

Barbara Klein (Hrsg.)

Künstliche Intelligenz im Healthcare-Sektor



Künstliche Intelligenz im Healthcare-Sektor

Inhaltsverzeichnis

1. Einleitung

- Künstliche Intelligenz im Healthcare-Sektor Seite 5
Barbara Klein
- KI an der Frankfurt UAS – Bestandsaufnahme und Zukunftsperspektive Seite 10
Susanne Rägler, Susanne Klüber

2. Die Technische und Produktperspektive

- Lernen, Sehen und Orientieren in autonomen Assistenzsystemen Seite 15
Ute Bauer-Wersing
- Informed Machine Learning für die Produktion von Vliesstoffen in der Medizintechnik Seite 20
Paulami Banerjee, Simone Gramsch
- KI als Chance im Gesundheitswesen: Von der Theorie zum maßgeschneiderten Produkt Seite 25
Diana Völz, Julia Schneider
- Kognitive Prozesse in der KI-unterstützten diagnostischen Entscheidungsfindung Seite 31
Ekaterina Jussupow
- Mensch-KI-Interaktion im Gesundheitswesen Seite 35
Luisa Pumplun

3. Die Versorgungsperspektive

- Soziale Arbeit*
- Wie wollen wir mit KI leben? – Systeme Künstlicher Intelligenz in der Sozialen Arbeit Seite 39
Anastasia Paschalidou
- Behindertenhilfe*
- Nutzer*innenakzeptanz als Erfolgsfaktor und Risiko bei KI-Anwendungen im Bereich Healthcare Seite 45
Annalies Baumeister, Olga Kozlova, Elizaveta Gardó, Patrizia Tolle
- Transformativ forschen mit Lern- und Experimentierräumen für KI-gestützte Assistenzsysteme in der beruflichen Rehabilitation von Menschen mit Behinderungen Seite 55
Barbara Lippa

Künstliche Intelligenz im Versorgungskontext von Menschen mit psychischen Erkrankungen: Innovative Ansätze für Betreuung und Unterstützung Leya Küsters, Melanie Schmidt, Barbara Klein	Seite 59
<i>Altenhilfe</i>	
Einsatzmöglichkeiten für KI in der ambulanten Pflege Miriam Peters	Seite 64
Versorgungsperspektive in der Altenpflege Lisa Luft	Seite 69
<i>Krankenhaus</i>	
Patient:innen und KI: Eine Frage der Perspektive bei der Bewertung von KI bei medizinischen Online-Diensten Eva Lermer, Susanne Gaube, Julia Cecil, Anne-Kathrin Kleine, Eesha Kokje, Dieter Frey, Matthias F. C. Hudecek	Seite 74
Neue methodische Ansätze zur Erforschung von Trends in der KI-gestützten mentalen Gesundheitsversorgung Anne-Kathrin Kleine	Seite 78
Einflussfaktoren auf die Nutzung KI-gestützter Technologien in der psychischen Gesundheitsversorgung Julia Cecil, Eesha Kokje, Susanne Gaube, Eva Lermer, Anne-Kathrin Kleine	Seite 83
Generative KI im Krankenhauswesen: Eine wertschöpfungskettenorientierte Betrachtung Christiane Saure	Seite 89
<i>Krankenversicherung</i>	
Relevanz von KI Anwendungen aus der Perspektive von Krankenversicherungen Sibel Altin	Seite 95
4. Die juristische und ethische Perspektive	
KI im Konfliktmanagement – rechtliche & ethische Aspekte Olivia Alig	Seite 101
Rechtlicher Rahmen des Einsatzes von KI in der Medizin Susanne Beck	Seite 107
Zur ethischen Evaluation von Technologien am Beispiel von ChatGPT Janina Loh	Seite 111
<i>Impressum</i>	Seite 116

Barbara Klein

Künstliche Intelligenz im Healthcare-Sektor

Multiple Krisen wie der Klimawandel, Kriege oder die Pandemie betreffen besonders stark vor allem die Schwächsten unserer Gesellschaft mit Folgen wie einer unzureichenden Versorgung, höherer Sterblichkeit, aber auch Einsamkeit und Isolation. Der Healthcare Sektor muss sich diesen Herausforderungen wie auch den schon lange bekannten Faktoren demografischer Wandel mit seinem Anstieg an Menschen mit chronischen Erkrankungen und/oder Behinderung sowie fehlendem Personal stellen.

Digitalisierung und insbesondere die Künstliche Intelligenz (KI) können hier neue Potenziale bieten, diesen vielfältigen Herausforderungen zu begegnen. Obwohl die Anfänge der KI in den 30er Jahre des letzten Jahrtausends liegen – der Begriff selbst wurde 1956 von John McCarthy vorgeschlagen und 10 Jahre später mit dem Programm ELIZA der erste Chatbot kreiert – gelang der öffentlichkeitswirksame Durchbruch erst am 30.11.2022 mit ChatGPT (Generative Pre-trained Transformer) der Firma Open AI. Binnen Kürze haben Millionen Menschen Erfahrungen damit gesammelt.

Während ein Großteil der Beschäftigten im Healthcare-Sektor Frauen sind, werden KI-Anwendungen – auch im Healthcare-Sektor – überwiegend von Männern entwickelt, was u.a. dazu führt, dass solche Anwendungen Frauen und diversitätsspezifische Aspekte nicht in einem ausreichenden Maß berücksichtigen mit den Folgen, dass die Produkte auf eine geringe Akzeptanz stoßen und sich auch auf dem internationalen Markt nicht durchsetzen können.

Die UNESCO-Empfehlung zur Ethik künstlicher Intelligenz formuliert weltweit akzeptierte Standards für KI-Technologien, zu denen sich 193 Mitgliedsstaaten bekannt haben. Ethische Regeln sind dabei mit menschenrechtlichen Verpflichtungen verknüpft und es wird ein Fokus auf die sog. Blind Spots wie z. B. KI

und Gender, Bildung, Nachhaltigkeit u. a. gelegt. Für Deutschland wird hier großer Handlungsbedarf bei der Gleichbehandlung und bei der Diversität der KI-Entwickler*innenteams gesehen (Kettemann, M. C. 2022). Diversität gilt als eine der Voraussetzungen, um eine entsprechende Berücksichtigung in der Programmierung von KI erhalten zu können.

Das Feld der Forscher*innen und Programmierer*innen im Bereich der künstlichen Intelligenz ist überwiegend männlich. Laut Wired & Element AI kamen im Jahr 2017 nur 12% der Beiträge auf den drei größten Konferenzen für Machine Learning von Frauen (Simonite, T. 2018). Ebenso sind Gesundheitskongresse weiterhin überwiegend männlich besetzt, wie eine Auflistung der Gäste auf Podiumsdiskussionen der wichtigsten Veranstaltungen aus dem Jahre 2019/2020 zeigt (Gramm, 2021). Dies führt dazu, dass Themen, die vorwiegend Frauen betreffen, weniger repräsentiert werden. Dieser „Gender-Data-Gap“, der beispielweise entsteht, wenn Daten, die nur Frauen betreffen, im geringeren Umfang erhoben werden oder Frauen in Studien unterrepräsentiert sind, führt zu einer Verzerrung in Diagnosen und Prognosen, die von KI-Systemen berechnet werden und somit zu Risiken in der Behandlung von Frauen (Steffens, B. 2021).

Dieses Buch entstand im Kontext des Projektes „Women in the Field of AI in Healthcare – Women AI Days“. Gefördert wurde dieses im Rahmen des Connectom Vernetzungs- und Innovationsfonds von Hessian AI. Hessian AI ist das Hessische Zentrum für Künstliche Intelligenz mit Sitz in Darmstadt, das exzellente Grundlagenforschung mit Praxisbezug betreibt und einen Transfer in Wirtschaft und Gesellschaft vorantreiben möchte. Ziel des Projektes war es Forscherinnen im Kontext KI und Healthcare zu identifizieren und miteinander zu vernetzen und ihnen auch ein öffentliches Gesicht zu verleihen. Dieses erfolgte im Rahmen eines Netzwerk-Workshops im September 2023 und weiteren Netzwerk-

veranstaltungen sowie öffentlichen Vorlesungen an der Frankfurt University of Applied Sciences (Frankfurt UAS), die medial über den Kooperationspartner HR-Info angekündigt wurden.

Für diesen Sammelband konnten 22 KI-Expertinnen gewonnen werden, die sich mit dem Thema aus unterschiedlichen Perspektiven auseinandersetzen. Susanne Rägler, Vizepräsidentin für Forschung, Weiterbildung und Transfer an der Frankfurt UAS skizziert zusammen mit Susanne Klüber die Strategien und Aktivitäten in Forschung und Lehre an der Hochschule im Kontext von KI sowie deren Anstrengungen Geschlechterstereotypen zu überwinden.

Im Themenbereich „Technische und Produktperspektive“ ermöglicht Ute Bauer-Wersing einen Einblick in KI-Methoden für autonome Assistenzsysteme.

Paulami Banerjee und Simone Gramsch setzen sich mit dem Informed Machine Learning als innovative Methode für den Einsatz von KI und Simulation auseinander, um Produktionsprozesse zur Herstellung von medizinischen Masken zu optimieren.

In der Orthopädietechnik werden für die Erstellung von Einlagen die Pedografie und das tradierte Wissen genutzt. Die Forschungsarbeiten von Diana Völz und Julia Schneider zeigen auf, dass KI-Methoden des maschinellen Lernens und der bildgebenden Diagnostik zu einem Mehrwert führen können und eine bessere Einbindung der Patient*innen und Berücksichtigung individueller Anforderungen zu maßgeschneiderten Einlagen, aber auch präventiven Maßnahmen führen.

Ekaterina Jussupow befasst sich mit der Unterstützung von Arzt*innen durch KI bei der medizinischen Diagnostik in der Radiologie.

Luisa Pumplun setzt sich mit den Voraussetzungen des maschinellen Lernens auseinander und wie das organisatorische und technische Umfeld effizient bei gleichzeitiger Berücksichtigung der Bedürfnisse der Menschen im Gesundheitswesen gestaltet werden kann.

Der Themenbereich „Versorgungsperspektive“ blickt auf den Einsatz von KI aus verschiedenen Handlungsfeldern

im Healthcare-Bereich: Von der Sozialen Arbeit, der Behinderten- und Altenhilfe, dem Krankenhauskontext bis hin zur Krankenversicherung werden Anwendungen, Pilotprojekte, Studien und relevante Erkenntnisse aus der Literatur diskutiert.

Anastasia Paschalidou thematisiert dabei in ihrem Beitrag den Einsatz realer sowie potenzieller Anwendungsmöglichkeiten der KI in der Sozialen Arbeit. Den Chancen z.B. mittels KI dem Fachkräftemangel zu begegnen oder auch auf Grund KI-gestützter Analysen Kindesmissbrauch zu erkennen oder präventive Ansätze zu ermöglichen, stellt sie die Entstehungsbedingungen und Reproduktion gesellschaftlicher Strukturen bezogen auf die KI gegenüber. Eine inklusive Praxis sowie das Prinzip der Ko-Produktion sind daraus abgeleitete Forderungen für den sozialgerechten Umgang mit KI-Systemen.

Annalies Baumeister et al. untersuchen im BMBF-Projekt „DoF Adaptiv“ die Nutzer*innenakzeptanz aus verschiedenen Perspektiven für einen KI-gestützten Roboterarm, der für Menschen mit schweren körperlichen Behinderungen als Greifhilfe dienen kann. Neben der Untersuchung der Akzeptanzfaktoren wird der aktuelle Rechtsrahmen und das Risikomanagement miteinbezogen.

Barbara Lipka berichtet aus den Erfahrungen des BMAS geförderten Projektes „KI.ASSIST“ bei dem Lern- und Experimentierräume das Erproben von Arbeitsinnovationen in der beruflichen Rehabilitation in einem geschützten Rahmen ermöglichen, ohne gleich in den Realbetrieb zu gehen.

Leya Küsters et al. setzen sich im Forschungsprojekt „TeilhabeAssistenz – Digitale Lösungen für betreute Wohnformen“ (Mittel des Landes Hessen, Distr@l-Förderlinie 2: Digitale Innovationsprojekte) mit verschiedenen Möglichkeiten KI-gestützter Anwendungen für einen Telepräsenzroboter für Menschen mit psychischen Erkrankungen systemisch auseinander und thematisieren neben erprobten, weitere potenzielle Anwendungen sowie erforderliche Rahmenbedingungen.

In Healthcare Bereich Altenhilfe zeigt Miriam Peters vier spannende Einsatzmöglichkeiten der KI für die ambulante Pflege auf, die zum Teil auch schon über die Erprobung hinausgehen.

Lisa Luft hat für die stationäre Langzeitversorgung Übersichtsarbeiten analysiert und stellt einen Kanon von Anforderungen auf, die für eine nutzungsfreundliche Entwicklung eine Rolle spielen können.

Eva Lermer et al. setzen sich mit der Bewertung von KI-gestützten medizinischen Online-Diensten auseinander, die häufig überraschend positiv ausfällt – gerade vor dem Hintergrund, dass von einer gewissen Technikresistenz ausgegangen wird.

Anne-Kathrin Kleine geht der Frage nach, welche KI-gestützten klinischen Entscheidungsunterstützungssysteme es in der mentalen Gesundheitsversorgung gibt und hat dazu Patent- und Onlinerecherchen durchgeführt.

Julia Cecil et al. unterscheidet in ihrem Beitrag patient*innenorientierte und therapeut*innenorientierte Tools und analysiert deren Nutzungsbedingungen auf der Basis einer Online-Studie.

Christiane Saure zeigt entlang der Wertschöpfungskette im Krankenhaus, welche Potenziale in der generativen KI liegen kann und berücksichtigt dabei die Primär- und patientenbezogenen sowie nicht medizinischen Sekundärprozesse.

Aus der Perspektive der Kostenträger betrachtet Sibel Altin die Einsatzmöglichkeiten der KI bei den Krankenversicherungen. Die Potenziale werden in der Verwaltung, der Leistungssteuerung sowie der Kundenkommunikation gesehen und könnte u.a. eine individuell angepasste – auch präventive Gesundheitsversorgung ermöglichen.

Im Themenbereich „Die juristische und ethische Perspektive“ diskutiert Olivia Alig KI im Kontext des Konfliktmanagements.

Susanne Beck blickt auf die rechtlichen Rahmenbedingungen und Herausforderungen des KI-Einsatzes in der Medizin. Neben einem Überblick über relevante rechtliche Aspekte geht sie auf das Konzept der „Meaningful Human Control“ ein, welches eine nicht unwesentliche Rolle bei der juristischen Verantwortung z.B. bei medizinischen Fehldiagnosen spielt.

Im Anschluss widmet sich Janina Loh den Fragen der ethischen Evaluation von KI am Beispiel Chat GPTs und stellt ein von ihr entwickeltes Vier-Dimensionen Modell dar, das in Ergänzung zu anderen Modellen der ethischen Evaluation von neuen Technologien genutzt und in Lehre und Forschung eingesetzt werden kann.

Die Beiträge verdeutlichen, dass zu den Erfolgsfaktoren für KI im Healthcare-Sektor neben den oben genannten Anforderungen weitere Faktoren gehören.

Zentral sind partizipative Ansätze – sowohl bei der Software- als auch bei der Anwendungsentwicklung – bei denen neben diversen Entwicklungsteams insbesondere auch die Anwender*innen und die relevanten Stakeholder miteinbezogen werden – z.B. in Form eines Ko-Produktionsprozesses. Für den Einsatz in der Arbeitswelt bieten sich Lern- und Experimentierräume an, um entsprechende Qualifikationen im Umgang mit der Technik, aber auch das Experimentieren mit den Veränderungen der Arbeitsaufgaben, -abläufen und -organisation zu ermöglichen, um den Roll-out reibungsärmer zu gestalten. Neben Datenschutz und Datensicherheit spielen ethische und regulatorische Aspekte eine zentrale Rolle. Fehlende Transparenz bei der KI-Entwicklung, der dem Healthcare-Sektor immanent zu Grunde liegende Zeitdruck und der Kostendruck sind nur einige Faktoren, die im Widerspruch zur Anforderung der letztendlichen Kontrolle des Menschen über KI stehen. Einerseits verlangen die zunehmende Komplexität und damit verbundenen Risiken eine (zu?) hohe Expertise im Umgang, die sich letztendlich mit den Produktionsbedingungen des Healthcare-Sektors – zu wenig Personal, zu viele Patient*innen, Klient*innen etc. – nicht vertragen. Andererseits zeigen die Beiträge in diesem Band, dass KI im Healthcare-Sektor voraussichtlich dazu beitragen kann, die Versorgung zu verbessern, neue Formen der Teilhabe und Inklusion zu ermöglichen, auch die Kosten zu senken und zur Effizienz des Sektors beizutragen. Entscheidend ist allerdings, dass diese Technologien verantwortungsbewusst unter Einbezug der weiblichen und Diversitäts-Perspektive sowie unter Berücksichtigung der ethischen und regulatorischen Rahmenbedingungen implementiert werden.

An dieser Stelle sei den engagierten Autorinnen und hessian.AI gedankt – ohne sie wäre diese Publikation nicht möglich gewesen.

Literaturangaben:

Gramm, F.: Keine Frau auf der Bühne – was es bedeutet, wenn Podiumsdiskussionen von Männern dominiert werden und was dagegen getan werden kann. In: Frauen in der digitalen Zukunft der Medizin und Gesundheitswirtschaft. Heidelberg 2021.

Kettemann, M. C.: UNESCO Empfehlung zur Ethik künstlicher Intelligenz. Bedingungen zur Implementierung in Deutschland. Bonn 2022. <https://www.unesco.de/wissen/wissenschaft/ethik-und-philosophie/studie-umsetzung-ki-ethik-empfehlung> [abgerufen am 28.12.2023]

Simonite, T.: AI Is the Future—But Where Are the Women? 2018. <https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/> [abgerufen am 28.12.2023]

Steffens, B.: Genderbetrachtung bei der Digitalisierung umfassend implementieren. In: Frauen in der digitalen Zukunft der Medizin und Gesundheitswirtschaft. Heidelberg 2021.

Prof. Dr. Barbara Klein

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Forschungszentrum FUTURE AGING
barbara.klein@fb4.fra-uas.de

Lebenslauf in Kürze

Prof. Dr. Barbara Klein forscht zur Entwicklung von nutzungsfreundlichen, nachhaltigen und sozio-technischen Lösungen für eine alternde Gesellschaft.

Nach ihrem Studium der Soziologie an den Universitäten Mainz, Frankfurt und London arbeitete Prof. Dr. Klein mehr als 20 Jahre bei der Fraunhofer-Gesellschaft, der größten Forschungsgesellschaft im Bereich der angewandten Forschung in Europa. Von 1994 bis 1995 forschte und lehrte sie im Rahmen eines Marie-Curie Stipendiums an der University of Stirling in Großbritannien. Nach ihrer Rückkehr zur Fraunhofer-Gesellschaft baute sie das Marktstrategie Team Public Health am Fraunhofer IAO in Stuttgart auf. Seit 2007 ist sie Professorin für Organisation und Management in der Sozialen Arbeit an der Frankfurt UAS. Seit 2018 ist sie Sprecherin des interdisziplinären Forschungszentrums FUTURE AGING und seit 2021 Dekanin des Fachbereichs Soziale Arbeit und Gesundheit. Sie ist Vice President 2024 der International Society for Gerontechnology und Präsidentin des Austrian-German-Swiss Chapters. Prof. Dr. Klein war Gastprofessorin an der Osaka-University in Osaka, Japan (2015, 2019-2021) sowie an der Northumbria-University in Newcastle, England (2013-2016). 2019 war sie Preisträgerin des Forschungspreises der Hessischen Hochschulen für Angewandte Wissenschaften. Sie ist Autorin und Mitautorin von mehr als 200 Publikationen.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist es zu innovativen Lösungen im Healthcare-Sektor für eine gute Versorgung und für gute Arbeit beizutragen und damit die Grenzen des Möglichen zu erweitern.



Susanne Rägler, Susanne Klüber

KI an der Frankfurt UAS – Bestandsaufnahme und Zukunftsperspektive

Die Frankfurt University of Applied Sciences (Frankfurt UAS) ist eine der größten Hochschulen für Angewandte Wissenschaften (HAW) in Deutschland. Durch die vier großen Fachbereiche und mit einer umfassenden Bandbreite an Disziplinen, von Architektur bis Soziale Arbeit, von Informatik bis BWL, ermöglichen die HAW-typische Praxisnähe und Anwendungsbezug sowie die konsequent interdisziplinäre Ausrichtung permanent neue Perspektiven für Forschung, Lehre und Transfer.

Wie für die meisten Hochschulen zutreffen dürfte, ist KI kein neues Feld für die Forschenden und Lehrenden der Frankfurt UAS. Insbesondere in der Informatik als Stammdisziplin der KI wurde und wird routinemäßig damit gearbeitet und daran geforscht. Dies gilt in besonderem Maße für die anwendungsbezogene Dimension der KI, aber auch darüber hinaus. Neu ist nun jedoch, dass wir KI an der Hochschule systematisch und über alle Fachbereiche hinweg in den Blick nehmen, verschiedene Perspektiven zusammenführen, eine Öffentlichkeit an der Hochschule über KI-Diskurse herstellen und damit den Erkenntnisgewinn in Sachen KI an der Frankfurt UAS gezielt fördern. Hierzu hat die Hochschule einen Prozess gestartet, an dessen Anfang eine systematische Erhebung der bereits vorhandenen KI-Kompetenzen stand. Ziel ist dabei auch, diese Kompetenzen mit festgestellten Bedarfen im Kompetenzaufbau bei Studierenden, professoralen Kolleg*innen und allen anderen Hochschulmitgliedern zusammenzubringen. Leitgedanke ist dabei immer: Wie können Akteur*innen, die bisher nicht zusammenarbeiten, zusammengebracht werden, um ihr Wissen auszutauschen, zu vergrößern und dadurch noch bessere (Anwendungs-)Lösungen für aktuelle, relevante Fragestellungen zu generieren?

KI an der Frankfurt UAS

Um die an der Hochschule bereits vorhandenen Ansätze zu strukturieren und erstmals sichtbar zu machen, wurde eine KI-Landkarte der Hochschule erstellt, auf der

die Expertisen in den Bereichen Human-AI Interaction, KI-Anwendungsgebiete, Ethische, soziale und rechtliche Implikationen, KI-Management, KI-Technologie und KI-Innovation geclustert sind (Abbildung 1).

In diesen Ansätzen werden beispielsweise Deep-Learning-Verfahren eingesetzt, um Steuerungstools in der Finanzindustrie schneller zu errechnen, die zur Einhaltung des 1,5-Grad-Ziels des Pariser Klimaabkommens beitragen. Ein weiteres Vorhaben, das sich KI-gestützt mit der Erreichung von Nachhaltigkeitszielen in Unternehmen auseinandersetzt, bereitet Daten aus Nachhaltigkeitsreports auf und stellt Transparenz und Vergleichbarkeit her. Technologische Anwendungen der KI umfassen beispielsweise Predictive-Maintenance-Verfahren sowie mittels KI entwickelte Prognosen zur Parksuchdauer, die zu einer verbesserten Verkehrsplanung beitragen. Daneben findet eine fundierte wissenschaftliche Auseinandersetzung mit möglichen Rechtsrahmen für KI und deren Anwendung statt. Komplettiert wird diese exemplarische Reihe durch die in diesem Band versammelten Ansätze zu KI im Healthcare-Sektor. Inzwischen sind an der Hochschule auch Rechenkapazitäten vorhanden, die es Forschenden ermöglichen, für ihre KI-Anwendungen und -Algorithmen Hardware vor Ort zu nutzen.

Auch im Bereich der Lehre ist die Hochschule seit der Nutzbarkeit von ChatGPT verstärkt für den Einsatz von KI sensibilisiert, beispielsweise mit Angeboten der Bibliothek zu KI im Studienalltag in Kooperation mit weiteren KI-versierten Teams an der Hochschule: Von der Literaturrecherche mit KI-Tools bis zur effektiven Nutzung von ChatGPT im Studium. So profitieren Studierende unmittelbar vom Transfer des an der Hochschule vorhandenen Wissens in die Lehre und in lehrbegleitende Veranstaltungen.

Um KI als Querschnittsaufgabe an der Hochschule noch stärker zu verankern, wird die Frankfurt UAS an jedem

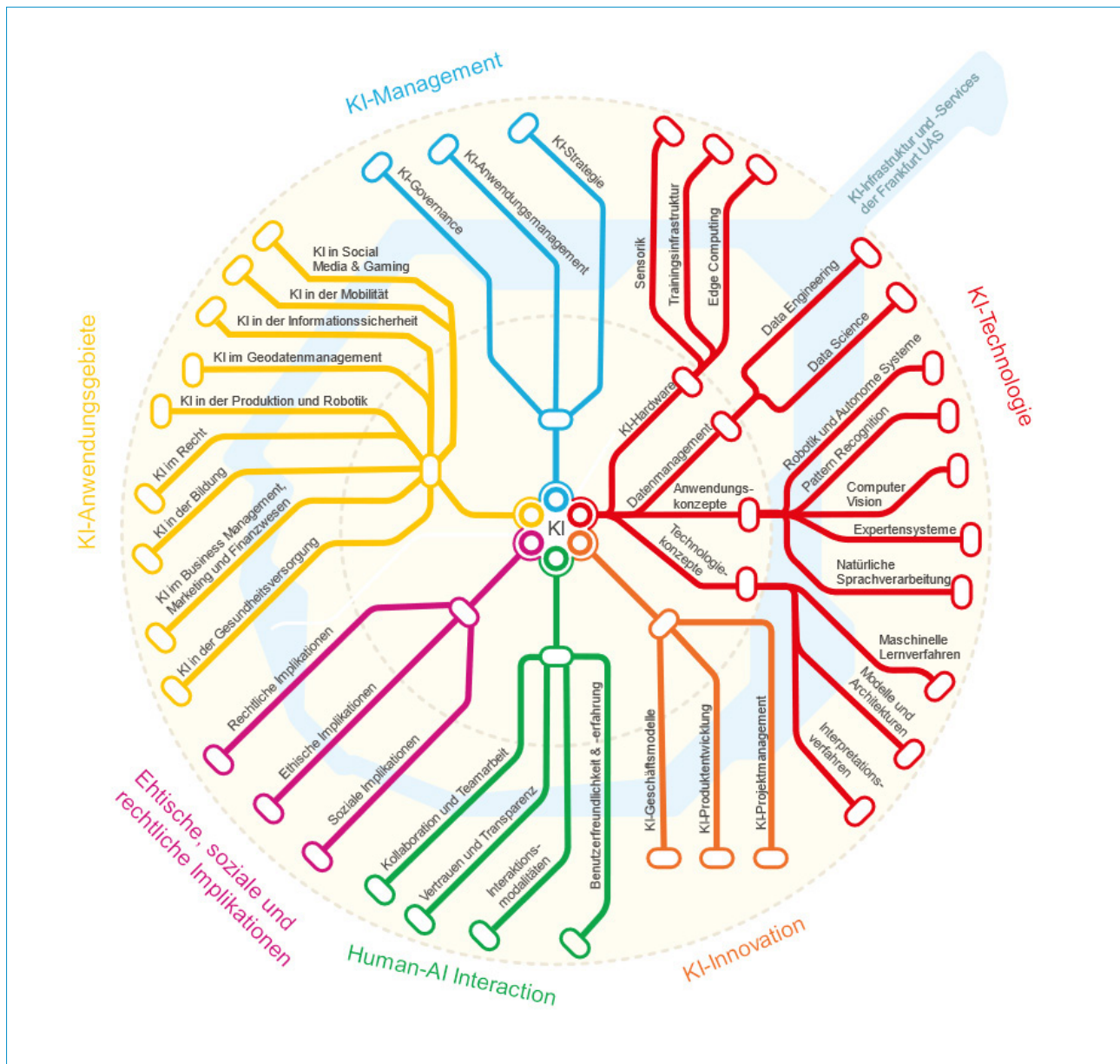


Abbildung 1: KI-Landkarte der Hochschule, auf der die Expertisen in den Bereichen Human-AI Interaction, KI-Anwendungsgebiete, Ethische, soziale und rechtliche Implikationen, KI-Management, KI-Technologie und KI-Innovation geclustert sind.

Fachbereich eine KI-Professur besetzen, die sich mit dem Thema nicht nur fachspezifisch, sondern als Quartett auch fachbereichsübergreifend auseinandersetzen. Alle diese Bestrebungen werden zukünftig in einem disziplin- und fachbereichsübergreifenden KI-Kompetenzzentrum gebündelt, in dessen Rahmen Forschung, Lehre und Transfer im Bereich KI zusammenlaufen. Das Zentrum bildet sodann den zentralen Knotenpunkt für alle KI-Aktivitäten an der Hochschule.

Die Frankfurt UAS hat die herausragende Rolle, die KI für viele wissenschaftliche und gesellschaftliche Bereiche in der Zukunft einnehmen wird, erkannt und strebt

danach, wissenschaftliche Erkenntnisse auf höchstem Niveau zu ermöglichen und gleichzeitig den Transfer in Lehre, Wirtschaft und Gesellschaft zu befördern. Wir nehmen für uns in Anspruch, Knotenpunkt in einem regionalen Netzwerk für KI zu sein.

Externe Vernetzung

Neben der internen Vernetzung ist bei dem gesellschaftlich relevanten Thema KI besonders auch die Einbindung in Netzwerke außerhalb der Hochschule im Fokus. An erster Stelle zu nennen ist hier hessian. AI, das in der hessischen Hochschullandschaft die bedeutendste Plattform, Impulsgeberin und Ermög-

licherin für KI ist. In diesem Zusammenschluss unter Beteiligung der drei hessischen Ministerien Hessisches Ministerium für Wissenschaft und Kunst, Hessische Ministerin für Digitale Strategie und Entwicklung sowie dem Hessischen Ministerium für Wirtschaft, Energie, Verkehr und Wohnen sind 13 hessische Hochschulen vertreten, um KI-Spitzenforschung in Hessen weiter voranzutreiben.¹ Aus hessian.AI-Mitteln wurde auch die erste der vier KI-Professuren am Fachbereich für Wirtschaft und Recht mit der Denomination *Wirtschaftsinformatik, insbesondere Künstliche Intelligenz und Entrepreneurship* eingerichtet. Die Mitgliedschaft der Frankfurt UAS bei AI FrankfurtRheinMain e.V. verfolgt ebenfalls das Ziel der Vernetzung in der wissenschaftlichen Community und insbesondere mit Wirtschaftsunternehmen, um so dem Transfergedanken konsequent Rechnung zu tragen.

Women in AI

Als öffentliche Bildungseinrichtung übernimmt die Frankfurt UAS gesellschaftliche Verantwortung. Dazu gehört auch, für Chancengerechtigkeit einzustehen. Im Bereich der KI-Forschung und -Anwendung sind Frauen heute unterrepräsentiert, was die Gefahr birgt, dass sich Gender Bias mitunter in Algorithmen einschreiben und die resultierenden Anwendungen die gesellschaftliche Benachteiligung von Frauen perpetuieren. Obwohl Frauen nicht per se technik-avers sind, stellt sich kein aus-

gewogenes Geschlechterverhältnis in der wissenschaftlichen Arbeit an und mit KI ein. Exemplarisch sei hierzu angeführt, dass auch im Jahr 2022 in Deutschland nur bei ungefähr einem Drittel der Publikationen im Bereich KI zumindest eine Wissenschaftlerin beteiligt war.²

Um der Unterrepräsentation von Frauen entgegenzuwirken, unterstützt die Frankfurt UAS in zahlreichen MINT-Initiativen Frauen und Mädchen durch Angebote zur Vernetzung, zum Erfahrungsaustausch und zur Stärkung von Diversität an der Frankfurt UAS.³ Algorithmen und Programme sind nicht objektiv und es ist von entscheidender Bedeutung, jetzt die Chance zu nutzen, geschlechterstereotype Wahrnehmungen nicht mit in die Zukunft unserer Gesellschaft zu nehmen. Das Bestreben der Frankfurt UAS ist es, disziplinenübergreifend Zugänge für Frauen und Mädchen zu schaffen, ausgehend davon, dass durch Zugänge auch Kompetenzen und durch Kompetenzen Verständnis und Gestaltungsmöglichkeiten folgen.

Die Frankfurt UAS begrüßt daher sehr die Initiative von Frau Prof. Dr. Barbara Klein, diese Publikation aufzusetzen, die die hochaktuellen Beiträge von Kolleginnen verschiedener Hochschulen in diesem Band mustergültig versammelt: interdisziplinär, von der technologischen Anwendung bis zur qualitativen Einordnung und quer durch die wissenschaftliche Community.

¹ Vgl. hessian.AI: Über uns. Online: <https://hessian.ai/de> [abgerufen am 13.12.2023].

² Vgl. Bundesministerium für Bildung und Forschung: Richtlinie zur Förderung von wissenschaftlichen Nachwuchsgruppen unter Leitung von Frauen im Bereich der Künstlichen Intelligenz, Bundesanzeiger vom 15.09.2023. Online: <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2023/09/2023-09-15-Bekanntmachung-Nachwuchsgruppen.html?view=renderNewsletterHtml> [abgerufen am 14.12.2023].

³ MINT-Koordinatorin, MINTvernetz, Mentoring Hessen, Vernetzungsmöglichkeiten für Studentinnen an unseren MINT-Fachbereichen 1 und 2. Vgl. auch Frankfurt UAS: MINT & Diversity. Online: <https://www.frankfurt-university.de/de/hochschule/einrichtungen-und-services/beauftragte-und-vertretungen/frauen-und-gleichstellungsbuero/mint-diversity/> [abgerufen am 15.12.2023].

Literaturangaben:

Bundesministerium für Bildung und Forschung: Richtlinie zur Förderung von wissenschaftlichen Nachwuchsgruppen unter Leitung von Frauen im Bereich der Künstlichen Intelligenz, Bundesanzeiger vom 15.09.2023. Online: <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2023/09/2023-09-15-Bekanntmachung-Nachwuchsgruppen.html?view=renderNewsletterHtml> [abgerufen am 14.12.2023].

Frankfurt UAS: MINT & Diversity. Online: <https://www.frankfurt-university.de/de/hochschule/einrichtungen-und-services/beauftragte-und-vertretungen/frauen-und-gleichstellungsbuero/mint-diversity/> [abgerufen am 15.12.2023].

hessian.AI: Über uns. Online: <https://hessian.ai/de> [abgerufen am 13.12.2023].

Prof. Dr. Susanne Rägler

Vizepräsidentin
Frankfurt University of Applied Sciences
vp-fwt@fra-uas.de

Lebenslauf in Kürze

Prof. Dr. Susanne Rägler ist seit 2004 Professorin für Allgemeine BWL, Steuern und Rechnungslegung und seit 2022 Vizepräsidentin für Forschung, Weiterbildung und Transfer der Frankfurt University of Applied Sciences. Nach ihrem Studium der Wirtschaftsinformatik schloss sie ihre Promotion im Jahr 1999 an der Universität Mannheim ab. Sie ist Steuerberaterin und arbeitete als solche bei Ernst & Young im Bereich Assurance/Wirtschaftsprüfung sowie in Industrieunternehmen, zuletzt als Leiterin der Steuerabteilung der MVV Energie AG.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

KI fasziniert mich bereits seit meinem Studium – meine Promotion habe ich über ein künstliches neuronales Netz zur Beurteilung des Fehlerrisikos in handelsrechtlichen Jahresabschlüssen geschrieben. Seither haben sich die Anwendungsmöglichkeiten von KI in den letzten Jahren vervielfältigt und die Nutzer*innenbasis signifikant verbreitert. Die gesellschaftliche Relevanz für die kommenden Jahre kann gar nicht hoch genug eingeschätzt werden. Deswegen möchte ich die Entwicklung auch in meiner Funktion als Vizepräsidentin für Forschung, Weiterbildung und Transfer an einer der großen hessischen Hochschulen für Angewandte Wissenschaften zum Wohle aller mitgestalten.

Susanne Klüber

Frankfurt University of Applied Sciences
susanne.klueber@hsl.fra-uas.de

Lebenslauf in Kürze

Susanne Klüber ist Persönliche Referentin der Vizepräsidentin für Forschung, Weiterbildung und Transfer der Frankfurt University of Applied Sciences. Nach ihrem Studium an der Goethe-Universität Frankfurt war sie im Rahmen des Lektorenprogramms der Robert Bosch Stiftung Lektorin an einer chinesischen Universität und hat anschließend ein Management-Trainee-Programm der Agentur für Arbeit absolviert. Vor ihrer Tätigkeit an der Frankfurt University of Applied Sciences hat sie ein BMBF-Projekt zum Deutscherwerb für erwachsene Geflüchtete bei der Stadt Frankfurt koordiniert.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Als Filmwissenschaftlerin denke ich bei KI zuerst an fantastische Zukunftsszenarien – von 2001 – *A Space Odyssey* über *Blade Runner* bis *Her*. Ich entscheide mich dafür, die zunehmende Relevanz von KI in allen Lebensbereichen als Chance für ihre Anwendungsbereiche von der medizinischen Versorgung bis zur eigenen Arbeitswelt zu begreifen und bin gespannt auf den Wandel, den wir in den kommenden Jahrzehnten erleben werden.

2. Die Technische und Produktperspektive

Ute Bauer-Wersing

Lernen, Sehen und Orientieren in autonomen Assistenzsystemen

Der Artikel gibt einen Überblick über die unterschiedlichen Herausforderungen, die bei der Entwicklung von autonomen Assistenzsystemen mittels KI-Techniken zu lösen sind. Anhand von zwei typischen Forschungsfeldern der KI werden Lösungsansätze beispielhaft umrissen, die in den vergangenen Jahren in der Forschungsgruppe „Machine Learning for Intelligent Systems“ erarbeitet wurden.

Autonome Assistenzsysteme werden in vielerlei Hinsicht eine große Rolle bei der Bewältigung des demographischen Wandels und den damit verbundenen Herausforderungen spielen. Die alternde Gesellschaft und der bereits existierende Fachkräftemangel, der sich in den nächsten Jahrzehnten noch verstärken wird, erfordern technische Lösungen, die in unterschiedlichsten Umgebungen robust und adaptiv einsetzbar sind. Aktuelle Forschung und Entwicklung im Bereich Robotik und künstlicher Intelligenz ermöglichen hier mittlerweile Innovationen, die sich auf viele Lebensbereiche positiv auswirken. Hier soll ein kurzer Einblick in zwei grundlegende Forschungsfelder im Bereich KI-basierte autonome Assistenzsysteme gegeben werden, die maßgeblich zum praktischen Einsatz dieser Technologien in unterschiedlichen Anwendungsfeldern beitragen, z.B. bei Unterstützung im Alltag, insbesondere von Personen mit körperlichen Einschränkungen oder bei der Effizienzsteigerung, um menschliche Arbeitskräfte zu entlasten oder um die Sicherheit z. B. im Straßenverkehr zu erhöhen.

