wie zum Beispiel das Detektieren von neuen Umgebungsinformationen oder Statusänderungen in missionskritischen Komponenten, beim Auftreten des jeweiligen Auslösers sofort umgesetzt. Darüber hinaus sind zeitgesteuerte, vordefinierte Intervalle relevant, die einen gleichmäßigen Informationsfluss aufrechterhalten. Ebenso sind Zustandsänderungen zu berücksichtigen, die den Übergang des Roboters von einem Betriebszustand in einen anderen umfassen.

Global betrachtet ein Roboter die allgemeinen Netzwerkbedingungen, um eine effiziente Kommunikation sicherzustellen. Ein wichtiger Aspekt ist die Netzwerklast, wobei die aktuelle Kommunikationslast des bestehenden Netzwerks überwacht wird, um eine Überlastung zu verhindern. Daraus wird

der Zeitpunkt der Kommunikation bestimmt, zu dem die zulässige Netzwerklast nicht überschritten wird.

Durch die sorgfältige Abwägung dieser lokalen und globalen Kriterien können die Roboter im Netzwerk effizient und zuverlässig kommunizieren, was die Gesamtleistung und Koordination der autonomen Roboteroperationen deutlich verbessert.

#### C. Ermittlung des Kommunikationsinhaltes (*Was?*)

Die Entscheidung, welche Informationen geteilt werden sollen, ist wichtig, um die Relevanz und Nützlichkeit der verteilten Informationen zu maximieren. Dies erfordert die Priorisierung der Informationen basierend auf ihrem potenziellen Einfluss auf die kollektiven Ziele des Roboterverbunds. Jeder Roboter bewertet die Relevanz der Informationen anhand verschiedener Kriterien. Dazu gehört die Ausrichtung der Informationen an den übergeordneten Zielen und Aufgaben der Mission, um sicherzustellen, dass die Informationen den Missionszielen dienen. Weiterhin werden die Neuartigkeit und Einzigartigkeit der Informationen im genutzten Weltmodell von jedem Roboter berücksichtigt, um Redundanzen zu vermeiden. Durch diese sorgfältige Bewertung der zu übertragenden Informationen können die Roboter im Netzwerk sicherstellen, dass nur relevante und nützliche Informationen geteilt werden, welche für die effiziente Ausführung der vorgegebenen Mission des Fahrzeugverbundes erforderlich sind.

#### D. Ermittlung der Kommunikationsempfänger (*An wen?*)

Die Wahl der relevanten Empfänger stellt sicher, dass Informationen die richtigen Roboter im Verbund erreichen, was eine effiziente Ausführung der vorgegebenen Mission des Fahrzeugverbundes ermöglicht. Die Auswahl der Empfänger wird durch verschiedene Faktoren bestimmt. Zunächst spielt die Entfernung zwischen den kommunizierenden Robotern eine wichtige Rolle. Die Kommunikation mit nahegelegenen Robotern wird priorisiert, um die lokale Zusammenarbeit zu fördern. Weiterhin sind die Rolle und Funktion der Roboter

entscheidend. Bestimmte Roboter werden gezielt angesprochen, basierend auf ihren Rollen und aktuellen Aufgaben innerhalb der Mission. Darüber hinaus kann die Netzwerktopologie berücksichtigt werden, um die Routing- und Verteilungswege zu optimieren. Diese systematische Herangehensweise bei der Auswahl der Empfänger, gewährleistet eine effektive und effiziente Informationsverteilung innerhalb des Verbundes autonomer Roboter.

#### IV. IMPLEMENTIERUNG

In diesem Kapitel wird die Implementierung der dezentralen Informationsverteilung in einer Simulationsumgebung beschrieben. Die Implementierung erfolgte in Python und umfasst eine Vielzahl von Metriken zur Bewertung der Kommunikationsansätze, die den Missionseffizient und die Belastung des Kommunikationsnetzwerks quantifizieren.

##### A. Simulationsumgebung

Die Simulationsumgebung besteht aus einem 60x60 m großen Bereich, in dem sich die Roboter, mit Ausnahme von Bereichen, die durch Hindernisse blockiert sind, in alle Richtungen frei bewegen können. Diese Hindernisse, die als 1x1 m große Rechtecke modelliert sind, werden von den Robotern mit einer Reichweite von 3 m detektiert. Der verwendete Pfadplanungsalgorithmus wählt jeweils den kürzesten Weg, basierend auf den aktuell bekannten Hindernissen, und stellt einen Mindestabstand zu den Hindernissen sicher. Durch die kontinuierliche Aktualisierung der Umgebungsinformationen und den Austausch dieser Informationen zwischen den Robotern reagiert der Algorithmus dynamisch auf Veränderungen und passt den optimalen Pfad an. Die Anordnung der Hindernisse und die daraus entstehende Umgebung, welche für alle Simulationsdurchläufe verwendet wurde, ist in Abbildung 3 dargestellt. Um den Schwerpunkt auf die Kommunikation zu legen, wird die Zielposition in einer Ecke der Umgebung platziert, um sicherzustellen, dass möglichst viele Roboter von einer Informationsverteilung profitieren können. Die Zielposition ist als schwarzes Rechteck in Abbildung 3 dargestellt. Das Ziel ist für alle Roboter das gleiche, mit dem gemeinsamen Ziel, dieses so schnell wie möglich zu erreichen. Um unabhängig von der verwendeten Hardware vergleichbare Simulationsergebnisse zu generieren, wird die interne Simulationszeit und nicht die reale Zeit gemessen. Es wurden verschiedene Simulationskonfigurationen mit  $N = 2$  bis 10 Robotern durchgeführt, wobei die Startpositionen der Roboter zufällig generiert wurden. Alle Simulationskonfigurationen wurden jeweils zehnmal wiederholt und daraus arithmetische Mittelwerte für die Auswertung generiert. Der notwendige Python-Code mit einer Kurzanleitung zur Ausführung der Simulation ist online frei zugänglich<sup>1</sup>.

##### B. Betrachteter Anwendungsfall

In dem betrachteten Anwendungsfall versuchen mehrere Roboter eine vorgegebene Zielkoordinate in einer unbekannt Umgebung möglichst schnell zu erreichen. Die Roboter navigieren durch die unbekannt Umgebung in Richtung der Zielkoordinate, während sie kontinuierlich ihre Umgebung und etwaige Hindernisse wahrnehmen. Diese Wahrnehmungen können sie mit dem Verbund teilen, um das Ziel effizienter zu erreichen. Die Kommunikation erfolgt über ein simuliertes 4G-Mobilfunknetz, wobei die Informationen „über das Internet“ übertragen werden. Es wird davon ausgegangen,

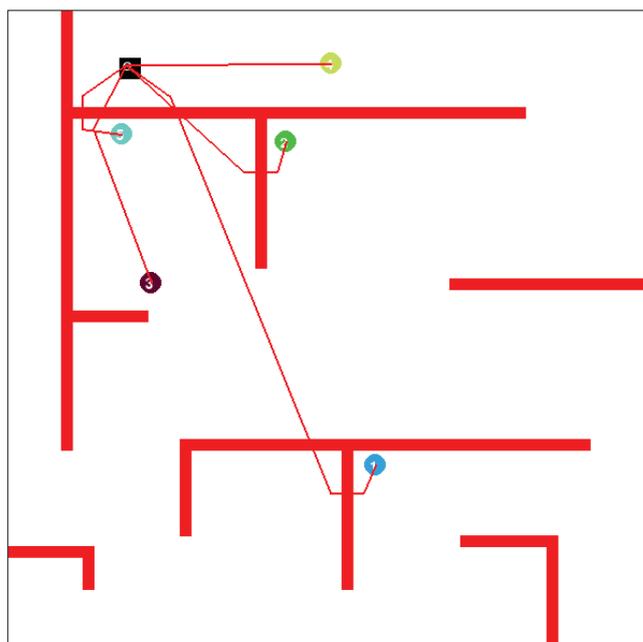


ABBILDUNG 3: DARSTELLUNG DER IN PYTHON IMPLEMENTIERTEN SIMULATION

dass die Internetverbindung zuverlässig ist (alle Nachrichten werden übermittelt), jedoch eine Verzögerung aufweist, die gleichmäßig zwischen 10 und 100 ms verteilt ist. Diese vereinfachte Python-Implementierung wurde gewählt, um den Fokus auf die verschiedenen Kommunikationsaspekte zu legen, ohne dabei den restriktiven Einschränkungen echter Robotikanwendungen ausgesetzt zu sein. Vielmehr ermöglicht diese Art der Implementierung eine präzise Kontrolle über die Kommunikationsaspekte und stellt somit eine geeignete Testumgebung dar, um die Methode unter Bedingungen zu evaluieren, wie sie in einem Framework wie dem Robot Operating System (ROS) schwer zu steuern sein könnte.

##### C. Metriken

Zur Quantifizierung und Bewertung der implementierten Szenarien der dezentralen Informationsverteilung wurden verschiedene Metriken verwendet, die in zwei Hauptgruppen unterteilt werden können: 1) Metriken Nutzens durch dezentrale Informationsverteilung und 2) Metriken zur Bewertung der Belastung des Kommunikationsnetzwerks.

###### 1) Missionsbezogene Metriken

Diese Metriken messen den Nutzen der dezentralen Informationsverteilung.

- **Benötigte Missionszeit:** Diese Metrik misst die Zeit, die die Roboter insgesamt benötigen, um ihre Mission abzuschließen. Je geringer die gemeinsam benötigte Zeit, je effizienter wurde die Mission ausgeführt.
- **Environmental Awareness Ratio (EAR):** Das EAR gibt das Verhältnis der bekannten Umgebungsinformationen eines Roboters zur Gesamtheit aller Umgebungsinformationen zu einem bestimmten Zeitpunkt an. Die Messung kann dabei auf einen bestimmten Bereich oder eine bestimmte Zeit begrenzt sein. Ein hohes

<sup>1</sup> [https://github.com/hsu-aut/RIVA\\_Decentralized\\_Communication](https://github.com/hsu-aut/RIVA_Decentralized_Communication)

EAR wirkt sich positiv auf die durchzuführende Mission aus.

## 2) *Kommunikationsbezogene Metriken*

Diese Metriken befassen sich mit den Kosten der Kommunikation einer dezentralen Informationsverteilung, d. h. der erzeugten Belastung des Kommunikationsnetzwerks.

- **Anzahl der übermittelten Nachrichten:** Eine Metrik zur Messung des Ressourcenbedarfs im Kommunikationsnetzwerk, indem die Menge der übertragenen Nachrichten gezählt wird. Je weniger Nachrichten versendet werden müssen, desto besser.
- **Nützlichkeit einer übermittelten Information:** Eine erhaltene Information wird als nützlich betrachtet, wenn sie zu einer sofortigen Änderung der eigenen, aktuellen Mission (z. B. des aktuellen Pfades) führt. Informationen, die keine sofortige Änderung bewirken, gelten als nicht nützlich. Zukünftige Auswirkungen einer erhaltenen Information werden aufgrund des nicht deterministischen Verhaltens nicht berücksichtigt.
- **Netzwerklast:** Die Netzwerklast wird vereinfacht über die Anzahl übermittelter Nachrichten im zeitlichen Verlauf ausgewertet. Eine geringe und vor allem gleichmäßige Netzwerkauslastung reduziert die Wahrscheinlichkeit verlorener Nachrichten und damit nicht übermittelter Informationen.

## D. Implementierte Szenarien

Die in Kapitel III vorgestellte Methode zur Erstellung einer dezentralen Informationsverteilung beinhaltet drei zu berücksichtigende Aspekte: *wann*, *was* und *an wen*. Der folgende Abschnitt beschreibt die spezifischen Implementierungen dieser Aspekte sowohl im Einzelnen als auch in einer integrierten Variante. Zusätzlich werden keine Kommunikation und eine uneingeschränkte Kommunikation als Vergleichswerte betrachtet.

### 1) *Keine Kommunikation*

Der einfachste Fall, bei dem keinerlei Kommunikation zwischen den autonomen Robotern stattfindet, dient als Basisvergleich für die anderen Kommunikationsstrategien.

### 2) *Zeitbasierte Kommunikation (Wann?)*

Mit der Methode *timing\_selective\_comm()* wird die Kommunikation basierend auf zeitlichen Kriterien gesteuert. Sender einer Nachricht können dabei immer nur verzögert neue Informationen versenden. Diese Strategie stellt sicher, dass Kommunikationsüberlastungen minimiert werden.

### 3) *Inhaltsbasierte Kommunikation (Was?)*

Die Methode *content\_selective\_comm()* bewertet bei mehreren möglichen Informationen die Relevanz hinsichtlich der durchzuführenden Mission. Im betrachteten Anwendungsfall werden Hindernisse, die sich dichter am Ziel befinden, höher bewertet und vorrangig ausgewählt, da sie für mehr Roboter potenziell relevant sind als weiter entfernte Hindernisse. Diese Methode sorgt dafür, dass nur wirklich nützliche Informationen im Netzwerk verbreitet werden.

### 4) *Empfängerbasierte Kommunikation (An wen?)*

Eine Bewertung und Selektion der Kommunikationsempfänger wird in der Methode *receiver\_selective\_comm()* durchgeführt. Dabei werden im betrachteten Anwendungsfall vorrangig Roboter ausgewählt, die sich weit vom Ziel entfernt be-

finden. Diese Strategie ermöglicht eine zielgerichtete Verbreitung von Informationen zu Verbundteilnehmern, die die Effizienz einer Mission am stärksten beeinflussen können.

### 5) *Integrierte Kommunikation (Wann, was und an wen?)*

Die Methode *integrated\_comm()* kombiniert zeitliche, inhaltliche und empfängerbasierte Kriterien. Diese integrierte Strategie selektiert Kommunikationsinhalte und -empfänger und steuert den Zeitpunkt zur Übermittlung von Informationen. Dadurch wird eine ausgewogene und effiziente Nutzung der Kommunikationsressourcen ermöglicht.

### 6) *Uneingeschränkte Kommunikation*

Mit der Methode *full\_comm()* wird ein weiterer Vergleichswert geschaffen, der jede neue Information sofort an alle Verbundteilnehmer sendet. Dieses Vorgehen erzeugt den höchsten Nutzen durch eine dezentrale Informationsverteilung bei gleichzeitig höchstem kommunikativen Aufwand.

## E. Resultate

Die Ergebnisse der implementierten Szenarien sind in Abbildung 4 für Simulationsdurchläufe mit jeweils zwei bis zehn Robotern dargestellt. Es zeigt sich, dass mit steigender Anzahl an Robotern die Anzahl der ausgetauschten Nachrichten stark zunimmt. Die Missionseffizienz (benötigte Zeit bis zum Erreichen des Ziels durch alle Roboter) kann um 35 % gesteigert werden, wenn Kommunikation eingesetzt wird (uneingeschränkte im Vergleich zu keiner Kommunikation). Dies geht jedoch mit einem sehr hohen Kommunikationsaufwand einher, da eine hohe Anzahl an Nachrichten ausgetauscht wird (bis zu 1.600 Nachrichten bei zehn Robotern). Die Berücksichtigung der relevanten Kommunikationsaspekte (*wann*, *was* und *an wen*) zeigen jeweils spezifische Verbesserungen. Diese umfassen die Reduktion der Nachrichtenanzahl, ein besseres Verhältnis von nützlichen zu nicht nützlichen Nachrichten sowie eine Verbesserung der Missionseffizienz. Insbesondere die inhaltsbasierte Kommunikation erzielt das beste Verhältnis von nützlichen zu nicht nützlichen Nachrichten.

Die integrierte Kommunikation übertrifft die Effizienz jedes einzelnen Aspekts und erzielt das beste Verhältnis aus Missionseffizienz und Kommunikationsaufwand. Bei Verwendung der integrierten Kommunikation wird die Effizienz um 27 % im Vergleich zu keiner Kommunikation gesteigert, was nur 8 % weniger als bei der vollständigen Kommunikation ist. Zudem wird diese Effizienzsteigerung mit einer deutlich reduzierten Anzahl an Nachrichten (um 34 %) und einem um 15 % höheren Prozentsatz nützlicher Mitteilungen erreicht. Eine vergleichende Darstellung der jeweiligen Missionseffizienz, gemessen in der durchschnittlich benötigten Zeit, wird in Abbildung 5 dargestellt.

Neben der reduzierten Anzahl an Nachrichten und dem erhöhten Verhältnis bei nützlichen Nachrichten verteilt der integrierte Ansatz die notwendigen Nachrichten am gleichmäßigsten über die Zeit und bedarf der geringsten Netzwerklast (vgl. Abbildung 6).

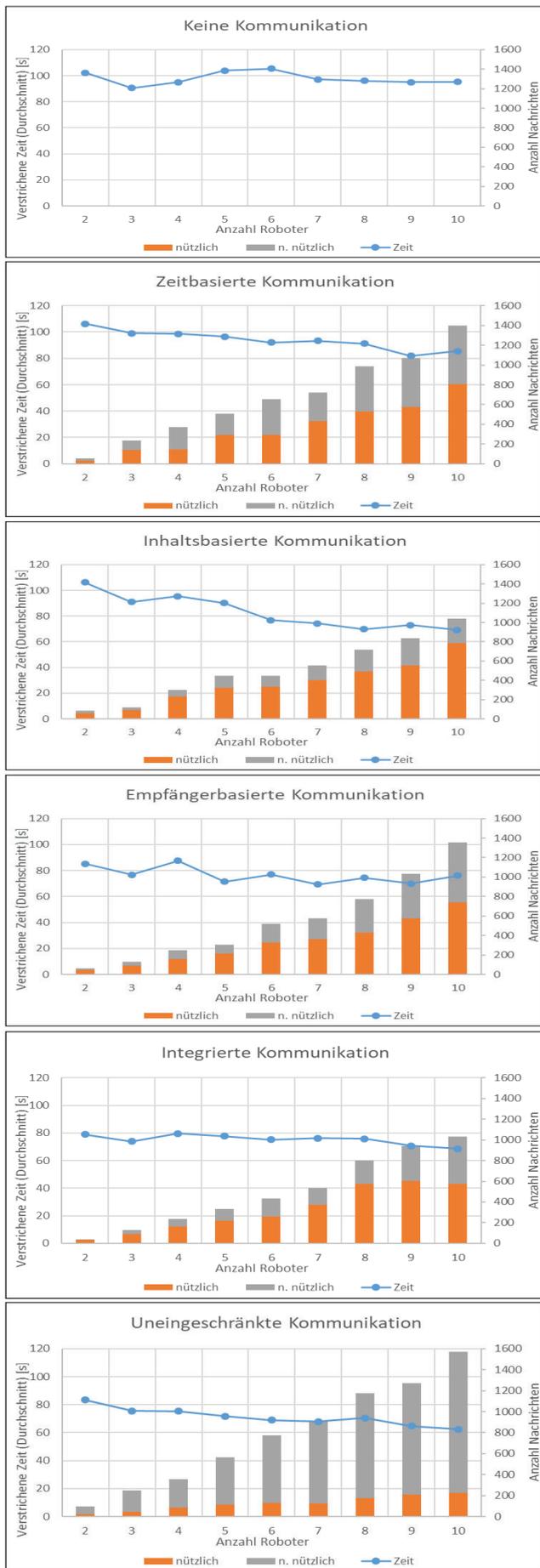


ABBILDUNG 4: SIMULATIONSERGEBNISSE

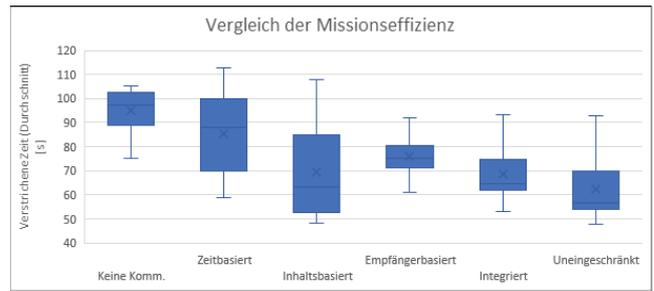


ABBILDUNG 5: MISSIONSEFFIZIENZ

Das EAR, dargestellt in Abbildung 7, erreicht nie den Wert eins, was auf den gewählten Anwendungsfall zurückzuführen ist, bei dem das Erreichen des Ziels wichtiger ist als das Entdecken der gesamten Karte. Das EAR bei vollständiger Kommunikation steigt schnell an, was zu einer hohen Netzwerklast insbesondere in einer frühen Phase der Mission führt. Die Bewertung der kommunikativen Aspekte *wann*, *was* und *an wen* zeigt qualitativ den gleichen Verlauf mit niedrigeren Werten über die Zeit, da nicht alle Informationen kommuniziert werden. Bei keiner Kommunikation steigt das EAR linear mit dem niedrigsten Endwert an, da jeder Roboter die Umgebung für sich selbst entdeckt. Das EAR bei integrierter Kommunikation zeigt trotz bester Missionseffizienz ein relativ schlechtes Ergebnis, was darauf hinweist, dass die reine Quantität von Informationen, welche das EAR misst, zur Bewertung unzureichend ist und die Qualität ebenfalls betrachtet werden muss.

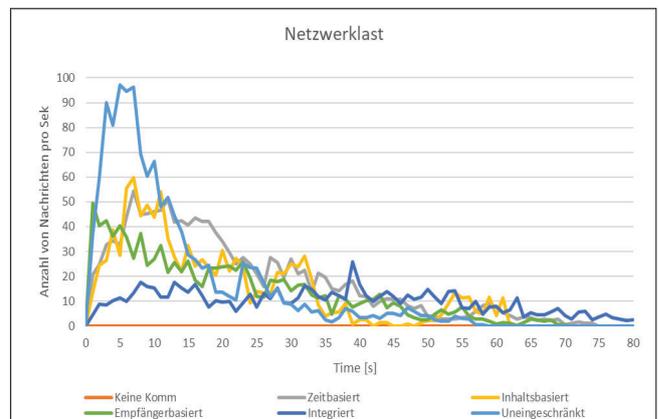


ABBILDUNG 6: NETZWERKLAST

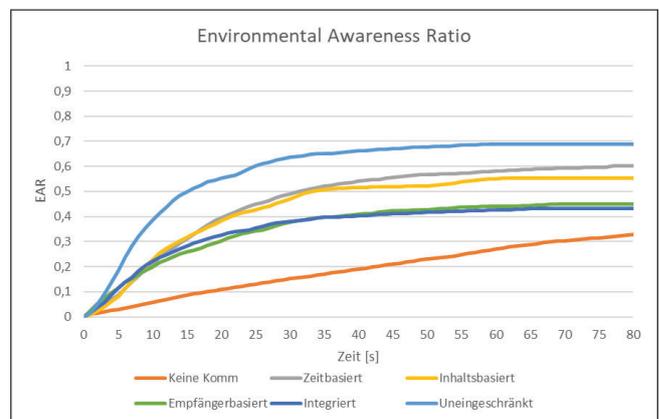


ABBILDUNG 7: ENVIRONMENTAL AWARENESS RATIO

## V. DISKUSSION

In diesem Beitrag wurde eine Methode entwickelt, die der dezentralen Informationsverteilung von kollaborierenden Robotern dient. Auf Basis des OODA-Zyklus entscheidet dabei jeder Roboter selbst, *wann* dieser *welche* Informationen und *mit wem* teilt. Zur Bewertung des Zeitpunktes, *wann* diese Informationen geteilt werden, werden sowohl ereignis- als auch zeitgesteuerte Auslöser berücksichtigt. Dabei erfolgt die Entscheidung sowohl unter Berücksichtigung eines gleichmäßigen Informationsflusses als auch einer vertretbaren Netzwerklast. Zur Bewertung des Kommunikationsinhaltes, *was* kommuniziert werden soll, wird die Nützlichkeit einer Nachricht im Hinblick auf das Missionsziel sowie die Neuartigkeit der Nachricht analysiert. Die Auswahl *an wen* kommuniziert wird, beruht zum einen auf der Entfernung zwischen den Robotern, zum anderen auf den jeweiligen Rollen und aktuellen Aufgaben der Roboter. Auf diese Weise wird sichergestellt, dass jeder Roboter auch nur für ihn relevante Informationen erhält.

Anhand der Ergebnisse aus der Implementierung in Kapitel IV wird deutlich, dass eine selektive Kommunikation durch einen autonomen Roboterverbund die Anzahl notwendiger Kommunikationen deutlich reduziert und gleichzeitig die Missionseffizienz nur geringfügig verringert. Durch die Reduktion der Nachrichtenanzahl sowie ein besseres Verhältnis aus nützlichen zu nicht nützlichen Nachrichten können autonome Roboterverbünde deutlich effizienter agieren. Darüber hinaus zeigt sich, dass der integrierte Ansatz für eine gleichmäßige Netzwerklast sorgt und gerade im Vergleich zur vollständigen Kommunikation erhebliche Vorteile bietet. Dies ist nicht zuletzt auch im Hinblick auf Katastropheneinsätze essenziell. Grundsätzlich adressiert die in diesem Beitrag vorgestellte Lösung vollständig die in der Einleitung definierten Hauptprobleme einer dezentralen Informationsverteilung im Gegensatz zur Vergleichsliteratur. Dies konnte durch verschiedene Metriken dargestellt und nachgewiesen werden. Einzig das verwendete EAR lässt sich zur Beurteilung der Effizienz der entwickelten Methode nur bedingt nutzen, da das Missionsziel in dem vorgestellten Szenario auf dem möglichst schnellen Erreichen eines Zeitpunktes und nicht auf die Erkundung einer Umgebung gelegt wurde.

Die hier vorgestellte Arbeit weist jedoch Limitationen auf. So wird beispielsweise ein idealisiertes System betrachtet, welches Kommunikationsfehler und -unterbrechungen vernachlässigt. Im Rahmen verschiedener Veröffentlichungen wird jedoch hervorgehoben, dass dies ein reales Problem ist, welches im Rahmen aufbauender Arbeiten ebenfalls in die Betrachtung inkludiert werden sollte. Es braucht folglich Mechanismen, die solche Fehler und Unterbrechungen erkennen und anschließend beheben können. Darüber hinaus wurde in der Simulation ein relativ simples Szenario entwickelt, welches sich darauf konzentriert, homogene Hindernisse zu erkennen und diese mit dem Fahrzeugverbund zu teilen. Beim Einsatz von autonomen Robotern in realen Umgebungen liegen jedoch eine Vielzahl heterogener Hindernisse – sogenannte Konfliktelemente – vor, die den Betrieb beeinflussen. Zu diesen Konfliktelementen zählen unter anderem andere Verkehrsteilnehmer, widrige Wettereinflüsse sowie Flugverbotszonen. Entsprechend sollte die Betrachtung dahingehend erweitert werden, dass heterogene Hindernisse ebenfalls inkludiert werden. Dies erfordert es anschließend auch eine entsprechende Sensor Fusion umzusetzen, um eine verlässige Navigation und Missionserfüllung auf Basis der konsolidierten Informationen zu gewährleisten. Zuletzt sollte auf Basis dieser

technischen Betrachtung auch analysiert werden, welche rechtlichen Rückschlüsse sich ziehen lassen. Eine vorherige Analyse im Bereich der interoperablen Kommunikation [34] zeigt, dass bisher kaum rechtlich bindende Vorgaben für die Kommunikation zwischen autonomen mobilen Robotern existieren. Dabei wurde gleichzeitig herausgestellt, dass sich dies unter gewissen Bedingungen anbieten würde. Daher empfiehlt es sich auch eine umfassende rechtliche Beurteilung im Bereich einer dezentrale Informationsverteilung durchzuführen, auf deren Basis dann die weitere Rechtsetzung geschehen kann.

## VI. ZUSAMMENFASSUNG UND AUSBLICK

In diesem Beitrag wurde eine Methode für eine dezentrale Informationsverteilung für autonome Roboter vorgestellt. Ausgehend von einer Literaturanalyse wurden bestehende Ansätze der Informationsverteilung bewertet und eine neue dezentrale Methode vorgestellt, die die relevanten Aspekte zur Bewertung möglicher Kommunikationen integriert. Die vorgeschlagene Methode berücksichtigt drei wesentliche Kriterien: den Zeitpunkt der Kommunikation, die Inhalte der zu übertragenden Nachrichten und die Auswahl der Empfänger. Durch diesen selektiven Ansatz kann die Anzahl notwendiger Kommunikationen deutlich reduziert und gleichzeitig die Missionseffizienz nur geringfügig verringert werden.

Die Methode basiert auf dem OODA-Zyklus (Observe, Orient, Decide, Act), bei dem jeder Roboter autonom entscheidet, *wann*, *was* und *mit wem* Informationen ausgetauscht werden. Dies umfasst sowohl ereignisgesteuerte als auch zeitgesteuerte Auslöser, um einen gleichmäßigen Informationsfluss und eine vertretbare Netzwerklast zu gewährleisten. Die Nützlichkeit und Neuartigkeit der Informationen werden analysiert, um sicherzustellen, dass nur relevante Informationen geteilt werden. Die Auswahl der Empfänger erfolgt basierend auf der Entfernung zwischen den Robotern, deren Rollen und aktuellen Aufgaben.

Die Simulationsergebnisse zeigen, dass die Anzahl notwendiger Kommunikationen deutlich reduziert und gleichzeitig die Missionseffizienz nur geringfügig verringert werden. Gleichzeitig wird das Verhältnis von nützlichen zu nicht nützlichen Nachrichten verbessert. Besonders in Katastropheneinsätzen mit eingeschränkten Netzwerken erweist sich diese Methode als vorteilhaft. Abschließend werden zukünftige Herausforderungen diskutiert, wie die Berücksichtigung realistischer Kommunikationsszenarien und die Notwendigkeit rechtlicher Rahmenbedingungen, um die praktische Anwendung dieser Methode weiter zu verbessern. Zukünftig sollen die Ergebnisse auf einen Anwendungsfall mit heterogenen mobilen Robotern angewendet werden.

## DANKSAGUNG

Diese Forschungsarbeit wird einerseits durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert (Projekt RIVA). dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert. Andererseits wird die Forschungsarbeit durch das Projekt Innovative luftgestützte urbane Mobilität (i-LUM) finanziert, welches vom Hamburger Landesforschungsfonds im Rahmen der Initiative Hamburg X gefördert wird.

## LITERATUR

- [1] S. Lee, D. Hong, J. Kim, D. Baek, and N. Chang, "Congestion-Aware Multi-Drone Delivery Routing Framework," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9384–9396, 2022, doi: 10.1109/TVT.2022.3179732.
- [2] A. Contini and A. Farinelli, "Coordination approaches for multi-item pickup and delivery in logistic scenarios," *Robotics and Autonomous Systems*, vol. 146, 2021.
- [3] K. Ahmad, H. Khujamatov, A. Lazarev, N. Usmanova, M. Alduailij, and M. Alduailij, "Internet of Things-Aided Intelligent Transport Systems in Smart Cities: Challenges, Opportunities, and Future," *Wireless Communications and Mobile Computing*, vol. 2023, pp. 1–28, 2023, doi: 10.1155/2023/7989079.
- [4] M. Noor-A-Rahim, Z. Liu, H. Lee, G. G. M. N. Ali, D. Pesch, and P. Xiao, "A Survey on Resource Allocation in Vehicular Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 701–721, 2022, doi: 10.1109/TITS.2020.3019322.
- [5] P. Wang, H. Deng, J. Zhang, L. Wang, M. Zhang, and Y. Li, "Model Predictive Control for Connected Vehicle Platoon Under Switching Communication Topology," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7817–7830, 2022, doi: 10.1109/TITS.2021.3073012.
- [6] F. Sherwani, M. M. Asad, and B. Ibrahim, "Collaborative Robots and Industrial Revolution 4.0 (IR 4.0)," in *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*, Karachi, Pakistan, 2020, pp. 1–5.
- [7] M. Knudsen and J. Kaivo-Oja, "Collaborative Robots: Frontiers of Current Literature," *Journal of Intelligent Systems: Theory and Applications*, pp. 13–20, 2020, doi: 10.38016/jista.682479.
- [8] Ī. Ūnal, O. Kabař, O. Eceođlu, and G. Moiceanu, "Adaptive Multi-Robot Communication System and Collision Avoidance Algorithm for Precision Agriculture," *Applied Sciences*, vol. 13, no. 15, p. 8602, 2023, doi: 10.3390/app13158602.
- [9] T. Elmokadem, "Distributed Coverage Control of Quadrotor Multi-UAV Systems for Precision Agriculture," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 251–256, 2019, doi: 10.1016/j.ifacol.2019.12.530.
- [10] M.-F. R. Lee and T.-W. Chien, "Artificial Intelligence and Internet of Things for Robotic Disaster Response," in *2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, Taipei, Taiwan, 2020, pp. 1–6.
- [11] M. N. Haddad, A. C. Santos, C. Duhamel, and A. A. Coco, "Intelligent Drone Swarms to Search for Victims in Post-Disaster Areas," *Sensors (Basel, Switzerland)*, vol. 23, no. 23, 2023, doi: 10.3390/s23239540.
- [12] Z. Li, G. Min, P. Ren, C. Luo, L. Zhao, and C. Luo, "Ubiquitous and Robust UxV Networks: Overviews, Solutions, Challenges, and Opportunities," *IEEE Network*, vol. 38, no. 2, pp. 26–34, 2024, doi: 10.1109/MNET.2024.3352691.
- [13] B. Mukherjee, M. F. Habib, and F. Dikbiyik, "Network adaptability from disaster disruptions and cascading failures," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 230–238, 2014, doi: 10.1109/MCOM.2014.6815917.
- [14] Department of Disaster Prevention and Mitigation, *Earthquake/Tsunami Victims Relief Efforts*.
- [15] T. Grove, J. Malsin, and R. Jones, *Earthquake Survivors' Time Is Running Out in Turkey and Syria*. [Online]. Available: [https://www.wsj.com/amp/articles/rescue-teams-race-against-time-to-find-survivors-11675840896?mod=Searchresults\\_pos14&page=5](https://www.wsj.com/amp/articles/rescue-teams-race-against-time-to-find-survivors-11675840896?mod=Searchresults_pos14&page=5)
- [16] F. Amigoni et al., "Online update of communication maps for exploring multirobot systems under connectivity constraints," in *Distributed Autonomous Robotic Systems: The 14th International Symposium*.
- [17] A. Soni, C. Dasannacharya, A. Gautam, V. S. Shekhawat, and S. Mohan, "Multi-Robot Unknown Area Exploration Using Frontier Trees," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Kyoto, Japan, 2022, pp. 9934–9941.
- [18] R. J. Marcotte, X. Wang, D. Mehta, and E. Olson, "Optimizing multi-robot communication under bandwidth constraints," *Autonomous Robots*, vol. 44, no. 1, pp. 43–55, 2020, doi: 10.1007/s10514-019-09849-0.
- [19] G. Best, O. M. Cliff, T. Patten, R. R. Mettu, and R. Fitch, "DecMCTS: Decentralized planning for multi-robot active perception," *The International Journal of Robotics Research*, vol. 38, 2-3, pp. 316–337, 2019, doi: 10.1177/0278364918755924.
- [20] P. Mazdin, M. Barciř, H. Hellwagner & B. Rinner, "Distributed task assignment in multi-robot systems based on information utility," *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, 2020, doi: 10.1109/CASE48305.2020.
- [21] B. Woosley, P. Dasgupta, J. G. Rogers, and J. Twigg, "Multi-robot information driven path planning under communication constraints," *Autonomous Robots*, vol. 44, no. 5, pp. 721–737, 2020, doi: 10.1007/s10514-019-09890-z.
- [22] M. Zager, C. Sieber and A. Fay, "Towards Semantic Interoperability: An Information Model for Autonomous Mobile Robots," *unpublished*.
- [23] P. Patr3n and D. M. Lane, "Adaptive mission planning: The embedded OODA loop," *Proceedings of the 3rd SEAS DTC Technical Conference*, 2008.
- [24] A. Hanna, P. L. G3tvsall, and Ekstr3m, M. & Bengtsson, K., "Requirements for designing and controlling autonomous collaborative robots system-an industrial case," *Advances in Transdisciplinary Engineering*, 2018.
- [25] M. Zager and A. Fay, "Design Principles for Distributed Context Modeling of Autonomous Systems," *IEEE Open J. Syst. Eng.*, vol. 1, pp. 179–189, 2023, doi: 10.1109/OJSE.2023.3342572.
- [26] F. A. Schiegg, D. Bischoff, and J. R. Krost & I. Llatser, "Analytical Performance Evaluation of the Collective Perception Service in IEEE 802.11p Networks," *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020.
- [27] G. Best, M. Forrai, R. R. Mettu, and R. Fitch, "Planning-Aware Communication for Decentralised Multi-Robot Coordination," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, QLD, 2018, pp. 1050–1057.
- [28] M. Barciř, A. Barciř, N. Tsiogkas, and H. Hellwagner, "Information Distribution in Multi-Robot Systems: Generic, Utility-Aware Optimization Middleware," *Frontiers in robotics and AI*, vol. 8, p. 685105, 2021, doi: 10.3389/frobt.2021.685105.
- [29] A. Alshehri, T. Miller, and L. Sonenberg, "Modeling communication of collaborative multiagent system under epistemic planning," *International Journal of Intelligent Systems*, vol. 36, no. 10, pp. 5959–5980, 2021, doi: 10.1002/int.22536.
- [30] N. Ruseno, C.-Y. Lin, and S.-C. Chang, "UAS Traffic Management Communications: The Legacy of ADS-B, New Establishment of Remote ID, or Leverage of ADS-B-Like Systems?," *Drones*, vol. 6, no. 3, p. 57, 2022, doi: 10.3390/drones6030057.
- [31] H. Ngo, H. Fang, and H. Wang, "Cooperative Perception With V2V Communication for Autonomous Vehicles," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 9, pp. 11122–11131, 2023, doi: 10.1109/TVT.2023.3264020.
- [32] A. Serra-Gomez, B. Brito, H. Zhu, J. J. Chung, and J. Alonso-Mora, "With Whom to Communicate: Learning Efficient Communication for Multi-Robot Collision Avoidance," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, 2020, pp. 11770–11776.
- [33] Z. Ma, Y. Luo, and J. Pan, "Learning Selective Communication for Multi-Agent Path Finding," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1455–1462, 2022, doi: 10.1109/LRA.2021.3139145.
- [34] M. Zager et al., "Interoperable communication for autonomous mobile robots: Enabling multimodal information exchange," *atp*, vol. 03, 2024.

# FROM MICROMECHANICS TO OPTIMAL SENSOR POSITIONING IN SHM APPLICATIONS – A CONCEPT APPLYING A SUBSTRUCTURE TECHNIQUE

Rasoul Najafi Koopas

*Chair of Solid Mechanics*

*Helmut Schmidt University /*

*University of the Federal Armed Forces*

Hamburg, Germany

najafikr@hsu-hh.de

Tim Suchan

*Chair of Structural Analysis*

*Helmut Schmidt University /*

*University of the Federal Armed Forces*

Hamburg, Germany

suchan@hsu-hh.de

Mathias Würkner

*Chair of Structural Analysis*

*Helmut Schmidt University /*

*University of the Federal Armed Forces*

Hamburg, Germany

mathias.wuerkner@hsu-hh.de

Natalie Rauter

*Chair of Solid Mechanics*

*Helmut Schmidt University /*

*University of the Federal Armed Forces*

Hamburg, Germany

natalie.rauter@hsu-hh.de

Wolfgang E. Weber

*Chair of Structural Analysis*

*Helmut Schmidt University /*

*University of the Federal Armed Forces*

Hamburg, Germany

wolfgang.weber@hsu-hh.de

Kathrin Welker

*Chair of Mathematical Optimization*

*TU Bergakademie Freiberg*

Freiberg, Germany

Kathrin.Welker@math.tu-freiberg.de

**Abstract**—In order to enhance Structural Health Monitoring of engineering structures, an appropriate modelling of the underlying structures as e.g. bridges or wings is necessary. Amongst other things this includes relevant (pre-)damages as cracks, delaminations, imperfect bonding, etc. which have to be incorporated at the so-called micro- or mesoscale of the structure. However, given the overall dimensions of typical engineering structures a discrete modelling of these (pre-)damages is not feasible at the macro-/structural scale. Thus, a scale-bridging is necessary to capture the structural behaviour. One promising approach to incorporate (pre-)damages at the microscale while maintaining a numerically manageable model of the overall structure is the sub-structure technique which will be used in the current project. Since a Structural Health Monitoring using the aforementioned numerical models strongly relies on useful measurement data it is of tremendous interest to determine the optimal number and the optimal position of the respective sensors. Hence, this topic is also addressed in the current contribution.

**Index Terms**—SHM, scale bridging, discrete modelling, sub-structure technique, optimal sensor positioning

## I. INTRODUCTION

Structural Health Monitoring (SHM) is a proper means to increase the service life of engineering structures such as bridges or wings. The continuous SHM allows for substantiated conclusions regarding the structure's current health state. With this knowledge an efficient maintenance work can be planned and performed. Beneath the increased service lifetime, the reliability of the structure is increased, too. However, this necessitates an appropriate modelling of the underlying structure.

Especially in civil engineering the health state of a structure, e.g. a bridge, is determined by regular inspections with fixed intervals. The findings are documented in so-called inspection reports and the respective bridge is rated according to the number and severity of the damages, if any. These written statements can be misinterpreted by readers other than the surveyors. Several approaches exist to automatically evaluate such inspection reports, e.g. [3]. However, the authors believe that a continuous monitoring of a structure and the automatic evaluation of the measurement data is more reliable. For this automatic evaluation and for deriving respective maintenance measures, a sophisticated (numerical) model of the bridge is needed. Amongst other things this includes relevant (pre-)damages as cracks, delaminations, imperfect bonding, etc. which have to be incorporated at the so-called microscale of the structure. Although much work has been conducted in recent years to refine material models or to develop numerical methods which are capable of capturing relevant effects at the microscale, there is still a lack in numerically efficient methods and in models incorporating multi-field mechanics.

As can be seen from FIGURE 1 a discrete modelling of the microscale phenomena such as cracks or the corrosion state is not feasible at the macroscale. This is due to the overall dimensions of typical (civil) engineering structures. Consequently, in order to capture the overall structural behavior a scale-bridging from the microscale to the macroscale is sought. Within this research project the sub-structure technique is applied to incorporate pre-damages, crack initiation and crack propagation at the microscale on the one hand, while

on the other hand ensuring a numerically manageable model of the structure at the macroscale. Once such a numerical model exists it allows substantiated conclusions and forecasts regarding the remaining service life with and without maintenance work. However, these conclusions and forecasts strongly depend on the measurement data obtained from the SHM. This necessitates research works to develop criteria for determining the optimal number and the optimal position of the sensors used for SHM.

In the precedent work [23] of some of the authors the general methodology for solving the aforementioned tasks was introduced. In the current contribution recent advances as well as the next scientific steps are described.

Based on the precedent motivation the outline of this contribution is as follows: Section II gives an overview of the discrete crack modelling at the micro- and mesoscale including crack growth due to traffic loads of a bridge. A methodology for incorporating this damage state at the macro-/structural scale is described in Sec. III. This scale-bridging allows for substantiated investigations regarding the optimal sensor positioning at a given bridge. The approach used within the project dtec.SHM is introduced in Sec. IV. Finally, a summary is provided in Sec. V.

## II. MICRO- AND MESOMECHANICAL MODELLING OF (PRE-)DAMAGED STRUCTURES

The adequate modelling of (pre-)damaged materials constituting engineering structures is of high scientific and practical interest. This includes modelling crack initiation and crack propagation [12], [13], [21] as well as a proper modelling of the (steel) reinforcing's corrosion state [20]. Depending on the aggregates used for the concrete mixture, stress corrosion cracking may occur [19]. An approach for adequately dealing with stress corrosion cracking phenomena is presented in [10].

Regarding the (pre-stressed) reinforcing of infrastructure buildings, such as e.g. bridges, several uncertainties exist. These are not limited to the current damage state but also extend to the actual cross-sectional area as well as the spatial distribution of the so-called interfacial transition zone (ITZ) between the reinforcing and the surrounding host material. Several approaches exist to describe the influence of the ITZ to the propagation of acoustic waves [24], [25] which are used in SHM applications.

In this work the focus for the numerical modelling of damage on the microstructural level lays on different fracture simulation methods to simulate and predict the crack propagation in a multiphase material such as concrete. Initially, four widely used fracture simulation methods are evaluated, comparing their computational expenses and implementation complexities within the Finite Element (FE) framework. This comprises the intrinsic Cohesive Zone Model (CZM) using zero-thickness cohesive interface elements (CIEs), the Standard Phase-Field Fracture Modelling (SPFM) approach, the Cohesive Phase-Field Modelling (CPFM) approach, and an innovative hybrid model. The hybrid approach combines the CPFM fracture method with an application of the CZM within the interface

zone. Within the context of concrete the Finite Element model is characterized by three phases, namely inclusions, matrix, and the interface zone. An example of a detailed modelling of an emerging crack utilizing different modelling approaches is provided in FIGURE 2.

The in-depth analysis of these modelling techniques reveals that the CPFM approach stands out as the most effective computational model in case the thickness of the interface zone is not significantly smaller than that of the other phases. In materials like concrete, which contain interfaces within their microstructure, the interface thickness is notably small when compared to other phases. This leads to the hybrid model standing as the most authentic finite element model, utilizing CIEs within the interface to simulate interface debonding.

A significant finding from this investigation is that the CPFM method is in agreement with the hybrid model when the interface zone thickness is not excessively small. This implies that the CPFM fracture methodology may serve as a unified fracture approach for multiphase materials, in case the interface zone's thickness is comparable to that of the other phases. For an indication of the interface's thickness it is referred to [16]. According to the measurement results presented therein, an interface thickness of 40% of the aggregate's radius might be a good first estimation.

Following this, the modelling approach on the microstructural level can serve as the basis for the incorporation of Neural Networks to predict the structural response at small scales for different damage scenarios. This is required in a subsequent step for the scale transition. Current literature demonstrates the effectiveness of new deep-learning-based surrogate models for fracture analysis in homogeneous and composite materials, e.g. [1], [2], [22]. Within this project a spatiotemporal UNet-based surrogate model is proposed, which is capable of predicting the homogenized stress-strain curve and final crack pattern of concrete microstructures. For this, a specially designed pipeline is developed to interpolate the FE data to a regular grid with high accuracy, eliminating the need for complex surrogate frameworks like Graph Neural Networks and making it possible to implement neural operator learning on the current model to increase accuracy for future development.

## III. SCALE-BRIDGING BY MEANS OF SUBSTRUCTURE TECHNIQUE

Bridge dimensions by several orders of magnitude differ from the dimensions of the bridge's microstructure and cracks, if any. As an example, it is referred to FIGURE 1 and FIGURE 2. Thus, modelling a bridge including its microstructure within one single FE model is not appropriate since the microstructural details would dramatically increase the FE model size. In practical applications such FE models can hardly be calculated due to a lack of computer memory space and computation time. A promising alternative is applying model order reduction techniques. Examples for physical subspace methods are modal subspace methods, the KRYLOV subspace method, and the so-called GUYAN reduction, cf. [4],

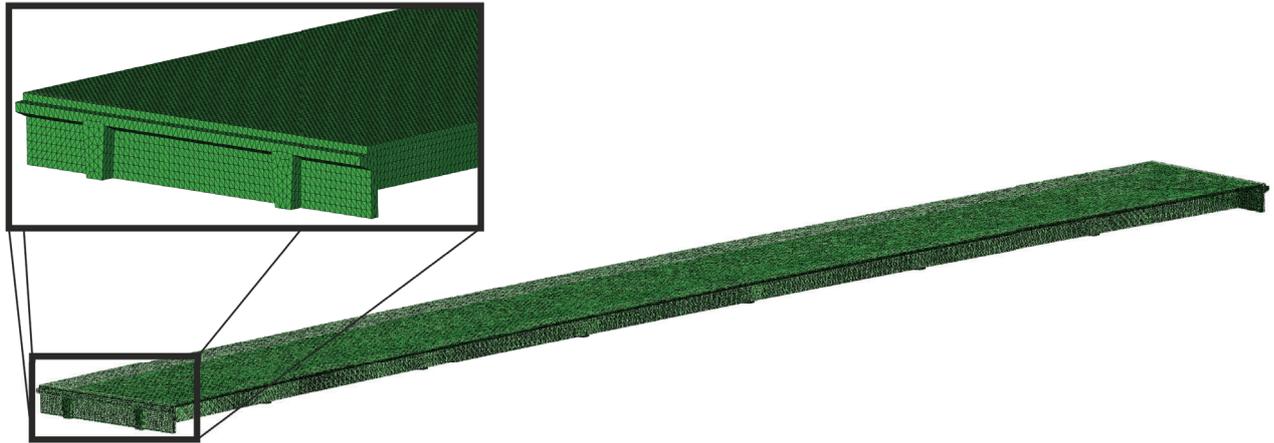


FIGURE 1. FE MODEL OF ONE SUPERSTRUCTURE OF THE BRIDGE STADER STRASSE, HAMBURG, GERMANY. CALCULATIONS ARE PERFORMED USING ABAQUS®

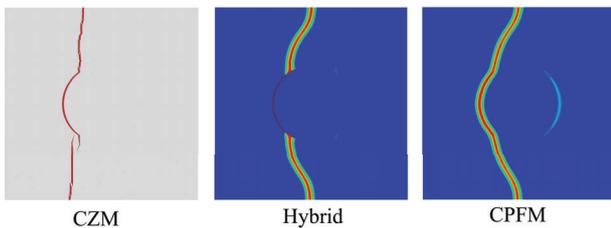


FIGURE 2. EXAMPLE FOR MODELLING OF A DISCRETE CRACK AT THE MICRO-SCALE WITH DIFFERENT FRACTURE SIMULATION METHODS, TAKEN FROM [14].

[5]. Other model order reduction techniques do not take into account the underlying physics. One example is the proper orthogonal decomposition [11]. However, main characteristics of the underlying mechanical system still need to be captured.

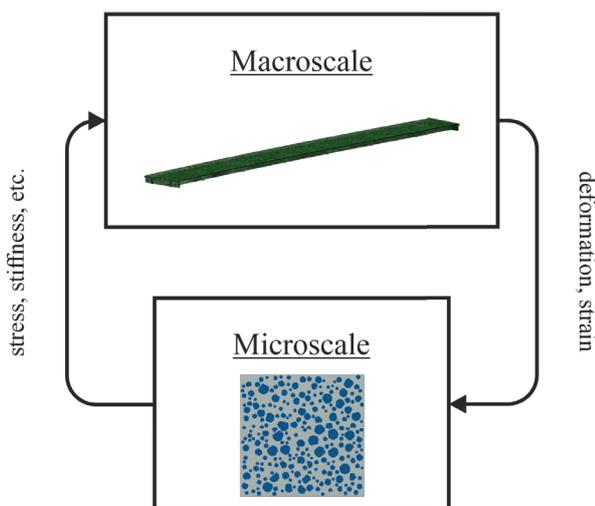


FIGURE 3. SKETCH OF SCALE BRIDGING USING SUBSTRUCTURE TECHNIQUE, INSPIRED FROM [6].

Another possibility to reduce the computational costs is to use scale transition approaches as e. g. the  $FE^2$  technique [17]. Herein, the basic concept is to generate two different FE models to deal with the different scales of the model: (i) the so-called microscale model, as it was described in Sec. II, takes into account the microstructural aspects as e. g. aggregates, ITZ, degradation (e. g. corrosion state), and cracks (if any); (ii) the so-called macroscale model describes the global, i. e. structural, behaviour and takes into account e. g. the loading conditions and the bearing conditions at the structural scale. In classical homogenization approaches both scales are related to each other in such a way that the average of the mechanical work on the microscale is equal to the mechanical work on the macroscale. This condition is known as the Hill-Mandel condition [7], [8].

In this contribution the focus is set on the  $FE^2$  approach, a sketch of which is given in FIGURE 3. Within this approach at each integration point of the macroscale model a microscale model has to be calculated. This microscale model consists of a so-called representative volume element (RVE), i. e. a volume element comprising the representative characteristics of the micromodel in the vicinity of the respective integration point at the macroscale. A detailed 3D FE model of a bridge will have at least several thousands of finite elements, see FIGURE 1. Thus, considering a RVE for each integration point may still lead to an unacceptable effort with respect to the computational cost. As an alternative, a subregion which is affected by the global loading conditions in a special manner can be identified and subsequently can be used to apply the  $FE^2$  method. If it is necessary to assign different RVEs to different integration points at the macroscale, the additional computational cost has to be taken into account. Only for the special case that linear elasticity is assumed and that each integration point is characterized by an unique RVE, six calculations of a microscale model should be sufficient. These calculations refer to six boundary value problems on the microscale which are characterized by linear independent

strain states on the macroscale (three extension strain states and three pure shear strain states). With these problems the effective (homogenized) elasticity tensor on the macroscale can be determined.

The boundary condition for the FE models at the microscale will be determined from an initially calculated macroscale model. For doing so, the deformation gradient or the strain are evaluated. After the microscale model(s) have been solved the results will be used to obtain macroscale quantities such as the tangent stiffness or the stresses by means of an averaging approach. Subsequently, these macroscale quantities are used to start a new iteration process on the macroscale model. Based on the evaluation of macroscale quantities such as displacements or velocities, in a next step optimal sensor positions can be identified from the numerical model. The respective procedure is described in the next section.

#### IV. OPTIMAL SENSOR POSITIONING AT THE STRUCTURAL SCALE

In engineering disciplines, such as e. g. aircraft engineering or wind engineering, the (micro-)sensors are optimised to be positioned within the respective material during the production process of the structure [9], [18]. However, in the project dtec.SHM focus is set on determining the remaining service life of already existing structures. Thus, sensors preferably need to be fixed at the surface of the material (i. e. steel-reinforced or pre-stressed concrete).

It is desirable to place the sensors in locations that provide the most information about damage and impact to the structure. It is not assumed that the location of the damage is known a priori. A non-negligible challenge in optimising sensor placement is the large number of possible locations. For practical applications the accessibility of the possible positions to place the sensors also needs to be taken into account. Determining the number and placement of sensors is essential for cost-efficient monitoring of the structure and for minimizing the amount of data. This is to be done in such a way that each sensor provides added value in terms of information about the condition of the structure. The problem can be solved numerically with the help of optimisation methods.

A first step to determine the sensor positions on a structure, e. g. a bridge or a building, is to set up a numerical model that should represent the structural behaviour of it. This can be performed by an optimisation approach that adjusts the local or global material parameters such that the measurement data and the results from numerical simulation are matched.

Then, one can employ optimal experimental design, which is a well-established technique to design and improve experimental setups in order to increase the accuracy of parameter identification in terms of a suitable optimality criterion in view of measurement errors. A problem, in which the parameter to be determined is the location of damage, needs to be modelled by the deviation from the material parameters. Measurements of the state can be made by a series of sensors. However, measurement errors have to be considered at some point. An

approach for automatically detecting sensor malfunctions from measurement data is presented in [15]. The overall aim is to optimise the position of the sensors such that the changes in measurements to changes in the material parameters are maximized. This can be realized with the help of the Fisher-information matrix. The approach for determining the optimal positioning of a sensor (network) is sketched in FIGURE 4.

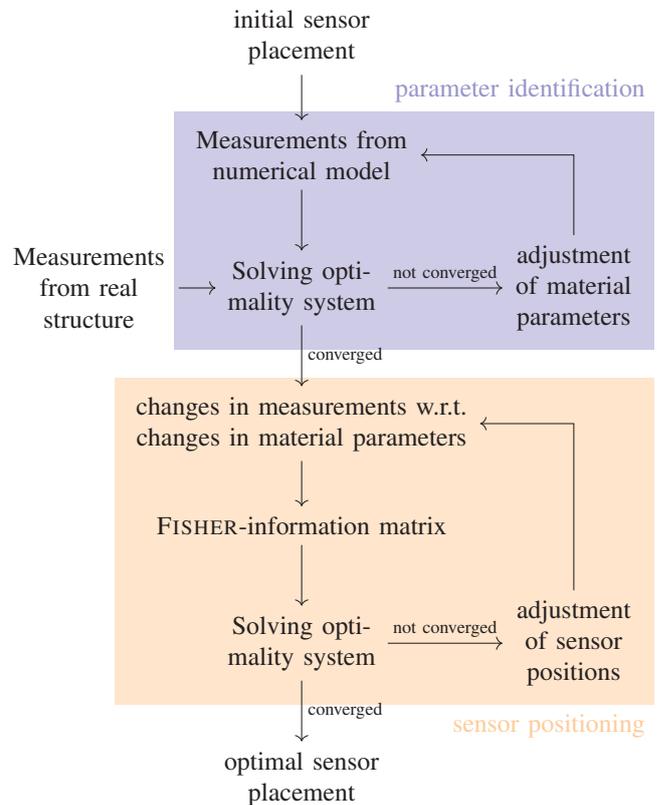


FIGURE 4. SKETCH OF THE CONCEPT AND PROCEDURES OF THE OPTIMAL DESIGN OF EXPERIMENTS AND THE PARAMETER ESTIMATION PROCEDURE TO DETERMINE OPTIMAL SENSOR POSITIONING.

#### V. SUMMARY

To deal with the challenges arising when planning, performing, and evaluating Structural Health Monitoring for e. g. bridges a three-step-approach is suggested. This approach includes (i) an adequate modelling at the microscale, (ii) a scale transition to the macroscale including numerical analyses at this macroscale, and (iii) based on the obtained findings an optimal positioning of the single sensors of a SHM sensor network.

At the microscale, several models have been developed by the authors. These models allow determining or capturing the effects of e. g. corrosion or crack initiation and propagation. With respect to the crack initiation and propagation, the Cohesive Phase-Field fracture model turned out to be a very effective approach for modelling these effects if the material combinations lead to comparable thicknesses of interfacial transition zones of the single constituents as it is the case for

e. g. bridges. The detailed knowledge of the material behaviour at the microscale is transferred to the macroscale by means of the FE<sup>2</sup> approach. To foster the numerical efficiency of the calculations at the macroscale, only those subregions of the structure are incorporated in the FE<sup>2</sup> approach which are highly loaded or/and which show existing or propagating cracks.

Furthermore, the procedure for the optimal positioning of sensors was described. Two optimisation problems are required. The first is used to build a numerical model which behaves similar to the structure and yields the associated material parameters. By perturbing these material parameters in a second optimisation problem, it is possible to maximize the effect of changes in material parameters—which are used to model the damage—on the changes in sensor signal. Thus, an optimal position for sensors can be obtained.

#### ACKNOWLEDGEMENT

This research work is funded by the project ‘SHM – Digitalisierung und Überwachung von Infrastrukturbauwerken’. This project is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

Natalie Rauter and Wolfgang E. Weber expressly acknowledge the financial support of the research work on this article within the Research Unit 3022 – Ultrasonic Monitoring of Fibre Metal Laminates Using Integrated Sensors by the German Research Foundation (Deutsche Forschungsgemeinschaft (DFG)).

#### REFERENCES

- [1] Z. Chang, Z. Wan, Y. Xu, E. Schlangen & B. Šavija, “Convolutional neural network for predicting crack pattern and stress-crack width curve of air-void structure in 3D printed concrete, *Engineering Fracture Mechanics*, vol. 271, pp. 108624, 2022, doi: 10.1016/j.engfracmech.2022.108624.
- [2] Y. Chen, T. Dodwell, T. Chuaqui & R. Butler, “Full-field prediction of stress and fracture patterns in composites using deep learning and self-attention”, *Engineering Fracture Mechanics*, vol. 286, pp. 109314, 2023, doi: 10.1016/j.engfracmech.2023.109314.
- [3] D.-C. Feng & W.-J. Wang & S. Mangaluthu & Z. Sun, “Condition Assessment of Highway Bridges Using Textual Data and Natural Language Processing- (NLP-) Based Machine Learning Models”, *Structural Control and Health Monitoring*, vol. 2023, pp. 9761154, 2023, doi: 10.1155/2023/9761154.
- [4] R. W. Freund, “Model reduction methods based on Krylov subspaces”, *Acta Numerica*, vol. 12(2), pp. 267–319, 2003, doi: 10.1017/S0962492902000120.
- [5] R. J. Guyan, “Reduction of stiffness and mass matrices”, *AIAA J*, vol. 3(2), pp. 380–380, 1965, doi: 10.2514/3.2874.
- [6] E. W. C. Coenen & V. G. Kouznetsova & M. G. D. Geers, “Multi-scale continuous-discontinuous framework for computational-homogenization-localization”, *Journal of the Mechanics and Physics of Solids*, vol. 60, pp. 1486–1507, 2012, doi: 10.1016/j.jmps.2012.04.002.
- [7] P. Suquet, “Elements of Homogenization for Inelastic Solid Mechanics”. In: E. Sanchez-Palencia & A. Zaoui (eds), “Homogenization Techniques for Composite Media”, Springer, Berlin, vol. 272, 1987, pp. 193–278, doi: 10.1007/3-540-17616-0\_15
- [8] R. Hill, “Elastic properties of reinforced solids: Some theoretical principles”, *Journal of the Mechanics and Physics of Solids*, vol. 11, pp. 357–372 doi: 10.1016/0022-5096(63)90036-X
- [9] J. N. Haus, L. Rittmeier, T. Roloff, A. Mikhaylenko, S. Bornemann, M. Sinapius, N. Rauter, W. Lang & A. Dietzel, “Micro-Oscillator as Integrable Sensor for Structure-Borne Ultrasound”, *Engineering Proceedings*, vol. 10(1), pp. 81, 2021, doi: 10.3390/ecsa-8-11313.
- [10] C. Kandekar, A. Ravikumar, D. Höche & W.E. Weber, “Mastering the complex time-scale interaction during Stress Corrosion Cracking phenomena through an advanced coupling scheme”, *Computer Methods in Applied Mechanics and Engineering*, vol. 428, pp. 117101, 2024, doi: 10.1016/j.cma.2024.117101.
- [11] G. Kerschen, J.-C. Golinval, A.F. Vakakis & L. A. Bergman, “The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview”, *Nonlinear Dynamics*, vol.41 (1–3), pp. 147–169, 2005, doi: 10.1007/s11071-005-2803-2.
- [12] R. N. Koopas, N. Rauter & R. Lammering, “Two-Dimensional Mesoscale Finite Element Modeling of Concrete Damage and Failure”, *Applied Sciences*, vol. 13, pp. 8971, 2023, doi: 10.3390/app13158971.
- [13] R. N. Koopas, N. Rauter & R. Lammering, “Two-dimensional meso-structural simulation of concrete fracture by the implementation of zero-thickness cohesive interface elements”, *PAMM*, vol. 22(1), pp. e202200020, 2023, doi: 10.1002/pamm.202200020.
- [14] R. N. Koopas, S. Rezaei, N. Rauter, R. Ostwald & R. Lammering, “Comparative analysis of phase-field and intrinsic cohesive zone models for fracture simulations in multiphase materials with interfaces: Investigation of the influence of the microstructure on the fracture properties”, preprint, 37 pp., doi: 10.48550/arXiv.2311.16826, 2024.
- [15] A. Liebert, W. Weber, S. Reif, B. Zimmering & O. Niggemann, “Anomaly Detection with Autoencoders as a Tool for Detecting Sensor Malfunctions”, 2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS), pp. 01–08, 2022, doi: 10.1109/ICPS51978.2022.9816908.
- [16] V. A. Matonis, “The interfacial stresses in particulate composite systems”, *Polymer Engineering and Science*, vol. 9(2), pp. 100–104, 1969, doi: 10.1002/pen.760090205.
- [17] L. Mester, V. Klempt, F. Wagner, S. Scheerer, S. Klarmann, I. Vakaliuk, M. Curbach, H.-G. Maas, S. Löhnert & S. Klinkel, “A Comparison of Multiscale Methods for the Modelling of Carbon-Reinforced Concrete Structures”, In: A. Ilki, D. Çavunt & Y. S. Çavunt (eds), “Building for the Future: Durable, Sustainable, Resilient”, fib Symposium 2023. Lecture Notes in Civil Engineering, vol. 350, Springer, Cham.
- [18] M. Rottmann, T. Roloff, N. Rauter, L. Rittmeier, M. Sinapius & W.E. Weber, “A numerical study on planar gradient acoustic impedance matching for guided ultrasonic wave detection”, *Journal of Vibration and Control*, vol. 30(3-4), pp. 697–710, 2024, doi: 10.1177/10775463221149764.
- [19] G. Schacht, M. Käding, G. Bolle & S. Marx, “Konzepte für die Bewertung von Brücken mit Spannungsrisskorrosionsgefahr”, *Beton- und Stahlbetonbau*, vol. 111(2), pp. 85–94, 2019, doi: 10.1002/best.201800087.
- [20] M. Shariati, W.E. Weber & D. Höche, “Parallel simulation of the POISSON-NERNST-PLANCK corrosion model with an algebraic flux correction method”, *Finite Elements in Analysis and Design*, vol. 206, pp. 103734, 2022, doi: 10.1016/j.finel.2022.103734.
- [21] T. Suchan, C. Kandekar, W.E. Weber & K. Welker, “Crack propagation in anisotropic brittle materials: From a phase-field model to a shape optimization approach”, *Engineering Fracture Mechanics*, vol. 303, pp. 110065, 2024, doi: 10.1016/j.engfracmech.2024.110065.
- [22] G. Wang, L. Zhang, S. Xuan, X. Fan, B. Fu, X. Xue & X. Yao, “An efficient surrogate model for damage forecasting of composite laminates based on deep learning”, *Composite Structures*, vol. 331, pp. 117863, 2024, doi: 10.1016/j.compstruct.2023.117863.
- [23] W. Weber, N. Rauter, R. Lammering & K. Welker, “Räumliche Auflösung des Schadenszustandes aus mechanischer und mathematischer Sicht”, In: D. Schulz et al. (eds), “dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg (Band 1)”, 2022, pp. 281–286, doi: 10.24405/14565.
- [24] W. Weber, U. Reuter & B.W. Zastrau, “An approach for exploring the dynamical behaviour of inhomogeneous structural inclusions under consideration of epistemic uncertainty”, *Multidiscipline Modeling in Materials and Structures*, vol. 9(1), pp. 81–99, 2013, doi: 10.1108/15736101311329179.
- [25] W. Weber & B. W. Zastrau, “Non-plane wave scattering from a single eccentric circular inclusion-Part I: SH waves”, *Journal of Theoretical and Applied Mechanics*, vol. 49(4), pp. 1183–1201, 2011.

## **Kapitel III**

### **Kompetenzen für die digitale Arbeitswelt (KoDiA)**

mit Beiträgen von

Deutsche Elektronen-Synchrotron DESY, Universität Hamburg

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Universität der Bundeswehr München

Universität Würzburg

## **Vorwort zum dtec.bw-Dachprojekt „Kompetenzen für die digitale Arbeitswelt (KoDiA)“**

Das Dachprojekt „Kompetenzen für die digitale Arbeitswelt (KODIA)“ unterstützt die zentrale Zielsetzung des Zentrums für Digitalisierungs- und Technologieforschung der Bundeswehr (dtec.bw), die digitale Souveränität Deutschlands und Europas zu stärken, aus verschiedenen Perspektiven. In einem interdisziplinären Forschungsansatz wollen wir einen Beitrag leisten, Verständnis, Bewältigung und Mitgestaltung einer beschleunigten Digitalisierung unter den Bedingungen der sicherheitspolitischen Zeitenwende zu fördern.

Hierzu wird in unterschiedlichen Einzelprojekten und in Kooperation mit zivilen und militärischen Stakeholdern geforscht:

Aus der Perspektive der Entwicklungspsychologie und der pädagogischen Psychologie wird im Projekt „AppLeMat – App für modulare Lerntherapie im Bereich Mathematik“ gemeinsam mit Lerntherapeuten der Praxis eine eigens entwickelte App zur Förderung mathematischer Kompetenzen erprobt. Die Forschung und Untersuchung von Motivation und Selbstregulation im Einsatz der App mit Schulkindern mit Rechenschwächen konnte durch wichtige Erkenntnisse aus der Praxisintegration erweitert und ergänzt werden. Im Folgenden kann dieses nachvollzogen werden.

Im Dachprojekt KoDiA ist auch die standortübergreifende „high performance computing Forschung“ an den Universitäten der Bundeswehr integriert, um eine Vielzahl von Projekten in den verschiedensten Disziplinen durch Bereitstellung von Rechenkapazität zu fördern. U. a. fand hierzu im Jahre 2023 ein Workshop zur erweiterten Bedarfserhebung statt, um die Möglichkeiten des high performance computing interessierten Forschenden zur Verfügung zu stellen (Einzelprojekt „hpc.bw – Competenceplatform for Software Efficiency and Supercomputing“).

Im mit dem Dachprojekt namensgleichen Einzelprojekt KoDiA („Kompetenzen für die digitale Arbeitswelt – Er-tüchtigung zur Digitalisierung“) wurden weiterhin in einem gestaltungsorientierten kontextualisierten Forschungsansatz unterschiedliche didaktisch methodische Instrumente entwickelt und erprobt. Im Vordergrund stand dabei die mündige und reflektierte Nutzung und Mitgestaltung einer digitalisierten Arbeits- und Lebenswelt. In den vergangenen zwei Jahren wurden dabei vermehrt Informationsgewinnung und -bewertung, Formen der Virtual und Augmented Reality sowie des Game-based Learning erforscht. Die daraus entstehenden „Dual-Use“-Ergebnisse kommen sowohl der beruflichen Bildung als auch zunehmend der Aus-, Fort- und Weiterbildung von Streitkräften zugute.

Zur mündigen Mitgestaltung brauchen Menschen belastbare und faktenbasierte Informationen und einen Referenzrahmen, um Menge und Vielfalt von diesen einordnen und bewerten zu können. Hierzu verfolgt KoDiA ein weiter entwickeltes Konzept Öffentlicher Wissenschaft, um Bürgerinnen und Bürgern aber auch Soldatinnen und Soldaten einen Referenzrahmen für Entscheidungen zugänglich zu machen.

Gemeinsam mit der Pädagogischen Hochschule Freiburg wurde darüber hinaus ein transdisziplinäres Forschungsnetzwerk aufgebaut, in dem zentrale Konstrukte aktueller wissenschaftlicher Forschung in einem multiperspektivischen Austausch mit dem Ziel der Förderung gemeinsamer Definitionen und Methoden untersucht werden.

Wir freuen uns auf Ihre interessierte Lektüre aber auch auf Ihre kritischen Rückmeldungen und Anregungen.

Hamburg, im November 2024

manuel schulz

Oberst i.G. Prof. Dr. habil. manuel schulz

Leiter Zentrum für technologiegestützte Bildung der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg und Leiter des dtec.bw-Dachprojektes KoDiA

# Entwicklung eines digitalen Mathematikförderprogramms für modulare Lerntherapie

Leon Skoba\*, Monika Daseking

*Entwicklungspsychologie und Pädagogische Psychologie  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
\*leon.skoba@hsu-hh.de*

Darius Endlich\*, Wolfgang Lenhard

*Lehrstuhl Psychologie IV  
Pädagogische Psychologie  
Universität Würzburg  
Würzburg, Deutschland*

**Kurzfassung** – Der vorliegende Beitrag skizziert die Entwicklung und den Aufbau eines digitalen Förderprogramms zur modularen Lerntherapie im Bereich Mathematik für Kinder mit Rechenschwierigkeiten. Darüber hinaus werden die Ergebnisse aus der Erprobung der ersten vorläufigen Version in den Jahrgangsstufen 2 bis 6 präsentiert. Basierend auf bisherigen empirischen Erkenntnissen wurden 15 verschiedene Aufgaben konzipiert, welche sich fünf zentralen Inhaltsbereichen der frühen mathematischen Kernkompetenzen zuordnen lassen: Mengenverständnis, Zählfertigkeiten und Zahlenkonzept, allgemeine Rechenfertigkeiten, visuell-räumliches Vorstellungsvermögen sowie Sachaufgaben und arithmetisches Faktenwissen. Des Weiteren wurde bei der Konzeption des Spiels auf eine adäquate methodische Darstellung sowie ein angemessenes Game-Design geachtet, welches sich an den spezifischen Bedürfnissen der Zielgruppe orientiert. Die Ergebnisse zeigen, dass das Spieldesign von den teilnehmenden Kindern positiv wahrgenommen wird. Zudem lassen sich mit der vorläufigen Version, bestehend aus drei Aufgaben, erste Hinweise auf die erwarteten Schwierigkeitsniveaus innerhalb der Aufgabenbereiche abbilden. Die ersten Ergebnisse erlauben zudem Aussagen bezüglich der potenziellen Umsetzbarkeit in der Praxis sowie der bevorstehenden Evaluation.

**Stichworte** – Computerbasierte Förderung, mathematische Kompetenzen, E-Learning, Lerntherapie

## I. EINLEITUNG

In zahlreichen Untersuchungen konnte die Wirksamkeit einzelner digitaler Anwendungen fürs Lernen nachgewiesen werden [1],[2],[3]. Auch für Kinder und Jugendliche mit Rechenschwierigkeiten liegen wissenschaftliche Belege vor, welche die Effektivität von digitalen Förderprogrammen nachweisen [4],[5], insbesondere für neuere Förderprogramme. Mittlerweile existiert eine Bandbreite zumeist kostenpflichtiger Mathematik-Fördersoftware: von einfachen Lernspielen mit einzelnen Übungen bis hin zu wissenschaftlich basierten Trainingskonzepten. Mögliche Rahmenbedingungen des Einsatzes der Förderprogramme sind mindestens genauso vielseitig und nicht immer klar definiert: ob zuhause, in der Schule oder in der Lerntherapie. Für die Anwender:innen (insbesondere Lerntherapeuten) lässt sich die Qualität der Förderprogramme auf den ersten Blick nicht erkennen. Daher wird häufig auf die ergänzende Anwendung der Förderprogramme verzichtet, da individuelle Erfahrungen gezeigt

haben, dass die Erwartungen im Hinblick auf Anwendung, Gestaltung oder inhaltliche Aspekte nicht erfüllt werden [6]. Zudem sollte ein Förderprogramm für die entsprechende Zielgruppe (Kinder und Jugendliche) so attraktiv gestaltet sein, dass das Interesse zur Nutzung des Programms geweckt wird – sowohl kurz- als auch langfristig [7]. Da sich digitale Anwendungen (u. a. Videospiele) für Kinder und Jugendliche im ständigen Wandel befinden und insbesondere die grafische Darstellung und spielerische Möglichkeiten kontinuierlich weiterentwickelten, könnten ältere Förderprogramme für Kinder und Jugendliche zudem weniger reizvoll gewesen sein. Zwar zeigten Benavides-Varela und Kollegen [4] in einer aktuellen Metaanalyse keinen signifikanten Vorteil von Förderprogrammen mit Videospielcharakter gegenüber einfachen digitalen Lernprogrammen ermitteln, jedoch vermuten die Autor:innen, dass sich dies in naher Zukunft durch die Weiterentwicklung der technologischen Möglichkeiten ändern könnte. Avila-Pesantez, Delgadillo und Rivera [8] kommen zu dem Schluss, dass Serious Games ohne ein klar umrissenes Spieldesign die festgelegten Förderziele allerdings nicht erreichen können. Die Autor:innen schlagen einen konkreten Ablauf bei der Entwicklung von Serious Games vor, um eine optimale und effiziente Lernumgebung zu ermöglichen. So müssen in den ersten Phasen unter anderem das Spielkonzept, Spielmechaniken, die Lerninhalte, -strategien und -ziele sowie die didaktische und pädagogische Umsetzung bei der Entwicklung mitberücksichtigt werden. Abschließend sollte mit Hilfe von mehreren Prototypen das Förderspiel auf verschiedene Aspekte wie z. B. Benutzerfreundlichkeit, technische Spielqualität und Erreichen von Lernzielen evaluiert werden.

Die Forschungsgruppe im Projekt AppLeMat – App für modulare Lerntherapie im Bereich Mathematik – hat sich zum Ziel gesetzt, ein digitales und wissenschaftlich fundiertes Mathematikförderprogramm in Form eines Tablet-basierten Spiels zu entwickeln, welches sowohl für die Lehrenden, die Lerntherapeuten, als auch für die Lernenden, insbesondere Kinder mit Rechenschwäche, ergänzend zur Lerntherapie eine förderliche Lernumgebung schafft. Auf dem Weg zur Veröffentlichung des Förderprogramms werden verschieden Studien mit vorläufigen Pilotversionen durchgeführt, um die Nutzerfreundlichkeit, Funktionalität, das inhaltliche Konzept und

Gütekriterien wie Aufgabenschwierigkeiten, Validität und Reliabilitätsmaße zu untersuchen, ebenso wie die Wirksamkeit des Programms.

## II. PROJEKTSKIZZE

Mit diesem Beitrag soll ein Einblick in den Entwicklungsprozess des Mathematikförderprogramms „Eckis Würfelkosmos – Abenteuer in der Zahlengalaxie“ gegeben werden: Zuerst wird sowohl das Spielkonzept als auch das Spieldesign dargestellt, mit dem primären Fokus auf die verschiedenen Inhaltsbereiche, die trainiert werden sollen. Anschließend sollen die Ergebnisse der Erprobung der ersten Pilotversion in Bezug auf Benutzerfreundlichkeit und Aufgabenschwierigkeiten dargestellt werden.

### A. Storyline

Das Spiel „Eckis Würfelkosmos – Abenteuer in der Zahlengalaxie“ ist für Kinder mit Rechenschwäche im Alter von acht bis zwölf Jahren konzipiert worden. Die Spielhandlung findet im Weltraum statt. Das Kind spielt eine:n Astronaut:in, die mit der Rakete eine Bruchlandung auf einem Würfel-Planeten erlebt. Auf dem Würfel-Planeten trifft der/die Spieler:in auf einen Nicht-Spieler-Charakter (non-playable character, NPC) Namens „Ecki“ – eine kleine Figur, die als Helfer das Kind dabei begleitet, die Rakete zu reparieren. Gemeinsam mit Ecki muss das Kind verschiedene Orte aufsuchen, weitere Helferfiguren kontaktieren und deren Missionen erledigen, um das Spiel erfolgreich abzuschließen und mit der Rakete nach Hause fliegen zu können. Zusammenfassend lässt sich die Spielmechanik wie folgt darstellen: Der oder die Nutzer:in bearbeitet Mathematikaufgaben, für die er anschließend eine Belohnung in Form von Würfeln und kleinen Gestaltungselementen erhält. Diese Würfel müssen aneinandergelagert werden, damit sich das Spieluniversum vergrößert, neue Inhalte freigeschaltet werden, sowie um in der Story voranzukommen. Sobald alle erworbenen Würfel verbraucht worden sind, müssen weitere Mathematikaufgaben bearbeitet werden, um wieder neue Würfel zu erhalten – der Spielzyklus wiederholt sich. Jede Aufgabe ist einem mathematischen Inhaltsbereich zugeordnet und sowohl grafisch als auch inhaltlich an das Weltraum-Schema angepasst. Dabei kommen neue Herausforderungen und Spielelemente dazu und die Aufgaben passen sich dem Schwierigkeitsniveau des Kindes an und variieren in den Inhaltsbereichen.

### B. Überblick über die Inhaltsbereiche

Es existiert eine Vielzahl an empirischen Untersuchungen darüber, welche Bereiche für das frühe Erlernen von Mathematikkompetenzen nützlich sind [9], [10], [11], [12], [13]. Zwischen den Autorinnen und Autoren gibt es Überschneidungen im Hinblick auf einzelne Faktoren (z. B. Zählfertigkeiten; vgl. Aunio und Räsänen [9] und Hirsch et al. [12], wobei sich die Zusammensetzung der Faktoren zum Teil stark unterscheidet, u. a. in Abhängigkeit vom Kontext. Fischer et al. [11] haben festgestellt, dass in den gängigsten deutschsprachigen Diagnostikinstrumenten teilweise andere Bereiche erhoben werden, als in den Förderprogrammen trainiert werden. Da Rechenschwäche kein einheitliches Störungsbild darstellt, empfehlen die Autor:innen, „dass theoriegeleitete Förderprogramme einem modularen Aufbau folgen, bei dem einzelne Bestandteile je nach Bedarf flexibel für die Förderung eingesetzt werden können“ [11, S.36].

Die Konstruktion der einzelner Trainingsbereiche orientierte sich an nationalen Bildungsstandards der Kultusministerkonferenz für das Fach Mathematik im Primarbereich [14]

sowie an lernpsychologischen Grundlagen und wissenschaftlichen Erkenntnissen zum Erlernen und Fördern von mathematischen Basisfertigkeiten und Rechengrundlagen – es ist eine Kombination aus curricularen und nicht-curricularen Aufgabenbereichen [15], die einem modularen Förderansatz folgt. Inspiriert wurde die Entwicklung von den beiden Förderprogrammen Rechengspiele mit Elfe und Mathis 1 und 2 [16], [17]. Die aktuelle Endversion des Spiels, die voraussichtlich im dritten Quartal des Jahres 2024 veröffentlicht wird, beinhaltet 15 Aufgaben, die theoriegeleitet in fünf Inhaltsbereiche eingeteilt werden können: (1) Mengenverständnis, (2) Zählfertigkeiten und Zahlenkonzept, (3) visuell-räumliches Vorstellungsvermögen, (4) allgemeine Rechenfertigkeiten und (5) Sachaufgaben und arithmetischen Faktenwissen.

#### 1) Mengenverständnis

Der Umgang mit Mengen stellt eine wichtige Voraussetzung für die mathematische Entwicklung in der Grundschule dar [18]. Darunter werden die Erfassung von Mengen und das Vergleichen von Mengen gefasst, ebenso wie das Aufteilen und Zusammenführen von Mengen sowie das Zuordnen der korrespondierenden arabischen Ziffer zu einer Mengendarstellung. Für den Inhaltsbereich *Mengenverständnis* sind drei Aufgaben angesetzt: (1) *Mengenvergleiche*, (2) *Zahlen zuordnen* und (3) *Simultane Mengenerfassung*.

In der Aufgabe *Mengenvergleiche* steht die Erfassung und der Vergleich der Mächtigkeit von Mengen im Vordergrund. Aufgabe der Kinder ist es, Kristalle von Asteroiden zu sammeln. Dabei soll stets jener Asteroid mit den meisten Kristallen ausgewählt werden. Um die Mengenvergleiche korrekt zu lösen, müssen nahe beieinander liegende Anzahlen unterschieden werden. Damit die Mengen visuell unmittelbare miteinander verglichen werden können, werden für die Darstellung von Mengen stets identische Objekte verwendet [19].