Damit autonome Systeme grundsätzlich in der Lage sind, diese Anforderungen zu erfüllen, müssen sie lernen sich in unbekanntem Umgebungen zu orientieren und zielgerichtet zu navigieren. Eine weitere Fähigkeit, die in den meisten Anwendungsszenarien unabdingbar ist, ist die Erkennung und Unterscheidung von verschiedenen Objekten, wobei hier der visuellen Objekterkennung eine wichtige Rolle zukommt. Zu bei-

den Aspekten, i.e. Sehen und Orientieren, wurde und wird aktiv geforscht und es stehen heute eine Reihe von unterschiedlichen KI-Methoden zur Verfügung, die sich als tragfähig herauskristallisiert haben.

Neben konventionellen KI-Methoden sind insbesondere auch sogenannte bio-inspirierte KI-Algorithmen für die Entwicklung autonomer Assistenzsysteme interessant. Die besondere Attraktivität dieses Ansatzes besteht in der Effizienz, Adaptierbarkeit und Robustheit, wie man sie auch in biologischen Systemen beobachtet. Der Versuch diese Prinzipien in KI-Systeme zu übernehmen zielt also darauf ab, genau diese Eigenschaften auch auf die Lösung komplexer technischer Probleme zu übertragen und zu erhalten. Im Folgenden werden anhand der bereits erwähnten Herausforderung KI-Methoden vorgestellt die durch biologische oder wahrnehmungspsychologische Phänomene motiviert sind.

Lokalisation und Navigation mit der Slow Feature Analysis

Die neuronalen Grundlagen, die uns eine Orientierung im Raum ermöglichen, zählen mit zu den am besten untersuchten kognitiven Leistungen des Gehirns. Die Entdeckung sogenannter Place Cells im Hippocampus von Nagetieren war der Ausgangspunkt für das Verständnis wie das Gehirn navigiert und räumliche Karten der Umgebung erstellt. Die Erforschung des als „inner GPS“ bezeichnete Systems aus verschiedenen für die räumliche Wahrnehmung spezialisierten Zellen wurde 2014 mit dem Nobelpreis für Medizin und Physiologie ausgezeichnet.

Es gibt eine Reihe mathematischer Modelle, die die Antworteigenschaften der beteiligten Neuronen nachbilden, u.a. die Slow Feature Analysis, kurz SFA, mit der sich das Antwortverhalten der Place Cells simulieren lässt. In Metka et al. (2018) konnte die SFA erst-

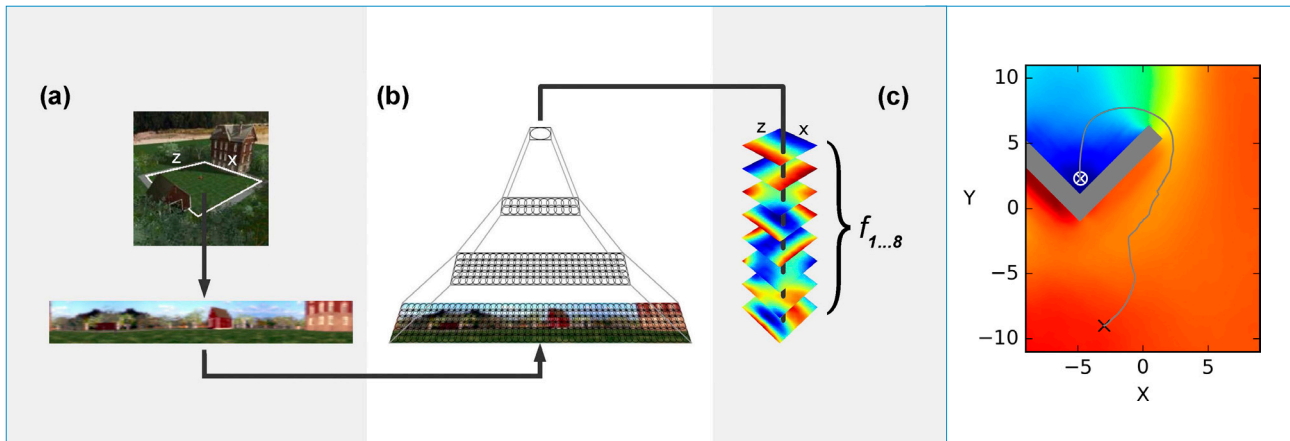


Abbildung 1, SFA-Netzwerk: In der Trainingsphase lernt das hierarchische Netz (b) aus den Bilddaten die der Roboter von der Umgebung aufnimmt (a) die Kodierung der Position in Form sogenannter Slow Features (c). Der rechte Teil der Abbildung zeigt beispielhaft die Navigations- und Positionstrajektorie des Roboters im SFA Raum. Von seiner Startposition, markiert durch ein Kreuz, erreicht der Roboter die Zielposition hinter dem Hindernis.

mals erfolgreich zum Lernen einer räumlichen Karte und zur Selbstlokalisierung eines mobilen Roboters eingesetzt werden. Das mathematische Modell lernt dabei in einer unüberwachten Trainingsphase aus Kamerabildern der Umgebung auf seine Position im Raum zurückzuschließen. Das Vorgehen ist überblicksartig in Abb.1 dargestellt. Dabei wird ein hierarchisches Netzwerk bestehend aus SFA-Knoten trainiert, so dass jede Position im Raum durch eindeutige Koordinaten in einem nieder-dimensionalen SFA-Raum kodiert ist.

Die mit dem SFA-Modell erreichte Lokalisationsgenauigkeit ist dabei vergleichbar mit anderen State-of-the-Art Algorithmen zum visuellen Self Localization and Mapping (SLAM). Darüber hinaus ist der Lernvorgang bei der SFA um zwei Größenordnung schneller als bei aktuelle End-zu-End Verfahren zur Posenschätzung. Ein weiterer Vorteil der SFA Kodierung zeigt sich im Hinblick auf eine selbstständige Navigation des mobilen Roboters in unbekanntem Umgebungen. In Metka et al. (2017) konnte gezeigt werden, dass sich der Roboter nur anhand eines Gradienten im SFA-Raum zu einem beliebigen Zielpunkt auch um vorhandene Hindernisse bewegen kann.

Wie im rechten Teil der Abb. 1 veranschaulicht, wählt der Roboter dabei eine effiziente Trajektorie vom Start zum vorgegebenen Zielpunkt. Um die Robustheit der Lokalisation zu erhöhen, wurde kürzlich in Haris et al. (2022) eine objekt-basierte Erweiterung des SFA-Ansatz vorgestellt. Ähnlich wie bei den in den letzten Jahren

entwickelten semantic SLAM Verfahren werden auf der Basis vortrainierter Objektdetektoren potentielle langzeitstabile Bereiche in der Umgebung ausgewählt. Dem Anwender wird so die Möglichkeit gegeben, basierend auf den erkannten Objekten, kooperativ mit dem System solche Bereiche im Bild auszuwählen, die für eine stabile Lokalisation über lange Zeitspannen hinweg geeignet sind. Das SFA-Netzwerk wird ausschließlich mit diesen Bildausschnitten trainiert. Somit erhält der Benutzer eine bessere Kontrolle und kann die Robustheit des Systems selbst kalibrieren.

Das Forschungsziel ein schnelles, effizientes und robustes räumliches Kartierungs-, Lokalisations- und Navigationsverfahren zu entwickeln, das auf kostengünstiger Hardware, wie sie im Endkundenmarkt üblich ist, eingesetzt werden kann, konnte damit erreicht werden. Das Verfahren eignet sich insbesondere für kleinere und mittelgroße Areale, die für den Einsatz von Assistenzrobotern im häuslichen Umfeld z.B. in der Wohnung aber auch im Außenbereich z.B. in Gärten typisch sind.

Erkennung verdeckter Objekte in komplexen visuellen Umgebungen

Ähnlich wie Menschen agieren Assistenzsysteme in komplexen Umgebungen, die durch eine Vielzahl von verschiedenen Objekten definiert sind, z.B. in häuslichen Umgebungen typische Haushaltgegenstände oder im Straßenverkehr typische Verkehrsteilnehmer. Will man diese Objekte automatisch erkennen und klassi-

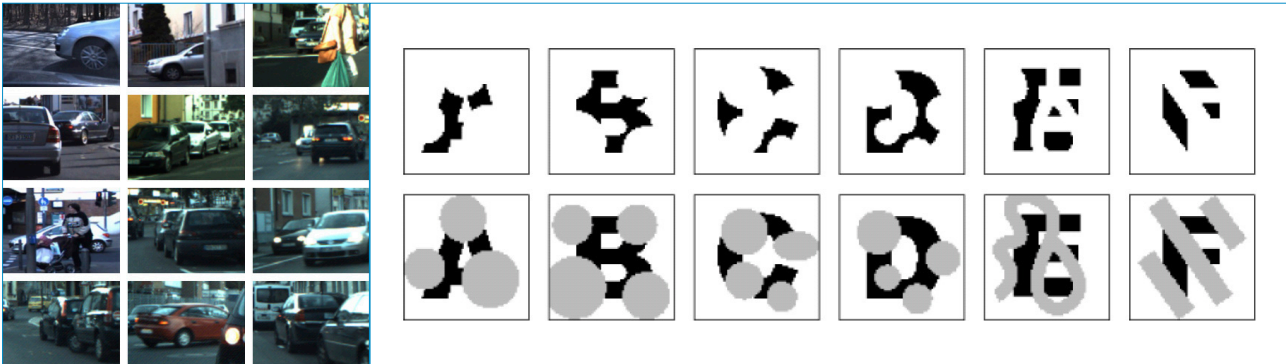


Abbildung 2, Erkennung verdeckter Objekte: Links eine Kollage mit Ausschnitten typischer Straßenszenen, in denen die Mehrzahl der Objekte von Verdeckungen betroffen sind, darunter auch mittel und hochgradig verdeckte Verkehrsteilnehmer. Rechts beispielhafte Illustration des wahrnehmungspsychologischen Effekts. In der oberen Bildreihe fällt es ersichtlich schwer, das verdeckte Objekt zu erkennen. Durch die Ergänzung und Erkennung der verdeckenden Strukturen werden in der unteren Sequenz die hochgradig verdeckten Objekte auf einmal erkennbar.

fizieren stellt man fest, dass ein Großteil der Objekte teilweise durch andere Objekte verdeckt werden. Die Untersuchung eines repräsentativen Datensatzes mit Aufnahmen von Verkehrsszenen hat ergeben, dass etwa 60 % der Pkws durch andere Verkehrsteilnehmer teilweise verdeckt werden. Gerade im Straßenverkehr ist die Erkennung von teilweise verdeckten Objekten aber kritisch. Die Einschätzung einer möglichen Gefahrensituation durch ein Fahrerassistenzsystem hängt stark von der visuellen Analyse der Szene ab. In Situationen wie sie z.B. durch die Erkennung eines teilweise verdeckten Kindes, eines schlecht sichtbaren Fahrradfahrers am Fahrbahnrand oder einer Reihe hintereinander geparkter Autos entstehen, können jederzeit zu einer potentiellen Gefahr bzw. Ausweichsituation für das Assistenzsystem führen. Im linken Teil der Abb. 2 finden sich typische Ausschnitte aus einem Trainingsdatensatz.

KI-basierte Ansätze zur Objekterkennung liefern für nicht bzw. schwach verdeckte Objekte mit einem Verdeckungsgrad bis ca. 40% meist brauchbare Erkennungsergebnisse, allerdings verschlechtert sich die Erkennungsleistung für mittel bis hochgradig verdeckte Objekte. Um die Erkennungsrate für diese Objekte zu steigern, haben wir bei der Modellierung auf Erkenntnisse der Wahrnehmungspsychologie zurückgegriffen. Abb. 2 beschreibt anschaulich das Phänomen, dass verdeckte Objekte einfacher zu erkennen sind, wenn man das verdeckende Objekt bereits erkannt hat.

Dieser Zusammenhang lässt sich beim Design von KI-basierten Ansätzen berücksichtigen z.B. in Form zwei-

er Klassifikatoren für unverdeckte und teilweise verdeckte Objekte respektive, die jeweils auf unterschiedlichen Trainingsdatensätzen trainiert werden. Die an die Wahrnehmungspsychologie angelehnte Idee des Algorithmus basiert in einem ersten Schritt darauf, zunächst alle unverdeckte Objekte in der Szene zu identifizieren und anschließend in der Umgebung der erkannten Objekte den spezialisierten Klassifikator anzuwenden.

Während für unverdeckte Objekte eine Reihe von geeigneten Datensätzen zur Verfügung steht, stellt die Bereitstellung von Trainingsdaten für verdeckte Objekte eine große Herausforderung dar. Daher wurde in Struwe et al. (2017) ein synthetischer Datensatz, der auf photo-realistisch gerenderten Szenen basiert, entwickelt und zum Trainieren des auf Verdeckung spezialisierten Klassifikators vorgeschlagen. Dieser Ansatz bietet im Vergleich zu realen Bilddatensätzen diverse Vorteile. Zum einen können verschiedene Verdeckungsszenarien einfach und ungebaised berücksichtigt werden und zum anderen entfällt das aufwendige Labeln der Daten per Hand, da Objektidentität, Position und Verdeckungsgrad (ground truth) bekannt sind.

Die beiden hier umrissenen Forschungsprojekte veranschaulichen Herausforderungen und beispielhafte Lösungsansätze, die bei der Entwicklung autonomer Assistenzsysteme eine Rolle spielen. Für ein System, das effektiv, effizient, robust und adaptiv ist, bieten sich bio-inspirierte Modelle an, da sie auf natürliche Weise diese wichtigen Eigenschaften mitbringen. Neben den angesprochenen Fragestellungen „Sehen“ und

„Orientieren“ erfordert die Entwicklung eines Assistenzsystems, das in variablen Umgebungen einsetzbar ist, je nach Anwendungsdomäne noch eine Reihe weiterer KI-basierter Methoden.

Aus technischer Sicht sind die aktuellen Fortschritte im Bereich Maschinelles Lernen und intelligente

Systeme rasant und es ist absehbar, dass sich unser Alltag im beruflichen und privaten Bereich mehr und mehr verändern wird. Dies wird dazu führen, dass KI-basierte Lösungen und autonome Assistenzsysteme insbesondere auch im Gesundheitswesen bei der Bewältigung der anstehenden Herausforderungen eine entscheidende Rolle zukommen wird.

Literaturangaben:

- Haris, M.; Franzius, M.; Bauer-Wersing, U.. Physical Interactive Localization Learning. International IEEE Conference on Advanced Robotics and Its Social Impacts. ARSO 2022, S. 1-6. Online: <https://doi.org/10.1109/ARSO54254.2022.9802969>
- Haris, M.; Franzius, M.; and Bauer-Wersing, U.. Robust Outdoor Self-localization In Changing Environments. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Macau, China 2019, S. 714-719. Online: <https://doi:10.1109/IROS40897.2019.8967549>.
- Metka, B.; Franzius, M.; Bauer-Wersing, U.. Bio-inspired visual self-localization in real world scenarios using Slow Feature Analysis. PLoS one, 13(9), e0203994. 2018. Online: <https://doi.org/10.1371/journal.pone.0203994>.
- Metka, B.; Franzius, M.; Bauer-Wersing, U.. Efficient navigation using slow feature gradients. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Vancouver, BC, Canada 2017, S. 1311-1316. Online: <https://doi.org/10.1109/IROS.2017.8202307>.
- Struwe, M.; Hasler, S.; Bauer-Wersing, U.. Rendered Benchmark Data Set for Evaluation of Occlusion-Handling Strategies of a Parts-Based Car Detector. In: Bräunl, T., McCane, B., Rivera, M., Yu, X. (eds) Image and Video Technology. PSIVT 2015. Lecture Notes in Computer Science(), vol 9431. Springer 2016, Cham. Online: https://doi.org/10.1007/978-3-319-29451-3_9.

Prof. Dr. Ute Bauer-Wersing

Frankfurt University of Applied Sciences
Fachbereich 2: Informatik und Ingenieurwissenschaften
ubauer@fb2.fra-uas.de

Lebenslauf in Kürze

Dr. Ute Bauer-Wersing vertritt seit 2007 als Professorin das Fachgebiet Datengetriebene Modellierung und Maschinelles Lernen am Fachbereich Informatik und Ingenieurwissenschaften der Frankfurt UAS. Sie lehrt in den Bachelor- und Masterstudiengängen des Clusters Informatik sowie im interdisziplinären Masterstudiengang Inclusive Design vornehmlich Inhalte der Programmierung, Mathematik und Künstlichen Intelligenz. In 2012 gründete sie die Forschungsgruppe Machine Learning for Intelligent Systems. Seither wurden unter ihrer Leitung eine Reihe von Drittmittelprojekten eingeworben und durchgeführt und bereits drei Promotionen erfolgreich abgeschlossen.



Als Gründungsmitglied ist sie seit 2018 im hochschulübergreifenden Promotionszentrum Angewandte Informatik aktiv und seither auch Mitglied im Promotionsausschuss des PZAI. Sie ist Autorin von mehr als 30 internationalen wissenschaftlichen Publikationen, war Mitglied im Programmkomitee mehrerer internationaler Konferenzen im Bereich KI und Robotik und ist an vier Patenten beteiligt. Vor Ihrem Ruf an die Frankfurt UAS besaß sie eine Professur an der Hochschule Trier für das Fachgebiet Mathematik und Informatik mit Schwerpunkt Data Mining in der Bioinformatik. Ergänzend zu Ihrer akademischen Laufbahn war Prof. Ute Bauer-Wersing mehrere Jahre leitend in anwendungs- und forschungsorientierten Positionen in der Industrie tätig. Darunter als IT Consultant und Projektleiter bei Daimler-Chrysler Interservices in München sowie als Gruppenleiterin Bioinformatik/Biomathematik bei den Biotechnologieunternehmen Aventis Research & Technologies GmbH & Co KG und Proteome Sciences Plc. in Frankfurt.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Bereits im Studium haben mich die Gebiete Maschinelles Lernen und künstliche Intelligenz fasziniert. Sowohl in meiner professionellen als auch akademischen Laufbahn konnte ich über viele Jahre in ganz unterschiedlichen Bereichen Methoden des maschinellen Lernens entwickeln und einsetzen z.B. im Gesundheits-Sektor und in der Entwicklung von autonomen Assistenzsystemen. Durch seine Vielseitigkeit hat das Feld für mich nichts an Faszination eingebüßt und ist aktuell, auch aufgrund der öffentlichen Aufmerksamkeit spannender denn je.

Paulami Banerjee, Simone Gramsch

Informed Machine Learning für die Produktion von Vliesstoffen in der Medizintechnik

Das Jahr 2020 markierte einen Wendepunkt sowohl für die Menschen als auch zahlreiche Unternehmen weltweit. Während viele Branchen aufgrund der COVID-19-Pandemie in Kurzarbeit gehen mussten, boomte die Produktion von medizinischen Masken. Um der Bevölkerung und dem medizinischen Personal ausreichend Schutzmasken zur Verfügung zu stellen, hat die Europäische Union allein im ersten Halbjahr 2020 Masken im Wert von 14 Milliarden Euro importiert, wovon 92 % aus China stammten. Im Vergleich dazu belief sich der Import im gleichen Zeitraum des Vorjahres auf nur 800 Millionen Euro. Trotz dieser enormen Steigerung um mehr als das 17,5-fache konnten anfänglich nicht genügend medizinische Masken bereitgestellt werden, um den Bedarf der Bevölkerung und vor allem des medizinischen Personals zu decken. Auch wurden durch die Corona-Pandemie Lieferketten unterbrochen, was sich insbesondere bei der Produktion von medizinischer Schutzausrüstung bemerkbar machte. Als Reaktion darauf initiierte das Bundeswirtschaftsministerium das Förderprogramm zur „Bundesförderung von Produktionsanlagen von Schutzausrüstung und dem Patientenschutz dienender Medizinprodukte sowie deren Vorprodukte“. Mit der Förderung sollte ein finanzieller Anreiz für mittelständische Unternehmen gesetzt werden, um Anlagen zur heimischen Produktion von FFP2/3 und medizinischen Masken aufzubauen. In der Folgezeit zeigte sich jedoch, dass der Aufbau deutscher Produktionskapazitäten herausfordernder als erwartet verlief, da der Herstellungsprozess für medizinische Schutzmasken äußerst komplex ist. Im Folgenden beschreiben wir den Herstellungsprozess für medizinische Schutzmasken und erläutern, wie mit Hilfe von Künstlicher Intelligenz und Simulationen der Produktionsprozess optimiert werden kann.

Herstellung von medizinischen Schutzmasken:

Spunbond- und Meltblownprozesse

Medizinische Schutzmasken wie FFP-Masken bestehen aus Vliesstoff und werden mit Hilfe von Gummibändern

über Mund und Nase gespannt. In Europa werden FFP-Masken in drei Klassen eingeteilt, wobei die FFP3-Maske die höchste Schutzwirkung aufweist und mindestens 99 % der Partikel bis zu einer Größe von maximal $0,6 \mu\text{m}$ filtern muss. Für die Filtration sorgen die inneren Lagen der Maske, die aus Meltblown-Vliesstoff bestehen. Die äußeren Lagen bestehen dahingegen aus einem sogenannten Spunbond-Vliesstoff.

Die Herstellungsprozesse von Spunbond- und Meltblown-Vliesstoffen ähneln sich und werden im Folgenden zusammen beschrieben. Beide Prozesse beginnen mit dem Schmelzen eines thermoplastischen Polymers wie beispielsweise Polypropylen. Die Polymerschmelze verhält sich von der Konsistenz ähnlich wie Honig und wird in einem zweiten Schritt durch eine Reihe von Spindüsen gepresst. Dabei entstehen Kunststoffäden ähnlich einem Honigfaden, der von einem Honiglöffel auf ein Brötchen tropft. Die Kunststoffäden kommen in einem dritten Schritt in eine Luftströmung. Die Temperatur der Luftströmung bestimmt den Prozess: höhere Temperaturen führen zum Meltblown- und niedrigere Temperaturen zum Spunbond-Prozess. Der Luftstrom zieht an den Kunststoffäden, macht sie dünner und kühlt sie gleichzeitig ab. Im vierten Schritt werden die Kunststoffäden verwirbelt und auf einem Transportband abgelegt. In einem nachgelagerten Prozessschritt werden die Kunststoffäden dann noch verfestigt, so dass aus den losen Fäden ein Vliesstoff entsteht. Je nach Anwendungszweck werden dann die Vliesstoffe weiterverarbeitet, beispielsweise zu FFP-Masken.

Optimierung ist eine Herausforderung

Auch wenn das Produktionsprinzip einfach anmutet, ist der optimale Betrieb eines Meltblown- oder Spunbondprozesses eine Herausforderung. Jeder einzelne Prozessschritt muss optimal auf die nachfolgenden Schritte ausgelegt werden. Insbesondere die Verwirbelungsphase der Kunststoffäden ist sehr stark durch die Turbulenz

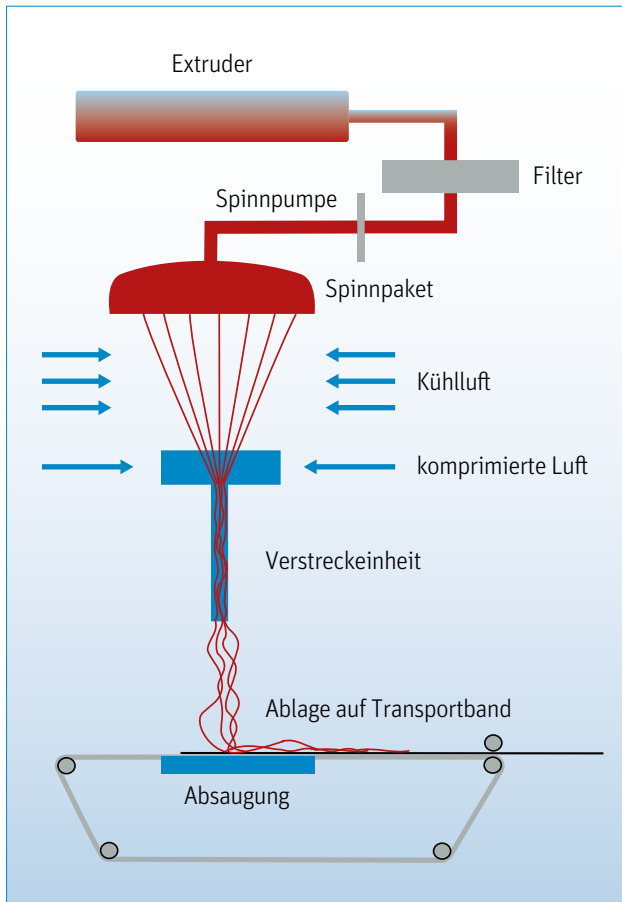


Abbildung 1: Schematische Darstellung von Meltblown- und Spunbond-Produktionsprozessen, Quelle: eigene Darstellung

der Luft geprägt – schließlich sorgt erst die Verwirbelung für die flächige Verteilung der Kunststoffäden. Aber wieviel Turbulenz sorgt für eine zufällige Verteilung der Fäden, so dass gleichzeitig die Fläche gleichmäßig bedeckt wird? Sollte an einigen wenigen, einzelnen Stellen weniger Fäden abgelegt werden, so leidet die Filtereffizienz an dieser Stelle. Diese FFP-Maske würde nicht ausreichend schützen, auch wenn die baugleichen Masken dieser Charge den Prüftest bestehen würden.

Eine Laboranlage hilft bei der Suche nach den optimalen Betriebseinstellungen wenig. Es ist bereits schwierig, aufgrund der sehr unterschiedlichen geometrischen Abmessungen die turbulente Luftströmung einer Laboranlage mit der Luftströmung einer Produktionsanlage zu vergleichen. Erschwerend hinzu kommt aber, dass die Kunststoffäden eine Rückwirkung auf die Luftströmung haben, da sie für die Luft einen Widerstand bilden. Damit ist die Wirkung der wenigen Kunststoffäden in der Laboranlage nicht vergleichbar mit der Wirkung der tausenden Kunststoffäden in der Produktionsanlage.

Eine Laboranlage kann daher nur sehr eingeschränkt zur optimalen Einstellung genutzt werden. Alle Versuche müssen an der Produktionsanlage durchgeführt werden, die in dieser Zeit für die eigentliche Produktion ausfällt.

Informed Machine Learning für die Auslegung der Produktionsprozesse von FFP-Masken

Da Versuche an Laboranlagen nur eine eingeschränkte Aussagekraft haben und Versuche an Produktionsanlagen zeitaufwändig und teuer sind, sind Forscherinnen und Forscher dazu übergegangen, die Produktion von Vliesstoffen mit Simulationen zu untersuchen. Seit über zwei Jahrzehnten entwickeln sie physikalische Modelle, basierend auf der sogenannten Cosserat-Rod-Theorie, zur Beschreibung der Bewegung der Kunststoffäden in Meltblown- und Spunbond-Anlagen (Marheineke & Wegener, 2006). Auch bei der Simulation der Kunststoffäden ist die Verwirbelungsphase eine große Herausforderung. Zwar gibt es physikalische Modelle, mit denen eine turbulente Luftströmung simuliert werden kann, wie beispielsweise die Reynolds-gemittelten Navier-Stokes-Gleichungen, doch werden bei solchen Simulationen die Turbulenzen der Luft als statistische Mittelwerte behandelt. Die spezifische Wirkung eines turbulenten Wirbels auf einen Kunststoffaden in einer industriellen Anlage wird nicht simuliert. Daher arbeiten Forscherinnen und Forscher in der Regel mit Ersatzmodellen für die Berechnung der turbulenten Luftkräfte auf die Kunststoffäden.

Diese Simulationen sind allerdings zeitaufwändig. Selbst die Simulation eines einzelnen Kunststoffadens kann eine ganze Woche in Anspruch nehmen. Bei Tausenden zu simulierenden Fäden in realen Produktionsprozessen stößt man schnell an die Grenzen der Rechenkapazität. Abhilfe sollen zukünftig Techniken des Informed Machine Learning schaffen. Bei Informed Machine Learning wird zusätzlich zum datenbasierten Training eines Machine-Learning-Modells vorhandenes a-priori-Wissen explizit in den Machine-Learning-Prozess integriert. Das explizite Wissen kann dabei auf vielfältige Weisen in den Prozess integriert werden. Die folgende Abbildung entnommen aus (Von Rueden et al., 2023) gibt eine Übersicht über die verschiedenen Wissensarten und ihre Repräsentation. Im Gegensatz zur ursprünglichen Abbildung sind hier die Integrationen von Wissen als Pfeil dargestellt, die für die Produktion von FFP-Masken relevant sind.

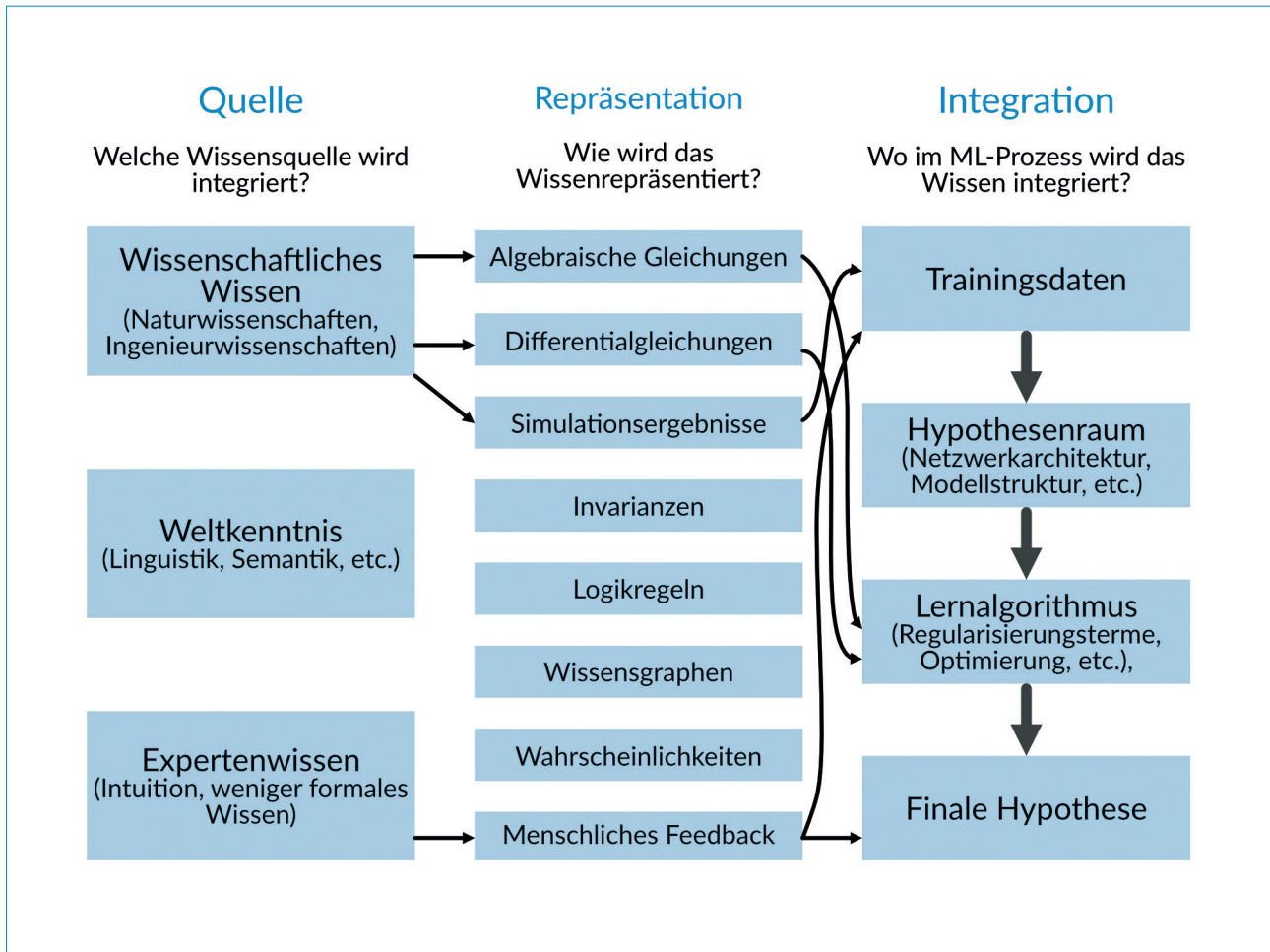


Abbildung 2: Konzept Informed Machine Learning; im Gegensatz zur ursprünglichen Abbildung sind hier nur Verbindungen gekennzeichnet, die bei der Optimierung von Spunbond- und Meltblown-Prozessen relevant sind; Quelle: In Anlehnung an Von Rueden et al. (2023)

Für die Analyse von Meltblown- und Spunbond-Prozessen ist die Ästhetik des Produkts, also des Vliesstoffes, sehr wichtig, die als menschliches Feedback in die Auswahl der Trainingsdaten und die finale Hypothese eingeht. Für die Integration des »Wissenschaftlichen Wissens« bieten sich mehrere Stellen an. Zum einen können Simulationsergebnisse genutzt werden, um die Datenbasis für das Training von Machine-Learning-Modellen zu erweitern. Zum anderen können physikalische Modelle in Form von algebraischen Gleichungen oder Differentialgleichungen auch direkt in die Machine-Learning-Modelle eingebaut werden, beispielsweise durch die Verwendung von Physics-Informed Neural Networks. Wie auch der Produktionsprozess selbst ist die Integration der Differentialgleichungen aufgrund ihrer stochastischen Natur eine große Herausforderung. Auch gibt es nicht das eine große physikalische Modell, das einen Meltblown- oder Spunbond-Prozess komplett beschreibt. Daher besteht der aktuelle Ansatz in der

Forschung darin, jeden Prozessschritt einzeln zu untersuchen, die experimentellen Daten mit Simulationen anzureichern und geeignete Modelle in die Machine Learning Pipeline zu integrieren. Beispielsweise entwickeln und visualisieren Victor et al. (2023) ein PINN-Modell (Physics-Informed Neural Network) für den zweiten Prozessschritt, bei dem die Kunststoffäden ausgesponnen werden. Banerjee & Padmanabha (2023) untersuchen wiederum, wie mit Hilfe von Machine Learning die optimalen numerischen Einstellungen einer Simulationssoftware gefunden werden können, um die Simulation der Luftströmung mit so wenig Rechenzeit wie möglich bei ausreichender Genauigkeit der Simulationsergebnisse durchführen zu können. Eine umfassende Simulationsstudie zum Aufbau von Trainingsdaten für die Bewegung und Ablage von Kunststoffäden in einer Spunbond-Anlage ist derzeit im Gange. Die beiden letztgenannten Forschungsaktivitäten werden durch das Mittelbauprogramm »Stärken stärken« der Frankfurt UAS gefördert.

Zusammenfassung und Ausblick

Die neue Technologie des Informed Machine Learning ist besonders für ingenieurwissenschaftliche Fragestellungen geeignet, bei denen die Verfügbarkeit von Experimenten begrenzt ist. Dadurch, dass nur wenige Trainingsdaten für herkömmliches Machine Learning zur Verfügung stehen, bietet die Integration physikalischer

Gesetzmäßigkeiten, wie Differentialgleichungen, in neuronale Netze einen signifikanten Vorteil. Ein prägnantes Beispiel für die Anwendung von Informed Machine Learning ist die Produktion von Meltblown- und Spunbond-Materialien für FFP-Masken. Es besteht die berechtigte Hoffnung, dass diese Technologie zukünftig zu effizienteren und optimierten Produktionsprozessen führen wird.

Literaturangaben:

- Marheineke, N.; Wegener, R.: Fiber dynamics in turbulent flows: General modeling framework, *SIAM Journal on Applied Mathematics*, 66(5), 1703 – 1726. 2006.
- Von Rueden, L.; Mayer, S.; Beckh, K.; Georgiev, B.; Giesselbach, S.; Heese, R.; Kirsch, B.; Pfrommer, J.; Pick, A.; Ramamurthy, R.; Walczak, M.; Garcke, J.; Bauckhage, C.; Schuecker, J.: Informed Machine Learning – A Taxonomy and Survey of Integrating Prior Knowledge into Learning Systems, *IEEE Transaction on Knowledge and Data Engineering*, 35(1). 2023.
- Victor, V. S.; Etmüller, M.; Schmeißer, A.; Leitte, H.; Gramsch, S.: VisForPINNs – Visualization for Understanding Physics Informed Neural Networks, Interaktiver Konferenzbeitrag *VisxAI*, 2023. Online: <https://observablehq.com/@visforpinns/visualization-for-understanding-pinns>.
- Präsentation Paulami Banerjee und Mohan Padmanabha: „Machine Learning-Optimized Approach for Parameter Selection in MESHFREE Simulations: Enhancing Accuracy and Efficiency“ im Rahmen der Konferenz „MathSEE Symposium on Applications of Mathematical Sciences“, Karlsruhe, Deutschland September 2023.

Paulami Banerjee

Frankfurt University of Applied Sciences
Fachbereich 2: Informatik und Ingenieurwissenschaften
paulami.banerjee@fb2.fra-uas.de

Lebenslauf in Kürze

Paulami schloss ihr Masterstudium der Informatik an der RPTU Kaiserslautern-Landau (ehemals TU Kaiserslautern) im Jahr 2022 mit dem Schwerpunkt Künstliche Intelligenz ab. Derzeit ist sie wissenschaftliche Mitarbeiterin am Fachbereich Informatik und Ingenieurwissenschaften der Frankfurt University of Applied Sciences, wo sie unter der Betreuung von Prof. Dr. Ute Bauer-Wersing und Prof. Dr. Simone Gramsch ihre Doktorarbeit über informierte maschinelle Lernansätze zur Optimierung von Vliesstoffproduktionsprozessen durchführt. Darüber hinaus ist sie in der Abteilung Transportvorgänge am Fraunhofer ITWM, Kaiserslautern, tätig.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegt darin, ihre Macht zu nutzen, um das Leben der Menschen zu verbessern und alltägliche Herausforderungen zu vereinfachen. Ich widme mich der Anwendung von KI, um Lösungen zu schaffen, die das Leben besser und zugänglicher für alle machen.

Prof. Dr. Simone Gramsch

Frankfurt University of Applied Sciences
Fachbereich 2: Informatik und Ingenieurwissenschaften
simone.gramsch@fb2.fra-uas.de

Lebenslauf in Kürze

Simone Gramsch hat Mathematik an der TU Kaiserslautern (heute RPTU Kaiserslautern-Landau) studiert und ebenfalls dort in Mathematik promoviert. Nach ihrer Promotion arbeitete sie als wissenschaftliche Mitarbeiterin am Fraunhofer-Institut für Techno- und Wirtschaftsmathematik in der Abteilung Transportvorgänge, wo sie u.a. als Gruppenleiterin sich mit der Modellierung, Simulation und Optimierung von Produktionsprozessen des textilen Maschinenbaus beschäftigte. Seit 2020 ist sie Professorin für Mathematik in den Ingenieurwissenschaften und forscht im Bereich Informed Machine Learning für Produktionsprozesse im Maschinenbau.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist es, Simulationen und maschinelles Lernen zu verbinden!