In der Aufgabe *Zahlen zuordnen* wird die Verbindung von Mengen und dazugehörigen Zahlen trainiert. Auf dem Bildschirm werden arabische Zahlen von 1 bis 10 dargeboten, die verschiedenen Mengendarstellungen zugeordnet werden sollen. Da die Anzahl der Punktmengen nicht direkt ersichtlich ist, müssen Kinder auf die Strategie des Abzählens zurückgreifen. Durch das wiederholte Abzählen von Punktmengen und der aktiven Zuordnung von arabischen Zahlen zu den Punktmengen wird das Verständnis gefördert, dass die zuletzt gezählte Zahl die Mächtigkeit der entsprechenden Menge repräsentiert [19]. Auf diese Weise erwerben die Kinder ein präzises Anzahlkonzept, das Kinder ohne Rechenschwäche üblicherweise bereits im Kindergarten entwickelt haben.

Mit den Aufgaben zur *Simultanen Mengenerfassung* soll die Fähigkeit trainiert werden, innerhalb kurzer Zeit (max. 2 Sekunden) eine geringe Menge an Objekten visuell zu erfassen, ohne dass auf die Zählfähigkeit zurückgegriffen werden kann [20, S.115ff]. Es werden Objekte in Form von Sternschnuppen präsentiert, die auf dem Bildschirm in Abhängigkeit von der Schwierigkeitsstufe zwischen einer und zwei Sekunden erscheinen werden; je nach Altersgruppe können bis zu sieben Elemente von Personen simultan erfasst werden (Anobile, Arrighi & Burr, 2019).

#### 2) Zählfertigkeiten und Zahlenkonzept

Das Ziel des Inhaltsbereichs *Zählfertigkeiten und Zahlenkonzept* „ist es, eine abstrakte mentale Repräsentation von Zahlen auszubilden und zu festigen.“ [21, S.81] Die Entwicklung eines Zahlenkonzepts sowie das Verständnis von Zahlen soll in diesem Bereich gefördert werden. Es ist, ähnlich wie

Mengenverständnis, ein grundlegender Bestandteil der Basiskompetenzen zum Erwerb der Rechenfertigkeiten, der sich in gängigsten Entwicklungsmodellen wiederfinden lässt [19], [22]. Der Inhaltsbereich Zählfertigkeiten und Zahlenkonzept wird durch folgende Aufgaben abgebildet: (1) *Zahlen verbinden*, (2) *Fehlende Zahl*, (3) *Zahlenstrahl* und (4) *Zählen*.

Bei der Aufgabe *Zahlen verbinden* werden die Kinder dazu aufgefordert Zahlen in einer vorgegebenen Reihenfolge zu verbinden. Es werden Punkte in Form von Sternen präsentiert, die nacheinander angetippt werden. Auf diese Weise entsteht am Ende ein fiktives Sternbild. Mit diesen Aufgaben soll das ordinale Zahlenverständnis trainiert werden, was zu den numerischen Kernkompetenzen zählt und in den gängigsten Förderprogrammen zu Rechenstörung wiederzufinden ist [11], [23].

Mit der Aufgabe *Zahlenstrahl* soll die zahlenräumliche Vorstellung trainiert und auf längere Sicht eine mentale Repräsentation eines relationalen Zahlenstrahls ausgebildet werden. Eine mentale Repräsentation des Zahlenstrahls dient den Kindern als eine wichtige Grundlage beim Rechnen [24], [25]. Kucian und Kollegen [24] konnten zeigen, dass sich bereits nach einem fünf-wöchigen Training mit einem Computerspiel zum Zahlenstrahl die Fehleranfälligkeit bei der Bearbeitung der Aufgabe verringert hatte, was für eine Entwicklung einer linearen Zahlenraumvorstellung spricht. Darüber hinaus konnte eine Verbesserung der Rechenfertigkeiten beobachtet werden.

In der Aufgabe *Fehlende Zahl* werden mehrere Zahlen gleichzeitig in einer bestimmten Reihenfolge visuell präsentiert und dabei mindestens eine Stelle ausgelassen (z. B. 2, 4, 6, X, 10). Die Aufgabe der Kinder ist es, eine Zahl zu bestimmen, die die abgebildete Zahlenreihe korrekt vervollständigt (in diesem Beispiel 8). Um diese Aufgabe bewältigen zu können, bedarf es sowohl eines ordinalen als auch eines relationalen Zahlenverständnisses. Während beim Zahlenstrahl die Linien und Markierungen dem Kind eine räumliche Orientierung geben, ist es bei dieser Aufgabe erforderlich die Intervallabstände zwischen den Zahlen zu bestimmen [26].

Die Fähigkeit des *Zählens* ist eine grundlegende Fertigkeit mit der auch die ersten Rechenoperationen, wie Addition und Subtraktion durchgeführt werden können [11], [19], [27]. Gleichzeitig müssen andere grundlegende Basisfertigkeiten vorab entwickelt sein, damit das Zählen fehlerfrei gelingt: Zum einen muss die Eins-zu-Eins-Zuordnung zwischen dem Zahlwort und Objekt sicher beherrscht werden, damit ein Abzählen von mehreren Dingen erfolgen kann. Zum anderen muss auch das Prinzip der Kardinalität verstanden werden – das Wissen darüber, dass die Zahl am Ende des Zählvorgangs für eine Größe steht.

### 3) Sachaufgaben und arithmetisches Faktenwissen

Mathematische Fähigkeiten beschränken sich nicht auf den isolierten Umgang mit Mengen und Zahlen. Darüber hinaus ist es wichtig, über mathematisches Faktenwissen zu verfügen und mathematische Fähigkeiten auf verschiedene Szenarien zu übertragen. Um diesem Umstand Rechnung zu tragen, widmet sich ein Inhaltsbereich den *Sachaufgaben und arithmetisches Faktenwissen*. Stern [28] stellt heraus, dass ein Training von Sachaufgaben dabei helfen kann, das mathematische Verständnis zu erweitern. Der Inhaltsbereich setzt sich zusammen aus Aufgaben zu den Bereichen (1) *Sachaufgaben* und (2) *Größenvergleich*.

Der Bereich *Sachaufgaben* dient dazu, mathematische Fertigkeiten in einem realistischen Alltagskontext anzuwenden. Im Gegensatz zu Textaufgaben, die zwar eine mathematisch sinnvolle Lösung erlauben, jedoch mitunter sehr realitätsfern konstruiert sind, ergeben Sachaufgaben stets ein sinnvolles episodisches Situationsmodell. Durch das Trainieren jener Anforderungen, die für eine Umsetzung von Texten in mathematische Situationsmodelle erforderlich sind, kann das mathematische Verständnis entschieden erweitert werden [28].

In der Aufgabe *Größenvergleich* werden zwei Mengenangaben einer physikalischen bzw. monetären Einheit präsentiert. Aufgabe der Kinder ist es, die beiden Mengen durch das Berühren der korrekten Operatoren ( $>$ ,  $<$ ,  $=$ ) zu vergleichen. Dabei lernen die Kinder, dass eine Zahl durch die Verbindung mit einer Maßeinheit eine Bedeutung erhält (obwohl 1.000 eine größere Zahl als 5 ist, sind 1.000 g weniger als 5 kg). Das Spiel orientiert sich am Themenbereich „Größen und Messen“ der nationalen Bildungsstandards für den Primarbereich [14].

### 4) Visuell-räumliches Vorstellungsvermögen

Obwohl in den S3-Leitlinien des AWMF [29] das visuell-räumliche Arbeitsgedächtnis als auch in den Bildungsstandards für das Fach Mathematik der KMK eines der zentralen inhaltlichen mathematischen Kompetenzen das räumliche Vorstellungsvermögen (Leitidee „Raum und Form“) [14] als ein wichtiger Förderbereich aufgeführt ist, so findet sich dieser kaum in den gängigsten deutschsprachigen Förderprogrammen für Kinder mit Rechenschwäche wieder [11]. Angestrebt sind 3 Aufgabenbereiche: (1) *Anzahl an Bausteinen*, (2) *Formen legen*, (3) *Visuelle Suche*.

In der Aufgabe *Anzahl an Bausteinen* werden einfach bis komplexe Figuren präsentiert, die aus einzelnen Würfeln aufgebaut sind; die Aufgabe des Kindes besteht darin die Anzahl der Würfel zu bestimmen. Hierfür müssen nicht sichtbare Würfel erschlossen und in der Vorstellung vervollständigt werden. Angelehnt ist die Aufgabe einerseits aus geläufigen Aufgabenstellungen aus dem Bereich der Intelligenzdiagnostik (Untertest „Dreiecke“ des K-ABC) [30], andererseits an den nationalen Bildungsstandards für den Primarbereich [14], um den Inhaltsbereich „Über räumliches Vorstellungsvermögen verfügen“ zu trainieren.

In der Aufgabe *Formen legen* sollen geometrische Formen aus Einzelementen zusammengefügt werden. Dafür werden eine Zielfigur und die einzelnen Bausteine der Zielfigur dargestellt. In einem Drag-und-Drop-Verfahren haben die Kinder die Möglichkeit, die einzelnen Elemente an vorgegebenen Positionen abzulegen und somit die Zielfigur nachzubauen. Den Empfehlungen der nationalen Bildungsstandard für den Primarbereich folgend [14], sollen die Kinder in der Aufgabe *Formen legen* die Inhaltsbereiche „Über räumliches Vorstellungsvermögen verfügen“, „Geometrische Figuren erkennen, benennen und darstellen“ sowie „Geometrische Abbildungen erkennen, benennen und darstellen“ trainieren.

Aufgaben zur *visuellen Suche* verfolgen den Ansatz der Förderung der visuelle Diskriminationsfähigkeiten und systemischen Suche. Innerhalb einer vorgegebenen Zeitspanne soll ein Kind entscheiden, welches der vorgegeben Auswahlmöglichkeiten sich vom Zielbild unterscheidet. Nur eine der Auswahlmöglichkeiten differenziert sich vom Zielbild, die Übrigen zeigen das Zielbild aus einer anderen Perspektive. Das Zielbild ähnelt einem Sonnensystem – eine Figur befindet sich in der Mitte, während andere Figuren (Kugeln verschiedener Farben und Positionen) auf ovalen Laufbahnen angebracht

sind. Die hier entwickelten Aufgaben zur visuellen Suche orientiert sich an der Mechanik von Testaufgaben zur Erfassung von visuell-räumlichen Fähigkeiten im Space-Untertest aus dem Developmental Test of Visual Perception – Third Edition [31].

#### 5) Allgemeine Rechenfertigkeiten

In dem Inhaltsbereich *allgemeine Rechenfertigkeiten* sollen die vier Grundrechenarten in drei Aufgaben trainiert werden: (1) *Addition und Subtraktion*, (2) *Multiplikation* und (3) *Division*.

Die drei Aufgaben in diesem Inhaltsbereich weisen eine Besonderheit auf, die sie von den übrigen Aufgaben aus den anderen Inhaltsbereichen unterscheidet. Die Aufgaben verfügen über eine spezielle Mechanik, die als Supporter-Level bezeichnet wird. In Supporter-Level werden verschiedene Rechenstrategien mit Hilfe von Beispielaufgaben dargestellt (z. B. Addition mit Hilfe von Mengendarstellung) [32], die die Kinder unterstützen sollen, die Aufgaben richtig zu lösen. Die Rechenstrategien werden dem Kind erst präsentiert, nachdem es eine vordefinierte Anzahl an Fehlern macht. Zusätzlich zu den Beispielaufgaben erhält das Kind von dem NPC „Ecki“ eine zusätzliche Hilfestellung zur richtigen Anwendung der Rechenstrategie. Erst nachdem mehrere Aufgaben im Supporter-Level richtig gelöst wurden, werden die Rechenstrategien ausgeblendet und das Kind kann die Einzelaufgaben ohne zusätzliche Unterstützung lösen. Die Supporter-Level dienen ausschließlich dem Verständniszweck und der Strategiefestigung, die Lösungswahrscheinlichkeit fließt nicht in die adaptive Schwierigkeitsanpassung der Aufgaben mit ein. So kann die Lerntherapeutik die Supporter-Level dafür nutzen, um mit dem Kind Lösungsstrategien zu festigen.

### C. Game Design

Eckis Würfelkosmos kann als ein Serious Game definiert werden: Neben den typischen unterhaltsamen Spielelementen wird als konkretes Ziel verfolgt, die Mathematikfähigkeiten von Kindern mit Rechenschwäche zu verbessern [33], [34]. Digitale Lernspiele, wie Serious Games, haben einen positiven Einfluss auf die Motivation zum Lernen und zum Üben von Mathematik [35], [36], [37], [38]. In neueren Studien konnte sogar ein stärkerer Einfluss als bei herkömmlichen Lernmethoden festgestellt werden, sofern eine ausgewogene Mischung aus (Video-)Spielelementen und pädagogischen Konzepten gegeben ist [39]. Digitale Förderprogramme können für Kinder mit Rechenschwierigkeiten eine geeignete Maßnahme darstellen, da diese häufig unter Motivationsproblemen leiden und sogar Ängste vor Mathematik entwickeln [40]. Um für diese spezifische Zielgruppe eine motivierende Lernumgebung zu kreieren, werden verschiedene Spielmechaniken eingebaut und das Design altersentsprechend gestaltet. Um die Grenzen dieser Arbeit nicht zu überschreiten, wird lediglich auf die zentralen motivationsfördernden Aspekte des Spiels eingegangen.

#### 1) Adaptive Lernumgebung.

Alle 15 Aufgaben sind in mehrere Schwierigkeitsstufen untergliedert. Je höher die Schwierigkeitsstufen, desto komplexer werden die Aufgabenanforderungen. In Abhängigkeit von der Lösungswahrscheinlichkeit passen sich die Schwierigkeitsstufen dem Fähigkeitsniveau des Kindes an. Damit das Kind in die nächsthöhere Stufe kommt, muss eine vordefinierte Anzahl an Aufgaben richtig gelöst werden. Umgekehrt, wenn eine gewisse Anzahl an Aufgaben falsch oder nicht in

der vorgegebenen Zeit gelöst wird, sinkt die Schwierigkeitsstufe und einfachere Aufgaben werden dem Kind präsentiert. Bevor ausreichend Daten vorliegen, um mit Hilfe der IRT die Adaptivität der Schwierigkeiten individuell anzupassen, wurde vorab durch das Projektteam ein allgemeiner Algorithmus festgelegt.

Neben der passiven adaptiven Schwierigkeitsanpassung innerhalb der Aufgabenbereiche während des Spielens besteht die Möglichkeit seitens der lerntherapeutischen Fachkraft unter Verwendung eines Backends manuell zwischen den Aufgaben eine Auswahl zu treffen. So kann eine erfahrene Lerntherapeutik entscheiden – in Abhängigkeit vom Fähigkeitsprofil des Kindes – welche Aufgaben intensiver und welche Aufgaben in bestimmten Lernphasen gar nicht geübt werden.

#### 2) Feedback und Belohnungssystem.

Die Feedbackqualität kann maßgeblich zum Erfolg oder Misserfolg von Spiel- und Lernerfahrung beitragen; insbesondere das elaborierte Feedback [41]. So muss auch die Art und Weise des Feedbacks in Abhängigkeit von der Dauer und dem Schwierigkeitsgrad des Lernspiels konzipiert werden [42]. Gezielte Hinweise und Instruktionen, die auf Vorwissen und bereits erlernten Strategien aufbauen, können das Lernen positiv beeinflussen [43]. „Eckis Würfelkosmos“ soll ergänzend zur Lerntherapie verwendet werden – so haben die Lerntherapeuten die Möglichkeit gemeinsam mit dem Kind die einzelnen Aufgaben zu bearbeiten und Strategien und Vorwissen aufzubauen, die anschließend im Lernspiel benötigt werden. Innerhalb einer Aufgabe wird nach einer vordefinierten Anzahl an Fehlern von dem NPC „Ecki“ ein aufgabenspezifischer Hinweis gegeben, der dem Kind helfen soll, die Einzelaufgabe richtig zu lösen. Je häufiger ein Kind Fehler innerhalb einer Aufgabe (z. B. Anzahl an Bausteinen) macht, desto spezifischer und konkreter werden die Hinweise; teilweise mit Hinweisen auf die benötigte Lösungsstrategie (= elaboriertes Feedback). Für die Aufgabe „Anzahl an Bausteinen“ erhalten Kinder, denen diese Aufgabe noch schwerfällt, bereits auf der 2. Ebene des automatischen Hilfsystems die Möglichkeit, die dargebotene Figur eigenständig zu rotieren. Auf diese Weise erhalten sie die Möglichkeit, die Figur von allen Seiten zu betrachten. Auf Ebene 3 des Hilfsystems werden die einzelnen Bausteine der Figur Zug um Zug abgezählt, sowohl visuell als auch auditiv.

Das Ziel des Förderprogramms ist es Kinder dazu zu motivieren, regelmäßig zu üben, um erlernte Inhalte zu automatisieren und zu festigen. Aus diesem Grund erhalten die Kinder nach jeder bearbeiteten Aufgabe eine Belohnung in Form von Bausteinen oder anderen Spielaccessoires, unabhängig von der Akkuratheit der bearbeiteten Einzelaufgaben. Zusätzlich bestärkt „Ecki“ das Üben des Kindes mit positiv formulierten Aussagen.

### III. ZIELESETZUNG UND FRAGESTELLUNGEN

Im Projektverlauf wurden verschiedene Versionen entworfen, erprobt und weiterentwickelt. Die erste Pilotversion (Stand vom November 2022) beinhaltet drei Aufgabenbereiche und eine rudimentäre Anlage der Spielwelt, in der die Spielenden mit den Belohnungen erste Elemente bauen konnten. Bei der Erprobung der ersten Pilotversion standen die Überprüfung der Benutzerfreundlichkeit für die entsprechende Zielgruppe sowie eine erste Beurteilung der Validität im Vordergrund. Die Validität des Förderprogramms soll anhand von Unterschieden in der Bearbeitungszeiten und Akkuratheit in Abhängigkeit von der Schwierigkeitsstufe innerhalb

des Spiels, sowie zwischen den Spielen und Spielern analysiert werden.

Folgende Fragestellungen lassen sich aus den Zielen der ersten Erprobung formulieren:

1. Benutzerfreundlichkeit – Wie benutzerfreundlich ist die Pilotversion für die Zielgruppe?

2. Überprüfung der Validität

2.1 Zeigen sich erwartete Unterschiede in der Bearbeitungszeit und in der Akkuratheit in Abhängigkeit von der Schwierigkeitsstufe innerhalb der Aufgaben?

2.2 Zeigen sich erwartete Unterschiede in der Bearbeitungszeit und in der Akkuratheit zwischen Kindern mit verschiedenen Leistungskonzepten?

2.3 Zeigen sich erwartete Unterschiede in der Akkuratheit zwischen den verschiedenen Aufgaben?

#### IV. METHODEN

##### A. Messinstrumente und Durchführung

###### 1) Erste Pilotversion

Die erste Pilotversion bestand aus drei Aufgaben: Mengenvergleiche (Spielname „Kristalle“), Simultane Mengenerfassung (Spielname: „Sternschnuppen“), Anzahl an Bausteinen (Spielname: „Blockwelten“). Jedes Kind bearbeitete 15 Einzelaufgaben pro Aufgabe. In Abhängigkeit von der Akkuratheit der Bearbeitung der Einzelaufgaben konnten die Kinder unterschiedlich hohe Schwierigkeitslevel erreichen. Jede Einzelaufgabe hat ein vorgegebenes Bearbeitungszeitlimit, auch in Abhängigkeit vom Schwierigkeitslevel. Wurde in der vorgegebenen Zeit keine Antwort gegeben, so wurde diese Einzelaufgabe als falsch gewertet und die nächste Einzelaufgabe präsentiert. Nach jeder vollständig absolvierten Aufgabe (= 15 Einzelaufgaben) erhielt das Kind eine festgelegte Anzahl an Bauelementen als Belohnung, die später im Baumodus eingesetzt werden konnten. Die Aufgabenbeschreibung wurde durch den Nicht-Spieler-Charakter (non-playable character, NPC) „Ecki“ erklärt. Nach Bearbeitung aller Einzelaufgaben aus allen drei Aufgaben konnte der Baumodus genutzt werden. In dem aktuellen Baumodus hat das Kind die Möglichkeit mit den erspielten Bausteinen und Gegenständen seine Welt individuell zu gestalten. Für jede Einzelaufgabe wurde deren Richtigkeit sowie deren Bearbeitungszeit erfasst. Das einmalige Absolvieren der drei Aufgaben inklusive Spielen im Baumodus nahm ca. 15 Minuten in Anspruch.

###### 2) Rechentest

Zur Messung der allgemeinen Rechenleistung wurden drei Untertests des Heidelberger Rechentests (HRT 1-4) [44] verwendet, der sich bereits in anderen Untersuchungen zur Evaluation von Förderprogrammen im Bereich Rechnen bewährt hatte [45]; die Güterkriterien sind zufriedenstellend (Retest-Reliabilität:  $r_{tt} = .87-.93$ ; kriteriumsbezogene Validität mit der Mathematikschulnote:  $r = -.67$ ) Folgende Untertests wurden durchgeführt: Addition, Subtraktion und Ergänzungsaufgaben. Das Leistungsniveau der Teilnehmenden kann anhand von normierten T-Werten bestimmt werden. Die Bearbeitung der Untertests inklusive Anleitung betrug ca. 10 Minuten.

###### 3) Fragebogen zur Benutzerfreundlichkeit

Die subjektive Einschätzung der Kinder zur Benutzerfreundlichkeit des Prototyps wurde mit einem selbstkonzipierten Fragebogen anhand einer bipolaren fünf-stufigen Likert-Skala, ergänzt durch unterstützende Verwendung von

Smileys, erhoben (1 = maximale Zustimmung; 5 = maximale Ablehnung). Dabei wurden die folgenden vier Inhaltsbereiche erfasst: Aufgabenschwierigkeit (z. B. „Wie schwierig oder wie leicht fandest du die erste Aufgabe „Kristalle“?“), Design (z. B. „Wie hat dir die Welt gefallen, die du selber bauen durftest?“), Handhabung/Verständlichkeit der App (z. B. „Wie gut hat dir Ecki die erste Aufgabe („Kristalle“) erklärt?“) und allgemeiner Spaßfaktor der App („Wie gerne würdest du die App noch mal spielen?“). Der Fragebogen wurde den Kindern vorgelesen und Zug um Zug gemeinsam bearbeitet, um sicherzustellen, dass alle Kinder die Fragen korrekt verstehen. Die Bearbeitungszeit des Fragebogens betrug ca. fünf Minuten.

###### 4) Ablauf

Die Daten der vorliegenden Studie wurden im Einzelsetting (für Kinder, die am lerntherapeutischen Angebot teilnahmen) sowie im Gruppensetting (in der Schulklasse) erhoben. Bei der Gruppentestung wurde die Klassenstufe als einzige demographische Variable erfasst, um den HRT auswerten zu können; darüber hinaus wurden keine weiteren personenbezogenen Daten erhoben (anonyme Untersuchung). Zu Beginn wurden die teilnehmenden Kinder von der jeweiligen Testleitung angeleitet, die drei Untertests des HRT zu bearbeiten. Anschließend wurde die Pilotversion auf einem Tablet präsentiert und die Kinder konnten eigenständig die Aufgabenbereiche (Reihenfolge der Aufgaben: Mengenvergleiche, simultane Mengenerfassung, Bausteine zählen) bearbeiten und anschließend im Baumodus ihre erspielten Bausteine verbauen. Zum Schluss wurde der Fragebogen zur Benutzerfreundlichkeit ausgefüllt. Die gesamte Erhebung dauerte ca. 30 Minuten.

##### B. Stichprobe

An der Erprobung der ersten Pilotversion nahmen insgesamt  $N = 29$  Kinder teil. Knapp ein Drittel der Kinder nahm am lerntherapeutischen Angebot teil (31 %). Der Großteil der Kinder besuchte die vierte Klasse (75,9 %), die übrigen Kinder sind über die zweite, dritte, fünfte und sechste Klasse verteilt. 20 Kinder besuchten die gleiche Grundschule in Unterfranken (Teilstichprobe 1), die restlichen Kinder wurden zufällig in lerntherapeutischen Praxen rekrutiert (Teilstichprobe 2). Für drei Kinder, die sich in Lerntherapie befanden, konnten keine T-Werte ermittelt werden, da diese eine höhere Klassenstufe besuchten, als in der Normierung des HRT vorgesehen ist. Sie wurden dennoch in die Studie einbezogen, da ihre erreichten Rohwerte in den Untertests deutlich niedriger waren als die Rohwerte der übrigen Kinder in Lerntherapie. Es kann daher angenommen werden, dass ihre Leistungen als unterdurchschnittlich eingeordnet werden können. Innerhalb der Gesamtstichprobe ( $W_{Add} = 0.99, p = 0.96$ ;  $W_{Sub} = 0.97, p = 0.52$ ;  $W_{Erg} = 0.95, p = 0.21$ ) sind die T-Werte für alle drei Untertest normalverteilt. Die Kinder, die an der Lerntherapie teilnahmen, wiesen im Durchschnitt niedrigere T-Werte in den Untertests des HRT 1-4 auf als die Kinder aus der Grundschule (Tabelle I). Der Mann-Whitney-U-Test ergab signifikante Unterschiede zwischen den beiden Gruppen für alle drei Untertests ( $z_{Add} = -3.42, p < .001$ ;  $z_{Sub} = -3.28, p < .001$ ;  $z_{Erg} = -3.29, p < .001$ ).

TABELLE I: MITTELWERTE UND STANDARDABWEICHUNGEN (IN KLAMMERN) FÜR DIE UNTERTESTS IM HRT (T-WERTE).

Gruppe	HRT 1-4 Untertests		
	Addition	Subtraktion	Ergänzungsaufgaben
Gesamtstichprobe <sup>1</sup> (N = 29)	50.46 (12.15)	53.58 (11.96)	53.38 (13.24)
Teilstichprobe 1 (N = 20)	54.9 (9.65)	57.85 (9.57)	57.7 (11.92)
Teilstichprobe 2 <sup>1</sup> (N = 9)	35.67 (6.53)	39.33 (7.12)	39.00 (3.63)

<sup>1</sup>Für drei Kinder konnten keine T-Werte ermittelt werden, da diese sich in höheren Klassen befanden, für die beim HRT 1-4 keine Normen vorliegen.

### C. Statistische Auswertung

Zu Beantwortung der Fragestellungen wurden deskriptive und inferenzstatistische Analysen durchgeführt. Im Hinblick auf die Untersuchung der Benutzerfreundlichkeit wurde sowohl die gesamte Stichprobe betrachtet als auch die Teilstichproben 1 und 2 separat. Aufgrund der Verletzung der Normalverteilungsannahme wurde zur Überprüfung der Fragestellung 1 der Wilcoxon-Test, ein nichtparametrischer Test, bei einer Stichprobegewählt. Dabei wird der hypothetische Median-Wert von 3 festgelegt, der innerhalb der fünf-stufigen-Likert-Skalierung einen mittleren Wert (neutraler Smiley) darstellt. Die Analyse erfolgt auf Item-Ebene, da die Items für die Untersuchung neu generiert wurden und sich nicht in Skalen zusammenfassen lassen.

In der deskriptiven Statistik zeichnete sich bereits ein deutlicher Unterschied in den Leistungen zwischen der Teilstichprobe 1 und 2 ab, so dass hier auch die Analyse zwischen den Gruppen im Hinblick auf die Bearbeitungszeiten und Akkuratheit verfolgt wurde. Die Bearbeitungszeiten für die Einzelaufgaben wurden vorab bereinigt: Bearbeitungszeiten für falsche Antworten wurden ausgeschlossen, ebenso wie Bearbeitungszeiten, die aufgrund technischer Probleme entstanden waren (z. B. Bearbeitungszeiten, die höher waren als die vorgegebenen Zeiten zum Beantworten einer Einzelaufgabe). Anschließend wurde für jede Aufgabe und für jede Schwierigkeitsstufe ein Mittelwert gebildet. Abschließend wurden die Bearbeitungszeiten mit Hilfe des natürlichen Logarithmus transformiert, um diese an Normalverteilung anzunähern und die Bedeutung von Ausreißerwerten zu reduzieren. Die Akkuratheit einer Schwierigkeitsstufe innerhalb einer Aufgabe lässt sich durch den Quotienten der Anzahl der korrekt gelösten Einzelaufgaben und der Gesamtanzahl der in der jeweiligen Schwierigkeitsstufe bearbeiteten Einzelaufgaben ermitteln.

Hinsichtlich der transformierten Bearbeitungszeiten lässt sich eine Normalverteilung für die einzelnen Schwierigkeitsstufen annehmen, während es bei den Akkuratheitswerten nicht der Fall ist. Daher wurde für die Überprüfung der Fragestellungen 2.1 und 2.2 mit transformierten Bearbeitungszeiten eine zweifaktorielle Varianzanalyse mit Messwiederholungen gerechnet, während für die Akkuratheit der Friedman-Test für abhängige Stichproben zum Einsatz kam. Da sich alle drei Aufgaben in ihrer Bearbeitungsart unterscheiden, erfolgt die inferenzstatistische Überprüfung der Fragestellung 2.3 lediglich für die Akkuratheit, da mögliche Unterschiede in den Bearbeitungszeiten auch auf die jeweilige Bearbeitungsart zurückzuführen sind. Für alle durchgeführten Tests innerhalb einer Fragestellung werden die Signifikanzen nach der Sidak-Korrektur angepasst.

## V. ERGEBNISSE

### A. Benutzerfreundlichkeit

#### 1) Deskriptive Statistiken

Fast alle Fragen wurden im Durchschnitt mit der höchsten Stufe bewertet; nur die Frage zu Schwierigkeitsempfindung für die Aufgabe Blockwelten wurde mit der zweithöchsten Stufe bewertet ( $M = 2.1$ ,  $SD = 0.72$ ). Nur zwei Kinder haben bei jeweils einer Frage eine negative Bewertung abgegeben (beide Kinder in Lerntherapie-Teilstichprobe 2); alle anderen Kinder beantworteten die Fragen positiv bis neutral. Zwischen den beiden Teilstichproben waren die Unterschiede der Mittelwerte marginal, so dass die interferenzstatistische Überprüfung anhand der Gesamtstichprobe stattfand.

#### 2) Inferenzstatistische Analysen

Der Wilcoxon-Test bei einer Stichprobe ergab für alle neun Items signifikante Unterschiede zu dem hypothetischen Medianwert von 3 (= neutrale Bewertung) in Richtung 1 (= maximale Zustimmung bzw. positive Bewertung). Somit wird die Pilotversion von den Testpersonen in der Wahrnehmung der Schwierigkeiten, dem Design, der Handhabung der Aufgaben und im Spaßfaktor positiv wahrgenommen.

### B. Bearbeitungszeiten & Akkuratheit

In der Tabelle II (Anhang) sind die durchschnittlichen Bearbeitungszeiten (transformiert und nicht transformiert) sowie die Akkuratheitswerte für alle Levelstrukturen der drei Aufgaben abgebildet, sowohl für die Gesamtstichprobe als auch für die beiden Teilstichproben.

#### 1) Deskriptive Beschreibung – Schwierigkeitsstufen

Es lässt sich ein Anstieg der Mittelwerte der Bearbeitungszeiten für die gesamte Stichprobe innerhalb der Aufgabe Kristalle sowie der Aufgabe Blockwelten beobachten. Bei der Betrachtung der beiden Aufgaben für die jeweiligen Teilstichproben bleibt der lineare Anstieg nur bei der Aufgabe Blockwelten bestehen; bei der Aufgabe Kristalle sinkt bei Kindern in Lerntherapie die Bearbeitungszeit von der ersten Schwierigkeitsstufe ( $M = 1603.8$ ,  $SD = 497.2$ ) auf die zweite Schwierigkeitsstufe ( $M = 1461.3$ ,  $SD = 543.3$ ), wohingegen bei der Teilstichprobe 1 die Bearbeitungszeit von der zweiten Schwierigkeitsstufe ( $M = 1810.48$ ,  $SD = 849.8$ ) auf die dritte Schwierigkeitsstufe ( $M = 1781.55$ ,  $SD = 1068.7$ ) abnimmt. In der Aufgabe Sternschnuppen verändert sich innerhalb der Gesamtstichprobe und in der Teilstichprobe 1 die durchschnittliche Bearbeitungszeit über die ersten beiden Schwierigkeitsstufen kaum; bei der dritten Schwierigkeitsstufe nimmt sie ab. Bei Kindern in Lerntherapie ist die Abnahme der Bearbeitungszeit über alle drei Schwierigkeitsstufen zu beobachten. Für die Akkuratheit lassen sich innerhalb einer Aufgabe kaum einheitliche Veränderungen erkennen; nur bei der Aufgabe Blockwelten zeigen Kinder in Lerntherapie mit zunehmender Schwierigkeit eine kontinuierliche Abnahme an Akkuratheit ( $M_{LVL1} = 0.962$ ,  $M_{LVL2} = 0.827$ ,  $M_{LVL3} = 0.645$ ). Bei den Kindern in der Schulklasse gibt es keine einheitliche Abnahme der Akkuratheit in der Aufgabe Blockwelten zu beobachten ( $M_{LVL1} = 0.842$ ,  $M_{LVL2} = 0.919$ ,  $M_{LVL3} = 0.861$ ). Jedoch lässt sich für alle Stichproben eine Tendenz feststellen: In der Akkuratheit lassen sich Unterschiede zwischen den Aufgaben beobachten; innerhalb der Gesamtstichprobe hat die Aufgabe Kristalle die höchsten Akkuratheitswerte, die Aufgabe Sternschnuppen etwas niedrigere und die Aufgabe Blockwelten die niedrigsten Werte.

## 2) Deskriptive Beschreibung – Stichprobenvergleich

In fast allen Aufgaben und Schwierigkeitsstufen weisen Kinder aus der Lerntherapie auf deskriptiver Ebene höhere Bearbeitungszeiten auf als Kinder aus der regulären Schulklasse. Ausnahme bildet die zweite Schwierigkeitsstufe der Aufgabe Kristalle; hier benötigen die Kinder aus der Schulklasse im Durchschnitt deskriptiv länger ( $M = 1810.48$ ,  $SD = 849.8$ ), als Kinder, die sich in Lerntherapie befinden ( $M = 1461.3$ ,  $SD = 543.3$ ). Bei der Betrachtung der Akkuratheitswerte haben Kinder in der Teilstichprobe 1 niedrigere Akkuratheitswerte für alle Schwierigkeitsstufen in der Aufgabe Sternschnuppe: Das bedeutet, dass sie im Durchschnitt in der Aufgabe mehr Fehler gemacht haben als Kinder in der Lerntherapie. Die Kinder in Lerntherapie haben außerdem in allen Aufgaben auf der ersten Schwierigkeitsstufe im Durchschnitt deskriptiv höhere Akkuratheitswerte (mehr Einzelaufgaben innerhalb einer Schwierigkeitsstufe gelöst) als Kinder aus der regulären Schulklasse.

## 3) Inferenzstatistische Analysen – Bearbeitungszeiten

Eine zweifaktorielle Varianzanalyse mit Messwiederholung mit dem Innersubjektfaktor Schwierigkeitsstufe und dem Zwischensubjektfaktor Gruppe (Teilstichprobe 1 & 2) ergab für die transformierte Bearbeitungszeiten nur für die Aufgabe Blockwelten signifikante Unterschiede für die Faktoren Schwierigkeitsstufe,  $F(2, 46) = 88.742$ ,  $p < .001$ ,  $\eta_p^2 = .794$  und Gruppe  $F(1, 23) = 8.05$ ,  $p = .009$ ,  $\eta_p^2 = .259$ , aber keinen signifikanten Effekt für Interaktion der beiden Faktoren  $F(2, 46) = .98$ ,  $p = .382$ ,  $\eta_p^2 = .041$ .

## 4) Inferenzstatistische Analysen – Akkuratheit

Die Ergebnisse des Friedman-Tests zeigen keine signifikanten Unterschiede zwischen den Schwierigkeitsstufen innerhalb einer Aufgabe in Bezug auf die Akkuratheit. Nach Sidak-Korrektur liegt das angepasste Signifikanzniveau bei  $p = .003$ . Dies trifft auf alle drei Aufgaben zu (Kristalle:  $\chi^2(2) = 10.18$ ,  $p = .006$ , Sternschnuppen:  $\chi^2(2) = 2.98$ ,  $p = .225$ , Blockwelten:  $\chi^2(2) = .53$ ,  $p = .768$ ). Dies deutet darauf hin, dass sich die Akkuratheit der Bearbeitung der Einzelaufgaben von Schwierigkeitsstufe zu Schwierigkeitsstufe innerhalb einer Aufgabe nicht signifikant verändert. Beim Vergleich der einzelnen parallelen Schwierigkeitsstufen zwischen den Aufgaben zeigen sich signifikante Unterschiede zwischen den Aufgaben in der Akkuratheit auf der zweiten und dritten Schwierigkeitsstufe (zweite Schwierigkeitsstufe:  $\chi^2(2) = 18.39$ ,  $p < .001$ , dritte Schwierigkeitsstufe:  $\chi^2(2) = 20.7$ ,  $p < .001$ ); die Post-hoc-Analysen mittels des paarweisen Vergleichs zeigen, dass die Unterschiede in der zweiten ( $p < .001$ ) und dritten ( $p < .001$ ) Schwierigkeitsstufe nur zwischen den Aufgaben Kristalle und Blockwelten signifikant sind. Für die erste Schwierigkeitsstufe ergibt der Friedman-Test keine signifikanten Unterschiede in der Akkuratheit zwischen den Aufgaben  $\chi^2(2) = 3.84$ ,  $p = .147$ . Das deutet daraufhin, dass die Personen auf der ersten Schwierigkeitsstufe in alle drei Aufgaben ähnlich häufig bzw. selten Fehler machen

## VI. DISKUSSION

Das primäre Ziel des vorliegenden Artikels ist es einen Einblick in die Entwicklung eines innovativen Förderprogramms für Kinder mit Rechenschwäche darzustellen. Nach Avila-Pesantez et al. [8] bedarf es eines umfangreichen Entwicklungsprozesses damit ein Serious Game auch die angestrebten Förderziele erreichen kann. Direkt im Anschluss an

die erste Phase, die Analyse, folgt in der Design-Phase die Planung und Entwicklung eines pädagogischen Konzepts, welches ein Spielkonzept (u. a. Spielszenario, Spielmechaniken) etabliert. Das Förderprogramm „Eckis Würfelkosmos – Abenteuer in der Zahlengalaxie“ verknüpft wesentlichen Inhaltsbereiche, die für den Erwerb von mathematischen Basisfertigkeiten notwendig sind [9], [10], [11], [12], [13] mit einem Spielumfeld, welches die Zielgruppe motivieren und zum Lernen anregen soll [39]. Die erste Pilotversion des Förderprogramms wurde mit dem Ziel entwickelt, die Eignung der Verknüpfung von Förderung und Spielsystem für die Zielgruppe zu eruieren.

### A. Einordnung der Ergebnisse.

Da es sich hierbei um eine explorative Untersuchung handelt, sollten die Ergebnisse mit Vorsicht interpretiert werden. Die Ergebnisse verdeutlichen, dass die erste Pilotversion bei allen Kindern auf verschiedenen Ebenen (Spieldesign, Handhabung/Verständlichkeit, Spaßfaktor, Aufgabenschwierigkeit) als positiv wahrgenommen wird. Dabei sollte nicht außer Acht gelassen werden, dass die Pilotversion nur 3 von 15 Aufgaben beinhaltet und der Bau- und Storymodus lediglich rudimentär Einblicke in die Grundfunktionen des Förderprogramms geben konnte. Vor diesem Hintergrund kann mit Vorsicht interpretiert werden, dass das Förderprogramm entsprechend der Zielgruppe benutzerfreundlich konzipiert wurde.

Auf die Frage nach der wahrgenommenen Aufgabenschwierigkeit konnte beobachtet werden, dass die Aufgabe Blockwelten (Inhaltsbereich: visuell-räumliches Vorstellungsvermögen) von den Kindern schwieriger eingeschätzt wurde als die beiden Aufgaben aus dem Inhaltsbereich Mengenverständnis. Die nachfolgenden Analysen der Bearbeitungszeiten und Akkuratheitswerte konnten die Wahrnehmung der Kinder bestätigen: In der Aufgaben Blockwelten (Anzahl an Bausteinen) haben die Kinder im Vergleich zu den Aufgaben Kristalle (Mengenvergleich) und Sternschnuppen (Simultane Mengenerfassung) im Durchschnitt mehr Zeit zur Bearbeitung der Einzelaufgaben benötigt (höhere Bearbeitungszeiten) als auch mehr Fehler gemacht (niedrigere Akkuratheitswerte). Während die höheren Bearbeitungszeiten möglicherweise auch auf die Bearbeitungsart der Aufgabe zurückgeführt werden können, zeigt die Abnahme der Akkuratheitswerte deutlich, dass die Aufgaben sukzessiv schwieriger werden. Dieses Ergebnis würde sich auch mit der theoretischen Annahme decken, dass Aufgaben aus dem Inhaltsbereich Mengenverständnis grundlegende Voraussetzungen zum Erwerb weiterer komplexerer mathematischer Kompetenzen sind [18], wohingegen die Aufgabe Blockwelten neben der visuell-räumlichen Komponente auch das Zählen als Kompetenz erfordert, die in der Regel auf dem Mengenverständnis aufbaut [11]. Auf die Frage, inwieweit sich die theoretisch hergeleiteten Schwierigkeitsstufen innerhalb einer Aufgabe auch tatsächlich in ihrer Schwierigkeit unterscheiden, wurden ebenfalls die Bearbeitungszeiten und die Akkuratheit von jeder einzelnen Aufgabe genauer betrachtet. Insbesondere bei der Aufgabe Blockwelten konnten signifikante Unterschiede zwischen den Schwierigkeitsstufen festgestellt werden: Mit höherer Schwierigkeitsstufe stieg auch Bearbeitungszeit. Bei der Aufgabe Kristalle haben die Bearbeitungszeiten zwar ebenfalls zugenommen, jedoch waren hier die Unterschiede zwischen den Schwierigkeitsstufen nicht ausreichend groß, um signifikante Effekte zu erzielen. Bei der Aufgabe Sternschnuppen haben die Bearbeitungszeiten mit zunehmender Schwierigkeit eher leicht abgenommen (auch wenn nicht sig-

nifikant). Gleichzeitig lässt sich feststellen, dass die Stichprobengröße mit zunehmender Schwierigkeitsstufe bei dieser Aufgabe kleiner wurde. Die Akkuratheit blieb dagegen konstant. Es lässt sich die Vermutung aufstellen, dass gerade Kinder, die in der Aufgabe weiterkommen, schneller die richtige Anzahl an Objekten wahrnehmen können als Kinder, die mehr Fehler machen und dem entsprechend noch in den niedrigen Schwierigkeitsstufen verbleiben. Beim Vergleich der beiden Gruppen lassen sich leichte Unterschiede in den Bearbeitungszeiten erkennen, was sich mit den Ergebnissen von Kuhn et al. [25] deckt: Kinder mit Rechenschwäche benötigen mehr Zeit als Kinder ohne Rechenschwäche. Insgesamt kann sich dies auf fast alle Bereiche übertragen lassen: Kinder in Lerntherapie waren im Vergleich zu den anderen Kindern im Durchschnitt langsamer; signifikant war der Effekt nur bei der Aufgabe Blockwelten. Es kann unter Vorbehalt angenommen werden, dass die Differenzierung der Aufgaben und der jeweiligen Schwierigkeitsstufen in Bezug auf unterschiedliche Leistungsniveaus mathematischer Kompetenzen adäquat ist. Um die aufgestellte These zuverlässig zu überprüfen, sind künftig gezielte Untersuchungen unter streng kontrollierten Bedingungen erforderlich, um potenzielle Schwächen auszuschließen.

### B. Limitationen.

Die vorliegende Studie ist mit einigen Limitationen behaftet, die bei der Interpretation der Ergebnisse berücksichtigt werden müssen und für die nachfolgenden Untersuchungen berücksichtigt werden sollten. Die gewählte Stichprobengröße und -auswahl ist noch deutlich begrenzt, was die Generalisierbarkeit der Ergebnisse einschränkt, insbesondere, da der Großteil der Kinder, dieselbe Klasse besuchte. Eine größere und diversere Stichprobe wäre notwendig, um umfassendere und repräsentativere Erkenntnisse zu gewinnen. An dieser Stelle ist anzumerken, dass die Klassenstufe bzw. das Alter ein erheblicher Faktor für die Bearbeitung der Aufgaben sein kann; Dreiviertel der Kinder besuchten zum Erhebungszeitpunkt die vierte Klasse. So können die fehlenden signifikanten Unterschiede bei der Aufgabe Kristalle dadurch entstanden sein, dass diese Aufgabe für die meisten Kinder sehr einfach zu lösen war; diese These wird durch die hohen Akkuratheitswerte (gleich 1, also keine Fehler) bekräftigt. Ähnlichen Einfluss kann die Lerntherapieerfahrung haben. Für die nachfolgenden Untersuchungen sollte kontrolliert werden, wie lange die Kinder an der Lerntherapie bereits teilnehmen.

Auch kann die Testsituation die Ergebnisse beeinflusst haben: Während die Teilstichprobe 1 im Gruppensetting das Förderspiel erprobt hatte, bearbeiteten die Kinder aus der Teilstichprobe 2 die Aufgaben im Einzelsetting. Die Gruppensituation birgt das Risiko einer Beeinträchtigung der Aufmerksamkeit der Kinder, beispielsweise durch soziale Ablenkungen. Dies könnte sich insbesondere auf die Akkuratheit bei der Aufgabe zur simultanen Mengenerfassung auswirken.

Die aktuelle Untersuchung basiert nur auf drei von 15 Aufgaben des gesamten Förderprogramms und betrachtet auch nur die ersten drei Schwierigkeitsstufen genauer. Die Ergebnisse dieser Studie sind daher möglicherweise nicht vollständig repräsentativ für das gesamte Förderprogramm und die einzelnen Aufgaben, die mehr als nur drei Schwierigkeitsstufen besitzen. Zudem konzentriert sich die Untersuchung auf eine kurzfristige Evaluation, wodurch langfristige Auswirkungen des Förderprogramms auf die mathematischen Fähigkeiten der Kinder nicht berücksichtigt werden konnten.

Trotz einiger Limitationen konnte diese Studie erste Einblicke in den Entwicklungsprozess einer neuen Fördersoftware „Eckis Würfelkosmos – Abenteuer in die Zahlengalaxie“ geben. Die gewonnenen Ergebnisse ermöglichen es, spezifischere Forschungsfragen weiterzuentwickeln, Aufgaben anzupassen und erste Aussagen über die Validität des Förderprogramms zu tätigen.

## DANKSAGUNG

Diese Forschungsarbeit in dem Forschungsprojekt *AppLeMat – App für modulare Lerntherapie im Bereich Mathematik* wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- [2] Karakoç, B., Eryılmaz, K., Turan Özpolat, E., & Yıldırım, İ. (2022). The Effect of Game-Based Learning on Student Achievement: A Meta-Analysis Study. *Technology, Knowledge and Learning*, 27(1), 207–222. <https://doi.org/10.1007/s10758-020-09471-5>
- [3] Ok, M. W., Bryant, D. P., & Bryant, B. R. (2020). Effects of Computer-Assisted Instruction on the Mathematics Performance of Students with Learning Disabilities: A Synthesis of the Research. *Exceptionality*, 28(1), 30–44. <https://doi.org/10.1080/09362835.2019.1579723>
- [4] Benavides-Varela, S., Zandonella Callegher, C., Fagioli, B., Leo, I., Altoè, G., & Lucangeli, D. (2020). Effectiveness of digital-based interventions for children with mathematical learning difficulties: A meta-analysis. *Computers & Education*, 157, 1–15. <https://doi.org/10.1016/j.compedu.2020.103953>
- [5] Li, Q., & Ma, X. (2010). A Meta-analysis of the Effects of Computer Technology on School Students' Mathematics Learning. *Educational Psychology Review*, 22(3), 215–243. <https://doi.org/10.1007/s10648-010-9125-8>
- [6] Löffler, L., Skoba, L., Richter, M., & Daseking, M. (2022). Einsatz und Praktikabilität von Applikationen und Software zur Förderung von Mathematikschwierigkeiten aus lerntherapeutischer Sicht. *Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw*, 1, 299–304. <https://doi.org/10.24405/14568>
- [7] Lamer, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., & Petridis, P. (2017). Essential features of serious games design in higher education: Linking learning attributes to game mechanics. *British Journal of Educational Technology*, 48(4), 972–994. <https://doi.org/10.1111/bjet.12467>
- [8] Avila-Pesantez, D., Delgadillo, R., & Rivera, L. A. (2019). Proposal of a Conceptual Model for Serious Games Design: A Case Study in Children With Learning Disabilities. *IEEE Access*, 7, 161017–161033. <https://doi.org/10.1109/ACCESS.2019.2951380>
- [9] Aunio, P., & Räsänen, P. (2016). Core numerical skills for learning mathematics in children aged five to eight years – a working model for educators. *European Early Childhood Education Research Journal*, 24(5), 684–704. <https://doi.org/10.1080/1350293X.2014.996424>
- [10] Braeuning, D., Ribner, A., Moeller, K., & Blair, C. (2020). The Multifactorial Nature of Early Numeracy and Its Stability. *Frontiers in Psychology*, 11, 1–13. <https://doi.org/10.3389/fpsyg.2020.518981>
- [11] Fischer, U., Roesch, S., & Moeller, K. (2017). Diagnostik und Förderung bei Rechenschwäche: Messen wir, was wir fördern wollen? *Lernen und Lernstörungen*, 6(1), 25–38. <https://doi.org/10.1024/2235-0977/a000160>
- [12] Hirsch, S., Lambert, K., Coppens, K., & Moeller, K. (2018). Basic numerical competences in large-scale assessment data: Structure and

- long-term relevance. *Journal of Experimental Child Psychology*, 167, 32–48. <https://doi.org/10.1007/s13138-018-0131-6>
- [13] Purpura, D. J., & Lonigan, C. J. (2013). Informal Numeracy Skills: The Structure and Relations Among Numbering, Relations, and Arithmetic Operations in Preschool. *American Educational Research Journal*, 50(1), 178–209. <https://doi.org/10.3102/0002831212465332>
- [14] KMK (2022). *Bildungsstandards für das Fach Mathematik. Primarbereich.* (Beschluss der Kultusministerkonferenz vom 15.10.2004, i.d.F. vom 23.06.2022). [https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2022/2022\\_06\\_23-Bista-Primarbereich-Mathe.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2022/2022_06_23-Bista-Primarbereich-Mathe.pdf)
- [15] Ise, E., Dolle, K., Pixner, S., & Schulte-Körne, G. (2012). Effektive Förderung rechenschwacher Kinder: Eine Metaanalyse. *Kindheit und Entwicklung*, 21(3), 181–192. <https://doi.org/10.1026/0942-5403/a000083>
- [16] Lenhard, W. & Lenhard, A. (2010). *Rechenspiele mit Elfe und Mathis I: Ein Mathematikförderprogramm für Kinder der ersten, zweiten und dritten Klassenstufe.* Göttingen: Hogrefe.
- [17] Lenhard, W., Lenhard, A. & Lingel, K. (2010). *Rechenspiele mit Elfe und Mathis II: Ein Mathematikförderprogramm für Kinder der dritten, vierten und fünften Klassenstufe.* Göttingen: Hogrefe.
- [18] Krajewski, K. & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. *Psychologie in Erziehung und Unterricht*, 53, 246-262.
- [19] Krajewski, K., Nieding, G., & Schneider, W. (2008). Kurz- und langfristige Effekte mathematischer Frühförderung im Kindergarten durch das Programm „Mengen, zählen, Zahlen“. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(3), 135–146. <https://doi.org/10.1026/0049-8637.40.3.135>
- [20] Landerl, K. & Kaufmann, L. (2008). *Dyskalkulie. Modelle, Diagnostik, Intervention.* Basel: Ernst Reinhard
- [21] Lenhard, A., Lenhard, W., Schug, M. & Kowalski, A. (2011). Computerbasierte Mathematikförderung mit den "Rechenspielen mit Elfe und Mathis I" - Vorstellung und Evaluation eines Computerprogramms für Erst- bis Drittklässler. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 79-88.
- [22] Von Aster, M. G. (2017). Dyskalkulie: Wenn Kinder nicht rechnen lernen. *Monatsschrift Kinderheilkunde*, 165(6), 482–489. <https://doi.org/10.1007/s00112-017-0289-x>
- [23] Freiberger, D., Aringer, C., Büchinger, J., & Furlan, P. (2020). *Training bei Dyskalkulie und Rechenstörungen.* Springer.
- [24] Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., Galli, M., Martin, E., & von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage*, 57, 782–795. <https://doi.org/10.1016/j.neuroimage.2011.01.070>
- [25] Kuhn, J.-T., Raddatz, J., Holling, H., & Dobel, C. (2013). Dyskalkulie vs. Rechenschwäche: Basisnumerische Verarbeitung in der Grundschule. *Lernen und Lernstörungen*, 2(4), 229–247. <https://doi.org/10.1024/2235-0977/a000044>
- [26] Fritz, A., Ehlert, A., & Leutner, D. (2018). Arithmetische Konzepte aus kognitiv-entwicklungspsychologischer Sicht. *Journal für Mathematik-Didaktik*, 1(39), 7–41. <https://doi.org/10.1007/s13138-018-0131-6>
- [27] Cordes, S., & Gelman, R. (2005). The young numerical mind: when does it count?. In J. I. D. Campbell (Hrsg.), *The handbook of mathematical cognition* (S. 127–142). Psychology Press
- [28] Stern, E. (2003). Früh übt sich: Neuere Ergebnisse aus der LOGIK-Studie zum Lösen mathematischer Textaufgaben in der Grundschule. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche – Lernwege, Schwierigkeiten und Hilfen* (S. 116–130). Weinheim: Beltz
- [29] AWMF - Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (Online, 2018). S3-Leitlinie: Diagnostik und Behandlung der Rechenstörung. Zugriff am 15.07.2024 unter <https://register.awmf.org/de/leitlinien/detail/028-046>
- [30] Kaufman, A. S. & Kaufman, N. L. (2015). *Kaufman assessment battery for children-II: KABC-II.* Pearson.
- [31] Hammill, D. D., Pearson, N. A. & Vorress, J. K. (2014). *Developmental Test of Visual Perception 3 (DTVP 3).* London: Pearson.
- [32] Kaufmann, S., & Wessolowski, S. (2021). *Rechenstörungen: Diagnose und Förderbausteine.* Klett/Kallmeyer.
- [33] Dörner, R., Göbel, S., Effelsberg, W., & Wiemeyer, J. (2016). *Serious Games.* Springer International Publishing. <https://doi.org/10.1007/978-3-319-40612-1>
- [34] Szczesna, A., Tomaszek, M., & Wieteska, A. (2012). The Methodology of Designing Serious Games for Children and Adolescents Focused on Psychological Goals. In E. Piętko & J. Kawa (Hrsg.), *Information Technologies in Biomedicine* (Bd. 7339, S. 245–255). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-31196-3\\_24](https://doi.org/10.1007/978-3-642-31196-3_24)
- [35] Barz, N., Benick, M., Dörrenbächer-Ulrich, L., & Perels, F. (2023). The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis. *Review of Educational Research*, 94 (2), 199-227. <https://doi.org/10.3102/00346543231167795>
- [36] Chao, T., Chen, J., Star, J. R., & Dede, C. (2016). Using Digital Resources for Motivation and Engagement in Learning Mathematics: Reflections from Teachers and Students. *Digital Experiences in Mathematics Education*, 2(3), 253–277. <https://doi.org/10.1007/s40751-016-0024-6>
- [37] Fadda, D., Pellegrini, M., Vivanet, G., & Zandonella Callegher, C. (2022). Effects of digital games on student motivation in mathematics: A meta - analysis in K - 12. *Journal of Computer Assisted Learning*, 38(1), 304–325. <https://doi.org/10.1111/jcal.12618>
- [38] Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>
- [39] Caserman, P., Hoffmann, K., Müller, P., Schaub, M., Straßburg, K., Wiemeyer, J., Bruder, R., & Göbel, S. (2020). Quality Criteria for Serious Games: Serious Part, Game Part, and Balance. *JMIR Serious Games*, 8(3), 1–14. <https://doi.org/10.2196/19037>
- [40] Baten, E., Pixner, S., Desoete, A. (2019). Motivational and Math Anxiety Perspective for Mathematical Learning and Learning Difficulties. In Fritz, A., Haase, V.G., Räsänen, P. (Hrsg.), *International Handbook of Mathematical Learning Difficulties* (S.457–467). Cham: Springer. [https://doi.org/10.1007/978-3-319-97148-3\\_28](https://doi.org/10.1007/978-3-319-97148-3_28)
- [41] Van Der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes: A Meta-Analysis. *Review of Educational Research*, 85(4), 475–511. <https://doi.org/10.3102/0034654314564881>
- [42] Imbellone, A., Marinensi, G., & Medaglia, C. M. (2016). Instructions and Feedback in Connection with the Duration and the Level of Difficulty of a Serious Game. In A. De Gloria & R. Veltkamp (Hrsg.) *Games and Learning Alliance* (S. 51–60). Springer.
- [43] Benton, L., Joye, N., Sumner, E., Gauthier, A., Ibrahim, S., & Vasalou, A. (2023). Exploring how children with reading difficulties respond to instructional supports in literacy games and the role of prior knowledge. *British Journal of Educational Technology*, 54(5), 1314–1331. <https://doi.org/10.1111/bjet.13318>
- [44] Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *Heidelberger Rechentest: HRT 1-4.* Göttingen: Hogrefe.
- [45] Wißmann, J., Heine, A., Handl, P., & Jacobs, A. M. (2013). Förderung von Kindern mit isolierter Rechenschwäche und kombinierter Rechen- und Leseschwäche: Evaluation eines numerischen Förderprogramms für Grundschüler. *Lernen und Lernstörungen*, 2, 91–109. doi:10.1024/2235-0977/a000033

## ANHANG.

TABELLE III: MITTELWERTE UND STANDARDABWEICHUNGEN (IN KLAMMERN) FÜR DIE BEARBEITUNGSZEITEN (BZ), TRANSFORMIERTE BEARBEITUNGSZEITEN (LOG BZ) UND DIE AKKURATHEIT INNERHALB DER DREI AUFGABEN DES PROTOTYPEN.

<i>Gesamtstichprobe</i>	<i>Aufgaben: Kristalle</i>			<i>Aufgaben: Sternschnuppen</i>			<i>Aufgaben: Blockwelten</i>		
	<i>LVL 1</i>	<i>LVL 2</i>	<i>LVL 3</i>	<i>LVL 1</i>	<i>LVL 2</i>	<i>LVL 3</i>	<i>LVL 1</i>	<i>LVL 2</i>	<i>LVL 3</i>
BZ in ms	1512.3 (595.3)	1702.1 (775.5)	1838.5 (973.1)	1309.5 (504.3)	1300.3 (498.3)	1154.8 (342.7)	5121.7 (1399.0)	7620.4 (2023)	10799.4 (2116.9)
Log BZ	7.26 (0.35)	7.34 (0.45)	7.40 (0.48)	7.12 (0.33)	7.11 (0.36)	7.01 (0.28)	8.51 (0.26)	8.91 (0.26)	9.27 (0.20)
Akkuratheit	0.9538 (0.089)	1	0.9872 (0.069)	0.9179 (0.124)	0.9343 (0.094)	0.9204 (0.117)	0.879 (0.189)	0.89 (0.128)	0.8112 (0.233)
N	29	29	29	29	28	23	29	29	25
<i>Teilstichprobe 1</i>									
BZ in ms	1471.17 (642.2)	1810.48 (849.8)	1781.55 (1068.7)	1208.2 (509.1)	1267.83 (512.1)	1126.69 (354.9)	4884.08 (1395.6)	6745.73 (1396.5)	10280.27 (1961.6)
Log BZ	7.22 (0.37)	7.39 (0.49)	7.35 (0.52)	7.04 (0.32)	7.07 (0.39)	6.98 (0.30)	8.46 (0.27)	8.80 (0.22)	9.22 (0.19)
Akkuratheit	0.947 (0.994)	1	1	0.881 (0.134)	0.917 (0.103)	0.908 (0.124)	0.842 (0.214)	0.919 (0.106)	0.861 (0.144)
N	20	20	20	20	19	15	20	20	20
<i>Teilstichprobe 2</i>									
BZ in ms	1603.8 (497.2)	1461.3 (543.3)	1965 (758.7)	1534.6 (438)	1369 (490.3)	1207.5 (335.3)	5650 (1330.2)	9564.2 (1878.7)	12876 (1368.4)
Log BZ	7.34 (0.31)	7.23 (0.35)	7.52 (0.39)	7.3 (0.29)	7.17 (0.32)	7.07 (0.25)	8.62 (0.22)	9.15 (0.19)	9.46 (0.10)
Akkuratheit	0.9689 (0.064)	1	0.9589 (0.123)	1	0.97 (0.062)	0.9438 (0.105)	0.9622 (0.075)	0.8267 (0.156)	0.645 (0.388)
N	9	9	9	9	9	8	9	9	5

Anmerkungen: Teilstichprobe 1: Kinder aus der vierten Klasse in Unterfranken, Teilstichprobe 2: Kinder in Lerntherapie LVL1: erste Schwierigkeitsstufe, LVL2: zweite Schwierigkeitsstufe, LVL3: dritte Schwierigkeitsstufe

# hpc.bw: An Evaluation of Short-Term Performance Engineering Projects

Johann Antonio Duffek,  
Imane Bechalaoui,  
Willi Leinen,  
Hauke Preuß,  
Simon Schlumbohn,  
Yannis Schumann  
*High Performance Computing  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany  
duffek@hsu-hh.de*

Fabian Dethof,  
Sylvia Keßler  
*Engineering Materials and  
Building Preservation  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany  
fabian.dethof@hsu-hh.de*

Marie Rathmann,  
Jessica Kleinschmidt,  
Alexander Kolling,  
Sabine Schmidt-Lauff  
*Continuing Education and  
Lifelong Learning  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany  
marie.rathmann@hsu-hh.de*

Andreas Fink  
*Business Administration, in  
particular Business Informatics  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany*

Marcus Stiemer  
*Theoretical Electrical Engineering and  
Numerical Simulations  
Helmut Schmidt University/  
University of the Federal Armed Forces  
Hamburg, Germany*

Matthias Mayr  
*Institute for Mathematics and  
Computer-Based Simulation  
University of the Bundeswehr Munich  
Munich, Germany*

Philipp Neumann  
*DESY, IT-Department and University Hamburg  
High Performance Computing & Data Science  
Hamburg, Germany*

**Abstract** — Increasing amounts of data and simulations in scientific areas enforce the need of improved software performance. The maintaining scientific staff is often not primarily trained for this purpose or lacks personnel and time to address software performance issues.

A particular aim of the dtec.bw-funded project hpc.bw is to tackle some of these shortcomings. A pillar of the hpc.bw agenda is the offer of a low-threshold consultancy and development support focused on performance engineering. This paper provides an insight on our related activities.

We illustrate the structure of our annual calls for short-term performance engineering projects, we outline our results at the example of the performance engineering project ‘benEFIT-Numerical simulation of non-destructive testing in concrete’, and we draw a first conclusion on the current procedure.

**Index Terms** — hpc.bw, HPC, performance engineering, parallel I/O, Fortran

## I. HPC.BW - MOTIVATION AND INTRODUCTION

The core objective in high performance computing (HPC) is the development of efficient software, that is often executed on supercomputers, in order to achieve computations within reasonable time frames and/or within resource bounds. Some relevant examples are optimisation problems in logistics and production, numerical simulations in engineering, or image

recognition/analysis via machine learning with real-time demands for medical diagnostics.

However, software is often not utilising the full potential of compute resources. Reasons for this are, e.g. limited time for software optimisations or missing expertise in HPC. This results in a reduced productivity in the use of digital research techniques.

Our project hpc.bw aims to strengthen HPC-related research at the two universities of the Federal Armed Forces and to enforce the transfer of HPC knowledge across the various fields. Based on these project activities, hpc.bw strives

- 1) to strengthen the research and development in respective research fields,
- 2) to promote the interdisciplinary exchange between HPC-related problems,
- 3) to derive and answer new HPC-related research questions from the different field specific problems, and
- 4) to establish a common HPC competence platform.

A key to strengthen research and development lies in performance engineering (PE), i.e. investigating and optimising programs for (parallel) performance.

In the following, after introducing the key activities of the project in 2023/2024 in Sec. II, the concept of short-term

performance engineering projects, that have been established in hpc.bw, is explained in Sec. III. A particular realisation of this concept is laid out in Sec. IV. We close with a short summary and outlook to future project activities in Sec. V.

## II. HPC.BW: PROJECT ACTIVITIES IN 2023/2024

In the first two years of the hpc.bw project a container-based HPC centre was installed, including supercomputer HSUper, communication channels and community-building measures (such as a multiplicator programme, newsletters, seminar series “Computation & Data”) were established and training concepts (such as workshops and performance engineering projects, cf. Sec. III for the latter) were created. The focus in the last one and a half years (2023-mid of 2024) was put on consolidation of these and running research activities (such as research on HPC for optimisation problems in logistics).

Various groups have adopted to using the HPC system HSUper, an initial TOP500 supercomputer (i.e. ranked 339 among the 500 fastest supercomputers in the world at the time of its inauguration). More than 180 users currently have access to the machine, more than 100 users rely on HSUper currently in their research and actively submit compute jobs. This also includes researchers funded directly from hpc.bw, conducting research on efficiently solving optimisation problems in logistics. This resulted in a total use of 73% of the available compute resources, that was measured over the past months (February–June 2024), demonstrating a currently well-defined size of the compute system and a functioning software stack for a wide range of applications (such as from fluid dynamics, materials science, algorithms research and numerical mathematics, or optimisation in logistics). To enable researchers to use HSUper, three workshops were conducted: two full-day workshops addressing beginners in HPC (held on 28 Sep 2023 and on 18 Apr 2024 at HSU) as well as one full-day workshop for more advanced users of HPC resources (held on 19 Apr 2024 at HSU).

To provide HPC resources also for researchers with less affinity to supercomputing and computer science in general, the *Interactive Scientific Computing Cloud (ISCC)* has been established in summer 2023. It consists of ten dual-socket Intel(R) Xeon (R) Platinum 8360Y nodes (which is the same architecture as used in HSUper) with 256 GB RAM each, as well as of two more nodes of the same type, but with 1TB RAM, additional 2TB local scratch storage and 8 NVIDIA A30 GPUs for more compute-heavy machine learning tasks. In contrast to HSUper, which runs a current Linux distribution and relies on job scheduling via SLURM, the ISCC uses a virtualisation environment built on top of VMware’s vSphere. This allows users to launch a virtual machine with Windows or Linux and run it including the entire desktop and GUI environments on (single) high performance compute nodes. In the mean time, more than 60 virtual machines have already been created and are in use.

Moreover, the HPC portal<sup>1</sup> has been launched to provide a virtual competence platform for the identified user groups at

<sup>1</sup><https://portal.hpc.hsu-hh.de/>

the universities at the Federal Armed Forces and beyond. The platform will be extended in the future with e-learning, more videos on selected HPC topics and continuous documentation of the compute platforms (such as HSUper and ISCC).

Moreover, hpc.bw staff actively reached out further to HPC and education communities: various renowned speakers were invited to the hpc.bw-empowered seminar series on “Computation & Data”, the project was represented, amongst others, at the NHR conference 2023, the Section Conference for Adult Education 2023 as well as at the ISC HPC congress 2024. From 24-26 June 2024, hpc.bw organised the European Trilinos User Group Meeting (EuroTUG) at HSU, with ca. 20 participants from more than ten European research institutions.

## III. SHORT-TERM PERFORMANCE ENGINEERING PROJECTS

The project hpc.bw organises short-term performance engineering projects in the form of easily requestable consultancy on HPC-related questions. Their aim is to provide help and support to users confronted with software performance issues. Figure 1 provides a schematic overview to the following project characterisation.

### A. Application Process

Eligible groups for the short-term PE projects are all departments at the Universities of the Federal Armed Forces. Annually, a call for PE projects is launched that groups can apply for. The application scheme is straightforward, with application forms not exceeding 2-3 pages. The application form asks to characterise

- 1) the applicant’s research project and its relation to HPC,
- 2) the existing software to be optimised,
- 3) the problems faced with the current software,
- 4) the anticipated gains in performance improvement for the software, required to answer related research questions, and
- 5) the necessary support needed as well as the estimated project duration.

Overall support must not exceed the equivalent of six person months of an HPC-affine scientific co-worker from the hpc.bw project. The applicants are additionally asked to categorise their HPC-related knowledge into ‘no’, ‘basic’, ‘advanced’, or ‘expert’.

This deliberately simple application procedure lowers the barrier for candidates without prior knowledge in HPC, and it provides us with a rough overview of the underlying problem and expected workload.

The applications are internally reviewed by the hpc.bw multiplicator team, which represents all involved institutions/departments of the Universities of the Federal Armed Forces, and they are externally reviewed by one more person. After review, projects are selected, and, if necessary, adapted in their extent and duration. Table I provides an overview on the granted projects, their extent, and involved personnel.

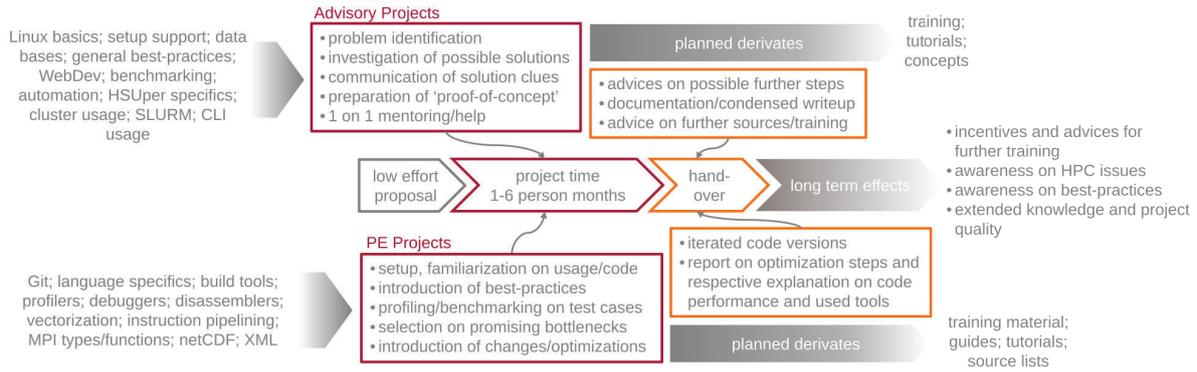


FIGURE 1. A SCHEMATIC OVERVIEW OF THE SHORT-TERM PERFORMANCE ENGINEERING PROJECTS OF 'HPC.BW'

TABLE I. Overview on the granted short-term PE projects. Project Count: total number of supported projects. Granted Person Months: total number of person months allocated in the respective call. Involved Staff: number of staff members from hpc.bw and the chair for HPC at HSU involved in the PE projects

Year	Project Count	Granted Person Months	Involved Staff
2022	7	12	4
2023	4	12	3
2024	6	10	6

### B. Project Implementation

Kick-off meetings are scheduled with the applicants of the accepted projects within a month after the submission deadline. Their purpose is to

- 1) introduce the involved personnel to each other,
- 2) gain a better insight into the outlined problem,
- 3) understand expectations and fix the scope of the project work,
- 4) establish and ask for further required prerequisites (such as software stacks, source code and compile scripts, test cases to evaluate performance gains),
- 5) define the communication channels form of joint project work (e.g. defining meeting schedules or organising time line of the project), and to
- 6) reiterate the project conditions.

The project conditions entail the contributions of material and abstracts for reports and publications as well as the acknowledgement of significant contributions.

Our staff – scientific co-workers from the project hpc.bw and members of the chair for HPC – then proceed to familiarise themselves with the project and work on the defined problems. A synchronisation is conducted according to the project nature and according to the preference of the involved people.

Each project is usually concluded by a spin-off meeting, presentation, and/or report to transfer our acquired findings to the project partner. An additional one-page report with a brief description of the measures and their outcome is written up by the involved project members to document the results.

Our intention is furthermore to simplify the typically encountered problems and to track lessons-learned in order to reuse them in training materials afterwards.

### C. Typical encountered Projects

The short-term PE projects can be usually classified into two kinds of projects.

#### a) The classical PE projects.

These are projects requiring the optimisation of some existing code. An example of such project and its outcome is given in Sec. IV. Their process is roughly as follows:

- 1) Request of source code access and local setup of the project.
- 2) Familiarisation with the source code and project.
- 3) Optional further/advanced preparation of the project.
- 4) Repeated iterations of profiling and adaptations of the program/source code/software environment.
- 5) Documentation of the changes and program hand-over.

The ideal outcome is a list of required (or already completed) program modifications and a measurable outcome/improvement. These improvements range from negligible changes (e.g. 5% runtime reduction) to runtime reductions of over 50%. The encountered problems range from a proper software setup, over the conscious use of libraries and system resources, to in-depth PE techniques applied to user-specific (or even third-party) code such as vectorisation and pipelining.

#### b) The advisory projects.

These projects have a wide variety of appearances and range from 'simple on-boarding help' to use an HPC cluster, over the setup and use of a particular compute- or data-intensive software, to the full build of a proof-of-concept. These are generally handed in by beginners.

The programming languages covered in our projects so far include, amongst others, Python, Matlab, Fortran, and C++ with project sizes from 400 up to 75,000 lines of code.

## IV. BENEFIT

The project 'benEFIT - Numerical simulation of non-destructive testing of concrete' is one of the classical PE

projects (see Sec. III-C) and was granted three person months of attention on our first short-term PE project call. Its subject was a self-written Fortran code of approximately 5,000 lines of code.

Its purpose is to simulate the elastic wave propagation during two different non-destructive testing methods (ultrasonic testing and impact echo testing). These methods are used to assess the condition of concrete structures, such as buildings or bridges, and detect possible defects inside the inspected concrete structures. The respective software elastodynamic finite integration technique (EFIT) utilises a finite difference staggered-grid method which is able to handle strongly heterogeneous domains more robustly.

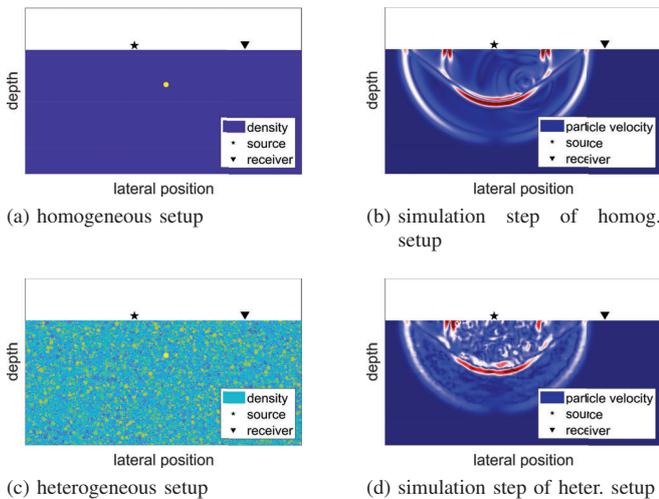


FIGURE 2. EXAMPLE SIMULATIONS CONDUCTED WITH BENEFIT

Figure 2 depicts two example calculations conducted with EFIT. Figure 2(a)-(b) depict a slice of a homogeneous simulation with embedded steel reinforcement (yellow point) and easily recognisable wave reflections. Figure 2(c)-(d) depict a heterogeneous - more realistic - scenario with a - therefore - very noisy response. EFIT is used to simulate such ensembles to examine the detectability of defects or other features in concrete using the two aforementioned non-destructive testing methods [1].

As the introduction of absorbing boundaries, here implemented with perfectly matched layers (PMLs), requires the use of many additional variables, two basic versions of the EFIT code exist (with and without PML).

### A. Initial State and Project Aim

Such simulations require fine domain resolutions due to the heterogeneous material and many time steps needed for the comparatively long simulated time interval, especially for impact echo testing, which results in high memory and runtime demands. Those had been already addressed by a distributed-memory parallelisation with MPI (Message Passing Interface). The declared project aim was the further reduction of runtime

and memory demands of EFIT, as well as the parallelisation of a separated accompanying domain creation tool.

### B. Initial Steps and Preparation

Initial steps involved the collapsing of several hard-coded program versions into one source code version, modifiable via configuration parameters. This reduced the lines of code considerably (see Fig. 3 baseline). A build setup via GNU Make was introduced to facilitate the build of the different program flavours and to incorporate instrumentation with profiling tools.

Those and all further changes were tracked and documented with the versioning system Git. This had the aim to acquaint our project partner with such a system and to separate the steps into comprehensible changes and document-able iterations (without the need of explicit version copies).

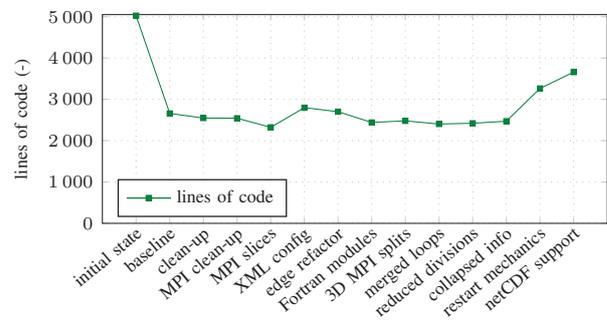
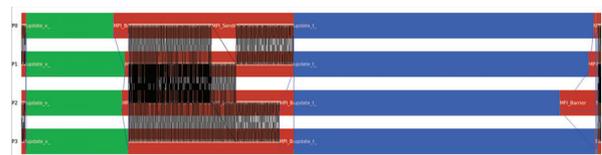


FIGURE 3. THE LINES OF CODE AT MAJOR PROJECT MILESTONES

### C. Initial Performance Engineering Process



(a) trace excerpt

Name	TSelf	TSelf	TTotal	TTotal	#Calls	#Calls	TSelf/Cal
All Processes							
MPI_Sendrecv	30.359e-3 s		30.359e-3 s		7957		3.81538e-6 s
exchng_t	1.17e-6 s		841.988e-6 s		0		n.a.
update_V	19.873e-3 s		19.873e-3 s		4		4.96825e-3 s
exchng_V	3.115e-3 s		32.749e-3 s		4		778.75e-6 s
update_t	57.9932e-3 s		57.9932e-3 s		4		14.4983e-3 s
User_Code	0 s		115.523e-3 s		0		n.a.
MAIN	94e-6 s		115.523e-3 s		0		n.a.
MPI_Barrier	3.9714e-3 s		3.9714e-3 s		13		305.492e-6 s

(b) accumulated time values of the trace excerpt

FIGURE 4. TRACING RESULTS OF ONE PROGRAM CYCLE

The previously introduced instrumentation was then utilised to acquire a deeper insight of the time-consuming program routines. Figure 4 displays the excerpt of one sampled program cycle (4 MPI processes) acquired via Intel® Trace Analyzer and Collector [2]. The two visible sub-compute steps (green and blue) are encased and interrupted by MPI information interchange steps (red). The table lists the accumulated time of the shown routines (TSELF).

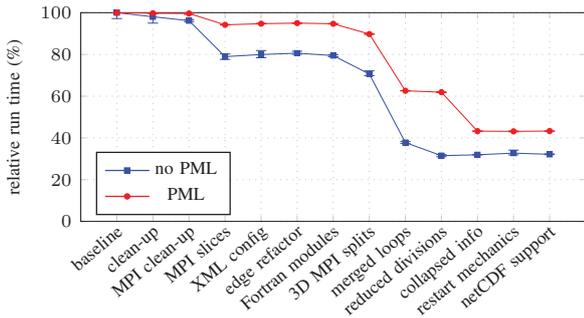


FIGURE 5. PROGRAM RUNTIME AT THE MAJOR PROJECT ITERATIONS AVERAGED OVER THREE RUNS (WHISKERS DENOTE MIN/MAX VALUES)

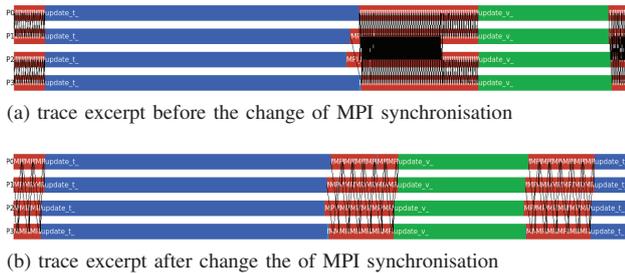


FIGURE 6. TRACE EXCERPT FOR THE NO PML PROGRAM VERSION, BEFORE AND AFTER THE STEP MPI SLICES

The relatively big amount of time spent in the update routines led to the decision to optimise these routines. Several attempts to apply loop fusion (the merging of different loops) and loop nesting optimisation (the blocking of n-dimensional loops to optimise memory access) led to no immediate significant improvement.

A second approach initiated a structural clean-up of the source code with a simplification of the MPI information exchange. These preceding steps enabled us to easily introduce a proper MPI information exchange for n-dimensional arrays leading to a noticeable runtime reduction (see Fig. 5 up to step MPI slices). This reduced runtime can be attributed to the reduction of MPI exchange operations and had a different impact on our two test cases (with/without PML). This is due to an increased share of computation to information exchange for the EFIT feature PML. The reduction and effect of this change is visualised in Fig. 6 (the black lines between the bars represent the exchange operations).

#### D. Interlude

We recommended and implemented an XML-based simulation configuration in order to avoid hard-coded program variants and program recompilation. This step has been undertaken to facilitate the use of EFIT within an extended/automated context, with the benefit of reproducible runs due to the separation of a static/fixed program from input and configuration files (simulation reproduction possible via configuration and input file). The mature Fortran XML-library TIXI [3] has been chosen as backbone for this implementation.

The changes encompassed refactoring (restructuring) steps of the source code and the transitioning to a more modern Fortran syntax. These changes were undertaken for the sake of a better clarity of the code base and led therefore to no significant performance changes.

The previous steps facilitated the introduction of a generalised domain splitting, which enabled us to choose a more cache-friendly domain splitting scheme (splitting along the dimension of the largest array stride). This led to an additional performance improvement (step 3D MPI splits).

#### E. Performance Engineering Process

A second attempt to merge simulation loops resulted in a very noticeable performance gain (step merged loops). This is due to a more informed approach about additional circumvention of intermediate variables. The results are a lower memory footprint, better cache utilisation (memory utilisation), and therefore an improved runtime.

Further optimisations included the avoidance of division operations by re-using results from prior calculations, and the reduction of artificially inflated information. The later change resulted in further significant runtime improvements for the PML case. This case requires more cache accesses and profits therefore from a reduced data volume. The division-reduction led to no noticeable improvement for the PML case as due to more prevalent memory instructions which masked this improvement. The less memory hungry case no PML in turn profits slightly.

#### F. Finishing Touches

A remark from our project partner regarding very long runtimes led to the implementation of a check-pointing mechanism with support of netCDF files [4] to EFIT.

This check-pointing enables the full dump of the state in a time step in the case of an externally triggered signal. This is very useful on compute job scheduled (i.e. Slurm managed) machines. It avoids the necessity to write frequent simulation dumps to the local storage, effectively reducing the potential runtime for cases which require a sparse recording of domain states.

NetCDF reduces the data written from and to disk even further. The ASCII based data-files formerly used by EFIT inflated the required data artificially (besides enforcing a text to data conversion during a read).

The impact of those changes had not been evaluated with the provided test cases. But measurements on in-production problems provided significant runtime improvements from e.g. 72h down to 15h. These effects originated primarily in the avoided heavy use of hard drive reads and writes (the file sizes were reduced by approximately 70%).

The project itself was concluded with a report providing an overview on the changes and their motivation. Additional hints on further steps such as MPI communication-hiding were provided as well.

Work on the domain generation program was skipped due to time constraints.

## V. CONCLUSIONS AND OUTLOOK

In this article, we provided an update on the project status of hpc.bw and detailed the short-term performance engineering projects, that are being carried out within the scope of our project. In particular, we demonstrated how hpc.bw can enable more efficient research at the example of the PE project benEFIT.

While benEFIT can be considered a success story, the overall chances to improve code performance depends on various aspects, ranging from code base and quality over program complexity and problem setting to be solved to actual skills available in the team. While the latter can, of course, be built up over time, the structure of performance engineering project calls does not put any constraints on software (open-source, closed-source, quality of documentation, etc.), programming language, and so forth. While this allows us to support an as heterogeneous HPC user group as possible, it increases the need for a wide range of expertise in the PE project team. Therefore, other project activities on building up an actual competence platform with online portal, documentation, and learning material is essential to help users directly and leverage PE projects for individual needs.

Another aspect to critically keep track of in the future is the overall time frame dedicated to the projects. Designed as *short-term* projects, first statistics have shown that the entire work time from hpc.bw PE project staff is distributed to ca. 70% on familiarisation with the considered software/ problem and its preparation for the actual performance engineering, to 10% on project result reporting and only to ca. 20% on actual performance optimisation tasks. The longer a project, the better the ratio of familiarisation-to-optimisation time becomes and the smaller the number of potentially supportable projects becomes (due to limits in the human resources). And the more researchers are already well-educated on programming (e.g. writing modular, well-understandable code with extensive documentation or using code versioning systems such as Git), the smaller the familiarisation phase typically becomes. Careful consideration of these aspects will be required to be monitored in the future.

### ACKNOWLEDGMENT

hpc.bw is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

### REFERENCES

- [1] F. Dethof & S. Keßler, “Design of Concrete Mock-Ups with Complex Defect Scenarios Using Numerical Simulations”, *Journal of Nondestructive Evaluation*, vol. 43(59), 2024, doi: 10.1007/s10921-024-01074-9.
- [2] “Intel® Trace Analyzer and Collector 2022.1.0”, Intel®, 2022.
- [3] M. Siggel et al., “TIXI 3.3.0”, DLR, 2022, available: [github.com/DLR-SC/tixi](https://github.com/DLR-SC/tixi).
- [4] Unidata, “NetCDF-Fortran 4.6.1”, UCAR/Unidata Program Center, 2023, doi: 10.5065/D6H70CW6.

# Das dtec.bw-Forschungsprojekt „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“ – gestaltungsorientierte Forschung für, in und mit den Streitkräften

Katharina Kreuzmann\*, Matthias Armster, Manuel Schulz

Zentrum für technologiegestützte Bildung

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Hamburg, Deutschland

\*kreutzmk@hsu-hh.de

**Kurzfassung** – Die Digitalisierung, verstanden als nächster großer Entwicklungsschritt der menschlichen Kommunikation und Interaktion, beeinflusst unsere Berufs- und Lebenswelt weitreichend. Die positiven, aber insbesondere die potenziell negativen Aspekte, die mit derart umfassenden und zunehmend beschleunigten Veränderungen einhergehen, können auf das einzelne Subjekt verunsichernd und in der Folge dessen destruktiv auf demokratische Gesellschaftsformen und ihre Arbeits- und Wirtschaftsstrukturen wirken. Um dieser Entwicklung zu begegnen, beforcht das dtec.bw-Projekt „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“ seit 2021 in drei Arbeitspaketen die Bedingungen, Möglichkeiten und Umsetzungsformen von Kompetenzerwerb in einer digitalen (Arbeits-)Welt. Angesichts der aktuellen Weltraumlage erfolgte im Laufe des Jahres 2022 auf Basis der bisherigen Forschungsergebnisse aus zivilen Bildungseinrichtungen eine Schwerpunktverlagerung der Forschungsarbeit auf die Aus-, Fort- und Weiterbildung von Streitkräften.

**Stichworte** – Bildung, Ausbildung, Kompetenz, Resilienz, Öffentliche Wissenschaft, Nachhaltigkeit, Bundeswehr, Streitkräfte, Einsatzfähigkeit

## I. EINLEITUNG

Das dtec.bw-Forschungsprojekt „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“ forscht seit 2021 in einem gestaltungsorientierten kontextualisierten Forschungsansatz in Kooperation mit Stakeholdern

- verschiedener Universitäten (Pädagogische Hochschule Freiburg, Universität der Bundeswehr München),
- Behörden (Bundesinstitut für Berufsbildung (BIBB), Hamburger Institut für Berufliche Bildung (HIBB)),
- sechs Hamburger Berufsschulen,
- der Logistiksule der Bundeswehr (LogSBw),
- der Schule für Feldjäger und Stabsdienst der Bundeswehr (SFJg/StDstBw),
- dem Streitkräfteamt der Bundeswehr (SKA),
- dem Bildungszentrum der Bundeswehr (BiZBw) und
- unter Beteiligung der fachlich zuständigen Referate des Bundesministeriums der Verteidigung (BMVg)

zu didaktisch-methodischen Verfahren, die den Kompetenzerwerb zur Bewältigung und Mitgestaltung der sich durch

die Digitalisierung exponentiell verändernden Berufs- und Lebenswelt fördern. Oberste Bedingung für die Forschung in diesem Feld ist die Interaktion in realen Bildungskontexten, um die Bedarfe sowie die Ausgangspunkte der darin agierenden Subjekte (Lernende, Lehrende, Lernberatende, Lernprozessorganisierende ...) aufzunehmen und zielgruppengerechte Szenarien zur Implementierung zu gestalten. Um im Forschungsfeld mögliche Lücken zwischen Theorie und Praxis zu schließen, werden die Erkenntnisse aus der Implementierungsforschung in die Theoriebildung zurückgespiegelt.

Das Projekt ist dazu in drei Arbeitspakete unterteilt:

Das **Arbeitspaket 1** fokussiert die Ertüchtigung des Subjekts. Studien zeigen, dass die aufwachsende Komplexität digitaler Technologien und Netzwerke Gefühle der Überforderung bei den Subjekten auslösen können. In Verbindung mit Fake News und fehlenden Kriterien zur Be- und Auswertung medialer Quellen sowie Inhalte besteht die Gefahr von Destabilisierungseffekten in Gesellschaften [1]. Diesen möglichen Effekten kann nur mit entsprechenden Bildungsszenarien begegnet werden, die Subjekten ermöglichen, ihre eine gesellschaftlich verantwortliche Balance zwischen Anpassung und Eigensinn sowie Vertrauen und Misstrauen zu finden. Dies gewann während der Projektlaufzeit in Anbetracht des völkerrechtswidrigen Angriffs der Ukraine im Februar 2022 und der damit eintretenden Zeitenwende an besonderer Bedeutung.

Ohne einen Referenzrahmen wird es jedoch schwierig den Subjekten dauerhaft die Möglichkeit zu bieten, ihre Einordnung zwischen den Polen Anpassung und Eigensinn sowie Vertrauen und Misstrauen vorzunehmen und auch ggf. gegen Anfeindungen durchzusetzen. Ein solcher Referenzrahmen muss sich – das ist entscheidend, um Polarisierungen jedweder Art noch vor dem eigenen Einordnungsprozess vorzubeugen – durch die Darbietung „gehärteter Informationen“ auszeichnen, bspw. durch entsprechend faktenbasierter Reviews, Kuratierungen der Inhalte oder Einordnungen in Diskussionsrunden. Dieser Aufgabe widmet sich das **Arbeitspaket 2**. Im Sinne einer „Öffentlichen Wissenschaft“ [2] ist in diesem Arbeitspaket der (Aus)Bildungskongress der Bundeswehr verortet, der den Subjekten mit seiner wissenschaftlichen Tagung und Fachausstellung als Plattform dienen soll, sich bezüglich neuer Entwicklungen und Erkenntnisse zu informieren und



ABBILDUNG 1: Arbeitsergebnisse

auszutauschen. So kann er eine Grundlage für die Entscheidungs- und Bewertungsprozesse sein, die jedes Subjekt als Mitglied und Mitgestalter der Gesellschaft durchläuft.

Die Herausforderung eines solchen Referenzrahmens liegt dabei u. a. in der zunehmenden Zergliederung wissenschaftlicher Fachgebiete. Dies kann zu Widersprüchen zwischen Disziplinen führen, die eine interdisziplinäre gemeinsame Beforschung dieser Bereiche/Begriffe lösen könnte. Genau dies erfolgt im **Arbeitspaket 3** in Kooperation mit der Pädagogischen Hochschule Freiburg in einem interdisziplinären wissenschaftlichen Forschungsnetzwerk, dem „Metaforum für Digitalisierung, Resilienz und Nachhaltigkeit (DfRN)“ [3]. Aktuell widmet sich die eingerichtete Arbeitsgruppe einer fachdomänen übergreifenden gemeinsamen Definition des Resilienzkonstruktes.

In den folgenden drei Kapiteln möchten wir einen kurzen Überblick über den aktuellen Erkenntnis- und Ergebnisstand in den jeweiligen Arbeitspaketen geben. Das abschließende Kapitel gibt einen Ausblick auf die zukünftige Projektarbeit.

## II. ARBEITSPAKET 1: BILDUNG FÜR DIE DIGITALE ARBEITSWELT – ERTÜCHTIGUNG DES SUBJEKTS

In Rückbezug auf die kommunikationssoziologischen Ausführungen Baeckers, der die Digitalisierung als nächsten bedeutsamen Entwicklungsschritt nach Sprache, Schrift und Buchdruck mit beweglichen Lettern sieht, steht die Menschheit mit der Digitalisierung erneut vor einer durch grundlegende Veränderungsprozesse geprägten Zeit [4]. Die besondere Herausforderung dabei liegt in der enormen Beschleunigung und der Erhöhung der Komplexität, unter denen diese Prozesse ablaufen [5]. Neben den positiven Aspekten, die die Digitalisierung mit sich bringt, ergeben sich daraus auch Einflüsse, die sich zunehmend verunsichernd auf die Subjekte auswirken. Dazu zählen insbesondere die aufwachsenden

weitreichenden Veränderungen von Berufsbildern durch Automatisierungsprozesse [6], der größer werdende Einfluss von Medien- und Filterblasen auf gesellschaftliche und politische Meinungen [7] und die möglichen Destabilisierungseffekte sozialer Medien auf demokratisch organisierter Gesellschaften durch einen unreflektierten Umgang mit Medienaussagen [1]. Auf dieser Grundlage sind im Arbeitspaket 1 mit sechs Hamburger Berufsschulen Lern-/Lehrsznarien entwickelt worden, die in unterschiedlichen fachlichen Bereichen angesiedelt und entsprechend individuell (zeitlicher Umfang, Einbindung digitaler Technologien ...) ausgestaltet sind. ABBILDUNG 2 gibt einen Überblick über die entwickelten Lernsznarien. Schon während der Konzipierungsphase bestätigten sich die zwei zentralen Forschungsannahmen: 1) Die Aushandlung der Positionierung zwischen den Polen *Eigensinn* und *Anpassung*, und zwar sowohl mit sich selbst als auch mit anderen, ist ein zentraler Bestandteil zur Bewältigung und Mitgestaltung der zunehmend digitalen Welt. 2) Eine entscheidende Fähigkeit für Informations-, Aushandlungs-, Bewertungsprozesse dieser und anderer Art wird sein, Inhalte und Quellen bezüglich ihrer Qualität, insbesondere hinsichtlich ihrer Vertrauenswürdigkeit, z. B. durch eine kurative redaktionelle Bearbeitung, beurteilen zu können [2]. Nach der Implementierungsphase, mehreren Unterrichtsbeobachtung und Reflexionsgesprächen mit den Lehrkräften sowie Schüler:innen haben sich die folgenden Fähigkeiten/Kompetenzen [8] als zentral zu fördernd herausgestellt, um den Herausforderungen der Digitalisierung resilient begegnen zu können:

- 1) Entwicklung und Anwendung von Kriterien für Informationen und Quellen (u.a. zur Vermeidung von affektiver Polarisierung)
  - zur Bewertung und Einordnung
  - zu Tiefe und Breite
  - zu Nutzung und Folgen

- 2) Kriteriengeleitete Übernahme von Gruppenrollen und Durchsetzung, u. a. zur Bewältigung der durch Digitalisierung entstandenen Arbeitsaufgaben  
→ Formen neuer (digitaler und gruppenorientierter) Arbeitskultur
- 3) Kriteriengeleitet Reflexion von Arbeits-/Gestaltungsprozessen (Bewertungskriterien: Produkt, Ablauf, Organisation, soziales Gefüge, Moral i. w. S.)

Ein Phänomen, das angesichts der zunehmenden Komplexität und Beschleunigung besonders in den Fokus unserer Forschungsarbeit rückte, ist das sog. „Social Loafing“. Soziales Faulenzen tritt insbesondere während Gruppenarbeitsphasen ein, in denen keine klaren Zuständigkeiten für die einzelnen Gruppenmitglieder definiert sind. Dies erfordert angesichts der Tatsache, dass viele Aufgaben nicht mehr allein lös-/leisbar sind und somit Teamarbeit zukünftig eine immer größer werdende Rolle spielen wird, eine klare Aufgaben- und insbesondere Verantwortungszuweisungskultur. Die Arbeit jedes einzelnen Gruppenmitglieds muss dabei zur Lösung der Gesamtaufgabe beitragen (→ echte Arbeitsteilung). „Social Loafing“ ist, wie Studien vermuten lassen, nämlich kein bewusster Prozess. Das „Herausziehen“ erfolgt zumeist unterbewusst, da das Gehirn mit seinen Ressourcen immer schonend umgeht. Sobald also der Nutzen der eigenen Leistung nicht deutlich ist, ist das Subjekt nicht motiviert und beteiligt sich nicht an der Aufgabe [10]. Dazu wurde in einer der konzipierten Lehr-/Lernszenarien die Methode EduScrum [11] in Rückbezug auf das Konzept „Experten unterschiedlicher Ausprägung“ [12] erfolgreich implementiert und erprobt, sodass sich zumindest feldexplorativ die positiven Effekte eines entsprechend ausgestalteten Arbeitsgruppenprozesses bestätigten.

Im Sinne eines Dual-Use war die Nutzbarmachung der Erkenntnisse aus den zivilen Bildungseinrichtungen von Projektbeginn an geplant, gewann durch den Angriffskrieg Russlands auf die Ukraine jedoch spätestens ab dem Frühjahr 2022 zunehmend an Bedeutung. Dazu wurden u. a. Kooperationen mit dem Streitkräfteamt der Bundeswehr, der Logistikschule der Bundeswehr und mit der Schule für Feldjäger und Stabsdienst der Bundeswehr geschlossen. Im Rahmen der bundeswehrgemeinsamen Arbeitsgruppe „Technologiegestützte Ausbildung“ des Streitkräfteamts der Bundeswehr, der Aus- und

Fortbildungsverantwortliche aller Teilstreitkräfte und Organisationsbereiche der Bundeswehr angehören, übernimmt das Zentrum für technologiegestützte Bildung (ZtB) schon viele Jahre die Rolle der wissenschaftlich beratenden Instanz in Bezug auf Regelungen, didaktische Grundausrichtung sowie organisatorische und technische Realisierung technologiegestützter Bildung im gesamten Geschäftsbereich des Bundesministeriums der Verteidigung.

Im Rahmen dieser Tätigkeit hat das ZtB das neue Ausbildungskonzept der eTrainer, eTrainer Professionals und eTrainer Experts beraten. Dieses Konzept basiert auf dem auch schon vom ZtB konzipierten Modellehrgang der sog. „Teletutoren“ [13], der im Hinblick auf aktuelle Entwicklungen (Digitalisierung, Bedeutung von Social Media, Einsatzfähigkeit bei Mangel an Personal und Gerät) zu einem mehrstufigen Rollenmodell für Lehrende in der technologiegestützten Ausbildung weiterentwickelt wurde.

Ein zentrales Element ist dabei die Einbindung von Verfahrens- und Handlungstrainern im Zuge der Dezentralisierung und Hybridisierung von Ausbildung u.a. auch zur Umsetzung des neuen Grundsatzes „Train while you fight!“. Die Weiterbildung der eTrainer Experts als höchste Stufe dieses Rollenmodells wird i. R. des dtec.bw-Projekts KoDiA durch das ZtB entwickelt und erprobt.

In der Logistikschule der Bundeswehr steht die Offizierausbildung im Fokus, insbesondere das Lernfeld zur materiellen Versorgung in der Landes- und Bündnisverteidigung – Stichwort „Logistische Drehscheibe Deutschland“. Hierfür wurde ein Simulationslabor aufgebaut, indem die Abläufe der Versorgung sowohl in Präsenz, als auch hybrid und unter Nutzung von Augmented sowie Virtual Reality trainiert werden, bevor die tatsächliche Übung im Feld erfolgt. Der Einsatz der Simulation soll bei den Teilnehmenden eine erste Form der Handlungssicherheit erzeugen.

An der Schule für Feldjäger und Stabsdienst der Bundeswehr wird der Fokus auf der Ausbildung von Feldwebeln und sog. eTrainern, eTrainer Professionals und eTrainer Experts liegen. Ziel in beiden Schulen ist, die in Ausbildung befindlichen Soldaten unter Einbezug vielfältiger digitaler Technologien und Inhalte zu einer reflektierten, fundierten und verant-

Ausbildungsberuf	Lernfeld gem. Rahmenlehrplan	Lehr-/Lernsituation
Maurer:in (BS 08)	Mauern eines einschaligen Baukörpers (LF 3)	Dreidimensionales Zeichnen (CAD-Software) ↳ Verzahnung praktischer und digitaler Arbeitswelt ↳ Entwicklung von Problemlösestrategien im digitalen Raum ↳ Kriterienentwicklung zur Auswahl u. Bewertung von Quellen u. Informationen
Kaufleute für Büromanagement (BS 18)	Büroprozesse gestalten und Arbeitsvorgänge organisieren (LF 2)	Workshopentwicklung / Projektmethode EduScrum ↳ Kriteriengeleitete Übernahme von Gruppenrollen (Formen neuer digitaler u. gruppenorientierter Arbeitskultur) ↳ Erfahrung als Entwicklerteams der eigenen Beruflichkeit
Zahnmedizinische Fachangestellte (BS 15)	Hygienemaßnahmen organisieren und Medizinprodukte aufbereiten (LF 3)	Erstellung Hygieneplan über „Erklärvideos“ im digitalen Raum und Umsetzung in „Virtueller Zahnarztpraxis“ ↳ Wahrnehmung der eigenen Beruflichkeit im digitalen Raum
Kaufleute für audiovisuelle Medien (BS 17)	Prozess der Leistungserstellung gestalten u. kontrollieren Endfertigung planen u. kontrollieren (LF 6+7)	Social-Media-Simulation (Fake News, Shitstorm) Panel-Diskussion ↳ Resilienzentwicklung in komplexen Social-Media-Szenaren
Tischler:in (BS 25)	Kleinmöbel herstellen (LF 4) bzw. Einzelmöbel herstellen (LF 5)	Vom Entwurf zum Produkt: alle Arbeitsschritte und Unterlagen organisiert über ein Projektmanagementtool ↳ Arbeitsprozess im digitalen Raum planen, organisieren und kontrollieren
Goldschmied:in (BS 19)	Kein konkretes Lernfeld, stattdessen eine Lerngebiete umfassende Lehr-/Lernsituation	Logoentwicklung zur eigenen Stempelgestaltung + interne Social-Media-Plattform zum Austausch (Expertennetzwerk) ↳ Auseinandersetzung mit Entrepreneurship im digitalen Raum ↳ Aufbau und Pflege hybrider Expertennetzwerke

ABBILDUNG 2: Entwickelte Lehr-/Lernsituationen in den Berufsschulen

wortungsbewussten Entscheidungsfindung in komplexen Lageszenarien zur Sicherung einer demokratisch verfassten Gesellschaft zu befähigen. Hybride, digitale und simulationsbasierte Ausbildungsszenare ermöglichen dabei die Herstellung einer Praxisnähe und gleichen ggf. bestehende Mängel an Einsatzgerät und Personal aus.

In beiden Projekten werden inzwischen gemeinsam auf Basis der Erkenntnisse aus den berufsbildenden Schulen in Hamburg entsprechende didaktische Konzepte kompetenzorientierter technologiegestützter Bildungsangebote entwickelt, implementiert, erprobt und beforscht. Dabei sind die folgenden Forschungsfragen untersuchungsleitend:

- Wie müssen e-Learning-Ausbildungsszenarien gestaltet sein, um einen dauerhaften und nachhaltigen Kompetenzerwerb, im Sinne eines lebenslangen und selbstverantwortlichen Lernens, durch die Ausbilder:innen anzustoßen, der über eine Teilhabe hinaus zur Mit-/Weitergestaltung von Ausbildung anregt?
- Wie müssen e-Learning-Ausbildungsszenarien (u.a. inhaltlich, didaktisch, infrastrukturell, medial) gestaltet sein, um als Ausbilder:innen tätige Einsatzkräfte auf eine selbständige, selbstbestimmte und sichere Aufgabenerfüllung vorzubereiten?
- Welche besonderen Aspekte sind bei der Konzeption von e-Learning-Ausbildungsinhalten mit Blick auf die aktuelle und zukünftige Lebens- und Arbeitswelt (u.a. demografischer Wandel, Diversität und Heterogenität der Lerngruppen – insbesondere auch bzgl. ihrer Bildungsbiografien, Arbeits- und Lernmodelle, Fake News, Zunahme der Komplexität und Beschleunigung der Arbeits- und Lebenswelt, sicherheitspolitische Manipulationsgefahren) und unter der Prämisse, die Verfügbarkeit von Personal und Material in der Truppe zu erhöhen, zu berücksichtigen?

Beispiele, Ergebnisse und Erkenntnisse sowohl aus den zivilen als auch aus den militärischen Bildungseinrichtungen werden jährlich auf dem (Aus)Bildungskongress der Bundeswehr präsentiert. Dieser wird im Fokus des folgenden Abschnittes stehen.

### III. ARBEITSPAKET 2: INNOVATIVE KOOPERATIONEN – ÖFFENTLICHE WISSENSCHAFT

Das Arbeitspaket 2 befasst sich mit dem bereits weiter oben erwähnten Referenzrahmen. Die zentrale Aufgabe dieses Referenzrahmens besteht darin, die Subjekte mit „gehärteten“, d.h. kuratierten, bestenfalls wissenschaftlich redaktionierten, Informationen und Fachwissen zu versorgen, damit mündige und selbstbestimmte Entscheidungen in einer zunehmend komplexen Arbeits- und Lebenswelt unterstützt werden können. Dies ist umso wichtig, da Desinformation und Manipulation in Sozialen Medien sowie der zunehmend beschleunigte technologische Fortschritt die Subjekte vor wachsende Herausforderungen stellt [1], [5]. Der (Aus)Bildungskongress der Bundeswehr, in dessen Vordergrund seit jeher der Informationsaustausch mit und der Dialog zwischen Expert:innen aus Wissenschaft, Wirtschaft, öffentlichem Dienst und Bundeswehr und Verantwortlichen für die Aus-, Fort- und Weiterbildung von Einsatzkräften steht, bietet als Use-Case im Rahmen „Öffentlicher Wissenschaft“ eine Plattform, die als ein solcher Informations-/Referenzrahmen dienen kann.

Inspiriert von der Idee Burawoys, die Soziologie in ein Gespräch mit Öffentlichkeiten zu bringen, kann der (Aus)Bildungskongress der Bundeswehr als „Agentur“ für den Austausch zwischen Wissenschaft und interessierter Öffentlichkeit, d. h. in diesem Fall vornehmlich Bildungsverantwortliche in den Einsatzkräften, dienen [14]. Dabei ist der Agenturgedanke nicht als einseitige Belehrung von Fachleuten gegenüber anderen zu verstehen, sondern vielmehr als Dialog mit und zwischen den Teilnehmenden [2].

Ein kontextualisierter Forschungsgegenstand ist hier den Dialog und die Mittlerfunktion des Agenturgedankens am Beispiel der Ausbildungsverantwortlichen im System Bundeswehr als Modell für eine „Öffentlichkeit“ zu erforschen. Anhand unterschiedlicher Kommunikationsformate, wie z. B. dem (Aus)Bildungskongress der Bundeswehr oder im Rahmen der hybriden Treffen der sogenannten KOA-Kaffeerunde (Kompetenzorientierte Ausbildung) wird die Umsetzung eines Agenturmodells erprobt. Die „Agentur“ nimmt in diesem Fall eine Mittlerfunktion ein, in welcher Fragen der Öffentlichkeit am o. g. Beispiel aufgenommen und im Austausch mit der Wissenschaft Lösungen erarbeitet werden und gleichzeitig auch Forschungsbedarfe aus der Praxis der Öffentlichkeit für die Wissenschaft aufgezeigt werden können [2]. Forschungsfragen zu möglichen Kommunikationsformaten und -räumen, den Eigenschaften und der Organisation der „Agentur“ sowie Einflüssen der Digitalisierung wie virtuelle Räume oder Fake-News spielen eine zentrale Rolle.

Der Kongress eröffnet als zentraler Use-Case somit nicht nur den Aufbau eines Referenzrahmens, sondern ermöglicht den Teilnehmenden durch die direkte Interaktion auf Augenhöhe auch die Teilhabe an der inhaltlichen Gestaltung dieses Referenzrahmens. Gleichzeitig besteht die Möglichkeit, für Fachleute aus Wissenschaft, Wirtschaft, öffentlichem Dienst und Bundeswehr aus der Teilnahme am Dialog Forschungsbedarfe zu identifizieren. Damit bietet der (Aus)Bildungskongress der Bundeswehr als Use-Case „Agentur“ nicht nur eine Grundlage für verlässliche Informationen für die eigene Bewertung, Entscheidung und Reflexion der Subjekte, sondern kann auf dieser Basis die Subjekte auch zur aktiven Mitgestaltung ihrer Berufs- und Lebenswelt über den Kongress hinaus befähigen.

Die Organisation und Durchführung des Kongresses erfolgt bundeswehrgemeinsam mit der Abteilung Ausbildung Streitkräfte im Streitkräfteamt der Bundeswehr, dem Bildungszentrum der Bundeswehr sowie der HSU/UniBw H als Gastgeber und wissenschaftlich verantwortlicher Instanz. Der Kongress zeichnet sich durch einen jährlich wechselnden thematischen Schwerpunkt aus dem Themenfeld der Aus-, Fort- und Weiterbildung von Einsatzkräften aus. Die aus dem jeweiligen Themenfeld resultierenden inhaltlichen Bedarfe werden im Dialog „Öffentlichkeit und Wissenschaft“ konkretisiert. Dazu dient das sog. „Programmkomitee“, das sich aus Vertreter:innen aus Wissenschaft und Bildungsverantwortlichen aus dem System der Bundeswehr zusammensetzt. Als Repräsentanz des oben beschriebenen Dialogs entscheidet es in mehreren hybrid ausgestalteten Beratungsrunden zu verschiedenen Zeitpunkten im Planungsprozess über die spätere Detailausplanung des Tagungsprogrammes.

Den Grundpfeiler des Kongresses bildet deshalb der durch das Programmkomitee auf Basis der Vorschläge aus dem Dialog „Öffentlichkeit und Wissenschaft“ entstandene Call for Papers. In diesem wird das jeweilige Thema mit Hintergrün-

den erläutert und Fachleute, Forschende sowie Wissenschaftler:innen dazu aufgerufen, sich anhand der Leitfragen an der inhaltlichen Ausgestaltung des Programms z. B. in Form eines Vortrags, Workshops, einer Posterpräsentation zu beteiligen. Im Rahmen eines Double-blind-Review-Auswahlverfahrens durch das Programmkomitee wird das wissenschaftliche Tagungsprogramm aus den Einreichungen des Call for Papers zusammengestellt. Ergänzt wird das wissenschaftliche Tagungsprogramm durch Beiträge der Streitkräfte, anderer Einsatzkräfte und der Wirtschaft. So soll ein breiter Dialog auf Basis der wissenschaftlichen und anderen Fachbeiträge gewährleistet werden. Zusätzlich besteht die Möglichkeit des Besuches der angeschlossenen Fachausstellung. Sie bietet Hands-on Erfahrungen und fördert insbesondere den praxisorientierten Dialog, wodurch der oft theoretisch geprägten Welt der Wissenschaft und Forschung die Möglichkeit geboten wird, relevante Forschungsfragen aus konkreten Praxiskontexten zu generieren.

Die Umsetzung des (Aus)Bildungskongress der Bundeswehr erfolgt hybrid. Neben einer wissenschaftlich geprägten Tagungsdokumentation, in Form eines Tagungsbandes, werden die Anteile des wissenschaftlichen Tagungsprogramms aufgenommen und sind so auch nach dem Kongress online verfügbar.

Der in Form des (Aus)Bildungskongress der Bundeswehr dialogisch gestaltete Referenzrahmen zur Unterstützung der Entscheidungs- und Bewertungsprozesse der Subjekte könnte somit ein Baustein zur Bewältigung der Herausforderungen der Digitalisierung sein. Entscheidend dabei ist – und dieser Aufgabe widmet sich das Arbeitspaket 3 – ein semantischer Konsens über grundlegende Begrifflichkeiten und Methoden bzw. zumindest ein entsprechender Hinweis bei entscheidenden Widersprüchen.

#### IV. ARBEITSPAKET 3: FORSCHUNGSNETZWERK – STAKEHOLDER BALANCING

Die zunehmende Zergliederung wissenschaftlicher Fachdomänen und -begriffe wurde im Rahmen des dtec.bw-Projekts insbesondere mit Blick auf den während der COVID-19-Pandemie omnipräsent gewordenen Resilienzbezug exemplarisch beforcht. In unserer interdisziplinären Zusammenarbeit sowohl auf institutioneller Ebene (Schulen, Behörden, Ämter, Abteilungen, Referaten, Universitäten, Ministerien) als auch auf fachlicher Ebene (Bildungswissenschaften, Psychologie, Wirtschaftswissenschaften, Ingenieurwissenschaften, Informatik) haben wir schnell festgestellt, dass teilweise sehr unterschiedliche Definitionen und Konzepte mit dem Begriff „Resilienz“ verknüpft sind. Darin begründet lag die Initiative, ein interdisziplinäres wissenschaftliches Forschungsnetzwerk ins Leben zu rufen. Unter Verantwortung unseres Kooperationspartner an der Pädagogischen Hochschule Freiburg fanden inzwischen sieben Netzwerktreffen des „Metaforum für Digitalisierung, Resilienz und Nachhaltigkeit (DfRN)“ [3] statt. Im Zuge der ersten Diskussionsrunden wurde eine Synopse als Zusammenschau der verschiedenen Definitionen von Resilienz aus den unterschiedlichen Disziplinen (Psychologie, Pädagogik, Soziologie, Elektrotechnik, Betriebswirtschaft, Informatik) erstellt. Auf Grundlage dieser wird zurzeit in hybriden Arbeitstreffen an einer gemeinsamen Veröffentlichung gefeilt, die einen ersten Schritt für eine interdisziplinäre Resilienzdefinition liefern soll. Die Besonderheit bei diesem Definitionsansatz ist, dass die einzelnen Fachdomänen ihre Beiträge nicht kumulativ erstellen, sondern vom ersten Moment an integrativ. Das bedeutet eine gemeinsame

Texterstellung im virtuellen Raum. Die Erscheinung ist für 2025 geplant.

#### V. AUSBLICK

Im Kontext der aktuellen Weltraumlage hat das Bundesministerium der Verteidigung bezüglich einer weiteren dtec.bw-Förderphase zur Bedingung gemacht, dass alle durch dtec.bw geförderten Projekte einen konkreten Bezug zum System Bundeswehr zukünftig nachzuweisen haben. Da in KoDiA die Beforschung der Aus-, Fort- und Weiterbildung von Streit- und anderen Einsatzkräften unter den Bedingungen der Digitalisierung schon in der ersten Projektphase immanenter Bestandteil der Forschungsarbeit war, ändert sich für die Ausrichtung von KoDiA in der zweiten Projektförderphase kaum etwas. Selbst gewählt wird sich KoDiA ab 2025 jedoch deutlich auf die Beforschung von Kompetenzen zum Erwerb von Handlungssicherheit und digitaler Souveränität (u.a. Verständnis, Bewältigung, verantwortliche Nutzung und Mitgestaltung der Digitalisierung) von Einsatzkräften in der Landes- und Bündnisverteidigung fokussieren. Dazu gehört u. a. die Verzahnung von Kompetenzen zur Bedienung und Bewertung technologischer Mittel zur Erschließung von Informationen mit Kompetenzen zur kriteriengeleiteten Reflexion der Quellen, Ergebnisse und Prozesse (digitale Kompetenzen). Die Integration von Handlungs- und Verfahrenssimulationen wird dabei ein zentraler Aspekt sein, um in Anbetracht von aufwachsendem Personal und Gerät die Verfügbarkeit des Personals in der Truppe zu erhöhen. Bislang liegen keine Forschungserkenntnisse zum Zusammenhang von Sicherheit (in einem weiten Verständnis menschlicher Sicherheit [9]), Demokratiebildung und Bildung zur mündigen und verantwortlichen Mitgestaltung der Digitalisierung im Kontext der sog. „Zeitenwende“ vor. Dieser Forschungslücke wird sich KoDiA ab 2025 intensiv unter den folgenden Fragenstellungen widmen:

- Welche besonderen Herausforderungen ergeben sich aus der Aufgabe der Landes- und Bündnisverteidigung zur Wahrung demokratischer Werte für die Ausbildung einsatzbereiter Streitkräfte und wie können diese realitätsnah trainiert werden (Handeln in gefährvollen Einsatzsituationen und unmittelbarem Entscheidungszwang)?
- Wie können diese Besonderheiten in bereits bestehende Ausbildungsszenarien (u.a. inhaltlich, didaktisch, infrastrukturell, medial) integriert werden, um auf eine selbständige, selbstbestimmte und sichere Aufgabenerfüllung vorzubereiten?
- Welche neu zu gestaltenden Ausbildungsbedarfe und -szenarien ergeben sich und wie müssen diese (u.a. inhaltlich, didaktisch, infrastrukturell, medial) gestaltet sein und umgesetzt werden, um Einsatzkräfte auf eine selbständige, selbstbestimmte und sichere Aufgabenerfüllung vorzubereiten?

Als weiterer Aspekt wird in KoDiA ab 2025 die Begleitforschung digitalisierter Unterrichtsräume im Rahmen von Lehrangeboten des Zentrums für technologiegestützte Bildung im Bereich der Interdisziplinären Studienanteile (ISA) dazukommen. Dabei stehen insbesondere die Auswahl, Didaktisierung und mediale Gestaltung der durch das ZtB angebotenen zunächst rein zivil erscheinenden akademischen Aus-

bildungsinhalte, -mittel und -räume im Zentrum, um die studierenden Offiziere optimal auf die Organisation, Evaluation und Weiterentwicklung von digitalisierter und hybridisierter Ausbildung als eine Kernaufgabe ihrer späteren Verwendung vorzubereiten. Ziel ist es, den Teilnehmenden die konkrete Verwendbarkeit der Fachinhalte ihres Studiums für ihre zukünftige Tätigkeit aufzuzeigen und sie zu interdisziplinären Netzwerkbildung zu ermuntern, da in derart interdisziplinären Netzwerken mit verteilten Arbeits- und Rollenaufgaben das Potential steckt, die auch zunehmend beschleunigten und komplexer werdenden militärischen Aufgaben zu bewältigen.

Der (Aus)Bildungskongress der Bundeswehr wird weiterhin als Use Case zur Beforschung des Konzepts „Öffentliche Wissenschaft“ [2] dienen. Als Plattform zur Erschließung wissenschaftlichen Wissens für die Öffentlichkeit der Sicherheits- und Einsatzkräfte und zum Austausch auf Augenhöhe zwischen Wissenschaftler:innen sowie Sicherheits- und Einsatzkräften soll er als belastbarer Referenzrahmen für informierte und faktenbasierte Entscheidungen den Subjekten behilflich sein.

### DANKSAGUNG

Diese Forschungsarbeit im Forschungsprojekt *Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung* wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGeneration EU finanziert.

### LITERATUR

- [1] Lorenz-Spreen, P. et al. (2022): A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nat Hum Behav* 7, 74–101. <https://doi.org/10.1038/s41562-022-01460-1>
- [2] Neusius, A. / Schulz, M. (2022): Das dtec.bw-Forschungsprojekt „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“ als Beispiel kontextualisierter Forschung. In: dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw. Band 1. Hrsg. von Fay, A. et al. dtec.bw. Hamburg. S. 311 – 320. In Anlehnung an: Aulenbacher, B. et al. (2017) (Hrsg.): *Öffentliche Soziologie. Wissenschaft im Dialog mit der Gesellschaft*. Campus. Frankfurt a.M.
- [3] <https://www.ph-freiburg.de/ibw/institut/berufliche-bildung-fuer-gesundheit-und-nachhaltigkeit/projekte/innolab-bbew/ueber-das-projekt.html> (abgerufen am 20.11.2024)
- [4] Baecker, D. (2017): „Wie verändert die Digitalisierung unser Denken und unseren Umgang mit der Welt? Ausgangspunkte einer Theorie der Digitalisierung.“ in *Gläß, R. (Hrsg.): Handel 4.0. Die Digitalisierung des Handels. Strategien, Technologien, Transformation.*, Berlin und Heidelberg, Springer Gabler, 2016. S. 3 - 24.
- [5] Stöcker, C. (2020): *Das Experiment sind wir: Unsere Welt verändert sich so atemberaubend schnell, dass wir von Krise zu Krise taumeln. Wir müssen lernen, diese enorme Beschleunigung zu lenken.* München: Karl-Blessing-Verlag in der Verlagsgruppe Random House GmbH.
- [6] Briggs, J.; Kodnani, D. (2023): *Global Economics Analyst: The Potentially Large Effects of Artificial Intelligence on Economic Growth.* In Goldman Sachs –Economics Research.
- [7] Eli Pariser: *Invisible sieve: Hidden, specially for you.* In: *The Economist*. 30. Juni 2011.
- [8] Schulz, M. / Neusius, A. (2023): *Bildung zur digitalen Kultur – Neue Kulturtechniken zur Bewältigung der Digitalisierung.* In: „Zeitenwende“ – Wie gelingt die Bewältigung der Herausforderungen? In: Schulz, M. (Hrsg.): „Zeitenwende“ – Wie gelingt die Bewältigung der Herausforderungen? (Aus)Bildungskongress der Bundeswehr 2023. Kongresskatalog. Bonn: Verlag managerseminare, S. 27 – 37. [https://www.hsu-hh.de/ztb/wp-content/uploads/sites/665/2023/09/2023-08-30\\_2023\\_Kongresskatalog.pdf](https://www.hsu-hh.de/ztb/wp-content/uploads/sites/665/2023/09/2023-08-30_2023_Kongresskatalog.pdf)
- [9] Schulz, M. / Neusius, A. / Kreuzmann, K. (2024): *Bildung durch Wissenschaft für Human Security – Offener Geist in geschlossener Gesellschaft?* In: Jung, T. / Sanders, O. (Hrsg.): *Universität und Militärische Sicherheit: Über den Streit um die drohende Einrichtung eines militärischen Sicherheitsbereichs an der Helmut-Schmidt-Universität.* Bielefeld: transcript, S. 159–194. <https://www.transcript-verlag.de/detail/index/sArticle/6768?number=978-3-8394-7023-7>
- [10] Karau, S. J. / Williams, K. D. (1993): *Social loafing: A meta-analytic review and theoretical integration.* In: *Journal of Personality and Social Psychology*. 65(4), 681–706.
- [11] vgl. <https://eduscrum.org>
- [12] Schulz, M. (2005): *Fernausbildung ist mehr... - Auf dem Weg vom technologischen Potenzial zur didaktischen Innovation.* In: Schulz, M. / Glump, H. (Hrsg.): *Fernausbildung ist mehr ... Auf dem Weg vom technologischen Potenzial zur didaktischen Innovation. Grundlagen der Weiterbildung, herausgegeben von Feuchthofen, J. / Jagenlauf, M. / Kaiser, A.. Augsburg: Ziel-Verlag, S. 10 – 37.*
- [13] Schulz, M. (2004): *Leitfaden für den Modelllehrgang Teletutor der Bundeswehr.* Wilhelmshaven.
- [14] Burawoy, M. (2015) *Public Sociology: Öffentliche Soziologie gegen Marktfundamentalismus und globale Ungleichheit.* Beltz Juventa. Weinheim/Basel.

# Komplexität als Kompetenzanforderung an Subjekte in organisationalen Kontexten unter Bedingungen der Digitalisierung

Andrea Neusius

Zentrum für technologiegestützte Bildung  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
neusius@hsu-hh.de

**Kurzfassung** – Subjekte schaffen Komplexität aufgrund ihres Angewiesen-Seins auf Kooperation und Sozialität. Die Bildung von Gemeinschaft, beginnend bei Interaktionsbeziehungen zwischen mindestens zwei Menschen bis hin zu einer vielfach differenzierten Weltgesellschaft bedingt auch das Eingebunden-Sein in unterschiedlichste organisationale Kontexte – sowohl indem wir Mitglieder von Organisationen sind, als auch indem wir als individuelles Subjekt mit Organisationen in Interaktionsbeziehungen treten. Eine wesentliche und notwendige Bedingung für das Entstehen von Gemeinschaft, Organisation und Gesellschaft ist ein Verständigungsmodus, der auf geteiltem Sinnverständnis beruht: Sprache – als gesprochene Sprache, als geschriebene Sprache und in Interaktionsbeziehungen begleitet von Körpersprache, Mimik und Gestik. Daraus entsteht notwendigerweise Komplexität, denn gegenseitiges Verstehen über das Medium Sprache birgt immer auch das Risiko des Nicht-Verstehens oder Miss-Verstehens. Zudem muss Sprache gelernt werden, sie stellt somit eine Kompetenzanforderung an die Menschen. Sie ist zugleich ein Mittel zur Subjektwerdung in dem Sinne, dass Subjekt-Sein erst in der Relation zu anderen möglich wird und hier Sprache auch ein Instrument des Selbst-Ausdrucks in Kommunikation mit anderen ist. Im Zuge der Digitalisierung erweitert sich der Interaktionsraum für uns als Subjekte um eine neue Dimension: den virtuellen Raum. Dort stehen uns verschiedene Formen der Kommunikation und Verständigung zur Verfügung. Wie sich dadurch neue Komplexität mit neuen Kompetenzanforderungen an die Subjekte ergeben, damit in einer digitalen Arbeits- und Lebenswelt Verständigung als geteiltes Sinnverständnis gelingen kann, soll im folgenden Beitrag erörtert werden.

**Stichworte** – Organisation, Subjekt, Komplexität, Kompetenz

## I. EINLEITUNG

Ausgangspunkt der Frage nach Komplexitätsanforderungen, die sich für Subjekte in organisationalen Kontexten im Zuge der Digitalisierung ergeben und danach, wie sich diese von Komplexitätsanforderungen in einer analogen Welt unterscheiden, ist ein Subjektverständnis, das an die Arbeiten des Sozialpsychologen Klaus Holzkamp anknüpft. Dieses Subjektverständnis zeichnet sich dadurch aus, dass es das Subjekt-Sein unter die Bedingung der Relation und Positionsbestimmung zu anderen stellt [1].

In diesem Subjektverständnis ist die Annahme angelegt, dass Vergemeinschaftung und Vergesellschaftung eine Grundbedingung menschlichen Daseins darstellen. Komplexität kennzeichnet menschliches Zusammenleben und damit

menschliche Existenz sui generis auf mehreren Ebenen: Das angewiesen Sein der Menschen auf Sozialität, dass eine anthropologische Grundbedingung unseres Lebens darstellt, erfordert, dass wir miteinander umgehen, miteinander interagieren [2].

Dafür bedarf es geteilter, gemeinsamer Zeichen, um Verständigung zu ermöglichen: Sprache als Zeichensystem erfüllt diese Funktion und erleichtert Kommunikation im Modus geteilter Sinnzusammenhänge. Gleichzeitig muss sie erlernt werden, damit eine Teilnahme an Kommunikation und Interaktion möglich wird [3]. Sie ist damit auch ein Beispiel für von uns Menschen selbst erzeugte Komplexität und begründet zugleich eine Kompetenzanforderung an uns.

Darüber hinaus ist sie Instrument zum Ausdruck von Identität im Abgleich mit Äußerungen anderer über sich, über die materielle und soziale Welt und den eigenen Standpunkt in dieser Welt. So schaffen wir quasi nebenbei auch eine Voraussetzung unseres Subjekt-Seins [1], [4]. Sie dient uns dabei auch als Medium, uns selbst gegenüber anderen zu erklären, uns also zu objektivieren. Dieser Aspekt ist gerade auch in Form geschriebener Sprache relevant, wenn es darum geht, wie wir Identität, Sozialität und geteilte Sinnzusammenhänge z. B. über soziale Medien, aber auch über Messenger-Dienste und Chats herstellen.

Das Eingebunden-Sein in organisationale Kontexte kann als Folge der Sozialität theoretisch gefasst werden: Menschliches Zusammenleben erfordert neben der grundsätzlichen Möglichkeit einer Verständigung und dadurch auf gemeinsame Ziele gerichteter Kommunikation und Interaktion Regeln. Je mehr einzelne Menschen in gemeinsame Interaktionsbeziehungen eingebunden sind, umso wichtiger wird es, den gemeinsamen Interaktionsrahmen durch Regeln zu definieren [5]. Organisationen sind durch komplexe, kodifizierte Regulationssysteme gekennzeichnet, die eine Ausrichtung des Interaktionsrahmens auf definierte Ziele ermöglichen und die Interaktionen auf diese Ziele hin arbeitsteilig optimieren sollen [6].

Durch die Digitalisierung von Arbeits- und Lebenswelt verändern sich Rahmenbedingungen, Prozesse und Regelsysteme, die den Interaktionskontext von Subjekten konfigurieren. Der Komplexität, die soziale Systeme von der Mikro- bis zur Makroebene kennzeichnet, wird damit eine zusätzliche Komplexitätsdimension hinzugefügt [6]. In organisationalen

Interaktionskontexten können dadurch Irritationen scheinbar selbstverständlicher Formen des Kooperierens, des Aushandelns und des Agierens entstehen. Ebenfalls können Strukturierungsprinzipien sowohl der Arbeits- und Interaktionsprozesse als auch der Interaktionsbeziehungen in Frage gestellt werden. Daraus ergeben sich Handlungsnotwendigkeiten, die abstrakt als organisationale Re-Strukturierungen gefasst werden können, konkret aber durch die Subjekte zu „erfinden“ und zu verwirklichen sind [3].

Im Beitrag sollen dazu folgende Leitfragen andiskutiert werden:

- Wie kann das Phänomen „Digitalisierung“ aus subjektwissenschaftlicher Perspektive beschrieben werden?
- Was macht Digitalisierung mit den Subjekten und was machen die Subjekte mit der Digitalisierung?
- Was sind Merkmale digitalisierungsbedingter Komplexität auf der Interaktionsebene?
- Welche Anforderungen an die Bewältigungs- und Gestaltungskompetenz der Subjekte folgen daraus?

## II. GRUNDANNAHMEN ZU SUBJEKT UND ORGANISATION

„Es gilt, den intersubjektiven Kontext einzubeziehen, durch den begreiflich wird, daß nicht nur ‚ich‘ und ‚die anderen‘ sich gegenüberstehen, sondern daß ich ‚für die anderen der andere‘ bin, womit ich von dem, was ich den anderen züfuge, auf einer verallgemeinerten Ebene mitbetroffen bin“ [1].

Wenn Menschsein als Subjekt-Sein verstanden wird, ist der Austausch, die Interaktion und die Kooperation mit anderen als Grundbedingung menschlichen Lebens zu fassen. Zugleich erfordert Subjekt-Sein in dieser Perspektive die Verständigung als geteiltes Verständnis von Sinnzusammenhängen, denn nur so ist Interaktion und Kooperation (als auf ein gemeinsames Ziel gerichtete Interaktion) möglich.

Verständnis und Verstehen gelingt uns wesentlich über das Medium Sprache, das wir von klein auf als Kulturtechnik erlernen – zunächst als gesprochene, später auch als geschriebene Sprache. Wir kommunizieren in Interaktionsbeziehungen aber nicht nur sprechend (oder schreibend, doch dazu später mehr), sondern auch über Körperhaltung, Mimik und Gestik [8].

Im Modus von Interaktion und Kooperation bilden wir Gemeinschaft, durch die Kodifizierung von Regeln der Interaktion schaffen wir Organisation und als Rahmenorganisation Gesellschaft, die wir heute auf einer Makroebene auch als „Weltgesellschaft“ erleben [9].

Organisationen entstehen zu einem definierten Zweck. Sie verfolgen Ziele und sind durch mehr oder weniger differenzierte Regeln, die in Satzungen, Geschäftsordnungen oder ähnlichem objektiviert sind, bestimmt. Den Menschen, die Mitglieder der Organisation sind, werden Aufgaben im Rahmen funktionaler Aufgaben- und Arbeitsteilung zugewiesen und ein Handlungsrahmen definiert, innerhalb dessen sie ihre Aufgaben erfüllen sollen [9], [10].

Dieser Handlungs- oder Interaktionsrahmen ermöglicht arbeitsteilige Kooperation, schafft aber auch einen Interpretationsspielraum, der einerseits den Ausdruck von Individualität bietet und andererseits der Ort ist, an dem sich das Subjekt in

Relation durch Selbst-Äußerung und Spiegelung der Selbst-Äußerungen der anderen positioniert.

Auch Subjekt und Organisation stehen miteinander in Beziehung [11]. Es wird hier insofern ein Organisationskonzept formuliert, dass Organisation als ein von Menschen – Subjekten – geschaffenes, gestaltetes Phänomen versteht. Unter dieser Prämisse kann die Gestaltung von Organisation als Prozess des Organisierens auch als organisationales Lernen beschrieben werden [12].

Das Erleben von Interaktion und Kooperation konfrontiert uns täglich in vielfacher Hinsicht mit Komplexität, mit der umzugehen wir im Laufe unseres Lebens erlernen. Dazu gehört immer auch die Erfahrung von Ohnmacht, wenn wir uns in Situationen befinden, deren „Spielregeln“ wir nicht durchschauen, deren Interaktionswirklichkeit mit den vermuteten Spielregeln nicht zusammenpasst oder deren Verlauf uns volatil erscheint [13].

Ein Erfahrungsfeld, in dem diese Phänomene besonders bewusst werden, sind organisationale Kontexte. Das können berufliche Organisationen wie Unternehmen oder Behörden sein, ebenso aber auch Jugendverbände, Sportvereine oder die Freiwillige Feuerwehr.

## III. DIGITALISIERUNG UND KOMPLEXITÄTSERLEBEN

Während der Covid-19-Pandemie haben sich für die meisten von uns über einen längeren Zeitraum die Interaktionsräume ihrer organisationalen Zugehörigkeiten und Einbindungen in den virtuellen Raum verlagert. Das Erfahrungsfeld „Organisation“ wurde damit um eine Dimension erweitert, die dank technischer Werkzeuge, wie virtueller Konferenzsysteme und netzgestützter, auch abgesicherter Zugriffsmöglichkeiten auf organisationsinterne Daten, Informationen und Arbeitsprodukte, vordergründig ein zielgerichtetes Weiterarbeiten an den gemeinsamen (organisationalen) Zielen in – so die Erwartung – mindestens gleicher Qualität ermöglichen, wie die persönliche Interaktion in der Zusammenarbeit vor Ort.

Aus vielfältigen Erfahrungsberichten und mittlerweile auch wissenschaftlicher Forschung lässt sich ableiten, dass die Erweiterung des Interaktionsraums organisationaler Kontexte intersubjektiver Interaktion und Kooperation vielfältige, im Vorfeld vielleicht so nicht erwartete Folgen hat.

Gehen wir zunächst von der Perspektive des Subjekts auf die Wirklichkeitserweiterung des Interaktionsrahmens um die Dimension der Virtualität ein, erscheint es hilfreich, an Meuelers Verständnis des Subjekt-Seins als Dialektik von Anpassung und Widerstand in drei Dimensionen des Erlebens anzuknüpfen:

- Die Äußere Natur: „Gefährdet durch die Möglichkeiten der Vernichtung erhält sich das Subjekt, indem es sich steigert. Das dazu benötigte Selbstbewußtsein ist ebenso Todesbewußtsein wie Freiheitsbewußtsein“ [14].
- Die Innere Natur: „Der mächtige Einfluss des Unbewußten [...] bestimmt alle Aussagen, alles Verhalten des bewußten Ichs“ [14].
- Die Soziale Welt: „Es sind die in lebenslangen Vergesellschaftungsprozessen verinnerlichten sozialen Deutungsmuster, die das Bewußtsein des einzelnen über die Wirklichkeit außerhalb seiner selbst und sein Leben in dieser Wirklichkeit bestimmen“ [14].

Diese theoretische Anknüpfung rekurriert zum einen auf die Dialektik von Bewusstsein und Unbewusstsein und zum anderen auf das ständige, mitunter widersprüchliche und von Zerrissenheit gekennzeichnete Verorten in einem Koordinatensystem aus Erfahrung und Wahrnehmung im Modus der Interaktion [4].

Die dadurch zusätzlich entstehende und uns permanent konfrontierende Komplexität ergibt sich aus der Notwendigkeit von Sozialität und potenziert sich in organisationalen Kontexten in der Auseinandersetzung mit von Menschen geschaffenen aber als Objektivationen außerhalb ihrer selbst stehenden Regulationssystemen.

Um zu verstehen, ob und ggf. wie Digitalisierung diesen Dimensionen von Komplexität eine (oder mehrere) hinzufügt, gilt es zunächst, ein Verständnis von Digitalisierung und in einem nächsten Schritt von Virtualität als Erscheinungsform von Digitalisierung in unserer alltäglichen Interaktion zu entwickeln.

Christian Stöcker beschreibt Digitalisierung als „Beliebige Arten von Daten, die in anderen Formaten vorliegen, in Abfolgen von Nullen und Einsen zu verwandeln und mit diesen dann mithilfe von Computern unterschiedlichste Dinge anzustellen“ [7]. Am Beispiel lernender Systeme in Algorithmen zeigt er unter anderem die Selektivität, die aufgrund der menschengemachten Quellen, aus denen sich diese Algorithmen ihr „Lernmaterial“ ziehen. Das führt dazu, dass mediale Repräsentationen uns diskriminierende Inhalte in die Timeline spülen, ohne dass wir uns dieser Selektivität bewusst sind.

Dirk Baecker definiert aus einer systemtheoretischen Perspektive das Phänomen „Digitalisierung“ grundsätzlicher und ordnet dieses in die mediale Evolutionsgeschichte der Menschheit ein. Dabei geht er von folgender Hypothese aus:

„Probleme der Digitalisierung entstehen daraus, dass elektronische Medien der Gesellschaft an der Schnittstelle von Mensch und Maschine einen Überschusssinn bereitstellen, auf dessen Bearbeitung bisherige Formen der Gesellschaft strukturell und kulturell nicht vorbereitet sind“ [3].

Er versteht dabei Digitalisierung als eine neue, unsere Wahrnehmung von Wirklichkeit erweiternde Kommunikationsform. Zunächst erweitert also Digitalisierung den Möglichkeitsraum der Kommunikation „und bedroht damit die bisherige Struktur und Kultur, die bisherigen Institutionen, Konventionen und Routinen, die auf die Modalitäten der älteren Verbreitungsmedien eingestellt sind“ [3]. Diese Bedrohung fasst Baecker in Anlehnung an Luhmann als „Überschusssinn“ und damit als eine neue Komplexität, die zu bereits zu bewältigender Komplexität hinzukommt [15]. Dabei argumentiert Baecker, dass auch der beobachtbare Reflex der Ablehnung neuer Kommunikationsmedien (er führt zum Beispiel den Buchdruck an) bereits eine Form der Auseinandersetzung und den Beginn der Bewältigung im Sinne der Nutzenerschließung darstellt [3].

Ohne den Ansatz von Baecker hier weiter auszuführen, kennzeichnet diese neue Komplexität eine Zunahme an Informationen bei gleichzeitiger Abnahme der Greifbarkeit ihrer Quellen, der Logik ihrer Kombination und damit dem bewussten Zugriff auf das, was sie mit uns als Subjekten macht.

Im Folgenden liegt der Fokus auf einem Teilphänomen von Digitalisierung, nämlich der Wirklichkeitserweiterung unseres Interaktionsraums um die Dimension der Virtualität.

Daraus ergibt sich zunächst die Frage „Was macht die Erfahrung virtuell vermittelter Welt mit uns als [...] Subjekte? [4]

Wikipedia liefert folgende Definition von „Virtualität“:

„Virtualität ist die Eigenschaft einer Sache, nicht in der Form zu existieren, in der sie zu existieren scheint, aber in ihrem Wesen oder ihrer Wirkung einer in dieser Form existierenden Sache zu gleichen. Virtualität bezeichnet also eine gedachte Entität, die in ihrer Funktionalität oder Wirkung vorhanden ist“ [16].

Virtualität ist insofern gedachte Wirklichkeit im Sinne eines „Als ob“, aber nicht im Sinne eines „Ist“. Sie ist eine Abstraktion des Möglichen, begegnet uns im Raum von Interaktion aber als scheinbar Greifbares, Reales. Wenn wir zum Beispiel die virtuell vermittelte Kommunikation über Videokonferenzsysteme im Rahmen mobilen Arbeitens betrachten, ist in diesem Verständnis unser:e Interaktionspartner:in zwar möglicherweise eine reale Person, sie erscheint uns auf dem Bildschirm aber als virtuelle Repräsentanz dieser Person. Vice versa gilt natürlich das Gleiche. Die konkrete Interaktion findet also zwischen virtuellen Repräsentanten der Interaktionspartner statt. Das wirkt sich auf unser Interaktionserleben aus, denn die virtuelle Repräsentation reduziert Aspekte von Wirklichkeit, die in realer, physischer Interaktion durchaus bedeutsam sind. Ein Beispiel dafür ist die Visualisierung auf einem 2D-Bildschirm gegenüber der physischen 3D-Wirklichkeit. Überdies sehen wir Mimik und Gestik nur in dem Ausschnitt, den die Kamera des Gegenübers erfasst in je nach Netzanbindung, Übertragungsrates und Bildschirmauflösung unterschiedlicher optischer und akustischer Qualität [4].

Um mögliche Folgen für die Interaktions- und Kooperationsqualität in virtuell vermittelter Kommunikation zu illustrieren, seien hier beispielhaft zwei empirische Arbeiten angeführt:

Weber et al. untersuchen aus sozialpsychologischer Perspektive die Auswirkungen des Arbeitens im Home Office auf die berufsbezogene Persönlichkeit und kommen dabei zu einer eher kritischen Einschätzung, vor allem mit Blick auf das Kontrollerleben in betrieblichen Arbeitsprozessen [17].

Brucks und Levav haben in einer Labor- und Feldanalyse in 5 Ländern in Europa, dem Mittleren Osten und Südasien Auswirkungen virtuell vermittelter Kommunikation auf die Qualität von Arbeitsprozessen und -ergebnissen untersucht. Dabei konnten sie zeigen, dass virtuell vermittelte Kommunikation die Entwicklung kreativer Ideen hemmt. Die Qualität der Auswahl weiter zu verfolgender Ideen ist gegenüber der physischen Interaktion vor Ort hingegen nicht schlechter. Die Autor:innen führen den Effekt auf die Kreativität darauf zurück, dass virtuell vermittelte Kommunikation den Fokus der Beteiligten auf den Bildschirm richtet und damit der Fokus auf kognitive Prozesse verringert wird [18].

Ergänzt werden diese und andere wissenschaftlich-empirisch erforschte Zusammenhänge durch anekdotische Erlebnisse, wie beispielsweise die Schwierigkeit, aus der pandemiebedingt in vielen Bereichen vollständigen Verlagerung aller beruflichen Interaktion und Kooperation in den virtuellen Raum, wieder in die Zusammenarbeit vor Ort zurückzukehren und Formen einer für die Subjekte individuell wie gemeinsam funktionalen Mischung virtuell vermittelter und physisch-realer Interaktion zu finden. Dabei scheint es sich nicht nur um eine Problematik zu handeln, die durch ein Regulationsvakuum zustande kommt. In vielen Organisationen, besonders

im behördlichen Kontext, wurden sehr schnell neue Regeln für genau dieses Problem verhandelt und kodifiziert. Trotzdem scheint es auf der konkreten Interaktionsebene zwischen den Subjekten weiterhin ein Feld für Konflikte, Unzufriedenheiten und mithin nicht bewältigte, vielleicht nicht bewältigbare, Komplexität zu bleiben. Mit diesem Beispiel wird deutlich, wie voraussetzungsvoll intersubjektiv geteilter Sinn für das Ausfüllen regulatorisch definierter Interaktionsrahmen zu sein scheint.

#### IV. KOMPLEXITÄTSANFORDERUNGEN AN DIE SUBJEKTE

Die, bis hierhin nur schlaglichtartig beschriebenen, komplexitätsstiftenden Phänomene, die Digitalisierung und Virtualität auf der Ebene intersubjektiver Interaktion hervorrufen, stellen uns vor die Aufgabe, diesen neuen Komplexitäten sowohl individuell als auch in organisationalen Kontexten handelnd zu begegnen. Der Handlungsbegriff impliziert dabei die Bezugnahme

- auf eine interpretierte materiale und soziale Wirklichkeit, einschließlich als gültig anerkannter Regulationen, als Ausgangspunkt des Handelns,
- auf eine Vorwegnahme der materialen und sozialen Handlungsfolgen im Sinne erwünschter oder nicht erwünschter Folgen und
- auf einen Prozess des Handlungsvollzugs im Zusammenwirken mit anderen.

In diesem Verständnis erfordert Handeln einen reflexiven Modus subjektiver Bewältigung von Komplexität [19], [20], [21].

Dabei geht es zunächst darum, die Wirklichkeitserweiterung durch Digitalität und Virtualität in ihrer fundamentalen Dimension für unser Weltverständnis zu erfassen. In einem zweiten Schritt geht es darum, als Subjekte die Möglichkeiten, die Verfügung über diese Welt (wieder) zu erlangen und zu erweitern.

Dies können wir als grundsätzliche Aufgabe von Bildung konzipieren, ausgehend von einer Definition des Bildungsbegriffs, die der Deutsche Ausschuss für das Erziehungs- und Bildungswesen bereits 1960 zeitlos formuliert hat:

Gebildet ist demnach der Mensch, „der in der ständigen Bemühung lebt, sich selbst, die Gesellschaft und die Welt zu verstehen und diesem Verständnis gemäß zu handeln“ [22].

Dieses Bildungsverständnis ist wesentlicher Ausgangspunkt des dtec.bw-Forschungsprojekts „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“. Es ist insofern Voraussetzungsvoll, als es den fundamentalen Kern des Verstehens einer durch Digitalisierung (und Virtualisierung) in vielfacher Hinsicht erweiterten Welt als notwendige Bedingung, in dieser Welt (weiterhin) als Subjekt im Zusammenwirken mit anderen Subjekten und in der Dialektik von Anpassung und Widerstand sein zu können, in den Fokus der Forschungs- und Bildungsarbeit stellt [23].

Forschungsgegenstand sind vor diesem Hintergrund verschiedene Dimensionen subjektiver Verfügungserweiterung von der reflektierten Erfahrung virtuell vermittelter Kommunikation und Interaktion bis hin zu neuen Formen von Kooperation in verteilten Teams.

Gemeinsam ist den verschiedenen Erfahrungskontexten von Digitalität und Virtualität, die das Projekt im Modus kontextualisierter Forschung in den Blick nimmt, das Bewusstsein für digitalisierte Datenverarbeitung, ihre Quellen und damit verbundene Interessen zu stärken, damit das bewusste Umgehen mit der virtuellen Repräsentanz so generierter Ausschnitte von Welt (im Sinne möglicher, aber nicht eins zu eins wirklicher Welt) zu ermöglichen und so die bewusste und verantwortliche Mitgestaltung dieser erweiterten Welt als Ideal mündiger Bürger:innen anzustreben [23].

#### DANKSAGUNG

Diese Forschungsarbeit im Forschungsprojekt *Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung* wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGeneration EU finanziert.

#### LITERATUR

- [1] K. Holzkamp, „Was heißt "Psychologie vom Subjektstandpunkt aus"? Überlegungen zu subjektwissenschaftlicher Theorienbildung,“ *Journal für Psychologie 1* (2), pp. 66-75, 1993.
- [2] R. Wrangham, Feuer fangen. Wie uns das Kochen zum Menschen machte - eine neue Theorie der menschlichen Evolution, München: Deutsche Verlags-Anstalt in der Verlagsgruppe Random House GmbH, 2009.
- [3] D. Baecker, „Wie verändert die Digitalisierung unser Denken und unseren Umgang mit der Welt? Ausgangspunkte einer Theorie der Digitalisierung,“ in *Gläß, R. (Hrsg.): Handel 4.0. Die Digitalisierung des Handels. Strategien, Technologien, Transformation.*, Berlin und Heidelberg, Springer Gabler, 2016, pp. 3 - 24.
- [4] A. Neusius, „Das Subjekt und die anderen - Implikationen von Virtualität,“ in *Schulz, M. / Büchter, K. / Neusius, A. / Pabst, A. / Kreuzmann, K. (Hrsg.): Perspektiven zur Subjektbildung*, Bielefeld, wbv-Verlag, 2024 (im Druck), p. o.S..
- [5] H. Kliemt, „Rahmen, Regeln und Regularitäten,“ in *Duschek, S. / Gaitanides, M. / Matiaske, W. / Ortman, G. (Hrsg.): Organisationen regeln. Die Wirkmacht korporativer Akteure*, Wiesbaden, Springer VS, 2012, pp. 111 - 128.
- [6] P. Preisendörfer, Organisationssoziologie. Grundalgen, Theorien und Problemstellungen. 2. Auflage, Wiesbaden: VS Verlag für Sozialwissenschaften, 2008.
- [7] C. Stöcker, Das Experiment sind wir, München: Karl-Blessing-Verlag in der Verlagsgruppe Random House GmbH, 2020.
- [8] U. / S.-L. H. Lubienetzki, Was wir uns wie sagen und zeigen: Psychologie der menschlichen Kommunikation, Berlin und Heidelberg: Springer, 2020.
- [9] C. Grey, A very short, fairly interesting and reasonably cheap book about studying organizations, London: Sage, 5th edition, 2022.

- [10] V. Tacke, „Formalität und Informalität. Zu einer klassischen Unterscheidung der Organisationssoziologie,“ in von Groddeck, V. / Wilz, S. M. (Hrsg.): *Formalität und Informalität in Organisationen*, Wiesbaden, Springer VS, 2015, pp. 37 - 92.
- [11] W. Bender, *Subjekt und Erkenntnis*, Weinheim: Deutscher Studien Verlag, 1991.
- [12] T. / M. S. Wendt, „Die Bildung der Organisation. Zur Komplexität von Management und Führung,“ *Gruppe. Interaktion. Organisation. Zeitschrift für angewandte Organisationspsychologie*, 15 November 2022.
- [13] M. Hennen, *Soziale Motivation und paradoxe Handlungsfolgen*, Opladen: Westdeutscher Verlag, 1990.
- [14] E. Meueler, *Die Türen des Käfigs. Wege zu Subjekt in der Erwachsenenbildung*, Stuttgart: Klett-Cotta, 1998.
- [15] N. Luhmann, *Die Gesellschaft der Gesellschaft*, Frankfurt / M.: Suhrkamp, 1997.
- [16] Virtualität [Online]. Available: <https://de.wikipedia.org/wiki/Virtualität>.
- [17] U. / K. K. / M. C. Weber, *Home Office. Auswirkungen auf die berufsbezogene Persönlichkeit*, Wiesbaden: Springer, 2022.
- [18] M. / L. J. Brucks, „Virtual communication curbs creative idea generation,“ *Nature, Vol 605*, pp. 108 - 112, 05 May 2022.
- [19] M. Schulz, „Fernausbildung geht weiter ... – Grundsätzliche Überlegungen und aktuelle Entwicklungen in Theorie und Praxis,“ in *Fernausbildung geht weiter... – Neue Beiträge zur Weiterentwicklung technologiegestützter Bildung aus pädagogisch-didaktischer Perspektive*, M. Schulz, H. Breyer und A. Neusius, Hrsg., Augsburg, Ziel-Verlag, 2006, pp. 9 - 46.
- [20] M. Schulz, „Fernausbildung verbindet... – Technologiegestützte Bildung im Spannungsfeld von Arbeiten und Lernen, Individuum und Gesellschaft,“ in *Fernausbildung fair-netzt alle... – Zum aktuellen Stand aus Forschung und Praxis der technologiegestützten Aus-, Fort- und Weiterbildung. Tagungsband zum 3. Fernausbildungskongress der Bundeswehr*, M. Schulz und A. Neusius, Hrsg., Augsburg, Ziel-Verlag, 2008, pp. 9 - 66.
- [21] M. Schulz, *Integrative Weiterbildung - Chancen und Grenzen. Konzeptionelle Überlegungen zur Integration allgemeiner, politischer und beruflicher Bildung*, Neuwied: Luchterhand, 1996.
- [22] D. A. f. d. E.-. u. Bildungswesen, *Zur Situation und Aufgabe der Deutschen Erwachsenenbildung*, Stuttgart: Klett-Cotta, 1960.
- [23] A. / S. M. Neusius, „Das dtec.bw-Forschungsprojekt „Kompetenzen für die digitale Arbeitswelt (KoDiA) – Ertüchtigung zur Digitalisierung“ als Beispiel kontextualisierter Forschung,“ in Fay, A. / Schulz, D. / Matiaske, W. / Schulz, M. (Hrsg.): *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, Bd. 1*, Hamburg, Eigenverlag HSU/UniBw H, 2022, pp. 311 - 320.

## **Kapitel IV**

### **Organisation – Personal – Arbeit – Leadership (OPAL)**

mit Beiträgen von

FernUniversität in Hagen

Harvard T. H. Chan School of Public Health

Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Institut für Angewandte Informatik (InfAI) e. V.

Katholieke Universiteit Leuven

Kompetenzzentrum für Public Management (KPM) – Universität Bern

Northern Business School

Q Agentur für Forschung

Pädagogische Hochschule Zürich

Universität der Bundeswehr München

vhw – Bundesverband Wohnen und Stadtentwicklung

WeTeK Berlin gGmbH

## **Vorwort zum dtec.bw-Dachprojekt „Organisation – Personal – Arbeit – Leadership (OPAL)“**

Bei der Konzeption der Forschungsstrategie im Rahmen des dtec.bw folgten wir der Leitannahme, dass Digitalisierung und Technik nicht nur Mittel für gesellschaftliche und mithin soziale Zwecke sind, sondern auch soziale Voraussetzungen haben. Vielfach beschworene, jedoch vage Begriffe wie ein innovationsfreundliches Klima oder Zukunftsfähigkeit galt es zu konkretisieren. Voraussetzungen in rechtlichen und administrativen Rahmenbedingungen waren es ebenso zu eruieren wie solche in Bildung und Organisation. Das Management von Innovationen in Betrieben und Verwaltungen sollte erkundet werden. Wirkungen und Folgen für Arbeit und Beschäftigung und darüber hinaus für Lebenslagen und Sozialstruktur waren in den Blick zu nehmen.

Es bedürfte längerer Ausführungen, ob und inwieweit die COVID-19-Krise unsere Gesellschaften unvorbereitet traf. Zumindest in den hier zur Diskussion stehenden Feldern konnte auf Instrumente, welche wie die Kurzarbeit in anderen Krisen erprobt waren, oder Maßnahmen, die wie das Homeoffice eine lange technologische und organisationale Inkubationszeit hatten, zurückgegriffen werden. Auf die Verknüpfung von Forschungsfragen zwischen den Technikwissenschaften und den sozialwissenschaftlichen Disziplinen im weitesten Sinne – von der Psychologie und der Pädagogik hin zu Ökonomie, Management, Recht und Soziologie – waren wir jedenfalls vorbereitet. Im Forschungscluster Organisation, Personal, Arbeit und Organisation (OPAL) ließen sich unmittelbar anschlussfähige Forschungsfragen und -projekte aktivieren oder entwickeln, deren Ergebnisse hier in Ausschnitten dokumentiert sind. Vielfach liegen bereits weitere Publikationen in Journalen und Fachzeitschriften vor, die nicht Eingang in unseren Übersichtsband finden konnten. Ferner sind die Projekte, auch wenn dieser Band zum Ende der ersten Förderperiode erscheint, noch nicht abgeschlossen.

Vielmehr blicken wir nach der erfolgreichen Evaluierung durch den Wissenschaftsrat auf die Phase der Verlängerung in den kommenden zwei Jahren. Dies hilft nicht nur Verzögerungen in vielen Projekten aufgrund der COVID-19-Krise auszugleichen. Vor allem können auch anschließende Forschungsfragen bearbeitet bzw. für eine Verstetigung von dtec.bw vorbereitet werden.

Die Rahmenbedingungen der Forschungsförderung wurden nicht zuletzt aufgrund der Zeitwende fokussiert. Dies ändert jedoch nichts an der Gültigkeit der eingangs formulierten Aussagen hinsichtlich der Verschränkung von technik- und sozialwissenschaftlichen Forschungsfragen. Das Forschungscluster OPAL ist mit früheren Projekten zum Management des Unerwarteten und aktuellen Datenerhebungen zur Resilienz von Organisationen, zur Cybersicherheit oder zum Katastrophenschutz auf konkrete Forschungsfragen im Verständnis der „human security“ justiert.

Hamburg, im November 2024

Wenzel Matiaske

Dachprojektleiter „Organisation – Personal – Arbeit – Leadership“

# Entwicklung eines offenen Lernangebots zur Vermittlung digitaler Kompetenzen für Studierende – Ein Ergebnis des Projekts DigiTaKS\*

Jan Schubert\*,  
Bettina Schasse de Araujo  
Digital Competence Center  
WeTeK Berlin gGmbH  
Berlin, Deutschland  
\*schubert@wetek.de

Mihail Miller\*, Stephan Klingner  
E.T. - Efficient Technology Integration  
Institut für Angewandte Informatik  
(InfAI) e. V.  
Leipzig, Deutschland  
\*miller@infai.org

Sabine Schmidt-Lauff  
Professur für Weiterbildung und  
lebenslanges Lernen  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
digitaks@hsu-hh.de

**Kurzfassung** – Das im Rahmen des Projekts DigiTaKS\* an der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg (HSU/UniBw H) entwickelte adaptierbare Lehr-/Lernpaket ComDigi S\* setzt einen innovativen technischen Ansatz zur Förderung digitaler Kompetenzen um. Es integriert fortschrittliche Funktionen einer modernen Learning Experience Plattform (LXP) mit dem etablierten und portablen SCORM-Standard, was eine hochgradig flexible und auf individuelle Lernbedürfnisse zugeschnittene Bildungserfahrung schafft, welche gleichzeitig maximale Kompatibilität bietet. Das hier skizzierte Teilprojekt (AP3) beleuchtet insbesondere den an DigComp 2.2 angelehnten und speziell auf die Zielgruppe der Studierenden zugeschnittenen Kompetenzrahmen und strukturiert Lernpakete nach den Prinzipien der Open Educational Resources (OER), um einen freien und umfassenden Wissensaustausch zu fördern. Weiterhin präsentiert dieser Beitrag konkrete Anwendungsfälle und Feedback von Nutzenden, um die Effektivität und Anwendbarkeit des ComDigi S\*-Ansatzes zu bewerten.

**Stichworte** – DigiTaKS\*, OER, DigComp, E-Learning, Hochschulbildung, digitale Kompetenzen, Lernplattform

## NOMENKLATUR

AP	Arbeitspaket
EU	Europäische Union
H5P	HTML5 (Hypertext Markup Language 5) Package
HSU/UniBw H	Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg
IT	Informationstechnologie
JSON	JavaScript Object Notation
KI	Künstliche Intelligenz
LMS	Lernmanagementsystem
LXP	Learning Experience Plattform
OER	Open Educational Resources
SCORM	Sharable Content Object Reference Model
xAPI	Experience API (Application Programming Interface)

## I. EINLEITUNG

Digitale Technologien haben sich in den letzten Jahren tiefgreifend auf die Bildungs- und Arbeitswelt ausgewirkt, so dass die Notwendigkeit, digitale Kompetenzen zu fördern, deutlich gestiegen ist – insbesondere in der Hochschulbildung [1]. Das Projekt DigiTaKS\* adressiert diesen Bedarf durch die Entwicklung verschiedener Konzepte und Werkzeuge – darunter das konfigurierbare und portable Lehr-/Lernpaket ComDigi S\*, das Test- und Trainingsmodule für die Vermittlung digitaler Kompetenzen bietet.

In diesem Beitrag werden die spezifischen Anforderungen an die Förderung digitaler Kompetenzen von Studierenden beschrieben und daraus die didaktischen und technischen Aspekte der Lösung ComDigi S\* abgeleitet. Die Lösung wird mit ihren Teilkomponenten vorgestellt, Erfahrungen aus ihrem praktischen Einsatz beschrieben, reflektiert und kritisch bewertet. Abschließend werden die Ergebnisse zusammengefasst und zukünftige Entwicklungsmöglichkeiten diskutiert.

## II. VERORTUNG IM RAHMEN DES PROJEKTES DIGITAKS\*

Der parallele und integrative Ansatz der Arbeitspakete im Projekt DigiTaKS\* (ausführliche Darstellung in [2]) ermöglichte eine zielgerichtete Konzeption und Entwicklung des Lernpakets ComDigi S\* (AP3) und förderte die Entstehung einer nachhaltigen, technologisch und sozial integrativen Lösung. Die Analyse der digitalen Praktiken der Studierenden, die in AP1 durchgeführt wurde, sowie die technischen Rahmenbedingungen der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg (HSU/UniBw H), die in AP2 als ‚digitales Basispaket‘ entwickelt und an die Studierenden ausgegeben wurden, hatten einen entscheidenden Einfluss auf die hier vorgestellten Entwicklungsarbeiten des AP3. Auch die Anforderungen an die Portabilität und Integrierbarkeit der Lösung, die sich aus den betrachteten Plattformen und Einsatzszenarien der beteiligten Partneruniversitäten – Julius-Maximilians-Universität Würzburg (AP4) und Universität Duisburg-Essen (AP5) – ergaben, mussten berücksichtigt werden. So sollte die in erster Linie für das Selbststudium konzipierte Lösung auch in hybriden Lehr-/Lernsettings sinnvoll einsetzbar sein, eine Verbreitung als Open Educational Resources (OER) ermöglichen und Aspekte der Barrierefreiheit berücksichtigen.

sichtigen. In dem agilen Entwicklungsprozess des Lehr-/Lernsystems wurden Synergieeffekte mit den Partneruniversitäten vor allem durch folgende Instrumente aktiviert und verstärkt:

#### A. Workshops

Die Anforderungsanalyse für das Lehr-/Lernpaket ComDigi S\* erfolgte in Workshops. Dabei wurden Rollen mit Anforderungen an die Lösung identifiziert (u. a. Studierende, Lehrende, Learning Designer:innen und Tutor:innen), „Personas“ [3] abgeleitet sowie deren Anforderungen an die Lösung gesammelt. Dies diente dazu, die Bedürfnisse und Erwartungen der verschiedenen Interessengruppen zu verstehen und in den Entwicklungsprozess zu integrieren. Darüber hinaus wurde über die Lösung innerhalb des Konsortiums in einem World Café-Format diskutiert. Dies förderte die kollektive Interaktion und den spontanen Austausch von Ideen, was zu einem tieferen Verständnis der kritischen strategischen Fragen führte [4]. Durch die Möglichkeit die Tische zu wechseln, konnten Ideen von einer Diskussionsrunde in die nächste getragen und die thematischen Erkenntnisse vernetzt und weiterentwickelt werden. Letztendlich förderte diese dynamische und partizipative Umgebung eine ganzheitliche Problembetrachtung sowie die darauf basierende Ableitung zielführender Designentscheidungen für ComDigi S\* (s. Kapitel IV).

#### B. Konsortialtreffen

In regelmäßigen Konsortialtreffen wurde das Lehr-/Lernpaket vorgestellt. Dabei wurde jeder Stand einer kritischen Reflexion durch das interdisziplinäre Konsortium ausgesetzt. Das Feedback diente als Basis für entsprechende Überarbeitungen.

#### C. Exposition und Evaluation

Durch den Einsatz der Module in Pilotstudien und realen Lernumgebungen konnten wertvolle Rückmeldungen aus der Praxis<sup>1</sup> gesammelt werden, welche eine laufende Optimierung der Lehr-/Lernpakets ermöglichten (s. Kapitel V).