Diana Völz, Julia Schneider

KI als Chance im Gesundheitswesen: Von der Theorie zum maßgeschneiderten Produkt

Künstliche Intelligenz (KI) bietet im Gesundheitswesen vielfältige Anwendungsmöglichkeiten. Diagnose- und Entscheidungsunterstützungssysteme können medizinischem Personal helfen, präzisere Diagnosen zu stellen und individualisierte Behandlungspläne zu entwickeln. KI unterstützt heutzutage bereits bei der Analyse medizinischer Bilder, wie Röntgenaufnahmen und MRTs, da sie Anomalien schneller und genauer erkennt. In der Präventivmedizin ermöglicht die KI die Identifizierung von Risikofaktoren und die Früherkennung von Krankheiten, was die Gesundheitsvorsorge revolutioniert. Parallel dazu kann KI auch bei der Entwicklung von Medizintechnikprodukten unterstützend eingesetzt werden und dies v.a. um maßgeschneiderte Produkte zu entwickeln, mit dem Ziel, die Lebensqualität der Patientinnen und Patienten erheblich zu verbessern. Die Kombination von KI mit moderner Medizintechnik eröffnet damit viele Chancen für die Gesundheitsversorgung hin zu einer passgenaueren Behandlung, die die Bedürfnisse jedes einzelnen Patienten berücksichtigt. Miniaturisierte Geräte, intelligente Implantate und tragbare Gesundheitssensoren sind Beispiele für innovative Medizintechnikprodukte, die auf dem Vormarsch sind.

Methodisch steckt die KI in der Produktentwicklung noch in den Kinderschuhen, gewinnt aber zunehmend an Bedeutung, um Produkte schneller und zielgerichteter an die Bedürfnisse der Kunden anzupassen. Ob es sich dabei um individualisierte Maschinenteile handelt, die auf spezifische Anforderungen abzielen, oder um personalisierte medizinische Hilfsmittel, die auf die Pathologien des Patienten zugeschnitten sind (Liu, 2022), spielt hierbei keine Rolle. In Designwerkzeuge integrierte generative Algorithmen ermöglichen beispielsweise heute schon das an Belastungsmodelle angepasste evolutionäre Wachsen von Strukturen, um optimierte Lösungen zu finden. (Garland, 2021) Ein weiteres Potential in der Produktentwicklung für

die Anwendung von KI bietet die Überprüfung von Simulationsergebnissen auf Plausibilität. (Bickel, 2019) KI-Anwendungen können Abweichungen von erwarteten Ergebnissen erkennen und Simulationsmodelle entsprechend anpassen, um sicherzustellen, dass die Produkte den Anforderungen entsprechen. Zusätzlich können KI-Algorithmen für das Erkennen von Mustern verwendet werden, wie das Suchen von ähnlichen Bauteilen. (Kamil, 2019)

KI als vielversprechende Methodik kann ein wichtiges Instrument in der Entwicklung von orthopädischen Hilfsmitteln sein. Medizinische Einlagen beispielsweise werden u.a. angewendet, um Fußfehlstellungen zu korrigieren oder spezifische Bereiche des Fußes je nach Bedarf zu entlasten oder zu stützen. Die Wahl des geeigneten Einlagentyps hängt jedoch nicht nur von dem Krankheitsbild ab, sondern ist eng mit den individuellen Patientenbedürfnissen und Empfindungen verknüpft. Passform und Tragekomfort der medizinischen Einlage spielen eine entscheidende Rolle in Bezug auf die Akzeptanz des Erkrankten, die Einlage zu nutzen.

Mit der Digitalisierung der zurzeit häufig noch handwerklichen Prozesse beim Design und der Fertigung bietet sich die Chance, medizinische Einlagen präziser an die individuelle Fußgeometrie und Druckbelastung des Patienten anzupassen. In einer digitalen Prozesskette wird mittlerweile über 3D-Scans der Fußgeometrie sowie der Aufzeichnung eines Ganganalysebildes (die sog. Pedografie) die Belastung an der Fußsohle beim Gehen ermittelt, um darauf basierend die Einlage mit dem Designwerkzeug zu modellieren, die dann im Anschluss 3D-gedruckt wird (vgl. Abbildung 1). Das Pedogramm erfasst die berechnete Druckverteilung und ermittelt die Abrollkurve sowie die Position des Körperschwerpunkts beim Gehen. Anhand dieser Daten wird in der Orthopädietechnik analysiert, welche

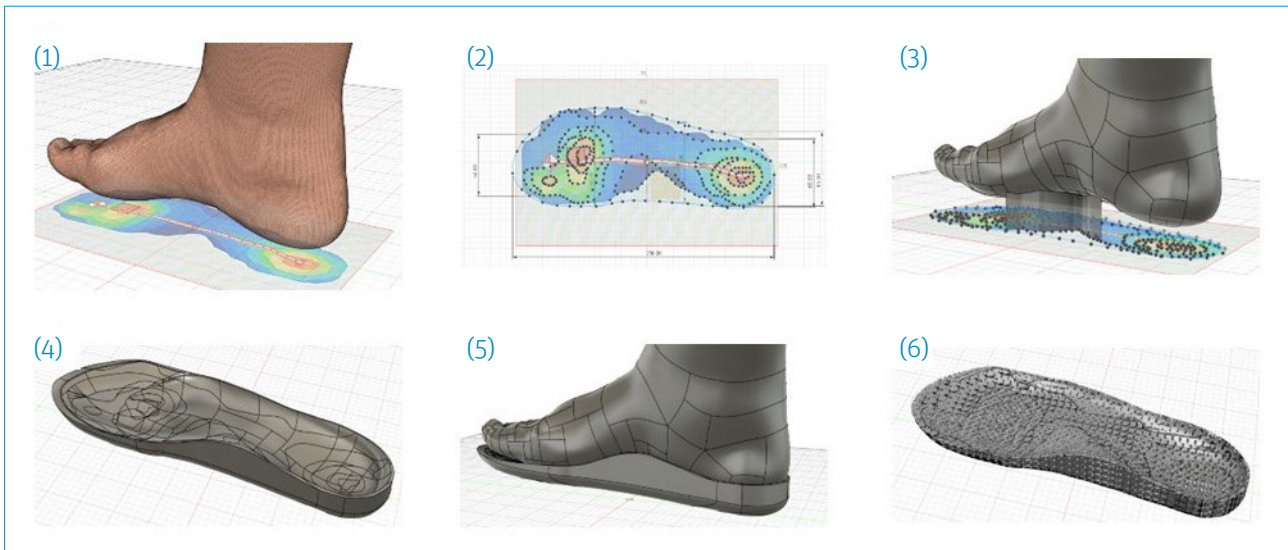


Abbildung 1: (1) überlagerte Daten (2) separierte Belastungsbereiche am Pedogramm (3) Sohlenerstellung durch 3D-Extrusion (4) 3D-Volumenkörper (5) Anpassungsprozess (6) Einlegesohle mit volumetrischen Gittern (Völz, 2023)

Fehlstellungen vorliegen, um die daraus resultierende skelettale Korrektur über die Einlegesohle vorzunehmen. Die Druckumverteilung am Fuß wird anschließend durch die Herstellung individueller Einlegesohlen vorgenommen.

Was nun die richtige Einlegesohle ist, hängt maßgeblich vom tradierten Wissen in der Orthopädiertechnik ab. Eine Analyse muss vielfältig sein, denn das Skelett besteht nicht nur biomechanisch gesehen aus einem System von Gelenken und Knochen, sondern auch Muskeln und Sehnen. Selbst die Struktur des Weichgewebes am Fuß hat, neben Krankheitssymptomen, wie z.B. eine vorliegende Nervenschädigung aufgrund von Diabetes, einen Einfluss auf das Gangbild. Grundsätzlich kann der Druck auf die Fußsohle nicht weggenommen, sondern nur auf andere Bereiche umverteilt werden. Die umfangreichen (Bild)-Daten können nun mit Methoden der KI vielversprechend für das Einlegesohlendesign genutzt werden.

Chancen für und mit maßgeschneiderten Einlegesohlen mittels KI

In den Ingenieurwissenschaften wird oft die KI-Methode des Maschinellen Lernens verwendet. Hier wird zwischen überwachtem, unüberwachtem und verstärkendem Lernen unterschieden. Beim überwachten Lernen ist klar, was richtig und was falsch ist. Das System lernt anhand von Beispielen. Diese werden basierend auf ihrem Inhalt klassifiziert und entsprechend gekenn-

zeichnet. Man spricht hier vom sog. Labeling. Diese Methode gilt deshalb als sehr sicher. In der Praxis werden zunächst klare Zielwerte, z.B. für gesund oder krank, angenommen, die Expertinnen und Experten auf Basis von Bilddaten beurteilen können. Pedogramme und skelettale Fehlstellungen können nun von Expertinnen und Experten eindeutig annotiert werden. Neue Bilder kann das System anschließend selbstständig zuordnen. Beim unüberwachten Lernen gibt es keine Label. Hier muss das System eigenständig Muster erkennen. Metriken von Ähnlichkeiten und Unterschiedlichkeiten dienen dazu, Daten Clustern zuzuweisen. Dabei ist das Ziel, Zusammenhänge zwischen Daten zu erkennen. Als Methode hierfür wird die Clusteranalyse, in der Clustering-Algorithmen Daten in Bezug auf ein Ergebnis kategorisiert, genutzt. Beide Verfahren, überwachtes und unüberwachtes Lernen, finden in der Orthopädiertechnik Anwendung. Ergebnisse können KI-unterstützte Korrekturen des Scans von Fußfehlstellungen oder die bildgebende Überführung von Pedogrammen von krank zu gesund sein, die dann im Designwerkzeug genutzt werden (vgl. Abbildung 2 und Abbildung 3).

Langfristig könnten mithilfe der KI anhand bildgebender Diagnostik und identifizierter Indikatoren auf Basis großer Datenmengen prädiktiv Erkrankungen erkannt und frühzeitig mit geeigneten orthopädischen Hilfsmitteln präventiv gegengesteuert werden. Dies kann u.U. invasive Eingriffe erheblich reduzieren. So könnte auch die Datenbasis durch Integration von spezifischen

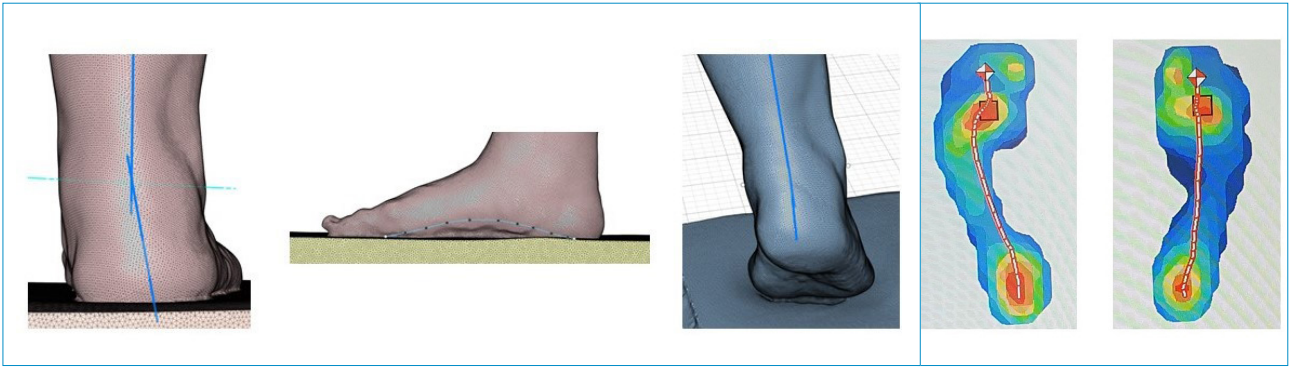


Abbildung 2: Diagnostik der Fehlstellung Knick-Senkfuß: links: axiale Beziehung zwischen unterem Bein und Fuß, Ansicht dorsal mittig: abgesenktes mediales Längsgewölbe eines rechten Fußes, Ansicht von medial rechts: Fersenhebetest, Ansicht dorsal. (Völz, 2023)

Abbildung 3: links: gesunder Fußabdruck, rechts Knick-Senkfuß

Anforderungen in den einzelnen Phasen des Gehens in den Anpassungsprozess integriert werden und wertvolle zusätzliche Informationen im Designprozess Anwendung finden (vgl. Abbildung 4). Die gewonnenen Datensätze werden miteinander kombiniert, um ein umfassendes Modell zu generieren.

lierung und dem Verhalten der Gitterstrukturen nach dem 3D-Druck nicht ausreichend vorhersagbar ist und die Varianz ihrer geometrischen Ausprägungen zu groß ist, wird hierfür das Verfahren der „Physics informed neural networks“ (PINNs) als KI Methode verwendet. Dieses erlaubt es, Modellwissen in den



Abbildung 4: Belastungen beim Gehen in den unterschiedlichen Gangphasen (Völz, 2023)

Eine Methode zum Design von Einlegesohlen stellt die Verwendung von Gitterstrukturbibliotheken dar. Basierend auf den individuellen Geometrieinformationen des Fußes wird ein volumetrisches Modell für die medizinische Einlage entwickelt. Dieses Modell ermöglichte die gezielte Segmentierung und Anpassung der Belastungsbereiche unter Verwendung von volumetrischen Gitterstrukturen mit variablen Parametern. Da die Verbindung zwischen der Model-

Trainingsprozess der Neuronalen Netzwerke zu integrieren, was bedeutet, dass neben Signalen, wie 3D-Drucktemperatur, Materialparameter und Gitterdesign, auch z.B. Materialgesetze Teil des Trainings der Netzwerke sind. Damit können der erforderliche Trainingsdatensatz minimiert und die Vorhersagegenauigkeit maximiert werden. Methodisch werden hier experimentelle Ergebnisse mit Simulationsmodellen verknüpft.



Abbildung 5: Strukturoptimierung von Stützstrukturen (eigene Darstellung)

Zusätzlich zur Perspektive des Erkrankten können KI Methoden beim Design von Produkten unterstützen. Generative Design Methoden, die 3D-Geometrien an Belastungsmodelle emergieren, sind bereits funktionaler Bestandteil von herkömmlichen Designwerkzeugen (vgl. Abbildung 5). Um die Akzeptanz der Produkte zu erhöhen, ist es sinnvoll, durch einen partizipativen Ansatz Patientinnen und Patienten besser in den Entwicklungsprozess einzubinden. Hier können Erfahrungsdaten, ausgewertet durch KI, unterstützen, den Anpassungsprozess zielgerichteter zu gestalten, indem Patienteneigungen anhand von Merkmalen erkannt und klassifiziert als Basis dienen. Zukünftig könnte auf Basis großer Datenmengen Prognosen zur personalisierten Anpassung und zum Verschleiß von orthopädischen Hilfsmitteln genutzt werden, um die Hilfsmittel rechtzeitig zu ersetzen, einen schlimmeren Krankheitsverlauf zu vermeiden und den Gesundheitssektor zu entlasten.

Die Ergebnisse zeigen Chancen für die Verwendung von KI zur Entwicklung von maßgeschneiderten Einlegesohlen auf. Diese Forschung liefert einen wichtigen Beitrag zur Weiterentwicklung der digitalen Prozesskette in der Orthopädietechnik und damit ein Mehrwert für das Gesundheitswesen. Trotz der vielversprechenden Potenziale gibt es jedoch auch Herausforderungen und Bedenken im Zusammenhang mit der Integration von KI in der Medizintechnik. Um statistische Unsicherheiten von KI-Algorithmen zu vermeiden, sind für das Training erhebliche Mengen an personalisierten Daten von sowohl gesunden Personen als auch Erkrankten notwendig. Datenschutz und Sicherheit sind wesentliche Anliegen, da sensible Gesundheitsdaten geschützt werden müssen. Die ethische Verantwortung bei der Nutzung von KI, insbesondere in der Entscheidungsunterstützung, erfordert klare Leitlinien und Überwachungsmechanismen. Außerdem müssen Ärzte und medizinisches Personal entsprechend geschult werden, um die KI-Systeme effektiv zu nutzen.

Literaturangaben:

Liu, C.; Tian, W.; Kan, C.: When AI meets additive manufacturing: Challenges and emerging opportunities for human-centered products development. In: *Journal of Manufacturing Systems*, Vol.64. Elsevier, 2022, S. 648–656. Online: <https://doi.org/10.1016/j.jmsy.2022.04.010>.

Garland, A.P.; White, B.C.; Jensen, S.C.; Boyce, B.L.: Pragmatic generative optimization of novel structural lattice metamaterials with machine learning. In: *Materials & Design*, Vol. 203, 109632. Elsevier, 2021. Online: <https://doi.org/10.1016/j.matdes.2021.109632>.

Bickel, S.; Spruegel, T.C.; Schleich, B.; Wartzack, S.: How Do Digital Engineering and Included AI Based Assistance Tools Change the Product Development Process and the Involved Engineers. In: *Proc. Int. Conf. Eng. Des.* 1 (1). Cambridge University Press, 2019, S. 2567–2576. Online: <https://doi.org/10.1017/dsi.2019.263>.

Židek K.; Lazorík, P.; Pite, J., Hošovský, A.: An Automated Training of Deep Learning Networks by 3D Virtual Models for Object Recognition, *Symmetry* 11, no. 4. MDPI, 2019, S. 496. Online: <https://doi.org/10.3390/sym11040496>.

Völz, D.; Schneider, J.; Wuttke, U.: Digital personalized medical insole process - A step forward to more personalized medical support,

HYPERLINK „<https://www.degruyter.com/journal/key/cdbme/html>“ *Current Directions in Biomedical Engineering*, 2023. Online: <https://doi.org/10.1515/cdbme-2023-1008>.

Prof. Dr.-Ing. Diana Völz

Frankfurt University of Applied Sciences
Fachbereich 2: Informatik und Ingenieurwissenschaften
voelz@fb2.fra-uas.de

Lebenslauf in Kürze

Prof. Dr.-Ing. Diana Völz ist seit 2015 als Professorin an der Frankfurt University of Applied Sciences im Fachbereich Informatik und Ingenieurwissenschaften tätig. Innerhalb der Lehrinheit Maschinenbau trägt sie die Verantwortung für die Bereiche Produktentwicklung, Konstruktion, CAD und Produktdatenmanagement. Ihr Studium absolvierte sie im Fach Maschinenbau an der Technischen Universität Darmstadt, wo sie auch ihre Promotion am Fachgebiet Datenverarbeitung in der Konstruktion erfolgreich abschloss. Vor ihrer Professur sammelte sie umfassende Berufserfahrung im Engineering bei der Firma Linde Material Handling in Aschaffenburg.



Die Forschungsschwerpunkte von Prof. Dr.-Ing. Diana Völz konzentrieren sich auf die Weiterentwicklung und Anwendung digitaler Werkzeuge in der Produktentwicklung, einschließlich Themen wie Digitale Zwillinge, Mixed Reality und Generatives Design. Die Erprobung dieser digitalen Werkzeuge findet in ihrem Mixed Reality Labor statt. Seit 2019 engagiert sie sich als Vorstandsmitglied im Forschungslabor „Personalized Biomedical Engineering“. In dieser Rolle verfolgt sie gemeinsam mit einer Gruppe von Forschenden das Ziel, die gesamte digitale Prozesskette von der Gestaltung bis zur Fertigung von Medizintechnikprodukten abzubilden und somit personalisierte Produkte anbieten zu können.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Die Anwendung von Künstlicher Intelligenz in der Produktentwicklung eröffnet für mich die Möglichkeit, nicht nur umfassende Analysen aus Erfahrungen mit Vorgängerprodukten in den Produktentwicklungsprozess zu integrieren, sondern auch proaktiv in die Nutzung einzugreifen. Auf diese Weise kann KI einen bedeutenden Beitrag leisten, indem die individuellen Bedürfnisse der Nutzerinnen und Nutzer in den Mittelpunkt der Produktentwicklung rücken.

Julia Schneider, M. Eng.

Frankfurt University of Applied Sciences
Fachbereich 2: Informatik und Ingenieurwissenschaften

Lebenslauf in Kürze

Julia Schneider absolvierte ihr Maschinenbaustudium an der Frankfurt University of Applied Sciences. Ihr Schwerpunkt lag hierbei in der Produktentwicklung. Seit April 2023 ist sie als wissenschaftliche Mitarbeiterin im Fachbereich Informatik und Ingenieurwissenschaften tätig. Im Rahmen dessen übernimmt sie Lehrtätigkeiten in unterschiedlichen Konstruktionsfächern innerhalb der Lehrinheit Maschinenbau. Berufserfahrung konnte sie bereits während ihres Studiums durch Lehraufträge sowie eine langjährige Tätigkeit als Werkstudentin in der Qualitätssicherung bei Rolls-Royce Deutschland Ltd. & Co. KG in Oberursel (Taunus) sammeln.

Julia Schneider promoviert in Kooperation mit der Universidad de Cádiz (Spanien) im Bereich der Entwicklung personalisierter Medizinprodukte. Dies erfolgt im Forschungslabor „Personalized Biomedical Engineering“ der Frankfurt University of Applied Sciences. Ziel ihrer Arbeit ist es, das Verhalten des Weichgewebes des menschlichen Fußes während des Gehens bei der Gestaltung orthopädischer Einlagen zu berücksichtigen. Eine additive Fertigung ermöglicht es dabei, die Einlage patienten- und pathologiespezifisch zu designen. So können je nach Bedarf bspw. einzelne Bereiche der Einlage stützend, andere entlastend wirken, ohne dass die Einlage aus unterschiedlichen Materialien zusammengesetzt werden muss.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Künstliche Intelligenz kann in vielen unterschiedlichen Bereichen eingesetzt werden. Vor allem in der Produktentwicklung eröffnet sie unglaubliche Möglichkeiten. Durch das Training mittels verhältnismäßig weniger Daten ist es möglich, Vorhersagen über das Verhalten neuer Produkte zu tätigen. Die Nutzung von KI kann daher dazu beitragen, neue Entwicklungen besser an verschiedene Anforderungen anzupassen und bereits vor der Fertigung von Prototypen zu optimieren. Auf diese Weise können Kosten sowie die Erzeugung unnötigen Abfalls reduziert werden.

Ekaterina Jussupow

Kognitive Prozesse in der KI-unterstützten diagnostischen Entscheidungsfindung

KI-Anwendungen unterstützen Ärzt:innen vermehrt bei der medizinischen Diagnostik. Ein Vorreiter ist die radiologische Bildgebung, in der Algorithmen Bilddaten, die durch Röntgen, Magnetresonanztomographie (MRT) oder Computertomographie (CT)-Verfahren erstellt werden, analysieren. Dabei werden verschiedene Verfahren des maschinellen Lernens eingesetzt. Hier lernen Algorithmen mit Hilfe von Trainingsdaten, die oftmals durch medizinische Experten vorbereitet werden, eine oder mehrere bestimmten Erkrankungen zu erkennen. Die Algorithmen können dann bestimmen, ob in einem bestimmten Bild eine bestimmte Erkrankung (wie z.B. ein Schlaganfall) vorliegt oder nicht. Diese Klassifikationen werden dann Radiolog:innen und anderen Ärzt:innen zur Verfügung gestellt, die dann basierend auf ihrer Expertise und der Einschätzung des Algorithmus ihre finale Diagnostikentscheidung treffen.

Da in der Medizin, die Diagnostikentscheidung von Menschen getroffen werden muss, müssen daher ausgebildete Radiolog:innen und Ärzt:innen die KI-Einschätzungen erst validieren. Daher spricht man nicht davon, dass in diesem Bereich Arbeit automatisiert wird, sondern eher davon, dass Menschen in ihrer Arbeit unterstützt werden. Im Folgenden, werde ich anhand von unseren Projekten aus der Radiologie kognitive Herausfordernden bei dieser KI-unterstützten Entscheidungsfindung beschreiben.

Nehmen wir an, dass eine KI einen bestimmten Diagnostikvorschlag generiert. Nun trifft eine solche KI-getroffene Einschätzung auf die Einschätzung von Radiolog:innen, dann können wir, wenn wir vereinfachend Alternativdiagnosen außen vor lassen, in Abbildung 1 zwischen vier möglichen Entscheidungsszenarien unterscheiden (Jussupow et al. 2021). Im

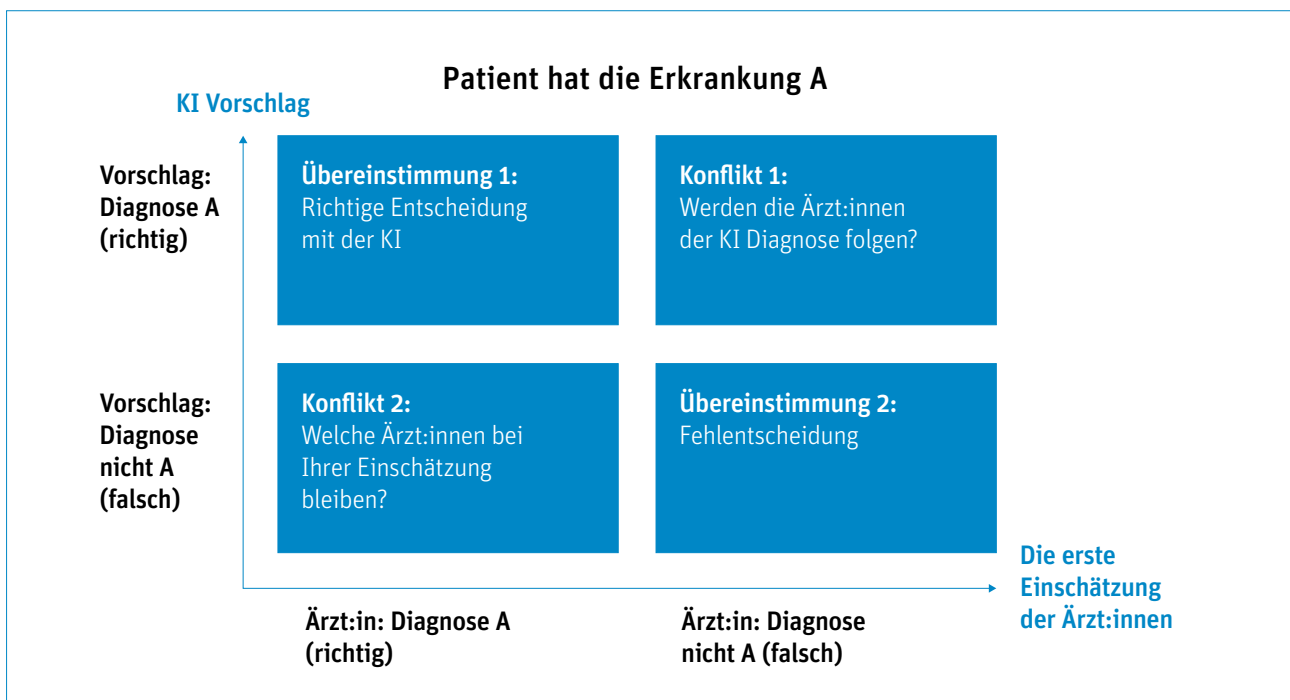


Abbildung 1: Entscheidungskonflikte in der Interaktion zwischen Ärzt:innen und KI-Systemen in der Diagnostik basierend auf der Tabelle 1 in Jussupow et al. (2021).

Folgenden werden die spezifischen kognitiven Herausforderungen und Chancen in jedem dieser vier Fälle erläutert.

Im Fall der Übereinstimmung 1, wird eine richtige Ersteinschätzung von Radiolog:innen durch ein KI-System bestätigt. Dies hat verschiedene Auswirkungen. Zum einen sind sich die bestätigten Radiolog:innen in ihrer Entscheidung sicherer als diejenigen, die keine KI-Unterstützung haben (Jussupow et al. 2021). Insbesondere Radiolog:innen mit wenig Erfahrung profitieren dabei sehr von der Zweitmeinung des KI-Systems (Gaubé et al. 2021). Dies hilft Radiolog:innen schneller Entscheidungen zu treffen. Wenn sie z.B. selbst nichts Auffälliges in der Diagnostik des Bildes sehen und das KI-System auch nichts entdeckt hat, können sie dann recht schnell zum nächsten Fall wechseln (Jussupow, Spohrer, and Heinzl 2022).

Insgesamt kann so die Unterstützung durch das KI-System zu schnelleren und besseren Entscheidungen beitragen, indem die Unsicherheit bei medizinischen Entscheidungen reduziert wird und die KI eine Zweitmeinung zur Verfügung stellt. Allerdings wird oft eine reine Bestätigung durch das KI-System als wenig hilfreich empfunden, vor allem, wenn dies nicht dazu beiträgt, dass man sich selbst sicherer fühlt (z.B. im Falle von höherer Expertise) (Jussupow, Spohrer, and Heinzl 2022). Ob die Systeme in der Praxis als hilfreich empfunden werden, ist dabei vor allem davon abhängig, wie die Radiolog:innen mit Fällen umgehen, in denen das KI-System einen Vorschlag macht, der nicht mit der Ersteinschätzung übereinstimmt. In diesen Fällen müssen die Radiolog:innen zusätzlichen kognitiven Aufwand betreiben, um das KI-System zu validieren und entscheiden, ob sie ihrer eigenen Einschätzung oder dem KI-System folgen (Jussupow et al. 2021).

Einige Experimente konnten zeigen, dass vor allem wenn ein KI-System eine Fehleinschätzung gibt, die diagnostische Genauigkeit drastisch leidet (Jussupow et al. 2021; Gaubé et al. 2021; Dratsch et al. 2023). Gleichzeitig führen Fehler von KI-Systemen dazu, dass Menschen sehr schnell das Vertrauen in den Algorithmus verlieren (Jussupow, Benba-

sat, and Heinzl 2020). Dennoch ist es wichtig, dass ein kognitiver Konflikt überhaupt entsteht, da dies den Radiolog:innen ermöglicht weitere Quellen hinzuzuziehen und zum Beispiel Kolleg:innen oder Expert:innen um Unterstützung zu bitten (Jussupow et al. 2021). Bedenklicher ist es, wenn der vierte Quadrant in der Abbildung 1 auftritt und kein Konflikt zwischen der eigenen Einschätzung und dem Vorschlag des KI-Systems entsteht. Dies ist z.B. der Fall, wenn schon die erste Einschätzung basierend auf dem Vorschlag des KI-Systems gebildet wird, da dies zu einem Automation Bias (Dratsch et al. 2023) führen kann. In diesem Fall wird dem KI-System zu sehr vertraut, dass nicht mehr validiert wird, wie dieses System zu Entscheidungen kommt.

Obwohl viele Studien darauf hindeuten, dass die diagnostische Genauigkeit zunimmt, wenn KI-Systeme die medizinische Entscheidungen unterstützen, stellt die Nutzung von KI-Unterstützung Radiolog:innen vor neue kognitive Herausforderungen. Sie können zwar schneller Entscheidungen treffen und erhalten eine zweite Einschätzung, die in manchen Fällen auch auf Dinge hinweisen kann, die sie selbst übersehen hätten – allerdings sind solche Systeme oft noch sehr spezialisiert auf kleine Aufgabenstellungen. Zudem ist die Akkuratheit der Systeme oft schwer nachzuvollziehen (Lebovitz et al. 2021), insbesondere da es oftmals unklar ist, auf welchen Daten solche Systeme trainiert werden und ob in den Trainingsdaten systematische Verzerrungen (Biases) vorkommen, die zu Fehleinschätzungen bei bestimmten demographischen Gruppen führen. Außerdem ist aktuell die Qualität der KI-Diagnostik oft noch von der technischen Ausstattung abhängig, da die Systeme abhängig von den eingesetzten bildgebenden Verfahren unterschiedliche Ergebnisse liefern.

Daher können KI-Systeme aktuell weniger erfahrene Radiolog:innen oder Ärzt:innen mit anderen Spezialisierungen unterstützen, um dem Personalangel in der Medizin entgegenzuwirken. Allerdings müssen in der Praxis Radiolog:innen oft Differentialdiagnosen stellen, verschiedene Körperteile und Bilderqualitäten befunden und verschiedene Ergebnisse miteinander integrieren, um gute Befunde und Behandlungspläne abzuleiten. Diese Tätigkeiten können

KI-Systeme aktuell nur bedingt ausführen. Ausgebildete Radiolog:innen können diese Systeme nicht ersetzen, auch weil sich die Rolle der Radiolog:innen durch die Technologie verändern wird. Die techni-

schen Entwicklungen (wie etwa Chat-GPT) haben allerdings das Potential die radiologische Arbeit neben der Diagnostik zu unterstützen und so langfristig die Arbeit von Ärzt:innen zu erleichtern.

Literaturangaben:

Dratsch, T.; Chen, X.; Mehrizi, M. R.; Kloeckner, R.; Mähringer-Kunz, A.; Püsken, M.; Baeßler, B.; Sauer, S.; Maintz, D.; Pinto Dos Santos, D.: Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance. *Radiology* 307 (4): e222176. 2023.

Gaube, S.; Suresh, H.; Raue, M.; Merritt, A.; Berkowitz, S. J.; Lerner, E.; Coughlin, J.F.; Gutttag, J.V.; Colak, E.; Ghassemi, M.: Do as AI Say: Susceptibility in Deployment of Clinical Decision-Aids. *NPJ Digital Medicine* 4 (31). 2021. Online:<https://doi.org/10.1038/s41746-021-00385-9>.

Jussupow, E.; Benbasat, I.; Heinzl, A.: Why Are We Averse Towards Algorithms? A Comprehensive Literature Review On. *Twenty-Eighth European Conference on Information Systems (ECIS2020)*. 2020. S. 1–16.

Jussupow, E.; Spohrer, K.; Heinzl, A.: Radiologists' Usage of Diagnostic AI Systems. *Business & Information Systems Engineering* 64 (3). 2022, S. 293–309.

Jussupow, E. Spohrer, K.; Heinzl, A.; Gawlitza, J.: Augmenting Medical Diagnosis Decisions? An Investigation Into Physicians' Decision Making Process With Artificial Intelligence. *Information Systems Research* 32 (3). 2021, S. 713–35.

Lebovitz, S.; Levina, N.; Lifshitz-Assa, H.: Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What. *MIS Quarterly* 45 (3). 2021, S. 1501–26

Prof. Dr. Ekaterina Jussupow

Technische Universität Darmstadt
Fachbereich 1: Rechts- und Wirtschaftswissenschaften
Fachgebiet Wirtschaftsinformatik

Lebenslauf in Kürze

Ekaterina Jussupow ist seit März 2023 Juniorprofessorin für Wirtschaftsinformatik an der Technischen Universität Darmstadt. Nach ihrem Studium in Psychologie und Literaturwissenschaften an der LMU München und der Universität Heidelberg wechselte Prof. Jussupow nach einer kurzen Zeit in der Industrie, u.a., bei der Boston Consulting Group, in die Wirtschaftsinformatik. Im Jahr 2021 promovierte sie in Wirtschaftsinformatik an der Universität Mannheim zur Rolle von KI-Systemen in der Radiologie und war danach als Post-Doktorandin ebenda tätig.



Mit ihrem Hintergrund in Psychologie konzentriert sie sich auf die menschliche Interaktion mit KI und untersucht, wie Individuen und Organisationen KI-Systeme in der Praxis nutzen. Sie versucht zu verstehen, wie KI-Systeme die menschliche Entscheidungsfindung unterstützen und herausfordern.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist es zu verstehen, wie Menschen bessere Entscheidungen treffen können.

Luisa Pumplun

Mensch-KI-Interaktion im Gesundheitswesen

Die Schnittstelle von Technologie und Menschen fasziniert mich seit jeher. Nur wenn wir verstehen, welchen Herausforderungen Nutzerinnen und Nutzer in ihrem privaten und professionellen Alltag gegenüberstehen und auf ihre Bedürfnisse eingehen, kann Technologie das volle Potenzial entfalten. Dies gilt insbesondere für den Bereich des Maschinellen Lernens, einem Teilgebiet der Künstlichen Intelligenz. Diese spezielle Technik kann Arbeitsprozesse abbilden, indem sie auf Daten und statistische Algorithmen zurückgreift, um Modelle der Wirklichkeit zu trainieren. Das macht sie zugleich sehr leistungsfähig, aber auch fehleranfällig. Eigenschaften, die sie für den Einsatz im Gesundheitswesen interessant machen, aber auch zu einer Herausforderung – gerade in der Diagnostik, die hohe Fehlerkosten verursachen kann – werden lassen.

In diesem Spannungsfeld gilt es, den Menschen in den Mittelpunkt der Entwicklung zu stellen. Meine Forschung nimmt daher eine zweiteilige Perspektive ein. Diese untersucht, wie Systeme rund um das maschinelle Lernen organisatorisch und technisch gestaltet werden sollten, um so effizient wie möglich zu sein und gleichzeitig die Bedürfnisse der Menschen im Gesundheitswesen zu berücksichtigen (z.B. Mesbah und Pumplun, 2020; Pumplun et al., 2021). Meine Studien befassen sich konkret mit den Voraussetzungen in Kliniken zur Einführung Maschinellen Lernens (z.B. Pumplun et al., 2021; Pumplun und Buxmann, 2020), der Erklärbarkeit darauf aufbauender Systeme („Explainable Artificial Intelligence“) (z.B. Pumplun et al., 2023), den spezifischen Anforderungen von Ärztinnen und Ärzten, Patientinnen und Patienten (z.B. Pumplun et al., 2021) sowie den Auswirkungen des Maschinellen Lernens auf Leistung und Wissen (z.B. Ellenrieder et al., 2023).