### III. KOMPETENZRAHMEN UND DIDAKTISCHER ANSATZ

Das Verständnis des Begriffs „digitale Kompetenzen“ wurde bis in die frühen 2000er Jahre eher vom Zugang zu digitalen Technologien und kundigem Umgang mit Hard- und Software geprägt. Mit dem größer werdenden Einfluss und der Ausbreitung des Internets geriet auch die Nutzung der Informations- und Kommunikationstechnologie in den Vordergrund [5]. Heute umfasst der Prozess der Digitalisierung nahezu sämtliche gesellschaftlichen Bereiche und kann nicht mehr einzelnen Personengruppen oder Fachbereichen zugeordnet werden [6]. Jede:r Einzelne ist mit der Aufgabe konfrontiert, sich kritisch mit den verändernden Lebensbedingungen auseinanderzusetzen. ComDigi S\* dient in diesem Sinne als Unterstützung bei der Orientierung in einer sich zunehmend digitalisierenden Welt. Bei der Entwicklung und Erstellung eines Instruments zur Förderung digitaler Kompetenzen muss daher zunächst die Frage geklärt werden, wie digitale Kompetenzen zu definieren sind. Eine bildungstheoretische Auseinandersetzung mit dieser Frage und der Entwicklung von Kompetenzrastern ist seit Mitte der 2000er Jahre zu beobachten [7]. Im Jahr 2013 wurde mit dem DigComp der Europäischen Kommission eine ausführliche Darstellung über

die bildungspolitisch als nötig erachteten digitalen Kompetenzen der Bürger:innen vorgelegt [8]. Im Jahr 2018 hat die EU digitale Kompetenzen in ihre überarbeitete Empfehlung zu den acht Schlüsselkompetenzen für lebenslanges Lernen aufgenommen [9] und umfassend definiert. In der neuesten Version 2.2 aus dem Jahr 2022 identifiziert der DigComp fünf Kompetenzbereiche mit 21 Teilkompetenzen und acht Kompetenzniveaus und bietet damit einen umfassenden Strukturierungsrahmen für die Arbeit mit digitalen Kompetenzen [10].

Das in AP3 von der WeTeK Berlin gGmbH und dem InfAI e.V. entwickelte Selbstlerntool ComDigi S\* orientiert sich in seinem Verständnis digitaler Kompetenzen ebenfalls am DigComp, berücksichtigt aber zusätzlich die spezifischen Interessen, Lebens- und Lernwelten und Herausforderungen der Studierenden. Ziel des Tools ist es, die Studierenden für die Rolle als sogenannte *Digital Change Agents* zu befähigen. Das sind Individuen, die digitale Prozesse nicht nur verstehen, sondern auch in gesellschaftliche, politische und globale Zusammenhänge einordnen können und damit über die notwendigen technischen, aber auch strategischen und ethischen Kompetenzen verfügen, um digitale Prozesse in ihrem privaten und beruflichen Alltag zu gestalten. Als zukünftige Fach- und Führungskräfte bringen die Studierenden somit nicht nur technisches Wissen mit, sondern ordnen dieses auch gesellschaftlich ein und entwickeln einen kritisch-reflektierten Umgang mit Technologien.

ComDigi S\* ist mehrdimensional strukturiert und besteht aus aufeinander aufbauenden Ebenen. Zunächst lassen sich, analog zum DigComp-Modell, unterschiedliche Kompetenzbereiche identifizieren, die in fünf **Modulen** verschiedene Handlungs- und Wissensfelder abdecken:

- (1) Umgang mit digitalen Informationen und Daten
- (2) Kommunikation und Zusammenarbeit
- (3) Erstellung digitaler Inhalte
- (4) Sicherheit
- (5) Problemlösen

Die ComDigi S\*-Module beinhalten **Kurse**, von denen jeder eine spezifische Kompetenz behandelt. Die **Lerneinheiten** innerhalb der Kurse beleuchten einen spezifischen Aspekt der Kompetenz im Detail. Lernressourcen in Form von Texten, Videos, Bildern, Übungen usw. werden als **Lerninhalte** bezeichnet (s. Abbildung 1).

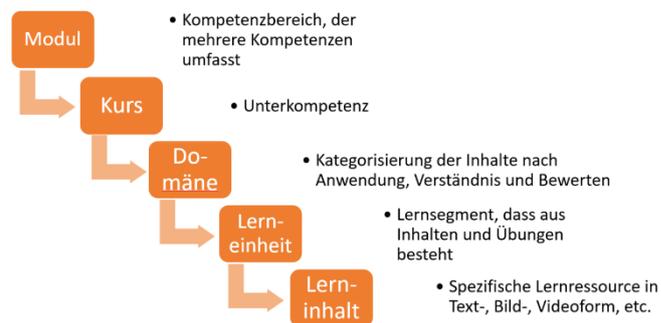


ABBILDUNG 1: EBENEN DES PÄDAGOGISCHEN MODELLS (COMDIGI S\*)

Anstelle der acht Kompetenzniveaus des DigComp werden in ComDigi S\* die drei kognitiven **Domänen** Verstehen, Anwenden und Bewerten in Anlehnung an die revidierte

<sup>1</sup> Neben der in Kapitel V erläuterten Evaluation im Hochschulkontext wurde die Lösung auch mit Externen im Rahmen von Weiterbildungsmaßnahmen und offenen Formaten pilotiert.

Bloom'sche Lernzieltaxonomie umgesetzt [11]. Diese Kategorisierung in Domänen verfolgt das Ziel, die Lerneinheiten interessengeleitet und unabhängig voneinander bearbeiten zu können. Die Domänen adressieren dabei unterschiedliche didaktische Ziele und sind zudem lebens- und lernweltspezifisch (hier: Hochschule/Studium/Alltag junger Menschen) orientiert:

- (1) **Verstehen** – In diesem Bereich liegt der Schwerpunkt auf der Vermittlung eines soliden Grundverständnisses digitaler Kompetenzen.
- (2) **Anwenden** – Dieser Bereich konzentriert sich auf die Anwendung von Werkzeugen und die Integration digitaler Technologien in den Alltag.
- (3) **Bewerten** – Im Rahmen dieses Bereichs werden die Studierenden zur kritischen Reflexion und Bewertung digitaler Technologien sowie zur Auseinandersetzung mit sozial-ethischen Implikationen angeregt.

Die Lerneinheiten sind so strukturiert, dass sie die Studierenden schrittweise an die jeweiligen Themen heranführen und zur kritischen Reflexion bewegen. Dabei werden verschiedene Lernressourcen wie Texte, Übungen, Videos und Bilder eingesetzt, die auf einschlägigen Fach- und Nachrichtenportalen, institutionellen Webseiten, Expert:innenforen und -blogs, Datenbanken und ähnlichem basieren. Dieses breite Spektrum an Quellen gestaltet den Lernprozess abwechslungsreich und motivierend.

Die didaktische Vermittlung selbst folgt den Prinzipien transformativer Lernprozesse, die in [2] und [12] näher eingeordnet werden. Im Verlauf der zu bearbeitenden Lerneinheit stoßen die Teilnehmenden immer wieder auf Reflexionsaufgaben, in denen sie aufgefordert werden, ihre eigenen Verhaltens-, Denk- und Handlungsweisen konstruktiv zu hinterfragen.

#### A. Auswahl und Relevanz der Inhalte

Relevante und für die eigenen Lebensumstände sinnvolle Inhalte fördern die Motivation und die individuelle Relevanz des Lernprozesses [13]. Im Vorfeld der Erstellung der Inhalte wurde sich der Zielgruppe – Studierende auch nicht-technikaffiner Fachbereiche – multimethodisch angenähert (s. Kapitel II), um darauf aufbauend die Inhalte mit den Bedarfen weiterer Stakeholder:innen, wie z. B. den militärischen Vorgesetzten, dem Verwaltungs- oder Lehrpersonal, in Einklang zu bringen. Auf diese Weise wurden Themenfelder identifiziert, die nicht nur im Studienkontext von Bedeutung sind, sondern weit darüber hinausgehen. In 50 Lerneinheiten werden Themen behandelt, die einerseits studienrelevante Tätigkeiten wie den Umgang mit Präsentations- und Textverarbeitungsprogrammen, Datenschutz, Arbeitsorganisation und Zeitmanagement aufgreifen. Andererseits reichen die Inhalte bis hin zur Vermittlung strategisch-globaler Zusammenhänge oder historisch-philosophischer Einordnungen der Informations- und Kommunikationstechnologie.

ComDigi S\* wurde mit KI-generierten Bildern und frei zugänglichen Inhalten und Materialien (OER) erstellt, die für den zielgruppenspezifischen Kontext didaktisiert wurden. Bei der Auswahl der Inhalte und der Gestaltung der Texte und Übungen wurden die Quellen stets sorgfältig auf ihre Unabhängigkeit und Aussagekraft geprüft bzw. entsprechend eingeordnet.

#### B. Kompetenzdiagnose und -auswertung

Um den Selbstlernprozess mit den Lerneinheiten anzuregen bzw. zu strukturieren, wird dieser sowohl durch einen Eingangs- als auch einen Abschlusstest in Form von Wissens- und Selbsteinschätzungsfragen eingrahmt. Nach Abschluss des Eingangstests erhalten die Studierenden eine Auswertung, die ihnen Orientierung und konkrete Hinweise zu relevanten Lerneinheiten und Domänen für den individuell gestaltbaren Lernweg gibt. Nach der Bearbeitung der Module wird ein Abschlusstest durchgeführt, der den Teilnehmenden sowie Vorgesetzten bzw. potentiellen Arbeitgeber:innen eine Einschätzung ihrer digitalen Kompetenzen ermöglicht.

#### IV. TECHNISCHE ASPEKTE

Aus den Workshops ergab sich eine Reihe funktionaler und nicht-funktionaler Anforderungen. Zum einen sollte die bereitgestellte digitale Technologie möglichst ohne Erweiterung der bestehenden Infrastruktur direkt und einfach integrierbar („Plug & Play“) und damit auch von nicht-technischem Personal bedienbar sein. Zum anderen sollte sie aber auch moderne Features wie Adaptierbarkeit und Interaktivität bieten. Die technische Umsetzung von ComDigi S\* umfasst mehrere Lösungsebenen, wobei durch die Kombination verschiedener Komponenten eine umfassende, bedarfsgerechte Umgebung geschaffen wird:

##### A. Authoring Tool (Adapt)

Das Adapt Authoring Tool basiert auf einem Open-Source-Framework [14] zur Erstellung von E-Learning Kursen. Es zeichnet sich durch mehrere Eigenschaften aus, die moderne Authoring Tools gemeinsam haben:

- (1) Unterstützung multimedialer Inhalte
- (2) Mobile Zugänglichkeit der Inhalte
- (3) Sowohl Online- als auch Offline-Nutzungsmöglichkeiten
- (4) Bereitstellung zur Integration in Lernmanagementsysteme (LMS) oder als eigenständige Web-Anwendung

Darüber hinaus unterstützt Adapt sowohl etabliertes SCORM-Tracking [15], als auch den moderneren Standard xAPI [16] (als Plugins). Letzterer hätte allerdings einen Learning Record Store vorausgesetzt, welcher in unserem Fall nicht vorhanden war (bzw. eine Erweiterung der technischen Infrastruktur speziell für ComDigi S\* erfordert hätte). Das Framework unterstützt den Aufbau von Kursen in einer hierarchischen Struktur mit Komponenten, Blöcken und Artikeln, die in einem klar strukturierten und gut lesbaren JSON-Format definiert sind. Durch den Einsatz des Adapt Authoring Tools können Kurse von mehreren Learning Designer:innen unabhängig voneinander gestaltet werden.

##### B. ComDigi S\* Edit|OER

Um eine kursübergreifende Sicht auf einzelne Lerneinheiten und Komponenten zu ermöglichen und damit Konsistenz, Wiederverwendbarkeit und Anpassung an spezifische Lehr-/Lernsettings zu gewährleisten, wurde ein zusätzliches Tool entwickelt: Edit|OER. Der Name spielt auf das Ziel der freien Verfügbarkeit der Materialien sowie deren Modifikation und Erweiterung an. Das Tool unterstützt die individuelle Zusammenstellung von Lernpaketen durch die Auswahl von

Kursen, die Zusammenstellung individueller Tests, die Bereitstellung von Druckversionen (Handouts) von Lerneinheiten und die Vorschau einzelner Lernkomponenten. Für letzteres wurde prototypisch die Übersetzung der Lernkomponenten in das H5P-Format erprobt [17]. Auf diesem Wege ist auch eine Verbreitung der Lernkomponenten in Micro-Lernszenarien denkbar [18]. Darüber hinaus erlaubt Edit|OER die kursübergreifende Pflege eines Glossars für Definitionen erklärungsbedürftiger Begriffe. Beim Export von Lernpaketen aus dem Tool werden die Glossareinträge bei Vorkommen in den Lerneinheiten erkannt und verknüpft.

Abbildung 2 zeigt beispielhaft, wie das Tool verwendet werden kann, um nach einem bestimmten Aufgabentyp zu filtern („matching“). Bei diesem Aufgabentyp muss die richtige Antwort aus einer Reihe von Optionen ausgewählt werden (Single Choice).

Frage (instruction)	Antwort
<input type="checkbox"/> Was ist Bias in KI-Systemen?	Ein Algorithmus, der Entscheidungen trifft <b>Eine Verzerrung, die in den Daten, Algorithmen oder Entscheidungen eines KI-Systems auftreten kann</b> Ein System, das Entscheidungen trifft, die frei von Vorurteilen und Diskriminierung sind Eine Technik zur Erkennung und Reduzierung von Bias in den Daten und Algorithmen

ABBILDUNG 2: FILTERN VON FRAGEKOMPONENTEN IN EDIT|OER

### C. ComDigi S\* Configurator

Darüber hinaus wurde ein weiteres Tool implementiert, das Anpassungsoptionen für exportierte Lernpakete bietet, wie z. B. die Beschränkung von Einstiegs- und Abschlusstests auf eine individuell festgelegte Anzahl von Fragen (s. Abbildung 3) sowie die Festlegung globaler Anzeigeoptionen.

Während Edit|OER die Zusammenstellung von Inhalten unterstützt, können Lehrende mit dem ComDigi S\* Configurator Aspekte der exportierten Lernpakete für unterschiedliche Lehr-/Lernsettings anpassen.

Richtige Antworten kennzeichnen

Auswahl kennzeichnen

Anzahl Fragen:

25

ABBILDUNG 3: KONFIGURATION VON TESTS IN COMDIGI S\*

### D. ComDigi S\* LXP

ComDigi S\* LXP vereint und bereichert die Lernerfahrung durch die Integration von Kursen und Gamification-Elementen (s. Abbildung 4), welche die Motivation der Lernenden erhöhen sollen [19]. Der Datenaustausch zwischen den einzelnen Kursen wird über einen API-Wrapper abgefangen und weitergeleitet. Ein kursübergreifendes Glossar ermöglicht die Suche nach bestimmten Begriffen und eine an den Informationsbedürfnissen orientierte Exploration der Lerninhalte. Das individuell konfigurierte Lernpaket lässt sich im ZIP-Format einfach in das LMS der Wahl hochladen und so neben anderen Lernmaterialien zur Verfügung stellen.

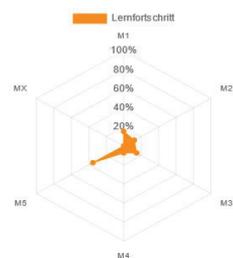


ABBILDUNG 4: „LERNRADAR“ IN COMDIGI S\* LXP

Zusammenfassend lässt sich sagen, dass die technische Umsetzung durch den Einsatz verschiedener Verfahren und einer modularen Architektur, die im Teilprojekt AP3 eruierten Anforderungen effektiv adressiert.

## V. PRAXISBEISPIELE UND ANWENDUNGSFÄLLE

Um die Praxistauglichkeit und Effektivität von ComDigi S\* zu testen, wurde es entlang des Entwicklungsprozesses in verschiedenen Szenarien erprobt. Die „Open Campus“-Veranstaltung an der HSU/UniBw H fand am 18. Juni 2022 statt und richtete sich in erster Linie an Studierende und Alumni der Universität sowie deren Familien. Den Besucher:innen bot sich die Möglichkeit, ihre Vorkenntnisse im Rahmen eines Einstiegstests zu überprüfen und aufbauend darauf ausgewählte Lerneinheiten auszuprobieren. Die Teilnahme erfolgte explorativ, wobei während der Teilnahme ein offener Dialog und nach der Nutzung Online-Umfragen Raum für Feedback boten. Insgesamt nahmen zwölf Teilnehmende am Einstiegstest und zehn an der Bearbeitung der Lerneinheiten teil. Die modulare Struktur wurde von den Interessent:innen positiv bewertet, insbesondere die Möglichkeit, sich auf Lerneinheiten zu konzentrieren, die aufgrund von Wissenslücken als sinnvoll erachtet wurden. Darüber hinaus gab es aber auch Anmerkungen zur Ergonomie und Übersichtlichkeit der Kursmaterialien, die wertvolle Hinweise zur weiteren Entwicklung des Lehr-/Lernpakets lieferten, insbesondere zur Verbesserung der Navigation, zur einfacheren Handhabung und zur inhaltlichen Verdichtung.

Die zweite Evaluationsphase von ComDigi S\* wurde im November 2022 an der HSU/UniBw H durchgeführt. Ziel war es, das Lernpaket in einem realistischen Bildungsszenario zu testen. Die Evaluation umfasste drei Lehrveranstaltungen mit unterschiedlichen Studierendengruppen, von Bachelor-Studierenden im ersten Trimester bis hin zu Master-Studierenden (n=50). Alle beteiligten Studierenden absolvierten zunächst einen Einstiegstest und bearbeiteten anschließend selbständig Lerneinheiten auf verschiedenen „Kursniveaus“. Die Lerninhalte wurden im SCORM 1.2-Format bereitgestellt und zuvor in ILIAS [20] hochgeladen. Die Datenerhebung erfolgte mittels quantitativer und qualitativer Methoden. Dazu gehörte die Erfassung von Interaktionen mit dem Lernpaket, eine Erhebung mittels strukturierter Fragebögen sowie offene Diskussionsrunden. Die Mehrheit der Studierenden würdigte den gut strukturierten und kompakten Inhalt des Kurses, die visuelle Gestaltung und die intuitiven, interaktiven Elemente des Lernpakets. Die Ergebnisse der Auswertung zeigten allerdings auch, dass die Mehrheit der Studierenden auf das Kursniveau A (Verstehen) eingestuft wurde. Dies deutete darauf hin, dass der Einstiegstest nicht optimal auf die Ausgangssituation bzw. Bedarfe der Studierenden abgestimmt war. Im Nachtrag wurde beschlossen, den zuvor erläuterten Begriff der „kognitiven Domänen“ statt „Niveaus“ zu nutzen, um den Grad der Wertung zu verringern (s. Kapitel III) sowie statt einer Emp-

## LITERATUR

fehlung bestimmter Lerninhalte, eine deskriptive Test-Auswertung zur Anregung des selbstgesteuerten Lernprozesses bereitzustellen. Technische Probleme sowie Verbesserungspotenziale in der Navigation und Interaktion mit dem Lernpaket wurden identifiziert und nachfolgend bearbeitet.

In der dritten Evaluationsphase von ComDigi S\* hatten die Studierenden vom 28.11. bis 20.12.2023 Zeit zum Selbststudium. Unter den 16 Studierenden, die das Lernpaket bearbeitet haben, gab es große Unterschiede, was die investierte Zeit und Anzahl der Interaktionen betreffen. Das Feedback, das in einer Online-Umfrage und mündlich in einer Seminarsitzung eingeholt wurde, beinhaltet positive Aspekte wie eine intuitive Benutzeroberfläche, abwechslungsreiches Lernmaterial und eine logische Lernstruktur. Kritisiert wurden Darstellungsprobleme auf Apple-Geräten sowie Unverständlichkeiten bei Testauswertungen und Navigationselementen. Verbesserungsvorschläge betrafen u. a. einen stärkeren Praxisbezug und die Möglichkeit, die Anredeform individuell anzupassen.

## VI. ZUSAMMENFASSUNG UND AUSBLICK

Mit ComDigi S\* wird ein System zur Verfügung gestellt, das dauerhaft wartbar, bedarfsgerecht anpassbar und erweiterbar ist. Das Softwarepaket wurde so konzipiert, dass es auf einer Vielzahl von Plattformen lauffähig ist. Das Paket befindet sich noch in der Entwicklung und erfährt durch kontinuierliche Updates Verbesserungen hinsichtlich Stabilität, Kompatibilität und Ergonomie. Die Verwendung standardisierter Datenformate wie SCORM 1.2 und zukünftig xAPI fördert die nahtlose Kommunikation zwischen technischen Plattformen und soll somit zu einer kohärenten Lernerfahrung beitragen.

ComDigi S\* wird nach Projektende als OER unter einer Creative Commons Lizenz der Öffentlichkeit zur Verfügung gestellt, um die Verbreitung und Anpassung der Inhalte für individuelle Lehr-/Lernsettings bzw. für weitere Zielgruppen in anderen Bildungsbereichen (bspw. Aus- und Weiterbildung) zu ermöglichen. Dies entspricht dem Selbstverständnis von ComDigi S\*, digitale Kompetenzen ganzheitlich und interdisziplinär zu vermitteln. Module, Kurse und Lerneinheiten können je nach Zielgruppe, persönlicher Präferenz oder Vorwissen der Teilnehmenden mittels der entwickelten Werkzeuge ausgewählt werden. Gleichzeitig ermöglichen sie Lehrenden, die Inhalte individuell anzupassen und für unterschiedliche Bedarfe und Einsatzgebiete weiterzuentwickeln. Durch diese Flexibilität soll sichergestellt werden, dass die Lernenden die für sie relevanten Kompetenzen sowohl gezielt wie auch spielerisch, motiviert und selbstgesteuert aufbauen und vertiefen können.

## DANKSAGUNG

Die Autor:innen bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. Das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr – ist ein von den Universitäten der Bundeswehr Hamburg und München gemeinsam getragenes wissenschaftliches Zentrum und Bestandteil des Konjunkturprogramms der Bundesregierung zur Überwindung der COVID-19-Krise. Es unterliegt der akademischen Selbstverwaltung. Die Mittel, mit dem das dtec.bw ausgestattet wurde, werden an beiden Universitäten der Bundeswehr zur Finanzierung von Forschungsprojekten und Projekten zum Wissens- und Technologietransfer eingesetzt. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

- [1] P. S. Petri, R. Krempkow, M. Ebner and B. Getto, "Editorial: Digitale Kompetenzen – Zur Rolle dieser „Future Skills“ im Hochschulkontext: Definition und Bedeutsamkeit, Messung und Förderung," *Zeitschrift für Hochschulentwicklung*, vol. 19, 2024.
- [2] S. Schmidt-Lauff, J. Schwarz, T. Rosemann, M. Rathmann and J. Schiller, "DigiTaKS\*-Digitale Schlüsselkompetenzen für Studium und Beruf," in *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw - Band 1 · 2022*, D. Schulz, A. Fay, W. Matiaske and M. Schulz, Eds., Hamburg, dtec.bw – Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, 2022, pp. 329-334.
- [3] W. W. Sim and P. S. Brouse, "Empowering Requirements Engineering Activities with Personas," *Procedia Computer Science*, vol. 28, p. 237–246, 2014.
- [4] J. Monforte, J. Netherway and B. Smith, "The world café is an unmethod within co-produced research," *Qualitative Research in Psychology*, vol. 20, p. 398–419, 2023.
- [5] R. Bachmann, F. Hertweck, R. Kamb, J. Lehner, M. Niederstadt and C. Rulff, "Digitale Kompetenzen in Deutschland – eine Bestandsaufnahme," *RWI Materialien*, vol. Heft 150, pp. 4-63, 2021.
- [6] S. Kistler, A. Puzio, A. M. Riedl and W. Veith, "Einleitung," in *Digitale Transformationen der Gesellschaft – Sozialethische Perspektiven auf den technologischen Wandel*, S. Kistler, A. Puzio, A. M. Riedl and W. Veith, Eds., Münster, Aschendorff Verlag GmbH & Co. KG, 2023, pp. 9-15.
- [7] N. Law, D. Woo, J. de la Torre and G. Wong, A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2, Montreal: UNESCO Institute for Statistics, 2018.
- [8] A. Ferrari, DIGCOMP: A framework for developing and understanding digital competence in Europe, Y. Punie and B. N. Brečko, Eds., Luxemburg: Amt für Veröffentlichungen der Europäischen Union, 2013.
- [9] Generaldirektion Bildung, Jugend, Sport und Kultur (Europäische Kommission), Key competences for lifelong learning, Luxemburg: Amt für Veröffentlichungen der Europäischen Union, 2019.
- [10] R. Vuorikari, S. Kluzer and P. Yves, DigComp 2.2: The Digital Competence Framework for Citizens – With new examples of knowledge, skills and attitudes, Luxemburg: Amt für Veröffentlichungen der Europäischen Union, 2022.
- [11] L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths and M. C. Wittrock, A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives, 2 ed., L. W. Anderson and D. R. Krathwohl, Eds., New York: Allyn & Bacon, 2001.

- [12] M. Singer-Brodowski, "Transformative Bildung durch transformatives Lernen. Zur Notwendigkeit der erziehungswissenschaftlichen Fundierung einer neuen Idee," *ZEP : Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik*, vol. 1'16, pp. 13-17, 2016.
- [13] H. Siebert, *Lernmotivation und Bildungsbeteiligung*, Bielefeld: Bertelsmann, 2006.
- [14] "Adapt Learning," [Online]. Available: <https://www.adaptlearning.org/>. [Accessed 30. April 2024].
- [15] "SCORM," [Online]. Available: <https://adlnet.gov/projects/scorm/>. [Accessed 30. April 2024].
- [16] "xAPI," [Online]. Available: <https://adlnet.gov/projects/xapi/>. [Accessed 30. April 2024].
- [17] "H5P," [Online]. Available: <https://h5p.org/>. [Accessed 30. April 2024].
- [18] K. Leong, A. Sung, D. Au and C. Blanchard, "A review of the trend of microlearning," *Journal of Work-Applied Management*, vol. 13, p. 88–102, 2020.
- [19] A. Khaldi, R. Bouzidi and F. Nader, "Gamification of e-learning in higher education: a systematic literature review," *Smart Learning Environments*, vol. 10, pp. 1-31, 2023.
- [20] "ILIAS," [Online]. Available: <https://www.ilias.de/>. [Accessed 30. April 2024].

# Digitale Transformation und hybride Führungsidentitäten in der Sozialen Arbeit

Dorina Kurta\*, Jaromir Junne  
Institut für Controlling und Unternehmensrechnung  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
\*kurtad@hsu-hh.de

**Kurzfassung** – Der vorliegende Beitrag untersucht die Herausforderungen, denen sich Führungskräfte in kleinen gemeinnützigen Organisationen der Sozialen Arbeit gegenübersehen. Der Beitrag zeigt, wie Führungskräfte, die aus der Sozialen Arbeit kommen, einen Hybridisierungsprozess durchlaufen, in dem sie ihre fachliche Expertise als Sozialarbeiter:innen mit hinzukommenden betriebswirtschaftlichen Managementanforderungen verbinden. In Interviews mit Führungskräften konnten wir drei typische Phänomene beobachten: professionelle Selbstbeschreibungen, die Soziale Arbeit und Managementaufgaben – nicht immer konfliktfrei – miteinander verbinden, neu gefasste Rollen und Rechenschaftspflichten durch die Nutzung von Fachsoftware und Veränderungen in den Beziehungen zwischen Führungskräften und Mitarbeitenden. Da digitale Instrumente immer auch soziale Rollen beeinflussen, erfordert die Nutzung von Fachsoftware daher nicht nur technisches Know-how, sondern auch soziales und organisatorisches Fingerspitzengefühl.

**Stichworte** – Digitalisierung, Fachsoftware, Führung, Soziale Arbeit, Non-Profit-Organisationen (NPOs), Hybridisierung

## NOMENKLATUR

DiWop            Projekt zur Förderung der Digitalisierung in der freien Wohlfahrtspflege

## I. EINLEITUNG

Die digitale Transformation stellt Führungskräfte in kleinen gemeinnützigen Organisationen der Sozialen Arbeit vor eine komplexe Herausforderung: Sie müssen ihre berufliche Identität zwischen fachlicher Expertise als Sozialarbeiter:innen und betriebswirtschaftlichen Managementanforderungen finden. Die Einführung von Fachsoftware im Zuge der Digitalisierung erfordert von ihnen Kompetenzen in Bereichen wie Budgetierung, Kostenrechnung und Preisgestaltung, während sie gleichzeitig ihre Kernexpertise als Sozialarbeiter:innen nicht vernachlässigen wollen. Dieser Balanceakt kann zu Spannungsfeldern und Konflikten führen, die das professionelle Selbstverständnis von Führungskräften berühren [1], [2]. Sie sehen sich mit der Aufgabe konfrontiert, einerseits ihre Organisation wirtschaftlich zu führen und andererseits professionelle Ansprüche aus der Sozialen Arbeit zu erfüllen und die Beziehung zu ihren Mitarbeitenden möglichst wenig zu belasten. Die Vereinbarung dieser beiden Ziele kann zu äußeren

und inneren Konflikten führen. Der vorliegende Beitrag untersucht diesen Hybridisierungsprozess, bei dem Fachexpertise und Managementaufgaben miteinander verschränkt werden, und beleuchtet die daraus resultierenden Herausforderungen für Führungskräfte in der Sozialen Arbeit. Dabei wird insbesondere der Einfluss der Digitalisierung als Katalysator für diese Entwicklung analysiert und diskutiert, wie Organisationen diesen Prozess gestalten können.

## I. HYBRIDISIERUNG VON FACH- UND MANAGEMENTAUFGABEN FÜR FÜHRUNGSKRÄFTE IN SOZIALEN ORGANISATIONEN

Dieser Beitrag orientiert sich am Begriff der „Hybridisierung“, der von Kurunmäki [3] für den medizinischen Bereich entwickelt wurde. Hybridisierung beschreibt den Prozess, bei dem Fachkräfte ihre berufliche Identität erweitern, um Managementaufgaben zu integrieren, ohne ihre ursprüngliche fachliche Identität aufzugeben. Der Begriff der Hybridisierung geht davon aus, dass Management- und Controllinginstrumente die Tendenz haben, sich in allen gesellschaftlichen Bereichen auszubreiten, insbesondere in solchen Bereichen, in denen bislang fachliche Expertise dominierte [4], [5]. Diese These findet sich auch in Arbeiten zum *new public management* [6], [7], wengleich inzwischen auch deutlich beobachtet wird, dass Fachexpertise und auch professionelle Identität gegen eine zu starke Management-Ausrichtung mobilisiert wird [8], [9], [10]

### A. Hybridisierung in der Sozialen Arbeit

Noordegraaf [1] und Noordegraaf & Steijn [11] haben das Konzept der Hybridisierung auf die Soziale Arbeit übertragen. Sie zeigen, dass Führungskräfte in diesem Bereich zunehmend gezwungen sind, Managementkompetenzen zu entwickeln und in den Beschreibungen ihrer professionellen Identität fachliche und betriebswirtschaftliche Anteile zunehmend miteinander verbinden und vermischen. Dieser Prozess verläuft jedoch nicht ohne Herausforderungen, da die Vereinbarung der unterschiedlichen Identitäten als Sozialarbeiter:in und Manager:in zu Rollenkonflikten und Spannungen führen kann. Führungskräfte müssen sich neue Kompetenzen in betriebswirtschaftlichen Bereichen wie Budgetierung, Kosten- und Leistungsrechnung und Preisgestaltung [12], aber auch

Mitarbeiterführung, Personalrecht und *human resource management* aneignen, welche die Organisation und ihre Dynamiken als Ganzes betreffen, während sie gleichzeitig ihre Kernexpertise als Sozialarbeiter:innen nicht vernachlässigen wollen, da sie auch für die Qualität der fachlichen Leistungen ihrer Organisation verantwortlich sind. Diese beiden Ziele können zu erheblichen Konflikten führen, da die Anforderungen der Managementaufgaben (z. B. Kostendruck) im Widerspruch zu den ethischen und professionellen Standards Sozialer Arbeit stehen können (z. B. intensive Beziehungsarbeit mit Klienten). Solche Spannungen können die berufliche Identität der Führungskräfte erheblich belasten [13].

### B. Digitalisierung als Treiber der Hybridisierung

Digitale Instrumente fungieren als Katalysator für Hybridisierungsprozesse in Non-Profit-Organisationen der Sozialen Arbeit, da sie den Fokus auf das Management von Daten und Prozessen lenken und die Rollen der verschiedenen Nutzer- und Administratorengruppen (z. T. neu) definieren. Dies greift immer auch in die fachliche Arbeit ein und entfaltet vielfältige, nicht immer vorhersehbare Wirkungen auf die Organisationen Sozialer Arbeit [14]. Digitale Instrumente verlagern Aufgaben nicht nur in Software, sondern re-priorisieren sie dabei auch und beeinflussen, wie Sozialarbeiter:innen ihre Arbeit ausführen und über sie denken [15]. Diese Wirkungen können Sozialarbeit auch behindern anstatt unterstützen [15], [16]. Parton [17] etwa beobachtet, wie die Verbreitung von Informationstechnologie in der Sozialen Arbeit über die letzten dreißig Jahre eine Verschiebung von einer narrativen hin zu einer stärker datenbankbasierten Denk- und Arbeitsweise mit sich gebracht hat. Er fragt, inwiefern Sozialarbeit eigentlich noch im Bereich von Menschen und ihren sozialen Beziehungen tätig sei, und nicht vielmehr im Bereich des Informationswesens.

Digitale Instrumente, meist in Form einer Fachsoftware, werden in der Sozialen Arbeit insbesondere für die Dokumentation von Klientenkontakten und von erbrachten Leistungen eingesetzt. Während Sozialarbeiter:innen sich traditionell vor allem ihren professionellen Werten und ihrem Selbstverständnis als Sozialarbeiter:innen verpflichtet fühlten, gab es immer schon auch bürokratische Rechenschaftspflichten innerhalb ihrer Organisationen und gegenüber den finanzierenden Behörden. Es wird jedoch international beobachtet, dass diese bürokratischen Rechenschaftspflichten mit digitalen Instrumenten intensiver werden [18] – und Rechenschaftspflichten wirken sich immer auch auf die Rollen und das (professionelle) Selbstverständnis der Beteiligten aus [19], [20], [3].

### C. Veränderungen der Beziehung zwischen Führungskräften und Mitarbeitenden

Die Einführung digitaler Instrumente bietet einen besonders günstigen Anlass zur Neuverhandlung professioneller Rollen. IT-Beauftragte und Prozessverantwortliche werden ernannt, Zugriffs- und Nutzerrollen in der Software definiert. IT-Affinität und -Expertise werden zu einer wertvollen professionellen Kompetenz.

Führungskräfte treiben die Digitalisierung voran, was Managementkompetenzen erfordert. Andererseits leisten sie Beziehungsarbeit mit ihren Mitarbeitenden, um die Akzeptanz und Nutzung der neuen Technologien zu gewährleisten. De Mesa [21] betont, dass unzureichende Kommunikation bei der Einführung von Fachsoftware die Beziehung zu den Mitarbeitenden belasten und Spannungen erzeugen kann. Insbesondere Befürchtungen neuer Formen von Transparenz und Kontrolle und Ängste vor einem Verlust professioneller Souveränität

lassen sich in vielen Digitalisierungsprozessen beobachten und beeinflussen die Beziehung zwischen Führungskräften und Mitarbeitenden. Frühere Studien fordern daher partizipativere Design-Ansätze für IT in der Sozialen Arbeit [16].

Vor diesem Hintergrund wollen wir in den folgenden Abschnitten auf die Erfahrungen von Führungskräften in sozialen Organisationen detaillierter eingehen.

## II. ERFAHRUNGEN VON FÜHRUNGSKRÄFTEN IN DER SOZIALEN ARBEIT

Im Rahmen des dtec.bw-Projektes „DiWop - Förderung der Digitalisierung in der freien Wohlfahrtspflege“ haben wir Interviews mit Führungskräften aus 13 kleineren gemeinnützigen Organisationen der Eingliederungshilfe und eine Fokusgruppe mit Sozialarbeiter:innen durchgeführt, die inzwischen in der Leitung ihrer Organisation arbeiten. In den folgenden Abschnitten stellen wir Beobachtungen und Ergebnisse aus diesen Gesprächen vor, insbesondere zu potenziellen Identitätskonflikten von Führungskräften, der Rolle von Fachsoftware, sowie Veränderungen in den Beziehungen zwischen Mitarbeitenden und Führungskräften. Die eingeflochtenen Zitate dienen der Illustration der jeweiligen Phänomene.

### A. Identitätskonflikte

Berufliche Identitätskonflikte können entstehen, wenn die berufliche Rolle und die damit verbundenen Aufgaben im Wandel begriffen sind. Besonders im Management sozialer Einrichtungen treten solche Konflikte auf, da die Führungskräfte überwiegend aus einem sozialarbeiterischen Hintergrund stammen und sich nun in Managementaufgaben wiederfinden. Ein Geschäftsführer beantwortet die Frage nach seiner professionellen Identität wie folgt:

*„Wenn ich gefragt bin, was ich beruflich mache, dann sage ich immer: 'Ich habe mal Sozialarbeiter gelernt, aber eigentlich bin ich das jetzt nicht mehr.' Also eigentlich ist es eine Managementaufgabe, die eben darin besteht, einmal zu gucken, dass wir nicht mehr Geld ausgeben, als wir einnehmen. Und dass wir die richtigen Mitarbeiter am richtigen Platz haben und sie befähigen, die Arbeit gut zu machen.“ (Geschäftsführer)*

Diese Aussage beschreibt einen Übergang von einer ursprünglichen beruflichen Identität in der Sozialen Arbeit hin zu Managementaufgaben, die „eigentlich“ eine neue professionelle Identität bildet. Ein Sozialarbeiter, der außerdem zunehmend die Aufgaben eines IT-Verantwortlichen übernommen hat, beschreibt seine inzwischen hybride Arbeit wie folgt:

*„Ich stelle mich als Sozialarbeiter [...] vor. Ich stelle auch immer vor, was wir machen in unserer Arbeit. Das stelle ich eigentlich immer vor, obwohl ich, wie gesagt, das ja nur noch mit gewissem Anteil meiner Arbeit mache. Aber das stelle ich schon noch immer vor. Ich bin noch Sozialarbeiter für mich.“ (IT-Beauftragter)*

Die Entscheidung, sich trotz eines inzwischen verlagerten Tätigkeitsfeldes als Sozialarbeiter vorzustellen, orientiert sich auch an den sozialen Reaktionen im jeweiligen Umfeld:

*„Die Reaktionen sind so geil unterschiedlich, ne? Wenn ich erzähle, was ich mache, Menschen mit Lerneinschränkungen behilflich zu sein, [...] ‚Oh geil!‘, das ist ja toll und so. Wenn ich erzähle: ‚Ja, ich mache immer mehr IT-Aufgaben in diesem Bereich‘, dann so: ‚Aha‘, langweilig eigentlich [lacht].“ (IT-Beauftragter)*

Die primäre Identifikation als Sozialarbeiter:in war allen Führungskräften, mit denen wir gesprochen haben, sehr wichtig. Eine Führungskraft beschreibt den Übergang zu mehr Leitungsaufgaben als Ergebnis eines längeren Prozesses:

*„Eigentlich war das eine schleichende Entwicklung, die über, ich sage mal, in den letzten 14 Jahren mindestens sechs, sieben kleine Wechsellpunkte hatte.“ (Geschäftsführerin)*

Die Beschreibung eines schleichenden Wandels spricht dafür, dass die Veränderung der eigenen Rolle nicht immer bewusst und in qualitativen Sprüngen geschieht. Alle Führungskräfte, mit denen wir gesprochen haben, begrüßen ihre berufliche Entwicklung grundsätzlich, bedauern aber zum Teil, dass sie so nun weniger Zeit für die direkte Arbeit mit Klienten haben, die ihnen am Herzen liegt, wenn Verwaltungs- und Managementaufgaben in den Vordergrund rücken.

*„Also ich versuche, mir das Menschliche zu bewahren. In dem ganzen Zeitalter der Digitalisierung, KI [künstliche Intelligenz] etc., versuche ich das. Ich glaube, das nehme ich noch mit, bis ich zur Rente gehe. Dass wird nicht aufhören, miteinander irgendwie in Kontakt zu sein. [...] Und das ist etwas, was ich irgendwie versuche, auch in meiner Position irgendwie selber auch nicht zu vergessen. Vor lauter Anforderungen oder Aufgaben, dass ich im Kontakt bleibe mit den Leuten und eigentlich mitkriege, was passiert da eigentlich?“ (Geschäftsführerin)*

Die meisten Führungskräfte haben dabei von Konflikten berichtet, die sich aus der wirtschaftlichen Verantwortung für die gesamte Organisation und professionellen Werthaltungen ergeben kann.

*„Das macht mir auch keinen Spaß, überhaupt nicht. Ganz im Gegenteil. Also, ich glaube, das unterschätzen immer ganz viele, was es bedeutet, Verantwortung für so ein Unternehmen zu haben. Und ich sitze jetzt auch nicht hier und habe Spaß daran, irgendwelche Entscheidungen zu treffen, wo ich auch selber denke: "Nee, ist jetzt blöd." Oder, wenn neue Anforderungen kommen, das zu übertragen. Also ich würde manche Dinge [aus sozialpädagogischer Sicht] gerne anders machen, ist ja logisch.“ (Geschäftsführerin)*

Eine solche oder ähnliche Ambivalenz tauchte in allen Gesprächen, die wir mit Führungskräften aus der Sozialen Arbeit geführt haben, auf. Meist spiegelten sich die Verschränkung von Management- und fachlichen Zielen auch in den Digitalisierungsprojekten der Organisationen wider, insbesondere bei der Einführung und Nutzung von Fachsoftware.

#### B. Neu gefasste Rollen und Rechenschaftspflichten durch Fachsoftware

Die Einführung neuer Technologien, wie beispielsweise einer neuen Fachsoftware, stellt soziale Einrichtungen gleich vor mehrere Herausforderungen. Gerade in kleinen Organisationen nehmen die involvierten Führungskräfte und Sozialarbeiter:innen bei der Einführung neuer digitaler Instrumente mehrere, z. T. für sie neue Rollen wahr, da es für diese Aufgaben kein extra Personal zur Verfügung steht.

*„Ich muss jetzt nicht nur Sozialarbeiter sein, sondern auch Buchhalter, IT-Experte und Projektmanager. Die neue Fachsoftware zwingt mich, mich mit Dingen auseinanderzusetzen, die weit über meine ursprüngliche Ausbildung hinausgehen.“ (Sozialarbeiterin)*

Fachsoftware ist in Hinblick auf Hybridisierungsprozesse besonders relevant, weil sie sowohl fachliche als auch betriebswirtschaftliche Aufgaben unterstützt und gleichzeitig ein zentrales Bindeglied beider Bereiche bildet. Sie kommt vor allem zur Arbeitsplanung, Dokumentation, Abrechnung und Controlling zum Einsatz und greift dadurch auch in eingespielte, soziale Rollen ein. Ein Geschäftsführer beschreibt die neue Fachsoftware seiner Organisation als Vermittlerin zwischen Verwaltung und Sozialer Arbeit:

*„Die Fachsoftware steht quasi zwischen uns und den Klienten. Einerseits ermöglicht sie eine effizientere Verwaltung, andererseits verändert sie die Art, wie wir Sozialarbeit betreiben. Wir müssen ständig zwischen diesen beiden Welten vermitteln.“ (Geschäftsführer)*

Dieses Zitat verweist auf die Notwendigkeit, die Tätigkeiten zwischen technischen Anforderungen, Verwaltungsprozessen und klientenzentrierter Sozialarbeit zu balancieren.

Fachsoftware, die meist für Dokumentation, Dienstplanung, Abrechnung und Auswertung der eigenen Tätigkeiten genutzt wird, definiert Zugriffsrechte und Rechenschaftsbeziehungen über die individuellen Nutzerkonten meist sehr explizit – z. T. deutlich expliziter und detaillierter als in analogen Verfahren. Damit kommt ihr eine potenziell mächtige Funktion zu und ihre Einsatzweise gibt in vielen Organisationen Anlass zu Diskussionen über Vertrauen und Transparenz, der Beziehung zwischen Führungskräften und Mitarbeitenden (s. nächster Abschnitt) und der Verbindung aus Managementanforderungen und fachlichem Nutzen. Dabei wird meist begrüßt, dass Fachsoftware ermöglicht, die individuelle Begleitung einzelner Klienten besser zu organisieren, insbesondere, wenn verschiedene Personen mit ihnen arbeiten. Gleichzeitig wird von den meisten unserer Interviewpartner:innen in Frage gestellt, ob die hohen Dokumentationsanforderungen letztlich zweckmäßig sind.

Während die meisten Dokumentationsanforderungen von Seiten der jeweiligen finanzierenden Behörden bzw. des Gesetzgebers gestellt werden, werden diese Fragen anlässlich der Einführung bzw. Nutzung von Fachsoftware regelmäßig innerhalb der gemeinnützigen Organisationen zwischen Führungskräften und Mitarbeitenden neu diskutiert und ausgehandelt.

#### C. Beziehungsdynamiken zwischen Führungskräften und Mitarbeitenden

Da Fachsoftware in die Rollendefinitionen und Rechenschaftspflichten von Führungskräften und Mitarbeitenden eingreift bzw. diese auf technischer Seite mit strukturiert, löst deren Einführung und Nutzung in der Regel Aushandlungsprozesse zwischen den beteiligten Statusgruppen und Stakeholdern nach sich. Die Führungskräfte haben insbesondere vier Befürchtungen beschreiben, die sie bei ihren Mitarbeitenden wahrnehmen: Angst vor (1) mehr Kontrolle, (2) Mittelkürzungen, (3) unsinniger Mehrarbeit und (4) technisch im Prozess abgehängt zu werden.

Als sensibelstes Thema wurde meist beschrieben, inwiefern durch Fachsoftware mehr Kontrolle der Arbeit durch Führungskräfte bzw. durch die Behörde ermöglicht wird. Die internen Diskussionen zwischen Geschäftsführungen und Mitarbeitenden bzw. ggf. dem Betriebsrat gehen dabei detailliert darauf ein, was, wie und wie oft dokumentiert wird, ob Führungskräfte die Arbeit ihrer Mitarbeitenden nun genauer kontrollieren können, und wer welche Rollen und Zugriffsrechte

in der Software formal zugewiesen bekommt. Viele Aspekte, die zuvor informell abgestimmt wurden, werden nun formalisiert. In einigen Organisationen wurden detaillierte Betriebsvereinbarungen erarbeitet, welche Funktionen der Software überhaupt genutzt werden dürfen. In einem Fall eskalierte der Streit bis vor das Arbeitsgericht.

Eine zweite Befürchtung besteht darin, dass Kontrollen bzw. Dokumentation als Argumentationsgrundlage für Mittelkürzungen dienen sollen. Diese Befürchtungen wurden uns berichtet von einigen Mitarbeitenden gegenüber ihren Geschäftsführungen, aber auch von den Geschäftsführungen gegenüber den finanzierenden Behörden, siehe auch [22].

*„(...) bei der Ambulantisierung ging es ja auch dann oft so darum, dass die Behörden dann sozusagen eine immer bessere, immer genauere Hilfeplanung [bzw. Dokumentation] machen wollten, eigentlich mit dem Ziel, dass die Hilfen noch bedarfsgerechter sind und vielleicht auch die Dauer der Hilfen begrenzt werden. Und dahinter ist aber immer spürbar, eigentlich geht es um Geld, ne?“ (Leitungskraft)*

Eine dritte Befürchtung richtet sich gegen unsinnige Mehrarbeit durch Fachsoftware, da sie immer technische Herausforderungen mit sich bringt und neue Anforderungen an den Prozess der Dokumentation vorgibt und z. T. zwingt, detaillierter und zeitnäher zu dokumentieren.

Als vierte Befürchtung wurde in unseren Gesprächen häufiger genannt, dass Mitarbeitende Sorge haben, bei der Digitalisierung nicht gut mit den neuen, z. T. ungewohnten Instrumenten zurecht zu kommen.

Diese vier Befürchtungen haben in einigen Organisationen zu Spannungen zwischen Führungskräften und Mitarbeitenden geführt und erfordern von Führungskräften in Organisationen der Sozialen Arbeit in diesen Fällen interne Beziehungsarbeit. Ein Geschäftsführer beschreibt seinen Ansatz so, dass sie „so lange schulen werden, bis keiner mehr Angst habe“. Die Einführung neuer Technologien erfordert nicht nur technisches Know-how, sondern auch ein hohes Maß an sozialer Kompetenz seitens der Führungskräfte. Auch wird deutlich, dass die Qualität der Beziehung zu den Mitarbeitenden entscheidend dafür ist, wie gut solche Prozesse verlaufen. Eine transparente Kommunikation und die Einbindung der Mitarbeitenden in Entscheidungsprozesse wurden als wesentliche Elemente genannt, um Vertrauen zu schaffen und Widerstände abzubauen.

Die potenziellen Befürchtungen zeigen auf, dass offene und frühzeitige Kommunikation entscheidend ist, um die Akzeptanz und Mitarbeit der Belegschaft zu sichern. Gleichzeitig verdeutlichen sie die finanziellen Spannungen, die oft im Hintergrund stehen und die Situation zusätzlich verkomplizieren. So könnte ein transparenter Umgang mit diesen finanziellen Realitäten helfen, das Vertrauen der Mitarbeitenden zu gewinnen. Es ist wichtig, dass Führungskräfte nicht nur die technischen, sondern auch die emotionalen und sozialen Aspekte solcher Veränderungsprozesse berücksichtigen und gleichzeitig die ökonomischen (um-)setzen und kommunizieren müssen, im Sinne einer *emotional labour*, siehe auch [23], [24]. Die Qualität der Beziehung zwischen Führungskräften und Mitarbeitenden scheint daher als wichtiger Faktor für das Gelingen von Veränderungsprozessen in sozialen Einrichtungen zu sein.

Eine Führungskraft bemerkt:

*„Also Veränderungen sind dann gut, wenn sie nicht aufgestülpt werden, wenn sie nicht irgendwie dirigiert werden, sondern wenn man die Begeisterung der Menschen dafür dann weckt. Ja, und das ist nicht immer einfach.“ (Leitungskraft)*

Ein weiterer Geschäftsführer betont:

*„Führungskräfte sind im Allgemeinen meistens eher veränderungsaffin und da muss man aufpassen, nicht zu schnell und zu viel auf einmal zu wollen. Da haben wir auch gelernt.“ (Geschäftsführer)*

Dies unterstreicht die Notwendigkeit, das Tempo von Veränderungen an die Bedürfnisse der Mitarbeitenden anzupassen.

Eine andere Führungskraft ergänzt:

*„Also das ist das Thema eben, dann alle mitnehmen und das ist inhaltlich, aber das ist auch vom Tempo her.“ (Leitungskraft)*

Diese Aussagen verdeutlichen, wie wichtig es ist, möglichst alle Mitarbeitenden bei Veränderungsprozessen einzubeziehen und dies bei Implementierungs- und Nutzungsprozessen stets zu integrieren. Die Einführung neuer Technologien erfordert somit nicht nur technisches Know-how, sondern auch ein hohes Maß an sozialer Kompetenz seitens der Führungskräfte. Ein Interviewpartner beschreibt die Herausforderung für Führungskräfte:

*„Also strategische Ausrichtung, konzeptionelle Vertiefung, Überblick behalten, so die wichtigen Punkte ansprechen und so weiter. Und dann so eine Adresse ist, wie so eine Art Generalbeschwerdestelle für alles, was nicht läuft.“ (Leitungskraft)*

Dies zeigt die Komplexität der Aufgaben, zwischen strategischen Zielen und der Rolle als Ansprechpartner für Probleme zu balancieren. Auch wird deutlich, dass die Qualität der Beziehung zu den Mitarbeitenden entscheidend dafür ist, wie gut solche Prozesse verlaufen. Eine transparente Kommunikation und die Einbindung der Mitarbeitenden in Entscheidungsprozesse sind wesentliche Elemente, um Vertrauen zu schaffen und Widerstände abzubauen. Führungskräfte sollten daher nicht nur auf die technischen, sondern auch auf die menschlichen Aspekte achten und entsprechende Unterstützung anbieten. Ein anderer Geschäftsführer hebt hervor:

*„Also bei uns würde ich sagen, Digitalisierung wird angenommen, wenn sie als Arbeitserleichterung erlebt wird. Was natürlich nicht immer spontan der Fall ist.“ (Geschäftsführer)*

Dies zeigt, wie wichtig es ist, den Nutzen von Veränderungen für die Mitarbeitenden deutlich zu machen. Zudem wird ein Wandel in der Haltung gegenüber Klienten beschrieben:

*„Dass man eher sozusagen sich auch als Dienstleistung versteht, auch bei Menschen, die sehr starke Einschränkungen haben. Wo es früher vielleicht auch immer nochmal hieß, irgendwie: 'Das sind doch keine Kunden.' Aber sind sie vielleicht eben doch, ne?“ (Geschäftsführer)*

Dieser Perspektivwechsel kann auch Auswirkungen auf die Beziehung zu Mitarbeitenden haben. Die Analyse der Zitate aus dem Fokusgruppeninterview zeigt, dass Identitätskonflikte, Spannungen bei der Einführung von Fachsoftware und

die Beziehungsqualität zu den Mitarbeitenden zentrale Herausforderungen im Management sozialer Einrichtungen darstellen. Ein bewusster, möglichst frühzeitiger Umgang mit diesen Aspekten und die Förderung einer offenen und unterstützenden Kommunikationskultur können helfen, diese Herausforderungen zu meistern und Veränderungen gelingender zu gestalten. Potentielle Konflikte zwischen fachlichen und betriebswirtschaftlichen Anforderungen können so leichter und früher sichtbar gemacht und dadurch besser bearbeitet werden.

### III. DISKUSSION

Der Hybridisierungsprozess, den Führungskräfte in kleinen gemeinnützigen Organisationen der Sozialen Arbeit im Zuge der Digitalisierung durchlaufen, stellt eine zentrale Herausforderung dar. Die Einführung von Fachsoftware erfordert nicht nur neue betriebswirtschaftliche Kompetenzen, sondern auch die Fähigkeit, diese mit der bestehenden fachlichen Expertise zu integrieren. Diese doppelte Verpflichtung führt oft zu Rollenkonflikten und Spannungen, die die berufliche Identität der Führungskräfte prägen. Die Analyse zeigt, dass die Digitalisierung als Katalysator für den Hybridisierungsprozess wirkt.

Diese Anforderungen können zu inneren Konflikten führen, da ethische und professionelle Standards der Sozialen Arbeit oftmals im Widerspruch zu betriebswirtschaftlichen Anforderungen stehen. Die Literatur und auch unsere Gesprächspartner:innen in den sozialen Organisationen betonen, dass unzureichende Kommunikation und mangelnde Unterstützung bei der Einführung von Fachsoftware die Beziehung zu den Mitarbeitenden belasten können, was zusätzliche Spannungen erzeugt [13], [21].

### IV. IMPLIKATIONEN

Um diesen Herausforderungen zu begegnen, können soziale Organisationen Räume für Aushandlungsprozesse der beruflichen Identitäten schaffen. Führungskräfte können von Unterstützungsangeboten profitieren, um die verschiedenen Identitätsanteile zu reflektieren. Organisationen können zudem transparente Kommunikationsstrategien entwickeln, um die Akzeptanz neuer Technologien zu fördern und die Beziehung zu den Mitarbeitenden zu stärken. Ein weiterer wichtiger Aspekt ist die Schaffung einer Organisationskultur, die sowohl die fachliche Expertise als auch die betriebswirtschaftlichen Anforderungen wertschätzt. Dies könnte beispielsweise durch die Förderung eines offenen Dialogs und die Einbindung der Mitarbeitenden in Entscheidungsprozesse erreicht werden.

### V. FAZIT

Der Hybridisierungsprozess stellt eine zentrale Herausforderung für Führungskräfte in der Sozialen Arbeit im Zeitalter der Digitalisierung dar. Die Digitalisierung verstärkt für Führungskräfte in der Sozialen Arbeit die Herausforderung, fachliche Expertise und Managementaufgaben zu vereinbaren. Dieser Prozess führt zu Spannungen zwischen den unterschiedlichen Rollen und kann die berufliche Identität erheblich beeinflussen. Organisationen können diesen Prozess aktiv gestalten, indem sie Räume für Aushandlungsprozesse schaffen und Führungskräfte bei der Integration der verschiedenen Identitätsanteile unterstützen.

Eine bewusste Gestaltung des Digitalisierungsprozesses, transparente Kommunikation und gezielte Unterstützungsmaßnahmen können Führungskräften helfen, ihre Rolle zwischen Fachlichkeit und Management auszubalancieren. Durch die Automatisierung administrativer Aufgaben können Fachkräfte mehr Zeit für ihre Kernaufgaben gewinnen, was die Facharbeit unterstützen könnte. Diese Perspektive zeigt auf, dass die Digitalisierung nicht nur Herausforderungen, sondern auch positive Aspekte für die berufliche Identität von Führungskräften in der Sozialen Arbeit bietet. Damit wird deutlich, dass Digitalisierung nicht nur eine Übertragung analoger Prozesse ins Digitale bedeutet, sondern immer auch eine mehr oder weniger starke Veränderung bzw. Transformation der Organisation mit sich bringt.

### DANKSAGUNG

Die Autoren bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

### LITERATUR

- [1] Noordegraaf M. Hybrid professionalism and beyond: (New) Forms of public professionalism in changing organizational and societal contexts. *Journal of Professions and Organization* 2015; 2: 187–206.
- [2] Shanks E, Lundström T, Wiklund S. Middle Managers in Social Work: Professional Identity and Management in a Marketised Welfare State. *The British Journal of Social Work* 2015; 45: 1871–1887.
- [3] Kurunmäki L. A hybrid profession—the acquisition of management accounting expertise by medical professionals. *Accounting, Organizations and Society* 2004; 29: 327–347.
- [4] Miller P, Power M. Accounting, Organizing, and Economizing: Connecting Accounting Research and Organization Theory. *The Academy of Management Annals* 2013; 7: 557–605.
- [5] Mennicken A, Miller P. Accounting, Territorialization and Power. *Foucault Studies* 2012; 0: 4–24.
- [6] Hood C. A Public Management for All Seasons? *Public Administration* 1991; 69: 3–19.
- [7] Hood C. The “new public management” in the 1980s: Variations on a theme. *Accounting, Organizations and Society* 1995; 20: 93–109.
- [8] Pflueger D. Quality improvement for all seasons: Administrative doctrines after New Public Management. *Financial Accountability & Management* 2020; 36: 90–107.
- [9] Junne B. *Accountability - Qualitätsmanagement in Hamburger Kindertagesstätten*. Wiesbaden: Springer, <http://link.springer.com/10.1007/978-3-658-13652-9> (2016, accessed 17 October 2016).
- [10] Lust M, Huber C, Junne J. Academic Identity as a Discursive Resource for Resistance: The Case of Quality Management in German Higher Education Institutions. *High Educ Policy* 2019; 32: 49–69.
- [11] Noordegraaf M, Steijn B (eds). *Professionals under Pressure: The Reconfiguration of Professional Work in Changing Public Services*. Amsterdam University Press. Epub ahead of print 31 December 2013. DOI: 10.1515/9789048518302.
- [12] Genth F, Jacobs CD. Serious play als edukativer Ansatz für die Weiterbildung von Fachpersonen. dtec-Publikationsband 2024.
- [13] Arnold M. Leading Digital Change and the Management of Hybridity in Social Work Organizations: In: Özsunur F (ed) *Advances in Public Policy and Administration*. IGI Global, pp. 55–73.
- [14] Gillingham P. Electronic Information Systems and Human Service Organizations: The Unanticipated Consequences of Organizational Change. *Human Service Organizations: Management, Leadership & Governance*, <https://www.tandfonline.com/doi/full/10.1080/23303131.2014.987412> (2015, accessed 17 July 2024).
- [15] Gillingham P. Technology Configuring the User: Implications for the Redesign of Electronic Information Systems in Social Work. *The British Journal of Social Work* 2016; 46: 323–338.

- [16] Gillingham P. Electronic Information Systems in Human Service Organisations: The What, Who, Why and How of Information. *The British Journal of Social Work* 2015; 45: 1598–1613.
- [17] Parton N. Changes in the Form of Knowledge in Social Work: From the ‘Social’ to the ‘Informational’? *The British Journal of Social Work* 2008; 38: 253–269.
- [18] Burton J, van den Broek D. Accountable and Countable: Information Management Systems and the Bureaucratization of Social Work. *The British Journal of Social Work* 2009; 39: 1326–1342.
- [19] Junne J. Enabling accountability: An analysis of personal budgets for disabled people. *Critical Perspectives on Accounting* 2018; 56: 46–62.
- [20] Roberts J. No one is perfect: The limits of transparency and an ethic for ‘intelligent’ accountability. *Accounting, Organizations and Society* 2009; 34: 957–970.
- [21] Castillo de Mesa J. Digital social work: Towards digital disruption in social work. *The Journal of Sociology & Social Welfare* 2021; 48: 117–138.
- [22] Linke V, Einhaus M, Hilgert C, et al. Prekäres Organisieren im subsidiären Wohlfahrtsstaat. Pandemisches Krisenmanagement als Spannungsverstärker oder -entschärfer für Einrichtungen der sozialen Daseinsvorsorge. In: Bode I, Jungmann R, Serrano Velarde K (eds) *Staatliche Organisationen und Krisenmanagement*. Nomos Verlagsgesellschaft mbH & Co. KG, pp. 110–134.
- [23] Hochschild AR. Emotion Work, Feeling Rules, and Social Structure. *American Journal of Sociology* 1979; 85: 551–575.
- [24] Hochschild AR. *The managed heart: commercialization of human feeling*. University of California Press, 1983.

# Serious Play als edukativer Ansatz für die Weiterbildung von Fachpersonen

Felix Genth

Institut für Controlling und Unternehmensrechnung  
Helmut Schmidt Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
felix.genth@hsu-hh.de

Prof. Claus D. Jacobs, Ph.D.

Kompetenzzentrum für Public Management  
Universität Bern  
Bern, Schweiz  
claus.jacobs@unibe.ch

**Kurzfassung** – Die vorliegende Arbeit analysiert, welchen Beitrag Serious Play zur Weiterbildung von Fachpersonen leisten kann. In Wohlfahrtsorganisationen sehen sich Mitarbeiter in Führungsverantwortung mit einer zunehmenden Hybridisierung ihrer Rolle konfrontiert. Die Hybridität ist dadurch gekennzeichnet, dass sie zwischen fachlichen und finanziellen Anforderungen navigieren müssen. In den von uns untersuchten Organisationen haben diese Hybrid Professionals häufig einen sozialarbeiterischen Hintergrund, müssen aber im Rahmen ihrer Führungsverantwortung auch die betriebswirtschaftliche Seite ihrer Organisation in den Blick nehmen. Dabei sollen sie im Rahmen des Projektes durch die Einführung neuer digitaler Tools unterstützt werden. Diese führen auch zu einer Erleichterung in der Erhebung und dem Umgang mit betriebswirtschaftlichen Kennzahlen, bieten aber für Personen, die mit vielfältigen komplexen Anforderungen umgehen müssen, keinen einfachen Einstieg in das Thema. Um diesen Personenkreis den Zugang zur BWL zu erleichtern, wurde im Rahmen des Forschungsprojektes DiWop von unserem Projektpartner Cord Solle von s-hvc eine BWL-Simulation angeboten und durchgeführt. Diese wurde wissenschaftlich in Form einer teilnehmenden Beobachtung begleitet. Die Auswertung der zwei Workshoptage hat ergeben, dass, wie in der Theorieliteratur zu Serious Play beschrieben, das spielerische Format dazu beitragen konnte, dass kognitive, soziale, materielle und emotionale Aspekte eines durch externe Anforderungen geleiteten Handelns verbunden werden konnten. Durch die spielerische Herangehensweise konnte die intrinsische Motivation der Teilnehmer gestärkt werden, sich mit BWL-Themen auseinanderzusetzen. Besonders das Gemeinschaftsgefühl durch Teamarbeit und das emotionale Hineinversetzen in einen Betrieb, durch das Durchspielen von Geschäftsjahren, hat die BWL für die Teilnehmer erfahrbar gemacht und somit den Zugang erleichtert.

**Stichworte** – *Serious Play, Hybridisierung, soziale Arbeit*

## NOMENKLATUR

DiWop	Projekt zur Förderung der Digitalisierung in der freien Wohlfahrtspflege
s-hvc	Schödelbauer – high value coaching
NPOs	Non-Profit-Organisationen

## I. HYBRIDE WOHLFARHTSORGANISATIONEN

Im Zuge des New Public Managements kommt es zu einer zunehmenden Hybridisierung von sozialen Organisationen.

Auch wenn diese Organisationen im Kern nicht gewinnorientiert arbeiten, wird ein wirtschaftliches Handeln vorausgesetzt. Die hauptsächlich vom Staat zur Verfügung gestellten Gelder sollen möglichst effektiv zur Erreichung der gesteckten Ziele eingesetzt werden. In Wohlfahrtsorganisationen wird diesbezüglich von einer Sozialrendite (Social Return on Investment) gesprochen. Auch die in unserem Projekt zur Förderung der Digitalisierung in der freien Wohlfahrtspflege (DiWop) begleiteten Organisationen sehen sich einem zunehmenden Druck ausgesetzt, wirtschaftlich und effizient zu handeln. Besonders Leitungskräfte und Geschäftsführer finden sich dadurch in einer zunehmend hybriden Rolle wieder. Sie müssen sozialarbeiterische Expertise und wirtschaftliches Denken miteinander vereinen. (Giacomelli, 2020) Sozialarbeiter haben dabei häufig eine idealistische Sicht auf ihre Arbeit, die sie durch wirtschaftliches Denken gefährdet sehen. (Jacobs et al., 2021) Die Literatur legt jedoch nahe, dass wirtschaftliches Denken einerseits die Leistungsfähigkeit von NPOs erhöhen (Mahmoud, 2012) und andererseits die Legitimität gegenüber der Umwelt verbessern kann. (Suykens et al., 2019) Das Vereinen dieser Perspektiven ist also lohnend, erfordert jedoch auch eine besondere Kompetenz und Fingerspitzengefühl bei potenziellen Konflikten zwischen wirtschaftlichen und fachlichen Anforderungen und Zielen. Unter anderem auch deshalb, weil in Wohlfahrtsorganisationen häufig Personen mit den Finanzen betraut sind, die keinen betriebswirtschaftlichen Hintergrund haben. Um diesen Personen den Zugang zu betriebswirtschaftlichen Grundlagen zu erleichtern, hat unser Projektpartner einen Workshop mit dem Titel: „Betriebswirtschaftliches Know-how in gemeinnützigen Organisationen“ angeboten, der aufgrund hoher Nachfrage zwei Mal durchgeführt wurde. Im Folgenden wird zunächst der theoretische Hintergrund zu Serious Play erläutert, um dann auf die Durchführung des Workshops einzugehen und wie sich die Vermittlung als Spiel auf die Wahrnehmung der Teilnehmer ausgewirkt hat.

## II. SERIOUS PLAY

Seit jeher und aus vielen unterschiedlichen Perspektiven wird Spiel, Spielen und spielerischem Handeln eine wichtige Rolle zugewiesen (Huizinga, 1955; Piaget 2001; Winnicott, 2001). Gleichwohl gibt es keine einheitliche, etablierte Definition für das Spielen als soziale Aktivität – erst recht nicht in Organisationen (Mainmelis & Ronson, 2006). Serious Play – das ernsthafte Spiel – definieren Statler, Heracleous und Jacobs (2011: 236) als eine soziale Praxis, die sich „bewusst auf

eine freudige, intrinsisch motivierende Aktivität als Mittel zur Erreichung eines ernsthaften, extrinsisch motivierten Arbeitszieles richtet“. Spielen als soziale Praxis anerkennt und verbindet diesbezüglich die kognitiven, sozialen, materiellen und emotionalen Aspekte dieses Handelns (Schatzki, 2001). Plato (1991) unterscheidet das frivol-leichtfertige Spiel vom ernsthaften Spiel, dass eine spielerische Form der Interaktion mit einer ernsthaften, instrumentellen Intention des Handelns verbindet. Diese zentrale Unterscheidung erlaubt auch, die tradierte Unterscheidung von Spiel und Arbeit zu überschreiten. Die Forschung zu Lernprozessen von Erwachsenen unterstreicht die Bedeutung von Emotionalität und Materialität für Lernerfolge (Rieber et al., 1998). Vor diesem Hintergrund ist es für eine Intervention der Aus- und Weiterbildung von Fachpersonen naheliegend, einen agogischen Ansatz zu wählen, der diesen Erkenntnissen Rechnung trägt und die spielerische Dimension (Emotionalität, Materialität) mit der ernsthaften Dimension (Intentionalität) verbindet: Serious Play, das ernsthafte Spiel.

### III. DIE BWL-SIMULATION

Im Folgenden wird der Ablauf des zweiten im Rahmen von DiWop durchgeführten Workshops zum Thema BWL in NPOs skizziert. Der Workshop fand 2024 an der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg mit 12 Teilnehmern statt. Die Teilnehmer setzten sich aus Geschäftsführern, Vorständen und Leitungskräften von kleineren Wohlfahrtsorganisationen der Eingliederungs- sowie der Kinder- und Jugendhilfe aus dem Hamburger Raum zusammen. Durchgeführt wurde der Workshop von Cord Solle, welcher langjährige Erfahrung in der Führung und Beratung von NPOs aufweist. Der Workshop basiert auf einer Organisationssimulation, die von Frank Schwedhelm von der nosolo Unternehmensberatung in Zusammenarbeit mit Cord Solle speziell für Non-Profit-Organisationen entwickelt wurde.

Zu Beginn des Workshops sollten die Teilnehmer ihr betriebswirtschaftliches Wissen auf einer Skala von 1 bis 10 einordnen. Fast alle Teilnehmer schätzten ihre Fähigkeiten sehr bescheiden zwischen 1 und 3 ein. Als Motivation an dem Workshop teilzunehmen, gaben die Teilnehmer unter anderem an, dass sie zusätzliche Leitungsverantwortung übernehmen werden und sich deshalb auch mit der Finanzierung ihrer Organisation auseinandersetzen wollen oder müssen. Andere Teilnehmer wollten lediglich ihr bereits vorhandenes Wissen auffrischen. Nach einer kurzen Vorstellungsrunde gab der Workshopleiter am Vormittag des ersten Tages eine kurze Einführung in betriebswirtschaftliches Grundwissen, dass er jeweils mit alltäglichen Beispielen illustrierte. Unter anderem erläuterte er, dass verschiedene Marktteilnehmer jeweils wirtschaftliche Interessen verfolgen, die einander z. T. widersprechen und verhandelt werden müssen. Des Weiteren wurde darauf eingegangen, dass der Markt über den Erfolg von Unternehmen entscheidet. Dieser unterliegt jedoch Schwankungen, weshalb Unternehmen zum Umgang mit finanziellen Risiken Rücklagen bilden müssen – auch NPOs. Im Anschluss wurden strukturelle Unterschiede von Non-Profit-Organisationen und For-Profit-Organisationen (FPOs) erläutert (siehe Abbildung 1).

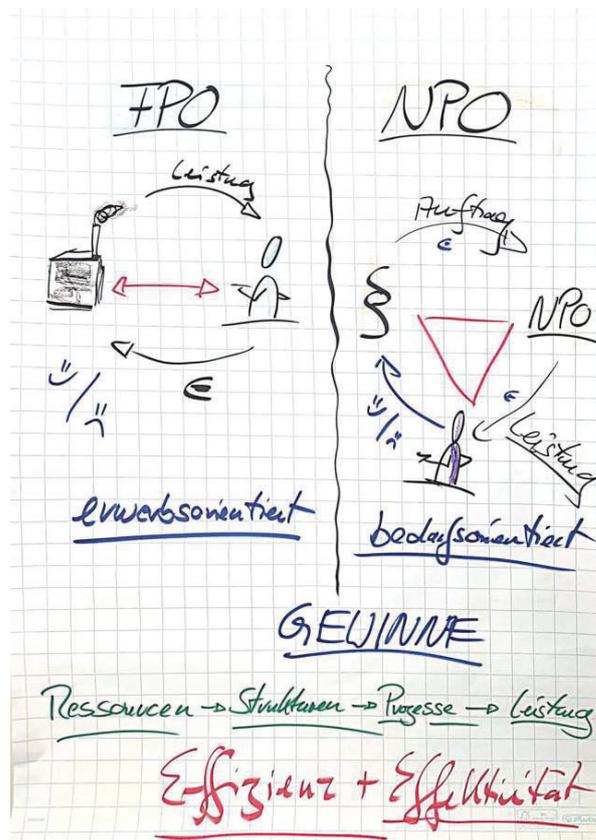


ABBILDUNG 1: UNTERSCHIEDUNG FPO UND NPO (CORD SOLLE)

Es wurde veranschaulicht, dass sich Non-Profit-Organisationen in einem sozialrechtlichen Dreiecksverhältnis befinden. Sie erhalten einen gesetzlichen Auftrag im Rahmen des Wohlfahrtsstaates, erbringen ihre Leistung dann an ihren Klienten und werden wiederum von dem Gesetzgeber finanziert. Demgegenüber gibt es bei FPOs eine direkte Handelsbeziehung zwischen dem Unternehmen und ihren Kunden. Das Unternehmen bietet eine Leistung an und der Kunde zahlt für diese Leistung (dies ist in NPOs nur selten der Fall, z. B. bei Klienten mit persönlichen Budgets). Des Weiteren seien FPOs in erster Linie erwerbsorientiert und NPOs bedarfsorientiert. Dass die Finanzen dennoch auch in NPOs eine entscheidende Rolle spielen, sollte im weiteren Verlauf des Workshops aufgezeigt werden. Da die Teilnehmer nun schon sehr gespannt auf die Simulation warteten, erläuterte der Workshopleiter im Anschluss die Grundlagen des Rechnungswesens, die zum Spielen benötigt werden. Er beschrieb den Aufbau einer Bilanz, Bedeutung und Aufbau von Aktiva und Passiva sowie der Gewinn- und Verlustrechnung (GuV) und der Liquiditätsrechnung, die im Anschluss für jedes Geschäftsjahr im Spiel aufgestellt werden sollten.

Um zu spielen wurden die Teilnehmer in Gruppen von 3-4 Personen aufgeteilt, die jeweils an einem Tisch um das Spielfeld (siehe Abbildung 2) Platz nahmen.

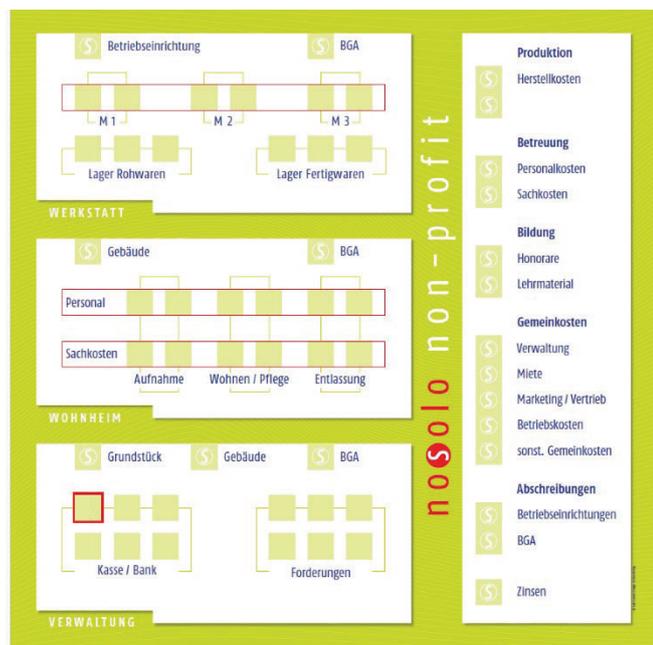


ABBILDUNG 2: NOSOLO NON-PROFIT SPIELFELD

Alle teilnehmenden Gruppen wurden mit der gleichen Startbilanz in Form von Münzen ausgestattet. Diese befanden sich in Plastikbehältern, die jeweils maximal 20 Münzen fassen und auf dem Spielfeld verschoben werden konnten. Das Spiel lässt sich über 10 Runden spielen, die jeweils für ein Geschäftsjahr stehen. Die simulierten Geschäftsjahre stehen jeweils unter einem eigenen Lernthema (z. B. Kosten versus Investition) und bilden praktische Geschäftsvorfälle einer NPO ab. Die Teilnehmer spielen die Geschäftsjahre nach, indem sie das Spielgeld entsprechend auf dem Feld verschieben bzw. „umbuchen“. Gibt es z. B. Ausgaben für Personal von drei, müssen drei Münzen von der Kasse zu Personalkosten verschoben werden. Auf diese Weise können die Spielenden haptisch und visuell nachvollziehen, wie das Geld zwischen den verschiedenen Positionen wandert. Am Ende eines jeden Geschäftsjahres können die Spieler durch Zählen der Münzen auf den jeweiligen Positionen eine Gewinn- und Verlustrechnung erstellen. Im Anschluss wird dann die Bilanz ermittelt, die in den späteren Runden durch die Liquiditätsrechnung ergänzt wird, um eine Insolvenz zu vermeiden. Den Spielern wird so veranschaulicht, welche unterschiedlichen Informationen die jeweiligen Rechnungen vermitteln. Sie lernen zum Beispiel, wie wichtig es ist, den Eigenkapitalanteil in den Passiva im Blick zu behalten. Im Spielverlauf weist jedes gespielte Geschäftsjahr spezifische Besonderheiten auf. So erläuterte der Spielleiter zum ersten Geschäftsjahr, dass der Betrieb aufgrund eines Vertrages mit dem Kostenträger dazu verpflichtet ist, immer eine Sockelliquidität von 10 aufrechtzuerhalten. Ansonsten verliert er seine Betriebserlaubnis. Die Sockelliquidität darf nur nach Beschluss des Vorstandes kurzfristig unterschritten werden. In einem späteren Geschäftsjahr, bei dem die Spieler für Investitionen so viel Geld aus der Kasse nehmen müssen, dass die Sockelliquidität kurzfristig unterschritten wird, tritt dieser Fall ein. Der Spielleiter fragte im Anschluss des Geschäftsjahres dann nach, ob die Spieler etwas Besonderes wahrgenommen haben. Dadurch, dass die Spieler selbst das Geld zwischen den verschiedenen Positionen verschieben und wahrnehmen können, wie die Kasse immer leerer wird, werden sie für diese Prozesse sensibilisiert.

Ein wichtiger Aspekt ist die Aufteilung des Spielfeldes (vgl. Abbildung 2). Auf der rechten Seite befinden sich die Ausgaben, die nach jedem Geschäftsjahr vom Spielleiter entfernt werden. Auf der linken Seite befindet sich das Vermögen und Kapital der Organisation. So wird in einem Geschäftsjahr z. B. angesagt, dass die Organisation in Gebäude investiert. Der Spieler kann dann nachvollziehen, dass dieses Geld sich weiterhin auf der linken Seite des Spielfeldes, also in den Aktiva, befindet und lediglich von der Kasse in die Position der Gebäude wandert. Dies veranschaulicht, dass in Gebäude investiertes Geld dem Unternehmen weiterhin als Vermögen zur Verfügung steht. Anders verhält es sich mit den Gehältern der Mitarbeiter. Diese werden von der Kasse auf der linken Seite auf die rechte Seite der Ausgaben verschoben und am Ende jedes Geschäftsjahres vom Spielleiter entfernt. Der Spielleiter führte hierzu erklärend aus, dass aus diesem Grund die Geschäftsführer lieber in Gebäude investierten als in Mitarbeiter. Das in Gebäude investierte Geld stünde der Organisation weiterhin als Vermögen zur Verfügung, während das in Mitarbeiter investierte Geld anschließend weg sei.

#### IV. SPIEL UND ERNST

##### A. Emotionalität

Bei dieser Abwägung, zwischen der Investition in Mitarbeiter oder in Gebäude, regte sich Widerstand unter den Teilnehmern. Die Mitarbeiter würden in gemeinnützigen Organisationen eine zentrale Rolle einnehmen und in sie investiertes Geld hätte auch eine positive Auswirkung auf die Klienten und die Leistung des Unternehmens. Das inhaltliche Argument des Spielleiters konnte dennoch nachvollzogen werden. Das Spannungsfeld zwischen ideellen Werten und finanziellen Vorteilen, in dem sich die Teilnehmer in ihrer hybriden Rolle bewegen müssen, konnte so anschaulich thematisiert werden. Dass die Reaktion auf das Beispiel recht emotional war, zeigt auch, dass das Durchspielen der Geschäftsabläufe einer konkreten Beispielorganisation, die häufig als trocken stigmatisierten betriebswirtschaftlichen Vorgänge, erfahrbar und nachfühlbar machen konnte. Diese emotionale Verknüpfung sorgt dafür, dass das Gelernte besser behalten werden kann. (vgl. Rieber et al., 1998)

##### B. Materialität

Die Simulation fand, ähnlich wie bei Monopoly, mit Hilfe von Spielgeld auf einem räumlich aufgeteilten Spielfeld statt, auf dem dieses Geld bewegt wird. Durch die Haptik und die Visualisierung auf dem Spielfeld lässt sich der Fluss des Geldes sehr gut nachvollziehen. (vgl. Rieber et al., 1998) Um die Spannung etwas zu erhöhen hat der Spielleiter zu Beginn erläutert, dass eine Spielmünze eine Million Euro symbolisiert. Wenn dann im Verlauf des Spieles eine Münze verloren geht und dadurch eine Million in der Bilanz fehlt, trug dies häufiger zur allgemeinen Belustigung bei. Der Verlust kann aber auch ernstgenommen werden. Das Spiel ist zwar nicht auf Konkurrenz ausgelegt, es wird jedoch trotzdem wahrgenommen, wenn an einem Tisch häufiger Geld verloren geht oder Rechnungen nicht stimmen als bei anderen. Wenn dann Mitglieder derselben Organisation an verschiedenen Tischen sitzen, wird ein Stück weit um die eigene Reputation gespielt. Fehlen an einem Tisch regelmäßig Millionen, kann es zur halb scherzhaft, halb ernst gemeinten Nachfrage kommen, ob diese Person wirklich mit den Finanzen betraut sein sollte.

### C. Sozialität

Im Gegensatz zu vielen Gesellschaftsspielen sind die um ein Spielfeld platzierten Mitglieder einer Gruppe keine Konkurrenten, sondern haben das gemeinsame Ziel, eine fehlerfreie Bilanz, GuV und Liquiditätsrechnung aufzustellen. Es entsteht eine eigene Gruppendynamik und individuelle Zuständigkeiten. Das gemeinsame Spielen erhöht die Motivation keine Fehler zu machen, ein korrektes Ergebnis zu ermitteln und so die Organisation wirtschaftlich zu führen. Diese Motivation war bei allen Teilnehmern spürbar. (vgl. Schatzki, 2001)

### D. Temporalität

Ergänzend zu den in der Literatur beschriebenen Eigenschaften des Spiels, der Emotionalität, Materialität und Sozialität (vgl. Rieber et al., 1998; Schatzki, 2001), spielt die Temporalität bei der BWL-Simulation eine entscheidende Rolle. Das Durchspielen von bis zu 10 Geschäftsjahren in zwei Workshoptagen lässt die Teilnehmer die Langzeitfolgen von wirtschaftlichen Entscheidungen nachvollziehen. So kann z. B. beobachtet werden, wie sich die Einführung eines neuen Geschäftsbereiches über mehrere Jahre auf die Wirtschaftlichkeit der Organisation auswirkt. Das Spielen ermöglicht es somit den Teilnehmern, die eigene Perspektive zu erweitern und ein auf langfristigen wirtschaftlichen Erfolg ausgerichteter strategisches Denken zu entwickeln.

### V. FAZIT

Die BWL-Simulation bietet einen besonderen Rahmen für die Vermittlung von betriebswirtschaftlichem Wissen an gemeinnützige Organisationen. Diese Organisationen sind professionell vor allem durch Sozialarbeiter und Sozialpädagogen geprägt, die in ihrer alltäglichen Praxis vielfältige Methoden anwenden, um einen Perspektivwechsel ihrer Klienten anzuregen. (vgl. Schwing und Fryszer, 2018) Sie sind also mit der spielerischen Erschließung neuer Perspektiven und Themen vertraut. Auch inhaltlich ist die BWL-Simulation auf diese Organisationen ausgerichtet. Die in den Geschäftsjahren durchgespielten Beispiele haben alle einen Bezug zur freien Wohlfahrtspflege. Der spielerische Ansatz nutzt dabei die besonderen Eigenschaften des Serious Play, um betriebswirtschaftliche Abläufe erlebbar zu machen. Dies verstärkt die Lernerfahrung und spiegelt sich auch in dem durchweg positiven Feedback der Teilnehmer wider, die berichteten, dass ihnen die Verbindung von Theorie und Praxis in der Simulation den Zugang zur BWL sehr erleichtert hat. Eine Teilnehmerin fasste ihre Erfahrung prägnant wie folgt zusammen: „Durch das Spielen hatte ich das Gefühl, dass es mein Gehirn richtig erreicht hat. Das ist bei den drei Fortbildungen, die ich vorher zu dem Thema gemacht habe, nicht so gewesen“ (Bereichsleitung, die demnächst die Geschäftsführung einer NPO übernimmt).

### DANKSAGUNG

Diese Forschungsarbeit im Forschungsprojekt DiWop – Förderung der Digitalisierung in der freien Wohlfahrtspflege wird durch dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

### LITERATUR

- [1] Abdulai Mahmoud, M., & Yusif, B. (2012), Market orientation, learning orientation, and the performance of nonprofit organisations (NPOs). *International Journal of Productivity and Performance Management*, 61(6), 624-652.
- [2] Giacomelli, G. (2020), The role of hybrid professionals in the public sector: a review and research synthesis, *Public Management Review*, 22:11, 1624-1651
- [3] Huizinga, J. (1955), *Homo Ludens: A Study of the Play Element in Culture*. Boston, MA: The Beacon Press.
- [4] Jacobs, C. D., Kreutzer, K., & Vaara, E. (2021), Political dynamics in organizational identity breach and reconstruction: Findings from the crisis in UNICEF Germany. *Academy of Management Journal*, 64(3), 948-980.
- [5] Mainemelis, C. and Ronson, S. (2006), 'Ideas are born in fields of play: towards a theory of play and creativity in organizational settings', *Research in Organizational Behavior*, 27: 81-131.
- [6] Piaget, J. (2001), *The Language and Thought of the Child*. London: Routledge.
- [7] Plato (1991), *The Republic of Plato*. Translated by A. D. Bloom. 2nd edn. New York: Basic Books.
- [8] Rieber, L. P., Smith, L., & Noah, D. (1998), The value of serious play. *Educational technology*, 38(6), 29-37.
- [9] Schatzki, T. (2001), 'Practice theory', in Schatzki, T. R., Knorr-Cetina, K., and von Savigny, E. (eds), *The Practice Turn in Contemporary Theory*. 1-14. London: Routledge.
- [10] Schwing, R., & Fryszer, A. (2018), *Systemisches Handwerk: Werkzeug für die Praxis*. Vandenhoeck & Ruprecht.
- [11] Statler, M., Heracleous, L. and Jacobs, C. D. (2011), 'Serious play as a practice of paradox', *Journal of Applied Behavioral Science*, 47/2: 236-56.
- [12] Suykens, B., De Rynck, F., & Verschuere, B. (2019), Nonprofit organizations in between the nonprofit and market spheres: Shifting goals, governance and management?. *Nonprofit Management and Leadership*, 29(4), 623-636.
- [13] Winnicott, D. W. (2001), *Playing and Reality*. London: Routledge Press.

# Zu den Funktionen und Folgen von Informationsmängeln in der Organisationsberatung

Lukas Lahme

*Institut für Controlling und Unternehmensrechnung  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
lukas.lahme@hsu-hh.de*

Christian Hilgert

*Institut für Soziologie  
Fernuniversität in Hagen  
Hagen, Deutschland  
christian.hilgert@fernuni-hagen.de*

**Kurzfassung** – In dem dtec.bw-Projekt *Digitalisierung in der Wohlfahrtspflege* wurden 13 Organisationen der sozialen Hilfe aufgenommen, um sie bei der Digitalisierung ihrer Organisation zu unterstützen. Dazu wurden vier Organisationsberatungen mit Erfahrungen im Bereich der sozialen Hilfe und Digitalisierungsprozessen beauftragt. Das ermöglichte die Untersuchung von Beratungsprozessen in verschiedenen Digitalisierungsprojekten. Hierbei fiel auf, dass einige Berater die Informationen, mit denen sie arbeiten, primär aus Interaktionen mit den Geschäftsführern sowie aus organisationsöffentlichen Workshops gewinnen. Aber wie für jede Organisation, die auf Arbeitsteilung besteht und sich vertikal ausdifferenziert, gilt: Informationen, die nach oben weitergegeben werden, sind stark gefiltert. In der Folge bekommen die Berater vor allem die formalen Strukturen der Organisation in den Blick – nicht aber die Informalität. Das kann sich, zumindest in einigen Fällen, negativ auf den Erfolg eines Beratungsprojekts auswirken. Gleichzeitig würde eine stärker auf die Informalität bezogene Beratung die Arbeit der Berater in verschiedenster Hinsicht erschweren. Erstens wird das Beratungsprojekt erheblich teurer, weil mehr Zeit in die Informationserhebung investiert werden müsste. Diese Kosten müsste der Finanzier, d. h. der Kunde, bereit sein, zu zahlen. Zweitens droht Beratern das Risiko, Überbringer schlechter Nachrichten zu werden, wenn aufgrund der höheren Informiertheit klar wird, dass bestimmte Vorhaben des Auftraggebers nicht sinnvoll umsetzbar sind. Drittens können Berater in Rollenkonflikte zwischen ihren Informanten und Auftraggebern geraten: Sie sind über informale Arbeitsweisen informiert, können sie zwecks des Friedens und der weiteren Zusammenarbeit mit ihren Informanten aber nicht an den Auftraggeber weiterleiten.

**Stichworte** – Beratung, Informationsasymmetrien, funktionale Analyse, Systemtheorie

## I. EINLEITUNG

Im Projekt „Förderung der Digitalisierung in der freien Wohlfahrtspflege“ geht es darum, ausgewählten Organisationen der sozialen Hilfe bei ihren Digitalisierungsvorhaben zu unterstützen und die organisationalen Schwierigkeiten von Digitalisierung in Wohlfahrtsorganisationen zu erforschen. Bei den Organisationen handelt es sich um Organisationen kleiner bis mittlerer Größe. Die Anzahl der Mitglieder variiert je nach Organisation zwischen 20 bis 300. Um die Unterstützung leisten zu können, wurden im Rahmen des Projekts einschlägige Organisationsberatungen beauftragt. Den Forschern ermöglicht das Forschungsprojekt daher nicht nur umfangreiche Einblicke in Organisationen der Sozialen Hilfe und die Heraus-

forderungen ihrer Digitalisierung, sondern auch in die entsprechende Beratungspraxis. So besteht im Projekt die Möglichkeit, an verschiedensten Beratungsinteraktionen (Abstimmungstermine, Ergebnispräsentationen, Workshops usw.) teilzunehmen und Beobachtungen anzustellen.

Eine dieser Beobachtungen war, dass einige Berater dazu tendieren, Informationen über die zu verändernde Organisation primär aus Gesprächen mit der Führungsriege – vor allem den Geschäftsführern – und aus organisationsöffentlich stattfindenden Workshops zu gewinnen. Aus unseren Beobachtungen lässt sich nicht ableiten, dass dies für *alle* Berater gilt und dass Berater *nur* so vorgehen. Einige Berater tendieren stärker als andere dazu, vereinzelt oder methodisch das vertrauliche Gespräch mit operativ arbeitenden Sozialarbeitern zu suchen. Und einige Berater variieren ihre Informanten je nach Arbeitsauftrag. Unsere nachfolgenden Überlegungen fokussieren auf Beratungsprozesse, deren Informationsbeschaffung dem obigen Muster folgt. Wir wollen uns die Frage stellen, inwiefern dadurch ein einseitiges Bild der Organisation konstruiert wird und ob und inwiefern sich dies auf die Ergebnisse des Beratungsprozesses auswirken kann. Um diese Frage zu bearbeiten, werden wir im Folgenden einige theoriegeleitete, organisationssoziologische Überlegungen zu ihrer spezifischen Selektivität dieser beider Formen der Informationsbeschaffungen machen. Abschließend wollen wir dann noch danach fragen, welche Funktionen diese Formen der Informationsbeschaffung – bzw. Nichtbeschaffung – erfüllen.

## II. INFORMATIONSMÄNGEL IM BERATUNGSPROZESS

Im Wesentlichen, so unsere These, erhalten Berater, die vorzüglich mit Leitungsebenen sprechen oder Workshops veranstalten, primär Informationen über die formale Seite der Organisation, während der Blick auf die informale Seite zwangsläufig verstellt bleibt. Formale Erwartungen sind solche, die durch die „Mitgliedschaftsregel gedeckt [sind], d. h. wenn erkennbar Konsens darüber besteht, daß die Nichtanerkennung oder Nichterfüllung dieser Erwartung mit der Fortsetzung der Mitgliedschaft unvereinbar ist.“ (Luhmann 1972, S. 38). Wir wollen zunächst ausführen, wodurch dieses verstellte Bild entsteht.

Wir wissen aus der systemtheoretischen Forschung, dass in arbeitsteiligen Organisationen Informationsasymmetrien entstehen: „Wenn die Informationen, die ein System zu verarbeiten hat, einen gewissen Umfang überschreiten, wird es un-

rationell, eine gleichmäßige Informiertheit aller Mitglieder anzustreben“ (Luhmann 1972, S. 179). Insbesondere gilt dies – wengleich nicht nur – in vertikaler Hinsicht, also hinsichtlich der Hierarchie der Organisation. Allein schon aus Gründen der begrenzten Informationsbearbeitungskapazitäten von Personen (Simon 1981, S. 116 ff.) lassen sich nicht alle Informationen der Basis an die Spitze weiterleiten. Würden sämtliche Organisationsmitglieder alle Informationen an die Geschäftsführung weiterleiten, wäre diese maßlos überfordert. Damit andere die Informationen zur Grundlage ihres Handelns machen können, müssen die ursprünglichen Informationsinhaber sie nach bestimmten Relevanzkriterien filtern, und sie in ihrer Komplexität und Widersprüchlichkeit reduzieren (Luhmann 1972, S. 174 f.). In der Organisation existieren dann, je nach Anzahl der Hierarchieebenen, unterschiedlich viele Filter. In unseren Fällen sind das in der Regel zwei Hierarchiestufen. Die Sozialarbeiter berichten also nicht direkt an den Geschäftsführer, sondern an ihre Teamleiter bzw. Fachbereichsleiter und diese wiederum filtern weiter Informationen, die sie dann an den Geschäftsführer weitergeben. Je größer die Organisation wird, desto eher werden Ausdifferenzierung der Organisation, konkret der Kontaktstrukturen, relevant. So berichtete ein Geschäftsführer, dass er vor einiger Zeit noch versuchte, an allen Teambesprechungen teilzunehmen – um so das Problem der Informationsasymmetrien abzuschwächen. Mit der wachsenden Zahl der Organisationsmitglieder hat sich das allerdings schlicht als nicht mehr opportun angeboten und wurde daher aufgegeben. Die Informationen, die oben ankommen, unterscheiden sich also mehr oder weniger stark von dem tatsächlichen Arbeitsalltag der Sozialarbeiter ab.

Relevant ist hier die Frage, nach welchen Kriterien die Informationen in der Hierarchie gefiltert werden. Das hängt von vielerlei ab. Etwa konnten wir beobachten, das oftmals von den Sozialarbeitern formal erwartet wird, ihre Klientenarbeit zu dokumentieren und einsichtig zu machen. Oder es gibt Super- bzw. Intervisionen, in denen Informationen zusammengetragen werden. Unabhängig von der Art und Weise, wie die Informationsweitergabe organisiert ist, lässt sich annehmen, dass Sozialarbeiter Hemmungen haben, informale Abweichungen oder gar brauchbare Illegalitäten (Luhmann 1972, S. 304 ff.; ausführlich Kühl 2020) an ihre Vorgesetzten mitzuteilen, um sich nicht selbst zu exponieren oder gar zu belasten. Damit sind Handlungen gemeint, die einerseits klar gegen die formalen Vorgaben der Organisation verstoßen, andererseits aber einen funktionalen Beitrag zur Leistungsfähigkeit des Betriebs erbringen. Wenn sich Berater also primär bei Führungskräften über die Arbeitsprozesse informieren, ist zu vermuten, dass sie sich ein unvollständiges Bild über die tatsächliche Praxis bilden, da ihnen die Dimension informaler Arbeitsweisen weitgehend verborgen bleibt.

Neben den Gesprächen mit der Geschäftsführung führen die Berater im Projekt aber auch Workshops mit operativ arbeitenden Sozialarbeitern durch, die im Beratungsprozess der Problembestimmung und Prozessaufnahme dienen. Mindestens hier, so ließe sich argumentieren, können die Berater aufgrund des unmittelbaren Austauschs mit den Sozialarbeitern, die den Arbeitsalltag an der Basis der Organisation kennen, Informationen über die informale Seite der Organisation sammeln. Dagegen lässt sich aber einwenden, dass diese Workshopformate keineswegs ‚sichere‘ Situationen sind, in denen die Teilnehmer bereit wären, sich ernsthaft zu exponieren. Denn die Öffentlichkeit dient als Zensurmechanismus, sodass die Anwesenden dazu tendieren, primär ungefährliche Informationen zu teilen, wie etwa wenig persönlich zurechenbare

Wertebekundungen (Kieserling 2002, S. 1 f.). Und zwar nicht allein aus dem Grund, weil Führungskräfte anwesend sind, sondern auch, weil ab einer bestimmten Gruppengröße die Frage relevant wird, ob man wirklich auf die Diskretion der Anwesenden vertrauen kann (Kieserling 2002, S. 1). Es gehört einiges an Mut – oder Unbedachtheit – dazu, vor versammelter Mannschaft zu berichten, dass man regelmäßig privat mit seinen Klienten Zeit verbringt, gegen das Arbeitszeitgesetz verstoßende Dienstpläne erstellt oder die Konto-Pin des Klienten kennt.

Problematisch kann dieser blinde Fleck bezüglich der informalen Seite der Organisation werden, wenn die Berater die formale Struktur der Organisation verändern bzw. dazu raten, ohne zumindest einige Folgen absehen zu können, die dadurch für die informalen Arbeitsweisen entstehen könnten (siehe zu dieser Schwierigkeit – allerdings ohne Beratungsbezug – auch Büchner 2020; Hilgert u. a. 2024). Solche informalen Arbeitsweisen werden allzu schnell als das Funktionieren der Organisation behindernde und deshalb zu beseitigende Missstände definiert. Aber so leicht, das hat die systemtheoretische Organisationsforschung ebenfalls gezeigt, ist es nicht. Organisationen sind mit einer Umwelt konfrontiert, die widersprüchliche Erwartungen an sie stellt und an die sich die Organisation zugunsten ihres Funktionierens anpassen muss (Luhmann 1972, S. 305). Auch die Sozialarbeitswissenschaft sieht diese widersprüchlichen Umwelterwartungen und greift zumindest eine davon immer wieder unter dem Begriff des Doppelmandats auf: Sozialarbeiter „sollen ihren Klienten helfen und sie sollen für die Gesellschaft die Funktion der sozialen Kontrolle ausüben (bspw. straffällig gewordene Jugendliche resozialisieren oder sterbende Menschen aus der Öffentlichkeit fernhalten“ (Dunkel 2011, S. 192). Allerdings lassen sich solche Widersprüchlichkeiten aus Konsistenzgründen nicht in die formalen Regeln übersetzen. Andernfalls wären die Regeln in sich widersprüchlich und ihre Funktion, falsches Handeln sanktionierbar zu machen, würde verloren gehen. Insofern können informale Arbeitsweisen durchaus funktional für die Organisation sein, weshalb seitens der Berater auch wohlüberlegt sein sollte, ob und inwieweit diese durch Veränderungen in der Formalstruktur irritiert oder sogar behindert werden könnten. Wir wollen das an einem Beispiel illustrieren.

In einer der Organisationen plant die Geschäftsführung – mithilfe der Beratung –, den telefonischen Kontakt zwischen Klienten und Sozialarbeitern zu reorganisieren, indem sie eine zentrale Telefonstelle einführen. Bisher besaß jeder Sozialarbeiter eine eigene Telefonnummer und war unter dieser für seine Klienten erreichbar. Der Geschäftsführung geht es dabei um zweierlei: Einerseits soll die zentrale Telefonstelle die Terminklärung für Sozialarbeiter und Klient übernehmen – bisher vereinbarten die Sozialarbeiter direkt Termine mit den Klienten. Der Kontext ist, dass die Klienten Termine oftmals absagen und die Sozialarbeiter die Termine nicht neu besetzen und einige stattdessen – so die Geschäftsführung – „Däumchen drehen“. Die Hoffnung ist also, die Arbeitszeit der Sozialarbeiter durch die zentralisierte Terminregelung effizienter nutzen zu können, indem für sie zentral neue Termine gebucht werden. Andererseits soll die Zentralisierung verhindern, dass die Sozialarbeiter ihren Klienten eine zeitlich zu extensive Betreuung zukommen lassen. Etwa sollen Sozialarbeiter den Klienten keine Hilfe außerhalb der dafür vorgesehenen Arbeitszeiten leisten – was laut verschiedener Sozialarbeiter unterschiedlicher Organisationen gängige Praxis im Feld ist.

Die Geschäftsführung definiert beide Punkte negativ und sieht Veränderungsbedarf. Gleichzeitig geht durch diese Perspektive unter, dass die bisherige dezentralisierte Arbeitsweise für die Organisation durchaus funktional sein kann. Natürlich kann es sein, dass die Sozialarbeiter in den ausgefallenen Terminen tatsächlich nur „Däumchen drehen“. Andererseits kann die Zeit aber auch dienstlich genutzt werden, zum Beispiel um Klientenarbeit zu dokumentieren. So wurde etwa in einer anderen Organisation in einer anonymen Mitarbeiterbefragung mitgeteilt, dass der Arbeitsplan im Alltag nicht genügend Zeit für die Dokumentation vorsehe. Denkbar wäre auch, dass solche freien Zeiten für fachliche Gespräche zwischen Tür-und-Angel genutzt werden (siehe hierzu etwa Ackermann 2017, S. 154). Auch der zweite Punkt bezüglich der Extensivierung der Arbeitszeit sollte nicht ausschließlich als dysfunktionale Abweichung abgetan werden. Durch die informale Entgrenzung der Arbeitszeit schaffen die Sozialarbeiter – auf eigene Kosten – Flexibilität, wodurch sie sich besser an ihre Klienten anpassen können. Diese Anpassung ist vor allem deshalb wichtig, weil das Vertrösten der Klienten auf einen späteren Zeitpunkt die zwischenmenschliche Beziehung belasten kann und somit die Mitarbeit bzw. Koproduktion (siehe hierzu Drepper/Tacke 2010, S. 268; Bommers/Scherr 2018, S. 258) der Klienten gefährdet (siehe hierzu konkret den empirischen Fall bei Streck/Unterkofler 2023, S. 130).

Fehlt es den Beratern nun an Informationen über solche informalen Arbeitsweisen, können sie offensichtlich nicht abschätzen, welche möglichen Folgeprobleme sie durch die Veränderung der formalen Regeln – hier: Einrichten einer zentralen Telefonstelle – erzeugen. Die potenziellen Folgeprobleme – Verlust an Flexibilität, Belastung der Sozialarbeiter-Klienten-Beziehung – könnten dann den nachhaltigen Erfolg des Beratungsprojekts beeinträchtigen. Und sei es nur, weil es Zeit bedarf, neue ‚workarounds‘ zu finden (Büchner 2020, S. 372; zu workarounds siehe auch die Studie von Huuskonen/Vakkari 2013). Die Vorteile der Umstellung – die ja durchaus vorhanden sein können – erkaufte man sich dann mit einem Bouquet an nicht absehbaren Folgeproblemen.

Zugegebenermaßen kann Beratung, die sich wenig für das Informale interessiert, auch relativ harmlos enden. Nämlich dann, wenn die Veränderungen ohnehin mehr oder weniger folgenlos für den Organisationsalltag sind. Das ist der Fall, wenn an Themen gearbeitet wird, die primär der Außendarstellung der Organisation dienen. Also etwa die Arbeit an wert aufgeladenen Leitbildern, die für den organisationalen Alltag geringe Relevanz besitzen, weil sie zu abstrakt sind, als dass aus ihnen konkrete Handlung abgeleitet werden könnten (Kühl 2017, S. 11 ff.). Ein anderer Fall wäre, dass zwar grundlegende strukturelle Veränderungen angestrebt werden, diese allerdings in der Formalstruktur so schlecht verankert sind, dass sie nicht zu tragen beginnen. Etwa beschließt man, zukünftig die Klientenarbeit zu dokumentieren, klärt aber nicht, wie diese Dokumentation aussehen soll und wer sie kontrolliert. Und noch ein letzter Fall: Die Beratung kann auch schlicht scheitern. Etwa wenn die Berater mit ihren Veränderungsabsichten starken Widerstand in der Belegschaft erzeugen.

### III. FUNKTIONEN INFORMALITÄTSBLINDER BERATUNG

Allerdings hat diese Form der Beratung, sei sie nun folgenreich oder folgenarm, auch durchaus seine Vorteile. Hierzu muss man lediglich das Referenzsystem wechseln: hin zu der Beratungsorganisation.

Zunächst lässt sich wohl unterstellen, dass eine stärker an Informalität interessierte Beratung schwieriger an den Kunden zu verkaufen ist. Denn die dafür nötigen Informationen müssen aufwendig erhoben werden. Ob die Berater das durch vertrauliche Interviews oder eher ethnografische Ansätze zu erreichen versuchen, ist zunächst einmal zweitrangig. Der Punkt ist, dass die Erhebung teuer zu bezahlender Zeit bedarf. Die Berater stehen dann vor der Schwierigkeit, diesen Preis vor dem Kunden zu rechtfertigen und ihn zur Zahlungswilligkeit zu bewegen.

Darüber hinaus beauftragen in der Regel Geschäftsführer Organisationsberatungen mit der Intention, bestimmte Probleme zu lösen oder bestimmte Vorhaben umzusetzen. Diese Vorhaben können – so kann sich bei der Beratung herausstellen – aber durchaus auch dysfunktionale Folgen für die Organisation haben. So laufen die Berater Gefahr, zum Überbringer schlechter Nachrichten zu werden – und dafür verantwortlich gemacht zu werden.

Für den letzten Vorteil der weniger an der Informalität interessierten Beratung können wir uns auf die Studie von Stefan Kühl (2007) beziehen. In dieser untersucht er ein Beratungsprojekt, in dem die Berater Wissen über die informalen Praktiken der Organisationsbasis sammeln konnten. Die Folge dieses Wissens war jedoch, dass die Berater in Rollenkonflikte zwischen ihren Informanten und der Führungsetage der Organisation gerieten: „Die Berater selbst werden jedoch schrittweise in die informellen Strukturen hineingezogen und tragen zum Schutz der Informalität gegenüber dem Top-Management bei“ (Kühl 2007, S. 274). Dieses Phänomen erklärt er damit, dass die Berater mindestens für potenzielle Folgeprojekte auf die Kooperation ihrer Informanten angewiesen sind und sie ihnen deshalb nicht schaden wollen (Kühl 2007, S. 285). Denn zur konstruktiven Kooperation der Informanten kann nicht ohne weiteres motiviert werden (zur theoretischen Herleitung dieses Problems siehe Weißmann 2023, S. 51–54). Die Sozialarbeiter und Berater teilen sich dieses Problem. Die Schwierigkeit solcher Rollenkonflikte können natürlich vermieden werden, indem man mit den Informanten gar nicht erst in Kontakt tritt.

### IV. FAZIT

Zusammengefasst stellen wir die These auf, dass eine über Informalität informierte Beratung vielleicht die für die zu beratende Organisation nachhaltigere Beratung ist, aus der Sicht der Berater aber auch einiges dafürspricht, diese gerade nicht zu ausgeprägt anzubieten. Nicht allein, weil sie – schon aufgrund der Informationsbeschaffung – die komplexere und aufwendigere Beratung ist. Sondern auch, weil diese Form der Beratung die Berater in ganz eigene Probleme führt. Man verkauft lieber ein kleineres, günstigeres Projekt als gar keins. Man riskiert lieber nicht den Unmut der Geschäftsführung, weil man schlechte Botschaften überbringen muss. Man meidet die Probleme, die sich aus Rollenkonflikten ergeben würden. Diese Interessenskonflikte zwischen Beratern und zu beratenden Organisationen sind Beratungsprozessen inhärent und nicht einfach zu lösen. Aber vielleicht müssen sie auch gar nicht immer gelöst werden. Immerhin hat auch die Beratung, die auf die umfangreiche Erhebung von Informationen verzichtet, ihre Funktion; und sei es nur die Legitimierung eines Vorhabens durch den Status der Berater (siehe hierzu Sturdy 1997, S. 513). Die Frage ist nur, wie nachhaltig nützt es auch der zu beratenden Organisation?

## DANKSAGUNG

Wir bedanken uns für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

(Hrsg.), *Organisierte Grenzrollen: Außendienst und Publikumskontakte in soziologischer Perspektive*. Wiesbaden: Springer Fachmedien, S.39–104, abrufbar unter: [https://doi.org/10.1007/978-3-658-40597-7\\_1](https://doi.org/10.1007/978-3-658-40597-7_1), letzter Zugriff am 8.1.2024.

## LITERATUR

- [1] Ackermann, Timo (2017): Über das Kindeswohl entscheiden: Zur Fallarbeit im Jugendamt, transcript Verlag, abrufbar unter: <https://www.degruyter.com/document/doi/10.1515/9783839437513/html>, letzter Zugriff am 14.2.2024.
- [2] Bommers, Michael/Scherr, Albert (2018): *Soziologie der Sozialen Arbeit Eine Einführung in Formen und Funktionen organisierter Hilfe 2., vollst. überarbeitete Auflage.*, Weinheim: Beltz Juventa.
- [3] Büchner, Stefanie (2020): Formalität und Informalität unter den Vorzeichen der Digitalisierung, in: Nadia Kutscher/Thomas Ley/UDO Seelmeyer/Friederike Siller/Angela Tillmann/Isabel Zorn (Hrsg.), *Handbuch Soziale Arbeit und Digitalisierung*. Weinheim Basel: Beltz Juventa.
- [4] Drepper, Thomas/Tacke, Veronika (2010): Zur gesellschaftlichen Bestimmung und Fragen der Organisation ›personenbezogener sozialer Dienstleistungen‹. Eine systemtheoretische Sicht, in: Thomas Klatetzki (Hrsg.), *Soziale personenbezogene Dienstleistungsorganisationen: soziologische Perspektiven*. Wiesbaden: VS Verlag für Sozialwissenschaften, S.241–284.
- [5] Dunkel, Wolfgang (2011): Arbeit in sozialen Dienstleistungsorganisationen: die Interaktion mit dem Klienten, in: Adalbert Evers/Rolf G. Heinze/Thomas Olk (Hrsg.), *Handbuch Soziale Dienste*. Wiesbaden: VS Verlag für Sozialwissenschaften, S.187–205, abrufbar unter: [https://doi.org/10.1007/978-3-531-92091-7\\_9](https://doi.org/10.1007/978-3-531-92091-7_9), letzter Zugriff am 29.11.2023.
- [6] Hilgert, Christian/Genth, Felix/Junne, Jaromir/Lahme, Lukas (2024): *Digitalität und Ambiguität. Organisationskulturen der sozialen Arbeit unter Druck.*, Hamburg: im Erscheinen.
- [7] Huuskonen, Saila/Vakkari, Pertti (2013): “I Did It My Way”: Social workers as secondary designers of a client information system, in: *Information Processing & Management* 49, S. 380–391.
- [8] Kieserling, André (2002): *Öffentlichkeit als Zensurmechanismus*, München: unveröffentlicht.
- [9] Köhl, Stefan (2007): Formalität, Informalität und Illegalität in der Organisationsberatung: Systemtheoretische Analyse eines Beratungsprozesses, in: *Soziale Welt* 58, S. 271–293.
- [10] Köhl, Stefan (2017): *Leitbilder erarbeiten*, Wiesbaden: Springer Fachmedien Wiesbaden, abrufbar unter: <http://link.springer.com/10.1007/978-3-658-13423-5>, letzter Zugriff am 25.6.2024.
- [11] Köhl, Stefan (2020): *Brauchbare Illegalität: Vom Nutzen des Regelbruchs in Organisationen*, Campus Verlag.
- [12] Luhmann, Niklas (1972): *Funktionen und Folgen formaler Organisation 2. Aufl.*, Berlin: Duncker & Humblot.
- [13] Streck, Rebekka/Unterkofler, Ursula (2023): Situative Herstellung des Persönlichen. Ethnografische Perspektiven auf Aushandlung von Beziehung in wenig formalisierten Settings der Sozialen Arbeit, in: *Österreichisches Jahrbuch für Soziale Arbeit*, S. 116–138.
- [14] Sturdy, Andrew (1997): The Dialectics of Consultancy, in: *Critical Perspectives on Accounting* 8, S. 511–535.
- [15] Weißmann, Martin (2023): Grenzrollen und Grenzsysteme: Funktionen und Strukturmerkmale des Handelns an den Grenzen sozialer Systeme, in: André Kieserling/Martin Weißmann

# Working from Home: Opportunities for Transformational and Health-Oriented Leadership & Specific Challenges Arising from Remote and Hybrid Work

Stephanie M. Neidlinger<sup>1,2\*</sup>, Dorothee Tautz<sup>1</sup>, Katharina Schübbe<sup>1</sup>, Annika Krick<sup>1</sup>, Sven Hauff<sup>1</sup> & Jörg Felfe<sup>1</sup>

<sup>1</sup> *Helmut Schmidt University/University of the Federal Armed Forces*  
Hamburg, Germany

<sup>2</sup> *Harvard T. H. Chan School of Public Health*  
Boston, MA, USA

\*neidlinger@hsu-hh.de

**Abstract** – The rapid evolution of remote and hybrid work arrangements has fundamentally reshaped leadership dynamics, presenting novel challenges and opportunities. As "new ways of working" increasingly become the norm, their impact on our professional lives promises to be long-lasting. This study delves into several critical areas centring around good leadership practices in increasingly digitalized and virtual settings: 1) the applicability of transformational and health-oriented leadership styles in remote and hybrid settings; 2) the impact of effective leadership on employee outcomes; 3) the dependency of leadership effectiveness in remote and hybrid environments on specific working conditions; 4) the benefits that leaders derive from employing healthy leadership styles; and lastly, 5) the distinctive challenges leaders face when managing teams remotely versus in traditional office settings. Our findings are drawn from data collected from two samples throughout Germany and across all sectors. Sample 1 consisted 1318 leaders 2180 employees surveyed in April 2021, sample 2 were 907 leaders and 2124 employees surveyed in September 2022. In our discussion, we outline practical implications for leaders and HR professionals, spotlighting strategies for integrating effective leadership practices within remote and hybrid work frameworks. This study not only highlights the evolving nature of leadership in the digital age but also offers insights into fostering a productive and supportive work environment, irrespective of the physical workspace.

**Keywords** – *Transformational leadership, health-oriented leadership, remote work, hybrid work*

## I. INTRODUCTION

### A. Remote & Hybrid Work Arrangements

It is widely agreed that the COVID-19 pandemic brought about fundamental technological changes in the workplace. Many experts say hybrid work models will be the future of work arrangements: in these models, both remote work and office-based work are possible for employees and leaders alike. Knowing the changing impacts of these differences in work location must be considered when deriving recommendations for leaders and organisations.

A study from our project "Digital Leadership & Health" in 2021 showed that 51% of respondents wished to continue working from home for two to three days after the Covid-19

pandemic. Furthermore, 21.4% even wished to work from home four days or entirely. Among leaders, however, the proportion of those wanting to work four or more days from home was lower, at 18.6% [1]. This indicates that hybrid work is the preferred option amongst most but seems to be more popular from employees', rather than leaders' perspective.

The increasing prevalence of hybrid work raises the questions of work outcomes and potential risks and benefits which can be brought about by remote and hybrid work. Many recent studies have focused on the effects of work arrangements on employee outcomes. Most previous works reach the conclusion that remote and hybrid work can be very beneficial for commitment [2], engagement [3], health [4], [5], work-life balance [6] and even organisational performance [7].

Although the benefits of remote and hybrid work are well-documented, the effects on employee outcomes are complex and can vary depending on several factors, leading to contradictory findings in some cases. Some studies such as Patel and colleagues [8] have identified increased presenteeism and burnout among remote workers due to blurred boundaries between work and personal life. Jensen and colleagues [9] suggest that remote work can lead to feelings of isolation and disconnection. While the effects of remote and hybrid work for employees have been in the centre of research attention in the past years, research regarding the effects of remote work on leaders and leadership are scarce.

Specifically, positive employee-oriented leadership styles might suffer in the remote context due to reduced contact and communication. With this study, we aim to disentangle whether transformational and health-oriented leadership work in remote settings and if they differ in effectiveness for remote vs. office contexts.

A previous interview study investigated the feasibility of transformational leadership and health-oriented leadership behaviours from both leader and employee perspective [10]. The study revealed five distinct challenges for both leadership styles: lack of social presence and interaction, lack of spontaneous and informal conversations, digital communication and technical problems, less trust and bonding, and less responsibility of the leader. Based on this it can be assumed that the

effectiveness of both leadership styles changes in remote working contexts compared to traditional office contexts. Next to these rather specific challenges for transformational leadership and health-oriented leaders, other research found further challenges for leadership such as the implementation of standard leadership tasks and functions (control, planning, motivation and coordination, team development, health promotion, etc.) [11].

Within this paper, we will provide an overview of leadership behaviours in remote vs. hybrid scenarios. Additionally, we will report some specific challenges for leaders in hybrid work settings who experience both on-site and remote leadership.

### B. Transformational Leadership

Transformational Leadership (TFL) [12] is a well-known and well-established leadership style in practice and leadership research. Its effectiveness on employee outcomes such as health, well-being, performance, or satisfaction has been demonstrated in various studies [13–16].

Transformational leadership comprises five sub-dimensions: (1) Idealized Influence Attributed, which refers to the leader acting as a role model and sharing values; (2) Idealized Influence Behaviour, which refers to the leader fostering team spirit; (3) Inspirational Motivation, which describes the leader inspiring enthusiasm for the future; (4) Intellectual Stimulation, which refers to the leader fostering creativity and stimulating new ways of thinking and (5) Individualized Consideration, which means that the leader takes employees' needs and personal strengths into account.

Research on TFL has mainly been conducted in traditional office contexts, while results regarding its effectiveness in remote contexts are scarce and inconsistent. While Purvanova & Bono [17] proved in their experimental study that transformational leaders were more effective for employee performance in the remote setting than in face-to-face (F2F) settings, Eisenberg and colleagues [18] and Hoch & Kozłowski [19] found the opposite in their survey studies. Purvanova & Bono [17] explain their finding by postulating that leaders reinforce their transformational leadership behaviour in uncertain, ambiguous contexts. Eisenberg and colleagues [18], on the other hand, postulate that the leaders' authenticity decreases due to remote work and fewer F2F contacts, so that TFL effectiveness decreases.

These contradictory findings indicate that further research is needed on the effectiveness of TFL and the boundary conditions that promote or inhibit its effectiveness in remote and hybrid work contexts. Moreover, previous research has mainly focused on employees who work fully remote and never meet with their leader in person. The question therefore arises as to whether these results can also be transferred to hybrid contexts.

Also, research questions arise about the effectiveness of TFL for different employee outcomes, as previous research has mainly focused on performance and the impact of the different sub-dimensions.

Based on this, we propose the following research questions:

**RQ A.1. Does transformational leadership work in remote settings?**

**RQ A.2. Is transformational leadership just as effective in remote settings as in traditional office settings?**

**RQ A.3. Does remote employees' perception of transformational leadership depend on other factors?**

### C. Health-oriented Leadership

The concept of health-oriented leadership (HoL) [20], describes leadership behaviours which put health and well-being of both leaders and their teams first and centre. Research underscores the critical role of leadership behaviour in influencing employee health and stresses the significance of adopting leadership styles that actively foster health [21].

The HoL model [20] encompasses two principal leader behaviours: StaffCare, which describes employee-directed behaviours and SelfCare, describing self-directed leader behaviours. In terms of StaffCare, leaders are encouraged to undertake proactive measures to enhance their team's health, such as identifying and mitigating health risks and improving working conditions to boost job satisfaction and engagement [22], [23], [24]. SelfCare is centred on the personal health of the leaders, highlighting the necessity for leaders to maintain their own health as a means to effectively support their teams [25], [26].

The interplay between SelfCare and StaffCare suggests that leaders who are adept at managing their own health are more capable of supporting their employees, thereby improving their health outcomes [27], [28]. Furthermore, studies indicate that HoL more comprehensively accounts for employee well-being compared to other leadership styles [29], [30].

Some recent studies have examined remote work as an antecedent of HoL: A study by Krick and colleagues [1] also explored the frequency of leaders' remote work as a predictor of their SelfCare and StaffCare behaviours. They found that higher frequencies of remote work supported both SelfCare and StaffCare, thus contributing positively to health-oriented leadership (HoL). Krick and colleagues suggested the reduced direct contact with employees may make leaders feel more responsible for their employees' health, though acting on these responsibilities from a distance can pose challenges.

An experimental study by Schübbe and colleagues [31] focused on leaders' ability to recognize health-specific warning signals from employees: They compared warning signal recognition across three communication settings – F2F, video-call and phone calls – and found that detection of employees' health-related warning signals was easier in face-to-face communication. Digital communication thus seems to pose a challenge for leaders' health-oriented leadership. Klebe and colleagues [32] investigated the role of information-communication technology (ICT) hassles on leaders' StaffCare behaviours via irritation in a sample of hybrid leaders who worked from home at least one day per week. They found that ICT hassles were related to lower amounts of StaffCare via irritation. Their study used a longitudinal three-wave design and delivers first empirical support for the effects of technological problems in leader-follower communication of HoL.

To further investigate this, we postulate the following research questions for health-oriented leadership:

**RQ A.4. Does HoL work in remote settings?**

**RQ A.5. Is HoL just as effective in remote settings as in traditional office settings?**

### **RQ A.6. Does remote employees' perception of HoL depend on other factors?**

Further questions regarding the effects of remote and hybrid work on leaders' health and SelfCare (the extent to which leaders take care of themselves) are relatively unknown. Recent work suggests that remote work can be beneficial for leaders in terms of their health and work-life balance [33]. It was also found that the effects of remote work on leaders' outcomes depend on other factors (e.g. personality). Based on this, we propose the following research questions:

### **RQ B.1. Do leaders benefit from implementing healthy leadership styles?**

### **RQ B.2. Is health-oriented self-leadership achievable in remote settings?**

### **RQ B.3. Does the extent of health-oriented self-leadership depend on other factors?**

#### *D. Specific Challenges for Leaders in Remote vs. office Settings*

Leaders consistently encounter demands at work which are different from the demands employees without leadership roles are facing. These include decision-making, planning, the need for constant availability, frequent interruptions [34], and prolonged work hours [35]. Such pressures can lead to increased health risks and potentially harmful levels of presenteeism [36], [37]. Thus, we want to focus on leaders' experienced work demands to shed some light on their specific issues.

For leaders in hybrid roles, there may be specific challenges associated with each workplace. Remote leadership can induce specific challenges due to the reduced opportunities for direct face-to-face interaction [10]. For instance, motivating teams, providing timely feedback, and monitoring the progress of tasks can become significantly more difficult in a remote setting. This is contrasted with in-office scenarios where spontaneous meetings and casual interactions can foster direct communication and quick problem solving. In hybrid environments, leaders navigate these varied settings. This duality can complicate leadership roles, as strategies which are effective in one context may not translate seamlessly to the other. For example, building team cohesion and maintaining open communication lines are straightforward when leaders can physically gather teams, but these become complex in a digital environment where engagement and non-verbal cues are limited. The same was found for health-related leadership behaviours, such as attentively looking for health-related cues in employees' speech and behaviours which are harder to observe in digital communication [31]. We want to answer the following question:

### **RQ C.1. Which specific challenges do hybrid leaders face in direct comparison of their remote and office workplaces?**

#### *E. Outline & Added Value of the Current Paper*

Overall, remote and hybrid settings seem to create a specific set of challenges for leaders. This also raises the question of whether traditional leadership styles such as transformational leadership [12], and more recent leadership styles, such as HoL [14] are transferable to remote and hybrid settings. Furthermore, we want to provide practitioners from various backgrounds with direct comparisons between leadership in

remote and office scenarios by deriving some specific recommendations for hybrid leaders.

With the current paper we aim to answer the following five key questions and shed some light on whether transformational and health-oriented leadership styles 1) can be implemented in remote and hybrid settings, 2) are equally effective in fostering positive employee outcomes, 3) whether their effectiveness depends on specific working conditions or other factors in remote contexts, 4) whether remote and hybrid leaders themselves benefit from engaging in these leadership behaviours and, 5) which specific challenges hybrid leaders face in their remote vs. office workplaces.

## II. METHOD

### *A. Sample & Procedure*

Sample 1 included 3715 German employees, of which 37.7% ( $N_{Leaders} = 1318$ ) were currently working in a leadership role and 62.3% were regular employees without a leadership role ( $N_{Employees} = 2180$ ). 50.3% of the sample identified as male (4 participants indicated their gender as "other"). Mean age was 43.62 years. The sample represented diverse sectors including the metal and electrical industry, chemical and pharmaceuticals, energy, construction, crafts, logistics, transport and traffic, banking and insurance, real estate, corporate, legal, personnel and tax consulting, advertising, trade, security, IT and telecommunications, education, training and science, medicine and healthcare, as well as public administration. Participants were surveyed about their remote/hybrid or office-based work arrangements as well as leadership and work and life outcomes. The study included participants in all work arrangements, ranging from fully office-based (16.2%), 1-2 remote days per week (24.5%), 3-4 remote days per week (28.5%), to fully remote (31%). Participants were surveyed once in April 2021. 24.3% ( $n = 320$ ) of leaders were fully remote, 10.2% ( $n = 134$ ) of leaders were fully office based and 65.5% of leaders ( $n = 899$ ) were in hybrid work arrangements (between 1 and 4 remote days per week). 34.8% ( $n = 759$ ) of employees were fully remote, 19.9% ( $n = 434$ ) of employees were fully office based and 45.3% of employees ( $n = 987$ ) were in hybrid work arrangements (between 1 and 4 remote days per week).

Sample 2 included  $N = 3031$  German employees, of which 29.9% ( $N_{Leaders} = 907$ ) were currently working in a leadership role and 70.1% ( $N_{Employees} = 2124$ ) participants were regular employees without leadership responsibilities. Sample 2 also included a wide range of sectors. Participants were surveyed about their remote/ hybrid or office-based work arrangements as well as leadership and work and life outcomes. Sample 2 also included participants in all work arrangements, ranging from fully office-based (27.1%), 1-2 remote days per week (31.4%), 3-4 remote days per week (24.6%), to fully remote (16.8%). Participants were surveyed once in September 2022. 50.8% of the sample identified as male. Mean age was 43 years.

### *B. Measures & Analysis*

The study utilized established scales to assess working conditions, transformational leadership (MLQ [38], [39]) health-oriented leadership (HoL [14]), engagement (UWES-3 [40]), well-being (WHO5 [41]), commitment (COMMIT [42]), job satisfaction (COPSOQ [43]), health (COPSOQ; [43]), stress (Irritation Scale [44]), and physical complaints (Health Impairment Scale [45]).

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	Age	–	-.08**	-.09**	.15**	–	.07**	–	.19**	.07*	-.04	.07*	-.22**	.10**
2	Gender	-.04	–	.03	.06*	–	.09**	–	-.06*	-.07*	-.11**	-.11**	.17**	-.10*
3	Remote work intensity	.05*	-.07**	–	.03	–	.06*	–	-.07*	-.01	.06*	.02	.09**	.01
4	Transformational Leadership (L)	–	–	–	–	–	.66**	–	.48**	.39**	.29**	.28**	-.14**	.38**
5	Transformational Leadership (E)	-.10**	-.08**	.08**	–	–	–	–	–	–	–	–	–	–
6	HoL StaffCare (L)	.08**	-.07**	–	–	.25**	–	–	.37**	.31**	.27**	.25**	-.15**	.35**
7	HoL StaffCare (E)	-.09**	-.12**	.07**	–	.75**	–	–	–	–	–	–	–	–
8	Commitment	.04	-.05*	.07**	–	.59**	–	.51**	–	.58**	.35**	.45**	-.25**	.46**
9	Engagement	.05*	-.03	.07**	–	.48**	–	.44**	.59**	–	.42**	.56**	-.19**	.59**
10	Health	-.07**	-.10**	.04*	–	.25**	–	.28**	.28**	.35**	–	.46**	-.41**	.55**
11	Job Satisfaction	.04	-.06**	.03	–	.42**	–	.41**	.48**	.53**	.41**	–	-.30**	.53**
12	Physical Complaints	-.07**	.24**	.12**	–	-.18**	–	-.24**	-.22**	-.31**	-.61**	-.41**	–	-.37**
13	Well-being	.07**	-.11**	-.05*	–	.32**	–	.33**	.35**	.54**	.56**	.53**	-.60**	–

Note: N = 1318 leaders, N = 2180 employees. Correlations for leaders are reported above the diagonal, correlations for employees are reported below the diagonal. \* p < .05. \*\* p < .01. (L) = rated by leaders, (E) = rated by employees. Gender was coded 1 = male, 2 = female. Due to the low number of participants who selected "other" as their gender (n = 1 leader, n = 3 employees), option "other" is not reflected in the correlation table.

TABLE 1: CORRELATIONS BETWEEN ALL STUDY VARIABLES FOR LEADERS AND EMPLOYEES (SAMPLE 1)

We used the Top-2-Boxes Method, correlation analyses and z-tests to compare group differences. The Top-2-Boxes method isolates the highest (or lowest) ratings on a Likert scale. All options on a scale (e.g. StaffCare) are categorized (e.g. values 4 and 5 representing "high" StaffCare) and their proportions are reported in a table containing other selected variables.

### III. RESULTS

Table 1 shows the correlations between study variables for Sample 1 for leaders (above the diagonal) and employees (below the diagonal), respectively.

For leaders in Sample 1, correlations between remote work and their self-rated transformational leadership behaviour, as well as their self-rated StaffCare, were slightly positive, but only the relationship between remote work and transformational leadership was significant. For employees, correlations between remote work and the perceived transformational leadership and StaffCare of their leader were slightly positive.

For leaders, correlations between self-rated transformational leadership behaviour and self-rated StaffCare with the outcomes were all moderate and significant. For employees, correlations between perceived transformational leadership and StaffCare of their leader with the outcomes were also moderate and significant.

Older leaders reported more self-rated transformational leadership behaviours and StaffCare than younger leaders. Younger employees perceived more transformational leadership and StaffCare than older employees. Female leaders reported higher transformational leadership behaviour. Male employees perceived more transformational leadership and StaffCare than female employees.

#### A. Leadership Styles

1) *Does transformational leadership work in remote settings? (RQ A.1.)*

When we look at the perception of leadership in Sample 1, fully remote employees (n = 759) generally perceive more transformational leadership compared to their office-based counterparts (n = 434). For instance, 47.1% of fully remote employees report that their leaders consider their individual needs and 48.9% say they promote team spirit. This is slightly higher than the 39.5% and 40% reported by fully office-based

employees. The differences, though minor, suggest that leadership quality may be seen more positively when remote work is an option. Remarkably, leadership quality is not rated lower for those working mainly or fully remotely; in fact, these work arrangements often show better outcomes.

This positive view is echoed in fully remote vs. fully office-based leaders' self-assessments from Sample 1, indicating fully remote leaders do not experience a decline in their leadership quality compared to their fully office-based counterparts.

When comparing leadership behaviour directly within Sample 2, employees in hybrid models (who experience leadership both on remote and office days) perceive transformational leadership as slightly weaker remotely, and stronger in the office. For example, only 51.6% felt their leader was highly confident in achieving goals when they worked remotely, versus 54.5% in the office setting. Also, it seems leaders are better at highlighting the importance of team spirit and understanding tasks, and recognizing individual needs and goals in the office (45%) compared to remotely (41.7%).

This leads to a somewhat paradoxical finding: remote work is rated more favourably when comparing different employee groups but rated slightly less favourably in direct comparison to the office context. However, it is noted that those in hybrid work arrangements experience more transformational leadership in both settings than those who are fully office-based.

Thus, employees tend to view leadership in a hybrid setup more favourably than without remote work. This may be due to additional positive factors that make also leadership seem better. Yet, leadership is perceived somewhat more positively in the office than remotely in a hybrid scenario. This indicates that the absence of direct contact might slightly hinder a positive leadership experience.

2) *Is transformational leadership just as effective in remote settings as in traditional office settings? (RQ A.2.)*

Transformational leadership showed a positive impact on employee engagement, well-being, commitment, job satisfaction, and health in Sample 1, even when working remotely. However, the influence of transformational leadership is slightly lower in remote work settings compared to office-based settings. We used correlation analyses to compare the

relationships between transformational leadership and the outcomes for  $n = 423$  employees who were fully office-based and  $n = 1746$  employees who worked remotely at least one day per week. Correlation analyses showed that transformational leadership varied in its effectiveness for increasing **engagement** ( $r = .45$  with remote work vs.  $r = .54$  without remote work,  $z = -2.25$ ,  $p = .024^*$ ), **commitment** ( $r = .56$  with remote work vs.  $r = .65$  without remote work,  $z = -2.63$ ,  $p = .009^{**}$ ), **job satisfaction** ( $r = .38$  with remote work vs.  $r = .50$  without remote work,  $z = -2.77$ ,  $p = .006^{**}$ ), and **health** ( $r = .22$  with remote work vs.  $r = .33$  without remote work,  $z = -2.22$ ,  $p = .027^*$ ). Transformational leadership was marginally related to reducing **physical complaints** ( $r = -.16$  with remote work vs.  $r = -.25$  without remote work,  $z = 1.75$ ,  $p = .080^\dagger$ ). Therefore, the effectiveness of transformational leadership is somewhat reduced with remote work.

3) *Does remote employees' perception of transformational leadership depend on other factors? (RQ A.3.)*

Neither individual factors (such as age, gender, personality), social interactions (i.e. frequency of contact, support from leaders), nor working conditions (i.e. flexibility, workload, constant availability) influence the perception of transformational leadership in Sample 1 when working remotely.

4) *Does HoL work in remote settings? (RQ A.4.)*

When comparing perceptions of health-oriented leadership (StaffCare), it appears that employees who work fully or at least partially remotely experience slightly more StaffCare than those who work exclusively from the office. For instance, 25.7% of respondents with remote work options report that their leaders "directly address them when they seem stressed and try to offer solutions" and 18.3% report their leaders "encourage them to ensure sufficient fresh air and physical activity". For employees without remote work options, these figures are somewhat lower at 20.5% and 11.8% respectively. Although the differences are only moderate, it can be inferred that leadership quality, in terms of health, is perceived better with remote work.

Leaders from Sample 1 themselves also report in their self-assessment that they promote and pay attention to the health of their employees even with remote work. Overall, they also state that it is even more successful with remote work than when they are fully office-based. For example, 62.9% of leaders with remote work options say they "directly address employees when they seem stressed and try to offer solutions" and 51.5% "encourage sufficient fresh air and physical activity". Among leaders without remote work options, these figures are slightly lower at 55.5% and 36.5% respectively.

However, when employees from Sample 2 who work in a hybrid model and who thus experience their leaders both remotely and at the office, directly compare leadership behaviour, it turns out that less health-promoting leadership behaviour is perceived remotely, and more health-promoting leadership behaviour is perceived at the office. For example, at the office, 20.7% of employees report that their leader immediately notices "if something is wrong with their health". In contrast, only 13.3% of respondents say the same about remote work.

Thus, as with transformational leadership, a contradictory finding appears: Remote work seems to be perceived more favourably when comparing different remote vs. office-based employee groups, but less favourably in direct comparison by hybrid employees. However, it is important to note that the

hybrid group experiences more health-promoting leadership in both settings than the office group.

Therefore, employees seem to rate leadership in a hybrid setup somewhat better than without remote work. Possibly, further positive conditions are responsible for the more positive assessment of leadership. Yet, in the hybrid scenario, leadership is perceived somewhat better in the office than remotely. This suggests that the lack of direct contact might slightly impair the leadership experience.

5) *Is HoL just as effective in remote settings as in traditional office settings? (RQ A.5.)*

From the employees' perspective, health-oriented leadership positively impacts health and well-being, as well as engagement, commitment, and job satisfaction of employees working remotely in Sample 1. The influence of StaffCare on employees' health and work-related attitudes is somewhat, but not significantly lower in remote work settings compared to non-remote settings. We used correlation analyses to compare the relationships between HoL and the outcomes for  $n = 423$  employees who were fully office-based and  $n = 1746$  employees who worked remotely at least one day per week. For example, HoL was not significantly related to reducing **physical complaints** ( $r = -.23$  with remote work vs.  $r = -.28$  without remote work,  $z = 1.00$ ,  $p = .317$ ) and in increasing **engagement** ( $r = .42$  with remote work vs.  $r = .49$  without remote work,  $z = -1.65$ ,  $p = .098^\dagger$ ) or **commitment** ( $r = .49$  with remote work vs.  $r = .53$  without remote work,  $z = -1.01$ ,  $p = .311$ ). The most significant difference was observed in **job satisfaction** ( $r = .38$  with remote work vs.  $r = .48$  without remote work,  $z = -2.30$ ,  $p = .021^*$ ). Therefore, the effectiveness of health-promoting leadership is slightly less with remote work than without.

Overall, the positive effects of health-oriented leadership on employee health [20] and well-being [27] demonstrated in previous studies can also be confirmed in remote work settings: Employees whose leaders actively promote employee health report better health conditions (better general health, higher well-being, fewer physical complaints, and psychological strains) even with increased remote work. Furthermore, positive effects of health-promoting leadership on employee attitudes towards work (engagement, commitment, and job satisfaction) are also valid in remote work settings.

6) *Does remote employees' perception of HoL depend on other factors? (RQ A.6.)*

Contrary to the perception of transformational leadership, which seems to be independent of working conditions and demographic factors, the perception of health-oriented leadership does depend on other factors. Whether and to what extent health-promoting leadership is exhibited in remote work depends on working conditions: **Communication** (synchronous vs. asynchronous) between leaders and employees and the **leaders' SelfCare** are crucial factors.

Leaders from Sample 1 report that they promote the health of their employees (**StaffCare**) more in remote work than in the office. This effect is even amplified when they communicate synchronously more often than asynchronously with their employees in digital collaboration (i.e. more via video or phone than email): Of the leaders surveyed whose communication with employees in remote work is predominantly synchronous, 60% report actively promoting their employees' health remotely. In contrast, only 27.7% of leaders who com-

municate asynchronously in remote work say the same. Therefore, the proportion of health-promoting leadership is 32% higher when leaders communicate synchronously in remote work. For digital collaboration in the office, synchronous communication is also crucial (41.7% StaffCare with synchronous vs. 23.1% StaffCare with asynchronous communication). However, the increase in StaffCare through synchronous communication in the office is less than in remote work. These differences are depicted in Figure 1. Since there is a lot of F2F contact in the office besides digital communication, synchronous communication in digital collaboration in the office seems less critical for healthy leadership. Conversely, with remote work, where there is less F2F contact, it seems helpful for leaders to maintain direct contact with their employees to show more StaffCare and act as role models.

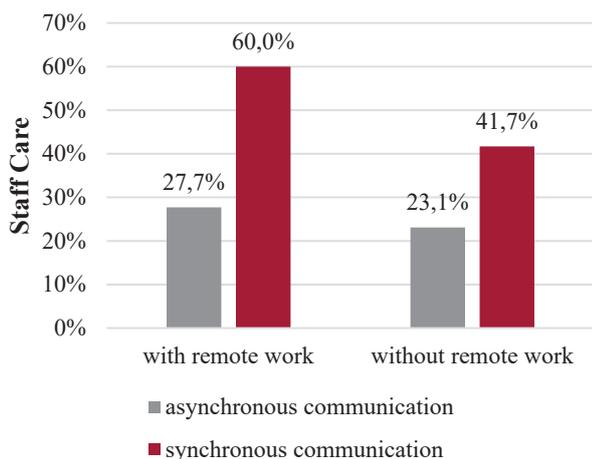


FIGURE 1: COMPARISON BETWEEN LEADERS SYNCHRONOUS VS. ASYNCHRONOUS COMMUNICATION WITH THE FOLLOWERS IN REMOTE VS. FULLY OFFICE-BASED SETTINGS FOR STAFFCARE.

Moreover, the **SelfCare** practices of leaders seem to amplify the positive effects of remote work on StaffCare: Leaders who pay more attention to their own health in remote work report promoting their employees' health more (74.1%) compared to leaders who pay less attention to their own health (9.4%). This corresponds to an increase in StaffCare of about 65% in remote work when leaders also practice SelfCare. In the office setting, those leaders who actively care for their own health report showing more StaffCare (61.5%) compared to those who do less for their own health (16.1%). However, the increase in StaffCare due to the leader's SelfCare is less in the office than in remote work (about 45% in the office vs. about 65% in remote work). These differences are depicted in Figure 2.

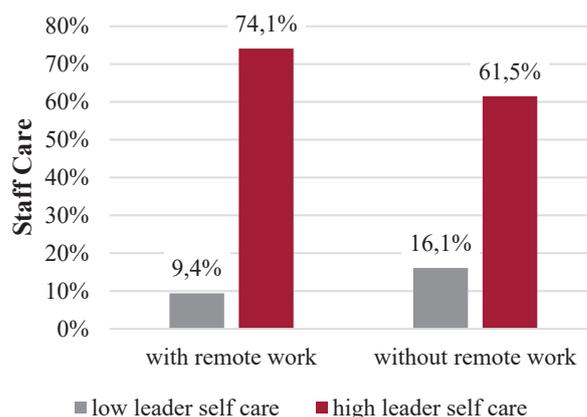


FIGURE 2: COMPARISON BETWEEN LEADERS SELF-CARE (HIGH VS. LOW) IN REMOTE VS. FULLY OFFICE-BASED SETTINGS FOR STAFFCARE.

## B. Leaders Benefit from Healthy Leadership Styles

### 1) Do Leaders benefit from implementing healthy leadership styles? (RQ B.1)

But do good leadership practices also create positive outcomes for leaders? We analysed the correlations between leaders' ( $N = 1318$ ) StaffCare from Sample 1 and their outcomes. Health-oriented leaders who report that they actively promote and support their employees' health in remote work also report better physical **health** ( $r = .27^{**}$ ), lower **stress** ( $r = -.23^{**}$ ), and higher **commitment** ( $r = .37^{**}$ ). Those who ensure health within the team ultimately experience less stress and are more satisfied. However, the opposite direction of the effect is also conceivable: Leaders with more psychological resources are better able to care for their employees' health. It is possible that health-oriented leadership behaviour and one's own psychological resources mutually influence each other.

### 2) Is health-oriented self-leadership achievable in remote settings? (RQ B.2.)

StaffCare was perceived more in remote settings, but what do leaders report about their own SelfCare? When comparing assessments of health-oriented self-leadership (**SelfCare**), it appears that both employees ( $N_{\text{Employees}} = 2180$ ) and leaders ( $N_{\text{Leaders}} = 1318$ ) from Sample 1 who are working remotely between 1 and 5 days per week can take better care of their health than those who are full office-based. For example, 51.1% of respondents with remote work options say they "reduce stress by optimizing their own work methods", "reduce stress by optimizing their working conditions" (48.6%), and "take care of their health during work hours" (36.3%). For employees without remote work options, these figures are significantly lower at 39.9%, 38%, and 26.5% respectively. Leader, however, seem to manage this even better than employees.

When employees and leaders from Sample 2 who are working in a hybrid model, are asked to directly compare health-oriented self-leadership when working remotely vs. when working at the office, we also found that SelfCare is rated slightly better when working remotely and somewhat weaker in the office. For instance, 47.7% of respondents say they "take care of their health during work hours" when working remotely. In the office, only 31.3% manage to do so.

Health-oriented self-leadership of employees positively impacts health and well-being, as well as engagement, commitment, job satisfaction, and the performance of employees in Sample 1. However, the influence of SelfCare is slightly

lower for employees with remote work than without remote work, in terms of reducing **physical complaints** ( $r = -0.36$  with remote work vs.  $r = -0.49$  without remote work,  $z = 2.96$ ,  $p = .003^{**}$ ) and **stress** ( $r = -0.36$  with remote work vs.  $r = -0.50$  without remote work,  $z = 3.21$ ,  $p = .001^{**}$ ). Thus, especially regarding the reduction of stress reactions, the effectiveness of SelfCare in remote work is slightly lower than without remote work. One reason for this could be that employees have more opportunities to support each other or cope with frustration when working in the office. Hence, social support at the office might increase SelfCare so that the negative consequences of stress are reduced. For **commitment** ( $r = .26$  with remote work vs.  $r = .24$  without remote work,  $z = 0.40$ ,  $p = .692$ ), and **engagement** ( $r = .30$  with remote work vs.  $r = .26$  without remote work,  $z = 0.81$ ,  $p = .419$ ), the relationship is slightly stronger in remote work, but the difference is not statistically significant.

The differences between leaders' SelfCare in Sample 1 (partial and fully) remote work vs. fully office-based work are somewhat more pronounced for leaders (reduction of **physical complaints**:  $r = -0.30$  with remote work vs.  $r = -0.51$  without,  $z = 2.75$ ,  $p = .006^{**}$ ; and **psychological stress**:  $r = -0.32$  with remote work vs.  $r = -0.66$  without,  $z = -5.02$ ,  $p < .001^{***}$ ; and improvement of **job satisfaction**:  $r = .38$  with remote work vs.  $r = .57$  without,  $z = -2.69$ ,  $p = .007^{**}$ ) than for employees.

Overall, the positive effects of health-oriented self-leadership can also be confirmed in remote work. Both surveyed leaders and employees reported a better state of health (better general health, higher well-being, fewer physical complaints, and stress) when they actively pay attention to their own health and take care of it while working remotely. Likewise, both leaders and employees reported a more positive attitude towards work (higher engagement, higher commitment, higher job satisfaction), more performance, and higher life satisfaction when they actively care for their own health in remote work.

### 3) Does the extent of health-oriented self-leadership depend on other factors? (RQ B.3.)

Whether and to what extent more health-promoting self-leadership is demonstrated in remote work also depends on some other conditions. The type of communication (synchronous vs. asynchronous) between leaders and employees, as well as stressors such as workload and interruptions, are crucial factors.

The fact that leaders show more StaffCare during synchronous communication in remote work seems to also affect employees' SelfCare. Employees from Sample 1 report that the effect of promoting their own health more in remote work than in the office (SelfCare) is amplified when they communicate synchronously with their leaders: 51.8% of employees whose communication with their leader in remote work is predominantly synchronous report actively taking care of their own health in remote work. In contrast, only 29.9% of employees who communicate asynchronously with their leaders in remote work report actively promoting their health. This corresponds to an increase in health-promoting self-leadership of about 22% when leaders and employees communicate mostly synchronously in remote work. A reason for that could be that leaders better perform as role models in synchronous communication channels and are able to better directly promote their employees' SelfCare (for example by suggesting breaks or discussing the workload and setting priorities). In the office, synchronous communication in digital collaboration is not

crucial for SelfCare (25% SelfCare with synchronous vs. 26.2% SelfCare with asynchronous communication). It can be inferred that the increase in StaffCare supported by synchronous communication in remote work also encourages employees to take greater care of their own health.

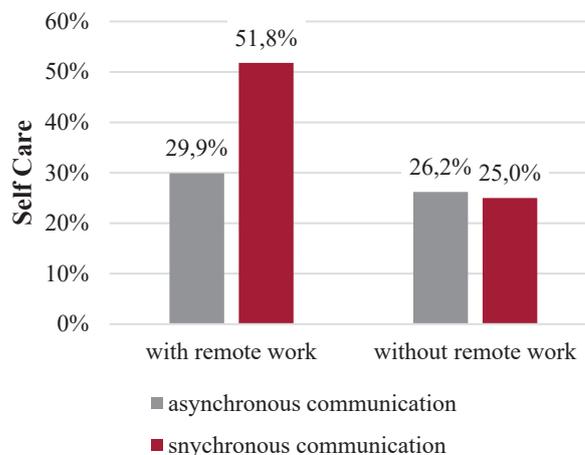


FIGURE 3: COMPARISON BETWEEN LEADERS SYNCHRONOUS VS. ASYNCHRONOUS COMMUNICATION WITH THE FOLLOWERS IN REMOTE VS. FULLY OFFICE-BASED SETTINGS FOR SELF-CARE.

However, as can be seen in Figure 4, a high workload and frequent interruptions are obstacles to SelfCare in remote work. The positive effect of remote work on SelfCare is mitigated by both factors. Employees who report a high workload with increased remote work only promote their health actively in terms of SelfCare 22.2% of the time. In contrast, 50.7% of those with a low workload in remote work report actively promoting their health. This corresponds to an increase in SelfCare in remote work of about 38% when employees are relieved by a lower workload. In the office, only 11.6% report SelfCare with a high workload, compared to 34.9% of employees with a low workload. However, the difference made by workload relief in remote work is more crucial for SelfCare than in the office.

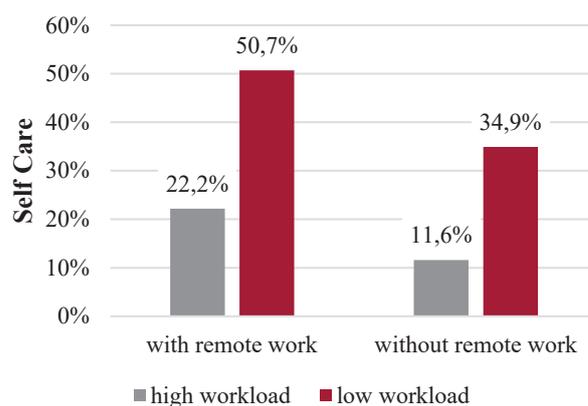


FIGURE 4: COMPARISON BETWEEN WORKLOAD (HIGH VS. LOW) IN REMOTE VS. FULLY OFFICE-BASED SETTINGS FOR SELF-CARE.

The same applies to work interruptions: While employees in the office report an increase in SelfCare from 20.7% to 34.4% due to a decrease in work interruptions, the SelfCare of

employees in remote work increases from 27.2% to 49.6% with fewer interruptions. This corresponds to an increase in SelfCare of about 22% due to fewer interruptions in remote work. Therefore, a reduction in workload and a decrease in interruptions seem to have stronger effects on SelfCare in remote work than at the office.

*C. Which specific challenges do hybrid leaders face in direct comparison of their remote and office workplaces? (RQ C.1.)*

Leaders from Sample 2 who work entirely or predominantly remotely report specific leadership challenges in areas such as supervision, control, support, conflict management, and communication, as well as in dealing with the team (coordination, support, team spirit), especially in comparison to leading in the conventional office routine.

Figure 5 depicts how specific leadership challenges apply to hybrid leaders' workplaces – they reported all challenges for their remote and their office-workplace, respectively.

The general trend shows that in almost every category, leaders find it more challenging to perform their leadership duties in a remote setting as opposed to on-site. Particularly, the tasks related to interpersonal relations and communication seem to be more difficult when managing teams remotely.

Notably, "having fun together" shows the most significant difference, with 38.9% of leaders finding it challenging remotely, versus 16.9% on-site, suggesting that creating an enjoyable work atmosphere is much tougher when the team is not physically together. Close behind are challenges such as noticing when someone needs support and understanding how much employees are working, both with over 40% finding it difficult remotely, compared to 22% on-site.

"Engaging in spontaneous communication" and "noticing employees' coordination within the team" are also substantially higher in the remote setting, both exceeding a 20% difference when compared to on-site, highlighting that casual interactions and observing team dynamics are less apparent when working remotely.

Specific Leadership Challenges - I find it difficult to...

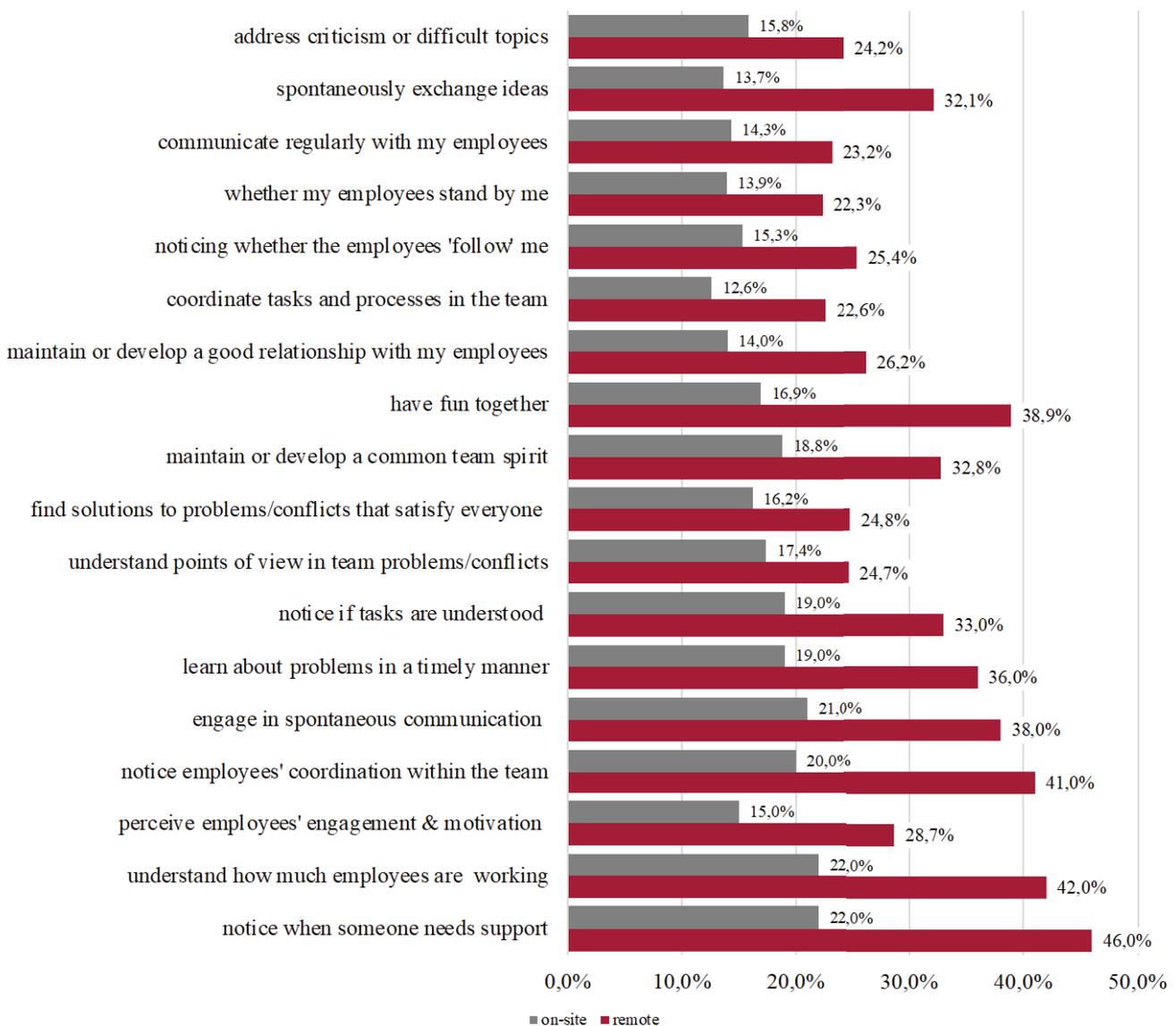


FIGURE 5: SPECIFIC LEADERSHIP CHALLENGES

In summary, Figure 5 indicates that remote leadership challenges are more pronounced in areas involving direct interaction, observation, and the informal aspects of team dynamics.

#### IV. DISCUSSION

This paper analysed the utility and effectiveness of leadership in remote and hybrid work settings. We found that hybrid and remote employees reported more transformational leadership behaviours from their leaders than office-based employees. Even though hybrid employees perceived transformational leadership as stronger in-office when directly comparing both settings. Hybrid employees generally viewed leadership more favourably in both settings than their office-only counterparts. The effectiveness of transformational leadership for engagement, well-being, and job satisfaction was positive in remote settings, but was found to be lower than in traditional office environments. Health-oriented leadership similarly had beneficial impacts on employees' health when working remotely, yet it was also perceived as more effective in the office.

Leaders who had implemented health-oriented leadership in remote and hybrid settings also reported experiencing personal health benefits for themselves. However, hybrid leaders also reported facing more significant challenges in remote contexts as compared to the office setting, especially in areas such as creating an enjoyable atmosphere, engaging in spontaneous communication, and noticing employees' needs. We conclude that while leadership quality was seen positively in remote settings with benefits stemming from transformational and health-oriented approaches, the lack of F2F interaction in remote work introduces specific challenges for leaders.

##### A. Implications for Leadership Practice

The aim of this paper was to provide leaders and organisations with specific recommendations to enhance work and well-being in remote and hybrid work arrangements. In the following, we provide guidelines for leaders to improve their leadership experience, their own challenges and their followers' benefit from transformational and health-oriented leadership.

###### 1) *Adaptation of Leaders' Communication Strategies for Effective Transformational and Health-Oriented Leadership*

Our findings suggest that leadership behaviours can be effective in remote settings, but leaders may need to adapt their communication strategies to bridge the gap where face-to-face interaction is limited. For instance, we recommend leaders to enhance their digital communication skills to effectively convey motivation and maintain team spirit when not physically present. Emphasizing clear, rich and structured communication can help mitigate the perceived reduction in leadership quality in remote settings. This could be achieved via structured weekly team meetings, bi-weekly 1:1 meetings with each team member and by supporting informal communication (via monthly knowledge sessions or bi-weekly coffee chats).

###### 2) *Emphasizing Synchronous Communication*

The effectiveness of health-oriented leadership is enhanced by synchronous communication (e.g. video calls or phone calls rather than e-mails). Leaders should strive to maintain regular live interactions with their teams to foster a sense of immediacy and connection, which is crucial for both observing and addressing team dynamics and individual needs

effectively. We recommend rich communication channels (videocalls) as they are better in helping leaders detect nuances in their employees' behaviours or potential health-related warning signals [31]. Klebe & colleagues [32] investigated the role of the effectiveness of HoL in digital vs. face-to-face contexts and found slightly weaker effects of HoL on engagement and satisfaction. This provides further evidence to emphasize synchronous and also face-to-face communication.

###### 3) *Fostering Engagement and Team Spirit Remotely*

Activities designed to build team cohesion and maintain engagement need to be creatively adapted for remote contexts. Leaders should consider virtual team-building activities and informal virtual meetups to replicate spontaneous interactions that occur naturally in an office setting.

###### 4) *Actively Planning & Designing Hybrid Work*

In hybrid contexts where leaders and employees have some office and some remote days, we recommend leaders to collaboratively design office and remote schedules for the team. This can ensure some face-to-face interaction if all team members commit to regular meetups at the office.

###### 5) *Proactive Role of Organisations*

Our focus on leadership within this paper is not intended to force full responsibility on leaders in creating functional and healthy remote and hybrid work environments. Instead, we want to put an emphasis on organisations to reflect upon current challenges faced by leaders and their ability to provide resources and guidance (i.e. leadership training), equipment (i.e. software, hardware, stable connection & sufficient bandwidth) and overall positive working conditions [46] which enable leaders to actively support in the design of remote and hybrid work practices without putting the sole responsibilities on them.

As part of our project, Digital Leadership and Health (DLaH), we are currently developing an online assistance platform (DigiLAP) for leaders which provides in-depth trainings for risks and benefits, leadership, digital communication, instant feedback and more relevant topics from the realm of remote and hybrid leadership.

##### B. Implications for Further Research

Further research is needed to examine antecedents and outcomes of remote and hybrid leadership practices. A detailed examination of short- and long-term effects via diary studies and larger longitudinal studies is called for.

We also recommend further experimental research to investigate the effects of leadership behaviours across different communication channels using varying technologies and communication types. Additionally, research needs to take into account new emerging technologies, like Virtual Reality and Augmented Reality and test the utility of these technologies for leadership and work scenarios.

Future studies should also put an emphasis on organisational level variables like the culture surrounding remote and hybrid work, organisational resources for remote work and human resource management to identify antecedents of successful remote work on the organisational level.

Further, boundary conditions regarding the effectiveness of transformational leadership and health-oriented leadership need to be further investigated. Our research demonstrated that leadership effectiveness is susceptible to the type of communication (synchronous vs. asynchronous) which is supported

by the findings of Schübbe and colleagues (under review [46]). However, there might be further boundary conditions such as relationship quality or tenure of relationship between employees and their leader that have an impact on how effective remote leadership is. These need to be further examined.

In addition, we call for practice-oriented papers like this one, which reports research findings in an understandable manner, thus serving as an accessible resource for leaders and practitioners in organisations (e.g. human resource departments) who want to access high-quality information but are not equipped to understand high-level methodology.

#### ACKNOWLEDGEMENTS

This research project “Digital Leadership and Health” is funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

The authors would like to express their thanks to their current student assistants, Mr. Johannes Damian Eiden, B.Sc., Ms. Jana Kiehn, B. Sc. & Mr. Thies Knudsen, B. Sc., for assistance in literature review.