Die Ergebnisse dieser Forschung zeigen, dass gerade im Gesundheitswesen spezielle Voraussetzungen für die Implementierung vom Maschinellen Lernen beste-

hen, seien es die Verfügbarkeit und Sensitivität von Patientendaten, der Mangel dedizierten Budgets für die Entwicklung Maschinellen Lernens in Kliniken oder medizinethische Aspekte wie das „Recht auf Nichtwissen“, u.v.m (vgl. Pumplun et al., 2021). Auch aus Sicht von Patientinnen und Patienten ist es essenziell, spezifische Voraussetzungen für die Anwendung vom Maschinellen Lernen zu berücksichtigen. So ist es möglich, dass das System nur sinnvolle Diagnosevorschläge ausgibt, wenn diesem die medizinische Vorgeschichte inklusive Medikation der Person als Kontext übergeben werden (vgl. Mesbah und Pumplun, 2020). Ärztinnen und Ärzte auf der anderen Seite haben einen hohen Bedarf an erklärbaren Systemen. Dabei ist es wichtig, im hektischen Klinikalltag die richtigen Erklärungen richtig zu präsentieren, um eine schnelle und präzise Unterstützung der Diagnostik durch ansonsten Black-Box-Systeme wie bspw. neuronale Netze zu erreichen. Dies kann durch einen sinnvollen Trade-Off zwischen der Bereitstellung von Erklärungen und dem kognitiven Aufwand, der mit dem Lesen derselbigen einhergeht, gelingen. Keine einfache Aufgabe sind Erklärungen für das Maschinelle Lernen doch oft komplex, statistisch geprägt und auf professionelle Entwicklerinnen und Entwickler ausgelegt anstatt auf Anwenderinnen und Anwender. Dem kann insbesondere durch die Einhaltung verschiedener Guidelines beim Design der Erklärungen entgegengewirkt werden. Zum Beispiel sollten verschiedene Arten von Erklärungen berücksichtigt werden, sowohl grundlegende Erklärungen über die Datenbasis, den angewendeten Algorithmus, etc. als auch Modell-bezogene Erklärungen wie der Einfluss von Eingangsdaten auf die Ausgabe des Systems. Zudem sollten den Anwenderinnen und Anwendern bekannte Arbeitsprozesse und Wissen bestmöglich berücksichtigt und deterministische Elemente, wie beispielsweise Biopsie Ergebnisse ähnlicher Fälle plus entsprechende Vorhersage durch Maschinelles Lernen, bei den Erklärungen eingebunden werden (vgl. Pumplun et al., 2023). Erklärbarkeit hat zudem

weitere positive Auswirkungen auf die Nutzerinnen und Nutzer von Maschinellern. Sie gewährleistet, dass Menschen, auch wenn sie mehr und mehr Reviewer als Akteure in der Zusammenarbeit mit Maschinellern werden, von den Systemen lernen können (vgl. Ellenrieder et al., 2023). Ein Aspekt, der in Anbetracht fortschreitender Automatisierung nicht zu vernachlässigen ist. Nur wenn Menschen die dahinterliegenden Prozesse weiterhin verstehen und entsprechende Daten generieren, kann auch das Maschinelle Lernen der Zukunft sinnvoll auf diesen trainiert werden.

Es ist also nicht zu vernachlässigen, welche Rolle der Mensch, und damit meine ich nicht nur die Entwicklerin oder den Entwickler des Maschinellen Lernens, für das Fortschreiten der Technologie spielt. Ein Verständnis über die Mensch-KI-Interaktion aufzubauen und diese bei der Entwicklung neuer Technologien zu berücksichtigen, wird maßgeblich dazu beitragen, dass Maschinelles Lernen im Gesundheitssektor sinnstiftend und langfristig eingesetzt werden kann und einen wirklichen Wert für Patientinnen und Patienten schafft.

Literaturangaben:

Ellenrieder, S.; Kallina, E.; Pumplun, L.; Gawlitza, J. F.; Ziegelmayr, S.; Buxmann, P.: Promoting Learning Through Explainable Artificial Intelligence: An Experimental Study in Radiology. International Conference of Information Systems, Hyderabad, India 2023. In Druck.

Mesbah, N.; Pumplun, L.: "Hello, I'm here to help you" – Medical care where it is needed most: Seniors' acceptance of health chatbots. European Conference of Information Systems, Virtual, 2020.

Pumplun, L.; Buxmann, P.: Intelligent Systems and Hospitals: Joint Forces in the Name of Health? Internationale Tagung Wirtschaftsinformatik, Potsdam, Germany 2020.

Pumplun, L.; Fecho, M.; Wahl, N.; Peters, F.; Buxmann, P.: Adoption of machine learning systems for medical diagnostics in clinics: qualitative interview study. Journal of Medical Internet Research, 23(10), e29301. 2021.

Pumplun, L.; Peters, F.; Gawlitza, J. F.; Buxmann, P.: Bringing Machine Learning Systems into Clinical Practice: A Design Science Approach to Explainable Machine Learning-Based Clinical Decision Support Systems. Journal of the Association for Information Systems, 24(4), 953-979. 2023.

Dr. Luisa Pumplun

Ehemalige Mitarbeiterin an der TU Darmstadt

Fachbereich: Wirtschaft und Recht

Fachgebiet Software & Digital Business

l.pumplun@t-online.de

Lebenslauf in Kürze

Dr. Luisa Pumplun ist seit 2022 im Bereich Data Science bei der Merck KGaA tätig. Zuvor promovierte sie am Fachgebiet Software & Digital Business, Fachbereich Wirtschaft und Recht, Technische Universität Darmstadt. Sie hat zudem einen Master-Abschluss in Wirtschaftsingenieurwesen mit Schwerpunkt Elektro- und Informationstechnik. Ihre Forschungsinteressen liegen in der Einführung und Nutzung datengetriebener Technologien wie dem Maschinellen Lernen auf individueller und organisationaler Ebene mit einem Schwerpunkt auf dem Gesundheitssektor.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist, neue technologische Entwicklungen im Bereich des Maschinellen Lernens zu verfolgen und dabei insbesondere den Menschen und seine Bedürfnisse im Blick zu behalten.



3. Die Versorgungsperspektive

Anastasia Paschalidou

Wie wollen wir mit KI leben? – Systeme Künstlicher Intelligenz in der Sozialen Arbeit

Aktuelle Anwendungsbereiche von Systemen Künstlicher Intelligenz (KI) in der Sozialen Arbeit sind seit Jahren Thema und Gegenstand von Diskussionen (vgl. Beranek/Hill u.a. 2019). Im Fokus dieses Beitrages stehen die Fragen nach möglichen Anwendungsbereichen sowie den Grenzen der Nutzung von KI-Systemen in der Sozialen Arbeit. Ich vertrete hier die Hypothese, dass wir KI-Systeme nur durch einen interdisziplinären Zugang in Ihrer „Potentialität“ erfassen werden und auch dadurch die vielfältigen Risiken wahrnehmen und etwas minimieren können. Hierbei gilt es menschenfreundliche Wege im Zusammenleben mit KI-Systemen zu erörtern und zu gestalten.

Disruptionspotenzial von KI-Systemen

Digitale Teilhabe ist zu einer existentiellen Voraussetzung sozialer Teilhabe geworden. Die soziale, zeitliche und räumliche Durchdringung digitaler Technologien in vielen Lebensbereichen verändert auch Sozialsysteme. Sie bilden eigene Regeln, Werte und Machtstrukturen heraus (Kutscher 2019). Dies gilt vor allem für disruptive Techniken. Das Disruptionspotenzial von KI-Systemen zwingt auch Fachkräfte in der Sozialen Arbeit sich damit auseinanderzusetzen. Bei der Auseinandersetzung mit KI-Systemen geht es nicht nur um einzelne „digitale Tools“, die zur Verfügung gestellt werden. Es geht nicht nur um Software oder Hardware, sondern auch um Menschenbilder und Menschenrechte sowie bestimmte Technikverständnisse (Beranek 2021). Bei der Entwicklung neuer KI-Systeme muss jedes Mal entschieden werden, was berechenbar sein soll und was wie formalisiert werden soll. Diese Vorstellungen sind entscheidend in der Entwicklung und in der Anwendung digitaler Innovationen (Allert u.a. 2017). Die Herangehensweise an KI-Systeme und die entsprechenden Deutungsmuster sind von grundlegender Bedeutung, um einen innovativen und auch verantwortlichen Umgang mit KI-Systemen zu eröffnen. Einige Deutungsmuster reduzieren die Komplexität

von KI-Systemen. Die Systemstruktur einer KI ist äußerst faszinierend. Wir sind nicht nur mit einem „technischen Werkzeug“ konfrontiert, obwohl vielfältige Werkzeugaspekte identifiziert werden können. KI-Systeme sind nicht nur überschaubare Bearbeitungsprogramme mit möglichen interaktiven Elementen. Beim Einlassen auf ein KI-System stabilisieren wir, dynamisieren und potenzieren wir dieses System in seiner technischen Potentialität. Diese KI-Systeme können nicht als isoliert zu betrachtende „Werkzeuge“ einer „Welt der Technik“ betrachtet werden, sondern sie werfen erkenntnistheoretische, soziale und ethische Fragen auf. Die Eigendynamik der Technikentwicklung eines KI-Systems und das technische Handeln von Individuen sind miteinander verknüpft. Diese Verknüpfung erzeugt spezifische Dynamiken, die wir auch gesellschaftlich zu verantworten haben. Um diese Dynamik zu steuern, bedarf es eines Technikverständnisses, welches die Komplexität technischen Wissens und „technischer Potentialität“ in den Blick bekommt (Poser 2007: 229). Der Geltungsbereich sowie die Grenzen der Nutzung von KI-Systemen müssen schon allein aus diesem Grund, auch aus der Perspektive der Sozialen Arbeit erforscht und diskutiert werden.

Fachkräftemangel und die Sehnsucht nach Entlastung durch KI-Systeme

KI-Systeme haben das Potential viele Aspekte unserer Lebenswelt zu verändern. Die KI-Technologie entwickelt sich von Jahr zu Jahr weiter, da KI-Systeme immer spezifischer zugeschnitten, zugänglicher und einfacher zu nutzen sind und die KI-Fähigkeiten allgemein weiter zunehmen, ist eine weitere Implementierung in die alltägliche Lebens- und Arbeitswelt von Menschen zu erwarten (vgl. IBM 2022). Da es vielfältige Definitionen von KI-Systemen gibt, beziehe ich mich in diesem Beitrag auf die aktualisierte Definition der OECD. Nach dieser Definition sind KI-Systeme maschinengestützte Systeme, welche aufgrund expliziter

oder impliziter Ziele aus den Eingaben, die diese Systeme erhalten, bestimmte Ausgaben wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen generieren und ableiten können. Diese KI-Systeme können Einfluss auf physische oder virtuelle Umgebungen haben. Es existieren viele unterschiedliche KI-Systeme, die sich in ihrem Grad an Autonomie und Anpassungsfähigkeit und nach dem Einsatz unterscheiden können (OECD 2023).

Der steigende Arbeitskräftemangel in vielen Arbeitsbereichen – auch in der Sozialen Arbeit – führt u.a. dazu, dass viele Unternehmen und auch Einrichtungen der Sozialen Arbeit die Einführung von KI-Systemen befürworten und vorantreiben. Laut einem IBM-Bericht vom Jahre 2022 wollen ca. 25% der Unternehmen durch die Einführung von KI-Systemen, den Fachkräftemangel besser bewältigen. KI-Systeme können Abläufe optimieren und diesen Mangel an Personal ausgleichen. In dem Maße, wie KI Systeme mittlerweile Wirtschaft und Gesellschaft durchdringen, stellt sich auch in der Sozialen Arbeit die Frage, wie KI-Systeme für die Praxis der Sozialen Arbeit entwickelt werden können und welche institutionellen Rahmenbedingungen hierfür notwendig sind. Der Fachkräfte- und Ressourcenmangel in der Sozialen Arbeit und der Bedarf, Entlastung durch KI-Systeme zu erfahren, ist sehr hoch. In Anbetracht des Fachkräftemangels in der Sozialen Arbeit mit einer Fachkräftelücke von ca. 20.500 Stellen, könnten KI-Assistenzsysteme entlastend wirken (Hickmann/Koneberg 2022: 1f). KI-Systeme könnten Daten analysieren, um festzustellen, wo Bedarf besteht und wie begrenzte Ressourcen am effektivsten eingesetzt werden können. Die Soziale Arbeit benötigt insbesondere sozialgerechte KI-Assistenzsysteme, die auch eine nachhaltige und zielgruppengerechte Umsetzung von Entlastung ermöglichen.

Vielfalt innovativer Anwendungen von KI-Systemen in der Sozialen Arbeit

Fachkräfte der Sozialen Arbeit und Adressat*innen werden durch KI Systeme mit einer Vielfalt innovativer Anwendungen konfrontiert. KI-Systeme können im Alltag und in belastenden Situationen konstruktiv assistieren und die Benachteiligung von Menschen in prekären Lebenslagen abbauen. KI-Systeme als Teil der Informatik können auf Programmen oder auf ma-

schinellem Lernen beruhen. In großen Datenmengen können sie Muster erkennen, schnelle komplexe Berechnungen und große Datenmengen verarbeiten und nach gewünschten Kriterien auch Daten überprüfen. Kontext- und handlungsbezogene Informationen können durch KI-Systeme zur Verfügung gestellt werden. Ebenso werden schon KI-Systeme in der Analyse der Dokumentation eingesetzt. Das Auswählen von Daten und Einpflegen bestimmter Inhalte ist mit vielen Entscheidungsprozessen verbunden. Jedoch sind solche Datenanalysen nicht neutral. KI-Systeme können nur mit bereits vorhandenen und ausgewählten Daten trainiert werden. Diese Daten spiegeln häufig vorhandene gesellschaftliche Strukturen wieder. So können KI-Systeme auf der einen Seite das Arbeitsleben entlasten und auf der anderen Seite Verzerrungen (Bias-Aspekte) produzieren und z.B. Sexismus und Rassismus fördern (vgl. Noble 2018; Beranek et al. 2019; Lucy/Bamman 2021). Auf diesen Zusammenhang gehe ich später noch einmal ein. Jetzt wenden wir uns erst einmal den vielfältigen Anwendungen zu.

KI-Systeme können verwendet werden, um Muster in sozialen und medizinischen Daten zu identifizieren und Frühwarnzeichen für Kindesmisshandlung und Vernachlässigung zu erkennen. Dies kann dazu beitragen, schneller auf potenzielle Gefahren zu reagieren. Hier können KI-Systeme zur schnellen Identifizierung von vorher definierten „Warnsignalen“ eingesetzt werden. In der Praxis der Sozialen Arbeit finden in Deutschland zurzeit überwiegend Pilotversuche statt, wie beispielsweise in der Kinder- und Jugendhilfe (Steiner/Tschopp 2022). Für die Bedarfsermittlung und Gefährdungseinschätzung in der Kinder- und Jugendhilfe existieren im deutschsprachigen Raum Pilotprojekte, die KI-Systeme implementieren (z.B. Predictive Modelling). Predictive Modelling existiert bereits seit den 1980er Jahren. Eingesetzt wird es insbesondere im Gesundheitsbereich. Hier sind wir mit einem Big Data basierten Expert*innensystem konfrontiert. Ein Verfahren der Analytik, das Daten mit statistischer Modellierung, Data-Mining-Techniken und maschinellem Lernen kombiniert (Schneider u.a. 2019; Schrödter u.a. 2020). Diskutiert wird, inwiefern es möglich sein kann, dass durch digitale Operationalisierung z.B. im Feld der Kindeswohlgefährdung eingesetzt werden kann (vgl. von Reder u. Koska

2022). Auf der Ebene der Intervention und Nachsorge können KI-Systeme unterstützend eingesetzt werden. Zum Beispiel bei der Identifikation von rückfallgefährdeten jungen Adressat*innen in der Suchthilfe, wie im Präventionsprojekt „Interact Blueprevent“ umgesetzt wird. So könnte auch die Zuweisung von Ressourcen in der Sozialen Arbeit „optimiert“ werden. Inwiefern dies für Adressat*innen der Sozialen Arbeit auch eine Verbesserung ihrer Lebenssituation eröffnen könnte, müsste sehr achtsam erprobt und erforscht werden. Lehrkonzepte und -materialien können evaluierungsbegleitend gemeinsam mit den Fachkräften entwickelt werden. KI-Chatbots werden schon in der psychischen Gesundheitsversorgung eingesetzt, um Menschen mit emotionalen Problemen und psychischen Erkrankungen Unterstützung und Ressourcen anzubieten. Solche lernenden Systeme in der Beratung werden auch im Feld der Sozialen Arbeit in Pilotprojekten entwickelt (vgl. CariFIX/Caritas). Sie können eine erste Anlaufstelle für Hilfe und Beratung sein. KI-Systeme können auch Sprachanalysen durchführen, um Anzeichen von Depressionen, Angstzuständen oder anderen psychischen Gesundheitsproblemen bei Text- oder Sprachnachrichten zu identifizieren.

In der Praxis der Sozialen Arbeit werden in Pilotprojekten neue KI Assistenzsysteme, welche pädagogische Prozesse „transparent“ dokumentieren sollen, erprobt. Solche Pilotprojekte werben mit dem Versprechen Fachkräfte in ihrer Arbeit zu entlasten. Dabei soll aber auch die pädagogische Arbeit mit der Unterstützung von KI-Systemen transparent, sozialgerecht orientiert und partizipativ umgesetzt werden. Diese vielfältigen Anwendungsmöglichkeiten von KI-Systemen in der Sozialen Arbeit sind beeindruckend. Wenn sie doch nur so einfach zu realisieren wären, ohne Risiken für Adressat*innen zu produzieren.

Bias-Aspekte reflektieren und entgegenwirken

Fachkräfte der Sozialen Arbeit tragen eine besondere Verantwortung, um Ihren Auftrag gerecht zu werden. Sie müssen sicherstellen, dass die Rechte der Adressat*innen geschützt und gestärkt werden, sowie die Vulnerabilität ihrer Adressatinnen berücksichtigt wird. Hierzu gehört es Risiken zu minimieren, denn die sogenannten „Fehlerkosten“ sind in der Sozialen Arbeit sehr hoch. Es geht um Menschen, die sich über-

wiegend in prekären Lebenssituationen befinden. Es geht nicht um Werkzeuge, die optimiert oder justiert werden müssen. Ja und es geht hierbei auch um Menschenrechte und Machtverhältnisse. Das Risiko der Entmündigung ihrer Adressat*innen (Kutscher 2019), die Gefahr der Verletzung der Menschenrechte (Bera- nek 2021) und die Gefahr einer falschen Vorverurteilung (Catanzaro 2020) sind keine Phantasien, sondern sie beruhen schon auf Erfahrungswerten mit KI-Systemen. Die Rechte der Adressat*innen in der Sozialen Arbeit, aber auch die der Fachkräfte, können nur durch einen reflektierten und verantwortlichen Umgang mit KI-Systemen gestärkt werden. Um diese Rechte auch zu schützen ist es notwendig, dass die Soziale Arbeit durch Kooperationen mit anderen Professionen (wie z.B. der Sozioinformatik und der Informatik) das KI-Anwendungspotential gemeinsam erforscht und bedarfsorientiert mitgestaltet. Die Risiken müssen interdisziplinär erforscht werden und minimiert werden. Bei einer Kooperation mit Unternehmen müssen Richtlinien vereinbart werden. Hier gilt es auch die Infrastruktur hinter KI-Innovationen zu kennen, um diese Richtlinien für eine Zusammenarbeit zu entwickeln.

Welche Brisanz dies hat, zeigt auch der IBM Bericht 2022. Er verweist darauf, dass die Mehrheit der Unternehmen keine wichtigen Schritte unternommen hat, um sicherzustellen, dass ihre KI-Systeme vertrauenswürdig und verantwortungsbewusst entwickelt werden, z. B. die Reduzierung von Verzerrungen (BIAS-Aspekte) (74 %), die Verfolgung von Leistungsschwankungen und Modelldrift (68 %) und die Sicherstellung, dass sie KI-gestützte Entscheidungen erklären können (61 %) (IBM 2022: 3). Diese Versäumnisse wirken sich auf die Konstruktion von KI-Systemen aus und vervielfältigen die Risiken für Nutzer*innen und somit auch die Risiken von Adressat*innen in der Sozialen Arbeit (IBM 2022). Eine Umfrage von Forbes Advisor zeigt, dass 76 % der Verbraucher über Fehlinformationen durch KI wie Google Bard, ChatGPT und Bing Chat besorgt sind (Haan 2023). Besorgt sind die Nutzer*innen nicht ohne Grund. Zum Beispiel haben Nicoletti und Bass bei bestimmten KI-Text-Bild-Modellen nachgewiesen, dass rassistische und sexistische Stereotypen reproduziert werden (vgl. Nicoletti/Bass 2023). Solche Verzerrungen könnten behoben werden, jedoch unterliegen vielfältige Anwendungen von KI-Systemen auch ökonomischen

Interessen. Diese ökonomischen Interessen erwachsen aus den Strukturen und Dynamiken einer bestimmten Gesellschaftsordnung. Hier gilt es einen verantwortlichen und sozialgerechten Umgang zu eröffnen und auch Richtlinien für KI-Systeme zu etablieren. Die Einhaltung ethischer Richtlinien durch unabhängige Ethikkommissionen und regelmäßige unabhängige Überprüfungen der KI-Systeme durch gemeinnützige Forschungs- und Advocacy-Organisationen (wie z.B. AlgorithmWatch) könnten hier wegweisend sein (Grimm u.a. 2019). Die gesetzlichen Grundlagen bei der Anwendung eines KI-Systems sollten Fachkräften in der Sozialen Arbeit bekannt sein. Auch das Europäische Parlament hat sich damit auseinandergesetzt und beschlossen KI-Systeme stärker zu regulieren. Dies ist das weltweit erste umfassende KI-Gesetz (Europ. Parl. 2023). Dieses KI-Gesetz soll gewährleisten, dass in der EU eingesetzte KI-Systeme sicher, transparent, nachvollziehbar, nicht diskriminierend und umweltfreundlich sind bzw. werden.

Weitere gesetzliche Regulierungen werden sicherlich folgen, denn Fachkräfte der Sozialen Arbeit und Adressat*innen werden durch die Verwendung von KI-Systemen leider mit einer Vielfalt verflochtener Diskriminierungsformen konfrontiert. Wie schon erwähnt, sind Datenanalysen nicht neutral. Sie spiegeln häufig vorhandene gesellschaftliche Strukturen wieder. Problematisch bei der Entwicklung von KI Systemen ist insbesondere die fehlende Diversität auf Seite der Entwickler*innen. Nur 10% der Angestellten im KI-Bereich von Google sind Frauen und gar nur 2,5% bzw. 3,6% sind Afroamerikaner*innen bzw. hispanischer Herkunft (West et al.: 2019). Dies gilt für die Entwickler*innen, die die Weichen für die KI Systeme stellen. Dies gilt nicht für die Zuarbeiter*innen, die Daten unter prekären Arbeitsverhältnissen einpflegen. Hinter vielen KI-Innovationen steckt auch eine Geschichte von Ausbeutung im globalen Süden. Zum Beispiel wurden für die Entwicklung des Textgenerators von ChatGPT von OpenAI kenianische Arbeiter*innen engagiert, die sogenannte Verzerrungen (Bias-Aspekte) reduzieren sollten. In diese KI-Innovation mussten von diesen Arbeiter*innen markierte Beispiele von Gewalt, Hassreden und sexuellem Missbrauch eingepflegt werden. Auf diese Weise lernte das KI-System, diese Formen zu erkennen und herauszufiltern, bevor

sie die Nutzer erreichen. Sie durften für weniger als 2 Dollar pro Stunde mit emotional belastenden Inhalten arbeiten. Supervision wurde hier nicht angeboten. Der Outsourcing-Partner von OpenAI in Kenia war Sama, ein in San Francisco ansässiges Unternehmen, das Mitarbeiter in Kenia, Uganda und Indien beschäftigt (vgl. Perrigo 2023). Die Geschichte der Arbeiter*innen, die ChatGPT möglich gemacht haben, bietet einen Einblick in die Arbeitsbedingungen in diesem oftmals nicht so erwähnten Aspekt der KI-Industrie. Dieser spielt jedoch eine wesentliche Rolle bei der Entwicklung von KI-Systemen für den öffentlichen Konsum. Nicht anders sieht es laut vielfältiger Recherchen bei anderen KI Global Player aus, wie die Harvard-Ökonomin Shoshana Zuboff eindrucksvoll erforscht hat (Zuboff 2018). Für die Nutzung von KI-Systemen in der Sozialen Arbeit müssen solche Verhältnisse thematisiert und reflektiert werden. Es gehört zum Auftrag der Sozialen Arbeit, Ausbeutungsverhältnisse zu reflektieren und diesen entgegenzuwirken. Dazu gehört es auch wie Roberto Simanowski erläutert, denjenigen Gehör zu verschaffen, die im gesellschaftlichen Diskurs bisher marginalisiert waren: Frauen, Migrant*innen, Homo- und Transsexuelle. Die Perspektive dieser Gruppen soll in den Trainingstexten der KI-Systeme laut Simanowski gestärkt werden, z.B. durch eine Art Quotenregelung für Minderheitsdaten (Simanowski 2020).

Hinzu kommt, dass ein KI-System auch nur so divers sein kann, wie die eingepflegte Datensammlung. Das führt dazu, dass KI-Systeme eine Vielzahl möglicher Verzerrungen (engl. Bias-Aspekte) aufzeigen können, da hier Vorurteile bewusst und unbewusst „miteingepflegt“ werden. Durch WEIRD (western, educated, industrialized, rich and democratic societies) Samples kam es z.B. im Falle Amazon dazu, dass eine zur Bewertung von Bewerber*innen eingesetzte KI-System hauptsächlich weiße Männer zur Einstellung vorschlug, da angeblich diese in der Vergangenheit am häufigsten erfolgreich in der Firma arbeiteten (vgl. Ming, 2019). Nach Simanowski ist auch politisches Engagement hier bedeutsam, welches beispielsweise die sozialen Folgen solcher BIAS- Aspekte oder prekären Arbeitsverhältnisse untersucht (Simanowski 2021: 91f.). KI-Systeme müssen transparent und Entschei-

dungsprozesse erklärbar sein. Anti-Bias-Algorithmen, die Bias mit korrigierenden Techniken minimieren und auch eine manuelle Überprüfung der Daten eröffnen, müssen weiterentwickelt werden.

Die Konzeptualisierung von KI-Systemen muss in konfliktiven Feldern stärker an eine inklusive Praxis rückgebunden werden. Sozialgerechte KI-Assistenzsysteme müssen gemeinsam mit den Menschen, die sie anwenden, und Menschen, die betroffen sind, interdisziplinär

entwickelt werden. In der Sozialen Arbeit unterliegen Methoden der Aushandlung (Prinzip der „Ko-Produktion“). Sie müssen prozessorientiert und machtsensibel sein. Methoden müssen evaluierbar sein und mit Blick auf die Adressat*innen auch ethisch gerechtfertigt sein (Legitimierung) (Galuske/Müller 2012). In einer Zeit multipler Krisendynamiken, wachsender sozialer und digitaler Ungleichheit sind besonders Fachkräfte der Sozialen Arbeit herausgefordert, hier einen sozialgerechten Umgang mit KI-Systemen zu entwickeln.

Literaturangaben:

- Allert, H.; Asmussen, M.; Richter, C. (Hrsg.): Digitalität und Selbst. Interdisziplinäre Perspektiven auf Subjektivierungs- und Bildungsprozesse. Bielefeld 2017.
- Beranek, A.; Hill, B.; Sagebiel, J.B.: Digitalisierung und Soziale Arbeit – ein Diskursüberblick, Soz. Passagen 11. 2019, S. 225–242.
- Beranek, A.: Künstliche Intelligenz und Menschenrechte. Politikum, 7(1). 2021, S. 4–12.
- Catanzaro, M.: In Spain, the VioGén algorithm attempts to forecast gender violence, veröffentlicht am 27.04.2020. Online: <https://algorithmwatch.org/en/viogen-algorithm-gender-violence/> [abgerufen am 18.12.2023].
- CariFIX, Caritas Projekt „Lernende Systeme in der Beratung“. Online: <https://www.caritas-digital.de/projekte/lernende-systeme-in-der-beratung/> [abgerufen am 18.12.2023].
- Europäisches Parlament -Aktuell: KI-Gesetz: erste Regulierung der künstlichen Intelligenz. 2023. Online: <https://www.europarl.europa.eu/news/de/headlines/society/20230601STO93804/ki-gesetz-erste-regulierung-der-kunstlichen-intelligenz> [abgerufen am 18.12.2023].
- Galuske M.; Müller W.: Handlungsformen in der Sozialen Arbeit, Geschichte und Entwicklung. In: Thole, W. (Hrsg.): Grundriss Soziale Arbeit. Wiesbaden 2012.
- Grimm, P.; Keber, T.; Zöllner, O. (Hrsg.): Digitale Ethik. Stuttgart 2019.
- Haan K.: Over 75% of Consumers are Concerned about Misinformation from Artificial Intelligence, Forbes Advisor survey. 2023. Online: <https://www.forbes.com/advisor/business/artificial-intelligence-consumer-sentiment/> [abgerufen am 18.12.2023].
- Hickmann, H.; Koneberg, F.: Die Berufe mit den aktuell größten Fachkräftelücken, IW-Kurzbericht, Nr. 67. 2022. Online: <https://www.iwkoeln.de/studien/helen-hickmann-fliz-koneberg-die-berufe-mit-den-aktuell-groesten-fachkraefteluecken.html> [abgerufen am 18.12.2023].
- IBM Global AI Adoption Index 2022, New research commissioned by IBM in partnership with Morning Consult, 2022, Corporation New Orchard Road Armonk, NY 10504. Produced in the United States of America May 2022. Online: <https://www.ibm.com/downloads/cas/GVAGA3JP> [abgerufen am 18.12.2023].
- Kutscher, N.: Digitalisierung in der Sozialen Arbeit und ethische Fragen. In: Kutscher et al. (Hrsg.): Handbuch Soziale Arbeit und Digitalisierung. Weinheim 2019, S. 41-54.
- Lucy, L.; Bamman, D.: Gender and Representation Bias in GPT-3 Generated Stories. In Association for Computational Linguistics (Hrsg.), Proceedings of the 3rd Workshop on Narrative Understanding. 2021, S. 48–55. Virtual: Association for Computational Linguistics.
- Ming, V.: Human insight remains essential to beat the bias of algorithms. Financial Times, 2019. Online: <https://www.ft.com/content/59520726-d0c5-11e9-b018-ca4456540ea6> [abgerufen am 18.12.2023].
- Noble, U.S.: Algorithms of Oppression: How Search Engines Reinforce Racism, NYU Press, 2018.
- Nicoletti, L.; Bass, D.: Humans Are Biased. Generative AI Is Even Worse: Stable Diffusion's text-to-image model amplifies stereotypes about race and gender – here's why that matters. 2023. Online: <https://www.bloomberg.com/graphics/2023-generative-ai-bias/> [abgerufen am 18.12.2023].
- Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD), Updated Definition AI. 2023. Online: <https://oecd.ai/en/wonk/ai-system-definition-update> [abgerufen am 18.12.2023].
- Perrigo, B.: Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. 2023. Online: <https://time.com/6247678/openai-chatgpt-kenya-workers/> [abgerufen am 18.12.2023].
- Poser, H.: Teleologie der Technik. Über die Besonderheiten technischen Wissens. In: Abel, G.; C. Renato; Hogrebe, W. u.a. (Hrsg.): Lebenswelten und Technologien. Berlin 2007, S. 217 – 233.
- Reder, M.; Koska, C.: Wie kann KI die Soziale Arbeit unterstützen? Porträt eines Pilotprojektes, Bayerisches Forschungsinstitut für digitale Transformation. 2022. Online: <https://www.bidt.digital/wie-kann-ki-die-soziale-arbeit-unterstuetzen/> [abgerufen am 18.12.2023].
- Schneider, D.; Seelmeyer, U.: Challenges in Using Big Data to Develop Decision Support Systems for Social Work in Germany. Journal of Technology in Human Service, 37(2–3). 2019, S. 113–128. Online: <https://doi.org/10.1080/15228835.2019.1614513>. [abgerufen am 18.12.2023].
- Schrödter, M.; Bastian, P.; Taylor, B.: Risikodiagnostik und Big Data Analytics in der Sozialen Arbeit. In N. Kutscher, T. Ley, U. Seelmeyer, F. Siller, A. Tillmann & I. Zorn (Hrsg.), Handbuch Soziale Arbeit und Digitalisierung. Weinheim/Basel 2020, S. 255–264.
- Simanowski, R.: Todesalgorithmus. Das Dilemma der künstlichen Intelligenz. Wien 2020.
- Simanowski, R.: Digitale Revolution und Bildung. Für eine zukunftsfähige Medienkompetenz. Weinheim/Basel 2021.
- Steiner, O.; Tschopp, D.: Künstliche Intelligenz in der Sozialen Arbeit. Sozial Extra 46. 2022, S. 466–471. Online: <https://doi.org/10.1007/s12054-022-00546-4> [abgerufen am 18.12.2023].
- West, S.M.; Whittaker, M.; Crawford, K.: Discriminating Systems: Gender, Race and Power in AI. AI Now Institute. 2019. Online: <https://ainowinstitute.org/discriminatingystems.html>. [abgerufen am 18.12.2023].
- Zuboff, S.: Das Zeitalter des Überwachungskapitalismus. Frankfurt/ New York 2018.

Prof. Dr. Anastasia Paschalidou

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
anastasia.paschalidou@fb4.fra-uas.de

Lebenslauf in Kürze

Anastasia Paschalidou ist Professorin für Außerschulische Jugendbildung, Politische Bildung, Jugendsozialarbeit und Jugendhilfe und Schule im Fachbereich Soziale Arbeit und Gesundheit an der Frankfurt University of Applied Sciences. Sie hat in Philosophie promoviert. In ihrer interdisziplinär angelegten Dissertation „Virtuelle Realität als existentielle Erfahrung. Ein philosophischer Versuch“, setzte sie sich mit medientheoretischen Konzepten und existentiellen Dimensionen virtueller Realitäten auseinander. Als Diplom-Pädagogin mit den Schwerpunkten Sozialpädagogik und Soziale Arbeit arbeitete sie über 20 Jahre in verschiedenen Arbeitsfeldern der Kinder- und Jugendhilfe. Zu Paschalidou`s Forschungsinteressen gehören diskriminierungskritische Bildungsstrategien, kritische Medienbildung, Umgang mit virtuellen Realitäten und künstlicher Intelligenz.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Mich faszinieren weiterführende Perspektiven im Umgang mit künstlicher Intelligenz. Der Geltungsbereich sowie die Grenzen der Nutzung von KI-Systemen interessieren mich. Meiner Ansicht nach sollten wir KI-Systeme nicht nur in eine „Welt der Technik“ verorten und dadurch nur einer bestimmten Wissenschaft überlassen. KI-Systeme verändern unsere Lebenswelt und eröffnen Handlungsspielräume. Nur durch einen verantwortlichen Umgang mit KI können wir auch die Potentiale in der Sozialen Arbeit erforschen und ausschöpfen. Ich setze mich damit auseinander, wie sich die Praxis der Sozialen Arbeit dadurch verändert und noch verändern wird. Hier gilt es für mich mitzugestalten und die Rechte unserer Adressat:innen zu stärken. Wir müssen entscheiden, wie wir mit KI leben wollen und dies muss interdisziplinär erforscht werden.

Annalies Baumeister, Olga Kozlova, Elizaveta Gardó, Patrizia Tolle

Nutzer*innenakzeptanz als Erfolgsfaktor und Risiko bei KI-Anwendungen im Bereich Healthcare

Künstliche Intelligenz (KI) gewinnt im Bereich Healthcare immer mehr an Bedeutung. Laut Pfannstiel ermöglicht der stetige Datenzuwachs im Gesundheitswesen einen zielgerichteten Einsatz von KI, der benutzer*innenfreundlich und flexibel ist (vgl. Pfannstiel 2022). Dabei kann es sich um KI-Anwendungen handeln, die „unsichtbar“ und weitestgehend unbemerkt bestimmte Funktionen übernehmen, darunter Diagnose-Tools oder Software zur Auswertung bildgebender Verfahren (z. B. CT oder MRT) (vgl. [isi.fraunhofer.de/de/blog/2023/kuenstliche-intelligenz-im-gesundheits-bereich.html](https://www.isi.fraunhofer.de/de/blog/2023/kuenstliche-intelligenz-im-gesundheits-bereich.html)), aber auch um „sichtbare“ Hilfen, bei denen den Nutzer*innen bewusst ist, dass sie es mit KI zu tun haben, wie z. B. Smart Speaker mit Spracherkennung (vgl. Wang, Preiniger 2019). Nach der Einschätzung von Panesar ist es zwar unwahrscheinlich, dass KI-Anwendungen Ärzt*innen und Pflegefachkräfte irgendwann gänzlich ersetzen werden, dennoch wird es zukünftig mehr smarte Technologien im Gesundheitswesen geben. Darunter auch autonome oder teilautonome Roboter zur Unterstützung von alten und behinderten Menschen (vgl. Panesar 2019; siehe auch Shaheen 2021).

Nutzer*innenakzeptanz als entscheidender Faktor für einen erfolgreichen Einsatz von KI-Anwendungen im Gesundheitswesen

Laut Arnold et al. ist die Zustimmung für KI-Anwendungen in den Bereichen Gesundheit und Mobilität überwiegend hoch – anders als in anderen Bereichen, wie Sicherheit oder Wirtschaft (vgl. Arnold et al. 2020). Dabei wird die Akzeptanz der Nutzer*innen von den folgenden Faktoren beeinflusst: Die wahrgenommene Nützlichkeit und die wahrgenommene Benutzer*innenfreundlichkeit der Technologie. Vertrauen in die Technologieanbieter*innen und die Technologie selbst. Erwartungen, was die Technologie leisten kann und wie anstrengend ihre Bedienung ist. Weitere

Faktoren sind soziale Einflüsse, Einstellung zur Technologie und erleichternde Bedingungen (vgl. Kelly et al. 2022; siehe auch Choung et al. 2022; Chew/ Achananurparp 2022). Entscheidend für die Akzeptanz von KI-Anwendungen ist dabei ein erkennbarer Nutzen. Nicht nur für den*die Einzelne*n, sondern auch für die Wirtschaft und die Gesellschaft als Ganzes. Dabei werden Nutzen und mögliche Risiken miteinander abgewogen. Weiterhin betonen Arnold et al. die Bedeutung des Faktors Kontrolle. Es sei wichtig, dass der Mensch in Kontrolle bleibt und die „Letztentscheidung“ trifft (vgl. Arnold et al. 2020). Im Bereich der Pflege werden KI-Anwendungen, u. a. auch der Einsatz von Robotern, als positiv betrachtet, so lange wie sie Pflegekräfte unterstützen, entlasten und ihnen mehr Zeit für zwischenmenschliche Kontakte ermöglichen. Dagegen sinkt die Akzeptanz deutlich, wenn droht, dass Menschen durch die KI ersetzt werden. Unterstützt KI bei Diagnosen oder der Ausgestaltung von Therapie, wird sich eine höhere Sicherheit versprochen. Dennoch wird KI-Anwendungen nicht vollumfänglich vertraut, weshalb die endgültige Entscheidung beim Menschen, z. B. bei den Ärzt*innen, verbleiben sollte, damit sie akzeptiert und letztlich auch sinnvoll eingesetzt werden (vgl. ebenda; siehe auch Kelly et al. 2022).

Nutzer*innenakzeptanz einer KI-basierten Sondersteuerung für einen Roboterarm

Im BMBF geförderten Projekt DoF-Adaptiv (Adaptive Freiheitsgradeinbettung als kooperatives Userinterface für einen Assistenzroboter) wird KI und maschinelles Lernen eingesetzt, um eine adaptive KI-basierte Sondersteuerung für assistive Roboterarme zu entwickeln. Das Projekt richtet sich an Menschen mit schweren körperlichen Beeinträchtigungen, die mit Funktionsverlusten in den Armen und Händen einhergehen. Dazu gehören Menschen mit einer hohen Quer-

schnittlähmung (Tetraplegie) oder degenerative Muskelerkrankungen wie Multiple Sklerose oder Spinaler Muskelatrophie Typ 1 und 2. Personen aus diesen Gruppen benötigen in der Regel bis zu 24/7 pflegende Angehörige, Assistenz- oder Pflegekräfte zur Unterstützung in ihrem Alltag. Assistive Roboterarme, wie der Jaco von Kinova, iArm von Assistive Innovations oder Bateo von Exxomove (alle drei Roboterarme sind auf dem deutschen Markt etabliert und verfügen über eine Hilfsmittelnummer), sollen ihren Nutzer*innen zu mehr Selbstständigkeit verhelfen und damit auch ihre

die Steuerung (vgl. Goldau et al. 2023). Das Projekt verfolgt bei der Entwicklung einen partizipativen Ansatz mit dem Ziel Forschung und Entwicklung an den Bedürfnissen der zukünftigen Nutzer*innen aber auch deren pflegenden Angehörigen, Assistenz- und Pflegekräfte auszurichten. Es stellt sich die Frage, welche Akzeptanz eine von KI unterstützte Sondersteuerung im Bereich der ambulanten Pflege findet und welche Anforderungen sie aus Sicht der Nutzer*innen und deren pflegenden Angehörigen, Assistenz- oder Pflegekräften gerecht werden sollte.

DoF-Adaptiv – Adaptive Freiheitsgradeinbettung als kooperatives Userinterface für einen Assistenzroboter

Bekanntmachung BMBF	Adaptive Technologien für die Gesellschaft – Intelligentes Zusammenwirken von Mensch und KI (MeKI)
Forschungsförderer	VDI/VDE Innovation + Technik GmbH
Fördernummer	16VS8564
Verbundkoordinator	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Bremen
Projektleitung Frankfurt UAS	Prof. Dr. Barbara Klein, Prof. Dr. Elizaveta Gardó, Prof. Dr. Patrizia Tolle (Stellv.)
Weitere Partner	TU-Darmstadt, Munevo GmbH, München
Laufzeit	02/2021 bis 01/2024
Website	www.interaktive-technologien.de/projekte/dof-adaptiv

Tabelle 1: Eigene Darstellung. Eckdaten zu DoF-Adaptiv.

Autonomie erhöhen. An einen E-Rollstuhl angebracht, werden die Roboterarme manuell über die gleichen Eingabegeräte gesteuert wie der E-Rollstuhl, d. h. in der Regel mit einem Joystick. Da die derzeit gängigen Steuerungskonzepte dieser Roboterarme noch mühsam und zeitraubend sind, wurde in DoF-Adaptiv an einer Lösung geforscht, wie die Bedienung eines Roboterarms vereinfacht und somit auch zugänglicher werden kann. Anhand von in Simulationen gewonnenen Trainingsdaten lernt die Software die Absicht der Nutzer*innen zu erkennen, z. B. dass auf ein Objekt gesteuert wird, um es zu greifen und „mitzudenken“. Den Nutzer*innen werden über Smartglasses Vorschläge zur Steuerung angezeigt, wie sie am effektivsten ihr Ziel erreichen. Die Nutzer*innen können diesen Vorschlägen folgen, müssen es aber nicht. So können Tätigkeiten mit dem Roboterarm schneller und mit weniger Anstrengung ausgeführt werden, aber die Nutzer*innen behalten jederzeit die Kontrolle über

Zu diesem Zweck wurde zu Beginn des Projekts gemeinsam mit zukünftigen Nutzer*innen Anwendungsszenarien für einen Roboterarm mit einer KI-basierten Sondersteuerung entwickelt. An der Ausarbeitung hatten sich sieben Menschen mit schweren körperlichen Beeinträchtigungen beteiligt, zwei pflegende Angehörige, zwei Physiotherapeut*innen und Vertreter*innen aller Projektpartner. In Workshops, Interviews und Fokusgruppen entstanden die vier Szenarien „Essen und Trinken“, „Türen öffnen und schließen“, „Supermarktregal / Objekte vom Boden aufheben“ und „Mikrowelle“ (vgl. Goldau et al. 2023). Weiterhin sollten aus Sicht der zukünftigen Nutzer*innen und deren Angehörige die folgenden Kriterien erfüllt sein, damit der Einsatz des Roboterarms mit der KI-basierten Sondersteuerung eine sinnvolle Unterstützung in ihrem Alltag ist:

| Die Nutzer*innen sollten spürbar mehr Autonomie und Freiheit im Alltag gewinnen.

| Der Roboterarm sollte für die Nutzer*innen ohne Unterstützung anderer nutzbar/ zu steuern sein. Dazu gehört, dass sie alltägliche Tätigkeiten direkt und allein durchführen können – ohne auf die Unterstützung oder den Dienstbeginn (bestimmter) Personen warten zu müssen. Vorbereitende Handlungen durch pflegende Personen, z. B. Essen kleinschneiden, werden aber akzeptiert.

| Der Einsatz von KI sollte ermöglichen, dass der Roboterarm zuverlässig und sicher genutzt werden kann und keine „Aufsichtsperson“ zugegen sein muss, damit die Nutzer*innen wieder Zeit für sich allein gewinnen können.

| Gleichzeitig sollte sich der Organisationsaufwand der Pflege verringern und die pflegenden Personen Entlastung erfahren.

Weiterhin wurde sich im Projekt auch aus der Perspektive des Risikomanagements mit dem Faktor „Nutzer*innenakzeptanz“ auseinandergesetzt. So wurde untersucht, welche Anforderungen die adaptive Sondersteuerung in Bezug auf Risikomanagement erfüllen muss und inwiefern die Akzeptanz (oder Ablehnung) der primären Nutzer*innen, als auch der pflegenden Angehörigen, Assistenz- und Pflegekräfte als sogenannte Sekundärnutzende das Risikomanagement beeinflussen.

Aktueller Rechtsrahmen

Die Integration von KI-Anwendungen im Gesundheitswesen verspricht bedeutende Fortschritte bei der Diagnose, Behandlung und Patientenversorgung. Allerdings gehen damit auch eine Vielzahl von Herausforderungen und Risiken einher, die sorgfältig analysiert und bewertet werden müssen. Obwohl viele KI-Medizinprodukte bereits zugelassen sind, bleiben viele regulatorische Fragen (noch) ungeklärt. Derzeit gibt es keine Gesetze und harmonisierte Normen, die speziell den Einsatz von KI in Medizinprodukten regulieren (Reinsch 2023). Noch fehlt es an klaren Richtlinien und etablierten Standards, um die Sicherheit und Wirksamkeit von KI-Systemen in der medizinischen Praxis zu gewährleisten. Um dieser Problematik zu begegnen, verständigten sich am 8.12.2023 Unterhändler*innen des Europaparlaments und der EU-Staaten in Brüssel nach langen Verhandlungen auf entsprechende Regeln. Nach Angaben des EU-Parla-

ments handelt es sich um das weltweit erste KI-Gesetz. EU-Binnenmarktkommissar Thierry Breton sprach von einer „historischen« Einigung (Spiegel 2023). KI-Systeme als Medizinprodukte werden dabei in der EU in die Risikokategorie „Hoch“ fallen. Zusätzlich zur EU Medical Devices Regulatory (MDR), sollen zukünftig diese Verordnungen in Form von Regulatorien für die Entwicklung von KI gelten.

Risikomanagement-Workshop in DOF-Adaptiv

Im Rahmen des Projektes DOF-Adaptiv wurde ein Risikomanagement-Workshop durchgeführt. Obwohl sich die Produktentwicklung noch in einem frühen Stadium befindet, wurden für den im Projekt entwickelten Roboterarm-Demonstrator sämtliche Risikomanagement-Aktivitäten vollständig umgesetzt, um Erkenntnisse für die spätere Produktentwicklung zu gewinnen. Darüber hinaus lag besonderes Augenmerk auf der Untersuchung, wie die Integration von KI in das Produkt die Risikobewertung beeinflusst. Insbesondere aus wissenschaftlicher Perspektive war es entscheidend, die Herausforderungen zu identifizieren und zu klären, welche Unterschiede sich im Risikomanagement ergeben, bedingt durch die Tatsache, dass es sich bei dem Produkt um ein KI-basiertes Medizinprodukt handelt.

Der Risikomanagement-Workshop gliederte sich in zwei Teile. Im ersten Teil wurde ein Risikomanagementprozess durchgeführt, der sich an bestehende Verordnungen orientierte. Dabei wurden die grundlegenden Anforderungen der EU-Richtlinien, Verordnungen und Normen berücksichtigt, darunter die Regulation EU 2017/745 (MDR), DIN EN ISO 14971:2019 und DIN EN ISO 13485:2021. Der Roboterarm-Demonstrator, ein KI-Medizinprodukt, wurde gemäß der Verordnung und den harmonisierten Normen dabei wie Software behandelt und musste die entsprechenden Anforderungen erfüllen. Dieser Teil des Workshops ergab, dass das Gesamtrisiko des Roboterarm-Demonstrators als „akzeptabel“ eingestuft wurde. Es wurden keine Risiken identifiziert, die den Nutzen des Produkts überwiegen würden.

Im zweiten Teil des Workshops lag der Fokus auf der Frage, welche Risiken speziell aus der Verwendung von KI resultieren und ob diese Risiken auch für den Roboterarm-Demonstrator relevant sind. Dabei wur-

den verschiedene KI-spezifische Risiken identifiziert, darunter beispielsweise mangelnde Akzeptanz, falsches Lernen, Probleme bezüglich Datenqualität und Datenschutz, Intransparenz dynamischer KI-Produkte, Ethik sowie Haftungs- und rechtliche Fragen. Für den Roboterarm-Demonstrator, der im abgeschlossenen Bereich trainiert wurde und im Feld nicht weiter lernen kann, wurde festgestellt, dass viele spezifische KI-Risiken entfallen. Allerdings wurde unter anderem die Nutzer*innenakzeptanz als ein wichtiges Thema angesprochen, das unabhängig davon, ob es sich um statische oder dynamische KI-Produkte handelt, eine besondere Rolle im Risikomanagement von KI-Produkten spielt. Diese Erkenntnisse bilden die Grundlage für weitere Analysen. Im Folgenden wird daher zunächst über die Unterschiede gesprochen, die sich daraus ergeben, ob es sich um ein statisches oder dynamisches KI-Produkt handelt und anschließend wird das Akzeptanzproblem als spezifisches KI-Medizinprodukt-Risiko analysiert.

Risiken und Zertifizierbarkeit von dynamischer KI im Vergleich zur statischen KI

Im Workshop wurden spezifische Risiken im Zusammenhang mit dynamischer KI identifiziert. Diese Risiken betreffen vor allem die Unvorhersehbarkeit von Verhaltensweisen, die mangelnde Erklärbarkeit von Entscheidungen, die kontinuierliche Anpassung an veränderte Umgebungen und die undurchsichtige Datenverarbeitung.

Im Allgemeinen besteht Konsens in der regulatorischen Gemeinschaft, dass KI, die bereits alle erforderlichen Kenntnisse für die Erfüllung ihrer spezifischen Zweckbestimmung erlangt haben, grundsätzlich gemäß der Verordnung 2017/745 zertifizierbar sind. Diese Erkenntnis wird unter anderem auch durch den Fragenkatalog „Künstliche Intelligenz bei Medizinprodukten“ der Interessengemeinschaft der Benannten Stellen für Medizinprodukte in Deutschland gestützt (vgl. Fragenkatalog „Künstliche Intelligenz bei Medizinprodukten“ 2023). Die gleiche Schlussfolgerung wurde auch in DoF-Adaptiv gezogen. In diesem Zusammenhang wird KI gemäß der Verordnung und harmonisierter Normen letztlich als Software betrachtet und muss daher die entsprechenden Anforderungen erfüllen.

Die Zertifizierung von dynamischer KI stellt eine grundlegende Herausforderung dar, da das System kontinuierlich im Betrieb lernt und sich weiterentwickelt. Im Gegensatz zur Zertifizierung statischer KI unterliegt dynamische KI keiner klaren Zertifizierbarkeit gemäß der Verordnung 2017/745, da die Verifizierung und Validierung gewährleistet sein muss. Unter anderem muss die Funktionsfähigkeit anhand des Verwendungszwecks validiert sein (vgl. Fragenkatalog „Künstliche Intelligenz bei Medizinprodukten“). Die Anforderungen an die Zertifizierung stehen im Widerspruch zur Veränderlichkeit der dynamischen KI, insbesondere in Bezug auf ihre spezifische Zweckbestimmung. Die fortlaufende Validierung und Überprüfung im Hinblick auf den Verwendungszweck, wie von der Medical Device Regulation (MDR) und relevanten Softwarenormen (IEC 62304) gefordert, gestalten sich daher als anspruchsvoll. Es entsteht die Situation, dass eine Software kontinuierlichen Veränderungen unterliegt und folglich fortlaufend validiert und bezüglich ihrer Zweckbestimmung überprüft werden muss. Im DoF-Adaptiv Risikomanagement-Workshop wurde die Frage erörtert, ob eine Lösung für Risiken, die sich aus der Veränderbarkeit des Produkts ergeben, darin besteht, die KI im abgeschlossenen Bereich spezifisch für die jeweiligen Aufgaben zu trainieren und im Feld keine weiteren Lernprozesse zuzulassen. Es wurde schließlich festgestellt, dass, obwohl dies das Risikomanagement erleichtern kann, die Gefahr besteht, wichtige Chancen zu verpassen, die sich aus der Veränderbarkeit des KI-Produkts ergeben. Im Leitfaden zur KI bei Medizinprodukten von der Johner Institut GmbH wird in diesem Kontext betont, dass, falls sich Hersteller*innen für ein dynamisches KI-Produkt entscheiden, quantitativ dargelegt werden muss, warum das Nutzen-Risiko-Verhältnis im Vergleich zu nicht kontinuierlich lernenden Systemen besser ist (vgl. Leitfaden 2023).

Im Zusammenhang mit dieser Thematik ergeben sich verschiedene Fragestellungen und Herausforderungen. Eine der Herausforderungen besteht in der möglichen Notwendigkeit der Neudefinition des Risikobegriffs spezifisch im Kontext von KI-basierten Medizinprodukten. Es gibt in der Literatur sehr viele Definitionen, wenn es um den Begriff Risiko geht. Im Risikomanagement Prozess zur Freigabe von Medizintechnik gemäß

MDR Annex-1 wird das Risiko als die Kombination aus der Wahrscheinlichkeit des Auftretens eines Schadens und der Schwere dieses Schadens definiert. Um eine präzisere und umfassendere Analyse der potenziellen Gefahren zu ermöglichen, ist es entscheidend, dass die Definition des Risikobegriffs an die Besonderheiten von KI-basierten Medizinprodukten überprüft und möglicherweise angepasst wird. Durch klare und einheitliche Definitionen wird die Kommunikation zwischen Hersteller*innen, Regulierungsbehörden und anderen Akteuren erleichtert, was zu einer effektiveren und transparenteren Risikobewertung beiträgt.

Akzeptanzproblem als spezifisches KI-Medizinprodukt-Risiko

Nutzer*innenakzeptanz spielt im Risikomanagement von medizinischen Produkten eine wichtige Rolle, da die Zustimmung und Akzeptanz seitens der Anwendenden erheblichen Einfluss auf die Sicherheit und Effektivität eines Produkts haben können. Usability-Probleme gemäß den Vorgaben der IEC 62366 müssen im Risikomanagement sorgfältig berücksichtigt werden. Besonders relevant erwies sich das Akzeptanzproblem bei KI-Medizinprodukten, da diese neu und intransparent für die Nutzer*innen sind. Daher muss die Nutzer*innenakzeptanz als integraler Bestandteil des Risikomanagements bei KI-Medizinprodukten betrachtet werden.

Die Johner Institut GmbH hat einen Leitfaden zur Anwendung von KI bei Medizinprodukten entwickelt. Dieser hat das Ziel, Hersteller*innen von Medizinprodukten eine praxisnahe Handlungsanleitung sowie eine konkrete Checkliste an die Hand zu geben, um das Fehlen einer harmonisierten Norm bestmöglich zu kompensieren. Der Leitfaden beschreibt verschiedene Anforderungen, die bei der Validierung der Benutzer*innenfreundlichkeit (Usability) beachtet werden sollten, und trägt somit dazu bei, die Nutzer*innenakzeptanz im Risikomanagement von KI-Medizinprodukten systematisch zu berücksichtigen (vgl. Leitfaden 2023).

Der Fragenkatalog „Künstliche Intelligenz bei Medizinprodukten“ der Interessengemeinschaft der benannten Stellen für Medizinprodukte in Deutschland hebt dabei insbesondere drei Punkte in Bezug auf die Usa-

bility-Validierung von KI-Medizinprodukten basierend auf IEC 62366-1 hervor (vgl. Fragenkatalog 2023):

1. Bewerten Hersteller*innen im Rahmen der Usability-Validierung, ob die Nutzer*innen die Gebrauchsanweisung verstehen?
2. Bewerten Hersteller*innen im Rahmen der Usability-Validierung, ob die Nutzer*innen dem Produkt blind vertrauen oder die Ergebnisse nachprüfen?
3. Bewerten Hersteller*innen im Rahmen der Usability-Validierung, ob die Nutzer*innen die Ergebnisse korrekt erkennen und verstehen?

Diese Fragen zielen darauf ab, sicherzustellen, dass die Usability von KI-Medizinprodukten gründlich überprüft wird, insbesondere im Hinblick auf die Verständlichkeit der Gebrauchsanweisung, das Vertrauen der Nutzer*innen in das Produkt sowie die korrekte Wahrnehmung und Interpretation der Ergebnisse. Der Leitfaden und der Fragenkatalog bieten somit wichtige Instrumente für eine effektive und praxisorientierte Umsetzung von Usability-Validierungen im Kontext KI in der Medizinproduktentwicklung.

Im Rahmen des Risikomanagement-Workshops in DoF-Adaptiv-Projekt wurde die Frage der Nutzer*innenakzeptanz diskutiert: Könnte beispielsweise die Verwendung des Begriffs „Programm“ anstelle von „KI“ die Akzeptanz verbessern? Es wurde jedoch betont, dass eine klare Benennung notwendig ist, um langfristig die Akzeptanz zu fördern und KI zur selbstverständlichen Realität werden zu lassen. Vielmehr geht es darum, dass in den Begriffen Klarheit herrscht und deutlich wird, was sich hinter dem Begriff KI verbirgt, welche Arten von KI existieren und was dies in jedem konkreten Fall bedeutet. Häufig werden die Begriffe KI, Machine Learning und Deep Learning unpräzise oder sogar synonym verwendet. Künstliche Intelligenz umfasst zahlreiche Verfahren, von denen Machine Learning nur einen Teil nutzt. Die neuronalen Netze (und somit Deep Learning) sind Teil des Machine Learning. Die Annahme, dass künstliche Intelligenz in der Medizin hauptsächlich auf neuronale Netzwerke zurückgreift, ist nicht korrekt. Generell gehören momentan die meisten angewandten Verfahren der künstlichen Intelligenz bei Medizinprodukten zur Kategorie „Machine Learning“, wobei neuronale Netz-

werke in der Anwendung an zweiter Stelle stehen (vgl. Reinsch 2023). Klarheit in Begriffen und das Verständnis dessen, was sich hinter diesen Begriffen verbirgt, können einen bedeutenden Beitrag leisten, um die Nutzer*innenakzeptanz von KI-Produkten zu fördern.

Evaluation der adaptiven KI-basierten Sondersteuerung mit zukünftigen Nutzer*innen

Im Sommer 2023 fanden vier Online-Workshops mit zukünftigen Nutzer*innen statt, in denen ethische und soziale Implikationen der in DoF-Adaptiv entwickelten KI-basierten Sondersteuerung diskutiert wurden. Insgesamt nahmen 13 Menschen mit schweren körperlichen Beeinträchtigungen teil, davon sechs Nutzer*innen assistiver Roboterarme. Außerdem ein*e pflegende*r Angehörige*r, zwei Vertreter*innen aus der Industrie und Vertreter*innen der Projektpartner*innen. Die Workshops werden derzeit qualitativ ausgewertet.

Erste Ergebnisse bestätigen, dass die Akzeptanz von KI vor allem davon abhängt, wie groß der wahrgenommene Nutzen ist und ob letztendlich die Kontrolle bei den Nutzer*innen bleibt (siehe dazu Arnold et al. 2020). So wurde in allen vier Workshops begrüßt, dass die adaptive Sondersteuerung die Bedienung eines Roboterarms erleichtert und sich für Tätigkeiten wie Essen und Trinken der Zeitaufwand verringert. Gleichzeitig war es den Teilnehmer*innen wichtig, dass trotz der Situationserkennung, der Roboterarm keine autonomen Handlungen ausübt, sondern die Kontrolle über die Steuerung bei den Nutzer*innen verbleibt. Auch, damit sie sich gegen die Vorschläge der KI entscheiden können, wenn sie den Eindruck haben, dass der KI eine Fehleinschätzung unterlaufen ist (z. B. ein Hindernis auf dem Fahrtweg des Roboterarms nicht erfasst wurde und mit der vorgeschlagenen Fahrtrichtung eine Kollision droht). „Also mir wäre nochmal wichtig zu betonen, dass ich ja bei aller KI und allem, was theoretisch vorgeschlagen werden könnte, bin ich ja noch Herr meiner Sinne und kann auch eine Aktion abrechnen oder einen Vorschlag ablehnen, wenn ich merke, dass es so nicht funktioniert“ (Teilnehmer*in aus Workshop 3).

Während der vier Workshops zeigte sich insgesamt eine hohe Akzeptanz für die im Projekt entwickelte

KI-basierte Sondersteuerung. Ein*e Teilnehmer*in sagte dazu: „Im Prinzip ist es ja so, dass wir als behinderte Menschen oftmals einen anderen Zugang und ein anderes Verständnis für Technik haben, weil wir merken, wie es uns hilft. Die Probleme bestehen dann meist mit dem Umfeld“ (Teilnehmer*in aus Workshop 2). Ein*e andere*r Teilnehmer*in betonte: „Auf jeden Fall würde man mit einer besseren Steuerung mehr machen. Ich wüsste nicht, wo das negative Punkte haben sollte“ (Teilnehmer*in aus Workshop 4). Und weiter kommentierte ein*e Teilnehmer*in zum Punkt Akzeptanz: „Wenn mit diesen runderen Bewegungen und besseren Bewegungsabläufen, die dadurch geschaffen werden können, das Ganze weiter nach vorne getrieben wird, dann steigt auch die Akzeptanz“ (Teilnehmer*in aus Workshop 4). Als akzeptanzfördernde Faktoren wurden vor allem ein Gewinn an Selbstständigkeit und Autonomie genannt. So sagte ein*e Teilnehmer*in: „Ich denke, dass es ja die Autonomie steigert, also man kann ja Dinge wieder selber machen und das auch in einer angemessenen Geschwindigkeit“ (Teilnehmer*in aus Workshop 4). Insbesondere in Bezug auf das Anwendungsszenario „Essen und Trinken“ wurde angemerkt: „Ich finde vor allem, die Möglichkeit zu haben, in seinem eigenen Tempo zu essen oder auch zu trinken, wann man möchte und nicht immer jemanden um Hilfe bitten muss, ist für mich zum Beispiel ein wichtiger Punkt“ (Teilnehmer*in aus Workshop 1). Außerdem wurde ein Zugewinn an Selbstständigkeit mit besseren Teilhabechancen, dem Wunsch nach Zugehörigkeit und einem höheren Selbstwertgefühl verknüpft. Ein*e Teilnehmer*in betonte wie wichtig es sei „... dass ich auch mal alleine raus in die Stadt kann, in ein Café gehen kann, weil ich dann keine Hilfe beim Trinken brauche. Dass ich einfach mal alleine in die Stadt gehen kann“ (Teilnehmer*in aus Workshop 3). Ein anderes Beispiel für Teilhabe und Zugehörigkeit wurde im 4. Workshop genannt: „Was mir gerade noch eingefallen ist, es gibt doch auch so Stadtreinigungsaktionen, wo die Menschen in der Stadt rumlaufen und den Müll vom Boden aufsammeln. Und da kann man ja dann auch mitmachen“ (Teilnehmer*in aus Workshop 4). Alltägliche Tätigkeiten (wieder) selbst ausführen zu können, ob für einen selbst (z. B. selbstständig trinken) oder um sich ehrenamtlich zu engagieren (z. B. Müll aufsammeln) wirkt sich laut der

Workshopteilnehmer*innen signifikant auf die Selbst- und Fremdwahrnehmung und das eigene Selbstbewusstsein aus: „(...) dann ist es etwas, was ich eigenständig tun kann und ja, dann natürlich auch irgendwo Selbstbewusstsein, Selbstwert einfach ein bisschen steigert oder eine Sache ist, wo man stolz darauf ist, dass man das selber machen kann“ (Teilnehmer*in aus Workshop 2).

Datenschutz spielte dagegen für die wenigsten Teilnehmer*innen eine Rolle, u. a. da in DoF-Adaptiv mit einer lokalen Lösung gearbeitet und keine Verbindung zum Internet benötigt wird, um das neue Steuerungskonzept anwenden zu können. Auch die Kamera, die am Roboterarm angebracht zur Situationserkennung benötigt wird, wurde als unbedenklich eingestuft, da sie keine Bilder von Personen erfasst oder speichert. Stattdessen wurde die Sorge geäußert, dass pflegende Angehörige oder Assistenzkräfte sich schwertun könnten, den Einsatz eines Roboterarms mit einer KI-basierten Sondersteuerung zu akzeptieren. Die Erfahrung der Teilnehmer*innen war bisher, dass pflegende Angehörige und Assistenzkräfte häufig einer neuen Technologie nicht vertrauen und den Nutzer*innen zu wenig zutrauen, diese Technologie zu beherrschen. So schilderte eine Teilnehmer*in: „Da sehe ich viel eher die Gefahr, dass diese Überfürsorge eintritt. [Imitiert eine pflegende Person] ‚Und dann komme ich lieber und mache dir das Essen in der Mikrowelle warm, weil, da passiert was, und KI und diese ganze Technik, und ich kann dir doch helfen.‘ Da sehe ich eher das Thema, dass da die Selbstbestimmung leidet“ (Teilnehmer*in aus Workshop 2). Ein*e Teilnehmer*in ergänzte: „Genau, da sehe ich die Gefahr wirklich. Und in dieser Kombi mit, es sind Ängste vor Technik und vor allem Ängste immer noch vor KI“ (Teilnehmer*in aus Workshop 2). Daher befürchten einige Teilnehmer*innen Konflikte mit ihrem sozialen Umfeld, wenn sie in ihrem Alltag einen Roboterarm mit einer KI-basierten Sondersteuerung nutzen wollen und dadurch mehr Autonomie (zurück) gewinnen könnten. „Ich kann mir vorstellen, dass da gerade auch bei Familienangehörigen das Problem ist, dass sie auch vielleicht loslassen müssen und einen auch selber machen lassen müssen, obwohl sie

das oft dann, glaube ich, nicht wollen oder dem Roboter vielleicht nicht vertrauen“ (Teilnehmer*in aus Workshop 1).

Zusammenfassung

In zunehmendem Maße etabliert sich Künstliche Intelligenz auch im Gesundheitswesen. Dennoch ist der aktuelle Rechtsrahmen für KI-Medizinprodukte noch nicht ausreichend festgelegt. Der Risikomanagement-Workshop im Rahmen des Projekts DOF-Adaptiv hat wichtige Erkenntnisse geliefert, insbesondere in Bezug auf die spezifischen Risiken von dynamischer KI im Vergleich zu statischer KI und das Akzeptanzproblem als spezifisches Risiko von KI-Medizinprodukten. Für den Roboterarm-Demonstrator, der ausschließlich im geschlossenen Bereich trainiert wurde und keine weiteren Lernprozesse im Feld durchführt, wurde festgestellt, dass viele spezifische KI-Risiken entfallen. Grundsätzlich könnte er gemäß den Regelungen für Software-Medizinprodukte zertifizierbar sein. Diese Beobachtung unterstützt die Aussage des Fragenkatalogs ‚Künstliche Intelligenz bei Medizinprodukten‘ der Interessengemeinschaft der Benannten Stellen für Medizinprodukte in Deutschland, dass statische KI im Einklang mit den bestehenden Verordnungen prinzipiell zertifizierbar ist. Im Gegensatz dazu stellt die Zertifizierung von dynamischer KI eine grundlegende Herausforderung dar, da diese Systeme kontinuierlich lernen und sich weiterentwickeln. Eine weitere Erkenntnis ist die Frage nach der möglichen Notwendigkeit der Neudefinition des Risikobegriffs, spezifisch im Kontext von KI-basierten Medizinprodukten. Diese Frage muss weiterhin erforscht und verfolgt werden.

Das Akzeptanzproblem stellt ein spezifisches Risiko dar, da Nutzer*innenakzeptanz eine entscheidende Rolle im Risikomanagement spielt, insbesondere bei neuartigen und intransparenten KI-Systemen. Im Rahmen des Workshops zum Projekt DOF-Adaptiv wurde der Schluss gezogen, dass es von großer Bedeutung ist, den Begriff KI klar zu definieren, um Missverständnisse zu vermeiden und die Nutzer*innenakzeptanz zu fördern. Eine präzise Kommunikation zwischen Hersteller*innen, Regulierungsbehörden und anderen Akteuren ist erforderlich, um eine effektive und transparente Risikobewertung von KI-Medizinproduk-

ten zu gewährleisten. Parallel dazu benötigt es auch eine gesellschaftliche Auseinandersetzung über den Einsatz von KI im Gesundheitswesen. Zum einen um über die neue Technologie ausreichend aufzuklären, zum anderen um Risiken aber auch Chancen ihres Einsatzes zu diskutieren. Hierbei ist es wichtig alle

Menschen zu beteiligen, die von KI-Anwendungen im Gesundheitswesen direkt betroffen sind, da sich KI-Anwendungen auf deren Alltag bzw. Arbeits-, und Pflegealltag, auswirken. Dazu gehört auch, die zukünftigen Nutzer*innen und weitere Stakeholder kontinuierlich in Forschung und Entwicklung einzubinden.

Literaturangaben:

- Arnold, N.; Frieß, H.-J.; Roose, J.; Werkmann, C.: Künstliche Intelligenz in Einstellungen und Nutzung bei unterschiedlichen Milieus in Deutschland. Hrsg. Konrad-Adenauer-Stiftung e. V., Berlin 2020. Online: <https://www.kas.de/de/einzeltitel/-/content/wenn-die-ki-unser-assistent-bleiben-kann-dann-koennen-wir-viel-draus-ziehen>.
- Chew, H. S. J.; Achananuparp, P.: Perceptions and needs of artificial intelligence in health care to increase adoption: scoping review. *Journal of medical Internet research*, 24(1), e32939, 2022.
- Choung, H.; David, P.; Ross, A.: Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 1–13. 2022.
- „EU einigt sich auf Regeln für künstliche Intelligenz“. Online: HYPERLINK „<https://www.spiegel.de/netzwelt/netzpolitik/ai-act-eu-erzielt-einigung-auf-regeln-fuer-kuenstliche-intelligenz-a-54d0ff03-d275-4c3a-b1b9-c491a0a258cd>“ <https://www.spiegel.de/netzwelt/netzpolitik/ai-act-eu-erzielt-einigung-auf-regeln-fuer-kuenstliche-intelligenz-a-54d0ff03-d275-4c3a-b1b9-c491a0a258cd> [abgerufen am 11.12.2023].
- Fragenkatalog „Künstliche Intelligenz bei Medizinprodukten“. Online: [https://www.ig-nb.de/?tx_epxelo_file\[id\]=861877&cHash=5c3e1b0889ef5e017c2ff309f4ea82a8](https://www.ig-nb.de/?tx_epxelo_file[id]=861877&cHash=5c3e1b0889ef5e017c2ff309f4ea82a8). [abgerufen am 8.12.2023].
- Goldau, F.; Baumeister, A.; Shivashankar, Y.; Frese, U.; Tolle, P.: DORMADL - Dataset of human-operated Robot Arm Motion in Activities of Daily Living. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2023.
- Kelly, S.; Kaye, S. A.; Oviedo-Trespalacios, O.: What factors contribute to acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, 101925. 2022.
- Leitfaden zur KI bei Medizinprodukten von Christian Johner, Christoph Molnar u.a. Online: HYPERLINK „https://github.com/johner-institut/ai-guideline/blob/master/Guideline-AI-Medical-Devices_DE.md“ https://github.com/johner-institut/ai-guideline/blob/master/Guideline-AI-Medical-Devices_DE.md [abgerufen am 8.12.2023].
- Panesar, A.: *Machine learning and AI for healthcare*. Coventry, UK Apress 2019, S. 1-73.
- Pfannstiel, M. A.: *Künstliche Intelligenz im Gesundheitswesen*. Wiesbaden 2022.
- Reinsch, D.: „Künstliche Intelligenz in der Medizin“. 2023. Online: HYPERLINK „https://www.johner-institut.de/blog/regulatory-affairs/kuenstliche-intelligenz-in-der-medizin/?fbclid=IwAR1GCV7lZurEWuvdmiTWz57LZRYWnYSPS6OF6Wj8ISYPTxvQxMJ6_VW9Fc4“ https://www.johner-institut.de/blog/regulatory-affairs/kuenstliche-intelligenz-in-der-medizin/?fbclid=IwAR1GCV7lZurEWuvdmiTWz57LZRYWnYSPS6OF6Wj8ISYPTxvQxMJ6_VW9Fc4 [abgerufen am 8.12.2023].
- Shaheen, M. Y.: *Applications of Artificial Intelligence (AI) in healthcare: A review*. *ScienceOpen Preprints*. 2021.
- Wang, F.; Preininger, A.: AI in health: state of the art, challenges, and future directions. *Yearbook of medical informatics*, 28(01), 016-026. 2019.

Annalies Baumeister, M.Sc.

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Forschungszentrum FUTURE AGING
annalies.baumeister@fb4.fra-uas.de

Lebenslauf in Kürze

Annalies Baumeister ist seit 2017 wissenschaftliche Mitarbeiterin im Forschungszentrum FUTURE AGING. Sie arbeitete als Heilerzieherin an einer Förderschule mit dem Schwerpunkt geistige Entwicklung, bevor Sie an der Frankfurt UAS zunächst einen Bachelor of Arts in Soziale Arbeit und dann den Master of Sciences in dem interdisziplinären Studiengang Barrierefreie Systeme – Case Management (Heute Inclusive Design) erwarb.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist die Perspektive der (zukünftigen) Nutzer*innen und die partizipative Einbindung dieser Menschen in die Forschung zu KI. Weiterhin beschäftigen mich vor allem Soziale und Ethische Implikationen von KI im Bereich Healthcare.



Prof. Dr. Patrizia Tolle

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Forschungszentrum FUTURE AGING
tolle@fb4.fra-uas.de

Lebenslauf in Kürze

Patrizia Tolle ist Professorin an der Frankfurt UAS und Mitglied im Forschungszentrum FUTURE AGING. Sie ist Diplom-Behindertenpädagogin und befasst sich mit Prozessen unterstützter Entscheidungsfindung als Mittel der Teilhabe von Menschen mit Behinderungen.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist die Erweiterung von Möglichkeitsräumen.



Prof. Dr. Elizaveta Gardó

Frankfurt University of Applied Sciences
Fachbereich 3: Wirtschaft und Recht
egardo@fb3.fra-uas.de

Lebenslauf in Kürze

Elizaveta Gardó ist seit 2018 Professorin für Wirtschaftsinformatik an der Frankfurt University of Applied Sciences. Sie hat an der Philipps-Universität Marburg studiert und promoviert. Ihre Forschungsschwerpunkte sind Mensch-Maschine Interaktion, Künstliche Intelligenz und Risikomanagement.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegt insbesondere in der Erforschung der Chancen, die Künstliche Intelligenz für unsere Gesellschaft im Allgemeinen und im medizinischen Bereich im Besonderen mit sich bringt. Die adäquaten regulatorischen Rahmenbedingungen unter Berücksichtigung der KI-spezifischen Risiken sind dabei unabdingbar.

Olga Kozlova, Dipl.-Wirtsch./ Dipl.-Bild. Kunst

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
olga.kozlova@fb4.fra-uas.de

Lebenslauf in Kürze

Olga Kozlova ist seit Januar 2023 wissenschaftliche Mitarbeiterin an der Frankfurt University of Applied Sciences. Sie ist Diplom-Wirtschaftswissenschaftlerin und verfügt über prägende Berufserfahrungen im Bereich Projektmanagement.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegt insbesondere in den Chancen, die sich durch die Integration von KI in kollaborative Prozesse zwischen Menschen und KI eröffnen. In diesem Kontext befasse ich mich intensiv mit Fragen zu den Risiken, die die weitreichenden Implikationen von KI im Allgemeinen sowie speziell im Bereich des Gesundheitswesens mit sich bringen.

Barbara Lippla

Transformativ forschen mit Lern- und Experimentierräumen für KI-gestützte Assistenzsysteme in der beruflichen Rehabilitation von Menschen mit Behinderungen

Die tiefgreifenden und langfristigen Veränderungsprozesse, die mit der Digitalisierung einhergehen und seit einigen Jahren auch durch Entwicklungen auf dem Gebiet der Künstlichen Intelligenz (KI) neu befeuert werden, betreffen alle gesellschaftlichen Bereiche. So vollzieht sich derzeit auch eine digitale Transformation der Arbeitswelt, die sich nicht nur in veränderten Geschäftsmodellen und neuen Tätigkeitsfeldern, sondern auch in Veränderungen von Arbeitsprozessen und Arbeitsorganisation manifestiert. Die Digitalisierung der Arbeitswelt bringt zwar neue Exklusionsrisiken mit sich, kann aber gleichzeitig neue Teilhabechancen für Menschen mit Behinderungen eröffnen. Vor allem digitale Assistenzsysteme, die in Abgrenzung zu substituierenden Technologien als Mensch-Technik-Systeme auf die Unterstützung menschlicher Arbeit zielen und die Zieldimensionen Gesundheit, Teilhabe und Arbeitsqualität adressieren (vgl. Apt et al. 2018), bieten große Chancen für die Gestaltung einer inklusiven Arbeitswelt. „Derartige Mensch-Technik-Systeme gehen mit einer Individualisierung von Arbeitsprozessen einher, welche die persönlichen Fähigkeiten der Beschäftigten berücksichtigt und somit Ansätze für die Gestaltung eines inklusiven Arbeitsmarktes bietet, indem etwa sprachliche, kognitive oder körperliche Einschränkungen ausgeglichen bzw. besondere Fähigkeiten genutzt werden“ (Apt et al. 2018, S. 7).