#### REFERENCES

- [1] A. Krick, J. Felfe, S. Hauff, and K.-H. Renner, “Facilitating Health-Oriented Leadership from a Leader’s Perspective: Antecedents at the Organizational, Workplace, and Individual Level,” *Zeitschrift für Arbeits- und Organisationspsychologie*, vol. 66, no. 4, pp. 213–225, 2022, doi: 10.1026/0932-4089/a000397.
- [2] T. Kortsch, R. Rehwaldt, M. E. Schwake and C. Licari, “Does remote work make people happy? Effects of flexibilization of work location and working hours on happiness at work and affective commitment in the German banking sector”, *International Journal of Environmental Research and Public Health*, vol. 19, no 15, pp. 9117, 2022, doi: 10.3390/ijerph19094955.
- [3] S. R. Sardeshmukh, M. Goldsby, and R. M. Smith, “Are work stressors and emotional exhaustion driving exit intentions among business owners?,” *Journal of Small Business Management*, vol. 59, no. 4, pp. 544–574, Jan. 2020, doi: 10.1111/jsbm.12477.
- [4] M. Lange and I. Kayser, “The Role of Self-Efficacy, Work-Related Autonomy and Work-Family Conflict on Employee’s Stress Level during Home-Based Remote Work in Germany,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 9, p. 4955, Apr. 2022, doi: 10.3390/ijerph19094955.
- [5] A: Krick and J. Felfe, “Health-oriented leadership in a digital world: a literature review and a research agenda”, in: *dtecbw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr*, D. Schulz, A. Fay, W. Matiaske and M. Schulz, Eds. Hamburg: Helmut-Schmidt-Universität Hamburg, 2022. pp. 347–357.
- [6] I. Efimov, V. Harth, and S. Mache, “Health-Oriented Self- and Employee Leadership in Virtual Teams: A Qualitative Study with Virtual Leaders,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, p. 6519, Sep. 2020, doi: 10.3390/ijerph17186519.
- [7] S. Chatterjee, R. Chaudhuri, and D. Vrontis, “Does remote work flexibility enhance organization performance? Moderating role of organization policy and top management support,” *Journal of Business Research*, vol. 139, pp. 1501–1512, Feb. 2022, doi: 10.1016/j.jbusres.2021.10.069.
- [8] C. Patel, M. Biron, S. C. Cooper, and P. S. Budhwar, “Sick and working: Current challenges and emerging directions for future presenteeism research,” *Journal of Organizational Behavior*, vol. 44, no. 6, pp. 839–852, Jul. 2023, doi: 10.1002/job.2727.
- [9] L. Petitta and V. Ghezzi, “Remote, Disconnected, or Detached? Examining the Effects of Psychological Disconnectedness and Cynicism on Employee Performance, Wellbeing, and Work–Family Interface,” *International Journal of Environmental Research and Public Health*, vol. 20, no. 13, p. 6318, Jul. 2023, doi: 10.3390/ijerph20136318.
- [10] D. C. Tautz, K. Schübbe, and J. Felfe, “Working from home and its challenges for transformational and health-oriented leadership,” *Frontiers in Psychology*, vol. 13, Dec. 2022, doi: 10.3389/fpsyg.2022.1017316.
- [11] K. Kirchner, C. Ipsen, and J. P. Hansen, “COVID-19 leadership challenges in knowledge work,” *Knowledge Management Research & Practice*, vol. 19, no. 4, pp. 493–500, Feb. 2021, doi: 10.1080/14778238.2021.1877579.
- [12] B. M. Bass and R. E. Riggio, *Transformational Leadership*. Psychology Press, 2006. doi: 10.4324/9781410617095.
- [13] K. A. Arnold, “Transformational leadership and employee psychological well-being: A review and directions for future research.,” *Journal of Occupational Health Psychology*, vol. 22, no. 3, pp. 381–393, Jul. 2017, doi: 10.1037/ocp0000062.
- [14] F. Franke and J. Felfe, “How does transformational leadership impact employees’ psychological strain?,” *Leadership*, vol. 7, no. 3, pp. 295–316, Aug. 2011, doi: 10.1177/1742715011407387.
- [15] G. Wang, I.-S. Oh, S. H. Courtright and A. E. Colbert, “Transformational leadership and 846 performance across criteria and levels: A meta-analytic review of 25 years of research”, *Group & Organization Management*, vol. 36, no. 2, pp. 223–270, 2011.
- [16] C. Deng, D. Gulseren, C. Isola, K. Grocutt, and N. Turner, “Transformational leadership effectiveness: an evidence-based primer,” *Human Resource Development International*, vol. 26, no. 5, pp. 627–641, Oct. 2022, doi: 10.1080/13678868.2022.2135938.
- [17] R. K. Purvanova and J. E. Bono, “Transformational leadership in context: Face-to-face and virtual teams,” *The Leadership Quarterly*, vol. 20, no. 3, pp. 343–357, Jun. 2009, doi: 10.1016/j.leaqua.2009.03.004.
- [18] J. Eisenberg, C. Post, and N. DiTomaso, “Team Dispersion and Performance: The Role of Team Communication and Transformational Leadership,” *Small Group Research*, vol. 50, no. 3, pp. 348–380, Feb. 2019, doi: 10.1177/1046496419827376.
- [19] J. E. Hoch and S. W. J. Kozlowski, “Leading virtual teams: Hierarchical leadership, structural supports, and shared team leadership.,” *Journal of Applied Psychology*, vol. 99, no. 3, pp. 390–403, 2014, doi: 10.1037/a0030264.
- [20] F. Franke, J. Felfe, and A. Pundt, “The impact of health-oriented leadership on follower health: Development and test of a new instrument measuring health-promoting leadership,” *Zeitschrift für Personalforschung*, vol. 28, 1-2, pp. 139–161, 2014.
- [21] S. der Kinderen, A. Valk, S. N. Khapova, and M. Tims, “Facilitating Eudaimonic Well-Being in Mental Health Care Organizations: The Role of Servant Leadership and Workplace Civility Climate,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 4, p. 1173, Feb. 2020, doi: 10.3390/ijerph17041173.
- [22] M. Arnold and T. Rigotti, “How’s the boss? Integration of the health-oriented leadership concept into the job demands-resources theory,” *Journal of Managerial Psychology*, vol. 38, no. 6, pp. 419–433, Aug. 2023, doi: 10.1108/jmp-01-2023-0030.
- [23] A. Santa Maria, C. Wolter, B. Gusy, D. Kleiber, and B. Renneberg, “The Impact of Health-Oriented Leadership on Police Officers’ Physical Health, Burnout, Depression and Well-Being,” *Policing: A Journal of Policy and Practice*, vol. 13, no. 2, pp. 186–200, Sep. 2018, doi: 10.1093/policing/pay067.
- [24] F. Teetzen, K. Klug, H. Steinmetz, and S. Gregersen, “Organizational health climate as a precondition for health-oriented leadership: expanding the link between leadership and employee well-being,” *Frontiers in Psychology*, vol. 14, Jun. 2023, doi: 10.3389/fpsyg.2023.1181599.
- [25] L. A. Grimm, G. F. Bauer, and G. J. Jenny, “Is the health-awareness of leaders related to the working conditions, engagement, and exhaustion in their teams? A multi-level mediation study,” *BMC Public Health*, vol. 21, no. 1, Oct. 2021, doi: 10.1186/s12889-021-11985-1.
- [26] F. Franke and J. Felfe, “Diagnose gesundheitsförderlicher Führung – Das Instrument „Health-oriented Leadership“,“ *Fehlzeiten-Report* 2011, pp. 3–13, 2011, doi: 10.1007/978-3-642-21655-8\_1.

- [27] A. J. Kaluza, F. Weber, R. van Dick, and N. M. Junker, "When and how health-oriented leadership relates to employee well-being—The role of expectations, self-care, and LMX," *Journal of Applied Social Psychology*, vol. 51, no. 4, pp. 404–424, Mar. 2021, doi: 10.1111/jasp.12744.
- [28] F. Pundt and J. Felfe, *HOL. An instrument to assess health-oriented leadership*. Göttingen: Hogrefe, 2017.
- [29] S. Pischel and J. Felfe, "Should I Tell My Leader or Not?—Health-Oriented Leadership and Stigma as Antecedents of Employees' Mental Health Information Disclosure Intentions at Work," *Journal of Occupational & Environmental Medicine*, vol. 65, no. 1, pp. 74–85, Sep. 2022, doi: 10.1097/jom.0000000000002688.
- [30] S. Vincent-Höper and M. Stein, "The Role of Leaders in Designing Employees' Work Characteristics: Validation of the Health- and Development-Promoting Leadership Behavior Questionnaire," *Frontiers in Psychology*, vol. 10, May 2019, doi: 10.3389/fpsyg.2019.01049.
- [31] K. Schübbe, A. Krick and J. Felfe, „Wahrnehmung von Warnsignalen im digitalen Kontext: Eine experimentelle Simulation zu den Unterschieden von F2F und digitaler Kommunikation“, *Symposium 2: Neue Perspektiven für Health-oriented Leadership: Bedingungsfaktoren für gesundheitsorientierte Führung im traditionellen und digitalen Arbeitskontext*, 13. Fachgruppentagung der AOW- und Ingenieurpsychologie, Kassel, 14.09.2023.
- [32] L. Klebe, J. Felfe, A. Krick, and D. Tautz, "Who has the most to lose? How ICT demands undermine health-oriented leadership," *Applied Psychology*, 2024, doi: 10.1111/apps.12530.
- [33] S. M. Neidlinger, J. Felfe, and K. Schübbe, "Should I Stay or Should I Go (to the Office)?—Effects of Working from Home, Autonomy, and Core Self-Evaluations on Leader Health and Work–Life Balance," *International Journal of Environmental Research and Public Health*, vol. 20, no. 1, p. 6, Dec. 2022, doi: 10.3390/ijerph20010006.
- [34] B. Pangert, N. Pauls and H. Schübach, „Consequences of Permanent Availability on Life-Domain-Balance and Health“, *Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA)*: Dortmund, Germany, 2013. Available online: <https://www.researchgate.net/publication/266851161>
- [35] L. L. Ten Brummelhuis, J. M. Haar, and M. Roche, "Does Family Life Help to be a Better Leader? A Closer Look at Crossover Processes From Leaders to Followers," *Personnel Psychology*, vol. 67, no. 4, pp. 917–949, Oct. 2013, doi: 10.1111/peps.12057.
- [36] S. Baeriswyl, A. Krause, A. Elfering, and M. Berset, "How workload and coworker support relate to emotional exhaustion: The mediating role of sickness presenteeism," *International Journal of Stress Management*, vol. 24, no. Suppl 1, pp. 52–73, 2017, doi: 10.1037/str0000018.
- [37] C. Dietz, H. Zacher, T. Scheel, K. Otto, and T. Rigotti, "Leaders as role models: Effects of leader presenteeism on employee presenteeism and sick leave," *Work and Stress*, vol. 34, no. 3, pp. 300–322, Feb. 2020, doi: 10.1080/02678373.2020.1728420.
- [38] B. M. Bass and B. J. Avolio, "Multifactor Leadership Questionnaire", 1990, Palo Alto, CA: Consulting Psychologists Press.
- [39] J. Felfe and K. Goihl, "Deutscher Multifactor Leadership Questionnaire (MLQ)", 2002, <https://doi.org/10.6102/ZIS22>
- [40] W. B. Schaufeli, A. Shimazu, J. Hakanen, M. Salanova and H. de Witte, "An Ultra-Short Measure for Work Engagement", *European Journal of Psychological Assessment*, 2017, 35(4), 577–591, doi:10.1027/1015-5759/a000430
- [41] "WHO-Five Well-being Index," *PsycTESTS Dataset. American Psychological Association (APA)*, 1998. doi: 10.1037/t36093-000.
- [42] J. Felfe and F. Franke, "COMMIT Commitment Skalen–Fragebogen zur Erfassung von Commitment gegenüber Organisation, Beruf/Tätigkeit, Team, Führungskraft und Beschäftigungsform", Hans Huber, 2012.
- [43] M. Nübling, U. Stöbel, H. Hasselhorn, M. Michaelis, and F. Hofmann, "COPSOQ: Ein Instrument zur Messung psychischer Belastungen am Arbeitsplatz," *Das Gesundheitswesen*, vol. 67, no. 07, Nov. 2005, doi: 10.1055/s-2005-920554.
- [44] G. Mohr, T. Rigotti, and A. Müller, "Irritation - ein Instrument zur Erfassung psychischer Beanspruchung im Arbeitskontext. Skalen- und Itemparameter aus 15 Studien," *Zeitschrift für Arbeits- und Organisationspsychologie*, vol. 49, no. 1, pp. 44–48, Jan. 2005, doi: 10.1026/0932-4089.49.1.44.
- [45] G. Mohr, "The assessment of mental health impairments of industrial workers", Lang, 1986, Frankfurt am Main, Germany.
- [45] J. Felfe, A. Krick, S. Hauff, K.-H. Renner, L. Klebe, K. Schübbe, D. Tautz, Y. Frontzkowski, P. Gubernator and S. M. Neidlinger, "Working from home: Opportunities and risks for working conditions, leadership and health" in: *dtecbw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr*, D. Schulz, A. Fay, W. Matiaske and M. Schulz, Eds. Hamburg: Helmut-Schmidt-Universität Hamburg, 2022, pp. 335-341.
- [46] K. Schübbe, A. Krick and J. Felfe, "They seemed just fine - Unraveling Leaders' Perception of Followers' Warning Signals Across Communication Channels", under review.

# Are Videocalls Outdated? A Comparison with Virtual Reality Meetings as a Future Perspective

Katharina Schübbe<sup>1\*</sup>, Yannick Frontzkowski<sup>1</sup>, Philip Gubernator<sup>1</sup>, L. Franziska Münstermann<sup>1</sup>, Jörg Felfe<sup>1</sup>, Sven Hauff<sup>1</sup>, Karl-Heinz Renner<sup>2</sup> & Annika Krick<sup>1</sup>

<sup>1</sup>Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg, Hamburg, Germany

<sup>2</sup>Universität der Bundeswehr, Munich, Germany

\*schuebbk@hsu-hh.de

**Abstract** – The rise of remote collaboration tools is transforming team collaboration. Despite its current limitations, Virtual Reality (VR) technology has the potential to overcome the challenges of traditional video calls and enhance remote meetings in the future. This study compares the effectiveness of VR and video calls (MS Teams) on team collaboration. A total of  $N = 90$  participants were performing a problem-solving task, and  $N = 127$  were conducting a creative task in both conditions (VR and video call). Measures of meeting evaluation, engagement, performance, and fatigue were assessed. The results reveal that video call collaboration is superior to enhancing performance. In comparison, VR offers significant benefits regarding comfort, social interaction, and engagement, particularly with creativity tasks. However, at the current stage of technical development, using VR, regardless of the task, is exhausting and causes fatigue. This research highlights the potential of VR as an effective tool for remote collaboration, emphasizing the importance of selecting the right platform based on the specific needs of the collaborative context.

**Keywords** – Virtual reality, digital collaboration, performance, fatigue

## NOMENCLATURE

$M$	Mean
HDMs	Head-Mounted Displays
VR	Virtual Reality
WFH	Working-From-Home

## I. INTRODUCTION

The increase in remote collaboration tools has prompted a need to understand how different communication platforms impact team dynamics and task performance. While video call platforms like Microsoft Teams and Zoom have become predominant in enabling remote interactions, they also present distinct challenges that can influence the outcomes of team collaboration. For one, face-to-face communication has reduced social presence, partly due to the lack of spatial positioning [1]. Additionally, the ability to detect subtle changes in facial expressions and mood is significantly limited [2]. This results in both verbal and non-verbal communication being restricted [1].

While Virtual Reality (VR) technology still has technical limitations, it can potentially solve these challenges in the future and may overcome the usual issues found in video calls. This is because VR offers immersive experiences that more closely replicate face-to-face interactions, allowing for more

natural communication and enhancing the sense of presence and engagement, which is often lacking in traditional video calls [3]. Therefore, also fewer distractions and stimuli lead to increased productivity [4]. Consequently, it makes sense to assess now whether there is potential for this technology in the future.

Although practical applications of VR have been explored in various sectors (e.g. such as retail experiences and tourism) so far [5], there have only been a few first studies investigating VR as a potential tool for team collaboration by directly comparing VR with video calls [3], [6], [7], [8]. However, the superiority or advantage of one communication tool (VR vs. video call) remains unclear and requires further research [3].

Focusing on further research, the type of task (task dependency) could be crucial in determining the effectiveness. Each task is unique, and comparing various tasks within each collaboration tool (VR vs. video call) may produce different and individual outcomes [9]. Due to the individuality of each task, they may impact fatigue differently. This calls for a better understanding of how different tasks influence outcomes such as fatigue but also engagement and performance [3], [8]. Also, understanding more about the mechanisms behind fatigue and exhaustion from using VR headsets, known in the literature as head-mounted displays (HDMs), is particularly important [3].

Therefore, this study aims to compare the efficacy of VR and video calls in supporting team collaboration across different types of tasks (problem-solving vs. creative) in a controlled lab setting. Problem-solving tasks require precise problem-solving and attention to detail, while creative tasks involve imagination, social stimulation, and brainstorming. Collaboration and interaction are essential in both tasks. By examining both subjective data (such as meeting evaluation and engagement) and objective data (including performance measures), we aim to explore how tasks differ between VR and video call settings.

This research contributes to the growing literature examining and comparing video and VR communication platforms [3]. Additionally, this initial study lays the foundation for further research that will delve into more detailed outcomes such as well-being, trust, and multitasking.

As teams increasingly rely on digital tools for complex tasks, leaders must identify each platform's strengths and limitations to optimize team synergy and productivity. Specifically, leaders must decide when to use each medium and be aware of potential challenges and risks.

## II. WHAT DO WE KNOW ABOUT VIDEO CALLS AND VR?

### A. Common Challenges in Video Call Conferences

Video platforms like Microsoft Teams (MS Teams), Webex, and Zoom are the go-to solutions, allowing team members to connect via video calls. These calls enable real-time visual and audio communication, offering the convenience of meeting from any location with an internet connection, thus saving time and resources otherwise spent on travel [10]. However, video calls also come with challenges, such as potential connectivity issues, the difficulty of engaging participants in the same way as in-person interactions, and the now well-recognized phenomenon of Zoom or video fatigue [1]. Furthermore, video calls can often limit the richness of communication to what is captured by the camera, potentially missing non-verbal cues [2].

These disadvantages present significant challenges for leaders, making it difficult to build and maintain relationships in a purely digital environment. The absence of spontaneous and informal interactions, standard face-to-face meetings, and limited verbal and nonverbal communication options can complicate communication. The lack of direct eye contact and body language further leads to misunderstandings and makes it particularly challenging for leaders to gauge moods and reactions accurately [11, 12].

### B. The Future Potential of Virtual Reality

The Metaverse, an innovative computer-mediated environment characterized by virtual, real-time interactions through avatars and interface technologies like VR headsets and glasses, represents a further evolution in digital communication [3]. VR provides an alternative to face-to-face communication by closely replicating the dynamics of in-person interactions. Compared to video calls, it offers high communication quality through avatars and other non-verbal cues [13].

Therefore, the 3D environment of VR allows for more natural and intuitive interactions, going beyond the flat screens of traditional video calls. Instead, VR creates a sense of presence, making users feel like they are indeed in the virtual environment and part of it. High-quality VR hardware allows users to perceive and translate their head movements, and sometimes body movements, into the virtual environment onto an avatar, enhancing this feeling of presence. The sense of social presence even exists when an avatar invades the personal space of another in the virtual world, prompting defensive reactions similar to those in the real world [6], [14], [15], [16].

Avatars enhance the sense of social presence, which can improve interaction despite challenges such as limited facial expressions and body movement inaccuracies. However, non-verbal cues in VR, including gestures, proximity, and haptic feedback, contribute to the realism of interactions, effectively mimicking real-life social dynamics. Therefore, facial expressions and lip-syncing of an avatar are crucial, significantly affecting user experiences by enhancing social presence and emotional perception. This has been found in the first study results showing that positive facial expressions in avatars can increase user comfort and encourage collaboration [13], [17].

While VR headsets (HDMs) offer an immersive experience, they also have their challenges. Non-verbal facial expressions are still largely absent in current VR technology. This limitation hinders the full range of human interactions in virtual environments, as facial cues play a crucial role in effective communication and emotional connection [18]. In the

future, advancements in VR technology will likely address the challenge of incorporating non-verbal facial expressions. Innovations such as improved facial tracking sensors and more sophisticated avatar animation software are expected to enable more accurate and real-time rendering of facial expressions, enhancing the realism and emotional depth of virtual interactions [19].

VR glasses are often heavy and can cause discomfort during extended use, which is particularly problematic given the high cost of quality systems. Technical issues such as latency, insufficient frame rates, and resolution can also impair the user experience. Additionally, the range of VR content is still limited, restricting the technology's applications and making it less appealing to the broader market [20].

While Virtual Reality (VR) technology currently has technical limitations, it can potentially address and possibly overcome the common challenges associated with video calls via platforms like MS Teams or Zoom. Alongside Augmented Reality (AR), VR offers a promising future alternative that can circumvent many issues inherent in traditional video conferencing. Currently, VR technology utilizes avatars that are capable of detecting and representing both hand gestures and the orientation toward conversation partners, enhancing the interaction quality significantly [13]. By now, VR systems have been explored across multiple domains, including clinical treatment [21], [22], education [23], [24], training [25] and creative collaboration [26]. However, their use for digital team collaboration in the workplace is insufficiently researched.

### C. Digital Team Collaboration in VR

Team members can use VR to extend their working-from-home (WFH) space into a virtual, three-dimensional office and collaborate on tasks. For instance, while traditional video calls are so far favoured for work tasks due to easier information exchange and note-taking, VR is preferred for team-building activities [6]. VR exhibits the most similarities to face-to-face interactions regarding spatial behaviour, actions, and facial expressions. Additionally, VR appears to have a positive effect on users' mood [13]. However, the high level of focus required in VR comes at a cost. Many people find spending time in VR demanding, mainly due to its newness and the pressure perception from the headset [14]. As users become more accustomed to VR, preferences may shift in the future [27]. In the context of workplace collaboration, a few types of tasks have been compared, but no definitive superior medium has been identified [3].

In addition with a few exceptions, there is a scarcity of experimental studies that analyse various dimensions of collaboration, such as performance, engagement, and fatigue, under controlled conditions [6], [7], [8], [9], [12], [28].

A comprehensive review of past research by Wei and colleagues [13] provides an overview of previous findings. They call for a deeper examination of communication processes, e.g. display of emotions and non-verbal cues. Although initial studies have explored the use of VR compared to video conferencing, the findings are inconsistent and often limited to specific use cases or short durations. Previous research has mainly focused on both general usage aspects and subjective perceptions of VR [3], [13]. For example, Hennig-Thurau and colleagues [3] focused on various tasks including brainstorming, product rating, presentation, and customer feedback. They found greater social presence and more exhaustion in VR in

most tasks, but no overall superiority of VR with regard to performance. They emphasize that future studies should pay more attention to individual differences and control individual variables.

To further investigate the difference between VR and video calls, we build upon the work of Hennig-Thurau and colleagues [3] by using different tasks and a modified research design: 1) To better control for potential confounding third variables, such as individual differences, motivation, or mental state, we used a within-subjects design. This procedure not only increases the statistical power but also controls the variability between the participants. 2) The serial processing of multiple tasks could have blurred the effects of individual tasks and the overall effect of collaboration tools. To address this, we focus on one task per study. 3) Generally, the type of task may cause different communication requirements. We assume that more complex communication requirements can be better managed with VR. Therefore, we used a more complex problem-solving task with distributed information to enhance interdependency and provoke more interaction and discussion among group members.

To sum up, we aim to directly compare VR and video call tools within the same individuals using a within-subjects design. Additionally, we conduct two separate studies with independent samples for each task. We will measure both subjective and objective outcomes to provide a comprehensive understanding of how these communication platforms influence collaboration.

Due to its advantages, VR might lead meeting participants to perceive meetings held in the VR environment as more favourable compared to video conference calls [29]. So far, few studies have fundamentally examined the direct comparison of media for collaboration (focusing on performance), so no superior tool has yet been identified. The lack of differences observed in previous studies could also be due to the lack of suitable tasks for comparison [30]. Therefore, other researchers call for further research to understand the barriers to adapting VR as one support for future remote work and to consider the impact of task types in VR [31].

In our studies, we aim to compare the conditions (VR vs. video call) and additionally examine how these differ depending on the task type (a problem-solving task in Study 1, a creative task in Study 2), as this could also be relevant and be a reason for inconsistent findings. Objective outcomes should be considered in addition to participants' subjective evaluations. Specifically, engagement, performance, and fatigue are key factors relevant to assessing the overall effectiveness of the interaction methods.

#### *D. What happens to Engagement and Performance?*

While VR might improve the "fun", "presence", and "naturalness" of interactions in collaborative meetings [29], it does not necessarily boost team effectiveness in terms of engagement and performance.

VR environments can minimize external distractions common in a home or office setting, potentially leading to more focused meetings. At the same time, VR might reduce the multitasking often observed during digital meetings via videoconferencing. This focused environment might help participants concentrate more on the meeting content as in traditional face-to-face settings, minimizing the distractions typically associated with other digital platforms [12], [32].

Furthermore, in previous experiments, participants experienced higher enjoyment, immersion, and temporal dissociation in VR, suggesting that VR can induce a more pronounced sense of flow [26], [33]. VR can have a positive impact on the flow experience and can demonstrate a mediating effect explained through motivation, curiosity, and cognitive benefits [34], it might also be positive for engagement in meetings [35].

VR might enhance their performance if participants feel more engaged in VR meetings. One advantage is that it works across multiple virtual synchronized screens and applications. This integration of software and VR headsets offers significant benefits for academic and work-related tasks [4]. Moreover, the first experimental results show that VR enhances team performance in logical tasks [36] but also for the creative process, though its impact on variety and novelty remains unclear [37]. Therefore, we believe VR could enhance participants' engagement and performance.

#### *E. Which collaboration tool is more exhausting?*

Zoom or video fatigue is the phenomenon of exhaustion known from the increased number of online meetings during the COVID-19 pandemic.

It describes the fatigue resulting from extensive use of virtual communication platforms, where symptoms can range from physical tiredness to psychological effects such as reduced concentration and heightened feelings of social disconnect [1], [38], [39]. This type of fatigue has also been observed in three-dimensional spaces, as evidenced by initial experiments [3].

Next to general fatigue, there is also visual fatigue, which often appears in digital contexts. Visual fatigue, usually called eyestrain, is discomfort or impairment arising from extended eye use. It is characterized by symptoms such as tiredness, burning eyes, blurred vision, and headaches. These symptoms are exacerbated by prolonged visual tasks such as staring at computer screens, reading, or driving [40], [41]. Thus, immersive interactions can also result in increased exhaustion.

Additionally, "Cybersickness" can occur in VR. This refers to a condition when exposure to virtual environments causes symptoms like motion sickness. These can include dizziness, nausea, and disorientation. Cybersickness is often triggered by a disconnect between what the user sees and what their body feels, especially in immersive VR settings [40], [42].

Other studies suggest that VR can enhance well-being in the workplace and serve as a tool for stress prevention and management [5], [43], [44]. However, given the multitude of findings and first results from prior studies [3], we expect that VR, due to or precisely because of its immersive nature, leads to greater fatigue in collaboration.

In summary, few comparative studies to date indicate a research gap for a small, controlled, artificial experimental setting. This setting is needed to compare collaboration across both mediums and tasks regarding outcomes.

### III. METHODS

To evaluate how collaboration in VR and video conferences differ in terms of subjective meeting assessments, engagement, performance, and fatigue, we compare these aspects in both conditions (VR vs. video call) and also in different types of tasks. Specifically, we planned to compare problem-solving

tasks (Study 1) and creative tasks (Study 2) to see how each medium supports or hinders these distinct activities.

#### A. Design

The experimental setup of both studies is structured as a 2x2 within design, aiming to directly compare two different mediums: virtual reality (VR) and video conference calls. These studies involve two distinct types of team-based tasks - one problem-solving (Study 1) and the other a creative (Study 2) task. Each session involves assembled teams of three participants.

#### B. Procedure and Materials

Both studies have the same procedure. Participants were randomly assigned to groups for one task on MS Teams or VR (via MetaQuest in WorkRooms Horizons). After performing their first task, they switched platforms for the other task.

#### C. Samples

The first sample for the problem-solving task (Study 1) consisted of  $N = 90$  participants, and the second sample for the creative task (Study 2) contained  $N = 127$  participants.

#### D. Outcome Measures

We examined how participants subjectively rated their collaborative experience after each meeting. We assessed six facets of the subjective collaboration evaluation regarding simplicity, productivity, enjoyment, comfort, sociability, and better-liked medium. Furthermore, we employed the scale developed by Schaufeli et al. to measure engagement during each meeting [45]. Performance evaluation was based on objective criteria. We matched the number of correctly and incorrectly placed seatings for the problem-solving task. For creative performance, we examined the relative frequency of each solution and calculated a score reflecting an overall performance. Lastly, to assess general and visual fatigue, we used the scale from Fauville et al. [46].

### IV. RESULTS

#### A. Evaluation of Meeting in Video Calls vs. VR

In evaluating meetings for the problem-solving task (Study 1), differences emerged across all dimensions between video calls and VR platforms. Participants were asked which medium they preferred, referring to different aspects of the medium. Video calls were rated higher in terms of 'simplicity' (video calls: 62.22%, VR: 37.78%) and 'productivity' (video calls: 64.44%, VR: 35.56%). Conversely, VR outperformed video calls in being 'more enjoyable' (video calls: 13.33%, VR: 86.67%), making participants feel 'more comfortable' (video calls: 33.33%, VR: 66.67%), being 'more sociable' (video calls: 24.44%, VR: 75.56%) and providing a 'the more liked medium' (video calls: 25.55%, VR: 74.45%).

Similar results can be found for the creative task (Study 2). The meeting evaluation showed differences across all dimensions. Video calls were rated significantly higher for 'simplicity' (video calls: 52.38%, VR: 47.62%) and 'productivity' (video calls: 53.96%, VR: 46.04%). On the other hand, VR received higher ratings for making participants feel more 'enjoyment' (video calls: 4.72%, VR: 95.28%), 'more comfortable' (video calls: 27.55%, VR: 72.45%) and for being 'more sociable' (video calls: 19.68%, VR: 80.32%) and providing a 'better liked medium' (video calls: 22.41%, VR: 77.59%).

#### B. Engagement and Performance in Video Calls vs. VR

For the problem-solving task (Study 1), when looking at meeting engagement, the difference between video calls ( $M = 3.25$ ) and VR ( $M = 3.31$ ) was not significant. On the other hand, for the creative task (Study 2), there was a significant difference in meeting engagement between the two platforms. VR saw higher engagement ( $M = 3.55$ ) compared to video calls ( $M = 3.13$ ),  $t(126) = 4.537$ ,  $p < 0.001$ ; indicating that the immersive qualities of VR may enhance participant involvement and engagement during creative, collaborative tasks.

The data for performance in Study 1 shows significant differences between video calls and VR in problem-solving tasks. Teams using video calls scored significantly higher on average ( $M = 23$ ) compared to those using VR ( $M = 19.23$ ), suggesting slightly better performance via video calls ( $t(82) = 4.273$ ,  $p < .001$ ). The number of errors was similar between the two platforms (video call:  $M = 2.07$ ; VR:  $M = 2.23$ ). Furthermore, in Study 2, for the creative task, the overall performance was significantly higher in video calls ( $M = 42.65$ ) compared to VR ( $M = 37.33$ ),  $t(114) = 2.251$ ,  $p = .026$ .

#### C. Fatigue in Video Calls vs. VR

In the problem-solving task (Study 1), comparing fatigue levels between video calls and VR revealed significant differences. General fatigue was significantly higher when meeting in VR ( $M = 2.45$ ) compared to video calls ( $M = 1.97$ );  $t(88) = 2.536$ ,  $p = .021$ . For the creative task (Study 2), the only significant difference was noted in the subscale visual exhaustion, where VR yielded a higher mean value ( $M = 1.94$ ) compared to video calls ( $M = 1.63$ );  $t(125) = -2.465$ ,  $p = .028$ . The general scale also showed slightly higher means for VR but without significant differences from video calls.

### V. DISCUSSION

The first conclusions emerge from this comparative experiment of VR and video calls as communication platforms for digital collaboration.

Independent of the task type (problem-solving vs. creative task), participants' subjective evaluations of collaboration revealed significant differences between the platforms (VR vs. video call). They favoured video calls for their simplicity and productivity, likely because they integrate more seamlessly into prior tool experience, established workflows, and mimic the structure of traditional work environments [6]. However, VR as a collaboration tool was preferred for comfort and social interaction, suggesting that its enveloping nature can foster a stronger sense of comfort and social connectedness, enhancing collaborative experiences in a digital working context [6].

This positive evaluation of VR meetings is also reflected in participants' engagement. Engagement was notably higher in VR, especially during creative tasks, likely due to its interactive and immersive nature [6], [47]. This distinction was not as pronounced as in problem-solving tasks, where engagement levels were comparable in both platforms (VR and video call). This suggests that task nature may influence how platform choice impacts user engagement.

Conversely, while VR enhances engagement, it does not have the same positive results in performance outcomes. Video calls facilitated higher achievement, which could be attributed to users' greater familiarity with the user-friendly interface designed for direct, goal-oriented tasks [47]. However,

it could also simply be due to the technology, which still needs further improvement. Upgrades like the Meta Quest 3 could potentially yield different results [6].

Furthermore, both tasks linked VR to greater general and visual fatigue. This may be attributed to participants' limited prior experience with the technology, exposing them to various novel stimuli [48], [49]. The immersive nature of VR demands more cognitive and visual effort, which can contribute to fatigue and current technical limitations [42], [50]. Additionally, the physical discomfort from wearing VR headsets for extended periods can further exacerbate fatigue [51]. These factors collectively might contribute to fatigue and, consequently, negatively influence performance outcomes.

Comparing our findings with the previous research of Hennig-Thurau and colleagues [3], our results confirm that participants experience more fatigue after meetings in VR. In line with Hennig-Thurau and colleagues [3], who found higher levels of social presence and positive emotions with VR, our studies also showed that participants experienced more sociability, enjoyment, and comfort in VR in both tasks.

However, our findings differed regarding the performance. In contrast to Hennig-Thurau and colleagues [3], in our studies performance was better in video calls in both tasks. Nonetheless, this could be due to different measures of performance. Whereas Hennig-Thurau and colleagues [3] used the amount of interactions in some tasks we used an objective indicator.

In our two studies, we also measured engagement, which had previously only been indirectly assessed through social presence and emotional connection, suggesting higher engagement [3]. Our findings showed that engagement was higher in the creativity task with VR but not in the problem-solving task. However, differences in engagement in the problem-solving task point in the expected direction.

Although we did not find a clear superiority of one tool over the other in terms of collaboration, the subjective evaluation of meetings in VR was more positive. Consequently, the results slightly favour VR, at least for tasks that require a high degree of interaction, such as creative tasks.

#### *A. Strengths and Limitations, and Future Directions*

It is important to consider this study's strengths and limitations. First, previous findings suggest that straightforward gameplay is especially attractive to first-time VR users [52]. The main issue is that it is unclear whether these positive results are due to a short-term effect or if they persist over a more extended period [53]. Currently, no studies examine the use of VR across multiple measurement points and over an extended period. Longitudinal studies would be critical in ruling out potential habituation effects.

Second, the mean age of the sample in this study was notably low. Considering that the employment rate is between the ages of 30 and 55, an older sample might exhibit a lower degree of familiarity with new technologies such as VR. This factor could potentially influence the outcomes of studies involving digital tools, as older participants may require more time to adapt to these technologies [54]. Future studies should consider extending the sample and allocating time for each task to give participants more opportunities to get accustomed to the medium.

Third, future studies should utilize the newer VR technology (e.g. Meta 3). This could resolve technical issues such as

the pixelated recognition of keyboards through the VR headset (HDMs). Additionally, the improved hardware and software capabilities of Meta 3 might enhance overall user experience and interaction quality, providing more reliable data on the effectiveness of VR in collaborative settings. These initial insights into problem-solving and creative tasks should be expanded with further studies involving different types of collaborative work in the workplace context [55]. Therefore, it would also be interesting to explore how these tasks impact not only engagement and performance but also social involvement and team cohesion.

Lastly, the intersection of VR and leadership has been scarcely explored. While there are preliminary first approaches to using VR for leadership assessment [56], [57] and development [58], there is no research on VR's effectiveness as a leadership tool. Future research should focus on digital interactions, examining aspects such as communication quality, social proximity, and familiarity to determine VR's potential as a leadership instrument.

#### *B. Practical Implications*

VR could enhance flexible work models and simplify digital collaboration across physical distances. These initial findings illustrate how different collaboration platforms can positively impact participants' engagement, potentially increasing social involvement and team cohesion. Using VR for collaborative tasks might simulate real-life interactions more effectively than traditional video calls.

The introduction of VR could offer employees the opportunity to be part of a modern and motivated workplace. However, it is crucial that they receive adequate training and support to adapt to the new technology. An open dialogue between employees and leaders is essential to ensure that VR technology is used in a way that promotes their well-being, productivity, and long-term engagement.

For companies, investing in VR technologies could be a strategic decision. Beyond enhancing digital meetings for collaboration, VR could also serve as a tool for professional development and team-building initiatives.

## VI. CONCLUSION

Our studies reveal that video calls are currently superior in digital collaboration when focusing solely on performance. However, VR offers significant benefits in comfort, social interaction, and engagement, particularly for creative tasks. At this stage of technical development, VR use remains exhausting and causes fatigue due to technical limitations, which may change in the coming years. Therefore, this research highlights the potential of VR as an effective tool for digital team collaboration.

#### ACKNOWLEDGEMENT

Yannick Frontzkowski and Philip Gubernator developed this experiment and fully supervised its execution.

The authors acknowledge funding from dtec.bw – Centre for Digitalization and Technology Research of the Bundeswehr. The dtec.bw is a scientific centre jointly supported by the Bundeswehr Universities of Hamburg and Munich and is part of the German Federal Government's economic stimulus program to overcome the COVID-19 crisis. It operates under academic self-administration. The funds provided to dtec.bw are

used at both Bundeswehr universities to finance research projects and knowledge and technology transfer projects. dtec.bw is funded by the European Union – NextGenerationEU.

## REFERENCES

- [1] J. N. Bailenson, "Nonverbal overload: A theoretical argument for the causes of Zoom fatigue," *Technology, Mind, and Behavior*, vol. 2, no. 1, 2021, doi: 10.1037/tmb0000030.
- [2] B. Matthews, Z. S. See, and J. Day, "Crisis and extended realities: remote presence in the time of COVID-19," *Media International Australia*, vol. 178, no. 1, pp. 198–209, 2021, doi: 10.1177/1329878X20967165.
- [3] T. Hennig-Thurau, D. N. Aliman, A. M. Herting, G. P. Cziehso, M. Linder, and R. V. Kübler, "Social interactions in the metaverse: Framework, initial evidence, and research roadmap," *J. of the Acad. Mark. Sci.*, vol. 51, no. 4, pp. 889–913, 2023, doi: 10.1007/s11747-022-00908-0.
- [4] L. Herrera et al., "Methodology for multitasking using virtual reality devices," in *International Conference On Signal Processing & Communication Engineering Systems: Spaces-2021*, Andhra Pradesh, India, 2024, p. 20026, doi: 10.1063/5.0117802.
- [5] L. Bodet-Contentin, M. Letourneur, and S. Ehrmann, "Virtual reality during work breaks to reduce fatigue of intensive unit caregivers: A crossover, pilot, randomised trial," *Australian critical care : official journal of the Confederation of Australian Critical Care Nurses*, early access. doi: 10.1016/j.aucc.2022.01.009.
- [6] K. Abramczuk, Z. Bohdanowicz, B. Muczyński, K. H. Skorupska, and D. Cnotkowski, "Meet me in VR! Can VR space help remote teams connect: A seven-week study with Horizon Workrooms," *International Journal of Human-Computer Studies*, vol. 179, p. 103104, 2023, doi: 10.1016/j.ijhcs.2023.103104.
- [7] A. Abdullah, J. Kolkmeier, V. Lo, and M. Neff, "Videoconference and Embodied VR: Communication Patterns Across Task and Medium," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, CSCW2, pp. 1–29, 2021, doi: 10.1145/3479597.
- [8] D. Oprean, M. Simpson, and A. Klippel, "Collaborating remotely: an evaluation of immersive capabilities on spatial experiences and team membership," *International Journal of Digital Earth*, vol. 11, no. 4, pp. 420–436, 2018, doi: 10.1080/17538947.2017.1381191.
- [9] A. D. Souchet, S. Philippe, D. Lourdeaux, and L. Leroy, "Measuring Visual Fatigue and Cognitive Load via Eye Tracking while Learning with Virtual Reality Head-Mounted Displays: A Review," *International Journal of Human-Computer Interaction*, vol. 38, no. 9, pp. 801–824, 2022, doi: 10.1080/10447318.2021.1976509.
- [10] A. Müller and A. Wittmer, "The choice between business travel and video conferencing after COVID-19 - Insights from a choice experiment among frequent travelers," *Tourism management*, early access. doi: 10.1016/j.tourman.2022.104688.
- [11] K. Kirchner, C. Ipsen, and J. P. Hansen, "COVID-19 leadership challenges in knowledge work," *Knowledge Management Research & Practice*, vol. 19, no. 4, pp. 493–500, 2021, doi: 10.1080/14778238.2021.1877579.
- [12] R. Riedl, "On the stress potential of videoconferencing: definition and root causes of Zoom fatigue," *Electronic markets*, early access. doi: 10.1007/s12525-021-00501-3.
- [13] X. Wei, X. Jin, and M. Fan, "Communication in Immersive Social Virtual Reality: A Systematic Review of 10 Years' Studies," in *Proceedings of the Tenth International Symposium of Chinese CHI*, Guangzhou, China and Online China, 2022, pp. 27–37, doi: 10.1145/3565698.3565767.
- [14] L. Karwan, "Virtual Reality in der Teamentwicklung," in *Hybrid Work*, J. Bath and K. Winkler, Eds., München: Haufe, 2023, pp. 229–236.
- [15] F. Moustafa and A. Steed, "A longitudinal study of small group interaction in social virtual reality," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, Tokyo Japan, S. N. Spencer, S. Morishima, Y. Itoh, T. Shiratori, Y. Yue, and R. Lindeman, Eds., 2018, pp. 1–10, doi: 10.1145/3281505.3281527.
- [16] J. Weidlich and T. J. Bastiaens, "Designing sociable online learning environments and enhancing social presence: An affordance enrichment approach," *Computers & Education*, vol. 142, p. 103622, 2019, doi: 10.1016/j.compedu.2019.103622.
- [17] Le Luo, D. Weng, N. Ding, J. Hao, and Z. Tu, "The effect of avatar facial expressions on trust building in social virtual reality," *Vis Comput*, vol. 39, no. 11, pp. 5869–5882, 2023, doi: 10.1007/s00371-022-02700-1.
- [18] R. E. Jack and P. G. Schyns, "The Human Face as a Dynamic Tool for Social Communication," *Current biology : CB*, vol. 25, no. 14, R621–34, 2015, doi: 10.1016/j.cub.2015.05.052.
- [19] S. Kimmel, F. Jung, A. Matvienko, W. Heuten, and S. Boll, "Let's Face It: Influence of Facial Expressions on Social Presence in Collaborative Virtual Reality," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany, A. Schmidt et al., Eds., 2023, pp. 1–16, doi: 10.1145/3544548.3580707.
- [20] A. Hamad and B. Jia, "How Virtual Reality Technology Has Changed Our Lives: An Overview of the Current and Potential Applications and Limitations," *International journal of environmental research and public health*, early access. doi: 10.3390/ijerph191811278.
- [21] F. Burrai, M. Sguanci, G. Petrucci, M. G. de Marinis, and M. Piredda, "Effectiveness of immersive virtual reality on anxiety, fatigue and pain in patients with cancer undergoing chemotherapy: A systematic review and meta-analysis," *European journal of oncology nursing : the official journal of European Oncology Nursing Society*, early access. doi: 10.1016/j.ejon.2023.102340.
- [22] J. Tesarz, H. Lange, M. Kirchner, A. Görlach, W. Eich, and H.-C. Friederich, "Efficacy of supervised immersive virtual reality-based training for the treatment of chronic fatigue in post-COVID syndrome: study protocol for a double-blind randomized controlled trial (IFATICO Trial)," *Trials*, early access. doi: 10.1186/s13063-024-08032-w.
- [23] T. Chandrasekera and S.-Y. Yoon, "Augmented Reality, Virtual Reality and Their Effect on Learning Style in the Creative Design Process," *Design and Technology Education*, vol. 23, no. 1, n1, 2018.
- [24] O. A. Meyer, M. K. Omdahl, and G. Makransky, "Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment," *Computers & Education*, vol. 140, p. 103603, 2019.
- [25] T.-W. Lui and L. Goel, "Learning effectiveness of 3D virtual reality in hospitality training: a situated cognitive perspective," *Journal of Hospitality and Tourism Technology*, vol. 13, no. 3, pp. 441–460, 2022.
- [26] L. J. Hagedorn, A. de Rooij, and M. Alimardani, "Virtual Reality and Creativity: How do Immersive Environments Stimulate the Brain during Creative Idea Generation?," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany, A. Schmidt, K. Väänänen, T. Goyal, P. O. Kristensson, and A. Peters, Eds., 2023, pp. 1–7, doi: 10.1145/3544549.3585848.
- [27] L. P. Berg and J. M. Vance, "An Industry Case Study: Investigating Early Design Decision Making in Virtual Reality," *Journal of Computing and Information Science in Engineering*, vol. 17, no. 1, 2017, Art. no. 011001, doi: 10.1115/1.4034267.
- [28] O. Torro, J. Holopainen, H. Jalo, H. Pirkkalainen, and A. Lähtevänoja, "How to get things done in social virtual reality—A study of team cohesion in social virtual reality-enabled teams," 09981331, 2022.
- [29] E. Yigitbas, S. Gorissen, N. Weidmann, and G. Engels, "Design and evaluation of a collaborative UML modeling environment in virtual reality," *Software and systems modeling*, early access. doi: 10.1007/s10270-022-01065-2.
- [30] S. Moore, M. Geuss, and J. Campanelli, "Communicating Information in Virtual Reality: Objectively Measuring Team Performance," in *Virtual, Augmented and Mixed Reality. Multimodal Interaction* (Lecture Notes in Computer Science), J. Y. Chen and G. Fragomeni, Eds., Cham: Springer International Publishing, 2019, pp. 473–489.
- [31] M. Sanaei, M. Machacek, S. B. Gilbert, P. Wu, and J. Oliver, "Comparing Perceptions of Performance Across Virtual Reality, Video Conferencing, and Face-to-Face Collaborations," in *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Honolulu, Oahu, HI, USA, 2023, pp. 4556–4561, doi: 10.1109/SMC53992.2023.10394218.
- [32] H. Cao et al., "Large Scale Analysis of Multitasking Behavior During Remote Meetings," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan, Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, and S. Drucker, Eds., 2021, pp. 1–13, doi: 10.1145/3411764.3445243.
- [33] Y.-C. Huang, L.-N. Li, H.-Y. Lee, M. H. Browning, and C.-P. Yu, "Surfing in virtual reality: An application of extended technology acceptance model with flow theory," *Computers in Human Behavior Reports*, vol. 9, p. 100252, 2023, doi: 10.1016/j.chbr.2022.100252.
- [34] C. R. Guerra-Tamez, "The Impact of Immersion through Virtual Reality in the Learning Experiences of Art and Design Students: The Mediating Effect of the Flow Experience," *Education Sciences*, vol. 13, no. 2, p. 185, 2023, doi: 10.3390/educsci13020185.
- [35] I. Valadez, S. Trullemans, and B. Signer, "Where Do We Meet? Key Factors Influencing Collaboration Across Meeting Spaces," Nov. 2023. [Online]. Available: <http://arxiv.org/pdf/2311.04707v1>
- [36] M. Aebersold, J. Rasmussen, and T. Mulrenin, "Virtual Everest: Immersive Virtual Reality Can Improve the Simulation Experience,"

- Clinical Simulation in Nursing*, vol. 38, pp. 1–4, 2020, doi: 10.1016/j.cens.2019.09.004.
- [37] Y.-Y. Wang, T.-H. Weng, I.-F. Tsai, J.-Y. Kao, and Y.-S. Chang, "Effects of virtual reality on creativity performance and perceived immersion: A study of brain waves," *Brit J Educational Tech*, vol. 54, no. 2, pp. 581–602, 2023, doi: 10.1111/bjet.13264.
- [38] A. Raake, M. Fiedler, K. Schoenenberg, K. de Moor, and N. Döring, "Technological factors influencing videoconferencing and zoom fatigue," arXiv preprint arXiv:2202.01740, 2022.
- [39] G. Fauville, M. Luo, A. C. Muller Queiroz, J. N. Bailenson, and J. Hancock, "Nonverbal mechanisms predict zoom fatigue and explain why women experience higher levels than men," Available at SSRN 3820035, 2021.
- [40] A. D. Souchet, D. Lourdeaux, J.-M. Burkhardt, and P. A. Hancock, "Design guidelines for limiting and eliminating virtual reality-induced symptoms and effects at work: a comprehensive, factor-oriented review," *Frontiers in psychology*, early access. doi: 10.3389/fpsyg.2023.1161932.
- [41] A. D. Souchet, W. Xie, and D. Lourdeaux, "Distinguishing Visual Fatigue, Mental Workload and Acute Stress in Immersive Virtual Reality with Physiological Data: pre-test results," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Christchurch, New Zealand, 2022, pp. 720–721, doi: 10.1109/VRW55335.2022.00211.
- [42] A. D. Souchet, D. Lourdeaux, A. Pagani, and L. Rebenitsch, "A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload," *Virtual Reality*, vol. 27, no. 1, pp. 19–50, 2023, doi: 10.1007/s10055-022-00672-0.
- [43] S. Riches, L. Taylor, P. Jeyarajaguru, W. Veling, and L. Valmaggia, "Virtual reality and immersive technologies to promote workplace wellbeing: a systematic review," *Journal of mental health* (Abingdon, England), early access. doi: 10.1080/09638237.2023.2182428.
- [44] A. de Carlo, F. Carluccio, S. Rapisarda, D. Mora, and I. Ometto, "Three uses of virtual reality in work and organizational psychology interventions. A dialogue between virtual reality and organizational well-being: Relaxation techniques, personal resources, and anxiety/depression treatments," *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, vol. 27, no. 1, 2020, doi: 10.4473/TPM27.1.8.
- [45] W. B. Schaufeli, A. Shimazu, J. Hakanen, M. Salanova, and H. de Witte, "An ultra-short measure for work engagement," *European Journal of Psychological Assessment*, 2017.
- [46] G. Fauville, M. Luo, A. C. M. Queiroz, J. N. Bailenson, and J. Hancock, "Zoom exhaustion & fatigue scale," *Computers in Human Behavior Reports*, vol. 4, p. 100119, 2021.
- [47] M. Bonfert et al., "Seeing the faces is so important—Experiences from online team meetings on commercial virtual reality platforms," *Front. Virtual Real.*, vol. 3, 2023, Art. no. 945791, doi: 10.3389/frvir.2022.945791.
- [48] D. D. Hodgson et al., "Visual feedback-dependent modulation of arousal, postural control, and muscle stretch reflexes assessed in real and virtual environments," *Frontiers in human neuroscience*, early access. doi: 10.3389/fnhum.2023.1128548.
- [49] E. Chang, H. T. Kim, and B. Yoo, "Virtual Reality Sickness: A Review of Causes and Measurements," *International Journal of Human-Computer Interaction*, vol. 36, no. 17, pp. 1658–1682, 2020, doi: 10.1080/10447318.2020.1778351.
- [50] S. H. Lee, M. Kim, H. Kim, and C. Y. Park, "Visual fatigue induced by watching virtual reality device and the effect of anisometropia," *Ergonomics*, early access. doi: 10.1080/00140139.2021.1957158.
- [51] E. Kim and G. Shin, "User discomfort while using a virtual reality headset as a personal viewing system for text-intensive office tasks," *Ergonomics*, early access. doi: 10.1080/00140139.2020.1869320.
- [52] J. S. Lemmens and C. F. von Münchhausen, "Let the beat flow: How game difficulty in virtual reality affects flow," *Acta psychologica*, early access. doi: 10.1016/j.actpsy.2022.103812.
- [53] Y. Jun, "The Differential Effects of Virtual Reality (VR) on the Novice and Experienced VR Users," *Asia Marketing Journal*, vol. 25, no. 2, pp. 61–70, 2023, doi: 10.53728/2765-6500.1610.
- [54] E. Vaportzis, M. G. Clausen, and A. J. Gow, "Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study," *Frontiers in psychology*, vol. 8, p. 1687, 2017, doi: 10.3389/fpsyg.2017.01687.
- [55] H. Lee and C. J. Bonk, "Collaborative Learning in the Workplace: Practical Issues and Concerns," *Int. J. Adv. Corp. Learn.*, vol. 7, no. 2, p. 10, 2014, doi: 10.3991/ijac.v7i2.3850.
- [56] M. Alcañiz, E. Parra, and I. A. Chicchi Giglioli, "Virtual Reality as an Emerging Methodology for Leadership Assessment and Training," *Frontiers in psychology*, early access. doi: 10.3389/fpsyg.2018.01658.
- [57] E. Parra, I. A. Chicchi Giglioli, J. Philip, L. A. Carrasco-Ribelles, J. Marín-Morales, and M. Alcañiz Raya, "Combining Virtual Reality and Organizational Neuroscience for Leadership Assessment," *Applied Sciences*, vol. 11, no. 13, p. 5956, 2021, doi: 10.3390/app11135956.
- [58] A. Gordon, M. van Lent, M. van Velsen, P. Carpenter, and A. Jhala, "Branching storylines in virtual reality environments for leadership development," in *Proceedings of the national conference on Artificial Intelligence*, pp. 844–851.

# Hybride Führung meistern: Einblicke in die Onlineplattform für Führungskräfte DigiLAP

Lene S. Fröhlich<sup>1\*</sup>, Annika Krick<sup>1</sup>, L. Franziska Münstermann<sup>1</sup>, Annabell Reiner<sup>1</sup>, Katharina Schübbe<sup>1</sup>, Karl-Heinz Renner<sup>2</sup> & Jörg Felfe<sup>1</sup>

<sup>1</sup> Helmut Schmidt University/Universität der Bundeswehr Hamburg, Hamburg, Deutschland

<sup>2</sup> Universität der Bundeswehr, München, Deutschland

\*froehlich.lene@hsu-hh.de

**Kurzfassung** – Im Rahmen des dtec.bw-geförderten Projektes „Digital Leadership and Health“ wurde eine wissenschaftlich fundierte und praxisnahe Lern- und Feedback-Plattform entwickelt, die Führungskräfte in ihrer digitalen Führung unterstützt. Die „Digital Leadership Assistance Platform (DigiLAP)“ ist ein Online-Tool, in dem sich Führungskräfte bezüglich der Chancen und Risiken des digitalen und hybriden Arbeitens, Führens und Zusammenarbeitens selbstständig weiterbilden können. In verschiedenen Basis- und Vertiefungsmodulen vermittelt DigiLAP Informationen zu relevanten Themen wie Gesundheit im Arbeitskontext, Führung, Zusammenhalt im Team oder Zusammenarbeit in Meetings. Führungskräfte haben zudem die Möglichkeit, ihre Situation bezogen auf die verschiedenen Themenbereiche anhand von Selbstchecks zu überprüfen und ihre Einschätzung durch Team-Befragungen mit der Sichtweise ihrer Mitarbeitenden zu vergleichen. Im „Cockpit“ können die Führungskräfte die Befragungsergebnisse anhand eines Ampelsystems im Blick behalten und Entwicklungsverläufe beobachten. Passend zu den Ergebnissen werden konkrete Handlungsempfehlungen und Übungen vorgeschlagen, welche die Führungskräfte dazu befähigen sollen, sich und ihr Team in den kritischen Bereichen weiterzuentwickeln. In einer ersten Testphase wurde DigiLAP von Führungskräften und FachexpertInnen aus Wissenschaft und Praxis ausführlich getestet und anhand einer quantitativen Befragung ( $N = 30$ ) sowie im Rahmen von Fokusgruppen ( $N = 18$ ) evaluiert. Die Ergebnisse zeigen, dass DigiLAP ein hilfreiches Instrument zur Weiterentwicklung digitaler Führungskompetenzen darstellt, welches Führungskräfte durch die eigenständige Bearbeitung sowie Möglichkeit der individuellen Rückmeldung flexibel und maßgeschneidert unterstützt. Nach Abschluss der Testphase wird die Plattform online (kosten)frei zugänglich sein, wodurch Führungskräfte im deutschsprachigen Raum von DigiLAP profitieren können.

**Stichworte** – Digitale und hybride Führung, Homeoffice, Gesundheit

## I. EINLEITUNG

Infolge der COVID-19-Pandemie und der damit einhergehenden Infektionsschutzmaßnahmen erhöhte sich die Zahl der Beschäftigten, die ihre Arbeit aus dem Homeoffice verrichteten, rapide. So arbeiteten laut dem ifo-Institut Ende 2020 etwa 76 % der Beschäftigten aus dem Homeoffice [1]. Auch nach der Pandemie gaben etwas über die Hälfte der Beschäftigten in einer repräsentativen Befragung den Wunsch an, weiterhin zwei bis drei Tage die Woche im Homeoffice arbeiten zu können [2]. Die Umstellung auf Homeoffice oder eine hybride Zusammenarbeit stellt vor allem Führungskräfte vor vielfältige Herausforderungen. Viele Führungskräfte sind verunsichert darüber, wie sie ihre Mitarbeitenden in digitalen und hybriden

Arbeitsformen erfolgreich führen können. So fällt es Führungskräften im Homeoffice schwerer einzuschätzen, wie es ihren Mitarbeitenden geht und ob diese Schwierigkeiten mit der Erledigung ihrer Aufgaben haben [3], [4].

Um Führungskräften eine Unterstützung bei den Herausforderungen der digitalen Führung zur Seite zu stellen, wurde im Rahmen des dtec.bw-geförderten Forschungsprojektes „Digital Leadership and Health“ an der Professur für Arbeits-, Organisations- und Wirtschaftspsychologie der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg eine wissenschaftlich fundierte und praxisnahe Onlineplattform für Führungskräfte entwickelt. Die „Digital Leadership Assistance Platform“ (kurz: DigiLAP) entstand in Zusammenarbeit mit verschiedenen Forschungseinrichtungen wie der Professur für Arbeit, Personal und Organisation der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg und der Professur für Persönlichkeitspsychologie und Psychologische Diagnostik der Universität der Bundeswehr München. Technisch umgesetzt wurde DigiLAP durch Zinal Consulting GmbH, die seit 20 Jahren Expertise im Bereich E-Learning-Management-Systeme haben.

## II. DIGITAL LEADERSHIP ASSISTANCE PLATFORM (DIGILAP)

### A. Ziele und Funktionen

Die Onlineplattform DigiLAP ermöglicht es Führungskräften, sich systematisch sowie orts- und zeitunabhängig mit Themen der digitalen und hybriden Führung und Zusammenarbeit auseinanderzusetzen. In diesem Sinne zielt DigiLAP darauf ab, Führungskräfte bei der digitalen und hybriden Führung zu unterstützen und sie dazu zu ermutigen, gemeinsam mit ihrem Team einen nachhaltigen Evaluations- und Optimierungsprozess zu gestalten.

Drei Hauptfunktionen stehen bei der Verwendung von DigiLAP im Vordergrund. DigiLAP ist gleichzeitig eine Lernplattform, ein Befragungs- und Feedbacktool und ein Assistenz- und Beratungstool für Führungskräfte. (1) Als **Lernplattform** strebt DigiLAP an, Führungskräfte zu Themen wie digitale und hybride Führung, Zusammenarbeit und Gesundheit zu informieren. Durch die Wissensvermittlung erhalten Führungskräfte die Möglichkeit, ihre Kompetenzen in den genannten Bereichen zu erweitern. (2) Zum anderen fungiert DigiLAP als **Feedback- und Befragungsplattform**. NutzerInnen können ihre eigene Situation durch Selbstchecks analysieren und sich anhand von Teambefragungen (Team Selbstchecks und Team Fremdchecks) Feedback von ihren Mitarbei-

tenden zu verschiedenen Themen rund um Führung, Zusammenarbeit und Gesundheit einholen. Im Anschluss an die Selbst- und Fremdchecks wird NutzerInnen automatisch ihre individuelle Auswertung in Form eines Ampelsystems angezeigt. Anhand dessen erfahren Führungskräfte, ob und in welchen Bereichen Handlungsbedarf besteht. Diese Analyse kommt einer „Gefährdungs- und Potenzialanalyse“ gleich und ermöglicht es Führungskräften, eigene Belastungen und Ressourcen sowie die ihres Teams besser einschätzen zu können. Das „Cockpit“ (s. Abbildung 2) auf DigiLAP bietet einen Überblick über alle Selbst- und Teambefragungen. Hier können Führungskräfte auf einen Blick sehen, welche Themen im „grünen Bereich“ sind und in welchen Bereichen noch Verbesserungspotenzial besteht. (3) Als **Beratungstool mit Assistenzfunktion** bietet DigiLAP gezielte Handlungsempfehlungen und konkrete Übungen, sodass Führungskräfte an ihrem Handlungsbedarf ansetzen, ihre Situation verbessern und Potenziale ausschöpfen können. Durch DigiLAP erhalten die NutzerInnen Hinweise, wie sie ihre eigene Situation und die ihres Teams verbessern können. Das Angebot gezielter Übungen und die Empfehlungen konkreter Maßnahmen helfen Führungskräften dabei, kritische Bereiche anzugehen und zu optimieren.

### B. Module und Themen

Ein Basis- und fünf Vertiefungsmodulbereiche geben Führungskräften auf DigiLAP die Möglichkeit, sich mit vielfältigen Fragestellungen und Themen wie Gesundheit und Führung systematisch und flexibel zu beschäftigen. Abbildung 1 verdeutlicht die Modulstruktur der Onlineplattform.

Der Basismodulbereich *Besonderheiten der Arbeit im Homeoffice verstehen* beschäftigt sich mit Chancen und Risiken der Arbeit im Homeoffice, Gesundheit, Leistung, Motivation sowie Führungsherausforderungen digitaler und hybrider Zusammenarbeit.

Im ersten Vertiefungsbereich *Effektives Führen im Homeoffice fördern* können Führungskräfte sich mit dem Konzept der Transformationalen Führung auseinandersetzen und lernen, wie Transformationale Führung im Homeoffice und in hybriden Arbeitsformen realisiert werden kann.

Der zweite Vertiefungsbereich *Gesundheitsorientiertes Führen im Homeoffice fördern* thematisiert den Ansatz Gesundheitsorientierter Führung (Health-oriented Leadership; [5]). Führungskräfte erfahren, wie sie ihre gesundheitsorientierte Selbstführung (SelfCare) und ihre gesundheitsorientierte Mitarbeiterführung (StaffCare) im Homeoffice und im Büro verbessern können.

Der Vertiefungsbereich *Individuelle Unterschiede bei der Arbeit im Homeoffice berücksichtigen* sensibilisiert Führungskräfte für individuelle Unterschiede und Voraussetzungen, durch die Mitarbeitende mehr oder weniger von der Arbeit im Homeoffice profitieren können.

Im vierten Vertiefungsbereich *Effektive Zusammenarbeit im Homeoffice fördern* können Führungskräfte sich mit Teamaspekten auseinandersetzen und lernen, wie Herausforderungen für hybride Meetings und digitale Zusammenarbeit bewältigt werden können. Außerdem wird thematisiert, wie die Zusammenarbeit verbessert und der Zusammenhalt im Team auch im Homeoffice gestärkt werden kann.

Der letzte Vertiefungsbereich *Unterstützende Rahmenbedingungen* vermittelt Führungskräften gesundheitsförderliche Personalmanagementstrategien. Führungskräfte erfahren,

welche Maßnahmen Unternehmen konkret ergreifen können, um die Gesundheit, Motivation und die Leistung ihrer Mitarbeitenden auch im Homeoffice langfristig zu stärken.

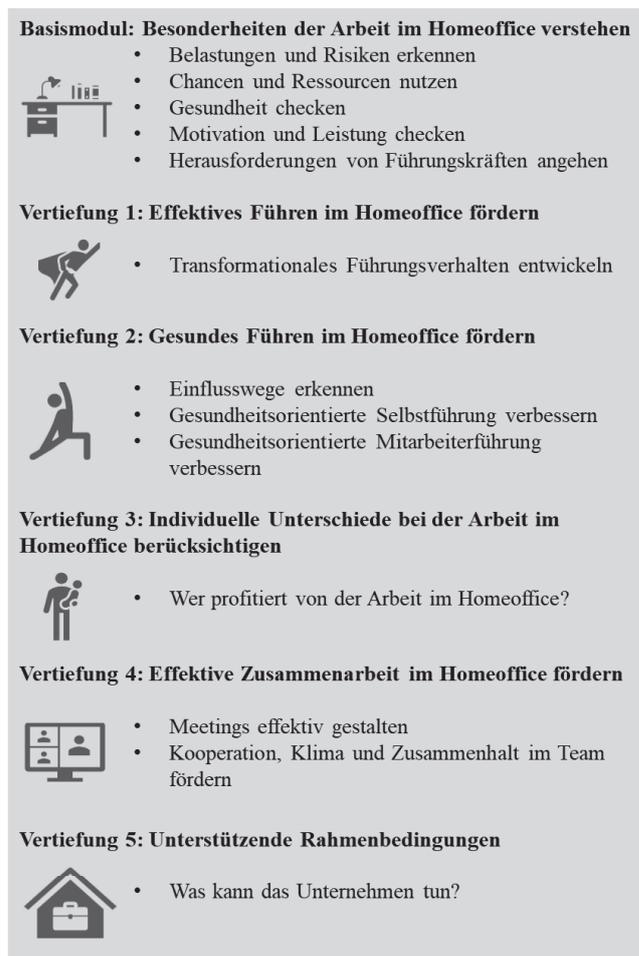


ABBILDUNG 1: MODULBEREICHE UND MODULE VON DIGILAP

### C. Nutzung von DigiLAP

Bei Beginn der Nutzung von DigiLAP können sich Führungskräfte im Reiter „Wie starte ich?“ informieren. Sie werden Schritt-für-Schritt angeleitet, wie sie DigiLAP möglichst systematisch und gewinnbringend verwenden können. Wir empfehlen den Führungskräften, ihren Mitarbeitenden im Rahmen eines Kick-Off-Meetings von ihrem Vorhaben, DigiLAP zur Weiterbildung zu nutzen, zu berichten. Dies kann dazu dienen, für Transparenz zu sorgen und die Motivation der Mitarbeitenden zu steigern, sich in den Prozess einzubringen und an den Befragungen teilzunehmen. Den Führungskräften werden verschiedene Nutzungsweisen von DigiLAP vorgestellt. Zum einen können die NutzerInnen sich Inhalte gezielt nach aktuellem Interesse auswählen oder sich für ihre Problemstellung die passenden Inhalte auswählen, wenn sie bereits ahnen, wo der „Schuh“ drückt. Hierfür steht eine Datei zur Verfügung, in der sich die NutzerInnen aus den Modulen ganz individuell ihren eigenen Ablaufplan erstellen können. Möchten NutzerInnen sich jedoch systematisch mit den Inhalten von DigiLAP auseinandersetzen, empfehlen wir die Module der Reihe nach durchzuführen.

### D. Aufbau von DigiLAP

Die Module folgen immer demselben Aufbau und bestehen aus den folgenden drei Kapiteln: „Worum geht es?“, „Wie sieht es bei mir und meinem Team aus?“ und „Was kann ich

tun?“. Im Kapitel „Worum geht es?“ erfahren die NutzerInnen anhand von Videos Genaueres zu dem jeweiligen Thema des Moduls.

Im Kapitel „Wie sieht es bei mir und meinem Team aus?“ können die NutzerInnen im Anschluss daran ihren persönlichen Selbstcheck machen. Die Selbstchecks fokussieren dabei auf unterschiedliche Themen rund um Homeoffice und hybrides Arbeiten. Jedes Modul hat einen eigenen Selbstcheck. Dies ist Schritt 1. Anschließend erhalten sie in Schritt 2 ihre persönliche Auswertung. Sie erhalten eine Gesamtpunktzahl für die Selbstchecks pro Modul und eine detaillierte Auswertung auf Ebene der einzelnen Handlungsfelder als Profil. Diese zeigt die gemittelten Ergebnisse für jeden Themenbereich bzw. jedes Handlungsfeld innerhalb eines Moduls. Anhand eines Ampelsystems wird gezeigt, wo Handlungsbedarf besteht. Während es bei „grünen“ Bereichen keinen Handlungsbedarf gibt, besteht bei „gelben“ oder „roten“ Bereichen mittleres bis hohes Verbesserungspotenzial und Handlungsbedarf. Zu den Themenbereichen bzw. Handlungsfeldern werden im Kapitel „Was kann ich tun?“ Handlungsempfehlungen, Tipps und Übungen angeboten. Neben Handlungsempfehlungen und Tipps finden die NutzerInnen hier konkrete Übungen, wie zum Beispiel Übungsblätter, Audios und Videos, die sie dabei unterstützen sollen, fit für das digitale und hybride Arbeiten zu werden.

Nach dem Selbstcheck und der Auswertung erhalten die Führungskräfte im Kapitel „Wie sieht es bei mir und meinem Team aus?“ außerdem in Schritt 3 die Gelegenheit, ihr Team mit Hilfe eines Teamchecks zu befragen. Hierfür kann eine E-Mail-Vorlage genutzt werden, die einen Einladungstext und den Link zur Befragung enthält. So können Führungskräfte ihre eigenen Perspektiven mit der Fremdsicht ihrer Mitarbeitenden abgleichen.

Im Cockpit erhalten die NutzerInnen eine Übersicht über alle Selbstchecks und Teamchecks (s. Abbildung 2). Hier werden alle Befragungsergebnisse von DigiLAP zusammengeführt. Mithilfe des Ampelsystems ist unmittelbar zu erkennen, in welchen Bereichen alles gut läuft und wo es Handlungsbedarf gibt. So bleiben die Führungskräfte stets über die wichtigen Themen der digitalen und hybriden Arbeit auf dem Laufenden und können ihre eigenen Einschätzungen anhand der Ergebnisse ihrer Selbstchecks für jedes Modul mit der Perspektive ihrer Teams anhand der Ergebnisse aus den Team Selbstchecks oder den Team Fremdchecks vergleichen. Aus Gründen der Anonymität werden die Befragungsergebnisse des Teams erst im Cockpit angezeigt, wenn mindestens fünf Mitarbeitende die Befragung ausgefüllt haben. Zusätzlich können Führungskräfte im Cockpit ihre Entwicklungsverläufe beobachten, indem die Ergebnisse früherer Befragungen aktuellen Befragungsergebnissen gegenübergestellt werden.

Anknüpfend daran können die Führungskräfte über den Reiter „Was nun?“ Hinweise erhalten, wie sie den DigiLAP-Prozess im Team begleiten können. Wir empfehlen einen Auswertungsworkshop zur Besprechung der Befragungsergebnisse gemeinsam mit dem Team durchzuführen und den Prozess der Befragung in regelmäßigen Abständen zu wiederholen. Zusätzlich können Führungskräfte sich im „Blog“ über aktuelle Forschungsergebnisse zu den Themen digitale Führung und Gesundheit informieren. Weitere Funktionen von

DigiLAP wie die Notiz- und die ToDo-Funktion unterstützen die Führungskräfte zusätzlich in ihrem Lernen und ihrer Handlungsplanung.

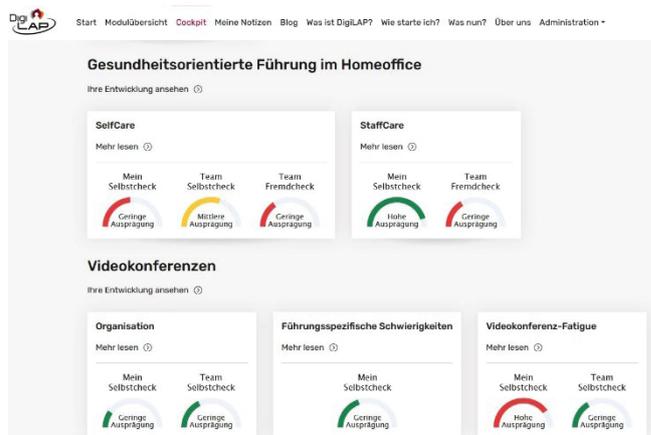


ABBILDUNG 2: EINBLICK IN DAS COCKPIT AUF DIGILAP

### E. Evaluation von DigiLAP

In einer ersten Testphase wurde DigiLAP von Führungskräften und FachexpertInnen aus Wissenschaft und Praxis ausführlich evaluiert. Die Testpersonen haben an einer quantitativen Befragung und an Fokusgruppen zur Bewertung der Plattform teilgenommen.

Zur Befragung wurde ein Evaluationsfragebogen entwickelt, der sich auf die wesentlichen Bereiche der Plattform (Module, Kapitel, Cockpit), das Design und die Usability bezog. Zusätzlich wurden soziodemographische Daten, wie die Homeoffice-Intensität, erhoben. Neben standardisierten Fragen, die auf einer Skala von 1 (trifft nicht zu) bis 5 (trifft zu) beantwortet wurden, enthielt der Fragebogen auch offene Fragen.

Insgesamt nahmen zwischen Dezember 2023 und März 2024  $N = 30$  Personen an der Befragung teil. Von den 30 Befragten waren 27 % Männer und 73 % Frauen. Der Altersdurchschnitt lag bei  $M = 40,2$  ( $SD = 8,9$ ). Die meisten Personen waren entweder in der Bildung, Erziehung und Wissenschaft (26,7 %), der öffentlichen Verwaltung (20 %) oder Pflege, Medizin und Gesundheit (20 %) tätig. 16,7 % der Personen arbeiteten fast immer in Präsenz bzw. im Büro. Ebenso viele (16,7%) arbeiteten eher in Präsenz bzw. im Büro. Die meisten jedoch (50 % der Teilnehmenden) arbeiteten gleichermaßen im Büro wie im Homeoffice. 16,7 % gaben an, eher im Homeoffice bzw. mobil zu arbeiten.

Die Ergebnisse der Befragung legen nahe, dass insbesondere das Design der Plattform, die Lernvideos und die Befragungs- und Feedbackfunktion von den Teilnehmenden sehr positiv bewertet wurden. Hinsichtlich des *Designs und der Nutzerfreundlichkeit* gaben zum Beispiel 93,3 % der Testpersonen an, dass DigiLAP ansprechend gestaltet ist<sup>1</sup>. Der Großteil der Testpersonen berichtete, dass DigiLAP ohne techni-

<sup>1</sup> Zur Auswertung wurden die Antwortoptionen 4 (trifft eher zu) und 5 (trifft zu) als Zustimmung und 1 (trifft nicht zu) und 2 (trifft eher nicht zu) als Ablehnung zusammengefasst.

sche Vorkenntnisse nutzbar ist (93,3 %), nutzerfreundlich gestaltet ist (80 %) und die Sprache klar und verständlich ist (96,7 %; s. Abbildung 3).

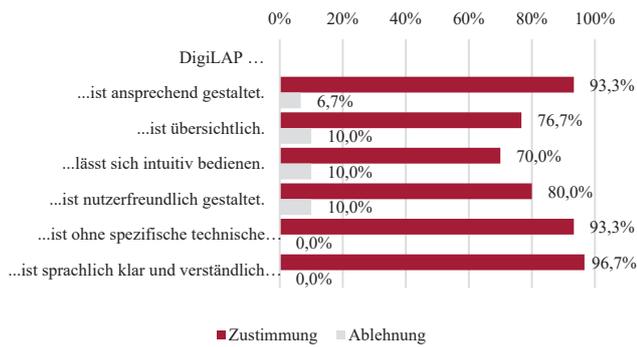


ABBILDUNG 3: ZUSTIMMUNG UND ABLEHNUNG DER TESTPERSONEN ZUM DESIGN UND ZUR NUTZERFREUNDLICHKEIT VON DIGILAP

Hinsichtlich der *Lernvideos* aus den *Kapiteln* „*Worum geht es?*“ gaben 96,7 % der Testpersonen an, dass diese gut aufbereitet sind. 86,7 % stimmten zudem zu, dass die Videos ausreichend über das jeweilige Thema des Moduls informieren.

Die Bewertung der Funktionen des *Kapitels* „*Wie sieht es bei mir und meinem Team aus?*“ ergab, dass auch hier der Großteil der Testpersonen die Selbst- und Fremdchecks als positiv bewerteten. Zum Beispiel gaben jeweils über 80 % an, dass der Selbstcheck ein hilfreiches Tool ist, um den eigenen Ist-Zustand zu erfahren (86,7 %), hilfreich ist, um die eigene Situation besser zu verstehen (83,3 %) und sie die Selbstchecks auch im Arbeitsalltag nutzen würden (80 %). Die Bewertung der Fremdchecks ergab ein ähnliches Bild. Ebenfalls positiv bewertet wurde die Möglichkeit, Selbst- und Fremdchecks zu vergleichen (Zustimmung über 70 %). Das Cockpit, in dem alle Ergebnisse im Überblick dargestellt sind, wurde von 90 % als positiv hervorgehoben. Alle Ergebnisse zur Bewertung der Selbst- und Fremdchecks, der Vergleiche und des Cockpits sind in Abbildung 4 dargestellt.

Das *Kapitel* „*Was kann ich tun?*“ wurde ebenfalls positiv bewertet. Hier gaben 90 % der Testpersonen an, dass die Handlungsempfehlungen hilfreiche Anregungen bieten, 80 % berichteten, dass diese sie konkret in ihrem Arbeitsalltag unterstützen. Jeweils 86,7 % hoben die Praxisnähe und die verständliche und ansprechende Gestaltung hervor. 73,3 % gaben an, dass die Handlungsempfehlungen sich gut im Arbeitsalltag anwenden lassen (s. Abbildung 5).

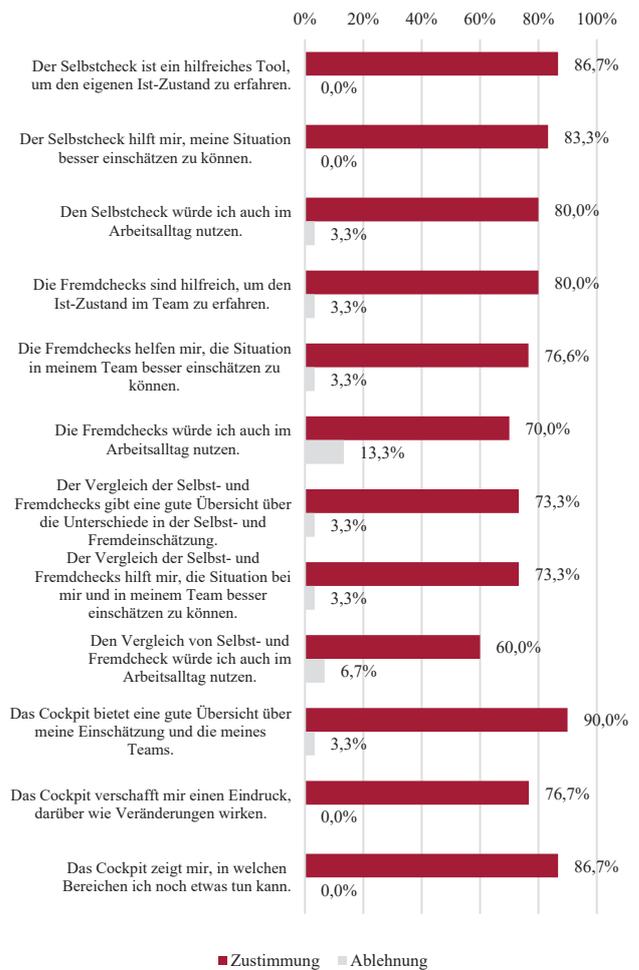


ABBILDUNG 4: ZUSTIMMUNG UND ABLEHNUNG DER TESTPERSONEN ZUM KAPITEL „WIE SIEHT ES BEI MIR UND MEINEM TEAM AUS?“ (SELBSTCHECKS, FREMDCHECKS UND COCKPIT)

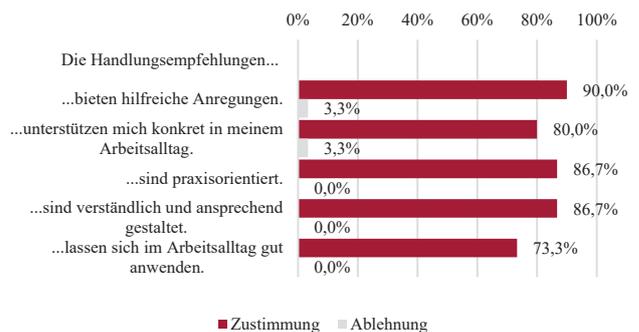


ABBILDUNG 5: ZUSTIMMUNG UND ABLEHNUNG DER TESTPERSONEN ZUM KAPITEL „WAS KANN ICH TUN?“ (HANDLUNGSEMPFEHLUNGEN)

Bei Betrachtung der allgemeinen Bewertung von DigiLAP zeigt sich, dass 76,7 % der Testpersonen DigiLAP zukünftig nutzen würden. 90 % denken, dass die Plattform auch für andere Führungskräfte nützlich sein wird und 83,3 % würden DigiLAP weiterempfehlen (s. Abbildung 6).

Im Durchschnitt lag die Bewertung der verschiedenen Bereiche der Plattform zwischen  $M = 3,97$  und  $M = 4,47$ , also im eher hohen bis sehr hohen Bereich. Insgesamt vergaben die

Teilnehmenden durchschnittlich  $M = 4,4$  Sterne von 5 für den Gesamteindruck der Plattform.

Abschließend konnten die Testpersonen anhand offener Fragen Verbesserungswünsche äußern (z. B.: Welche Features wären noch sinnvoll? Welche Inhalte wären noch wichtig?). Beispielsweise empfahlen einige Testpersonen, die Handlungsempfehlungen etwas kürzer aufzubereiten und in „kleinere Häppchen“ zu verpacken. Als zukünftige Features für DigiLAP wünschten sich die Testpersonen zum Beispiel einen Chatbot, eine noch stärkere Individualisierung, eine zusätzliche mobile App oder mehr Gamification. Hieraus wurde eine Liste an Weiterentwicklungsideen und Verbesserungsbedarf erstellt. Viele der Verbesserungsvorschläge konnten bereits jetzt umgesetzt werden. Andere Punkte können erst in einer späteren Projektphase umgesetzt werden, da es hierfür einer technischen Weiterentwicklung der Plattform bedarf.

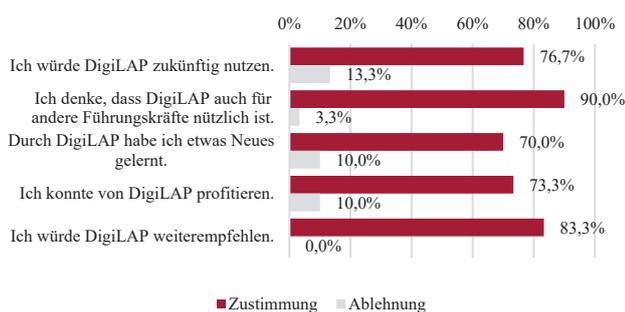


ABBILDUNG 6: ZUSTIMMUNG UND ABLEHNUNG DER TESTPERSONEN ZUR WEITEREMPFEHLUNG, ZUM PROFITIEREN UND ZUR ZUKÜNFTIGEN NUTZUNG

Um die konkreten Verbesserungsvorschläge besser diskutieren zu können, wurden Fokusgruppen durchgeführt, an denen insgesamt 18 Personen teilnahmen.

Auf Basis der Befragungsergebnisse wurden spezifische, offene Fragen formuliert, die zur Diskussion anregen sollten. Methoden wie *Mentimeter* wurden unterstützend eingesetzt, um die Fokusgruppen so interaktiv wie möglich zu gestalten. Die Fragen wurden in vier Blöcke unterteilt: (1) generelle Nutzung von DigiLAP (Warum sollte jemand DigiLAP nutzen? Welche Benefits bietet DigiLAP? Welche Nutzungsmöglichkeiten gibt es? Was könnte einer Nutzung entgegenstehen?), (2) Benutzeroberfläche der Plattform (Intuitive Gestaltung), (3) Motivation der NutzerInnen (Wie können NutzerInnen unterstützt werden DigiLAP langfristig zu nutzen?) und Weiterentwicklung sowie (4) Vermarktung der Plattform.

Am häufigsten gewünscht wurden kürzere Bearbeitungszeiten der Module und „kleinere Häppchen“, was sich insbesondere auf die Videos und Handlungsempfehlungen aus Kapitel 1 und 3 bezog. Dieser Punkt wurde in einer anschließenden Überarbeitungsphase priorisiert und umgesetzt. Hierfür wurden die Videos in kürzere Sequenzen aufgeteilt und Handlungsempfehlungen noch besser und übersichtlicher aufbereitet.

Darüber hinaus wurde ein zusätzliches offenes Fragetool gewünscht, mit Hilfe dessen NutzerInnen ergänzend schnell eigene Fragen stellen können und eine Auswertung erhalten. Dieser Punkt wurde in der nächsten Weiterentwicklungsphase umgesetzt.

Weiterhin wünschten sich die Teilnehmenden eine Unterstützung, um sich auf der Plattform besser zurecht zu finden.

Hierfür wurde der Bereich „Wie starte ich?“ auf DigiLAP integriert. Eine kurze Anleitung zeigt, welche Schritte als Erstes unternommen werden können. Zusätzlich werden die Inhalte (Welche Module gibt es?) sowie Funktionsweisen (Wie funktioniert DigiLAP?) übersichtlich und verständlich anhand von Erklärvideos dargestellt. Außerdem wird ein Bereich mit FAQs verfügbar sein.

Insgesamt zeigen die Ergebnisse der Testphase, dass die Plattform viel Potenzial für Führungskräfte und Unternehmen bzw. Behörden bietet. Die Testpersonen hoben vor allem die folgenden Punkte positiv hervor:

- die flexible Nutzung von DigiLAP durch E-Learning,
- die wissenschaftliche Fundierung,
- die Handhabung über die Kapitelstruktur,
- das Befragungs- und Feedbacktool durch Selbst- und Fremddchecks mit einer Echtzeitauswertung,
- aber auch Selbstreflexionsmöglichkeiten durch Übungen,
- die inhaltliche Verständlichkeit der komplexen Themen,
- den Wissenstransfer und die damit verbundene Weiterbildungsmöglichkeit,
- die praktischen Übungen und
- die Cockpitfunktion mit Verlaufskontrolle.

Trotz der Fülle an Informationen wurden die Materialien und Übungsblätter als informativ und hilfreich angesehen. Die Videos und insbesondere der „Gamification-Faktor“ von DigiLAP wurden als besonders ansprechend wahrgenommen, ebenso wie die Optik und die Funktionen des Cockpits.

#### F. Anwendungsbereiche

DigiLAP kann als digitales Führungsinstrument eingesetzt werden, um Führungskräfte bei den Herausforderungen der digitalen und hybriden Führung zu unterstützen. Gleichzeitig kann es als Tool innerhalb der Personal- und Führungskräfteentwicklung oder als Angebot des Betrieblichen Gesundheitsmanagements angewendet werden.

### III. AUSBLICK

Die Ergebnisse der Testphase werden genutzt, um DigiLAP weiterzuentwickeln und zu verbessern. Im Anschluss daran wird DigiLAP für alle Führungskräfte deutschlandweit online verfügbar sein. Das Führungsinstrument kann in verschiedenen Bereichen wie Entscheidungsfindung, Gesundheit, Motivation, Selbstbewertung und Evaluation genutzt werden. Auch im Rahmen des Betrieblichen Gesundheitsmanagements oder der Personalentwicklung kann DigiLAP Anwendung finden. Vor dem Hintergrund der vermehrten Nutzung von Arbeit im Homeoffice sowie hybriden Arbeitssettings und der voranschreitenden Digitalisierung ist es wichtig, Führungskräfte mit dem relevanten Wissen und den benötigten Kompetenzen auszustatten. Mit DigiLAP ist ein Online-Tool für Führungskräfte entwickelt worden, das sie dabei unterstützt, ihr Team auch in digitaler und hybrider Zusammenarbeit erfolgreich und gesund zu führen.

## DANKSAGUNG

Die AutorInnen bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. Das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr – ist ein von den Universitäten der Bundeswehr Hamburg und München gemeinsam getragenes wissenschaftliches Zentrum und Bestandteil des Konjunkturprogramms der Bundesregierung zur Überwindung der COVID-19-Krise. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

Des Weiteren bedanken sich die AutorInnen bei ihren derzeitigen studentischen Hilfskräften, Melanie Reule, B. Sc., Flora-Marta Blömer, B. Sc., Philip Georg Werner, B. Sc., Falk Döring und Maximilian Horn, für Ihre Unterstützung bei der Entwicklung und Evaluation von DigiLAP.

## LITERATUR

- [1] K. Demmelhuber, F. Englmaier, F. Leiss, S. Möhrle, A. Peichl, and T. Schröter, “Homeoffice vor und nach Corona: Auswirkungen und Geschlechterbetroffenheit,” *Ifo Schnelldienst digital*, vol. 1, no. 14, pp. 1-14, 2020.
- [2] A. Krick, J. Felfe, S. M. Neidlinger, L. Klebe, D. Tautz, K. Schübbe, Y. Frontzkowski, P. Gubernator, S. Hauff, and K.-H. Renner, *Auswirkungen von Homeoffice: Ergebnisse einer bundesweiten Studie mit Führungskräften und Mitarbeitenden*. [Online] Available: <https://www.hsu-hh.de/psyaow/newsblog-aus-unserem-dtec-projekt-digital-leadership-and-health/> (accessed: Apr. 19 2024).
- [3] A. Krick, J. Felfe, L. Klebe, and D. Tautz, “Hybrides Führen: Führen in Zeiten von Homeoffice,” in *Fehlzeiten-Report 2023: Zeitenwende – Arbeit gesund gestalten*, B. Bandura, A. Ducki, H. Schröder and M. Meyer, Eds.: Springer, 2023, pp. 271-286.
- [4] A. Krick, J. Felfe, and K. Schübbe, “Führung im Homeoffice,” in *Handbuch Mitarbeiterführung: Wirtschaftspsychologisches Praxiswissen für Fach- und Führungskräfte*, J. Felfe and R. van Dick, Eds.: Springer, 2023, pp. 659–675.
- [5] F. Franke, J. Felfe, and A. Pundt, “The impact of health-oriented leadership on follower health: Development and test of a new instrument measuring health-promoting leadership,” *Zeitschrift für Personalforschung*, vol. 28, 1-2, pp. 139–161, 2014.

# Praxisforschung zwischen Anspruch und Wirklichkeit. Über die Umsetzung von Partizipation und Wissenstransfer

Alexander Degel, Katharina Liebsch\*  
Fakultät für Geistes- und Sozialwissenschaften  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
\*liebschk@hsu-hh.de

Antje Masemann\*, Eric Sons,  
Abteilung Soziale Arbeit  
Northern Business School  
Hamburg, Deutschland  
\*EricSons@gmx.de

**Kurzfassung** – Der Beitrag nimmt in Auseinandersetzung mit der Tradition der Praxisforschung die kooperative Forschungspraxis zwischen den im dtec.bw-Projekt „Digitalisierung der psychosozialen Beratung im Feld der Familien- und Lebenshilfe“ beteiligten Konsortialpartnern selbstreflexiv in den Blick. Er veranschaulicht, wie Praktiker:innen an Durchführung und Ergebnis-Auswertung von Forschung beteiligt sind und beleuchtet die Prozesse der gemeinsamen Produktion von Einsichten, Produkten und Beziehungskonstellationen sowie die Grenzen dieser Zusammenarbeit.

**Stichworte** – partizipative Forschung, transdisziplinäre Zusammenarbeit, soziale Interaktion der Kooperation

## I. EINLEITUNG

Die Zusammenarbeit von Wissenschaft und Wirtschaft mit dem Anliegen Wissen zu erarbeiten, das dazu beiträgt, die jeweilige Praxis zu verbessern, ist konstitutiv für alle dtec.bw-Forschungsprojekte. In den Sozialwissenschaften firmiert diese Erwartung unter dem Label „Praxisforschung“ und wird von Anbeginn an methodologisch-kritisch diskutiert. In Auseinandersetzung mit dieser Tradition nehmen wir hier die kooperative Forschungspraxis zwischen den im dtec.bw-Projekt „Digitalisierung der psychosozialen Beratung im Feld der Familien- und Lebenshilfe“ beteiligten Konsortialpartnern selbstreflexiv in den Blick. Indem wir exemplarisch veranschaulichen, wie die Praktiker:innen an Durchführung und Ergebnis-Auswertung beteiligt sind, werden die Prozesse der gemeinsamen Produktion von Einsichten, Produkten und Beziehungskonstellationen, aber auch die Grenzen dieser Zusammenarbeit deutlich.

## II. PRAXISFORSCHUNG – IDEEN UND KONZEPTE

Mit Praxisforschung verbindet sich eine Forschungshaltung und Forschungsstrategie, deren Ursprünge in der sozialwissenschaftlichen Aktions- bzw. Handlungsforschung der 1930er und der 1970er Jahre liegen. Aktionsforschung will grundlagen- und praxisorientiertes Wissen hervorbringen, das dazu beiträgt, Lebenswelten sowohl zu verstehen als auch sie zu verändern. Ziel ist die Veränderung des Forschungsgegenstands während des Forschungsprozesses [1]. Dafür ist es erforderlich, die Akteur:innen der Praxis als Personen mit eigenen Erkenntnis- und Lernprozessen anzusehen und legitime Deutungen der erhobenen Empirie nicht ausschließlich bei den

forschenden Wissenschaftler:innen anzusiedeln [2]. Vielmehr sollen „sich die Beforschten in den Analysen und Interpretationen auch wiederfinden“ [3, p. 105].

Der methodische Ansatz auf Aktivierung von Betroffenen bei der Erkenntnisgewinnung ist historisch verbunden mit der sozialkritischen Gemeinwesenarbeit in der Nachfolge von Jane Adams, ein Ansatz, der sich in den 1930er Jahren in den USA zum „community organizing“ radikalisierte. Während dieser Zeit betrieb auch Kurt Lewin [4] in den USA eine Aktions- und Handlungsforschung, die Sozialwissenschaften als Mittel verstand, um soziale Konflikte und Probleme zu lösen. In Untersuchungen zu Arbeitsbedingungen von Fabrikarbeitenden sollten diese nicht nur als zu beforschendes Objekt verstanden werden, sondern als gleichberechtigte Akteure kooperativ miteinbezogen werden.

Diese Tradition verknüpfte sich in den 1970er Jahren mit der wachsenden Artikulation anti-kolonialer Theorie insbesondere in Lateinamerika wie auch mit der „Pädagogik der Unter-drückten“ von Paulo Freire, die in den Ländern des globalen Südens auch feministisch gewendet wurde. Im Umfeld westdeutscher pädagogischer Modellprojekte der frühen 1970er Jahre, insbesondere im Schulbereich (Gesamtschulen, Innovationen im Unterricht) wurde die Idee von Handlungsforschung als partizipative und formative wissenschaftliche Begleitung aufgegriffen und erprobt. Dabei war den Reform:innen die Nähe Kurt Lewins zum „social engineering“ im Auftrag der Industrie unangenehm, sie sahen ihren Ort vielmehr in der gesellschaftskritischen Diskussion der Sozialwissenschaften nach 1968 und bezeichneten ihren Ansatz als „aktivierende Sozialforschung“ [5].

Entsprechend ist auch die Praxisforschung im Bereich Sozialer Arbeit darauf ausgerichtet, Strukturen und Prozesse professioneller sozialpädagogischer Praktiken fachlich zu verbessern [6]. Hier sind die Praktiker:innen im Rahmen ihrer Tätigkeit forschend aktiv, indem sie selbst oder zusammen mit Wissenschaftler:innen Feld- und Fallstudien erarbeiten [7, p. 177ff.]. Methodisch gewann die teilnehmende Beobachtung einen hohen Stellenwert; die Ergebnisse der Evaluation sollten durch die direkte Rückkoppelung an die gesamte Projektgruppe für Verbesserung der Praxis nutzbar werden.

Heute findet sich der Begriff der Aktionsforschung im Bereich der Organisationsberatung mit dem Anspruch, die Belegschaft mitzunehmen, indem ihr Wissen in die Problemlösungen einfließt. Diese Konzepte lassen aber keine Beziehung

mehr zu einer politischen Kritik (und darauf bezogenen Zielen) an Macht- und Unterdrückungsverhältnissen erkennen. Je nach Schwerpunktsetzung und Perspektivierung nimmt Praxisforschung die Form einer emanzipatorisch-politischen Aktionsforschung, Evaluations- oder Auftragsforschung oder Begleit-Forschung an [8]. Während bei der Evaluationsforschung die Einbindung von Praktiker:innen am geringsten ist, fällt sie in der Aktionsforschung am höchsten aus. Dabei ist deren Einbindung in die Interpretation der Daten voraussetzungsvoll und benötigt erhebliche Ressourcen für entsprechende Schulungen und Qualifikationen. Die unterschiedlichen Typen und Formen von Praxisforschung unterscheiden sich hinsichtlich ihres Erkenntnisinteresses und des Verhältnisses von Theorie, Problembezug und Praxis: Wird eine wissenschaftliche, d. h. analytisch-theoretische Orientierung verfolgt (z. B. Design-Based Research-Konzept oder Grounded Theory) oder soll vorrangig den (Erkenntnis)-Interessen der Praktiker:innen gefolgt und handlungsbezogenes Praxiswissen produziert werden, das möglichst zeitnah in die Praxis implementiert werden soll?

### III. PRAXISFORSCHUNG ZUR DIGITALISIERUNG VON SOZIALUNTERNEHMEN

Das Erfordernis neuen Praxiswissens bildete den Ausgangspunkt des dtec.bw-Forschungsprojekts „Digitalisierung psychosozialer Beratung im Feld der Familien- und Lebenshilfe“. Hier brachte es die Covid-19-Pandemie mit sich, die sozialarbeiterischen Angebote ad hoc auf digitalisierte Formate umzustellen, trotzdem kaum zeitliche, personelle und informationelle Ressourcen dafür vorhanden waren, passgenaue und nachhaltige Lösungen für die Umstellung von analoger zu digitaler Interaktion zu entwickeln. So war es Ziel des Projekts, in enger Kooperation mit zwei ausgewählten Organisationen der Praxis Sozialer Arbeit einen Forschungs- und Entwicklungsprozess als exploratives und dialogisches Unterfangen der Problembeschreibung, Bedarfsanalyse und Entwicklung von Strategien samt korrespondierender Umsetzung einer „Digitalisierung der psychosozialen Beratung im Feld der Familien- und Lebenshilfe“ zu realisieren. Dabei galt es auch, Digitalisierung von Beratung nicht nur als Kompetenz-Erweiterung von Einzelnen, sondern zudem als arbeitspraktischen Prozess organisationalen Lernens eines gesamten Unternehmens bzw. gar eines gesamten, Bundesländer übergreifenden Verbandes zu verstehen.

Die beiden Projektpartner aus dem Bereich der Sozialwirtschaft sind Unternehmen in freier Trägerschaft, die im Hinblick auf die Digitalisierung ihrer psychosozialen Beratungstätigkeiten erst am Anfang stehen und zu Projektbeginn gerade damit begonnen hatten, die Digitalisierungsstrategie ihrer gesamten Organisation systematisch zu bedenken, ohne dass sie dabei auf vorhandene Expertise und organisationsinternes Know-how zurückgreifen konnten. Entsprechend bildeten die Nöte, Erfordernisse und Hemmnisse zur Digitalisierung den Ausgangspunkt für die gemeinsame Konzeptualisierung des Forschungsgegenstands, der mittels eines offenen, induktiven Vorgehens erarbeitet und bearbeitet wurde.

Das Projektkonsortium – bestehend aus zwei wissenschaftlichen Mitarbeiter:innen, zwei Professor:innen (Helmut-Schmidt-Universität Hamburg/Universität der Bundeswehr Hamburg und Northern Business School Hamburg) sowie zwei durch Projektmittel finanzierte Digitalisierungsbeauftragte der Praxispartner und deren jeweilige Geschäftsführer:innen – etablierte einen monatlichen Jour Fixe. Dort wurden die Themen und Fragen der unterschiedlichen Projektarbeitsgruppen diskutiert und bearbeitet. Es gab die *Projekt-AG*

*Milestones und Methoden*, in der die wissenschaftlichen Projektpartner die Arbeitsphasen planten und berieten, die *Projekt-AG Digitalisierung vor Ort*, in der die wissenschaftlichen Mitarbeiter:innen zusammen mit den Digitalisierungsbeauftragten der Praxispartner die Erhebungen und Arbeitsschritten konzipierten und reflektierten, sowie die *Projekt-AG Steuerung*, in der die Geschäftsführungen der Praxispartner und die Projektleitung zusammenarbeiteten.

Regelmäßigen Kontakt zwischen Wissenschaft und Praxis gab es, wenn die wissenschaftlichen Mitarbeiter:innen an thematisch einschlägigen Veranstaltungen, Diskussionsrunden und Ergebnisvorstellungen der Praxispartner ethnographisch beobachtend teilnahmen. Wechselseitigkeit und Rollentausch zwischen Wissenschaft und Praxis fanden statt, etwa wenn die wissenschaftliche Mitarbeiter:in bei der Erstellung eines Projektantrags des Sozialträgers mitarbeitete oder wenn die Digitalisierungsbeauftragten methodische Expertise für Erhebungsverfahren und organisationale Planungen in den Forschungsprozess einbrachten.

Für die Diskussion und Weiterentwicklung von Hypothesen standen die Wissenschaftler:innen mit den Digitalisierungsbeauftragten der Praxispartner zeitweise in einem regen, intensiven und frei gestalteten Austausch. Die Gespräche dienten zudem der Reflexion der konkreten Digitalisierungsvorhaben und der Vorbereitung zur Mediation von Konflikten, welche die organisationale Digitalisierung mit sich brachte. Damit realisierte sich die handlungs- und problembezogene Reflexion und Neugestaltung der sich digital wandelnden sozialarbeiterischen Arbeit und Beratungspraxis. Darüber hinaus machte die Theorie geleitete Auswertung der empirischen Daten der systematischen Erhebungen (teilnehmende Beobachtungen und mehrere Befragungen) insbesondere die mit organisationalen Digitalisierungsprozessen verbundenen neuen Kontroll- und Psycho-Dynamiken sowie die Veränderung von sozialen Interaktionsformen und organisationalen Aushandlungsprozessen deutlich.



ABBILDUNG 1: NÄHE UND FERNE

### IV. ÜBER DEN UMGANG MIT DYNAMIKEN DER PRAXISFORSCHUNG

#### A. Den Forschungsgegenstand neu konzeptualisieren

Mit Beginn des Projektes wurde mittels standardisiert-quantitativer Verfahren (Abfrage per Fragebogen) und qualitativ rekonstruktiver Methoden (Interviews und Gruppendiskussionen) das organisationale Wissen und die Praktiken zum Thema Digitalisierung und Online-Beratung erhoben. Die Bedarfsanalyse zeigte, dass das primäre Anliegen und der vorrangige Bedarf der Praxispartner in der Digitalisierung der internen Kommunikation und des Personal- und Verwaltungsmanagement lagen. Zum Thema Online-Beratung hingegen

gab es bei den beteiligten Sozialträgern einen eher geringen Entwicklungs- und Reflexionsbedarf. Entsprechend rückten dann Vorgehensweisen zur Weiterentwicklung organisationaler Digitalisierung in den Mittelpunkt der Forschung, während die Gestaltung und Verbesserung der Online-Beratungspraxis in den Hintergrund traten. Für die Praxis-Partner:innen klärten sich durch die Bedarfsanalyse die Prioritäten ihrer Anliegen in Sachen Digitalisierung. Die beteiligten Wissenschaftler:innen hingegen waren nun gehalten, die in der Phase der Beantragung des Projekts entwickelten Grundlagen und Hypothesen neu auszurichten.

### *B. Fremdheit zwischen Wissenschaftler:innen und Praktiker:innen überwinden*

Auf Seiten der Praxispartner herrschte bei einigen Mitarbeitenden, die selbst nicht studiert hatten, zunächst Zurückhaltung im Kontakt und Berührungangst in der Zusammenarbeit mit den Wissenschaftler:innen aus der „fremden, akademischen“ Welt. So gab es beispielsweise die Frage, ob man die Projektleitung duzen oder siezen sollte, die Unterstellung eines Kontrollansinnens von Befragungen und teilnehmender Beobachtung oder auch die Mitarbeiter:innen, die vor dem Expert:innen-Interview ihre Vorgesetzte fragten: „Was sollen wir dem eigentlich erzählen?“ und die dann während des Interviews sozial erwünscht antworteten und die vorherrschenden Konflikte in der Organisation bewusst und gezielt nicht thematisierten. Umgekehrt bewirkte die Erfahrung des Anders-Seins bei einigen Forschenden, sich zunächst mit den Logiken der sozialen Gruppe der Praktiker:innen stark zu identifizieren und die eigene Rolle als Forschende zu minimieren, so dass sie erst im späteren Verlauf des Forschungsprozesses als Forschende aktiv und sichtbar wurden.

Damit diese kulturelle Fremdheit nicht zum Bremsklotz für Kooperation und Datenerhebung wird, erwies es sich zur Bearbeitung des Themas und dessen Überwindung als hilfreich, Ergebnisse der Beobachtungen und Befragungen bei Teamtreffen der Sozialträger vorzustellen und dort ausführlich zu diskutieren. Darüber hinaus wirkte sich die Arbeit mit dem methodischen Konzept der Situationsanalyse [9] als vorteilhaft für die Zusammenarbeit von Wissenschaft und Praxis aus. In der Situationsanalyse benennen die Praktiker:innen konkrete Erfahrungen, deren thematische Relevanz in sogenannten Mappings von Erlebnissen und Ereignissen gemeinsam herausgearbeitet wird.

### *C. Unterschiedliche Interessen, Kompetenzen und zeitliche Ressourcen in Einklang bringen*

Die enge Kooperation mit den Praktiker:innen während des Forschungsprozesses bringt es mit sich, die Perspektiven und Interessen von Wissenschaft und Praxis abzugleichen und miteinander zu vermitteln. Nicht selten verbinden Praktiker:innen mit Wissenschaft Evaluation und Optimierung von Arbeitsprozessen, während Forschende auch daran interessiert sind, einen empirischen Zugang ins Feld zu bekommen, um neben wissenschaftlicher Begleitung und Problemlösung auch die Theoriebildung des Fachs datenverankert voranzutreiben.

So waren wir im Projektverlauf mit konkreten Ungleichheiten konfrontiert, etwa dass die Wissenschaftler:innen über mehr Zeit für Reflexionsprozesse verfügen, im methodischen Denken geschult und es gewohnt sind, detailliert mittels der akademischen Sprache über Sachverhalte zu sprechen und zu debattieren. Beispielsweise kommentierte eine Praxispartnerin den Workshop, bei dem Forschungsergebnisse vorgestellt wurden, dass es interessant gewesen sei, sie aber „wenig von dem Gesagtem verstehe“. Diese Einschätzung aufnehmend

haben wir gemeinsam um Übersetzungen fachspezifischer Termini und unterschiedlicher Denklogiken gerungen, mit dem Ziel, eine gemeinsame Sprache zwischen Praxis und Wissenschaft zu entwickeln.

Insgesamt gelang es, die Zusammenarbeit regelhaft, verbindlich und auf Augenhöhe zu gestalten. Dabei verlief der Informationsfluss überwiegend von der Praxis zur Wissenschaft und die inhaltliche Verknüpfung zu wissenschaftlichen Erkenntnisinteressen und Fragestellungen hatte im Rahmen der Projektmeetings deutlich weniger Raum. Dies war häufig der Zeitknappheit und der Fülle von Themen geschuldet. Entsprechend waren die wissenschaftlichen Perspektiven und ihre Bedeutung für die Praxisreflexion eher selten regelmäßiger Bestandteil von Projekttreffen. Vielmehr wurden sie fokussiert in dafür vorgesehenen Einzel-Veranstaltungen eingebracht und werden sich mit dem Ende der Projektlaufzeit intensivieren.

### *D. Nähe und Distanz austarieren*

Der Spagat zwischen dem konkreten praktischen Handeln einerseits und der distanzierten Wahrnehmung und Reflexion dieser Handlungspraktiken andererseits stellt die Praxispartner:innen vor Herausforderungen. Sie müssen professionsbezogene Selbstbilder, eingespielte Routinen und Hierarchien in Frage stellen und neue Entwürfe bzw. Leitbilder für Subjektivierungsweisen und die Organisation insgesamt produzieren. Dafür braucht es ausreichend Zeit, kollegiales Miteinander und Expertise. Im Rahmen unseres Projekts konnten wir dies weitenteils realisieren. Indem Schulungen bezahlt, Weiterbildungen ermöglicht und Organisationsentwicklungsprozesse initiiert wurden, vergrößerte sich auch die Kompetenz aller Beteiligten, zur eigenen Arbeitsorganisation in Distanz zu treten und sie wohlwollend-kritisch zum Gegenstand des organisationalen Lernens zu machen.

Auf Seiten der beteiligten Wissenschaftler:innen war schnell klar, dass eine an der Aktionsforschung ausgerichtete Fokussierung auf die Bedarfe und Optimierungsbedürfnisse der Praxis es mit sich bringen kann, die Wissenschaft auf die Rolle des Dienstleisters einzuengen. Deswegen galt es, eine eigene, von der Praxis unabhängige wissenschaftliche Agenda zu entwickeln und dabei das Engagement für die Praxis sicherzustellen, ohne das eigene Erkenntnisinteresse aus den Augen zu verlieren.

Die Verstrickung in die Logiken der Praxis stellte vor dem Hintergrund der engmaschigen Kooperation von Wissenschaft und Praxis eine Verführung zur Komplizenschaft dar, die in der wissenschaftlichen Projekt-AG thematisiert und reflektiert wurde. Gegenstand von Reflexion war beispielsweise, dass eine wissenschaftliche Mitarbeiter:in zeitweise die Sprache und die Begrifflichkeiten des Organisationsberaters übernahm, der den Praxispartner beriet, und sie zur Beschreibung der eigenen wissenschaftlichen Beobachtungen verwendete. Identifiziert mit den Perspektiven und Logiken des Praxispartners ging die kritisch-wissenschaftliche und reflektierende Distanz zwischenzeitlich verloren. Ein weiteres Thema war, dass die regelmäßige, mehrmals wöchentlich stattfindende Teilnahme einer wissenschaftlichen Mitarbeiter:in bei Terminen der Praxispartner zwar eine große Menge an interessanten empirischen Daten erbrachte, aber zugleich dazu führte, dass die wissenschaftliche Mitarbeiter:in ihre Rolle als Wissenschaftlerin sukzessive verließ. Dies machte es erforderlich, den Austausch und die Beziehung zwischen Wissenschaft und Praxis neu zu justieren und seitens der Forschenden die Anwesenheit bei den Praxispartnern vor Ort punktuell auch einzuschränken.

Bestandteil dieser Pendelbewegungen zwischen Nähe und Distanz war zudem, dass die notwendige explorative Offenheit der wissenschaftlichen Fragestellung des Projekts andauernd gestaltet, neu justiert und verändert werden musste. Dazu gehörte auch, sich kontinuierlich damit zu beschäftigen, welche Theorien und Konzepte das Forschungsthema aufzuschlüsseln und zu bereichern vermögen.

#### V. FAZIT: PRAXISFORSCHUNG ZWISCHEN WISSENSCHAFTLICHER BEGLEITUNG, KO-PRODUKTIVER ERKENNTNISERZEUGUNG UND DER BEARBEITUNG WISSENSCHAFTLICHER FRAGEN

Die Bildung von Konsortialpartnerschaften zu Forschungszwecken erfolgt unter der Annahme, dass die transdisziplinäre Zusammenarbeit von Expert:innen verschiedener Bereiche zur forschersischen Produktivität beiträgt: Die Forschung findet dort statt, wo ihre Ergebnisse Anwendung finden sollen, Unterschiede und Gemeinsamkeiten von Wissenschaft und Praxis werden aktiv gestaltet, auftretende Schwierigkeiten sind für alle Beteiligte wahrnehmbar und können so zügiger bearbeitet werden.

Um die Zusammenarbeit fruchtbar werden zu lassen, ist Kommunikation – also Zeit für Austausch und Verständigung und Übersetzungen der unterschiedlichen Logiken und Bezeichnungspraktiken – unerlässlich. Diese wird zumeist von Seiten der Wissenschaft koordiniert und geht als Bestandteil des Forschungsprozesses auch in das Ergebnis der Forschungsleistung ein. Von Seiten der Praxis ist dafür das Interesse und die Bereitschaft erforderlich, sich als Ko-Forschende zu beteiligen. Mit dieser Beteiligung ist verbunden, dass Praktiker:innen befähigt und legitimiert werden für die gemeinsame Erforschung sozialer Wirklichkeiten und um Veränderung sozialer Praxis im Verlauf des Forschungsprozesses gezielt zu gestalten. Dafür ist es auch erforderlich, dass beide Seiten die im gemeinsamen Forschungsprozess sichtbar werdenden Ungleichheiten und Machtkonstellationen kritisch reflektieren.

Die Herstellung von weitgehend macht- und hierarchiefreien Räumen gemeinsamen Arbeitens ist auch deshalb eine Herausforderung, weil hier Wissensformen aufeinandertreffen, die in einem hierarchischen Verhältnis zueinander stehen. Theoretisches Wissen, Erfahrungswissen, situiertes Wissen über Beziehungen, Interaktionen, Gemeinschaften sind asymmetrisch im Hinblick auf Relevanz und Status sowie im Hinblick auf die Beanspruchung von Deutungsmacht und Wissenshoheit. Erst wenn nicht der Status von Wissen, sondern die Funktion von Wissen ins Zentrum rückt, können alle Beteiligten zu Expert:innen werden.

Im Feld der deutschsprachigen Sozialen Arbeit wird diese Art der Forschungskoooperation seit den 1980er und 1990er Jahren als *Praxisforschung* bezeichnet und sie ist zumeist auf betriebliche Organisations- und Qualitätsentwicklungen konzentriert [10, p. 8]. Die Sozialarbeitswissenschaftlerin Maja Heiner [11] unterscheidet drei Modelle der Praxisforschung im Hinblick auf die Form der Zusammenarbeit zwischen Wissenschaft und professioneller Praxis und dem forschersischen Anteil der Fachkräfte: Im ersten Modell kooperieren Vertreter:innen der Praxis mit Forscher:innen, wobei sie selbst keine forschende Agenda übernehmen. Insofern bleibt die übliche Arbeitsteilung bestehen. Im zweiten Modell forschen sowohl Vertreter:innen der Praxis als auch Forscher:innen.

Mitarbeiter:innen und Leitungspersonen sind hier beispielsweise in die Datenerhebung eingebunden. Im dritten Modell besteht die zentrale Aufgabe der Wissenschaftler:innen in der Beratung der forschenden Praktiker:innen.

Beteiligung von Praktiker:innen als Ko-Forschende bedeutet demzufolge, dass weder *für* noch *über*, sondern *gemeinsam mit* ihnen geforscht wird. Dies kann in allen oder in ausgewählten Phasen des Forschungsprozesses erfolgen und reicht von der Anhörung über die Mitbestimmung bis hin zur Entscheidungsbefugnis der Ko-Forschenden. Dies kann nur gelingen, wenn die verschiedenen Bezugssysteme und Sachzwänge der Beteiligten, die unterschiedlichen Erwartungen, Interessen und Deutungen abgeglichen und konstruktiv miteinander verschränkt werden.

Dabei ist es naheliegend, dass diese Kooperation Konflikte mit sich bringt – eine Dynamik, die auch als ein Qualitätsmerkmal verstanden werden kann, als ein Hinweis darauf, dass alle Beteiligten ihre Anliegen einbringen, verhandeln und sich für ihre Interessen einsetzen [10, p. 86f.].

Für die ko-forschenden Praktiker:innen sind Selbstbestimmung und Eigenregie wichtig, um den zirkulären Prozess in Gang zu setzen, in dem die gewonnenen Informationen und Daten sowohl in den Forschungsprozess eingespeist als auch dazu genutzt werden, praxisrelevante Handlungsstrategien zu entwickeln. Für die beteiligten Wissenschaftler:innen bedeutet Eigenregie, den schmalen Grat zwischen praktischem Problembezug und wissenschaftlicher Perspektive auszuhalten und auszutarieren. Dazu gehört es auch, Verstrickungen in die Praxis und mit der Praxis abzuklären und die Forschungsagenda, die sich im Zuge empirischer Erhebungen und Ergebnisse immer wieder verändert, nicht aus dem Blick zu verlieren.

Nur so ist es möglich, schlussendlich Konzepte und Kategorien empirisch gesättigt zu identifizieren, die sowohl wissenschaftlich als auch praktisch von Bedeutung sind und die das weitere Nachdenken über Forschungsgegenstand und Praxisprobleme produktiv werden lassen.

#### DANKSAGUNG

Die Autor:innen bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. Das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr – ist ein von den Universitäten der Bundeswehr Hamburg und München gemeinsam getragenes wissenschaftliches Zentrum und Bestandteil des Konjunktur-programms der Bundesregierung zur Überwindung der COVID-19-Krise. Es unterliegt der akademischen Selbstverwaltung. Die Mittel, mit dem das dtec.bw ausgestattet wurde, werden an beiden Universitäten der Bundeswehr zur Finanzierung von Forschungsprojekten und Projekten zum Wissens- und Technologietransfer eingesetzt. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

LITERATUR

- [1] J. Rieger und G. Straßburger, „Warum Partizipation wichtig ist. Selbstverständnis und Auftrag sozialer Berufe,“ *Straßburger, Gaby/Rieger, Judith (Hrsg.): Partizipation kompakt. Für Studium. Lehre und Praxis sozialer Berufe*, Bd. Weinheim: Beltz Juventa, p. 42–51, 2014.
- [2] K. Holzkamp, *Grundlegung der Psychologie*, Frankfurt/M./New York: Campus, 1983.
- [3] P. Mayring, *Einführung in die qualitative Sozialforschung*, München: Psychologie Verlags Union, 1990.
- [4] K. Lewin, „Group Decision and Social Change,“ *Eleanor E. Maccoby, Theodore M. Newcomb, Eugene E. Hartley (Hrsg.): Readings in Social Psychology*, p. 459–473, 1952.
- [5] T. Heinze, *Handlungsforschung im pädagogischen Feld*, München: Juventa, 1975.
- [6] E. Steinert, *Sozialarbeitsforschung für Studium und Praxis.*, Frankfurt/M.: Peter Lang, 2008.
- [7] B. Hill, „Forschung in der Kulturellen Bildung,“ *B. Hill/T. Biburger/A. Wenzlik (Hrsg.): Lernkultur und Kulturelle Bildung*, p. 174–187, 2008.
- [8] H. Moser, *Instrumentenkoffer für die Praxisforschung*, Freiburg: Lambertus, 2015.
- [9] A. Clarke, *Situationsanalyse. Grounded Theory nach dem Postmodern Turn*, Wiesbaden: SpringerVS, 2012.
- [10] H. v. Unger, *Partizipative Forschungspraxis. Einführung in die Forschungspraxis*, Wiesbaden: Springer, 2014.
- [11] M. Heiner, *Praxisforschung in der Sozialen Arbeit*, Freiburg: Lambertus, 1988.

# Reconfiguring Teachers' Work Through Automated Evaluation

Samira Alirezabeigi<sup>a</sup>, Mathias Decuyper<sup>a,b</sup>

<sup>a</sup>*KU Leuven*

Leuven, Belgium

<sup>b</sup>*Zurich University of Teacher Education*

Zurich, Switzerland

\* samira.alirezabeigi@kuleuven.be

\* mathias.decuyper@phzh.ch

Sigrid Hartong

*Helmut-Schmidt-University/*

*University of the Federal Armed Forces*

Hamburg, Germany

\*hartongs@hsu-hh.de

**Abstract** – Over the past years, the potential benefits and risks of educational technologies (EdTech) in schools have been increasingly debated. At the same time, research has only started to develop a more nuanced understanding of the *manifold pedagogical effects* of these technologies in school and classroom practice. Within the broader context of the dtec.bw project SMASCH (Smart Schools), which aims at supporting schools in their digital transformation, this chapter focuses on concrete manifestations and effects of automated evaluation practices on teachers' work. More specifically, the study analysed the screen recordings of ten digital tests from the moment of their creation to their archiving and is equally based on interviews with teachers about their evaluative practices. The findings of the study indicate different and ambivalent processes that reconfigure the relations between teachers and technology, which we exemplify using two cases – when teachers make digital tests for later automation, and when teachers supervise the machine's actual grading and feedback provision practices. Our findings not only offer important conclusions regarding how automated evaluation can, and should, be more strongly framed in a pedagogical manner, but they can also inform the need for a careful response to the more recent rise of GenAI technologies.

**Keywords** - Digitization of schools, teachers' work, automated evaluation practices, grading, SMASCH (Smart Schools)

## I. FRAMING THE RESEARCH: THE SMASCH PROJECT

The project SMASCH (Smart Schools) was initiated in the beginning of 2021 as part of the *Centre for Digitization and Technology Research* (dtec.bw). The aim of the project, which brings together research from education science, sociology, and organizational studies, is to support schools in their digital transformation. The project hereby seeks to overcome a technological focus on digitization (e.g. bringing tablets into schools), but instead approaches school transformation from a 'culture of digitality'-perspective [1]. This means that schools are supported in critically dealing with, and creatively responding to, wider socio-technical transformations of society in relation to their specific local context and organizational culture (see also [2]). A key part of the project and its aim to experimentally integrate digital technologies is therefore the discussion, visualisation and transformation of organisational culture. From a research perspective, an interest in the overall organisational dynamics caused by digitization is hereby brought together with the critical investigation of EdTech design (see also [3]).

As the literature clearly shows, EdTech enactments and their ambiguous effects oftentimes stand in (sharp) contrast to hopes that are commonly associated with the technologization of education, including the enhancement of educational equality, the personalization of education, or increased accessibility to knowledge [4]. For instance, structural limitations of schools (outdated ICT infrastructures, insufficient funding, lack of time or skills to adapt products) has regularly led to a rather unreflecting EdTech usage (see also [5]). Schools have, in that regard, commonly used EdTech products that are free of charge and easily usable, but, at a closer glance, neither safe nor pedagogically sound [6], [7], [8]. Moreover, the relative absence of qualitative evidence on the efficacy and/or performativity of EdTech applications has made it increasingly hard for schools to know which EdTech products to procure [9]. Lastly, the literature clearly shows that digitization, if it is to be successfully implemented in schools, needs to be systematically tied to a comprehensive view on school leadership and school culture [10]. That is to say, when considered through a school development lens, the literature clearly shows the advantage and necessity of adopting a *whole school approach* towards digitization, instead of giving this responsibility into the hands of specific singular actors – such as ICT coordinators or principals – alone [11].

In line with this argumentation, the SMASCH project has been working with 13 schools in Germany (Hamburg) and Belgium (Flanders) over the past four years, including both primary and secondary schools, schools with diverging socio-economic status backgrounds and locations (rural, suburban, urban), as well as different scope of experience regarding the implementation of digital technologies. Despite these differences, the focus shift from digitization as (fast) instrumental implementation and (easy) usability of technologies towards more experimental forms of digitization as ongoing and conscious reflection of school culture, values and practices, formed unfamiliar territory for most schools. Consequently, the project aimed at providing extensive support for the schools, which in both countries included co-design research and different forms of events, and – in the case of Hamburg – additional support from school development coaches, IT experts and media educators.

In fact, since the initiation of the project, the schools have not only changed significantly with regards to how they approach digitization in relation to their organisational culture, but they equally identified different transformation areas that

appeared most suitable to address their specific local needs (see for an overview [2]). The research (and wider support) team, consequently, focused in a locally specific manner on these areas, and co-designed different scenarios of experimentation together with the schools. These areas range from the (re)design of classrooms or school buildings, over the (re)design of learning management systems [12] or new apps, topic-related multimodal project weeks (e.g. a VR-supported project week around bees), to the more specific socio-technical (re)design and reflection of individual pedagogical contexts. Hereby, the research in SMASCH always integrated perspectives on governance (e.g. digitization policies) [13]. In doing so, different studies in SMASCH show how Edtech always carries pedagogical conceptions and ideas with it that are ‘baked’ into how applications (e.g. learning management systems) are designed, and that influences the schools’ ‘world-making’ and pedagogical decision-making in multiple ways [14], [15].

In what follows, we illuminate one of the many studies conducted in SMASCH over the past years, which is situated in the broader context of how teachers’ pedagogical practices and decision-making become reconfigured through platformization: automated evaluation.

## II. AUTOMATED EVALUATION AS MULTIFACETED SOCIO-TECHNICAL PRACTICES

When digital evaluation is implemented in schools, automation most visibly manifests itself, and is being discussed, in relation to *grading*. This is especially so since grading practices have commonly been considered as one of the most labour-intensive and repetitive tasks for teachers. In this respect, EdTech promoters oftentimes argue that automation of grading can liberate teachers from this time-consuming task, so that they can concentrate on educationally more valuable activities [16]. In addition to time-efficiency, automated grading is commonly associated with the enhancement of fairness and consistency in relation to the subjective practices of humans [17].

At the same time, automated grading is always related to, and affected by, broader practices of (digital) evaluation. This means that grading in a school context is always embedded within broader “constellations of valuation” [15] that is, understandings of what is of (which) value when judging educational performance against the background of students’ individual situatedness. Shifting the form of evaluation (and grading as a key component of evaluation) into a digital form, thus, plays an important role in transforming (or reproducing) such understandings.

In close relation to the former point, automated grading and evaluation are commonly embedded within the broader environment of online learning platforms and their specific logics. In the case of the research illustrated here, the school under study used *Moodle* as their Learning Management System (LMS). As such, the features and logics according to which Moodle orders and allows certain activities (or not), substantially preconfigures automation processes around grading.

Lastly, as with other EdTech implementation contexts, digitizing and automating evaluation inevitably results in a redistribution and reconfiguration of tasks between (here) the teacher and the platform through which the automated grading is mediated. Consequently, even though these redistributions might suggest an increased efficiency in one task, such as

‘faster’ or ‘fairer’ grading, they simultaneously reorganize teachers’ time and effort in other aspects of their work – as such reshuffling the time they spend on different tasks, but not necessarily freeing it up.

## III. STUDY DESIGN

As a theme of the SMASCH project, automation of evaluation – and automated grading in particular – was brought up by one of the schools that was, at that time, implementing such automated grading for the end-of-year examination; that is, the summative assessment of students. In the study, we conducted an in-depth ethnographic investigation, tracing teachers’ evaluative practices as they worked with Moodle to better understand the reconfigured socio-technical interrelations between teachers and the platform. While some teachers at that school were already using digital tests for formative assessments and keen to expand this experience to summative assessment contexts, others were more worried about how digital evaluation could impact the main summative evaluation approach of the school. For these reasons, the school decided to in the first year give teachers the choice between the digital, the paper-based, or a hybrid form of evaluation. This diversity and multimodality of different approaches towards evaluation, in turn, formed a rich and complex research basis, as multiple evaluative practices could be observed and compared.

To address the plurality of evaluative practices, we split the research into two sub studies with two different foci. In the first study, we asked ten teachers to record their screens (in case they were working digitally with the test) and fill out a diary whenever they worked on one digital test in moments of 1) making a digital test, 2) students taking the test, 3) grading the test, and, ultimately, 4) providing feedback to students. The diary was aimed to document time and place of working on a test, as well as the interruptions and the off-screen materials deployed, and activities conducted. On the one hand, these screen recordings allowed teachers to stay close to their usual work routines. On the other hand, the recordings provided detailed insights into how teachers’ activities over time became redistributed whilst working with digital tests. The video recordings were afterwards annotated and were followed by a group conversation with teachers, in which we presented our analysis and further discussed, and deepened our understanding of, their practices.

In the second sub study, we interviewed ten teachers during the final examination period to focus even more explicitly on the aspect of grading. Two of the ten teachers hereby did not use digital evaluation for the summative assessment. The interviews were accompanied by ethnographic observations during the examination period. The combined methods of ethnographic presence and interviews allowed for meticulous investigation of teachers’ practices, linking moments of (on-screen) grading to a broader practice of evaluation that was happening in the school at that moment [17].

## IV. FINDINGS

The findings of the study can be broadly structured in two parts, both illuminating different aspects of how teachers’ work becomes reconfigured through automated evaluation and grading. While one set of findings shows how this reconfiguration happens when teachers *make tests for later automation on the platform*, the second set of findings focuses more on teachers *supervising the machine’s actual grading and feedback provision practices*.

### A. *When a test becomes a data object: teachers producing tests for automation*

In order to disentangle how automation of evaluative practices (such as grading) becomes enacted, and how it impacts the work of teachers, we first scrutinized how automatable tests are being made on the Moodle platform. Analysing teachers' diaries hereby demonstrated the many implications that come with the challenge of 'datafying' test questions. This means that for tests to become automatable, teachers need to transfer individual test questions into data objects. Data objects are hereby understood as material-semiotic things that carry (relational) meanings and properties but are simultaneously bound to the material conditions in which they are being produced [18]. More specifically, in the Moodle platform, 'question data objects' must be assigned with different properties, including a visual tag indicating the *type of answers* a student can give, namely a word, a number, a text, etc. Another group of tags categorizes the *type of knowledge or competence* the question addresses, offering choices between understanding, self-reflection, analysis, etc. The data object also includes the number of times a question has been used in previous tests (if the question is completely new, that property is automatically zero), or the distribution of points (i.e. tags of grading already inscribed into the data object).

As these different examples show, the properties which teachers need to assign to questions when preparing them for automation – including significantly *pre-structuring the grading practices* – turn those questions into complex and multi-layered data objects. When analysing our screen recording material, we indeed noticed that teachers dedicated an extensive amount of time to (re)tagging questions. Since (only) successfully created questions could be stored in a question pool and (re)used for future tests, the laborious work of tagging formed the precondition for test questions to move through time (e.g. being reused later) and space (e.g. being used by other teachers in other schools). Bates and colleagues (2016) denote this aspect as the *mutability of data objects*, which points to the re-purposing, adaptation, and remixing of the data objects for different ends. Here, the intimate connection between teachers' data work and the mutability of questions as data objects is rendered clearly visible.

After having tagged and categorized the test question, the teachers need to add the answers to the question data object, as well as define the feedback students would receive for each answer. The following video annotation demonstrates an example of such a practice, which in this case is a teacher who is actually reusing an already existing question for her new test, for which she recodes the answers: "The question appears on the screen as a mix of codes and texts, which is due to the fact that the teacher has a specific interface view that is different from what students eventually get to see. As the question is in the editing mode, the teacher adds a sub-question: 'What is the British way of writing the date below? -31/09'. After all the tagging and categorization work, this is the first time the full text of a question appears on the screen. Responses to the questions are coded in front of the question text. The teacher searches for the formulation of the code in other questions, copies it and finally pastes and adds the correct answer for her newly added question. She then codes other possible answers to that question and attributes half points to those answers. Finally, she codes in the American style of the date as a wrong answer and adds a code which sends a feedback comment to those who wrote this response". By adding responses in form of code, the teacher again makes it possible for the question to

be automatically corrected later. Moreover, in cases where answers can be written in different ways, the teacher predicts and codes *possible* answers into the platform, while equally attributing points each answer would get.

As noted above, teachers equally include templated responses – i.e. automated feedback – that students will receive immediately after the test. In doing so, however, the actual characteristic of feedback changes: while feedback in the original sense of the word is only possible after a particular student response has been given, in the automated grading case, feedback equally becomes premade and, in that sense, turns into anticipated, codable *feedforward*. For the teacher, it means that making a test not only includes creating (or adapting) automatable questions, but equally requires a reconfiguration of their own feedback practices into feedforward practices. This reshuffling simultaneously enacts a *temporal folding of making and judging the test*, condensed in the initial work of tagging and coding. This also means, however, that it is only after these time-consuming tasks that the immediacy and time-efficiency promise of automated evaluation and grading can happen. As such, the 'speed' allegedly inherent to the automatic grading mostly refers to when data objects, then, can be stored in the question pool for later reuse, and it is counterweighted by the many practices of making question data objects.

Lastly, teachers report that the intense work of data object creation strongly nudges them to reduce the number of open (not automatically gradable) questions in favour of short answers as well as multiple-choice questions. On the Moodle platform we observed, open questions cannot be fully graded automatically. The only thing teachers could do to accelerate grading of these questions, was to code in the platform the different words that they assigned as keywords to appear in the answer. These keywords, then, would be labelled correct if they were used in students' answers (see also next section). In turn, for the more sceptical teachers, it is exactly this push away from open questions that evokes them *not* to use digital evaluation altogether. As a geography teacher mentioned during the interview "If you use digital evaluation, you do it to save time, and you're mostly going to try things that the computer can correct by itself. So, I think the questions really change if you go for a full digital evaluation. You're going to automatically ask less questions in which the student needs to formulate an answer themselves from scratch, you're going to guide them much more".

In sum, we see how teachers' grading and feedback practices are, through automation, being reshuffled and to large parts integrated into the *making of question data objects*. This does not mean, however, that teachers are afterwards passively watching the algorithmic processing. Rather, we found that teachers intensively supervise the machine during grading and feedback practices. It is this practice to which we turn next.

### B. *Educating the machine: teachers supervising automated grading and feedback practices*

When it comes to the actual grading process, the Moodle interface divides questions of a test into two categories of automatically graded questions and questions that need manual grading. This way, the interface 'orders' teachers' tasks in terms of what is to be graded. For automatically gradable questions, the Moodle interface assigns question-answer-blocks with a colour coding in order to create a visual distinction between correct answers (green), partially correct answers (orange), and wrong answers (red) (see also [19]).

To begin with, all teachers of our sample unanimously reviewed all red and orange colours to make sure the system made a fair judgement. In this context, we were surprised to see how often teachers, then, *changed the judgement of the system*, for instance by unmarking spelling mistakes flagged as false answers by Moodle. In such cases of rejudging, the teachers referred to their in-depth knowledge about the students: “In this example, I gave a full point to one student, even though he wrote the wrong letter L. This is something that, for some students, I would still give a ‘zero’. But with him, I made it into a ‘one’ point, because he is a student with another language at home, for example... and that’s something that... if it’s only one letter wrong and it more or less sounds the same, I still give the point”. This example by the history teacher shows how the teacher’s rejudging practice deliberately refrains from accuracy and objectivity, but is rather legitimized by referring to contextual, professional knowledge that the teacher has gained over time from her students, and which does not fit into the automated grading logic [20]. In turn, the next example illuminates how teachers also try to find ways to *educate the machine* about ‘their way’ of grading, that is, to bring more human and educational professional judgement into the system. “I can also code a lot of correct answers for example, like, [I can] code one, two, three, four, five, correct answers, and two additional answers which I grade 50% of a point. I add these right answers during the correction phase as well. Because I see that the first ten students gave a bit of a different answer than I predicted, but these answers are correct as well, I insert those answers as correct, so that the next students with same answers automatically get the point. The more the computer can do, the less I have to do”.

Here, we see an interesting dynamic, namely that teachers use given students’ answers that do not match the prestructured ones in order to *feed back*, yet not to the students, but to the question’s data object, here by coding spelling mistakes as correct answers. On the one hand, this serves for the future use of the question in other tests; on the other hand, it educates the machine to incorporate the teachers’ pedagogical ideas and their sense of professional judgement into the platform.

Related to this active correction of machine judgement, we also identified struggles about the (re)distribution of responsibility between the teacher and the platform. A math teacher expressed that “I sometimes feel a little bit guilty that I don’t check every question, because I feel somehow that I should do that. I struggle between the feeling of it’s a lot of work that I don’t get anything out of it, and the sense of responsibility as a teacher to have to check every answer. On the one hand, I find it a lot of work to open every student’s test and to look at it when it is already corrected automatically. On the other hand, I think that it’s my responsibility to do that, so that’s a double feeling. I feel that I am delegating my responsibility to the machine and so I am happy with the paper part [that is not automatically gradable, authors], because that’s the one I really correct, and then I feel better about myself for not checking automatically graded questions once again myself”.

As the quote shows, the teacher problematizes that the responsibility for the evaluation cannot be delegated to the platform. Reviewing all the partially correct and incorrect graded answers equally signifies how teachers take up the responsibility for grading according to their own pedagogical ideas and principles. Interestingly, the result is that the only part of correction that does *not* get reevaluated by the teacher is correctly

marked answers. It is such correct answers that are being filtered out before the teacher’s intervention comes into play. As such, even though the embeddedness of automated processes in grading students’ examination suggests a distribution of responsibility from the teacher to the machine, allegedly prioritizing the automatic decisions over the professional judgement of teachers [21], we see how teachers strongly resist this prioritization by reviewing the work of the machine and by inserting their pedagogical ideas and values into it, *yet only in relation to the incorrectly marked answers*.

A last observation in the context of teacher practices during the grading process refers to those remaining open questions that teachers sometimes still put into a test, but that cannot, at least not in the used Moodle system, be graded automatically. Interestingly, we can still observe that the logic of automation affects those grading practices. More specifically, the Moodle interface allows to order answers per question, thus displaying all answers to one question on one page. The idea hereby is that teachers can scroll and comparatively grade open answers per question for all students simultaneously. While this automated reordering of answers, in the view of the teachers, appears as speeding up the grading process, the digital format simultaneously requires teachers to reconfigure how they practice grading. While in the case of paper tests, teachers commonly use spontaneous text annotations such as underlining, quick jotting down of notes in the margins, etc., they now need to laboriously insert text boxes for each response the students give. This, again, nudges teachers to use the templated, automatable feedback structure available while *making* the test, rather than the individual feedback while *correcting* the test. In other words, also in the case of open questions, the platform clearly triggers an efficiency-oriented view on feedback provision.

## V. DISCUSSION AND OUTLOOK

The aim of this chapter was to provide insights into the dtec.bw project SMASCH (Smart Schools), which aim it is to approach digitization of school organisations not from a technology-centred, but rather from a more holistic (whole school approach), critical-analytical perspective. The open project design hereby facilitated to generate a broad scope of different research areas that navigate between practical problems or needs of the project schools on the one hand, and researchers’ interests on the other hand. One of these research areas has hereby emerged around matters of automated evaluation and grading.

Exemplifying this research around automated evaluation and grading in this chapter, we showed how EdTech implementation has evoked a substantial reconfiguration of teachers’ work. However, rather than diminishing teachers’ workload, we argued that teachers’ practices are being reshuffled in orchestration with *multiple and very specific kinds of automation* that take shape through (in our case) the Moodle platform. As a result, specific types of question-making as well as evaluative practices are favoured, whilst others are clearly discouraged. In contrast to these *substantial impacts of the technology on teachers*, the findings also clearly indicate that *teachers actively educate and influence the machine* in how to execute particular tasks [22]. In this sense, it can be argued that teacher subjectivities emerge that are not the only source of evaluation and judgement (anymore), but that (still) remain responsible for that evaluation. This, however, blurs the line between repetitive and administrative aspects of grading on the one hand, and an *educationally meaningful* practice of evaluation on the

other hand, which is based on teachers' professional judgement and sensemaking [23], [24].

We would like to conclude this chapter by reflecting on potential implications of our study with regards to the more recent upsurge of GenAI (Generative Artificial Intelligence) technologies. The recent breakthrough of applications such as ChatGPT, Google Gemini, or Microsoft Copilot, which allow to produce 'conversations' with students and teachers based on a combination of Large Language Models (LLMs) and prompting techniques, undoubtedly marks a substantial chance of educative practices, whether they take place in- or outside of schools. Over the last years, various GenAI-based applications have been adapted to, and developed further for, the education system, discursively framed by well-known promises and hopes that strongly echo the automation discourse as we have outlined in this chapter (e.g. relieving teachers' workload, fostering personalized learning, increasing the fairness of assessment, etc.) [25]. At the same time, many critical voices have been raised that call for a careful adoption of AI in education, or even suggest abolishing it altogether [26], [27].

Against the background of our study on automated evaluation, we resonate such critical voices in the sense that 'hopes' of automation and (now) AI commonly underestimate the complex character of educational practices, which change through, and simultaneously resist, technology in multiple ways. This multidimensionality can only be understood by carefully disentangling the specific moments, processes or dimensions of automation (e.g. task making, task reusage, grading, etc.). Impacts of these changes and resistances are hereby neither good nor bad as such, but rather always carry with them positive *and* negative impacts [28]. In other words, the better we understand these manifold reconfigurations, the better we can estimate the specific demands on how GenAI systems should at all be implemented within educational practices [29]. What has already become clear, and what this paper has equally shown, is that the laborious work of reviewing technological output is unlikely to diminish, but rather seems to further increase when teachers are expected to stay responsible for educationally relevant and pedagogically meaningful judgements.

#### ACKNOWLEDGEMENTS

This research was funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] F. Stadler and C. Kuttner, "Schule in der Kultur der Digitalität – Schule als Reflexionsraum. im Gespräch mit Felix Stalder," in *Praxistheoretische Perspektiven auf Schule in der Kultur der Digitalität*, vol. 62, Springer, 2022, pp. 3-19.
- [2] S. Hartong, A. Loft-Akhoondi, N. Brandau, B. Junne, I. Czarnojan and T. Scheytt, "Auf dem Weg zur Digitalität in Schule: Zwischenbericht über Interventionen und Explorationen im Rahmen des Projekts "SMASCH - Smarte Schulen"," *Working Papers*, vol. 10, 2023.
- [3] N. Brandau and S. Alirezabeigi, "Critical and participatory design in-between the tensions of daily schooling: working towards sustainable and reflective digital school development," *Learning Media and Technology*, vol. 48, no. 2, pp. 337-349, 2022.
- [4] F. Macgilchrist, "Cruel optimism in edtech: When the digital data practices of educational technology providers inadvertently hinder educational equity," *Learning, Media and Technology*, vol. 44, no. 1, pp. 77-86, 2019.
- [5] C. Adams and S. Groten, "A TechnoEthical Framework for Teachers," *Learning, Media and Technology*, pp. 1-18, 2023.
- [6] M. Decuyper, E. Grimaldi and P. Landri, "Introduction: Critical studies of digital education platforms," *Critical Studies in Education*, vol. 62, no. 1, pp. 1-16, 2021.
- [7] N. Kerssens and J. van Dijck, "Governed by edtech? valuing pedagogical autonomy in a platform society," *Harvard Educational Review*, vol. 92, no. 2, pp. 284-303, 2022.
- [8] S. Livingstone, K. Pothong, A. Atabey, L. Hooper and E. Day, "The Googilization of the classroom: Is the UK in protecting children's data and rights?," *Computers and Education*, 2024.
- [9] C. Ortegón, M. Decuyper and B. Williamson, "Mediating educational technologies: Edtech brokering between schools, academia, governance, and industry.," *Research in Education*, 2024.
- [10] I. Martínez Lunde, *Digitization in School Leadership and Educational Governance*, 2022.
- [11] H. Reinius, I. Kaukinen, T. Korhonen, K. Juuti and K. Hakkarainen, "Teachers as transformative agents in changing school culture, , Volume 120, 2022.," *Teaching and Teacher Education*, vol. 120, 2022.
- [12] N. Brandau, "Was steckt in Lernmanagementsystemen und welche Implikationen ergeben sich hieraus für eine digital mündige Gestaltung im Schulkontext?: Eine Studie zu LMS.lernen.hamburg," in *Digitale Lerntechnologien. Von der Mystifizierung zur reflektierten Gestaltung von EdTech*, Digitale Lerntechnologien, 2024, pp. 63-86.
- [13] T. Tierens, M. Decuyper, S. Alirezabeigi and S. Hartong, "Governing through time(s): temporal modes of governance in digital education policy," *British Journal of Sociology of Education*, pp. 1-16, 2024.
- [14] S. Hartong and A. Renz, *Digitale Lerntechnologien: Von der Mystifizierung zur reflektierten Gestaltung von EdTech*, Bielefeld: transcript verlag, 2024.
- [15] M. Decuyper, S. Hartong, N. Brandau, L. Joecks, A. Loft-Akhoondi, C. Ortegón and L. Vanerman, "Maneuvering constellations of valuation: a critical investigation of the edtech startup sector," *Critical Studies in Education*, pp. 1-20, 2024.
- [16] R. Zhuang, Y. Zhuang, D. Zou, X. Qin and P. L.H. Yu, "AI-assisted automated scoring of picture-cued writing tasks for language assessment," *Education and Information Technologies*, vol. 28, pp. 7031-7063, 2022.
- [17] N. Selwyn, T. Hillman, A. Bergviken-Rensfeldt and C. Perotta, "Making Sense of the Digital Automation of Education. In *Postdigital Science and Education*," *Postdigital Science and Education*, vol. 5, pp. 1-14, 2022.
- [18] J. Law, "Material Semiotics," 2019. [Online]. Available: [www.heterogeneities.net/publications/Law2019MaterialSemiotics.pdf](http://www.heterogeneities.net/publications/Law2019MaterialSemiotics.pdf). [Accessed 10 06 2024].
- [19] H. Friis Ratner, "Automation, agencies and aesthetics: the politics of data visualizations in configuring teachers' expertise," *Discourse: Studies in the Cultural Politics of Education*, 2024.
- [20] C. Perrotta, "Afterword: Platformed professional(itie)s and the ongoing digital transformation of education," *Tertium Comparationis*, vol. 29, no. 1, pp. 117-130, 2023.
- [21] B. Williamson, K. N. Gulson, C. Perrotta and K. Witsenberger, "Amazon and the New Global Connective Architectures of Education Governance," *Harvard Educational Review*, vol. 92, no. 2, pp. 231-256, 2022.
- [22] C. Perrotta, K. N. Gulson, B. Williamson and K. Witsenberger, "Automation, APIs and the distributed labour of platform pedagogies in Google Classroom," *Critical Studies in Education*, vol. 62, pp. 97-113, 2021.
- [23] N. Wagener-Böck, F. Macgilchrist, K. Rabenstein and A. Bock, "From Automation to Symmation: Ethnographic Perspectives on What Happens in Front of the Screen," *Postdigital Science and Education*, vol. 5, pp. 136-151, 2023.
- [24] S. Hartong and M. Decuyper, "Editorial: Platformed professional(itie)s and the ongoing digital transformation of education," *Tertium Comparationis*, vol. 29, no. 1, pp. 1-21, 2023.

- [25] R. Luckin, M. Cukurova , C. Kent and B. du Boulay, "Empowering educators to be AI-ready," *Computers and Education: Artificial Intelligence*, vol. 3, 2022.
- [26] G. Ritzer, J. M. Ryan, S. Hayes, M. Elliot and P. Jandrić , "McDonaldization and Artificial Intelligence," *Postdigital Science and Education*, 2024.
- [27] B. Williamson, A. Molnar and F. Boninger, "Time for a Pause: Without Effective Public Oversight, AI in Schools Will Do More Harm Than Good," National Education Policy Center, , 2024.
- [28] S. Hartong and I. Sander, "Critical Data(Fication) Literacy In und Durch Bildung," in *WHITEPAPER DATENKOMPETENZ*, Weizenbaum Institut, pp. 19-21.
- [29] W. Holmes, . K. Porayska-Pomsta, K. Holstein, E. Sutherland, T. Baker, S. Buckingham Shum, O. C. Santos, M. T. Rodrigo, M. Cukurova, . I. I. Bittencourt and K. R. Koedinger , "Ethics of AI in Education: Towards a Community-Wide Framework," *International Journal of Artificial Intelligence in Education*, vol. 32, no. 3, p. 504–526, 2022.

# Digitalisierung kommunaler Verwaltungen – ein drei Säulen Modell

Thomas Kuder  
vhw – Bundesverband Wohnen und  
Stadtentwicklung  
Berlin, Deutschland  
tkuder@vhw.de

Gary S. Schaal\*, Christina Schaefer,  
Sebastian Dumm, Anne Hoss,  
Bastian Manteuffel  
Helmut-Schmidt-Universität/  
Universität der Bundeswehr Hamburg  
Hamburg, Deutschland  
\*gschaal@hsu-hh.de

Gero Serfas  
Q Agentur für Forschung  
Mannheim, Deutschland

**Kurzfassung** – Die digitale Transformation kommunaler Verwaltungen in Deutschland verläuft langsam und wird insbesondere im internationalen Vergleich stark kritisiert. Der Beitrag stellt das Forschungsprojekt „Smarte Systeme“ vor, das darauf abzielt, Erfolgsfaktoren für eine effektive Digitalisierung kommunaler Verwaltungen zu identifizieren. Anhand einer Fallstudie zur Digitalisierung des Bauamts in Ludwigsburg wird ein Mixed-Methods-Ansatz verwendet, der Theorie und empirische Analyse kombiniert. Das Projekt betont die Bedeutung eines ganzheitlichen Ansatzes, der smarte Systemarchitektur, smarte Verwaltungsarchitektur und smarte Mehrebenen-Governance integriert, um digitale Souveränität auf kommunaler Ebene zu erreichen. Die Ergebnisse zeigen, dass es keine universelle Lösung gibt, aber ein differenziertes analytisches Rahmenwerk strategische Entscheidungen im Transformationsprozess erleichtern kann. Das normative Ziel des Projektes ist es, das Potenzial der Digitalisierung, das über die bloße Umwandlung analoger Prozesse in digitale hinausgeht, auszuschöpfen. Der Beitrag fokussiert auf die Chancen und Herausforderungen der smarten Multilevel-Governance für kommunale Verwaltungsdigitalisierung.

**Stichworte** – Verwaltungsdigitalisierung, Multilevel-Governance, smarte Systemarchitektur, digitale Transformation, digitales Bauamt

## NOMENKLATUR

EfA	Einer für Alle
TVL	Tarifvertrag für den Öffentlichen Dienst der Länder
FONA	Strategische Forschungs- und Innovationsagenda

## I. EINLEITUNG

Die digitale Transformation hat in den vergangenen Jahren auch die deutsche Kommunalverwaltung erfasst und stellt diese vor vielfältige Herausforderungen [1]. Die mangelnde Geschwindigkeit der digitalen Transformation kommunaler Verwaltungen und ihrer Dienstleistungen in Deutschland ist dabei Gegenstand wissenschaftlicher Kritik – auch und gerade im internationalen Vergleich [2], [1], [3], [4], [5], [6], [7].

Das dtec.bw-Forschungsprojekt *Erfolgsfaktoren Smarter Systemarchitektur für kommunales E-Government (Smarte Systeme)* ist vor diesem thematischen Hintergrund verortet.

Als interdisziplinär orientiertes Forschungsprojekt ist es an den Schnittstellen zwischen Verwaltungs-, Politik- und Sozialwissenschaften, Anthropologie und Stadtplanung verortet. Sein Ziel ist es, die Nutzungsqualität und -intensität von E-Government und digitaler Verwaltung auf lokaler Ebene zu erhöhen, um so einen Beitrag zur Steigerung der digitalen Souveränität Deutschlands auf kommunaler Ebene zu leisten.

Hierzu wurde versucht, Erfolgsfaktoren und Gelingensbedingungen kommunaler Verwaltungsdigitalisierung anhand eines Praxisbeispiels, der Digitalisierung des Bauamts in Ludwigsburg, zu identifizieren (vgl. mit Blick auf die OZG-Leistungen insgesamt [8], [4]). In einem rekursiven Prozess der Vermittlung von Theoriebildung und empirischer Analyse konnte im Projekt unter Nutzung eines Mixed-Methods-Ansatzes ein erstes Analyse- und Beratungsdesign für kommunale Verwaltungsdigitalisierung entwickelt werden, das aktuell noch validiert wird.

Zusammenfassend existiert zwar kein Patentrezept für erfolgreiche Verwaltungsdigitalisierung auf kommunaler Ebene, keine One-size-fits-it-all-Lösung, aber ein differenziertes Raster von Faktoren und strategischen Entscheidungen, die die Transformation erleichtern können.

Erfolgreiche kommunale Verwaltungsdigitalisierung umfasst für das Projekt *Smarte Systeme* deutlich mehr als die bloße Überführung analoger in digitale Prozesse. Normatives Ziel eines Digitalisierungsprozesses sollte es sein, das Potenzial der Digitalisierung zu erschließen. Hierfür ist ein erweitertes Verständnis notwendig, das wir als *Digitalität* bezeichnen. Es umschließt jene Mehrwerte, die nur aus *digitaler* Verwaltung resultieren können und die sich nur dann einstellen, wenn nicht nur analoge in digitale Prozesse eins-zu-eins umgesetzt werden. *Digitalität* bezeichnet somit die emergente Qualität eines *Outputs*.

Die Voraussetzung hierfür ist eine *smarte* Systemarchitektur. **Smartness ist im Rahmen unseres Ansatzes eine emergente Systemeigenschaft, die auf smarterer Digitalisierung basiert, sich aber nicht in ihr erschöpft.** Nur Smartness auf der Ebene der Systemarchitektur kann zu Digitalität kommunaler Verwaltungsdigitalisierung führen.

Unser Ansatz geht von der These aus, dass eine *digitale Systemarchitektur* drei Säulen, die sowohl politisch als auch wissenschaftlich häufig isoliert betrachtet wurden und werden,

holistisch zusammenführen muss, um kommunale Verwaltungsdigitalisierung erfolgreich aufzusetzen und genuine digitale Mehrwerte zu generieren (vgl. [9] für eine theoretische inspirierte Diskussion des Konzepts der Architektur in diesem Kontext). Diese drei Säulen sind a) *smarte Systemarchitektur* (technische Systeme (IT)), b) *smarte Verwaltungsarchitektur* (organisationale Struktur, MitarbeiterInnen-Motivation und deren Digitalkompetenz) sowie c) *smarte Multilevel-Governance*.

Um die vielschichtigen Herausforderungen zu bewältigen, müssen die Kommunen einen ganzheitlichen und kooperativen Ansatz querschnittlich über die drei Säulen verfolgen, der eine enge Zusammenarbeit zwischen verschiedenen Interessengruppen, einschließlich aller Governance-Ebenen, PartnerInnen aus dem Privatsektor und BürgerInnen, erfordert. Wenn sie die Governance-, Technologie- und Organisationale-/Humankapitalaspekte der kommunalen Digitalisierung angehen, können die Kommunen in Deutschland das volle Potenzial der digitalen Transformation ausschöpfen und die Effizienz, Reaktionsfähigkeit und Widerstandsfähigkeit der kommunalen Dienstleistungen verbessern (vgl. bereits [10]).

Im Folgenden wird der theoretische und konzeptionelle Ansatz dargestellt (II). Im Anschluss erfolgt ein empirischer Deep Dive in die Multilevel-Governance anhand unserer Fallstudie Ludwigsburg (III). Auf dieser Grundlage werden abschließend erste Gelingensbedingungen kommunaler Verwaltungsdigitalisierung im Feld der Multilevel-Governance identifiziert (IV). Der Aufsatz schließt mit einigen Ausführungen zu weiteren Forschungsbedarfen (V).

## II. KONZEPTIONELLER UND METHODISCHER ANSATZ

### A. Datengrundlage

Die Datengrundlage für unsere hier präsentierten Überlegungen bilden Interviews mit Gemeinderäten der Stadt Ludwigsburg sowie mit ausgewählten Fachbereichs-, Referats- und Dezernatsleitungen, die am Projekt „Digitales Bauamt“ beteiligt waren. Hinzu kommen Interviews mit ausgewählten Akteuren, die in die bundesweiten Aktivitäten zur Verwaltungsdigitalisierung einbezogen waren, darunter Akteure aus Verbänden, Dachverbänden, Landesministerien und nachgeordneten Einrichtungen, die im Herbst 2021 bis Sommer 2022 geführt wurden (vgl. [11]). Den InterviewpartnerInnen wurde Anonymität zugesagt, deshalb werden sie im Folgenden mit Kürzeln zitiert. In Anlehnung an die Grounded Theory [12] erfolgte die Erarbeitung von Forschungsergebnissen durch Erhebung von Interviewdaten, durch Kategorienbildung sowie durch deren wechselseitige Weiterführung bzw. -entwicklung (vgl. auch grundlegend zur Methodik [11]).

### B. Konzeptioneller Rahmen

Die Analyse der Digitalisierung von Verwaltung ist national wie international ein etabliertes Forschungsfeld [13], das aufgrund seiner Querschnittlichkeit in diversen Disziplinen verortet ist und jeweils unter spezifischen Erkenntnisinteressen adressiert wird, so u.a. in der Verwaltungswissenschaft als Verwaltungsmodernisierung [13], [14], [15], in der Politikwissenschaft als Multi-Level-Governance Herausforderung [16], [17], [18], [20] und im Digitalisierungsdiskurs unter dem Label „Smart City“ [3].

Ausgehend von einer systematischen Sichtung der thematisch relevanten akademischen Literatur und Diskurse konnte das Projekt die zentralen Erfolgsfaktoren für die kommunale Verwaltungsdigitalisierung deduktiv herausarbeiten. Diese

theoretisch abgeleiteten Faktoren wurden dann in einem weiteren Schritt mit den empirischen Erkenntnissen aus den durchgeführten Interviews kontrastiert und ergänzt. Auf diese Weise wurde ein ganzheitliches Verständnis der Gelingensbedingungen für eine erfolgreiche Digitalisierung der kommunalen Verwaltung entwickelt.

Auf einer Makroebene lässt sich festhalten, dass die drei oben genannten „Säulen“ erfolgreicher kommunaler Verwaltungsdigitalisierung in der einschlägigen Literatur und auch in den Interviews aufgegriffen werden. Allerdings findet sich dort oftmals eine „versäulte“ Betrachtungsweise, die der jeweiligen disziplinären Verortung geschuldet ist und meistens auf die IT-Dimension abzielt. Angesichts der zentralen Rolle der IT-Infrastruktur für die Digitalisierung der Verwaltung ist diese Fokussierung gerechtfertigt.

Zugleich vernachlässigt eine solche Priorisierung jedoch, dass Digitalisierung sich nicht auf Technologie oder Technik reduzieren lässt. Viele Science- und Technology-Studies verweisen eindrucksvoll darauf, Digitalisierung als *sozio-technisches System* mit *Prozesscharakter* zu konzeptualisieren [19]. Entsprechend liegt der Fokus unserer interdisziplinären Forschungsarbeit darauf, den *emergenten Mehrwert* einer *holistischen Konzeptualisierung* von kommunaler Verwaltungsdigitalisierung herauszuarbeiten.

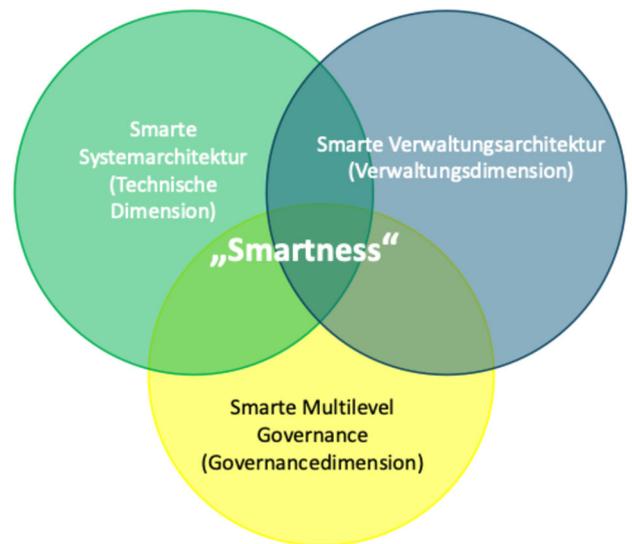


ABBILDUNG 1: SMARTNESS DER SMARTEN DIGITALEN VERWALTUNGSSYSTEMARCHITEKTUR.

Nachfolgend werden die Chancen und Herausforderungen der drei Säulen einzeln erörtert. Anschließend folgt eine vertiefende Analyse der Multilevel-Governance-Säule.

#### 1) *Smarte Systemarchitektur (IT)*

Basal für den Erfolg kommunaler Verwaltungsdigitalisierung ist die IT-Infrastruktur. Diese umfasst sowohl die technische Dimension (Hardware) als auch die Software und die Spezifikation von Schnittstellen, um Interoperabilität zu gewährleisten. Herausfordernd ist dabei, dass sich die Kommunen mit den *rasanten* und kontinuierlichen Fortschritten der digitalen Technologien auseinandersetzen müssen, die ein umfassendes Verständnis und die strategische Integration einer breiten Palette von digitalen Lösungen erfordern (vgl. Organisationstruktur und Digitalisierungskompetenz unten).

Auf technischer Ebene erfordert die Digitalisierung umfassende Investitionen in die Modernisierung der IT-Infrastruktur, die Einführung von E-Government-Lösungen sowie die Gewährleistung von Datenschutz und Cybersicherheit, sodass die Resilienz kommunaler Verwaltung und deren Dienstleistungen gestärkt wird. Eine performante IT-Infrastruktur benötigt zudem ein Datenmanagement, das kontinuierlich gepflegt wird, denn der Mehrwert kommunaler Verwaltungsdigitalisierung erschließt sich über die Digitalisierung vormals analoger Daten, deren Zusammenführung und Analyse. Datenschutzrechtliche Aspekte können hierbei für die Datenanalyse eine Herausforderung darstellen.

Die Finanzierung der IT-Infrastruktur – Hardware, Lizenzverträge, Wartung, etc. – kann insbesondere kleinere Kommunen vor große Herausforderungen stellen, da es sich um kontinuierliche Kosten handelt, da man – was insbesondere vor dem Hintergrund von wachsender Cyberkriminalität und hybriden Angriffen zwingend notwendig erscheint – keine veralteten Systeme – sei es in Hard- oder Software – nutzen sollte.

Für eine erfolgreiche Digitalisierung der kommunalen Verwaltung sind drei Schlüsselfaktoren entscheidend: 1. Investitionen in die geeignete technologische Infrastruktur; 2. Regelmäßige Instandhaltung dieser Infrastruktur; 3. Sicherstellung der Interoperabilität zwischen den verschiedenen Systemen.

In diesem Kontext muss schließlich eine zentrale Motivation der Verwaltungsdigitalisierung auf kommunaler Ebene kritisch adressiert werden: Die Einführung von Digitalisierung wird häufig – zumindest „politisch“ – mit Kostenersparnissen begründet [20]. Unsere Forschung deckt diese Hoffnung nicht. Zumindest kurz- bis mittelfristig resultieren aus Digitalisierungsprozessen für die Kommunen eher *Mehrkosten*, da mehrere Strukturen parallel entwickelt und betrieben werden müssen (analoge und digitale Technik; alte und neue IT, etc.). Da sich die Updatezyklen mit hoher Wahrscheinlichkeit verkürzen werden, steigen auch die kontinuierlichen Infrastrukturkosten. Mittel- und langfristig ist Verwaltungsdigitalisierung trotzdem „alternativlos“, da auch die öffentliche Verwaltung in Zukunft vor einer zunehmend größeren Herausforderung steht, Personal zu akquirieren. Hierzu ist jedoch Digitalisierung allein nicht ausreichend – auf der Ebene der Systemarchitektur muss die Nutzung von künstlicher Intelligenz bereits heute konstitutive Berücksichtigung finden.

## 2) *Smarte Verwaltungsarchitektur (Organisation und Humankapital)*

Über die technische Implementierung und Erneuerung der IT-Infrastruktur hinaus sind auch organisatorische Aspekte, das Führungsverhalten gegenüber den MitarbeiterInnen sowie deren digitale Weiterbildung entscheidende Erfolgsfaktoren für eine erfolgreiche digitale Transformation der kommunalen Verwaltung.

Verwaltungen sollen spezifische Leistungen erbringen. Im Idealfall folgt die Organisationsstruktur einer Verwaltung dem Prozess einer konkreten Leistungserbringung [19]. Daraus folgt auch, dass sich Organisationsstrukturen verändern sollten, wenn sich durch die Umstellung von Analog in Digital die Prozesse der Leistungserbringung verändern. Aus technischen Innovationen sollten mithin Veränderungen der organisationalen Struktur folgen. Dies bezieht sich sowohl auf die vertikale Dimension der Leistungserbringung *innerhalb* einer (Fach-)Abteilung einer Verwaltung als auch auf die *horizontale Di-*

*mension* der Koordination von Prozessen über (Fach-)Abteilungen hinweg. Horizontale und vertikale Dimensionen organisationaler Anpassung sollten gleichzeitig konzeptualisiert werden – dies ist in der empirischen Praxis jedoch selten der Fall. Zu den organisationalen Anpassungen gehört häufig auch die Implementation eines IT-Beauftragten oder einer IT-Abteilung. Hierzu zwei Einsichten: Erstens sind IT-Beauftragte Personen mit Querschnittsaufgaben und sollten auch institutionell entsprechend verortet sein. Zweitens bestehen Herausforderungen im Prozess der Verwaltungsdigitalisierung häufig im Bereich der „kommunikativen Anschlussfähigkeit“. Reine InformatikerInnen verstehen die Fachprozesse selten, während reine VerwaltungsmitarbeiterInnen selten über hinreichende Kenntnisse im Feld der IT verfügen. Benötigt wird also eine rare Doppelkompetenz, um Fachprozesse und IT anschlussfähig zu gestalten.

Damit ist der Bogen zu den VerwaltungsmitarbeiterInnen bereits geschlagen. Die erfolgreiche Umsetzung digitaler Initiativen hängt stark von den Fähigkeiten und Kompetenzen der kommunalen MitarbeiterInnen ab. Unsere Interviews – auch jenseits von Ludwigsburg – haben verdeutlicht, dass die MitarbeiterInnenmotivation und deren digitaler Kompetenz häufig entscheidende Variablen für die Tiefe und den Umfang konkreter Digitalisierungsprozesse darstellen. Oberflächliche Digitalisierung erzeugt jedoch Pfadabhängigkeiten (ggf. sogar *critical junctures* im Sinne des Neo-Institutionalismus), die – auf lange Sicht – nicht nur mit höheren Kosten als Digitalisierung „aus einem Guss“ einhergehen, sondern auch anspruchsvolle kommunale Digitalisierung massiv zeitlich verzögern kann, sodass Kommunen langfristig ins Hintertreffen geraten können, wenn sie kurzfristig zu adaptiv auf vorhandene (begrenzte) Humankapitalressourcen eingehen.

Kommunen stehen jedoch vor der Herausforderung, Talente mit dem erforderlichen digitalen Fachwissen zu gewinnen und zu halten, da sie auf einem hart umkämpften Arbeitsmarkt mit dem privaten Sektor und anderen Organisationen konkurrieren müssen. Weiterbildungs- und kontinuierliche Lernprogramme für bestehende MitarbeiterInnen sind von entscheidender Bedeutung, um die digitale Kompetenzlücke zu schließen und eine Kultur der Innovation und Anpassungsfähigkeit innerhalb der kommunalen Belegschaft zu fördern. In der Praxis erweist sich der TVL und seine nicht mehr aktuellen Stellenbeschreibungen als Nadelöhr für die Akquise von neuen MitarbeiterInnen.

## 3) *Multilevel-Governance*

Während Bundes- und Landesregierungen Initiativen zur Förderung der Digitalisierung ergriffen haben, liegt die Umsetzung und Koordinierung dieser Bemühungen häufig auf den Schultern der Kommunen. Die Sicherstellung einer effektiven Zusammenarbeit zwischen den verschiedenen Ebenen, die Abstimmung von Strategien und die Sicherstellung der notwendigen Finanzierung und Ressourcen, um die digitale Transformation voranzutreiben, kann für Kommunen eine gewaltige Aufgabe sein. Auf der Ebene der kommunalen Governance müssen Städte und Gemeinden ihre Strukturen, Prozesse und Entscheidungswege an die veränderten Rahmenbedingungen anpassen.