Assistenzsysteme, die sich künstlicher Intelligenz bedienen, um die berufliche Teilhabe von Menschen mit Behinderungen zu unterstützen, wecken, sowohl auf Seiten von Menschen mit Behinderungen selbst (vgl. Lippla & Stock 2022) als auch bei anderen Akteuren aus dem Kreis der Leistungserbringer beruflicher Rehabilitation, der Kostenträger oder von Unternehmen auf der Suche nach Fachkräften großes Interesse (vgl. Lippla 2022). Da-

bei werden schnell hohe Erwartungen an die Technologien aufgebaut, die mit ihrer Einführung und Verbreitung zusammenhängenden Veränderungsprozesse treffen in der Praxis allerdings auf vielfältige Hemmnisse. Für den Gesundheitsbereich lässt sich zeigen, dass Hemmnisse für die Nutzung, Verbreitung, Transfer und Nachhaltigkeit von digitalen Technologien mit der Komplexität der Technologien zunehmen, wobei sich die Komplexität aus Abhängigkeiten und Wechselwirkungen unterschiedlichster individueller und organisationaler Faktoren sowie Rahmenbedingungen ergibt (vgl. Kunze 2022).

KI-gestützte Assistenzsysteme können im Sinne des Sozialgesetzbuch Neuntes Buch (SGB IX) als Hilfsmittel oder technische Arbeitshilfen begriffen werden oder sie werden im Sinne des Arbeitsschutzes zur barrierefreien Gestaltung des Arbeitsumfeldes eingesetzt. Jedoch weisen sie im Vergleich zu den herkömmlichen Hilfsmitteln einige Besonderheiten auf (vgl. Lippla 2022):

- | KI-Technologien sind datenverarbeitende Systeme. Um gut zu funktionieren, brauchen sie relevante und qualitativ hochwertige Daten. Je nach Einsatzszenario handelt es sich dabei um personenbezogene Daten, z.B. Gesundheitsdaten wie Atemfrequenz oder Herzratenvariabilität, Kommunikationsinhalte oder unternehmensbezogene Daten, z.B. über Prozessabläufe.
- | Mit ihrer Anwendung auf mobilen Endgeräten wie Smartphones oder Tablets verschwimmen die Grenzen zwischen privater und beruflicher Nutzung digitaler Hilfsmittel. Die Systeme können bei Freizeitaktivitäten ebenso wie beim Lernen oder Arbeiten unterstützen.
- | Bei KI-gestützten Assistenzsystemen handelt es sich um Software-Produkte, deren Lebenszyklen wesentlich kürzer ausfallen können als die der herkömmlichen

Hilfsmittel. Wie die Entwicklung der letzten Jahre im Bereich der sogenannten Large Language Models am Beispiel von ChatGPT gezeigt hat, können die Innovationssprünge enorm ausfallen. Die Systeme werden laufend verbessert und verändert, um Fehler zu beheben oder neue Funktionen zu implementieren und werden dann in neuen Versionen angeboten. Damit veralten softwarebasierte Hilfsmittel schneller und müssen ggf. öfter ersetzt werden.

| Die KI-gestützten Assistenzsysteme bedürfen oft einer intensiven und fortlaufenden Anpassung an neue Nutzer*innengruppen mit ihren individuellen Bedarfen, an neue Tätigkeitsfelder oder neue Arbeitsumgebungen. Und auch die Kompetenzen im Umgang mit den Technologien, sowohl bei Menschen mit Behinderungen als auch bei Fachkräften in der beruflichen Rehabilitation oder auf dem allgemeinen Arbeitsmarkt müssen mitunter intensiver geschult werden. Das geht mit einer Verschiebung von reinen Anschaffungskosten des Hilfsmittels hin zu Folgekosten einher.

Ob und unter welchen Rahmenbedingungen KI-Technologien die berufliche Teilhabe von Menschen mit unterschiedlichen Behinderungen verbessern können, lässt sich kaum aus einer distanziert analytischen Perspektive erfassen. Das „Übereinkommen über die Rechte von Menschen mit Behinderungen“ (UN-Behindertenrechtskonvention, UN-BRK) weist der Wissenschaft eine aktive Rolle bei der Förderung der Verfügbarkeit und der Nutzung neuer, die Inklusion unterstützender Technologien zu. Um die digitale Transformation der beruflichen Rehabilitation und der Arbeitswelt im Sinne der Inklusion aktiv mitzugestalten, werden Forschungsformate gebraucht, die es den Forschenden wie auch den Stakeholdern (Menschen mit Behinderungen, Leistungserbringern beruflicher Rehabilitation, Kostenträgern, Arbeitgebern, beratende Akteuren u.a.) erlauben, bedarfsorientiert Veränderungen anzustoßen, um sich KI-gestützten Assistenztechnologien systematisch, aber zunächst nur „auf Probe“ zu nähern. Dabei können Wissen für den Einsatz und die (Weiter-)Entwicklung von KI-Technologien und den organisationalen Wandel generiert und notwendige Kompetenzen erworben werden.

In dem durch das Bundesministerium für Arbeit und Soziales (BMAS) aus Mitteln des „Ausgleichsfonds für überregionale Vorhaben zur Teilhabe schwerbehinder-

ter Menschen am Arbeitsleben“ geförderten Verbundprojekt „KI.ASSIST“, das von April 2019 bis März 2022 unter der Koordination des Bundesverbandes Deutscher Berufsförderungswerke (BV BFW) zusammen mit den Bundesarbeitsgemeinschaften der Berufsbildungswerke (BAG BBW) und der Werkstätten für behinderte Menschen (BAG WfbM) sowie dem Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) durchgeführt wurde, wurde der Ansatz der betrieblichen Lern- und Experimentierräume des BMAS adaptiert, der genau dies ermöglicht hat. Das BMAS fördert solche betrieblichen Lern- und Experimentierräume seit 2017 zur Erprobung von Arbeitsinnovationen. Hierzu wurden in neun Einrichtungen der beruflichen Rehabilitation (drei Berufsförderungswerke, drei Berufsbildungswerke, drei Werkstätten für behinderte Menschen) sowie in einem Unternehmen Lern- und Experimentierräume (kurz LER) zur Erprobung von KI-gestützten Assistenzsystemen eingerichtet. Von Interesse war zum einen, wie solche Systeme Menschen mit Behinderungen beim Lernen und Arbeiten unterstützen können. Zum anderen war von Interesse, wie organisationale Transformationsprozesse angestoßen und gestaltet werden können, die mit der Einführung und Verstetigung von KI-Technologien einhergehen.

Die LER im Projekt „KI.ASSIST“ zeichneten sich insbesondere durch folgende Charakteristika und gleichzeitig Gelingensbedingungen aus (vgl. Thieke-Beneke et al. 2022):

| Praxisorientierung und Authentizität: Die Erprobung der KI-gestützten Assistenzsysteme fand bei den Praxispartnern vor Ort im realen Betrieb und nicht unter künstlich hergestellten Laborbedingungen statt. Die LER bildeten reale Arbeits- und Lernsituationen ab und setzten an konkreten Bedarfs- und Problemlagen der Einrichtung und der Menschen mit Behinderungen an.

| Lösungsorientierung und Ergebnisoffenheit: Für den Einsatz von KI-gestützten Assistenzsystemen gibt es noch keine Blaupausen im Kontext beruflicher Rehabilitation. Die Problemdefinitionen, Zielsetzungen und Lösungsvorschläge richteten sich stark an den jeweiligen Bedarfen und Gegebenheiten in den Einrichtungen der beruflichen Rehabilitation aus. Sie wurden nicht vom Forschungsteam vorgegeben, gleichwohl unterstützte

das Projektteam die Einrichtungen bei deren Erarbeitung. Die erarbeiteten Lösungen konnten entweder verfestigt oder aber auch verworfen werden.

| Partizipationsprinzip: Die Wünsche und die Expertise der betroffenen Akteure tragen dazu bei, dass relevante und akzeptierte Lösungen erarbeitet werden. Dies schließt nicht nur die Beteiligung von Menschen mit Behinderungen ein. Auch Fach- und Führungskräfte, aber auch Kostenträger und andere Stakeholder müssen mitgenommen werden, wenn die gefundenen Lösungen im System der beruflichen Reha nachhaltig verankert werden sollen.

| Fehlerkultur und rekursives Lernen: Es ist kennzeichnend für Innovationsprozesse, dass Akteure nur über vorläufiges Wissen verfügen. Zum einen kann sich die Technologie selbst verändern, z.B. wenn Funktionen hinzukommen oder wegfallen. Zum anderen lassen sich hemmende, aber auch fördernde Faktoren oft erst im praktischen Tun erkennen. Man nähert sich in einzelnen Schritten und mehreren Reflexionsschleifen der Lösung an. Mit dem Gelernten können die Lösungen dann weiterentwickelt und verbessert werden.

Für die am Projekt beteiligten Einrichtungen der beruflichen Rehabilitation, stellten die LER einen „geschützten Raum“ dar, in dem innovative Konzepte für das Lernen und Arbeiten ausprobiert werden konnten, ohne dass dazu bereits in den Organisationen Struktur- und Prozessveränderungen größeren Ausmaßes vorgenommen werden mussten (vgl. Thieke-Beneke et al. 2022).

Das Projekt „KI.ASSIST“ ermöglichte mit der Erprobung von KI-gestützten Assistenzsystemen in den Einrichtungen der Beruflichen Rehabilitation eine erste Auseinandersetzung mit Chancen und Risiken von KI für Menschen mit Behinderung für die Teilhabe am Arbeitsleben. Hierbei wurde teils literatur-, teils evidenzbasiert Wissen generiert, das in zahlreiche Ergebnisberichte und erste praktische Handlungsemp-

fehlungen eingeflossen ist (alle Ergebnisse sind unter www.ki-assist.de abrufbar).

Für die nachhaltige Verankerung von KI-Unterstützung im Alltag der beruflichen Rehabilitation, für den Transfer in andere Bildungs- und Arbeitsbereiche einer Einrichtung sowie zu anderen Leistungserbringern erwies sich die dreijährige Projektlaufzeit als zu kurz. Auch mittelfristige Effekte des Einsatzes KI-gestützter Assistenzsysteme wie Rückkehr bzw. Übergang von der Rehabilitations-Einrichtung auf den allgemeinen Arbeitsmarkt ließen sich in der Projektlaufzeit von „KI.ASSIST“ nicht näher betrachten. Hierzu bedarf es zudem einer stärkeren Einbeziehung von Unternehmen als Praxispartner. Diese und weitere offene Fragestellungen werden in dem seit Oktober 2022 (und noch bis September 2027) laufenden Projekt „KI-Kompass Inklusiv“, das wie das Vorgängerprojekt „KI.ASSIST“ aus Mitteln des Ausgleichsfonds des BMAS gefördert wird, weiterverfolgt. In dem Projekt, das im Verbund der vier weiter oben genannten Projektpartner aus „KI.ASSIST“ unter Koordination des DFKI durchgeführt wird, wird ein Kompetenzzentrum für KI und Inklusion in der Arbeitswelt aufgebaut, das neben aktuellen Informationen zur Verfügbarkeit KI-gestützter Assistenzsysteme, unterschiedliche Schulungs- und Beratungsangebote für unterschiedliche Zielgruppen bereithält. Daneben spielen aufbauend auf den Erfahrungen aus den Lern- und Experimentierräumen auch transformative Forschungsformate in Form von Praxislaboren mit Arbeitgebern, mit Technologie-Forschenden und Technologie-Herstellern sowie mit Kostenträgern und beratenden Akteuren eine zentrale Rolle. Hier können KI-Technologien im Kontext des allgemeinen Arbeitsmarktes erprobt, Konzepte zur partizipativen Technikentwicklung und -anpassung sowie Empfehlungen zur Weiterentwicklung von Rahmenbedingungen, Strukturen und Prozessen für den Einsatz von KI-Technologien für die berufliche Teilhabe von Menschen mit Behinderungen erarbeitet werden.

Literaturangaben:

Apt, W.; Bovenschulte, M.; Priesack, K.; Weiß, C.; Hartmann, A.: Einsatz von digitalen Assistenzsystemen im Betrieb. Expertise im Auftrag des Bundesministeriums für Arbeit und Soziales (Forschungsbericht 502). Institut für Innovation und Technik. Berlin 2018.

Kunze, C.: Digitale Transformation und Teilhabe. Chancen und Herausforderungen aus Perspektive der Technikgestaltung. In: Teilhabe. Die Fachzeitschrift der Lebenshilfe 2/2022, Jg. 61. 2022, S. 62-66.

Lippa, B.: Inklusive Arbeitswelt mit Künstlicher Intelligenz. Impulse aus der projektbegleitenden Arbeitsgruppe. Ergebnisbericht des Projekts KI.ASSIST. Bundesverband Deutscher Berufsförderungswerke e. V. Berlin. 2022.

Lippa, B.; Stock, J.: Selbstbestimmte Teilhabe am Arbeitsleben durch KI-gestützte Assistenztechnologien? Überlegungen und Erfahrungen aus dem Projekt KI.ASSIST. Ergebnisbericht des Projekts KI.ASSIST. Bundesverband Deutscher Berufsförderungswerke e. V. Berlin. 2022.

Thieke-Beneke, M.; Stock, J.; Lippa, B.; Biedermann, J.; Stähler, L.; Feichtenbeiner, R.: Die KI.ASSIST Lern- und Experimentierräume zur Erprobung KI-gestützter Assistenztechnologien. Von der Konzeption bis zur Umsetzung. Ergebnisbericht des Projekts KI.ASSIST. Bundesverband Deutscher Berufsförderungswerke e. V. Berlin. 2022.

Barbara Lippa

Bundesverband Deutscher Berufsförderungswerke e. V.
b.lippa@bv-bfw.de

Lebenslauf in Kürze

Barbara Lippa studierte Kommunikationswissenschaft und Soziologie an der Philipps-Universität Marburg und der Technischen Universität Berlin. Seit 2006 forscht und publiziert sie mit dem Schwerpunkt digitale Partizipation und Teilhabe, u.a. in den Themenfeldern E-Partizipation, Open Government, Open Data und digitale Teilhabe älterer Menschen. Seit 2019 befasst sie sich als Senior Referentin und Projektleiterin im Bereich Forschung & Innovation des Bundesverbandes Deutscher Berufsförderungswerke e. V. mit den Möglichkeiten und Rahmenbedingungen des Einsatzes von KI-Technologien für die berufliche Teilhabe von Menschen mit Behinderungen. 2022 hat sie an der Universität Hamburg eine Weiterbildung zur Change Managerin mit dem Schwerpunkt digitale Transformation absolviert.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... gilt der Betrachtung und Begleitung von Veränderungsprozessen, die KI-Technologien aber auch die Digitalisierung insgesamt derzeit insbesondere in der Arbeitswelt anstoßen. Hier entstehen gerade für Inklusion von Menschen mit Behinderungen auf dem allgemeinen Arbeitsmarkt vielfältige Chancen. Die Gestaltung einer inklusiven Arbeitswelt betrachte ich dabei als eine gemeinsame Aufgabe von Forschung und Praxis. Insofern gilt meine Leidenschaft auch der transdisziplinären und transformativen Forschung, die Veränderungen nicht nur als Außenstehende untersucht, sondern auch aktiv mitgestaltet. Dabei kommt es auf die Kooperation von Forschungs- und Praxisakteuren an, bei der in partizipativen Formaten alle Interessen mitgenommen und nachhaltige Lösungen erarbeitet werden.

Leya Küsters, Melanie Schmidt, Barbara Klein

Künstliche Intelligenz im Versorgungskontext von Menschen mit psychischen Erkrankungen: Innovative Ansätze für Betreuung und Unterstützung

Die Gesundheitsversorgung von Menschen mit psychischen Erkrankungen in Deutschland unterliegt verschiedenen Einflussfaktoren, wie dem Personalmangel in den Gesundheitsberufen aber auch der Digitalisierung und Technisierung. Um eine qualitativ hochwertige Versorgung trotz Unterbesetzung bei zeitgleicher Überbelastung der Fachkräfte zu gewährleisten, können innovative Ansätze hilfreich sein. Vor allem in den vergangenen Jahren hat sich der technologische Fortschritt rasant beschleunigt, sodass neuartige Lösungen viele Lebens- und Arbeitsbereiche durchdringen (Birkner 2021). Intelligente Assistenzsysteme wie smarte Uhren, die über Sensoren Stürze erkennen und kontinuierlich die Herzfrequenz analysieren oder Rollatoren mit automatisierter Geschwindigkeitseinstellung und Navigationsunterstützung nehmen an Wichtigkeit zu (Klein et al. 2023). Dabei können neue Technologien und Entwicklungen aus dem Bereich der künstlichen Intelligenz bei stark strukturierten oder sich wiederholenden Tätigkeiten sowie zur Dokumentation und Kontrolle von Gesundheitsdaten eingesetzt werden (Budde et al. 2023). Die Nutzung dieser unterstützenden Technologie soll Fachkräfte von Routinetätigkeiten entlasten, sodass die gewonnene Zeit den Patientinnen und Patienten zugutekommen kann (Bures et al. 2023). Des Weiteren kann dadurch die digitale Teilhabe der Anwenderinnen und Anwender gefördert werden. Dieser Ansatz wird im Forschungsprojekt „TeilhabeAssistenz – Digitale Lösungen für betreute Wohnformen“ verfolgt, was aus Mitteln des Landes Hessen im Rahmen der Distr@I-Förderlinie 2: Digitale Innovationsprojekte gefördert wird.

Im Forschungsprojekt wird der Einsatz von Telepräsenzrobotern im Versorgungskontext von Menschen

mit psychischen Erkrankungen in betreuten Wohnformen untersucht. Dadurch sollen digitale Teilhabe und der Erwerb digitaler Kompetenzen ermöglicht werden. Die Fachkräfte der psychiatrischen Versorgung können dabei die Möglichkeiten der digitalen Assistenz und Betreuung für Menschen mit psychischer Erkrankung mithilfe des Telepräsenzroboters erproben. Das Verbundvorhaben ist am Forschungszentrum FUTURE AGING der Frankfurt University of Applied Sciences angesiedelt. Beteiligt sind das Softwareunternehmen PureSec sowie der Anwendungspartner Vitos Begleitende Psychiatrische Dienste gGmbH. Während der Projektlaufzeit von 29 Monaten wurde der Einsatz von Telepräsenzrobotern vorbereitet, sodass diese nach befürworteter Stellungnahme der Ethikkommission über einen Zeitraum von 6 bis 8 Monaten in betreuten Wohnformen erprobt werden. Die Interaktion der Klientinnen und Klienten sowie der Fachkräfte mit dem Roboter wird sozialwissenschaftlich begleitet, beispielsweise durch teilnehmende Beobachtungen, Interviews und standardisierte Befragungsinstrumente (Döring & Bortz 2016).

In den Einrichtungen der Vitos Begleitende Psychiatrische Dienste gGmbH leben Menschen mit psychischen Erkrankungen in unterschiedlichen Wohnsituationen: Dazu zählen Einzelwohnungen, Wohngemeinschaften, in der zwei bis drei Personen zusammenleben sowie Wohnverbände mit mehreren Klientinnen und Klienten. Der Betreuungsbedarf der einzelnen Klientinnen und Klienten ist individuell unterschiedlich, wobei immer ein ganzheitlicher Ansatz verfolgt wird: Der Mensch wird als Einheit von Körper und Psyche betrachtet und nicht auf behandlungsbedürftige Symptome reduziert (Gold et al. 2018). Nach dem Empower-

ment-Ansatz, verstanden als Selbstbefähigung, eigene Stärken zu nutzen, sollen Menschen mit psychischen Erkrankungen ermuntert werden, möglichst eigenständig Entscheidungen zu treffen und selbstverantwortlich Lösungen zu finden (Schädle-Deininger 2010). Basierend auf Vorgesprächen und einer ausführlichen Ermittlung der Anforderungen, bei der die zukünftigen Nutzenden (von Unger 2014) involviert wurden, konnten spezifische Inhalte aus drei Hauptanwendungsfeldern identifiziert und konkret ausgestaltet werden: Die digitale Betreuung bietet den Fachkräften, die Klientinnen und Klienten in externen Wohnungen betreuen, die Option der Videokommunikation. Dadurch wird eine ortsunabhängige Kommunikation außerhalb der regulären Präsenzbesuche möglich und den Klientinnen und Klienten ein höheres Sicherheitsgefühl vermittelt. Ein weiteres Anwendungsfeld, die digitale Assistenz, besteht beispielsweise aus regelmäßigen

Erinnerungen, wie die tägliche Medikamenteneinnahme, durch den Telepräsenzroboter. Das dritte Anwendungsfeld stellt die digitale Teilhabe dar. Dazu zählen Elemente, die auf dem Telepräsenzroboter abrufbar sind, wie digitale Bewegungs- und Entspannungsübungen, individualisierte Rätsel, kognitives Training und Apps zum Musikhören.

Für die Erprobungsstudie wurde in einem Wohnverbund mit 16 Klientinnen und Klienten sowie in zwei Einzelwohnungen jeweils ein Telepräsenzroboter Temi (Abbildung 1) eingesetzt. Telepräsenzroboter können eine Bild- und Tonübertragung über ein Display liefern, welches ein Wahrnehmen der Umgebung ermöglicht, ohne physisch anwesend zu sein (Klein 2023). Temi kann autonom durch einen Raum navigieren. Hierfür werden die über Sensoren definierten Flächen digital kartiert und darin festgelegte Ziele nach Befehlen angefahren. Dadurch kann der Telepräsenzroboter beispielsweise zu festgelegten Tageszeiten vordefinierte Punkte im Wohnverbund anfahren und an Essenszeiten oder die Medikamenteneinnahme erinnern. Die Navigation wird durch Lidar Technologie (Light Detection and Ranging) realisiert. Mithilfe von Lichtstrahlen wird die Umgebung gescannt und es werden dreidimensionale Karten erstellt. Dazu werden künstliche Intelligenz (KI) und maschinelles Lernen eingesetzt, um beispielsweise Objekte zu erkennen: KI-Berechnungsverfahren können Lidar-Daten verarbeiten, um Menschen, Gegenstände und Ähnliches in Echtzeit zu erkennen und einzuordnen. Zeitgleich sorgen Entscheidungen der KI-Systeme dafür, dass Zusammenstöße vermieden werden, indem sie Hindernisse identifizieren und alternative Fahrstrecken planen. Durch maschinelles Lernen kann außerdem die effiziente Bewegung durch den Raum umgesetzt werden, indem optimale Fahrtrouten geplant und an verändernde Bedingungen angepasst werden. Die Kombination der Lidar-Technologie mit künstlicher Intelligenz ermöglicht es autonomen Systemen, ihre Umgebung besser zu verstehen, fundierte Entscheidungen zu treffen und sich sicher durch komplexe Orte zu bewegen. Auch das im Projekt genutzte Sprachmodell von Temi basiert auf KI. Ein Sentence-Transformer wandelt gesprochene Sätze um, indem aus einzelnen Satzbausteinen semantische Ähnlichkeiten zwischen den Eingaben der Nutzenden und den Beispieltextrn errech-



Abbildung 1: Telepräsenzroboter Temi (robotemi Ltd.)

net werden (Reimers & Gurevych 2019). Damit kann bestimmt werden, welche der vorgegebenen Dialoge am besten zur Aussage passt. Einfach ausgedrückt, wird das Sprachmodell genutzt, um die ähnlichste der Beispieleingaben zu jeder beliebigen Nutzeneingabe zu finden. Zusätzlich kann das Sprachmodell mit einem Chatbot mit KI-Elementen kombiniert werden, wie ChatGPT. Mithilfe solcher hinterlegten Sprachmodelle wird das Themenspektrum des Roboters über die hinterlegten Dialoge hinaus erweitert. Weitere vorinstallierte Anwendungen des Telepräsenzroboters können über die Sprachsteuerung oder das Touchdisplay aufgerufen werden. Das sind beispielsweise Apps aus den Bereichen Lernen, TV-Mediatheken, kognitives Training sowie Kochanleitungen. Zudem kann Temi mit einem Smartphone über die Temi-App verbunden werden und lässt sich so während der Videotelefonie via Smartphone durch einen Raum steuern. Voraussetzung für die Nutzung der meisten Funktionen ist eine stabile Internetverbindung. Einige der auf dem Roboter vorinstallierten Anwendungen wie der Sprachassistenten Alexa wurden aus Datenschutzgründen nicht genutzt, sodass sie entweder datenschutzkonform neu programmiert oder deaktiviert wurden.

Aufgrund der Ersterprobung von Telepräsenzrobotern bei der beschriebenen vulnerablen Zielgruppe wurde bei der Auswahl der Inhalte des Telepräsenzroboters bewusst auf komplexe Anwendungen verzichtet, um eine bestmögliche Akzeptanz der Technologie durch eine zuverlässig nutzbare Anwendungsoberfläche zu gewährleisten. Zudem wurden alle Funktionen des Telepräsenzroboters vorab aus ethischer und datenschutztechnischer Sicht geprüft und ausschließlich vordefinierte Anforderungen umgesetzt. Nach erfolgreicher Erprobung dieser Anforderungen sind weitere Anknüpfungspunkte mit Anwendungen, die auf KI basieren, integrierbar. Dazu zählt beispielsweise die KI-basierte Gesichtserkennung, die bereits als Funktion auf dem Telepräsenzroboter verfügbar ist. Während der Pilotphase des Projekts wurde die Gesichtserkennung jedoch aus Gründen der Datensicherheit deaktiviert. Gesichtserkennung könnte beispielsweise verwendet werden, um Nutzende zu erkennen, persönlich anzusprechen und individualisierte Aussagen zur Tagesroutine zu kommunizieren. Ein weiteres Anwendungsszenario wäre die Erweiterung der Telepräsenzroboterfunktionen um eine KI-gestützte

sprachgesteuerte Pflegedokumentation. Die bisherige Pflegedokumentation ist häufig sehr zeitintensiv und beansprucht einen bedeutenden Anteil der täglichen Arbeitszeit (Seibert 2021). Unter Verwendung einer direkten Spracherkennung könnten pflege- und therapie-relevante Dokumentationen vor Ort digital abgespeichert werden, sodass zum einen sichergestellt ist, dass alle relevanten Informationen festgehalten wurden, zum anderen führt die KI-basierte Spracherkennungsdokumentation durch die Zeitersparnis der wegfallenden Arbeiten am Computer zu einer effizienten Unterstützung der Fachkräfte (Budde et al. 2023). Während der Erprobung der Telepräsenzroboter im Projekt „TeilhabeAssistenz – Digitale Lösungen“ für betreute Wohnformen haben Erfahrungen der Nutzenden gezeigt, dass die Menüführung über das Touchdisplay zum Teil als schwierig wahrgenommen wurde. Das liegt beispielsweise daran, dass Symbole von aufgerufenen Internetseiten sehr klein dargestellt sind. Für Menschen mit Sehbeeinträchtigungen stellt das eine Nutzungsbarriere dar und mindert ggf. das Nutzungsverhalten und die Akzeptanz der Technologie. Auch hier können Lösungen, die auf KI basieren, Abhilfe schaffen. In einem Forschungsprojekt wird beispielsweise ein Softwareinstrument entwickelt, was Internetseiten KI-gestützt und automatisiert in eine vereinfachte Benutzeroberfläche umwandelt, sodass diese auch von Personengruppen mit geringen Technikerfahrungen gut nutzbar ist (Hessisches Digitalministerium & Lylu GmbH, o. D.). KI bietet auch in der häuslichen Versorgung weitere Möglichkeiten. Mithilfe maschinellen Lernens können Sensordaten interpretiert werden, also eine Beziehung zwischen den digitalen Sensordaten und der realen Situation im häuslichen Umfeld eines Menschen hergestellt werden (Kunhardt 2022). Daten aus verschiedenen Quellen zu verknüpfen, würde einen Mehrwert schaffen, um eine personalisierte und individualisierte Versorgung zu ermöglichen. Robotik kann dabei in Ambient Assisted Living Lösungen integriert werden, also Systemen, die das alltägliche Leben von Menschen mit Beeinträchtigungen situationsabhängig unterstützen (Barber et al. 2022).

Um sicherzustellen, dass der Einsatz neuer Technologien bestmöglich gelingt und den Bedürfnissen der Nutzenden entspricht, ist es unerlässlich, alle Beteiligten frühzeitig einzubeziehen (Bures et al. 2023). Die dabei stattfindende wechselseitige Beeinflussung der sozia-

len Wirklichkeiten der Entwickelnden, Nutzenden und Forschenden soll ein gegenseitiges Verständnis schaffen und ein Empowerment erzielen (von Unger 2014). Der gemeinsame Entwicklungsprozess nimmt Einfluss auf die Akzeptanz und Nutzung der Technologie, da diese u. a. von der wahrgenommenen Benutzungsfreundlichkeit abhängt (Venkatesh & Bala 2008). Bei der bisherigen Verwendung und Beurteilung von KI-Technologien in der Pflege wurde ein partizipatives Vorgehen jedoch kaum umgesetzt, wie eine Übersichtsarbeit zeigt (Bures et al. 2023). Eine Ursache kann in Herausforderungen bezüglich der Nutzenden liegen: Ein partizipatives Vorgehen kann beispielsweise bei Menschen mit demenziellen Erkrankungen schwierig umsetzbar sein und wirft zusätzliche ethische Fragestellungen, auch in Bezug auf die Einwilligungsfähigkeit, auf (von Laufenberg 2022).

Zusammenfassend lassen sich Potenziale durch den Einsatz von KI in der Pflege ableiten, wie eine Entlastung

des Personals, die Verbesserung der Versorgungsqualität sowie die Chancen, individualisierte Pflegeprozesse umzusetzen (Bures et al. 2023). Damit der Einfluss der neuen Technologien gelingen kann, ist eine flexible und bedachte Anpassungsfähigkeit im Transformationsprozess notwendig (Birkner 2021). Grundsätzlich sind neben den technologischen Entwicklungen ethische, rechtliche und rahmenbedingende Faktoren zu berücksichtigen. Hierfür kann ein Konzept zur Einbettung von KI-Systemen in der Pflege genutzt werden (Wolf-Ostermann 2021). Für den (zukünftigen) Einsatz KI-basierter Technologien bedarf es neben der partizipativen Einbindung relevanter Personengruppen auch die Vermittlung eines grundlegenden Verständnisses der KI-Technologie (Seibert 2021). Zudem sind weitere Forschungsarbeiten notwendig, um den Einsatz, tatsächlichen Nutzen und die Auswirkungen von KI auf die Zielgruppe umfassender zu beleuchten und zu bestehenden oder alternativen Lösungen in Bezug zu setzen (Seibert et al. 2021).

Literaturangaben:

- Barber, R.; Ortiz, F. J.; Garrido, S.; Calatrava-Nicolás, F. M.; Mora, A.; Prados, A.; Vera-Repullo, J. A.; Roca-González, J.; Méndez, I.; Mozos, Ó.M. : A Multirobot System in an Assisted Home Environment to Support the Elderly in Their Daily Lives. *Sensors* 22(20):7983. 2022. Online: <https://doi.org/10.3390/s22207983>.
- Birkner, F.: Der Mensch im digitalen Wandel. In: Hildebrandt, A., Landhäußer, W. (Hrsg.) CSR und Digitalisierung. Management-Reihe Corporate Social Responsibility. Springer Gabler, 2021. Online: https://doi.org/10.1007/978-3-662-61836-3_44.
- Budde, K.; Hiltawsky, K.; Eskofier, B.; Heismann, B.; Kirchner, E.; Klevesath, M.; Lang, M.; Loskill, H.; Neumuth, T.; Schapranow, M.-P.; Schmidt-Rumpoosch, A.; Susec, B.; Welskopp-Deffaa, E. M.; Wolf-Ostermann, K. : KI für Gesundheitsfachkräfte – Chancen und Herausforderungen von medizinischen und pflegerischen KI-Anwendungen. Whitepaper aus der Plattform Lernende Systeme. 2023. Online: https://doi.org/10.48669/pls_2023-2.
- Bures, D.; Hosters, B.; Reibel, T.; Jovy-Klein, F.; Schramm, J.; Brendt-Müller, J.; Sander, J.; Diehl, A.: Die transformative Wirkung von künstlicher Intelligenz im Krankenhaus. *Innere Medizin* 64, 1025–1032. 2023. Online: <https://doi.org/10.1007/s00108-023-01597-9>.
- Döring, N.; Bortz, J.: Datenerhebung. In: Döring, N., Bortz, J. (Hrsg.), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer-Lehrbuch, 2016. Online: https://doi.org/10.1007/978-3-642-41089-5_10.
- Gold, K.; Schlegel, Y.; Stein, K.-P. (Hrsg.): *Pflege konkret. Neurologie Psychiatrie*. 6. Auflage. Elsevier, 2018.
- Hessisches Digitalministerium & Lylu GmbH (Hrsg.). (o. D.). ABUS. LIDIA-Plattform Hessen. Abgerufen am 9. Dezember 2023, von <https://www.lidia-hessen.de/projekte/abus/>.
- Klein, B.; Graf, B.; Röhrich, K.; Ringwald, M.; Schmidt, M.; Schlömer, I. F.; Roßberg, H.: *Robotik in der Gesundheitswirtschaft: Einsatzfelder und Potenziale*. 2. überarbeitete und erweiterte Auflage. Medhochzwei, 2023.
- Kunhardt, H.: Smarthome und Smartliving – Möglichkeiten und Grenzen der KI im Projekt DeinHaus4.0. In: Pfannstiel, M.A. (Hrsg.) *Künstliche Intelligenz im Gesundheitswesen*. Springer Gabler, Wiesbaden 2022. Online: https://doi.org/10.1007/978-3-658-33597-7_41.
- von Laufenberg, R.: KI-Systeme in Pflegeeinrichtungen – Erwartungen, Altersbilder und Überwachung. In *Künstliche Intelligenz, Demokratie und Privatheit*. Nomos Verlagsgesellschaft mbH & Co. KG, 2022, S. 353-376.
- Reimers, N.; Gurevych, I.: Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint arXiv:1908.10084. 2019.
- Seibert, K.; Domhoff, D.; Bruch, D.; Schulte-Althoff, M.; Fürstenau, D.; Biessmann, F.; Wolf-Ostermann, K.: Application Scenarios for Artificial Intelligence in Nursing Care: Rapid Review. *Journal of medical Internet research*, 23(11), e26522. 2021. Online: <https://doi.org/10.2196/26522>.
- Schädle-Deiniger, H.: *Fachpflege Psychiatrie*. Mabuse, 2010.
- von Unger, H.: Partizipative Ansätze. In: von Unger (Hrsg.), *Partizipative Forschung*. Springer VS, 2014. Online: https://doi.org/10.1007/978-3-658-01290-8_2.
- Wolf-Ostermann, K.; Fürstenau, D.; Theune, S.; Bergmann, L.; Biessmann, F.; Domhoff, D.; Schulte-Althoff, M.; Seibert, K.: Konzept zur Einbettung von KI-Systemen in der Pflege. 2021.

Leya Küsters

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Forschungszentrum FUTURE AGING
Leya.kuesters@fb4.fra-uas.de

Lebenslauf in Kürze

Leya Küsters ist Rehabilitationspädagogin und schließt aktuell den interdisziplinären Masterstudiengang Inclusive Design – Digital Health und Case Management an der Frankfurt University of Applied Sciences ab. Seit 2021 arbeitet sie in verschiedenen Projekten mit Fokus auf Technologien für Menschen mit Beeinträchtigungen am Forschungszentrum FUTURE AGING, u. a. im Projekt TeilhabeAssistenz – Digitale Lösungen für betreute Wohnformen. Außerdem ist sie Digital Health und Education Multiplikatorin der Robert Bosch Stiftung und Careum Stiftung. Während ihrer ersten akademischen Ausbildung unterstützte sie in einem Forschungsprojekt zur digitalen Förderung von Schulkindern mit Teilleistungsstörungen.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegt darin Barrieren, mithilfe dieser abzubauen. Das setzt voraus, dass KI für alle zugänglich und nutzbar ist.

Melanie Schmidt

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Forschungszentrum FUTURE AGING
melanie.schmidt@fb4.fra-uas.de

Lebenslauf in Kürze

Melanie Schmidt ist gelernte Krankenschwester und war lange Jahre in der Zentralen Notaufnahme der Universitätsklinik Frankfurt tätig. 2005 bis 2011 studierte sie Wissenschaftsjournalismus (B. A.), schloss 2021 zudem ein Studium der Pädagogik für Pflege- & Gesundheitsberufe (B. A.) sowie 2023 ein Masterstudium in Erwachsenenbildung (M. A.) an der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau ab. Von 2014 bis 2017 war sie als wissenschaftliche Mitarbeiterin im BMBF-Projekt Main-Career an der Offenen Hochschule Frankfurt UAS tätig, im Anschluss als HMSI-Projektmitarbeiterin des Hessischen Kindervorsorgezentrum der Universitätsklinik Frankfurt. Seit Juli 2021 ist sie wissenschaftliche Mitarbeiterin am Forschungszentrum FUTURE AGING an der Frankfurt University of Applied Sciences.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... beginnt dann, wenn künstliche Intelligenz dazu beiträgt, die Versorgungssituation von Menschen zu verbessern. Führt KI beispielsweise dazu, dass Pflegefachpersonen mehr Zeit für Patientinnen und Patienten haben, weil Routinetätigkeiten automatisiert werden, bleibt der Mensch im Fokus.

Miriam Peters

Einsatzmöglichkeiten für KI in der ambulanten Pflege

1. Herausforderungen und Charakteristika der ambulanten Pflege

Knapp 5 Mio. pflegebedürftige Personen gab es im Jahr 2021 in Deutschland (Pflegebedürftige nach Versorgungsart, Geschlecht und Pflegegrade – Statistisches Bundesamt (destatis.de)). 4, 17 Mio. Personen davon werden zuhause versorgt. 1,05 Mio. Personen wurden im Jahr 2021 von ca. 15.400 ambulanten Pflegedienste betreut, 2,65 Mio. Personen allein durch Angehörige. (Beschäftigte in der Pflege (bundesgesundheitsministerium.de). Einer Vorausberechnung der Pflegestatistik zufolge werden 5,6 Millionen pflegebedürftige Personen Ende 2035 zu versorgen sein, diese Zahl wird auf 6,8 Millionen Ende 2055 ansteigen und schließlich im Jahr 2070 bei 6,9 Millionen Pflegebedürftigen liegen. Wird angenommen, dass die Pflegequoten infolge des 2017 weiter gefassten Pflegebedürftigkeitsbegriffs noch bis 2027 zunehmen, so werden 2035 rund 6,3 Millionen und 2055 rund 7,6 Millionen Pflegebedürftige ausgewiesen (Zahl der Pflegebedürftigen steigt bis 2070 deutlich an – Statistisches Bundesamt (destatis.de)). Zehn Jahre zuvor waren es knapp ein Fünftel (18,6 %) weniger (Schwinger und Tsiasioti 2018).

Ambulante Pflege zeichnet sich durch Flexibilität und den Fokus auf die häusliche Betreuung von Patienten aus. Im Mittelpunkt steht die individuelle, bedarfsgerechte Unterstützung, die es ermöglicht, Menschen in ihrer gewohnten Umgebung zu versorgen. Die Herausforderungen liegen dabei in der Vielfalt der Patientenbedürfnisse und der Notwendigkeit einer effizienten Ressourcenallokation. Büscher definiert häusliche Pflege wie folgt: „die Gesamtheit der pflegerischen Unterstützung in der häuslichen Umgebung eines pflegebedürftigen Menschen“ (Büscher 2020, S. 56). Diese wird sowohl durch Angehörige, Freunde, Bekannte oder andere Personen im Rahmen ihrer sozialen Beziehung zum pflegebedürftigen Menschen

als auch durch professionelle Leistungserbringer, wie ambulante Pflegedienste erbracht. Häufig wird mit dem Begriff der „pflegenden Angehörigen“ eine Pflegeperson vor Ort in räumlicher Nähe der Hilfebedürftigen assoziiert, obwohl eine zunehmende Bedeutung von familiärer Unterstützung aus räumlicher Entfernung heraus („distance caregiving“) beobachtet werden kann (vgl. Wagner et al. 2019, S. 529). Die Unterstützung durch Angehörige und andere Bekannte wird auch als informelle Pflege bezeichnet. Formelle Pflege bezeichnet die Versorgung durch ambulante Pflegedienste und weitere professionelle Leistungserbringer.

Mit der Pflegeversicherung wurde auch das Ziel „ambulant vor stationär“ eingeführt. Häusliche Pflegearrangements haben Anspruch auf Unterstützung durch die Hauptleistungen der Pflegeversicherung, die Sachleistung für die Unterstützung durch einen Pflegedienst und einer Geldleistung für selbst beschaffte Pflegehilfen. Weiterhin werden Leistungen zur Sozialversicherung für Angehörige und Angebote zur Teilnahme an Pflegekursen durch die Pflegeversicherung übernommen, sowie Beratung zu Pflegefragen (vgl. Büscher 2020, S. 57).

Häusliche Pflegearrangements entstehen auf zwei Wegen, entweder schleichend durch zunehmenden Unterstützungsbedarf oder durch Unfälle oder andere plötzliche Ereignisse (vgl. Büscher 2020, S. 57). Neben der Unterstützung von körpernaher Pflege werden beispielsweise Hilfen im Haushalt, das Management von Behördengängen für den älteren Angehörigen, Gespräche mit der Pflegeberatung etc. übernommen (Wagner et al. 2019). Die Leistung, die durch ambulante Pflegedienste erbracht wird, wird als Sachleistung in § 36 SGB XI beschrieben. An zeitlichem Umfang stellt die Leistung der professionell erbrachten Dienstleistung etwa ein Zehntel der erforderlichen

Pflegezeit dar (vgl. Räker et al. 2020, S.79). Vor dem Hintergrund der beschriebenen Herausforderungen werden zunehmend digitale, KI-gestützte Technologien entwickelt, die Pflegearrangements in der Häuslichkeit unterstützen und entlasten sollen. Weiterhin soll KI zur „Attraktivitätssteigerung in der professionellen Pflege beitragen, zum anderen die Pflegebedürftigen in ihren alltäglichen Aufgaben unterstützen, um die Selbstständigkeit zu erhalten. Diese Notwendigkeit besteht insbesondere, da ältere Pflegebedürftige zunehmend alleine zu Hause leben und pflegende Angehörige mit steigender Tendenz in Erwerbsverhältnissen eingebunden sind“ (Karl 2022, S. 6). Allerdings gibt es bislang vergleichsweise wenige Studien zum Technikeinsatz und seinen Folgen für die Arbeitsorganisation und Arbeitsqualität in der ambulanten Pflege (Berger et al. 2017; Daxberger 2018). Wichtige Forschungserkenntnisse aus verschiedenen Studien sind, dass „Technikeinsatz den Interaktionsprozess zwischen Pflegekraft und Klienten oder Klientinnen nicht stören darf, wenn Pflegekräfte die Technik positiv annehmen sollen“ (Pöser und Bleses 2018, S. 9–10).

2. Begriffsbestimmung Künstliche Intelligenz in der Pflege

Karl (2022) beschreibt KI wie folgt: „Handlungen, Problemlösungsstrategien oder Entscheidungsprozesse, die bis vor Kurzem noch Menschen vorbehalten waren, sollen in Zukunft durch Computer auf Grundlage eines soliden Datenpools generiert werden. Im Idealfall sollen die Aktionen, die durch den Computer errechnet werden, von höherer Qualität und Sicherheit geprägt sein, als würden die Vorgänge durch einen Menschen ausgeführt“ (Karl 2022, S. 7). Unterschieden werden dabei maschinelles Lernen, Deep Learning und Big Data (vgl. Lernende Systeme- Die Plattform für Künstliche Intelligenz 2019, S.6). Maschinelles Lernen ist dabei die Schlüsseltechnologie der Künstlichen Intelligenz. Auf Basis großer Beispieldatensätze werden Muster erkannt und durch Algorithmen Modelle entwickelt. Unterschieden wird dabei zwischen überwachtem, unüberwachtem und verstärkendem Lernen. Für das überwachte Lernen wird dem Algorithmus neben den Rohdaten auch ein erwartetes Ergebnis zur Verfügung gestellt, während beim unüberwachtem Lernen kein Prognoseziel for-

muliert wird. Der Algorithmus identifiziert dann Gemeinsamkeiten. Beim verstärkenden Lernen „trifft ein lernendes System Entscheidungen, auf deren Basis es anschließend handelt“ (Lernende Systeme- Die Plattform für Künstliche Intelligenz 2019, S. 7).

Eine weitere Form stellt das Deep learning dar, das auf Basis sogenannter künstlicher neuronaler Netze arbeitet. Diese bestehen aus verschiedenen Schichten – der Eingabeschicht, der Ausgabeschicht und den verborgenen Schichten. Die Anzahl und Verbindung der Neuronen können elektrische Potenziale aufnehmen, verarbeiten und weiterleiten. Darüber hinaus charakterisiert sich die Architektur des künstlichen Neuronalen Netzes über die Anzahl der Neuronen. Neuronale Netze werden eingesetzt, um komplexe Berechnungen durchzuführen, zur Bilderkennung, zum Verarbeiten von natürlicher Sprache in Textform oder zur Spracherkennung in auditiver Form (Karl 2022, S. 8).

3. Technologien, die derzeit in der ambulanten Pflege eingesetzt werden

Die Plattform Lernende Systeme (Budde et al. 2023) nennt als am häufigsten auftretende Einsatzfelder von KI-Anwendungen zwei Anwendungsfelder in der Pflege:

- | Tracking/Monitoring/Klassifizierung von Aktivität und Gesundheit (Umgebungssensoren zur Analyse von Bewegungsdaten; körpernahe Sensoren; Analyse von Vitalwerten zur Detektion/Klassifikation von Zustandsveränderungen). Wolf-Ostermann et al. (Wolf-Ostermann et al. 2021) geben in einem Projektbericht, der unter anderem eine systematische Literaturrecherche beinhaltet an, dass ca. 30% der gefundenen Literatur die Erfassung und das Monitoring von Aktivitäten und gesundheitsbezogenen Parametern adressiert.
- | Koordination und Kommunikation: Planung von Versorgungs- und Arbeitsprozessen; Optimierung von Informationsflüssen. Dazu zählt unter anderem auch die Tourenplanung in der ambulanten Pflege.

Personalplanung und Ressourcenallokation

Konventionelle Tourenplanung ist zeitintensiv und komplex. Gaugisch et al. (2023) formulieren den Einsatz von KI in der Tourenplanung folgendermaßen: „Im zugehörigen ambulanten Pflegedienst wird die KI

in der Tourenplanung eingesetzt. Hier optimiert die KI über das Matching aus Mitarbeitenden- und Kundenwünschen sowie Qualifikationen und Kompetenzen hinaus die Tour so, dass eine effiziente, gesundheitsförderliche, soziale, ökologisch und ökonomisch sinnvolle Route vorgeschlagen wird“ (Gaugisch et al. 2023, S. 30).

KI analysiert dabei historische Daten und Prognosen, um den Bedarf an Pflegepersonal genau zu bestimmen. Der Dienstplan folgt dann nicht mehr einem starren System, sondern wird an zeitlichen und inhaltlichen Bedarfen orientiert. Ziel ist es dabei physische und psychische Belastungen bei den Pflegenden zu reduzieren, indem Arbeitsspitzen identifiziert werden und Personal zusätzlich eingeplant werden kann. Die KI soll zudem die verfügbaren Ressourcen bestmöglich nutzen, bei gleichbleibender oder verbesserter Pflegequalität. Wolf-Ostermann et al. (Wolf-Ostermann et al. 2021) weisen noch darauf hin, dass neben individuellen Versorgungswünschen der Pflegebedürftigen, der Qualifikation und Verfügbarkeit der Mitarbeitenden auch das Verkehrsaufkommen und weitere Informationen zu den Versorgungsarrangements nötig sind, um KI-basierte Tourenplanung für eine verbesserte Qualität einerseits und höheren Zufriedenheit andererseits zu nutzen.

Patientenüberwachung

Sensoren und Wearables erfassen Echtzeitgesundheitsdaten, während KI-Algorithmen Abweichungen erkennen und Frühwarnungen generieren. Dazu gehören unter anderem Systeme, die Stürze erkennen können. Karl (2022) beschreibt dazu die App Lindera, die anhand eines Videos sowie eines psychosozialen Tests eine Prognose über das Sturzrisiko erstellt. Über das Video wird durch ein neuronales Netz ein 3D Bild des Skeletts erstellt. In der App werden dann auf Basis der Ergebnisse der Gangbildanalyse verschiedene Parameter, wie Schrittlänge, -zeit, -höhe, Geschwindigkeit sowie die Kadenz ermittelt, um die Ergebnisse der Analyse im letzten Schritt in einem Bericht zusammenzufassen. Anhand des Ergebnisses können dann präventive Maßnahmen eingeleitet werden. Ein positiver Effekt für das Gesundheitssystem könnte die Vermeidung von Folgekosten darstellen. Das Verfahren wurde an der Charité bereits erprobt (vgl. Karl 2022, S. 11).

4. Chancen und Herausforderungen des Einsatzes von KI-basierten Technologien in der Pflege

Insgesamt kann der Einsatz KI-basierter Technologien derzeit noch als unzureichend bewertet werden (Zerth et al. 2021). Becka und Evans beschreiben als Ursachen eine „heterogene bzw. unzureichende digitale Infrastrukturausstattung in der Fläche, In-sellösungen, unklare (Re-)Finanzierungsstrukturen, Interoperabilitätsprobleme vorhandener Systeme, die die Übertragung von Daten an den Schnittstellen zwischen Sektoren und Leistungserbringern hemmen, aber auch Defizite in der Evaluation und mangelnde Transparenz bezüglich der Output- und Outcome-bezogenen Nutzeneffekte“ (Evans und Becka 2020, S. 8). Benötigt werden Studien, die Aspekte wie Wirtschaftlichkeit, Bedarfsgerechtigkeit, Versorgungs- und Arbeitsqualität gleichermaßen in den Blick nehmen und die Auswirkungen sowohl auf die betroffenen pflegebedürftigen Menschen (z. B. Akzeptanz, psychosoziale Gesundheit, Sicherheit) als auch die organisationale Perspektive (z. B. Nutzungsintensität, Akzeptanz oder Sicherheit im Versorgungsprozess) in den Blick nehmen (Evans und Becka 2020).

Darüber hinaus bestehen Herausforderungen bei den Themen Datenschutz, Transparenz in den Entscheidungsprozessen der Algorithmen, ethische Bedenken sowie der Wahrung der zwischenmenschlichen Dimension in der Pflege.

Chancen werden unter anderem in einer höheren Selbständigkeit und besserer Vernetzung für Menschen mit Pflege- und Unterstützungsbedarf sowie allen an der Versorgung Beteiligten, der Unterstützung professionellen Handelns und Verbesserung des Versorgungsalltags, und schließlich in der frühzeitigen Erkennung von physischen oder kognitiven Veränderungen und Komplikationen gesehen (Budde 2023, S. 23).

5. Fazit

Die Integration von KI in der ambulanten Pflege könnte zu einer transformativen Verbesserung der Pflegequalität und -effizienz beitragen. Dazu wären die oben beschriebenen Herausforderungen zu überwinden und KI Technologien an die Bedürfnisse der ambulanten Pflegedienste sowie der zu versorgenden

Personen anzupassen. Ein ausgewogener Ansatz, der Technologie und Menschlichkeit vereint, wird entscheidend sein, um die Potenziale von KI voll auszuschöpfen und gleichzeitig die Integrität und Fürsorge in der Pflege zu bewahren. Becka und Evans (2020) verweisen darauf, dass die digital gestützte Reorganisation pflegespezifischer Verantwortungs-, Aufgaben- und Tätigkeitsbereiche nicht nur auf Entlastungsef-

fekte abzielt, sondern in diesem Kontext auch die Bedeutung pflegespezifischen Erfahrungswissens, von Qualifikation und Kompetenz neu verhandelt wird. Sie argumentieren, dass eine Aufwertung und Entlastung beruflich Pflegenden durch digitale Technik vor allem vom Ziel des Einsatzes und den Möglichkeiten der Pflegefachpersonen im Sinne einer aktiven, professionellen Steuerung abhängen.

Literaturangaben:

- Karl, A.: Potenziale und Anwendungsszenarien künstlicher Intelligenz in häuslichen Pflegearrangements im Kontext einer alternden Gesellschaft. Bayerisches Forschungszentrum. Kempten (BZPD Working Paper) 2022.
- Berger, R.: Deutsches Institut für angewandte Pflegeforschung; Vallendar, Philosophisch-Theologische Hochschule: Informations- und Kommunikationstechnologie für die Pflege. Studie im Auftrag des Bundesministeriums für Gesundheit. 2017. [zuletzt geprüft am 23.07.2017].
- Blumenberg, P.; Krebs, M.; Niemann, L.-M.; Stehling, H.; Stomberg, D.: Expertenstandard Förderung der Mundgesundheit in der Pflege. Osnabrück 2023. Online: <https://d-nb.info/1287938442/04>.
- Budde, K.; u.a.: KI für Gesundheitsfachkräfte: Chancen und Herausforderungen von medizinischen und pflegerischen KI-Anwendungen. Lernende Systeme – Die Plattform für Künstliche Intelligenz, Geschäftsstelle c/o acatech. München 2023.
- Büscher, A.: Bedarfslagen in der häuslichen Pflege. In: Neuausrichtung von Versorgung und Finanzierung. Springer Open, Berlin 2020.
- Daxberger, S.: Neue Technologien in der ambulanten Pflege. Wie Smartphones die Pflegepraxis (mit-) gestalten. Mabuse-Verlag, Frankfurt am Main 2018.
- Evans, M.; Becka, D.: Pflege und Digitalisierung. In: Tanja Klenk, Frank Nullmeier und Göttrik Wewer (Hg.): Handbuch Digitalisierung in Staat und Verwaltung. Springer Fachmedien, Wiesbaden 2020, S. 1–12.
- Gaugisch, P.; Risch, B.; Stolze, D.; Strunck, S.: Future Care and Services Attraktive und adaptive Pflege. Chancen und Potenziale von New Work in der Altenpflege. 2023. Online: https://pflege.iao.fraunhofer.de/wp-content/uploads/FuCaSe.ChancenundPotenzialevonNewWorkinderAltenpflege_oi24098fu4.pdf.
- Lernende Systeme- Die Plattform für Künstliche Intelligenz: Lernende Systeme im Gesundheitswesen. Grundlagen, Anwendungsszenarien und Gestaltungsoptionen. 2019.
- Pöser, S.; Bleses, P.: Digitalisierung der Arbeit in der ambulanten Pflege im Land Bremen. Praxis und Gestaltungsbedarfe digitaler Tourenbegleiter. 1. Auflage. Institut Arbeit und Wirtschaft Arbeitnehmerkammer, Bremen (Arbeit und Wirtschaft in Bremen, Ausgabe 25) 2018. Online: <http://hdl.handle.net/10419/179518>.
- Räker, M.; Schwinger, A.; Klauber, J.: Was leisten ambulante Pflegehaushalte? Eine Befragung zu Eigenleistungen und finanziellen Aufwänden. In: Neuausrichtung von Versorgung und Finanzierung. Springer Open, Berlin 2020.
- Schwinger, A.; Tsiasioti, C.: Pflegebedürftigkeit in Deutschland. In: Klaus Jacobs, Adelheid Kuhlmeier, Stefan Greß, Jürgen Klauber und Antje Schwinger (Hg.): Qualität in der Pflege: mit 24 Tabellen und 43 Abbildungen. Springer Open, Pflege-Report, Berlin 2018. Online: https://link.springer.com/chapter/10.1007/978-3-662-56822-4_16.
- Wagner, M.; Franke, A.; Otto, U.: Pflege über räumliche Distanz hinweg. In: Zeitschrift für Gerontologie und Geriatrie 52 (6). 2019, S. 529–536. Online: <https://doi.org/10.1007/s00391-019-01605-4>.
- Wolf-Ostermann, K.; Fürstenau, D.; Theune, S.; Bergmann, L.; Biesmann, F.: Konzept zur Einbettung von KI-Systemen in der Pflege. Sondierungsprojekt zu KI in der Pflege (SoKIP), 2021.
- Zerth, J.; Jaensch, P.; Müller, S.: Technik, Pflegeinnovation und Implementierungsbedingungen. In: Klaus Jacobs, Adelheid Kuhlmeier, Stefan Greß, Jürgen Klauber und Antje Schwinger (Hg.): Pflege-Report 2021: Sicherstellung der Pflege: Bedarfslagen und Angebotsstrukturen. Springer, Berlin/Heidelberg 2021, S. 157–172.

Prof. Dr. Miriam Peters

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Miriam.peters@fb4.fra-uas.de

Lebenslauf in Kürze

Prof. Dr. Miriam Peters ist seit 2022 als Professorin für klinische Pflege an der Frankfurt University of Applied Sciences. Ihre Arbeitsschwerpunkte liegen im Bereich der Entwicklung von Aufgabenprofilen für akademisch qualifiziert Pflegende, Fallarbeit in der Pflege, Neue Technologien in Bildungs- und Versorgungspraxis sowie methodenbasierter Theorieentwicklung in der Pflege. Zuvor hat sie am Bundesinstitut für Berufsbildung ein Forschungsprogramm zur Pflegebildung und zum Pflegeberuf (vgl. § 60 Absatz 4 PflAPrV) geleitet. Nach einem Studium zur Diplom-Kauffrau schloss sie eine Ausbildung zur Gesundheits- und Krankenpflegende ab und war einige Jahre in verschiedenen Versorgungsbereichen der Intensivpflege tätig. Das Master- und Promotionsstudium im Bereich Pflegewissenschaft hat sie an der Philosophisch-Theologischen Hochschule Vallendar absolviert. 2021 schloss sie ihre Promotion (Dr.rer.cur.) zum Thema „Digitale Bildung in der Altenpflege“ ab.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Mitgestaltung von KI Einsatzmöglichkeiten in verschiedenen Settings der ambulanten Pflege sowie die Anbahnung von Kompetenzen im Umgang mit künstlicher Intelligenz in der Pflege.



Lisa Luft

Versorgungsperspektive in der Altenpflege

Im vorliegenden Beitrag werden mit Blick auf die aktuelle Studienlage sowie anhand von exemplarischen KI-basierten Anwendungsbeispielen aus dem Bereich der stationären Langzeitpflege aktuelle Fragestellungen beleuchtet und Impulse für eine nachhaltige und zielgruppengerechte Umsetzung abgeleitet. Da es sich um einen auszugshaften Einblick in aktuelle Forschungs- bzw. Projektergebnisse und keine umfassende oder gar abschließende Betrachtung der Thematik handelt, wurden als Diskussionsgrundlage insbesondere Übersichtsarbeiten und Sammelwerke aufgegriffen. Zudem gilt anzumerken, dass im Rahmen einer kritischen Betrachtung auch datenschutzrechtliche Aspekte von hoher Bedeutung sind, die im vorliegenden Beitrag allerdings nicht thematisiert werden.

Besonderheiten der Versorgungsperspektive „stationäre Langzeitversorgung“ und Implikationen für die Forschung.

Die Versorgungsperspektive „stationäre Langzeitversorgung“ zeichnet sich durch die Besonderheit aus, dass Überlegungen zur Implementierung KI-gestützter Anwendungen eine komplexe Betrachtungsweise erfordern, die multiperspektivisch sowie interdisziplinär ausgerichtet sein sollte (Deckert et al., 2 ff.); denn: „Gute Technik umfasst mehr als bloße Technik“ (Weber 2021, 21). So ist insbesondere der soziale und emotionale Kontext zu betrachten und Technik als ein Bestandteil in diesem soziotechnischen System zu verstehen (ebd.). Zudem werden sowohl volkswirtschaftliche Aspekte wie Kosten, Effizienz, Wirtschaftlichkeit und Nachhaltigkeit betrachtet, die im Setting der stationären Langzeitpflege zunehmend in den Fokus rücken (Deckert et al. 2021, 2ff.) als auch industriepolitische Fragen und Geschäftspotenziale (Weber 2021, 22).

Demgegenüber steht die Perspektive der Menschen, die in Einrichtungen der stationären Langzeitpflege

leben und die aufgrund ihrer körperlichen und/oder kognitiven Einschränkungen vorwiegend ein erhöhtes Schutzbedürfnis aufweisen bzw. ihre eigenen Rechte unter Umständen nicht mehr in vollem Maße vertreten können. So werden neben den oben genannten wirtschaftlichen Aspekten auch ethische Fragestellungen bedeutsam sowie die Einhaltung der Charta der Rechte hilfe- und pflegebedürftiger Menschen. Übergeordnetes Ziel ist es dementsprechend die Würde, Selbstbestimmung, Autonomie, die individuellen (Grund-) Bedürfnisse, das Wohl (Deckert et al. 2022, 2 ff.) und die Lebensqualität der Menschen in den Fokus der Handlungen zu stellen (ebd., 13). Hinzu kommt, dass sich die pflegerische Versorgung im Setting der stationären Langzeitpflege komplex gestaltet. Es handelt sich in der Regel um sehr sensible und intime Versorgungssituationen, die körpernahe Tätigkeiten erfordern. Zudem befinden sich zu pflegende Menschen häufig in existenziell bedrohenden Lebenssituationen, erleben beispielsweise Trauer, sodass soziale Aspekte wie Beziehungsaufbau und -pflege, Kommunikation, professionelle Nähe sowie ein vertrauensvoller Umgang einen zentralen Stellenwert einnehmen (Deckert et al. 2022, 13 ff.). Vor dem Hintergrund dieser Komplexität werden zugleich auch mögliche Grenzen KI-basierter Anwendungen deutlich. So können beispielsweise humanoide Roboter zwar Routinen KI-gestützt erlernen und Anomalien erkennen, jedoch nicht kreativ und flexibel anpassbar handeln, sodass sie in ihrer Anwendung als eine entlastende und qualitätsverbessernde Maßnahme unter Berücksichtigung des jeweiligen Einsatzszenariums und damit verbundenen Bedarfen kritisch zu prüfen sind und Pflegepersonal nicht ersetzen können (ebd., 34).

Ziel sollte eine partizipative und „vorausschauende“ Technikentwicklung sein, die laut Weber (2021) trotz der tendenziell als zunehmend wahrgenommenen Anwendung im Rahmen von BMBF Förderlinien in aktu-

ellen Forschungsvorhaben bislang nur bedingt umgesetzt wurden (ebd., 22). Im Zuge dessen nimmt auch die Perspektive der Pflegefachpersonen in ihrer Rolle als Nutzer*innen bzw. Anwender*innen einen zentralen Stellenwert ein. Mit Verweis auf das Scoping Review von Vetter & Cerullo (2021) merkt Weber (2021) weiterhin an, dass auch die Festlegung einer allgemein anerkannten Definition aussteht, was unter einer altersgerechten Ausgestaltung von Assistenzsystemen zu verstehen ist. Diese unklare Begriffslage führt beispielsweise dazu, dass sich auch in den Studien eine heterogene Kategorisierung der angewendeten Technologien abzeichnet. Somit ist die Datenlage als unzureichend belastbar einzustufen, sodass eine wichtige Grundlage fehlt, um Assistenzsysteme bedarfsgerecht (weiter-)entwickeln und eine Aussage zu Möglichkeiten der Anwendung treffen zu können (ebd., 17).

Blick in die aktuelle Studienlage zur nationalen Umsetzung und KI-basierte Anwendungsbeispiele

Die Ergebnisse des vom BMBF geförderten Projekts »Diffusion altersgerechter Assistenzsysteme – Kennzahlenerhebung und Identifikation von Nutzungshemmnissen« weisen darauf hin, dass trotz hoher Förderaufwendungen für altersgerechte Assistenzsysteme und einer tendenziell hohen Akzeptanz vonseiten der Praxis eine nachhaltige Implementation kaum stattgefunden zu haben scheint (Weber 2021, 11). So kommen Vetter & Cerullo (2021) in ihrem Scoping Review beispielsweise zu dem Ergebnis, dass in den 11 von ihnen analysierten Artikeln (Stand der Literaturrecherche: 2020, nationale Artikel: n=10) am häufigsten die elektronische Pflegedokumentation und assistive Technologien (bspw. Hebehilfen) Anwendung gefunden haben. Die Daten deuten insgesamt auf eine eher zögerliche Umsetzung hin, was insbesondere den Einsatz von Robotik betrifft. Auffällig war außerdem, dass fast ausschließlich Prototypen zur Anwendung kamen und dass es sich überwiegend um Pilotprojekte handelte. Im Bereich der Robotik wurden fast ausschließlich Serviceroboter eingesetzt. Kritisiert wird insbesondere, dass die Perspektive der Pflegefachpersonen trotz signalisiertem Interesse bislang nur bedingt berücksichtigt wurde. Fokussiert wurde hingegen meist die „reine“ Technik.

Als Gründe für die zögerliche Umsetzung wird von den Autorinnen herausgearbeitet, dass in den nationalen

stationären Langzeitpflegeeinrichtungen zunächst Rahmenbedingungen geschaffen, finanzielle Fragen geklärt und Schulungsangebote für Mitarbeitende etabliert werden müssten (ebd., 178 ff.). Auch Deckert et al. 2022 beschreiben suboptimale Rahmenbedingungen wie ein unstabiles und vor allem datenschutzrechtlich unsicheres W-Lan (ebd., 18) sowie einen hohen Implementierungsaufwand (ebd., 1). Weber (2021) nennt als weitere hemmende Faktoren die hohen Entwicklungskosten der Technik und damit verbundene Verzögerungen in der (Weiter-)Entwicklung sowie fehlende Marktreife (ebd., 17 f.).

Die Akzeptanz gegenüber dem Einsatz KI-basierter Anwendungen wird (ausgenommen gegenüber der Robotik) insgesamt als hoch eingeschätzt, obgleich auch ein Festhalten an veralteten Leitbildern und Professionsverständnissen thematisiert wird (Weber 2021, 18). Als akzeptanzförderliche Maßnahme wird beschrieben, alle am Prozess beteiligte Personen über die Anwendung ausreichend zu informieren und die Sinnhaftigkeit der KI-basierten Anwendungen zu begründen (ebd., 11 & 39). Vonseiten pflegender Personen wurden als positive Aspekte insbesondere die mit der elektronischen Pflegedokumentation verbundenen Zeitersparnisse hervorgehoben, jedoch zugleich auch der Wunsch nach mobilen Geräten geäußert (Vetter & Cerullo 2021, 178 ff.). Auch Deckert et al. (2022) beschreiben auf Basis von Projekterfahrungen, dass KI-gestützte Dokumentationssoftware als Arbeitserleichterung wahrgenommen wird, da eine dreifach schnellere Texterfassung möglich ist. In ihrer Publikation stellen sie Software vor, die in der Lage ist, fehlerhafte oder unvollständige Eingaben zu erfassen und erkannte Begrifflichkeiten in abrechenbare Leistungen zu überführen. Eine weitere Software passt sich zudem durch einen integrierten Spracherkennungsdienst an die Anwender*innen und die jeweilige Umgebung an. Auch hier werden Spracheingaben analysiert und in Form von Einträgen in die Pflegedokumentation überführt. Dies ermöglicht u.a. eine zeitlich flexible Pflegedokumentation per Smartphone (ebd., 21 ff.).

Als ein weiteres Anwendungsfeld KI-gestützter Technik im Setting der stationären Langzeitversorgung stellen Deckert et al. (2022) Telecare/Augmented Reality vor. Datenbrillen werden bislang überwiegend im Fort- und

Weiterbildungsbereich eingesetzt. Ziel ist es insbesondere, die Umsetzung von Pflegeprozessen durch das Zuschalten von Expert*innen und/oder eingeblendete Informationen wie Pflegestandards und/oder Auszüge aus Pflegedokumentationen zu unterstützen. Möglich ist außerdem eine Barcodeidentifizierung, Sprachsteuerung oder Fotodokumentation (ebd., 25 ff.).

Im Hinblick auf Einsatzfelder der Robotik wird in Klein et al. (2023) ein umfassender Überblick mit vielfältigen Anwendungsbeispielen, Potenzialen und Handlungsempfehlungen gegeben. Soziale bzw. humanoide Roboter werden insbesondere zur Förderung der Interaktion, beispielsweise im Bereich der Versorgung von Menschen mit Demenz eingesetzt. Gefördert werden können u.a. die Selbstständigkeit und Mobilität durch Erinnerungen, Handlungsanleitungen und Gedächtnistraining. Zudem können KI-basierte Lösungen zur Initiierung von Handlungsänderungen entwickelt und eingebracht werden. Auch ein Kontakt zu Angehörigen kann hergestellt, zu pflegende Menschen in ihren Zimmern „besucht“ und angesprochen werden. Möglich ist auch eine Ortung zum Schutz der zu pflegenden Menschen und die Übertragung von Daten in die Pflegedokumentation (Deckert et al. 2022, 29).

Abschließend sind auch KI-basierte Assistenzsysteme wie die Sturzsensorik zu nennen (ebd., 17 f.). Zur Anwendung kommt hier insbesondere die automatische Sturzerkennung und es werden Anomalitäten im Mobilitätsverhalten auf Basis von Aktivitätsanalysen erkannt. Weiterhin werden bei Bedarf Alarme ausgesendet und/oder Beleuchtung automatisch aktiviert (ebd., 34). Ein weiteres Beispiel ist der im Projekt TrinkTracker entwickelte „intelligente Pflegebecher“. Durch eine sensorgestützte automatische Erfassung sowie Dokumentation der Trinkmenge ist der Trinkbecher bei unzureichender Flüssigkeitszufuhr in der Lage eine Trink-Erinnerung abzusenden¹.

Aktuelle Fragestellungen sowie Denkanstöße zur Unterstützung der nachhaltigen und zielgruppenorientierten Umsetzung

Auf Basis der im Beitrag aufgezeigten exemplarischen Einblicke in aktuelle Forschungs- und Projektergebnisse wird weiterer dringender Forschungsbedarf zur Thematik in vielfacher Hinsicht deutlich.

| Die nur bedingt belastbare nationale Datenlage sowie die zurückhaltende Umsetzung von KI-basierten Anwendungsmöglichkeiten kann aufgrund von fehlenden Erfahrungen laut Weber (2021) zu einer Verstärkung von Vorurteilen und/oder zu hohen Erwartungen führen (ebd., 40).

| Eine Erweiterung der Forschungsperspektive ist dringend erforderlich, um insbesondere den sozialen und emotionalen Kontext perspektivisch stärker in den Blick zu nehmen.

| Der Schutz und das Wohl der zu pflegenden Menschen ist im Setting der stationären Langzeitversorgung in einer ganz besonderen Weise zu berücksichtigen. Daher sollte eine kritische Abwägung von Nutzen und Risiken erfolgen und vor allem individuelle Bedürfnisse beachtet werden (Wahrung von Autonomie und Selbstbestimmung).

| Auch die Anwender*innen sind perspektivisch stärker in die Entwicklungsprozesse einzubeziehen (partizipativen Technologieentwicklung). Hierzu ist es beispielsweise empfehlenswert zunächst einrichtungsspezifische Bedarfe, Rahmenbedingungen und Arbeitsbelastungen aus Sicht der Pflegefachpersonen zu erheben und Szenarien zu identifizieren, die sich für KI-bezogene Anwendungen eignen könnten. So resümiert Weber (2021) sehr eindrucksvoll: „In der Pflege sollte jedoch der Einsatz von Robotern dem Bedarf folgen und nicht umgekehrt“ (ebd., 33).

| Als zentraler Aspekt wird auch die Begleitung der Anwender*innen im Zuge der Einführung KI-basierter Anwendungen beschrieben (technischer Support, Bereitstellung von Ansprechpersonen, Eröffnung „geschützter“ Räume zur Erprobung, Angebot von Schulungen, Klärung von Bedenken) (Deckert et al. 2022, 21 ff.).

| Auch im Rahmen von Aus-, Fort- und Weiterbildung sollte die Vermittlung digitaler Kompetenzen bspw. durch die Integration in Rahmenlehrplänen erfolgen und nicht erst auf betrieblicher Ebene (Weber 2021, 40).

| Zudem sind in Einrichtungen der stationären Langzeitpflege häufig zunächst Rahmenbedingungen zu schaffen (bspw. verlässliches und stabiles W-Lan).

| Die Verzögerungen in der nationalen (Weiter-)Entwicklung von altersgerechten Assistenzsystemen führen nicht nur zu Abhängigkeiten (Weber 2021, 18), sondern auch dazu, dass eine zielgruppenorientierte Ausgestaltung erschwert wird.

Ziel sollte ein wohlüberlegtes „sinngabendes Zusammenwirken von Menschen und Maschinen“ (Deckert et al. 2022, 11-12) sein. Somit nehmen auch die konzeptionelle Einbettung KI-basierter Anwendungen sowie die interdisziplinäre Betrachtungsweise einen

wichtigen Stellenwert ein. Zudem gilt es, eine kritisch reflektierte Haltung einzunehmen bzw. beizubehalten, um nicht Gefahr zu laufen, die eigene Verantwortung sowie das selbstständige professionelle Handeln an KI-basierte Anwendungen abzugeben.

Literaturangaben:

Deckert, R.; Rascher, I.; Recken, H.: Digitalisierung in der Altenpflege, Analyse und Handlungsempfehlungen. Springer, Wiesbaden 2022.
Online: https://doi.org/10.1007/978-3-658-38973-4_1.

Klein, B.; Graf, B.; Schlömer, I. F.; Roßberg, H.; Röhrich, K.; Baumgarten, S.: Robotik in der Gesundheitswirtschaft: Einsatzfelder und Potenziale/herausgegeben von Stiftung Münch. 2. überarbeitete und erweiterte Auflage. Medhochzwei, Heidelberg 2023.

Weber, K.: Gute Technik für ein gutes Leben?“ In: Frommeld, D.; Scorna, U; Haug, S.; Weber, K. (Hg.): Gute Technik für ein gutes Leben im Alter? Akzeptanz, Chancen und Herausforderungen altersgerechter Assistenzsysteme. Alter - Kultur - Gesellschaft. Band 4. Transcript Verlag, Bielefeld 2021, S. 11-26.

Weber, K.: Altersgerechte Assistenzsysteme: Ein Überblick. In: Frommeld, D.; Scorna, U; Haug, S.; Weber, K. (Hg.): Gute Technik für ein gutes Leben im Alter? Akzeptanz, Chancen und Herausforderungen altersgerechter Assistenzsysteme. Alter - Kultur - Gesellschaft. Band 4. Transcript Verlag, Bielefeld 2021, S. 27-64.

Vetter, M.; Cerullo, L.: Die tatsächliche Nutzung digitaler Assistenzsysteme in der Altenpflege. Ein Scoping Review. In: Frommeld, D.; Scorna, U; Haug, S.; Weber, K. (Hg.): Gute Technik für ein gutes Leben im Alter? Akzeptanz, Chancen und Herausforderungen altersgerechter Assistenzsysteme. Alter - Kultur - Gesellschaft. Band 4. Transcript Verlag Bielefeld 2021, S. 161-184.

Prof. Dr. Lisa Luft

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
Forschungszentrum FUTURE AGING
l.luft@fb4.fra-uas.de

Lebenslauf in Kürze

Frau Prof. Dr. Lisa Luft ist Diplom-Pflegewirtin, Gesundheits- und Krankenpflegerin und hat einen Masterstudiengang im Bereich Public Health abgeschlossen. Seit 2012 ist sie in vielfältigen Forschungsprojekten mit einem hohen Bezug zum praktischen Feld der Pflege tätig. Ein weiterer Fokus lag auf der Verbesserung der Versorgung von zu pflegenden Menschen im Setting der stationären Langzeitpflege. Frau Prof. Dr. Lisa Luft promovierte an der Philosophisch Theologischen Hochschule Vallendar und ist seit Sommer 2023 als Geschäftsführung des Forschungszentrums FUTURE AGING tätig. Auch die Förderung der Akademisierung der Pflege in Hessen ist eines ihrer zentralen Anliegen.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Erforschung praktischer Anwendungsmöglichkeiten und Erprobung von technischen Innovationen sowohl im interdisziplinären Forschungsteam FUTURE AGING als auch im Rahmen kritischer Reflexionen mit Studierenden. Zudem sind für mich qualitative Forschungsansätze zur Wahrung der Bedürfnisse von zu pflegenden Menschen und Anwender*innen von zentraler Bedeutung.



**Eva Lermer^{1,2}, Susanne Gaube³, Julia Cecil², Anne-Kathrin Kleine²,
Eesha Kokje², Dieter Frey², Matthias F. C. Hudecek^{4,5}**

Patient:innen und KI: Eine Frage der Perspektive bei der Bewertung von KI bei medizinischen Online-Diensten

Die rasante Entwicklung digitaler Technologien führt zu enormen Veränderungen im Gesundheitswesen (Gaube et al., 2023; Gaube & Lermer, 2023). Dazu gehören unter anderem das zunehmende Aufkommen von Online-Medizinplattformen, um medizinische Beratungen für Patient:innen anzubieten. Der Begriff Online-Medizinplattform bezieht sich im Allgemeinen auf alle Online-Gesundheitsdienste, die medizinische Beratung durch Informationstechnologie anbieten (Jiang et al., 2021). Diese Plattformen, manchmal auch als Telemedizin bezeichnet, bieten eine Vielzahl von Dienstleistungen an, einschließlich medizinischer Online-Konsultation, Beratung und Gesundheitsmanagement (z. B. El-Sherif et al., 2022; Liang et al., 2021). Online-Medizinplattformen bieten mehrere Vorteile, wie bequeme Zugänglichkeit und Kosteneffizienz im Vergleich zu Praxisbesuchen vor Ort (Bharti et al., 2020). Trotz der wachsenden Beliebtheit dieser Plattformen gibt es noch wenig Forschung dazu, wie Patient:innen diese akzeptieren und mit ihnen interagieren (Nadarzynski et al., 2020).