Governance in Mehrebenen-Systemen entsteht, wenn sich politische Prozesse durch Aufteilung von Macht und Kompetenzen auf mehr als nur eine räumlich-administrative Ebene beziehen, z. B. die Gebietskörperschaften Bund, Länder und Kommunen in der Bundesrepublik Deutschland. Sind diese nicht unabhängig voneinander oder sind diese z. B. über die

hohe Bindewirkung von Gesetzen eng miteinander verflochten, müssen Entscheidungen aufgrund von gegenseitigen Interdependenzen und Verflechtungen zwischen den Ebenen sorgfältig koordiniert werden. Dabei entstehen in der Regel komplexe und intergouvernementale Konfigurationen aus institutionellen Regelsystemen, die auf den verschiedenen beteiligten Ebenen jeweils den eigenen Funktionslogiken folgen und deren Handlungsoptionen bestimmen [21].

In der Politikwissenschaft findet sich die Position, dass es aufgrund dieser komplexen Mehrebenen-Verflechtungen vor allem in Reformprozessen zu erheblichen Steuerungsdefiziten, Blockaden, suboptimalen Entscheidungen oder fehlenden Bindewirkungen kommen kann [22]. Daraus lässt sich für komplexe Prozesse der Transformation, wie im Falle der Digitalisierung der kommunalen Verwaltung, ein besonderes Forschungsinteresse ableiten, dem in diesem Beitrag beispielhaft nachgegangen werden soll.

Ein auf die Mehrebenen-Verflechtung der Bundesrepublik für übertragbar erachtetes Analysekonzept für die internationale Politik ist der „linkage approach“ [23]. Dabei stehen die Kommunikation und das Handeln der politisch Verantwortlichen im Mittelpunkt der Betrachtung, die einerseits ihr Land international repräsentieren, andererseits aber auch für die Politik im eigenen Land die politische Verantwortung tragen [21]. In diesem Zwei-Ebenen-Spiel steht die Kongruenz der Kommunikation und des politischen Handelns auf den zwei beteiligten Ebenen auf dem Prüfstand. Dieses Konzept soll nachfolgend für die Bund-Länderebene sowie die kommunale Ebene nutzbar gemacht werden. Ziel der empirischen Anwendung dieses theoretischen Ansatzes ist es, die Digitalisierungsprozesse in der Stadt Ludwigsburg dahin gehend zu analysieren, ob Steuerungs- und Koordinationsdefizite zwischen den räumlich administrativen Ebenen von Bund, Ländern und Kommune bestanden, ob sich Ursachen und Folgen benennen und sich ggf. Ansätze für die Überwindung von Steuerungsdefiziten, Blockaden, suboptimalen Entscheidungen und Akzeptanzproblemen ableiten lassen.

In der wissenschaftlichen Literatur wurde die Relevanz der kommunalen Ebene bei der Verwaltungsdigitalisierung in Deutschland immer wieder betont, zugleich aber auch deren Unterrepräsentation in den Arbeits- und politischen Entscheidungsgremien bei Bund und Ländern problematisiert [24]. Die Dezentralisierung im föderalen System der Bundesrepublik, der resultierende Koordinationsaufwand und die fehlende horizontale Kooperation der Kommunen untereinander wurden als wesentliche Herausforderungen dieses Reformvorhabens konstatiert oder als wesentliche Ursachen für das Zurückbleiben Deutschlands bei der digitalen Transformation im internationalen Vergleich benannt. Eine besondere Chance sieht man in der einzigartigen Zusammenarbeit mit externen Akteuren, Stakeholdern sowie auch der Bürgerschaft. Insbesondere die Rahmen setzende Governance wurde dabei als wichtiger Einflussfaktor der digitalen Transformation hervorgehoben [24], [1].

Bei den empirischen Beiträgen, die sich qualitativ mit Digitalisierungsprozessen befassen, ist die kommunale Ebene wenig vertreten. Zudem wurden die Kommunen dabei tendenziell als operative Umsetzungsebene aus einer übergeordneten Perspektive heraus betrachtet [24], [17]. Das Thema der Multilevel-Governance im Zuge der digitalen Transformation wurde in der Regel als Ursache für das langsame Vorschreiten der Bundesrepublik benannt. Es fand insbesondere

bei Analysen zum Zusammenspiel und zu den Koordinationsbemühungen zwischen Bund und Ländern einige Beachtung, wobei die nachgeordnete kommunale Ebene weitgehend ausgeblendet blieb.

### III. CASE STUDY: MULTILEVEL-GOVERNANCE IN LUDWIGSBURG

Der Prozess der kommunalen Digitalisierung in der Stadt Ludwigsburg wurde über viele Jahre durch unterschiedliche Entwicklungsphasen, Strategien, Herangehensweisen und intra-gouvernementale Verflechtungen geprägt. Diese haben zahlreiche, sich im Laufe der Zeit erheblich wandelnde Koordinationsherausforderungen auf kommunaler Ebene hervorgebracht, die es zu bewältigen galt [18]. Über die internen kommunalen Verflechtungen hinaus haben die Digitalisierungsprozesse jedoch auch noch erhebliche Koordinationsherausforderungen in komplexen Mehrebenen-Konstellationen und Multilevel-Governance-Prozessen aufgeworfen, die nachfolgend analysiert werden. Ziel ist, die Bedeutung der Multilevel-Governance für eine erfolgreiche kommunale Verwaltungsdigitalisierung exemplarisch zu verdeutlichen.

#### A. Die Stadt Ludwigsburg in der Mehrebenen-Verflechtung

Bereits kurz nach dem Millennium hatte das Bundesland Baden-Württemberg mit den ersten Vorarbeiten zu einer Service.bw-Plattform zur Verwaltungsdigitalisierung begonnen, die für alle Kommunen anschlussfähig sein sollte. Um diese Zeit begann in der Stadt Ludwigsburg das Nachdenken über die Digitalisierung. In Vorbereitung der Masterplanung zur nachhaltigen Stadtentwicklung von 2006 brachten dann namhafte Unternehmen das Thema Smart City ins Gespräch und stießen bei einer innovationsoffenen, experimentierfreudigen und wirtschaftsfreundlichen Stadtpolitik und Verwaltung auf großes Interesse.

Eine ehemalige Gemeinderätin erinnert sich an eine Ratsitzung, bei welcher der OB „Geschäftsführer von allen großen Firmen, Bosch, Porsche und so eingeladen“ habe (IV14:62).

Dem eingangs benannten, auf bundesdeutsche Verhältnisse übertragenen Zwei-Ebenen-Spiel [23] folgend, zeigt sich das erste Jahrzehnt der Digitalisierungsbemühungen in Ludwigsburg seitens der politisch Verantwortlichen von einer eindeutigen strategischen Orientierung auf allen Ebenen geprägt: Sowohl im Zusammenspiel mit Bundes- und Landesbehörden einerseits, als auch beim Versuch der intrakommunalen Übersetzung andererseits, agierten die Verantwortlichen im Sinne der dominanten Hightech-Strategien des Bundes einheitlich. So suchten die Verantwortlichen bereits früh als innovative Vorreiter im Hinblick auf Digitalisierung einen intensiven Kontakt zu den Bundesministerien. Neben der anfangs erfolgreichen Teilnahme an zahlreichen zukunftsorientierten Bundes- und Landeswettbewerben, beteiligten sie sich u. a. auch intensiv an der Erarbeitung der im Jahr 2017 veröffentlichten Smart City Charta des Bundes (IV06:1ff.).

Zudem hatten sie jedoch auch die Aufgabe, die visionären Ideen und die Erkenntnisse aus dieser Zusammenarbeit auf kommunaler Ebene zu institutionalisieren bzw. zu operationalisieren. Solche Zwei-Ebenen-Spiele führen häufig, so der konzeptionelle Ansatz, zu inkonsistenten Verhaltensmustern auf beiden Bezugsebenen [25]. Im vorliegenden Fallbeispiel äußerte sich dies allerdings weniger im inkonsistenten politischen Handeln. Vielmehr erwies sich die schwierige Übertragbarkeit und Implementierung komplexer Zielvorstellungen

und abstrakter Projektideen der Bundes- und Landesebene als ein Problem, selbst wenn die anspruchsvollen Projekte, z. B. ein digital gesteuertes Verkehrsleitsystem, auf städtische Anwendungen heruntergebrochen wurden. In Ludwigsburg äußerte sich dies in der Dominanz visionärer Leitbilder, langfristiger Entwicklungsziele und weitreichender technologischer Projekte. Einer leistungsfähigen Operationalisierung und hinreichend tragfähigen Institutionalisierung auf kommunaler Ebene wurde dagegen zunächst weitgehend vernachlässigt und in einer späteren Phase vergeblich zu implementieren versucht.

Verstärkend kamen vertikale Abstimmungsprobleme zwischen Politik und einer sektoral strukturierten Verwaltung hinzu. So wurde die Initiative für eine digitale Vorreiterrolle vom bundesweit vernetzten, jedoch mit wenig Ressourcen für ein solches Vorhaben ausgestatteten Fachbereich „Nachhaltige Stadtentwicklung“ betrieben, während die gesamte organisatorische Verantwortung für die Digitalisierung im Hauptamt der Stadt verortet war. Das Hauptamt hatte jedoch andere, z. B. haushaltsbezogene, Herausforderungen zu bewältigen und verfolgte lange Zeit keine (eigene) Digitalisierungsagenda (IV06:2).

Allmählich wurden allerdings auf Bundesebene die visionären Handlungsfelder, die es zu bearbeiten galt, auf die städtische Ebene heruntergebrochen. So übertrug die Strategische Forschungs- und Innovationsagenda (FONA) des Bundes wesentliche Aspekte der Hightech-Strategie auf die Städte und formulierte technologische Anforderungen, um eine „Zukunftstadt“ realisieren zu können [30]. Zu diesem Zweck wurde 2015, initiiert durch das Bundesministerium für Bildung und Forschung (BMBF), die „Nationale Innovationsplattform Zukunftstadt“ (IPZ) gestartet, der sich bald auch die Stadt Ludwigsburg mit eigenen projektierten Beiträgen anschloss. Digitalisierung sollte damit in das Feld urbaner Nachhaltigkeit eingebettet werden, mit der Vision einer CO<sub>2</sub>-neutralen, energie- und ressourceneffizienten sowie klimaanangepassten Stadt von morgen [17].

Ebenfalls 2015 entwickelte man in Ludwigsburg unter Leitung des Director Smart Cities eines ansässigen Unternehmens das Living Lab, ein öffentlich gefördertes Innovationsnetzwerk für nachhaltige und digitale Stadt- und Verwaltungsentwicklung. In Reallaboren suchte die Stadtpolitik und Verwaltung die Kooperation mit Partnern aus Wirtschaft, Industrie und Wissenschaft, um Impulse für die technologische Entwicklung und deren städtische Nutzung zu gewinnen [25]. Über das Living Lab wurden Fördermittel für visionäre Projekte akquiriert, z. B. im Wettbewerb „Digitale Stadt“ des Digitalverbands Bitkom und des Deutschen Städte- und Gemeindebunds oder im Landesförderprogramm „Städte und Gemeinden 4.0“.

Erstmals kommt mit der extern besetzten Leitung des Living Lab ein Faktor zum Tragen, der sich auch im weiteren Verlauf der Digitalisierungsprozesse als bedeutsam erwies: die einzigen Akteure, die in komplexen Mehrebenen-Konstellationen oder auch „Zwei-Ebenen-Spielen“ unabhängig und frei von engen institutionellen Bindungen agieren können und sich daher besonderer Beliebtheit erfreuen, so der konzeptionelle Ansatz der Mehrebenen-Analyse, sind externe Expertinnen und Experten [21]. Man erhofft sich von diesen vor allem, dass sie auch hartnäckige institutionelle Limitierungen und Denkbarrieren auf den jeweiligen räumlich-administrativen Ebenen durchbrechen können. Man setzte also auf ökonomisches und technologisches Know-how, vernachlässigte dabei

jedoch die sozio-technologische Komponente von Digitalisierung [17].

Nationale Plattformen, Bundes- und Landeswettbewerbe, Innovationsnetzwerke, Masterpläne und Vorzeigeprojekte – die Digitalisierungsbestrebungen in Ludwigsburg waren zu dieser Zeit, aus der Mehrebenen-Perspektive heraus betrachtet, von dem Bestreben geprägt, erfolgreich an die zukunftsweisenden Aktivitäten von Bund und Ländern zur digitalen Transformation anzuknüpfen. Aus der Anknüpfung an diese weitreichenden Ziele ließen sich zwar, selbst wenn sie mit anspruchsvollen Projekten auf die kommunale Ebene heruntergebrochen wurden, gemeinsame, technologisch geprägte Zielorientierungen für die eigene Stadt ableiten, diese entfalteten allerdings trotz anfänglicher Erfolge nicht die Wirksamkeit, die man sich davon versprach. So gelang es den Verantwortlichen in diesem Zwei-Ebenen-Spiel trotz gleicher Sprache und Zielorientierung in beide Richtungen nicht, die anspruchsvollen, an Bund und Ländern orientierten Zielvorstellungen auf die Ebene der Kommunalverwaltung zu übertragen, zu operationalisieren und in institutionelle Strukturen und Regelungen zu überführen.

„Man hat Fördermittel akquiriert und hat dann relativ zusammenhanglos versucht ohne klare digitale Agenda einfach immer dort, wo man an Fördermittel gekommen ist, ein Stück weit voranzuschreiten.“ (IV17:9)

Die im Zitat aus Verwaltungskreisen benannten strukturellen Defizite verweisen darauf, dass man vornehmlich in nicht integrierten, instrumentellen, von außen aufgesetzten Foren mit externer Expertise agierte [17]. Zu dieser Zeit entwickelte sich bei Gemeinderäten und in der Stadtverwaltung erstmals eine gewisse Entfremdung zwischen den abgehobenen Zielorientierungen der Stadtpolitik als bundesweiter „Vorreiter der kommunalen Digitalisierung“ einerseits und einer bodenständigen internen Sichtweise andererseits. Folgende Interviewaussage mag dies illustrieren:

„Im Bereich [...] der Verwaltung, ist man ein Stück weit verzweifelt [...], weil die Äußerungen in der Öffentlichkeit und in der Presse mit den technischen Rahmenbedingungen, [...] Arbeitsweisen [und] den Notwendigkeiten, die man in diesen Fachdezernaten [...] erkannt hat, eben nicht übereingestimmt hat. Und so ist eigentlich ein super Bild nach außen entstanden, das in der Verwaltung selbst in keiner Form bestätigt werden konnte.“ (IV07:22)

#### *B. Zunehmender Handlungsdruck durch das Onlinezugangsgesetz*

Auf Bundesebene dauerten grundlegende Institutionalisierungsbemühungen mehr als ein Jahrzehnt. Mit am Anfang stand im Jahr 2010 z. B. die „Nationale E-Government-Strategie“ und die Gründung des „IT-Planungsrats“ für die politische Mehrebenen-Koordination der digitalen Transformation sowie der IT-Staatsvertrag, der 2019 – rund neun Jahre später – zur Gründung der FITKO überarbeitet wurde. Die „Föderale IT-Kooperation“ (FITKO), die im Jahr 2020 von Bund und Ländern gemeinsam eingerichtet wurde, unterstützt den IT-Planungsrat und dessen politische Aufgaben vorwiegend bei technischen, anwendungsbezogenen IT-Fragen und der Entwicklung föderaler IT-Strukturen und -Standards.

Unterdessen gewannen im Jahr 2017 die Digitalisierungsprozesse in Deutschland auch durch die Formulierung von Leitlinien, z. B. die Smart City Charta des Bundes, vor allem

aber durch verbindliche Regulierungen, d. h. durch institutionelle Settings, deutlich an Gestalt. Dementsprechend wuchs auch der administrative Handlungsdruck zur baldigen Realisierung von konkreten Vorhaben. Vor allem wurde die Digitalisierung durch das Onlinezugangsgesetz (OZG) vom 14. August 2017 befeuert, das Bund und Ländern ermöglichen sollte, die Digitalisierung der Verwaltung bundesweit voranzubringen (BGBl. I S. 3122, 3138). Das Gesetz verpflichtete Bund und Länder mit einer Fristsetzung bis Ende des Jahres 2022 dazu, für nahezu 600 Verwaltungsleistungen digitale Zugänge anzubieten, eingebettet in eine vernetzte IT-Infrastruktur, die den Nutzern einfache Zugriffe ermöglichen sollten – für Kommunen, die explizit keine Erwähnung im Gesetz finden und von den Ländern beteiligt werden sollten, aufgrund der hohen Bindewirkung von gesetzlichen Regelungen trotz alledem ein Ansporn zur Eigeninitiative.

„2017 wurde dann gestartet in vielen Kommunen. Also, die, die dann innovativ unterwegs sind, haben sich Gedanken gemacht ‚Wie können wir das lösen, wie können wir das dann vielleicht auch in Eigenregie entwickeln?‘ Das ist auch alles gut und richtig, aber es ist nun mal so, dass die Entwicklungskraft und dann auch teilweise die Innovationskraft, die ja hinter so einer bundesweiten Entwicklung im EfA-Kontext steckt, doch nochmal eine andere ist [...]“ (IV18:10)

Auf der Länderebene in Baden-Württemberg zeigten sich die aufwendigen und mehrjährigen Vorarbeiten zur institutionellen Verankerung der Digitalisierung z. B. anhand der lang andauernden Suche nach einer tragfähigen landesministeriellen Organisationsstruktur und deren Implementierung (IV21), aber auch anhand der umfassenden Digitalisierungsstrategie der Landesregierung Baden-Württemberg aus dem Jahr 2017 (Ministerium für Inneres, Digitalisierung und Migration) sowie den Aktivitäten, die sich seitdem schrittweise entwickelt haben. Sie alle sollten der Digitalisierung in der Fläche, wie es mitunter heißt, zum Durchbruch verhelfen: so z. B. seit 2017 der Aufbau einer Digitalakademie@bw, die mehrjährig geförderte Ausbildung von kommunalen Digitallotsen in der Digitalakademie, die gute Praxis, die über Jahre im Landeswettbewerb „Digitale Zukunftskommune@bw“ in den Siegerstädten Heidelberg und Ulm erarbeitet wurde und auch im Jahr 2024 in den Veranstaltungen „Digitale Zukunftskommune@bw“ präsentiert wird. Auch die Herausbildung einer Reihe von Barcamps seit 2022 zählt dazu, die auf die Initiative der Stadt Ludwigsburg zurückgeht und inzwischen auf drei selbst organisierte Barcamps in Baden-Württemberg zur Bearbeitung von Fragen des digitalen Bauantrags mit zuletzt knapp 100 Teilnehmenden angewachsen ist.

Währenddessen erzielten auch die weiteren Versuche der Stadtpolitik in Ludwigsburg, das Zwei-Ebenen-Spiel mit der zuvor eingeschlagenen, zukunftsorientierten Top-Down-Strategie sowie der Expertise von außen zum Erfolg zu verhelfen, keinen Durchbruch. So beauftragte man im Sinne einer Leitvorstellung für die Stadt die Erarbeitung einer in ihrer Wirkung weitgehend verpufften „Digitale Agenda Ludwigsburg“, wofür erneut externe, ökonomisch versierte Expertinnen und Experten mit viel Handlungsfreiheit beauftragt wurden.

„Es gab viel Geld zu verteilen. Und das ist ja immer toll. Man hat ein Projekt, was man vorzeigen kann und so. Das hört sich jetzt vielleicht alles ein bisschen, wie soll ich sagen, negativ an. Das ist nicht schlecht, was da auf dem Papier [der digitalen Agenda] entstanden ist. Aber es ist halt nichts passiert“ (IV06:5).

Auch bei der im selben Jahr erstmals in Angriff genommenen organisatorischen Verankerung der Digitalisierung wurde von stadtpolitischer Seite auf einen freischaffenden, gegenüber der Verwaltung vollständig freigestellten Experten zurückgegriffen. Als CDO mit einer Stabsstelle beim Oberbürgermeister sollte er die Verantwortung für die Auswahl und Initialisierung aller Projekte mit digitalen Komponenten übernehmen. Sein zweiköpfiges Team sollte alle Verwaltungsbereiche verknüpfen, sämtliche IT- und Digitalprozesse steuern und mit dem „Bürgerkonto“ ein innovatives Vorzeigeprojekt realisieren. Damit sollte es Bürgerinnen und Bürgern ermöglicht werden, alle Behördengänge digital abzuwickeln, darunter den Bezug von Formularen, die Antragsstellung und später auch den Empfang von Dokumenten.

### C. Strategische Neustrukturierung

Nach anfänglichen konzeptionellen Erfolgen geriet die Digitalisierung in Ludwigsburg, wie analytisch gezeigt wurde, mit ihrer ausgeprägten Top-Down-Strategie und ökonomisch sowie technologisch ausgerichteten externen Expertise gegen Ende des Jahrzehnts zunehmend ins Stocken und erfuhr einhergehend mit einem stadtpolitischen Kurswechsel in den Folgejahren eine umfassende Neustrukturierung.

Zunächst konnte sich die Stadt aufgrund fehlender Implementierungsnachweise und Erfolge auf kommunaler Ebene in mehreren Wettbewerben, wie dem Bitkom-Wettbewerb „Digitale Stadt“, „Digitale Zukunftskommune@bw“ oder „Zukunftsstadt“, nicht mehr behaupten, wodurch die daran geknüpften Digitalisierungsprozesse einen Dämpfer erfuhren.

Bei den OB-Wahlen im Juni 2019 wurde ein neuer Oberbürgermeister ins Amt gewählt, was unmittelbar die Kündigung des CDO nach sich zog. Dieser war erheblich in die öffentliche und verwaltungsinterne Kritik geraten [39]. Der Stand der Dinge wurde später später offensichtlich, als die Stadt, die stets damit geworben hatte, im Hinblick auf Digitalisierung „vorn mitzuspielen“, in einem Vergleich digitaler Dienstleistungen und Bürgerservices in den 100 größten deutschen Städten nur mittlere und abgeschlagene hintere Plätze belegte, womit sich, so die Presse, das Mittelmaß bei der Digitalisierung und die unterdurchschnittliche Leistung beim Bürgerservice offenbare [40]. Die zuvor unternommenen Versuche der strukturellen und institutionellen Verankerung der Digitalisierungsprozesse, darunter das gesamte Team des CDO und die Vorzeigeprojekte „Smart City Cloud“ und „Bürgerkonto“, waren damit gescheitert und wurden im Zuge des stadtpolitischen Kurswechsels vollständig abgewickelt.

Unter der neu gewählten politischen Führung wurden zunächst eine Bestandsanalyse und Organisationsentwicklung beauftragt. Spätestens nach dem Abklingen der Coronapandemie wurde dann, als Konsequenz aus früheren Misserfolgen, von der Top-Down-Strategie weitgehend Abstand genommen und mit Gründung des Fachbereichs „Digitalisierung und Informationstechnik“ eine administrativ handlungsfähige Verankerung in der Verwaltung hergestellt. Ein 27-köpfiges Team mit neu gewählter Leitung übernahm das Aufgabengebiet. Dort werden heute, gemeinsam mit dem Fachbereich Personal und Organisation, alle Digitalprojekte der Stadt gesteuert. Ein Kurswechsel erfolgte zudem bei der fachspezifischen Umsetzung: Erstmals wurden mit einer an verwaltungsinternen Notwendigkeiten orientierten Bottom-Up-Strategie die erforderlichen Ressourcen und die administrative Verantwortung für die Entwicklung digitaler Fachverfahren in die Hände der zuständigen Fachbereiche gelegt.

#### D. Die Genese des digitalen Bauantragsverfahrens in Ludwigsburg

Mit der Verabschiedung des Onlinezugangsgesetzes OZG von 2017, das Bund und Länder dabei unterstützen sollte, die Verwaltungsdigitalisierung voranzubringen, wurde recht schnell offensichtlich, dass im Sinne der Mehrebenen-Verflechtung der Bundesrepublik Deutschland der weitaus größte Teil der Verwaltungsleistungen auf kommunaler Ebene zu erbringen sein wird. Dies führte im Sinne des Grundrechts auf kommunale Selbstverwaltung sowie vermuteten Koordinationsdefiziten zwischen Bund-Ländern und Kommunen zu erheblichen Auseinandersetzungen:

„Beim OZG [...] steht ganz vorne drin: ‚Das OZG verpflichtet Bund und Länder.‘ Das war ein großer Konflikt am Anfang, weil dann in der Denke von Bund und Ländern natürlich die Kommunen mit verpflichtet sind, was die Kommunen aber nie so verstanden wissen wollten und sich da auch sehr gegen gewehrt haben, weil der Bund nicht direkt die Kommunen verpflichten kann.“ (IV19:3)

2019 wurde dann, auch als Reaktion auf das OZG, die Landesbauordnung Baden-Württemberg novelliert. Damit wurde zum zweiten Mal auf übergeordneter Ebene eine verbindliche Vorgabe für die Kommunen festgesetzt, indem ein gesetzlicher Anspruch für die Einreichung digitaler Antragsdokumente geschaffen wurde, sodass in Baden-Württemberg ab 2022 keine analogen Bauanträge mehr eingefordert werden sollten (IV12:82). Angesichts der Bindungswirkung von Gesetzen mit unmittelbarer Wirkung auf das kommunale Handeln erhöhte sich noch einmal der Handlungsdruck mit enger Fristsetzung für die Kommunen.

In Ludwigsburg beförderte der wachsende Zeitdruck und die Last der Verantwortung für die neue Gesetzeslage, die kommunale Neuausrichtung der Verwaltungsprozesse sowie das darüber in Gang gesetzte Projekt des digitalen Bauamtes. Im zuständigen Bürgerbüro Bauen, nunmehr mit der nötigen kommunalen Rückendeckung ausgestattet, entschied man sich, diese „Chance“ „proaktiv“ nutzen zu wollen (IV09:4), digitale Baugenehmigung „größer [zu] denken“ und sich nicht auf Gesetzesvorgaben zu begrenzen (IV13:18 und 20).

„Die LBO - Landesbauordnung für Baden-Württemberg wurde im Sommer 2019 dann recht kurzfristig und überraschend novelliert. Der gesetzliche Anspruch auf Einreichen der Bauantragsunterlagen in digitaler Form wurde da dann auch mit aufgenommen. Das hat die Baurechtsbehörden relativ kalt erwischt. So auch das Bürgerbüro Bauen in der Stadt Ludwigsburg.“ (IV09:1)

Kleinteilige Digitalisierungsprojekte hatte es in den Fachbereichen der Verwaltung schon zuvor gegeben, doch erst nach der Neustrukturierung erhielten diese die politische Rückendeckung, mehr Ressourcen für die Entwicklung digitaler Vorhaben aufzuwenden. Mit breiter Unterstützung durch Politik und Verwaltung und in enger Kooperation zwischen Bürgerbüro Bauen und den zuständigen Fachbereichen für Digitalisierung sollte die digitale Bereitstellung des Bauantragsverfahrens fristgerecht vorgebracht und den Bürgerinnen und Bürgern damit die vollständige digitale Bearbeitung von Bauanträgen ermöglicht werden.

#### E. Koordinationsdefizite der Mehrebenen-Verflechtung

Mit den Gesetzesvorgaben des Bundes, den Verordnungen der Länder und den parallel vorangetriebenen Prozessen der Entwicklung digitaler Instrumente auf beiden Ebenen wurde

die Koordination vor allem von einzelnen lokalen Digitalisierungsprojekten im Mehrebenen-Gefüge der Bundesrepublik zu einer erheblichen Herausforderung.

In der ersten, weniger erfolgreichen Phase der Digitalisierung hatte in der Stadt Ludwigsburg eine doppelte Top-Down-Strategie das „Zwei-Ebenen-Spiel“ dominiert. Aus diesen resultierten zahlreiche, komplexe und intra-gouvernementale Koordinationserfordernisse [17]. Abgelöst wurde diese Variante des Zwei-Ebenen-Spiels durch eine gesetzlich, von Bund und Ländern fristgerecht eingeforderte Top-Down-Strategie, mit einem – der hohen Bindewirkung entsprechenden – kommunalen Handlungsdruck. Dem stand allerdings inzwischen eine auf den lokalen Erfahrungen der Vergangenheit basierende, kommunalpolitisch für richtig erachtete und neu strukturierte Bottom-Up-Initiative der Digitalisierung gegenüber. In Folge entwickelten sich erhebliche Koordinationsdefizite zwischen Bund und Ländern einerseits, und der innovationsfreudigen Kommune Ludwigsburg andererseits, wie nachfolgend zu zeigen sein wird.

Diese angesprochenen Koordinationsdefizite kulminierten insbesondere im Jahr 2021, nachdem das Ludwigsburger Vorhaben, den digitalen Bauantrag ohne Medienbrüche annehmen, bearbeiten und genehmigen zu können, nach einer zweijährigen, ressourcen- und aufwandsreichen Entwicklungsphase bereits große Fortschritte erzielt hatte und vor der erfolgreichen Implementierung stand. Mitte des Jahres 2021 wurde dann auf Bundes- und Länderebene eine „Einer-für-Alle“-Lösung (EfA) zur Bearbeitung digitaler Bauanträge bereitgestellt – gerade einmal vier Monate, bevor die Kommunen eine erfolgreiche Digitalisierung von Verwaltungsleistungen im Sinne des OZG vermelden sollten.

Die Idee einer EfA-Lösung war zuvor im IT-Planungsrat, dem zentralen Beratungsgremium auf Bundesebene, entstanden. Dort hatte man Themenfelder und dafür zuständig zeichnende Bundesländer definiert, die digitale Verfahren erproben und ihre Erfahrungen später der Allgemeinheit zur Verfügung stellen sollten. Aus dieser Idee heraus entwickelte sich dann – bereits einige Jahre nach Verabschiedung des OZG – die Idee, dass in den Themenfeldern doch auch „Einer-für-alle“ – Lösungen erarbeitet und bundesweit verfügbar gemacht werden könnten (IV21:3). Die EfA-Lösung zum digitalen Bauantrag wurde, so der Eindruck von Befragten, wenig mit der kommunalen Ebene abgestimmt, im Bundesland Mecklenburg-Vorpommern entwickelt und sollte mit nachfolgender Zustimmung der Landesregierung auch in Baden-Württemberg übernommen werden.

„Das ist der Ansatz von Bund und Ländern: Wenn man so ein Bund-Länder-Gremium macht, dann sind die Kommunen ja automatisch mit dabei, weil die Länder ja am Tisch sitzen. Das ist die Position von oben. Und von unten sind die Kommunen natürlich nicht einverstanden, weil das eine sehr indirekte Vertretung ist und in der Theorie die Länder die Kommunen fragen müssten“ (IV19:2).

„Also aus unserer Wahrnehmung her sind die ganzen oder sind ein Großteil der EfA-Nachnutzungsprojekte jetzt so im letzten viertel- oder halben Jahr auf den Tisch gekommen. Also relativ spät, wenn man die ursprüngliche Umsetzungsfrist OZG 31.12.2022 sieht. Jetzt ist es eher so, dass gefühlt die Länder ihre Leistungen noch „losbekommen“ wollen, so lange es noch die Bundesförderung gibt.“ (IV20:4)

Die sich im betrachteten Mehrebenen-Gefüge aufzeigenden Koordinationsdefizite zwischen Bund, Ländern und Kommune führen im untersuchten Fallbeispiel dazu, dass die Digitalisierungserfolge, die im Bauantragsverfahren der Stadt Ludwigsburg erzielt wurden, sich heute zwar als eine sehr erfolgreiche, funktionierende Individuallösung, aber auch als ein isolierter Entwicklungspfad darstellen, der mittel- und langfristig von übergeordneter Seite infrage gestellt und durch bundesweit einheitliche Vorgaben und Schnittstellen zunehmend mehr konterkariert werden dürfte. Gleiches gilt im Übrigen auch für den letztlich ebenfalls isolierten technologischen Entwicklungspfad, der mit Service.bw – zusätzlich zu den Entwicklungen auf Bundesebene sowie der Arbeit in Ludwigsburg – verfolgt wurde. Dort wurden, kurz nach dem Millennium beginnend, die Bereiche Information, Kommunikation und Online-Dienste systematisch aufgebaut und erweitert. Nicht zuletzt wurde noch im Jahr 2020 in Kooperation mit einem externen Dienstleister als Pionierleistung der Digitalisierung eine provisorische, den Anforderungen des OZG jedoch nicht genügende Antragsstrecke für die digitale Einreichung kommunaler Bauanträge in Baden-Württemberg erarbeitet (IV20:3).

„Wir haben aber jetzt gerade das Thema [...] digitales Baugenehmigungsverfahren, wo es aktuell zwischen uns [...] und dem Land große Diskussionen gibt. Weil das Land eine Lösung aus Schleswig-Holstein [Mecklenburg-Vorpommern A. d. V.] nachnutzen will. Und wir [...] sagen, na ja, wir haben ja schon funktionierende Prozesse auf Service.bw. Warum nutzt man die nicht weiter? Das ist jetzt gerade so ein bisschen die Schwierigkeit vor dem Hintergrund, dass wir mehr oder weniger vor vollendete Tatsachen gestellt worden sind.“ (IV20:3)

#### IV. TENTATIVE SCHLUSSFOLGERUNGEN AUS DER CASE STUDY

Das eigentliche, weitgehend ungelöste, Problem, so erscheint es bei nachträglicher Erörterung, bestand an erster Stelle darin, dass Bund und Länder mit dem OZG eine strenge Fristsetzung für die ersten Schritte zur Verwaltungsdigitalisierung – und damit u. a. auch für den digitalen Bauantrag – vorgegeben haben, die sich angesichts fehlender Strukturen, Institutionen und angesichts komplexer Koordinationsbedarfe im vielschichtigen Mehrebenen-Gesamtgefüge der Bundesrepublik auch für Bund und Länder selbst als überaus anspruchsvoll und vielleicht auch als unrealistisch erwies. So galt es Organisationen, Institutionen, Regelwerke sowie Verantwortlichkeiten und Koordinationsmöglichkeiten, vorrangig zwischen Bund und Ländern, zu entwickeln und zu etablieren, um in eine erfolgreiche Verfahrensentwicklung für die anwendungsbezogenen EfA-Lösungen gelangen zu können. Ein langer Weg, an dessen Ende man dann auch indirekt die Kommunen erreichen wollte, ohne ihnen aufgrund ihrer unantastbaren Grundrechte auf kommunale Selbstverwaltung „zu nahe zu treten“.

So konnten die potenziell möglichen, einheitlichen Vorgaben wie die EfA-Lösungen, an denen sich die Kommunen hätten orientieren können, teilweise erst spät, beim Bauantragsverfahren deutlich zu spät, bereitgestellt werden. Den Ausführungen der beteiligten kommunalen Interessenvertretungen zufolge, fand dabei die kommunale Ebene, auf denen ja die Hauptlast bei der Umsetzung von digitalen Verwaltungsverfahren ruhen sollte, zu wenig Berücksichtigung (IV19, IV20). Kommunen verfügen in der Regel im System der Mehrebe-

nen-Verflechtung über geregelte Beteiligungs- und Berücksichtigungsrechte bei der Erarbeitung von räumlich-administrativ übergreifenden Institutionen seitens des Bundes und der Länder. Im Gegenzug müssen sie sich an entsprechenden rechtlichen Vorgaben halten. Im vorliegenden Fall war die Beteiligung der Kommunen vorwiegend über Interessenvertretungen, wie Bundes- und Landesverbände von Städten und Gemeinden, in beratender Funktion gewährleistet (Deutscher Städtetag, Deutscher Städte und Gemeindebund u. a.). An erster Stelle stand dabei die Mitwirkung im IT-Planungsrat von Bund und Ländern sowie an deren Arbeitsgruppen (IV19:1). Belange von einzelnen innovativen Städten, die sich auf den Weg der Digitalisierung gemacht hatten, fanden dabei zwangsläufig keine Berücksichtigung. Auch scheinen die Möglichkeiten der Einflussnahme der kommunalen Interessenvertretungen auf das OZG auch im IT-Planungsrat eher begrenzt gewesen zu sein, wie das Zitat vermuten lässt:

„[...] in diesem ‚jetzt müssen wir einen ganz großen Schritt machen und das ist dann das OZG‘ waren die Kommunen nicht eingebunden. Aber irgendjemand muss sich Gedanken gemacht haben ‚Wie komme ich einen relativ großen Schritt weiter?‘ Er hat aber nicht in erster Linie an die Kommunen gedacht.“ (IV19:5)

Die OZG-Fristsetzung und die nachfolgende Novelle der Landesbauordnung Baden-Württemberg setzte zugleich diejenigen Kommunen unter einen enormen Handlungsdruck, die, wie die Stadt Ludwigsburg, gewillt waren, der veränderten Rechtslage aufgrund der davon ausgehenden Bindungswirkung von Verordnungen fristgerecht nachkommen zu wollen. Für Ludwigsburg kamen die Vorgaben von Bund und Ländern jedoch deutlich zu spät.

Die gesetzlich fundierte, hohe Bindungswirkung, die in Ludwigsburg zu konstatieren war, fand allerdings nicht immer und überall in den bundesdeutschen Städten und Gemeinden Berücksichtigung in Form von einschlägigen Entwicklungsaktivitäten. Vielmehr wurde recht unterschiedlich auf die Fristsetzungen reagiert und es kam insbesondere in vielen kleineren Städten und Gemeinden, die sich auf die eine oder andere Weise überfordert fühlten, zum Verlust der Bindungswirkung von Gesetzen und Verordnungen sowie zu Blockaden bei den Bestrebungen zur digitalen Transformation, wie folgendes Zitat verdeutlichen mag:

„Wir haben eine Inhomogenität bei den Städten, wie weit die bei der Umsetzung OZG sind. Also manche sind da wirklich sehr, sehr weit vorangeschritten. Das sind auch diejenigen, die selbst Prozesse erstellen und sich da so ein Stück weit beholfen haben mit eigenen Mitteln. Aber für viele gerade kleine Gemeinde und Städte ist das, die stehen halt da komplett noch am Anfang. Und die wissen eigentlich nicht so wirklich, wo sie anfangen sollen.“ (IV20:6)

Die Stadt Ludwigsburg machte sich jedoch – auch angesichts fehlender Orientierung und fehlender erkennbarer Fortschritte auf Bundes- und Landesebene – unverdrossen ans Werk und vollendete mit großem Aufwand ihr eigenes, weitgehend funktionsfähiges digitales Bauantragsverfahren. Erst in der Zukunft wird es sich allerdings zeigen müssen, ob sich das eigenständige digitale Bauantragsverfahren der Stadt Ludwigsburg als kompatibel und durchsetzungsfähig erweist gegenüber einem ggf. landesweit implementierten EfA-Verfahren sowie gegenüber zunehmend vereinheitlichten Schnittstellen, Portalzugängen, Registern und Datenarchiven auf den verschiedenen räumlich-administrativen Ebenen u. v. a.

Wenn sich aus den bisher gewonnenen Erkenntnissen über mehrseitige Koordinationsdefizite im dargelegten Fallbeispiel der kommunalen Digitalisierung in Ludwigsburg tentative Schlussfolgerungen ableiten lassen, dann dürfte an erster Stelle die Berücksichtigung der zeitlichen Notwendigkeiten, Ressourcen und Transferbedarfe stehen, die es zur Erbringung der institutionellen und organisatorischen Rahmenbedingungen für solche umfassenden Reformvorhaben braucht, insbesondere in komplexen Gemengelagen der Mehrebenen-Verflechtung. Dies gilt gleichermaßen für Bund und Länder, die nach heutigem Wissen einen mehrjährigen Vorlauf gebraucht hätten, um das OZG fundiert vorzubereiten und zu einem Erfolgsmodell zu machen.

Diese naheliegende Erkenntnis gilt allerdings auch für die Stadt Ludwigsburg. Dort lassen sich ebenso Steuerungs- und Koordinationsdefizite zwischen gesamtgesellschaftlichen Prozessen wie der digitalen Transformation auf Bundes- und Landesebene sowie der sich als schwierig erweisenden Implementierung der Verwaltungsdigitalisierung auf kommunaler Ebene am Beispiel von Ludwigsburg aufzeigen.

Es hat sich im Lauf der Lernprozesse dort erwiesen, dass Ludwigsburg weit weniger Gewicht auf übergeordnete Zielvorstellungen, Vorzeigeprojekte und extern engagierte Expertise hätte legen sollen, als vielmehr sehr viel Gewicht auf eine Tragfähigkeit versprechende Institutionalisierung der Digitalisierung auf kommunaler Ebene. Das Zwei-Ebenen-Spiel in Ludwigsburg hat die Schwierigkeiten und die immensen Herausforderungen verdeutlicht, die es beim Transfer von abstrakten, visionären Zielvorstellungen auf Bundes- und Länderebene bei der Strukturbildung, Institutionalisierung und Implementierung auf der konkreten kommunalen Ebene zu bewältigen gilt. Dabei hat sich an dieser Stelle nach jahrelangen Versuchen, an die weitreichenden Zielvorstellungen von Bund- und Ländern anzuknüpfen, die konsequente Anwendung des Subsidiaritätsprinzips innerhalb der Kommunalverwaltung, nämlich die eigenständige Erarbeitung von einschlägigen Fachverfahren in den dafür zuständigen Fachämtern, bei gesamtstädtischer Koordination, als erfolgreich erwiesen.

Für sinnvoll zu erachten bleibt last but not least, in einem komplexen Mehrebenen-System wie der Bundesrepublik und bei einer weitreichenden gesellschaftlichen Reform wie der Digitalisierung, ein von Beginn an im Konsens unter allen Akteuren festgelegtes Prozess- und Verfahrensmanagement mit einem gewissen professionellen Abstand zu den alltäglichen Herausforderungen der Digitalisierung. Dazu gehört z. B. die gemeinsame Verständigung aller beteiligten Ebenen über zeitliche Abläufe und Prozessschritte. Dabei dürfte ein befristeter zeitlicher Vorlauf für die Erfordernisse der Institutionalisierung der Digitalisierungsprozesse bei Bund (Gesetze), Ländern (Verordnungen, Organisationsstruktur, Transfer-Infrastruktur) und Kommunen (Leitbilder, Organisationsstruktur, Umsetzungsrichtlinien) zum allseitigen Vorteil beigetragen haben. Auch hätte es vielleicht gemeinsamer deliberativer Verhandlungen „auf Augenhöhe“ über die Top-Down-Entwürfe von verbindlichen Vorgaben durch Bund und Länder bedurft, die für alle gemeinsam als Orientierung für die regionale und kommunale Weiterarbeit dienen sollten. Und es hätte eine Verständigung mit den Kommunen über nachfolgende, in der Verantwortung der kommunalen Selbstverwaltung liegende Bottom-Up-Entwicklungsziele gebraucht.

Das mag vor allem im Nachgang einfach klingen, ist es aber nicht, wie das nachfolgende Zitat zur Arbeit des IT-Planungsrates verdeutlichen mag:

„Ein Problem ist vielleicht aus meiner Wahrnehmung, dass das große Ganze manchmal etwas aus dem Blick gerät. Dass man da 23 Tagesordnungspunkte hat und verliert sich im Kleinklein. Aber die großen Linien, wenn wir jetzt beim OZG sind, wohin wollen wir eigentlich, wir wollen, dass die Verwaltung durchgängig digital wird, das verliert man dann manchmal.“ (IV19:1)

Große Hoffnung macht heute allerdings, so lässt sich abschließend festhalten, die allseits spürbare Dynamik, die sich nach schwerfälligen Jahren des Reformbeginns mittlerweile auf allen beteiligten Ebenen – Bund, Länder und Kommunen, aber auch in Wissenschaft und Gesellschaft – herausgebildet hat, sodass sich schrittweise in aller Breite erfolgreiche Ordnungsstrukturen zu etablieren beginnen, Ungereimtheiten, Widersprüche und Hemmnisse bearbeitet und allmählich ausgeräumt werden sowie gemeinsam gangbare Wege zur weiteren digitalen Transformation der Gesellschaft aufscheinen, auch wenn dieser Weg noch ein sehr weiter sein wird.

## V. FORSCHUNGS-AUSBLICK

Der vorliegende Beitrag argumentierte, dass erfolgreiche kommunale Verwaltungsdigitalisierung als sozio-technisches System konzipiert werden muss, das einem holistischen Ansatz folgen sollte. Dieser verbindet smarte Systemarchitektur auf der IT-Ebene und smarte Verwaltungsarchitektur mit smarterer Multilevel-Governance. Die emergenten Effekte einer integrierten Ausarbeitung der drei Säulen erzeugen die „Smartness“ kommunaler digitaler Verwaltung.

In einem Deep Dive wurde die Geschichte der Digitalisierung der Verwaltung in der Stadt Ludwigsburg, mit Fokus auf das digitale Bauamt, aus einer Multilevel-Governance Perspektive rekonstruiert. Deutlich wurde dabei, dass der Erfolg kommunaler Verwaltungsdigitalisierung von der Qualität (horizontalen) (stadt-)interner Governance und der Qualität der (vertikalen) Multilevel-Governance abhängt. Die Analyse der Multi-Level-Governance verdeutlichte erstens nicht nur, wie beschränkt kommunale Handlungsspielräume bei der Verwaltungsdigitalisierung häufig sind, sondern auch, wie optimierungsbedürftig die *bi-direktionale Kommunikation* – also top-down und bottom-up – empirisch mitunter ist. Deutlich wurde zweitens, dass der *Wettbewerbsföderalismus*, der in vielen Policy-Feldern ein Innovationsmotor sein kann, im Feld der Digitalisierung diese Funktion höchstens eingeschränkt wahrnehmen kann. Aufgrund seiner Querschnittlichkeit und der Notwendigkeit potenziell europaweiter Interoperabilität sind Insellösungen in der IT eher innovationshemmend als -fördernd. Die massiv parallele Entwicklung digitaler Lösungen auf kommunaler Ebene verbrennt daher unnötig finanzielle Ressourcen. Der „Einer für Alle“-Ansatz (EfA) erweist sich in der Theorie als vielversprechend, scheitert in der Praxis jedoch noch an intransparenten Verfahren und dem Beharrungswillen von Kommunen, ihre – häufig kostspielig selbst entwickelte IT-Lösung – weiter nutzen zu wollen, statt eine neue (EfA) Lösung aufzusetzen (vgl. für eine positivere Einschätzung [26]).

Neben einem Kommunikationsdefizit scheint hier ein ideatives Defizit auf: Es fehlt – jenseits der (vermeintlichen) Kosteneinsparung – an einem Narrativ, einem Telos, mit dem die Verwaltungsdigitalisierung aufgesetzt wird. Ein solches Telos könnte im Rahmen partizipativer und inklusiver demokratischer Beteiligungsprozesse entwickelt werden (vgl. [27]). In dieser Perspektive bietet Verwaltungsdigitalisierung eine Option, eine *realistische demokratische Utopie* zu entwickeln.

## DANKSAGUNG

Die Autor:innen bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. Das dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr – ist ein von den Universitäten der Bundeswehr Hamburg und München gemeinsam getragenes wissenschaftliches Zentrum und Bestandteil des Konjunkturprogramms der Bundesregierung zur Überwindung der COVID-19- Krise. Es unterliegt der akademischen Selbstverwaltung. Die Mittel, mit dem das dtec.bw ausgestattet wurde, werden an beiden Universitäten der Bundeswehr zur Finanzierung von Forschungsprojekten und Projekten zum Wissens- und Technologietransfer eingesetzt. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

## LITERATUR

- [1] S. Kuhlmann und M. Heuberger, „Digital transformation going local: implementation, impacts and constraints from a German perspective“, *Public Money & Management*, Nr. Online First: <https://doi.org/10.1080/09540962.2021.1939584>, 2021.
- [2] J. Marienfeldt, J. Kühler, S. Kuhlmann und I. Proeller, „Kommunale Verwaltungs-digitalisierung im föderalen Kontext: ein europäischer Ländervergleich“, *dms*, Bd. 17, Nr. 1, pp. 35-59, 2024.
- [3] B. Distel, „Digitalwüste Deutschland. Digitalisierung im internationalen Vergleich“, *APuZ*, pp. <https://www.bpb.de/shop/zeitschriften/apuz/digitale-gesellschaft-2022/505684/digitalwueste-deutschland/>, 2022.
- [4] J. Redmann und D. Rückel, „Die digitale Transformation kommunaler Einrichtungen – Herausforderungen und Erfolgsfaktoren“, *HMD Praxis der Wirtschaftsinformatik*, Bd. 58, pp. 978-990, 2021.
- [5] C. Fischer, M. Heuberger und M. Heine, „The impact of digitalization in the public sector: a systematic literature review“, *dms*, Bd. 14, Nr. 1, pp. 3-23, 2021.
- [6] J. Boger und C. Wölbert, „Digitalisierung der Verwaltung: 32 Städte im Ranking“, *c't*, pp. <https://www.heise.de/hintergrund/Digitalisierung-der-Verwaltung-32-Staedte-im-c-t-Ranking-7465640.html>, 23 Januar 2023.
- [7] F. Felden, T. Zeit, P. Bauer und S. Siegert, „Schneller, einfacher, nutzerorientierter – drei zentrale Forderungen an die Digitalisierung der deutschen Behördendienste“, BCG, 2023. [Online]. <https://web-assets.bcg.com/81/0d/1f7eae994e27ad941daf3a47edc2/bcg-digital-government-survey-2023.pdf>. [Zugriff am 2024 6 18].
- [8] S. Halsbenning, „Digitalisierung öffentlicher Dienstleistungen: Herausforderungen und Erfolgsfaktoren der OZG-Umsetzung in der Kommunalverwaltung“, *HMD Praxis der Wirtschaftsinformatik*, Bd. 58, pp. 1038-1053, 2021.
- [9] U. Lohmann, *Architekturen der Verwaltungsdigitalisierung*, Springer, 2021.
- [10] M. Groß und A. Krellmann, „Verwaltungsdigitalisierung integriert betrachten“, in *Handbuch E-Government*, Springer, 2019.
- [11] T. Kuder, B. Manteuffel, G. Schaal, C. Schaefer, S. Dumm, T. Perry und G. Serfas, „Spannungsfelder und Erfolgsfaktoren kommunaler Digitalisierung: Erste Ergebnisse der Einführung des digitalen Bauantrages in Ludwigsburg“, in *dtec.bw-Beiträge des HSU*, Bde. %1 von %2368-377, Hamburg, 2023.
- [12] P. Muckel, „Die Entwicklung von Kategorien mit der Methode der Grounded Theory“, *Historical Soccial Research, Supplement*, Bd. 19, pp. 211-231, 2007.
- [13] A. Schmidt, *Verwaltung, eGovernment und Digitalisierung*, Springer, 2019.
- [14] M. Fischer und J. Weiß, „Stand und Treiber der Verwaltungsdigitalisierung in deutschen Kommunalverwaltungen“, in *Neue Herausforderungen im Verwaltungs- und öffentlichen Sicherheitsmanagement*, Nomos, 2023.
- [15] I. Mergel, „Digitale Transformation als Reformvorhaben der deutschen Verwaltung“, *dms*, Bd. 12, Nr. 1, pp. 162-171, 2019.
- [16] B. Friedländer und S. Vedder, „Innovation und Verwaltungsreform“, *dms*, Bd. 17, Nr. 1, pp. 3-12, 2024.
- [17] J. Bogumil und P. Gräfe, „Verwaltungsverflechtungen im föderalen System“, *PVS*, pp. <https://doi.org/10.1007/s11615-023-00525-8>, 2023.
- [18] T. Kuder und B. Manteuffel, „Koordinationsherausforderungen der Digitalisierung in der Kommunalverwaltung“, *dms*, Bd. 17, Nr. 1, pp. 148-175, 2024.
- [19] P. Gräfe, „Die Kommune im Geflecht der Online-Verwaltungsportale: A Tale of Too Many Portals“, *dms*, Bd. 17, Nr. 1, pp. 125-147, 2024.
- [20] M. Rackwitz, T. Hustedt und G. Hammerschmid, „Digital transformation: from hierarchy to network-based collaboration? The case of the German “Online Access Act“,“ *dms*, Bd. 14, Nr. 1, pp. 101-120, 2021.
- [21] L. Dorenbusch, G. Auth und C. Pflüger, „Prozessorientierte Vertrauensniveaubestimmung für digitale Verwaltungsleistungen der kommunalen Ebene“, *HMD Praxis der Wirtschaftsinformatik*, pp. <https://link.springer.com/article/10.1365/s40702-023-00950-8>, 2023.
- [22] J. Hofmann, „Verwaltungsdigitalisierung als politisches Sehnsuchtsprojekt“, in *Regieren in der Transformationsgesellschaft*, Springer, 2023, pp. 187-195.
- [23] A. Benz, *Regieren in komplexen Regelsystemen*, Wiesbaden, 2004.
- [24] A. Benz, „Politische Steuerung in lose gekoppelten Mehrebenensystemen“, in *Gesellschaftliche Komplexität und kollektive Handlungsfähigkeit*, Frankfurt/New York, Campus, 2000, pp. 97-124.

- [25] R. Putnam, „Diplomacy and domestic politics: the logic of two-level games,“ *IO*, Bd. 42, Nr. 3, pp. 427-460, 1988.
- [26] I. Mergel, „Digitale Transformation als Reformvorhaben der deutschen Verwaltung,“ Bd. 12, Nr. 1, pp. 162-171, 2019.
- [27] Stadt Ludwigsburg, „Innovationsnetzwerk Living Lab Ludwigsburg. Bericht 2015 bis 2019,“ 2019. [Online]. Available: [https://www.ludwigsburg.de/site/Ludwigsburg-Internet-2020/get/params\\_E-1021459599/18244955/LivLab\\_Brosch%C3%BCre\\_DRUCK\\_v1.pdf](https://www.ludwigsburg.de/site/Ludwigsburg-Internet-2020/get/params_E-1021459599/18244955/LivLab_Brosch%C3%BCre_DRUCK_v1.pdf). [Zugriff am 2024 4 20].
- [28] H. Frieling, R. Nell, E. Ottendörfer, L. Helmrich, E. Schmitz, *Interkommunale Zusammenarbeit*, Fraunhofer IAO, 2022.
- [29] S. Stängle, „Die Erstellung kommunaler Digitalisierungskonzepte,“ in *Smart Cities, Digitalisierung und Bürgerbeteiligung*, Springer, 2023.
- [30] B. Bundesministerium für Bildung, Forschung, „Zukunftsstadt: Strategische Forschungs- und Innovationsagenda,“ 2015. [Online]. Available: <https://www.fona.de/medien/pdf/Zukunftsstadt.pdf>. [Zugriff am 18 4 2023].
- [31] B. f. U. N. B. u. BMUB, „Smart City Agenda,“ 2017. [Online]. Available: <https://www.smart-city-dialog.de/wp-content/uploads/2019/12/smart-city-charta-langfassung.pdf>. [Zugriff am 18 4 2023].
- [32] J. Bogumil, S. Kuhlmann, M. Heuberger und J. Marienfeldt, „Bürgernahe Verwaltung digital?,“ 2022. [Online]. Available: <https://library.fes.de/pdf-files/a-p-b/19351.pdf>. [Zugriff am 18 4 2023].
- [33] A. Hettich, „Der neue OB will in Ludwigsburg nicht alles anders machen,“ 2022. [Online]. Available: <https://www.stimme.de/ueberregional/baden-wuerttemberg/nachrichten/pl/der-neue-ob-will-in-ludwigsburg-nicht-alles-anders-machen-art-4216737>. [Zugriff am 18 4 2023].
- [34] T. Kuder, *Bürgerbeteiligung im Prozess der Digitalisierung. Das Modellprojekt Ludwigsburg.*, Berlin: vhw werkSTADT Nr. 22, 2018.
- [35] T. Kuder, „Herausforderungen der Digitalisierung für benachteiligte Stadtquartiere,“ in *Kursbuch Bürgerbeteiligung #3*, Berlin, 2018.
- [36] T. Helfen, T. Kuder und B. Manteuffel, *Herausforderungen der Digitalisierung für benachteiligte Stadtquartiere*. vhw-Schriftreihe Nr. 17, Berlin, 2020.
- [37] T. Kuder, „Demokratie 4.0 - Beteiligung und Mitverantwortung im Zeichen der Digitalisierung,“ in *Demokratie 4.0*, Berlin, 2021, pp. <https://gemeinsamberlin.blog/publikation-demokratie-4-0/>.
- [38] T. Kuder und B. Manteuffel, „Urbane Daten in kommunalen Spannungsfeldern,“ *Forum Wohnen und Stadtentwicklung*, Bd. 1, pp. 17-20, 2023.
- [39] V. Mayer, „Noch ein Abgang in Ludwigsburg. Der neue Digital-Offizier geht von Bord,“ 2019. [Online]. Available: <https://www.stuttgarter-zeitung.de/inhalt.noch-ein-abgang-in-ludwigsburg-der-neue-digital-offizier-geht-von-bord.52691312-a8fb-482d-936f-9b4a08ea9860.html> [Zugriff am 05 11 2024].
- [40] C. Walf, „Schlechte Noten für Bürgerservice,“ 2020. [Online]. Available: [https://www.lkz.de/lokales/stadt-ludwigsburg\\_artikel-schlechte-noten-fuer-buergerservice-\\_arid,584442.html](https://www.lkz.de/lokales/stadt-ludwigsburg_artikel-schlechte-noten-fuer-buergerservice-_arid,584442.html) [Zugriff am 05 11 2024].

# Representativeness in Business Surveys and Linked Employer Employee Data: Evidence from the SOEP-LEE2 Data for Germany

Wenzel Matiaske\*, Christoph Halbmeier, Martina Maas, Torben Dall Schmidt  
*Helmut Schmidt University/University of the Federal Armed Forces Hamburg*  
 Hamburg, Germany  
 \*matiaske@hsu-hh.de

**Abstract** – A central goal of the dtec.bw project SOEP-LEE2 is to provide high-quality data for research in organisational, business, and economic studies. To this end, the project conducted two multi-year surveys among German establishments. The first survey, SOEP-LEE2-Core, is linkable to employee data from the Socio-Economic Panel (SOEP) and is characterized by a complex sampling procedure. The second survey, SOEP-LEE2-Compare, lacks this linkage possibility and has a simpler sampling design. In this article, we evaluate the sampling designs of the two surveys in term of their statistical representativeness. Our results indicate that the more complex sampling design of SOEP-LEE2-Core is susceptible to selective participation in some of the analysed dimensions, resulting in an overly large proportion of employees from the public sector. In contrast, the simpler design of SOEP-LEE2-Compare is less prone to this type of selection and yields more representative data.

**Keywords** – Survey design, data linkage, employee-first sampling, representativeness, Socio-Economic Panel (SOEP)

## NOMENCLATURE

FEA	Federal Employment Agency
HR	Human resources
IT	Information technology
SOEP	Socio-Economic Panel
SOEP-LEE2	Socio-Economic Panel Linked Employee-Employer Version 2

## I. INTRODUCTION

Socio-economic research increasingly analyses data that are obtained from various sources and linked together. Data linkage enriches the information available in single datasets, enhancing their research potential. If datasets with comparable content are linked, data linkage can also help to assess and improve data quality. The benefits of data linkage can come at the cost of losing statistical representativeness. Not every observation of an initially representative dataset may be linkable. If these observations have different characteristics than the linkable ones, the linked dataset is not representative any longer. As a consequence, analytical findings cannot be easily extrapolated to the initial population, weakening the generalizability of the analysis. The causes, magnitude, and consequences of selective linkage depend on the specific case

at hand and require empirical assessment to determine suitable remedies.

In this article, we assess the statistical representativeness of two establishment surveys that are central to the dtec.bw project SOEP-LEE2 [1], [2]. The first survey, SOEP-LEE2-Core, is a linked employer employee survey with a sampling design that follows the employee-first approach [3], [4]. The survey collects data from establishments that employ individuals taking part in the German Socio-Economic Panel study (SOEP-Core). Every employee in SOEP-Core was asked for contact data of their employer and, if provided, employers were invited to participate in the SOEP-LEE2-Core survey. The sampling procedure is more complex than those of typical establishment surveys and involves different stages of selective participation that can erode statistical representativeness.

We assess the representativeness of the SOEP-LEE2-Core data by comparing them to administrative data from the employment statistics of the Federal Employment Agency (FEA). We find that the SOEP-LEE2-Core data are representative in some, but not all, dimensions. Employees in SOEP-LEE2-Core distribute across East and West Germany in proportions that are comparable to the FEA data. Likewise, they present a good representation of the population in terms of the establishment size, although employees from very large establishments are somewhat underrepresented. In contrast, employees from the public sector are relatively abundant in the SOEP-LEE2-Core data, suggesting that the survey's sampling and linking approach is susceptible to this type of selection.

To understand the selection mechanisms further, we analyse the representativeness of the project's second survey, SOEP-LEE2-Compare. This survey was conducted among establishments drawn as a random sample from the FEA register data. Different to SOEP-LEE2-Core, the sampling procedure did not include the selective step of collecting employer contact information. Fewer steps of selection should, all else equal, lead to more representative data. Indeed, we find that the establishments surveyed for SOEP-LEE2-Compare resemble the population of the FEA data very well. Although the SOEP-LEE2-Compare data also contain a relatively large proportion of public sector establishments, the imbalance is substantially smaller than in SOEP-LEE2-Core.

Our findings are corroborated by an additional analysis of the SOEP-LEE2-Core survey in [5]. There we find that

employees working in the public sector are more willing to reveal employer contact information. In addition, public sector employers are more likely to take part in the subsequent survey. The two effects add up, affecting the representativeness of SOEP-LEE2-Core in this dimension.

In the next section, we provide an overview of the related literature on representativeness in business surveys. In section III, we describe the sampling procedure of SOEP-LEE2-Core and show how the collected data compare to the FEA data. In section IV, we examine the sampling design and representativeness of the SOEP-LEE2-Compare survey. In the last section, we provide some concluding remarks.

## II. THE REPRESENTATIVENESS OF BUSINESS SURVEYS IN THE LITERATURE

In order to generalize the results of a survey to the population, the participants in the survey must be representative of the population. The risk of systematic bias in the sample increases with the proportion of non-participants in a sample [6]. This risk has increased in recent years as the willingness to participate in business surveys has been declining for several years [7], [8]. However, the declining willingness to participate does not necessarily lead to sample bias, but only if non-participation depends on the characteristics of the establishment and is therefore not random. There are two forms of unit nonresponse: non-participation can be due to failure to contact the selected participant and non-participation can be a result of refusal to participate [9], [10]. Studies on non-response indicate that the probability of participation in a survey depends on several factors. These factors fall into four categories: 1. environment, 2. participant, 3. survey design, 4. interviewer [11], [12], [10]. The researcher can influence the reduction of non-response, particularly through the choice of survey design and the selection and training of interviewers [13], [14], [15]. At the environmental or organisational level, the researcher has no influence. Factors at the organisational level that have been identified as influencing non-response include the organisational structure, corporate culture, the internal distribution of information, and the sector [16], [17]. In this context, studies have shown that older and smaller establishments are more likely to participate in a survey than younger and larger establishments [8]. The fact that larger establishments are less likely to participate in surveys is due to the coordination effort required, as several departments and levels of hierarchy are involved in responding to the questionnaire, depending on the focus of the topic [18]. Willingness to participate in surveys also increases in good economic times [19]. With regard to the respondent, studies show that the likelihood of participating in a survey depends on the respondent's knowledge, qualifications, attitudes, and commitment [10], [20]. Respondents can include senior management as well as managers and staff from specific departments such as HR or IT. In some cases, especially when the survey covers several topics, it may be necessary to involve several employees with specific expertise, which can increase the risk of non-response [18]. It should be noted that sampling bias is not uncommon in surveys and that the investigation of non-response requires special attention in personnel and organizational research.

## III. SOEP-LEE2-CORE

### A. Sampling design of SOEP-LEE2-Core

The sampling design of the SOEP-LEE2-Core survey is closely linked to the SOEP-Core study. SOEP-Core is a panel study of German individuals that every year collects survey data about topics such as income, educational attainments, health, and well-being. SOEP-Core regularly receives new samples to compensate for panel attrition and, in some cases, to collect data from specific populations of interest. The various samples are weighted with survey weights to obtain data that are representative for the adult population living in Germany [21].

In its 2021 and 2023 waves, SOEP-Core provided the sampling frame for SOEP-LEE2-Core, following the employee-first sampling procedure [3], [4]. This consisted in asking dependently employed individuals in SOEP-Core about the name and address of their employer with the aim to subsequently survey them. Employees in establishments with fewer than five employees were not considered for sampling due to privacy reasons. In the 2021 wave, the procedure yielded a sample of 1,063 individuals with a successful linkable employer interview, starting from a total of 8,343 eligible individuals. The reduction in sample size was due to two factors. First, some employees had to be excluded because they refused to provide employer contact information. Second, the response rate in the establishment survey was 17.8 percent, resulting in a further reduction of the sample size.

### B. Representativeness of SOEP-LEE2-Core

The diminution of the initial sample begs the question of whether the remaining individuals still represent the total population of employees in Germany well. We assess this by comparing the SOEP-Core and SOEP-LEE2-Core 2021 data to aggregate statistics from the establishment file of the Federal Employment Agency (FEA data for short) [22]. The FEA data are a reliable benchmark because they are assembled from the obligatory employment notifications that employers send to the social security system. Therefore, they comprise all German establishments with at least one employee subject to social security contributions. For the comparison, we use the data as of 30 June 2021, and we exclude establishments with fewer than five employees beforehand. In this section, we use the person-level version of the data, which is the level SOEP-LEE2-Core was designed to be representative for.

The comparison is affected, to a minor degree, by small differences between the FEA data and the sampling frame of SOEP-LEE2-Core. SOEP-LEE2-Core includes civil servants, persons in marginal employment, and family members working in family businesses, whereas the FEA data are limited to employed persons subject to social security contributions [22], [23]. Moreover, SOEP-LEE2-Core includes individuals who lost their job within the six months prior to the SOEP-Core interview, while the FEA data do not contain these recent unemployed. Lastly, there are minor differences in the reference period. SOEP-Core interviews took place between May 2021 and January 2022, whereas the FEA data refer to 30 June 2021. The minor incongruencies are reflected in the total counts of employees in both datasets. The FEA data cover 31,339,827 employees in establishments with at least five employees, whereas SOEP-Core comprises a weighted number of 33,356,544 employees. Despite the differences, we consider the FEA data to be a suitable

benchmark due to its administrative nature and coverage of the grand majority of employees in Germany.

In Table I, we show how employees are distributed in the different datasets. The first column contains the distribution in the FEA data, the second column shows the distribution among those 8,343 SOEP-Core individuals eligible for sampling, and the third column refers to SOEP-LEE2-Core, that is, individuals with a successful and linkable establishment interview. The data shown in columns (2) and (3) are weighted with SOEP survey weights to compensate for the disproportional sampling of SOEP-Core.

TABLE I: DISTRIBUTION OF EMPLOYEES IN THE FEA, SOEP-CORE, AND SOEP-LEE2-CORE DATA IN PERCENT.

	(1)	(2)	(3)	(4)
	<i>FEA data</i>	<i>SOEP-Core</i>	<i>SOEP-LEE2-Core</i>	<i>Difference (1) - (3)</i>
<b>Establishment size</b>				
5-9	8.1	7.4	9.6	-1.5
10-49	25.6	25.2	25.0	0.6
50-249	30.0	27.0	31.8	-1.8
250-499	11.6	10.4	11.0	0.6
500 and more	24.7	26.1	21.2	3.5
Information missing	0.0	4.0	1.5	-1.5
Total	100.0	100.1	100.1	
<b>Region</b>				
East Germany incl. Berlin	18.3	19.1	20.2	-1.9
West Germany	81.7	80.9	79.8	1.9
Total	100.0	100.0	100.0	
<b>Industry</b>				
Agriculture, forestry, fishing	0.6	0.5	0.4	0.2
Mining, energy and water supply, sewage	1.8	1.7	2.0	-0.2
Food, beverages, tobacco, consumer goods	3.6	3.1	3.7	-0.1
Production goods, capital goods, and consumer durables	17.5	14.6	11.6	5.9
Construction	5.4	3.3	3.0	2.4
Wholesale trade	6.1	1.7	3.9	2.2
Retail trade	7.0	6.1	5.6	1.4
Transportation and storage	5.8	4.4	3.3	2.5
Accommodation and food service activities	2.5	1.5	2.1	0.4
Information and communication, financial and insurance activities, business, scientific and professional services	20.8	12.6	12.8	8.0
Public administration, defence, social security, education, human health, and social work	26.0	28.7	43.2	-17.2
Other services	2.9	2.0	4.1	-1.2
Information missing	0.1	19.6	4.3	-4.2

	(1)	(2)	(3)	(4)
	<i>FEA data</i>	<i>SOEP-Core</i>	<i>SOEP-LEE2-Core</i>	<i>Difference (1) - (3)</i>
Total	100.0	99.8	100.0	

Note: In case of SOEP-Core, the region of residence is shown, whereas for the other datasets, the region of the workplace is shown.

The upper pane of the table shows the distribution across the establishment size. In most size categories, the share of employees is very similar in the three datasets. For instance, in the category of establishments with 10 to 49 employees, the share is 25.6 percent according to the FEA data, and 25.0 percent in the SOEP-LEE2-Core data. In the category of establishments with 500 and more employees, the SOEP-LEE2-Core data have a share that is by 3.5 percentage points lower than that of the FEA data, indicating that employees from very large establishments are slightly underrepresented in SOEP-LEE2-Core. This is probably acceptable for most analyses and overall, the SOEP-LEE2-Core data are a good representation of the employee population in terms of the establishment size.

The distribution of employees across regions is again similar in all three datasets. 18.3 percent of employees work in East Germany according to the FEA data, 19.1 percent according to SOEP-Core, and 20.2 percent according to SOEP-LEE2-Core. The differences in percentage shares may relate to the fact that establishments in East Germany have on average fewer employees than West German establishments. Overall, the difference is relatively small, indicating a good regional balance in the SOEP-LEE2-Core data.

The imbalances are more substantial in the industry categories. Most prominently, it is in the category of “public administration, defence, social security, education, human health and social work” industries, where the largest differences between the datasets exist. According to the FEA data, 26.0 percent of employees work in these industries, whereas the SOEP-LEE2-Core data contains 43.2 percent of them. In SOEP-Core, the industry variable has a large share of missing values and provides evidence that is not fully conclusive. However, SOEP-Core shows a share that is much closer to the FEA data (28.7 percent), suggesting that the imbalances are introduced by the sampling design of SOEP-LEE2-Core. Indeed, in [5] we show that employees from the public sector are significantly more likely to report their employer contact information. In addition, public sector employers are also more likely than their private sector counterparts to participate in the subsequent establishment survey, resulting in significant differences in the industry composition. This is something that has to be taken into account when analysing the SOEP-LEE2-Core data.

#### IV. SOEP-LEE2-COMPARE

The SOEP-LEE2 project comprised a second establishment survey, SOEP-LEE2-Compare. It aimed at, among other things, assessing the design and representativeness of traditional establishment surveys with a simple sampling design compared to linkable surveys and employer-first sampling.

### A. Sampling design of SOEP-LEE2-Compare

Different to SOEP-LEE2-Core, SOEP-LEE2-Compare used traditional random sampling and forwent the additional stage of selection inherent in the employer-first sampling. Establishments were drawn at random from the establishment file of the FEA described in the previous section. Before sampling, establishments with less than five employees were excluded for comparability with SOEP-LEE2-Core that did not contain these establishments either. The remaining establishments were divided into 120 strata along three dimensions: 1) five categories of firm size 2) two regions, and 3) 12 categories of industries as depicted in Table I and II. Larger establishments were drawn with disproportional higher probability as we expected them to respond less often than small ones.

The drawn sample consisted of 25,973 establishments that subsequently were contacted and asked to participate in the survey. As in SOEP-LEE2-Core, establishments could choose between answering via telephone or web interview. The survey questions were identical to those of the SOEP-LEE2-Core survey and a total of 2,026 establishments responded, corresponding to a response rate of 7.8 percent.

### B. Representativeness of SOEP-LEE2-Compare

Does a simpler survey design lead to more representative data? To answer this question, we assess the representativeness of SOEP-LEE2-Compare by comparing the surveyed establishments to the total of establishments in the FEA data. The comparison is straightforward because establishments were sampled from these data, and we can compare them in the same dimensions that were used for stratification and are available for all establishments. To take into account the disproportional sampling, we weight the surveyed establishments with the inverse of the sampling probability. Then, establishments in SOEP-LEE2-Compare should show the same distribution of characteristics as all FEA establishments if their response to the survey is not selective.

The results of the comparison are presented in Table II. The upper part of the table shows that establishments surveyed for SOEP-LEE2-Compare have a very similarly distribution of the establishment size than those in the FEA data. For instance, 42.3 percent of the establishments in SOEP-LEE2-Compare belong to the category of establishments with 5 to 9 employees. The percentage is 43.4 in the FEA data. Similarly, there are only very minor deviations between the two datasets among larger establishments, indicating that the surveyed establishments are very representative in terms of size. The regional distribution of establishments resembles each other in the two datasets as well, underscoring the representativeness of the data.

In terms of industries, the results indicate that the SOEP-LEE2-Compare data are somewhat skewed to industries in the category “public administration, defence, social security, education, human health and social work”. They constitute 25.9 percent of the surveyed establishments relative to 19.6 percent in the FEA data. Similar to SOEP-LEE2-Core, the difference suggests that public sector establishments were more likely to consent to the survey. However, the difference is not as large as in SOEP-LEE2-Core, where the extra step of collecting employer contact information additionally distorted the data towards public sector employers. Yet, analyses based

on the SOEP-LEE2-Compare may take into account this distortion towards establishments from the public sector.

## V. CONCLUSION

In this article, we evaluate how the sampling design of different establishment surveys affects their respective statistical representativeness. The first survey, SOEP-LEE2-Core, is a linked employer employee survey whose sampling design involves two stages of selective participation. We show that the collected linkable data is representative in some, but not all dimensions, by comparing them to data from an administrative register. The second survey, SOEP-LEE2-Compare, has a simpler sampling design, but otherwise features the same design characteristics as SOEP-LEE2-Core. The collected data match the distributional characteristics of the register data better than SOEP-LEE2-Core, suggesting that a simpler sampling design is beneficial for statistical representativeness.

TABLE II: DISTRIBUTION OF ESTABLISHMENTS IN THE FEA AND SOEP-LEE2-COMPARE DATA IN PERCENT.

	(1)	(2)	(3)
	<i>FEA data</i>	<i>SOEP-LEE2-Compare</i>	<i>Difference (1) - (2)</i>
<b>Establishment size</b>			
5-9	43.4	42.3	1.1
10-49	44.2	44	.2
50-249	10.5	12	-1.5
250-499	1.2	1.1	.1
500 and more	.7	.6	.1
Total	100	100	0
<b>Region</b>			
East Germany incl. Berlin	19.7	20.1	-.4
West Germany	79.2	79.9	-0.7
Total	100	100	0.0
<b>Industry</b>			
Agriculture, forestry, fishing	1.3	1.6	-0.3
Mining, energy and water supply, sewage	1.2	1.1	0.1
Food, beverages, tobacco, consumer goods	2.8	3.9	-1.1
Production goods, capital goods, and consumer durables	8.3	8.2	0.1
Construction	11.3	8.6	2.7
Wholesale trade	7.9	6.1	1.8
Retail trade	11.8	9.5	2.3
Transportation and storage	4.8	2.9	1.9
Accommodation and food service activities	5.6	5.2	0.4
Information and communication, financial and insurance activities, business, scientific and professional services	20.7	20.0	0.7

	(1)	(2)	(3)
	FEA data	SOEP-LEE2-Compare	Difference (1) – (2)
Public administration, defence, social security, education, human health, and social work	19.6	25.9	-6.3
Other services	4.7	6.9	-2.2
Information missing	0.1	0.0	0.1
Total	100	99.9	0.0

Our results have a number of important implications for users of linked datasets as well as researchers who want to design linkable surveys. First, data linkage often requires the consent of the persons whose data should be linked. Giving consent can be selective, underscoring the importance of careful formulation of data linkage consents. Second, in many cases the linked sample will be smaller than the original sample. The selection process should be analysed thoroughly. Statistical weighting can rebalance the linked sample, which should be feasible, as rich data for modelling the selection process are typically available. Third, there are instances where none of the above is possible, which does not invalidate the analyses on the linked dataset. However, generalizations to the overall population should be made with great caution.

#### ACKNOWLEDGEMENT

This research and the SOEP-LEE2 project are funded by dtec.bw – Digitalization and Technology Research Centre of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

#### REFERENCES

- [1] W. Matiaske, T. D. Schmidt, C. Halbmeier, M. Maas, C. Schröder, T. Böhm, A. S. Kritikos and S. Liebig, „The Socio-Economic Panel Linked Employer- Employee Survey Version 2 (SOEP-LEE2): Overview and Results from the First Wave,“ in *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg: Forschungsaktivitäten im Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr dtec.bw*, Bd. 1, D. Schulz, A. Fay, W. Matiaske and M. Schulz, Hrsg., Hamburg, 2022, pp. 378-382.
- [2] W. Matiaske, T. D. Schmidt, C. Halbmeier, M. Maas, D. Holtmann, C. Schröder, T. Böhm, S. Liebig and A. S. Kritikos, „SOEP-LEE2: Linking Surveys on Employees to Employers in Germany,“ *Jahrbücher für Nationalökonomie und Statistik Data Observer*, 2023.
- [3] M. Weinhardt, P. Jacobebbinghaus and S. Liebig, „Linked-Employer-Employee (LEE) Daten in der Organisationsforschung,“ in *Handbuch empirische Organisationsforschung*, S. Liebig, W. Matiaske and S. Rosenbohm, Hrsg., Wiesbaden, Springer Gabler, 2016, p. 1–28.
- [4] M. Weinhardt, A. Meyermann, S. Liebig and J. Schupp, „The Linked Employer–Employee Study of the Socio-Economic Panel (SOEP-LEE): Content, Design and Research Potential,“ *Jahrbücher für Nationalökonomie und Statistik*, Bd. 237, Nr. 5, pp. 457-467, 2017.
- [5] W. Matiaske, T. D. Schmidt, C. Halbmeier and M. Maas, „Nonresponse and Non-contact in Linked Employee-Employer Surveys: Evidence from the German Socio-Economic Panel Linked Employer-Employee Survey (SOEP-LEE2),“ *Working Paper*, 2024.
- [6] P. L. Alreck and R. B. Settle, *The Survey Research Handbook: Guidelines and Strategies for Conducting a Survey*, Irwin Professional Publishing, 1995.
- [7] D. I. Pielsticker and M. R. Hiebl, „Survey Response Rates in Family Business Research,“ *European Management Review*, Bd. 17, Nr. 1, p. 327–346, 2020.
- [8] C. König and J. W. Sakshaug, „Nonresponse trends in establishment panel surveys: findings from the 2001–2017 IAB establishment panel,“ *Journal for Labour Market Research*, Bd. 57, Nr. 23, 2023.
- [9] R. M. Groves and E. Peytcheva, „The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis,“ *Public Opinion Quarterly*, Bd. 72, Nr. 2, p. 167–189, 2008.
- [10] F. Janik and S. Kohaut, „Why don’t they answer? Unit non-response in the IAB establishment panel,“ *Quality & Quantity*, Bd. 46, p. 917–934, 2012.
- [11] R. M. Groves and M. P. Couper, *Nonresponse in Household Interview Surveys*, New York: John Wiley & Sons (Wiley series in survey methodology), 1998.
- [12] D. K. Willimack, E. Nichols and S. Sudman, „Understanding unit and item nonresponse in business surveys,“ in *Survey nonresponse*, R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little, Hrsg., John Wiley & Sons, 2002, p. 213–227.
- [13] National Research Council, *Nonresponse in Social Science Surveys: A Research Agenda*, R. Tourangeau and T. J. Plewes, Hrsg., Washington, D.C.: The National Academies Press, 2013.
- [14] D. K. Willimack and J. L. Dalzell, „An Examination of Non-Contact as a Source of Nonresponse in a Business Survey,“ *AAPOR - ASA Section on Survey Research Methods*, 2006.
- [15] P. Gleiser, J. W. Sakshaug, M. Volkert, P. Ellguth, S. Kohaut and I. Möller, „Introducing Web in a Mixed-Mode Establishment Survey: Effects on Nonresponse,“ *Journal of the Royal Statistical Society Series A (Statistics in Society)*, Bd. 185, Nr. 3, p. 891–915, 2022.
- [16] D. Tomaskovic-Devey, J. Leiter und S. Thompson, „Organizational Survey Nonresponse,“ *Administrative Science Quarterly*, Bd. 39, Nr. 3, pp. 439-457, 1994.
- [17] D. Tomaskovic-Devey, J. Leiter und S. Thompson, „Item Nonresponse in Organizational Surveys,“ *Sociological Methodology*, Bd. 25, pp. 77-110, 1995.
- [18] D. K. Willimack und G. Snijkers, „The Business Context and its Implications for the Survey Response Process,“ in *Designing and Conducting Business Surveys*, G. Snijkers, G. Haraldsen, D. Willimack und J. Jones, Hrsg., Hoboken, NJ, John Wiley & Sons, 2013, p. 39–82.
- [19] C. Seiler, „The determinants of unit non-response in the Ifo Business Survey,“ *ASta Wirtschafts- und Sozialstatistisches Archiv*, Bd. 8, p. 161–177, 2014.
- [20] B. R. Fulton, „Organizations and Survey Research: Implementing Response Enhancing Strategies and Conducting Nonresponse Analyses,“ *Sociological Methods & Research*, Bd. 47, Nr. 2, p. 240–276, 2018.
- [21] J. Goebel, M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder und J. Schupp, „The German Socio-Economic Panel (SOEP),“ *Jahrbücher für Nationalökonomie und Statistik*, Bd. 239, Nr. 2, pp. 345-360, 2019.
- [22] U. Brix und M. Fritsch, „Die Betriebsdatei der Beschäftigtenstatistik der Bundesanstalt für Arbeit,“ in *Das Gründungsgeschehen in Deutschland. Darstellung und Vergleich der Datenquellen*, M. Fritsch und R. Grotz, Hrsg., Heidelberg, Physica-Verlag, 2002, pp. 55-78.
- [23] Bundesagentur für Arbeit, *Grundlagen: Qualitätsbericht – Statistik der sozialversicherungspflichtigen und geringfügigen Beschäftigung. Version 7.12*, Nürnberg: Bundesagentur für Arbeit, 2022.



gefördert durch



[dttecbw.de](https://dttecbw.de)

Druckversion: ISBN 978-3-86818-329-0 · Digitale Version: ISBN 978-3-86818-317-7

DOI: <https://doi.org/10.24405/16768>