Auch hier kommt Künstliche Intelligenz (KI) immer mehr zum Einsatz (Haupt, 2019). Mit der zunehmenden Verbreitung von KI-Technologie im Gesundheitswesen ist es entscheidend, die Reaktionen der Patient:innen auf diese Dienste zu antizipieren, da angenommen werden kann, dass sie ein bedeutender Aspekt der medizinischen Versorgung werden (Richardson et al., 2022). Dies ist besonders kritisch, da Vorbehalte der Patient:innen gegenüber KI die Akzeptanz und Nutzung dieser Technologien behindern könnten. Studienergebnisse zu Anwendungen von KI zeigen, dass die Wahrnehmung der KI in der Öffentlichkeit von Sorgen über den Kontrollverlust über KI oder ethischen Bedenken bis hin zu Hoffnungen auf KI im Gesundheitswesen reichen kann (Fast

& Horvitz, 2017). Diese Ergebnisse unterstreichen die Bedeutung der Einbeziehung von Patient:innen, um sicherzustellen, dass KI-Technologie auf eine Weise in das Gesundheitswesen integriert wird, die öffentliches Vertrauen fördert und potenzielle Ängste lindert (Richardson et al., 2021).

Eine Reihe von Studien aus dem medizinischen Kontext deutet jedoch darauf hin, dass Menschen im Allgemeinen eine positive Einstellung gegenüber KI haben. In einer deutschen Studie beispielsweise bewerteten mehr als die Hälfte der Teilnehmer:innen (N = 462 Patient:innen) den Einsatz von KI in der Medizin als positiv oder sehr positiv, während nur ein kleiner Prozentsatz (4,77 %) negative oder sehr negative Wahrnehmungen äußerte. Die befragten Personen äußerten keine größeren Bedenken bezüglich KI, stimmten aber stark der Aussage zu, dass Ärzt:innen die Kontrolle über KI-Technologie behalten sollten (Fritsch et al., 2022). Mit einer qualitativen Studie (N = 48 Patient:innen aus Großbritannien), die sich auf Hautkrebs-Screening konzentrierte, kamen Nelson et al. (2020) zu ähnlichen Ergebnissen. Während 75 % der Patient:innen angaben, sie würden KI ihren Freunden und ihrer Familie empfehlen, betonten 94 % die Bedeutung einer symbiotischen Beziehung zwischen Mensch und KI (Nelson et al., 2020). Darüber hinaus zeigt ein kürzlich erschienener Übersichtsartikel, dass Patient:innen und die Öffentlichkeit den Einsatz von KI im Gesundheitswesen im Allgemeinen unterstützen. Viele Personen äußerten jedoch auch Vorbehalte und gaben einer menschlichen Betreuung den Vorzug (Young et al., 2021).

Zusammenfassend scheint es eine positive Einstellung gegenüber KI im Gesundheitswesen zu geben, aber

auch eine Präferenz für menschliche Ärzt:innen gegenüber KI-Technologie. Bislang noch nicht untersucht wurde jedoch, ob die allgemein positive Einstellung zu KI-Technologien im Gesundheitswesen in Situationen variiert, in denen Patient:innen entweder selbst von der Technologie betroffen sind oder wenn sie eine andere Person betrifft. Studien zur Risikobewertung haben gezeigt, dass Risikowahrnehmungen variieren, je nachdem, wer betroffen ist. Zum Beispiel fanden Lerner et al. (2013), dass Risiken, die einen selbst betreffen, als weniger wahrscheinlich wahrgenommen werden, als wenn sie andere Personen betreffen (z. B. einen durchschnittlichen Bürger). Eine Annahme ist, dass das Phänomen des unrealistischen Optimismus (Harris & Hahn, 2011) entscheidend für diesen Effekt ist. Getreu dem Motto „Mir wird schon nichts passieren“, besteht die Tendenz dazu, die Wahrscheinlichkeit, dass ein negatives Ereignis eintritt, für sich selbst als geringer einzuschätzen, als wenn es jemand anderen betrifft (Wills, 1981). Entsprechend der Construal Level Theory (CLT; Trope & Liberman, 2010) jedoch kann angenommen werden, dass Risiken, die einen selbst betreffen, höher eingeschätzt werden als Risiken, die andere betreffen, da hier die psychologische Distanz (PD) geringer ist. Zu den Annahmen der CLT zählt, dass die subjektive PD zu einem Ereignis oder Objekt beeinflusst, ob man darüber eher konkret oder abstrakt nachdenkt (Lerner et al., 2015, 2016b). Studien zur Risikowahrnehmung haben gezeigt, dass konkretes Denken zu höheren Risikobewertungen führt als abstraktes Denken (Lerner et al., 2016a). Es ist unklar, ob diese Forschungsergebnisse auf den Kontext von Online-Medizinplattformen angewendet werden können. Unrealistischer Optimismus würde darauf hindeuten, dass die Risikobewertung für einen selbst niedriger ist als für andere. Daher sollte, wenn von der Nutzung von KI ein Risiko erwartet wird, dieses Risiko für andere höher eingeschätzt werden als für einen selbst. Die gegenteilige Annahme würde jedoch gemäß der CLT gemacht werden.

Eine aktuelle Studie (Hudecek et al., 2023) fokussiert auf diese Differenzierung in der Perspektive und den Einfluss von KI mit folgenden zwei Forschungsfragen:

Welchen Einfluss haben die Perspektive (ich bin selbst betroffen vs. eine andere Person ist betroffen) und die Beratungsquelle (KI vs. menschliche:r Ärzt:in) auf die Bewertung der Diagnose, Behandlungsempfehlung und Risikowahrnehmung? Die Studienergebnisse zeigen, dass die Perspektive (selbst vs. andere Person) und die Beratungsquelle (KI vs. menschlicher Arzt/Ärztin) signifikante Einflüsse auf die Bewertung der Diagnose, der Behandlungsempfehlung und der Risikowahrnehmung haben. Patient:innen bevorzugen menschliche Ärzt:innen gegenüber KI wenn es um sie selbst geht, zeigen jedoch keine Präferenz wenn es um andere geht. Dies deutet auf eine kritischere Wahrnehmung von KI-gestützten Diagnosewerkzeugen bei persönlicher Betroffenheit hin.

Aus den Ergebnissen der Studie lässt sich schließen, dass die persönliche Betroffenheit eine Schlüsselrolle bei der Akzeptanz und dem Vertrauen in KI-gestützte medizinische Beratung spielt. Dies hat praktische Implikationen für die Entwicklung und Implementierung von KI in der Gesundheitsversorgung. Es zeigt, dass es wichtig ist, Vertrauen in KI-Systeme aufzubauen und die menschliche Interaktion in der medizinischen Beratung zu berücksichtigen, um eine breite Akzeptanz und effektive Nutzung dieser Technologien zu gewährleisten. Diese Erkenntnisse sind relevant für Anbieter von Gesundheitsdiensten, Entwickler von KI-Systemen und politische Entscheidungsträger im Gesundheitsbereich.

Diese Art der Forschung ist wichtig, da KI-fähige digitale Gesundheitstechnologie im Gesundheitswesen zunehmende Verbreitung findet (Kleine, Kokje, Lerner, et al., 2023; Kleine, Kokje, Hummelsberger, et al., 2023), aber vergleichsweise wenig darüber bekannt ist, wie Patient:innen die Nutzung dieser Werkzeuge sehen. Das Verständnis der Präferenzen und Bedenken der Patient:innen ist entscheidend für die Gestaltung KI-fähiger Werkzeuge, die ihren Bedürfnissen und Erwartungen entsprechen. Insgesamt zielt diese Forschung darauf ab, ein besseres Verständnis der Patient:innen auf moderne digitale Gesundheitstechnologie zu erlangen, indem relevante Faktoren identifiziert werden, die ihre Wahrnehmungen beeinflussen.

Literaturangaben:

- Bharti, U.; Bajaj, D.; Batra, H.; Lalit, S.; Lalit, S.; Gangwani, A.: Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19. 2020 5th International Conference on Communication and Electronics Systems (ICCES), 870–875, 2020. Online: <https://doi.org/10.1109/ICCES48766.2020.9137944>.
- El-Sherif, D. M.; Abouzid, M.; Elzarif, M. T.; Ahmed, A. A.; Albakri, A.; Alshehri, M. M.: Telehealth and Artificial Intelligence Insights into Healthcare during the COVID-19 Pandemic. *Healthcare*, 10(2), 385, 2022. Online: <https://doi.org/10.3390/healthcare10020385>.
- Fast, E.; Horvitz, E.: Long-Term Trends in the Public Perception of Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1), 2017. Online: <https://doi.org/10.1609/aaai.v31i1.10635>.
- Fritsch, S. J.; Blankenheim, A.; Wahl, A.; Hetfeld, P.; Maassen, O.; Deffge, S.; Kunze, J.; Rossaint, R.; Riedel, M.; Marx, G.; Bickenbach, J.: Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *DIGITAL HEALTH*, 8, 2022. Online: <https://doi.org/10.1177/20552076221116772>.
- Gaube, S.; Lermer, E.: KI-gestützte Entscheidungsunterstützungssysteme im Gesundheitswesen. In L. Kovács (ed.) *Künstliche Intelligenz und menschliche Gesellschaft*, De Gruyter Oldenbourg, 2023. Online: <https://doi.org/10.1515/9783111034706-009>.
- Gaube, S.; Suresh, H.; Raue, M.; Lermer, E.; Koch, T.K.; Hudecek, M.F.C.; Ackery, A.D.; Grover, S.C.; Coughlin, J.f.; Frey, D.; Kitamura, F.C.; Ghassemi, M.; Colak, E.: Non-task expert physicians benefit from correct explainable AI advice when reviewing X-rays. *Scientific reports*, 13(1), 1383, 2023.
- Harris, A. J. L.; Hahn, U.: Unrealistic optimism about future life events: A cautionary note. *Psychological Review*, 118(1), 135–154, 2011. Online: <https://doi.org/10.1037/a0020997>.
- Haupt, C.: Artificial Professional Advice. *Yale Journal of Health, Policy, Law, and Ethics*, 18(3), 55–77, 2019. Online: <https://ssrn.com/abstract=3400898>.
- Hudecek, M. F. C.; Lermer, E.; Gaube, S.; Kleine, K.; Kokje, E.; Cecil, J.; Batz, F.: Perception of online medical platforms: Testing the influence of the source of diagnosis (human vs. artificial intelligence) on different target persons (self vs. average person), 2023. Retrieved from osf.io/6kzxf
- Jiang, X.; Xie, H.; Tang, R.; Du, Y.; Li, T.; Gao, J.; Xu, X.; Jiang, S.; Zhao, T.; Zhao, W.; Sun, X.; Hu, G.; Wu, D.; Xie, G.: Characteristics of Online Health Care Services From China's Largest Online Medical Platform: Cross-sectional Survey Study. *Journal of Medical Internet Research*, 23(4), e25817, 2021. Online: <https://doi.org/10.2196/25817>.
- Kleine, A.-K.; Kokje, E.; Lermer, E.; Gaube, S.: Attitudes Toward the Adoption of 2 Artificial Intelligence-Enabled Mental Health Tools Among Prospective Psychotherapists: Cross-sectional Study. *JMIR Human Factors*, 10, e46859, 2023. Online: <https://doi.org/10.2196/46859>.
- Kleine, A.-K.; Kokje, E.; Hummelsberger, P.; Lermer, E.; Gaube, S.: AI-Enabled Clinical Decision Support Tools for Mental Healthcare: A Product Review. Preprint, 2023. Online: <https://osf.io/ez43g>.
- Kleine, A.-K.; Lermer, E.; Cecil, J.; Heinrich, A.; Gaube, S.: Advancing mental health care with AI-enabled precision psychiatry tools: A patent review. *Computers in Human Behavior Reports*, 12, 100322, 2023. Online: <https://doi.org/10.1016/j.chbr.2023.100322>.
- Lermer, E.; Streicher, B.; Sachs, R.; Frey, D.: How risky? The impact of target person and answer format on risk assessment. *Journal of Risk Research*, 16(7), 903–919, 2013. Online: <https://doi.org/10.1080/13669877.2012.761267>.
- Lermer, E.; Streicher, B.; Sachs, R.; Raue, M.; Frey, D.: The effect of construal level on risk-taking: Construal level and risk. *European Journal of Social Psychology*, 45(1), 99–109, 2015. Online: <https://doi.org/10.1002/ejsp.2067>.
- Lermer, E.; Streicher, B.; Sachs, R.; Raue, M.; Frey, D.: Thinking Concretely Increases the Perceived Likelihood of Risks: The Effect of Construal Level on Risk Estimation: Effect of Construal Level on Risk Estimates. *Risk Analysis*, 36(3), 623–637, 2016a. Online: <https://doi.org/10.1111/risa.12445>.
- Lermer, E.; Streicher, B.; Sachs, R.; Raue, M., & Frey, D. (2016b). The Effect of Abstract and Concrete Thinking on Risk-Taking Behavior in Women and Men. *SAGE Open*, 6(3), 215824401666612. Online: <https://doi.org/10.1177/2158244016666127>.
- Liang, X.; Kong, Y.; Peng, G.: Online Medical Platform Oriented Identification of User Needs Related to Alzheimer's Disease and Ontology Construction. In N. Streitz; S. Konomi (Eds.), *Distributed, Ambient and Pervasive Interactions*. Springer International Publishing, 2021, S. 316–330. Online: https://doi.org/10.1007/978-3-030-77015-0_23.
- Nelson, C. A.; Pérez-Chada, L. M.; Creadore, A.; Li, S. J.; Lo, K.; Manjaly, P.; Pournamdari, A. B.; Tkachenko, E.; Barbieri, J. S.; Ko, J. M.; Menon, A. V.; Hartman, R. I.; Mostaghimi, A.: Patient Perspectives on the Use of Artificial Intelligence for Skin Cancer Screening: A Qualitative Study. *JAMA Dermatology*, 156(5), 501, 2020. Online: <https://doi.org/10.1001/jamadermatol.2019.5014>.
- Richardson, J. P.; Curtis, S.; Smith, C.; Pacyna, J.; Zhu, X.; Barry, B.; Sharp, R. R.: A framework for examining patient attitudes regarding applications of artificial intelligence in healthcare. *DIGITAL HEALTH*, 8, 2022. Online: <https://doi.org/10.1177/20552076221089084>.
- Richardson, J. P.; Smith, C.; Curtis, S.; Watson, S.; Zhu, X.; Barry, B.; Sharp, R. R.: Patient apprehensions about the use of artificial intelligence in healthcare. *Npj Digital Medicine*, 4(1), 140, 2021. Online: <https://doi.org/10.1038/s41746-021-00509-1>.
- Trope, Y.; Liberman, N.: Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–463, 2010. Online: <https://doi.org/10.1037/a0018963>.
- Wills, T. A.: Downward comparison principles in social psychology. *Psychological Bulletin*, 90(2), 245–271, 1981. Online: <https://doi.org/10.1037/0033-2909.90.2.245>.
- Young, A. T.; Amara, D.; Bhattacharya, A.; Wei, M. L.: Patient and general public attitudes towards clinical artificial intelligence: A mixed methods systematic review. *The Lancet Digital Health*, 3(9), e599–e611, 2021. Online: [https://doi.org/10.1016/S2589-7500\(21\)00132-1](https://doi.org/10.1016/S2589-7500(21)00132-1).

Prof. Dr. habil. Eva Lermer

Technical University of Applied Sciences Augsburg
Department of Business Psychology
eva.lermer@tha.de

Lebenslauf in Kürze

Eva Lermer ist Psychologin und Soziologin und hat zum Thema Entscheidungen und Umgang mit Risiken in Kooperation mit der Munich Re an der Ludwig-Maximilians-Universität (LMU) München promoviert und habilitiert. Sie ist Forschungsprofessorin an der Technischen Hochschule Augsburg (THA) an der sie das Projekt PSY-A-EYE leitet. Hierbei geht es um die Erforschung der Mensch-KI-Interaktion im Bereich der Augenheilkunde. Parallel zu ihrer Tätigkeit an der THA leitet Eva Lermer zusammen mit Dr. Anne-Kathrin Kleine das von der VolkswagenStiftung geförderte Projekt „Human-AI-Interaction in Healthcare“ an der LMU München. Ziel dieser Forschung liegt in der Optimierung der Zusammenarbeit mit KI.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... die Art und Weise zu erforschen und zu gestalten, wie Menschen und KI-Systeme interagieren, um die menschliche Entscheidungsfindung zu unterstützen und zu optimieren.



Anne-Kathrin Kleine

Neue methodische Ansätze zur Erforschung von Trends in der KI-gestützten mentalen Gesundheitsversorgung

Etwa ein Drittel aller Menschen weltweit leidet irgendwann im Leben an einer psychischen Erkrankung, und nur ein Drittel der Menschen mit psychischen Problemen erhält eine adäquate Behandlung. Die fehlende zeitnahe Diagnose und angemessene frühe Intervention bei psychischen Erkrankungen können zu einer erheblichen Verschlimmerung des Krankheitszustands, langen Genesungszeiten, einer verminderten Lebensqualität und insgesamt zu hohen Gesundheitskosten führen. Die Nutzung von Künstlicher Intelligenz (KI) in der psychischen Gesundheitsversorgung verspricht, einigen dieser Engpässe zu begegnen. KI-gesteuerte Algorithmen können beispielsweise diagnostische und therapeutische Empfehlungen auf Basis von Mustern in Sprach- und Videodaten erstellen. So kann die Nutzung verschiedener Arten von Patient:innendaten zu einer frühen Erkennung psychischer Erkrankungen und einer zeitnahen Intervention beitragen (Graham et al., 2019).

KI-gestützte Klinische Entscheidungsunterstützungssysteme (KI-KEUS) können in der klinischen Praxis zur Verbesserung der Patient:innenversorgung eingesetzt werden (Gaubé & Lermer, 2023), beispielsweise bei der Erkennung oder Vorhersage psychischer Erkrankungen oder der Identifizierung optimaler Behandlungsansätze. Trotz des Potenzials von KI-KEUS für die Verbesserung psychiatrischer und psychotherapeutischer Versorgung kommen sie in der klinischen Praxis noch wenig zum Einsatz. Es ist unklar, an welchem Punkt der Einsatzbereitschaft von KI-KEUS wir aktuell stehen. Werden KI-KEUS nur in der Forschung behandelt? Gibt es überhaupt kommerziell verfügbare KI-KEUS, die bereits zur klinischen Entscheidungsfindung genutzt werden können?

Es existieren einige wissenschaftliche Übersichtsarbeiten zu KI-KEUS für die psychische Gesundheits-

versorgung. So haben Tornero-Costa et al. (2023) 129 wissenschaftliche Artikel identifiziert, in denen KI-Systeme zur Diagnose und Behandlungsauswahl bei psychischen Erkrankungen entwickelt wurden. Um einen genauen Überblick über den Stand und die Anwendung von KI-KEUS in der psychischen Gesundheitsversorgung zu erhalten, müssen jedoch auch Datenquellen aus der Praxis genutzt werden. So können beispielsweise Patentdaten Aufschluss über neue Technologien und Anwendungsverfahren geben. Sie beinhalten zudem technische Informationen, die in wissenschaftlichen Artikeln nicht verfügbar sind. Zusätzlich sind Daten von Anbieterseiten und offiziellen Regulationsdokumentationen relevant, um einen Überblick über kommerziell verfügbare Produkte zu erhalten.

Die Nutzung von Patent- und Anbieterdaten zur Erforschung des Entwicklungsstandes klinischer Entscheidungsunterstützungssysteme für die psychische Gesundheitsversorgung

Im Rahmen eines Patent-Reviews wurde eine umfassende Suche patentierter KI-KEUS in der Derwent Innovation (DI) Patent-Datenbank durchgeführt (Kleine et al., 2023b). Insgesamt konnten 279 relevante Patente gefunden werden. Um einen Überblick über das technologische Spektrum der Patente zu erhalten, wurde eine Konvergenzanalyse auf der Grundlage von Cooperative Patent Classification (CPC) Codes durchgeführt. Das CPC-System ist hierarchisch aufgebaut und gruppiert Patente in verschiedene Technologiebereiche. So wird zum Beispiel die Technologie-Subklasse G16H "Gesundheitsinformatik" der Technologie-Klasse G16 "Informations- und Kommunikationstechnologie für verschiedene Anwendungsfelder" zugeordnet. Für die Konvergenzanalyse wurden die Verbindungen zwischen den CPC-Kodierungen in-

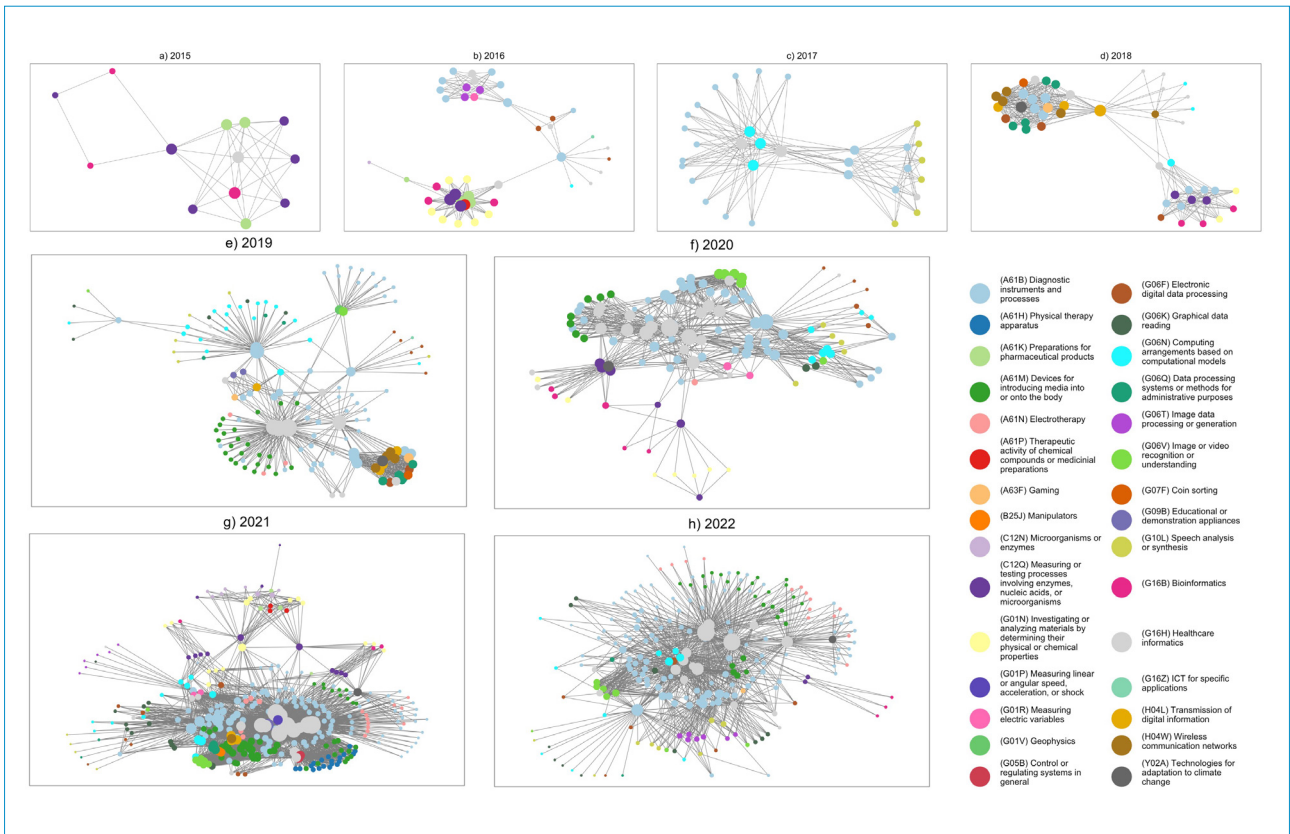


Abbildung 1: Die Abbildung zeigt die Verbindungen zwischen CPC-Unterklassen innerhalb von Patenten. Die Größe der gefüllten Kreise gibt Aufschluss über die Häufigkeit einer Unterklasse über alle Patente hinweg. Je näher zwei Kreise beieinander liegen, desto häufiger wurden ihre jeweiligen Technologie-Unterklassen zusammen erwähnt. Abbildung übernommen aus Kleine et al. (2023b).

nerhalb von Patenten grafisch dargestellt. Das Ergebnis der Konvergenzanalyse ist in Abbildung 1 zu sehen. Wie in Abbildung 1 zu sehen ist, ist die Rate der Technologiekonvergenz nach 2015 deutlich gestiegen. In der Anfangsphase der Entwicklung von KI-KEUS (2015 bis 2016) sind wenige Technologie-Unterklassen und wenige Verbindungen zwischen Technologien zu erkennen. Das Tempo der Konvergenz verschiedener Technologien erhöht sich in der nachfolgenden Wachstumsphase (ab 2016). Die Entwicklung von KI-KEUS wird durch die Einführung neuer Technologien und neuer Verbindungen zwischen existierenden Technologien vorangetrieben. Die Häufigkeit des Auftretens einzelner und die Verbindungen zwischen den Technologie-Unterklassen ermöglichen Rückschlüsse auf Schlüsseltechnologien und Innovationstrends. Ergänzend zur grafischen Darstellung der Konvergenzanalyse können die Beschreibungen von Patenten mittels Methoden zur Verarbeitung natürlicher Sprache analysiert werden, um die Relevanz spezifischer Themen und ihre zeitlichen Entwicklungen besser zu verstehen (siehe Kleine et al., 2023b).

Basierend auf den Daten aus dem Patent-Review und Online-Recherchen haben Kleine et al. (2023a) sieben KI-KEUS identifiziert, die reguliert und aktuell kommerziell verfügbar sind. Aufgrund ihrer Eigenschaft als Medizinprodukt müssen KI-KEUS von offiziellen Regulierungsbehörden begutachtet und als sicher und zuverlässig eingeordnet werden, bevor sie kommerziell verfügbar gemacht werden dürfen. Von den sieben Produkten ließen sich fünf Patenten ordnen. Abbildung 2 verdeutlicht, dass es Überschneidungen zwischen Forschungsarbeiten, Patentdaten und kommerziell verfügbaren Produkten gibt. Viele Forschungsarbeiten werden jedoch nie patentiert oder kommerzialisiert. Zudem sind nicht alle verfügbaren KI-KEUS patentiert und nicht alle patentierten Innovationen werden kommerzialisiert.

Die ermittelten Produkte lassen sich drei Hauptkategorien zuordnen: Die Diagnose von Autismus-Spektrum-Störungen bei Kindern, basierend auf Verhaltensdaten, die Diagnose unterschiedlicher mentaler Störungen mittels Analyse von Daten aus Chatbot-Kon-

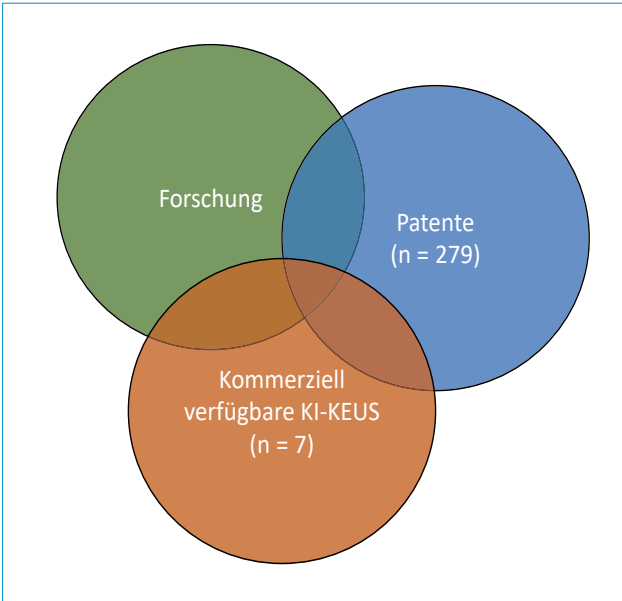


Abbildung 2: Venn Diagramm zur Verdeutlichung der Überschneidungen und Diskrepanzen zwischen Forschungsarbeiten, Patenten und kommerziell verfügbaren Produkten im Bereich KI-gestützter klinische Entscheidungsunterstützungssysteme (KI-KEUS) in der psychischen Gesundheitsversorgung.

versationen und schließlich die Auswahl von Medikamenten zur Behandlung von Depressionen, gestützt auf klinische und genetische Daten. Abbildung 3 bietet eine Darstellung der Daten zum Markteintritt der Anbieter, der Daten der Produktzertifizierungen und der Veröffentlichungen von wissenschaftlichen Arti-

keln zur externen Validierung. Es ist zu beobachten, dass der zeitliche Abstand zwischen Firmengründung und Kommerzialisierung über die Jahre scheinbar abnimmt. Dies könnte auf die Tatsache zurückzuführen sein, dass neuere Produkte bereits existierende ähnliche Produkte als Referenz nutzen und dadurch schneller reguliert werden können. Zusätzlich wurden Pressemitteilungen und Online-Artikel zu nicht patentierten KI-KEUS untersucht. Dabei zeigt sich, dass zahlreiche KI-KEUS keine Zulassung zur kommerziellen Nutzung erhalten haben, weil klinische Studien und allgemein der Nachweis ihrer Überlegenheit im Vergleich zu herkömmlichen Diagnose- und Therapieverfahren noch fehlten (Kleine et al., 2023a).

Fazit

Die systematische Nutzung von Patent- und Anbieterdaten kann wertvolle Einblicke in den aktuellen Stand und die zukünftigen Trends in der Entwicklung von KI-KEUS für die psychische Gesundheitsversorgung bieten. Die Verwendung unterschiedlicher Datenquellen und der Einsatz vielfältiger Methoden zur Datenaufbereitung können wichtige Informationen für Forschende, Praktiker:innen und Entscheidungsträger:innen liefern. Die Nutzung unterschiedlicher Datenquellen ermöglicht das Erkennen potenzieller Innovationsfelder und die

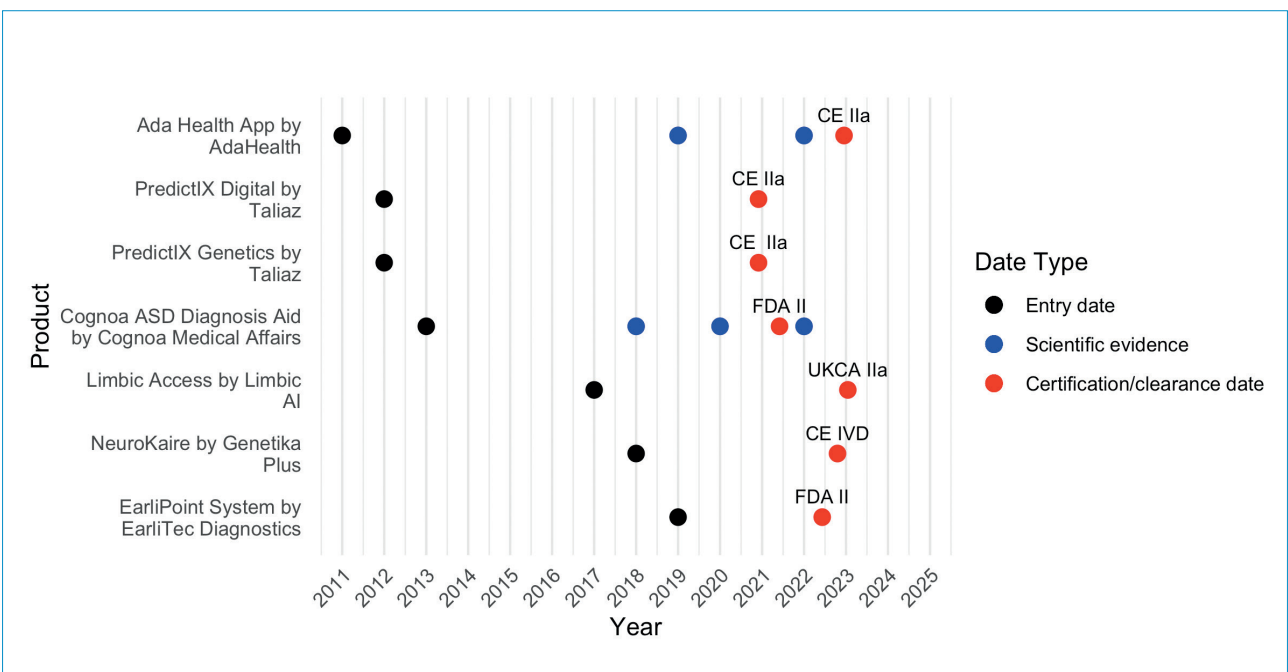


Abbildung 3: Die Abbildung zeigt das Datum des Markteintritts, der Veröffentlichung wissenschaftlicher Evidenz und der Zertifizierung für alle identifizierten KI-KEUS für die psychische Gesundheitsversorgung. Abbildung übernommen aus Kleine et al. (2023a).

Identifikation von Herausforderungen im Kommerzialisierungsprozess. Darüber hinaus können die so gewonnenen Erkenntnisse dazu beitragen, effektive und regulierte Anwendungen ausfindig zu machen. Insgesamt kann die kombinierte Betrachtung und

Analyse von Forschungs- und Praxisdaten dabei helfen, die Bedürfnisse beider Bereiche besser zu verstehen. Dies kann schlussendlich dazu führen, dass KI-KEUS langfristig effizienter in der psychischen Gesundheitsversorgung angewendet werden können.

Literaturangaben:

Gaube, S.; Lerner, E.: KI-gestützte Entscheidungsunterstützungssysteme im Gesundheitswesen. 2023, S. 121–136. Online: <https://doi.org/10.1515/9783111034706-009>.

Graham, S. A.; Depp, C. A.; Lee, E. E.; Nebeker, C.; Tu, X. M.; Kim, H.-C.; Jeste, D. V.: Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports*, 21(11). 2019. Online: <https://doi.org/10.1007/s11920-019-1094-0>.

Kleine, A.-K.; Kokje, E.; Hummelsberger, P.; Lerner, E.; Gaube, S.: AI-Enabled Clinical Decision Support Tools for Mental Healthcare: A Product Review. 2023a, December 9. Online: <https://osf.io/ez43g>.

Kleine, A.-K.; Lerner, E.; Cecil, J.; Heinrich, A.; Gaube, S.: Advancing mental health care with AI-enabled precision psychiatry tools: A patent review. *Computers in Human Behavior Reports*, 12, 100322. 2023b. Online: <https://doi.org/10.1016/j.chbr.2023.100322>.

Tornero-Costa, R.; Martinez-Millana, A.; Azzopardi-Muscat, N.; Lazari, L.; Traver, V.; Novillo-Ortiz, D.: Methodological and quality flaws in the use of artificial intelligence in mental health research: Systematic review. *JMIR Mental Health*, 10(1), e42045. 2023. Online: <https://doi.org/10.2196/42045>.

Dr. Anne-Kathrin Kleine

Ludwig-Maximilian-Universität München
Department Psychologie
Anne-Kathrin.Klein@psy.lmu.de

Lebenslauf in Kürze

Anne-Kathrin Kleine ist Psychologin und hat an der Universität Groningen zum Umgang von Menschen in beruflichen Umbruchsphasen promoviert. Derzeit konzentriert sie sich auf die Untersuchung der Interaktion von Menschen mit künstlicher Intelligenz in verschiedenen beruflichen Kontexten. Sie leitet als Postdoktorandin gemeinsam mit Eva Lerner das von der Volkswagen-Stiftung geförderte Projekt „Human-AI-Interaction in Healthcare“ an der Ludwig-Maximilians-Universität München. Neben ihrer Forschungsarbeit setzt Frau Kleine sich aktiv für die Implementierung von Open-Science-Prinzipien in der Wissenschaft ein und bietet Workshops zur Datenanalyse mit der Programmiersprache R sowie der Replizierbarkeit von Forschungsergebnissen in den Sozialwissenschaften an.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegt darin, die komplexen Abläufe in der Interaktion zwischen Menschen und KI-Systemen besser zu verstehen. Dabei beschäftige ich mich intensiv damit, wie diese Technologien die menschliche Entscheidungsfindung optimieren können, ohne dabei die kritischen Aspekte aus den Augen zu verlieren. Wo liegen die Grenzen der Automatisierung, wo die Risiken und welche neuen Möglichkeiten bieten sich uns?

Julia Cecil¹, Eesha Kokje¹, Susanne Gaube², Eva Lermer^{1,3}, Anne-Kathrin Kleine¹

Einflussfaktoren auf die Nutzung KI-gestützter Technologien in der psychischen Gesundheitsversorgung

Die fortschreitende Entwicklung künstlicher Intelligenz (KI) hat in den letzten Jahren auch im Gesundheitswesen Einzug gehalten. Im Bereich der psychischen Gesundheitsversorgung eröffnen KI-gestützte Technologien neue Möglichkeiten für eine verbesserte Patient:innenversorgung, die besonders vor dem Hintergrund der vorhandenen Versorgungsengpässe von Bedeutung ist: Die durchschnittliche Wartezeit auf einen Therapieplatz in der ambulanten Versorgung in Deutschland beträgt knapp fünf Monate (Deutscher Bundestag, 2022). Obwohl KI-gestützte Technologien dieses Problem nicht lösen können, können sie doch erste Abhilfe schaffen, indem sie bestimmte Arbeitsprozesse erleichtern oder übernehmen und so den bestehenden Engpässen entgegenwirken (Blease et al., 2020).

1. Einsatzbereiche KI-gestützter Systeme in der psychischen Gesundheitsversorgung

Die Einsatzbereiche KI-gestützter Systeme in der psychischen Gesundheitsversorgung sind vielfältig. Sie reichen von der Diagnose und Behandlung psychischer Erkrankungen bis hin zu Feedback-Systemen für Anwender:innen sowie der Übernahme organisatorischer und administrativer Aufgaben durch Praxismanagement-Systeme (Kellog & Sadeh-Shavitt, 2022; Kleine, Kokje, Lermer, et al., 2023; Kleine Kokje, Hummelsberger, et al., 2023). Die zugrundeliegenden Daten werden dabei vor allem durch Verhaltensanalysen (z.B. Video- und Audioaufzeichnungen), biologische Messungen (z.B. Blutproben, Herzrate) und bildgebende Verfahren (z.B. EEG oder MRT) gewonnen (Kleine, Lermer, Cecil, et al., 2023). Grundsätzlich lässt sich zwischen Tools, die sich auf die Diagnose und Therapie von Patient:innen (sog. patient:innenorientierte Tools) und solchen, die sich auf die Verbesserung der Qualität therapeuti-

scher Leistungen (sog. therapeut:innenorientierten Tools) konzentrieren, unterscheiden.

1.1 Patient:innenorientierte Tools in der psychischen Gesundheitsversorgung

KI-gestützte Anwendungen, die primär auf Patient:innen ausgerichtet sind, umfassen Diagnose- und Behandlungstools. In der Diagnostik haben KI-gestützte Anwendungen das Potenzial, die Genauigkeit und Effizienz von Diagnosen zu verbessern, indem sie eine Vielzahl von Patient:innendaten berücksichtigen (Kellog & Sadeh-Shavitt, 2022). Auch die Behandlung von psychischen Erkrankungen kann von KI-gestützten Anwendungen profitieren, da individuellere und präzisere Behandlungsvorhersagen möglich werden (Aafjes-van Dorn et al., 2020). Im Bereich depressiver und anderer affektiver Störungen wurden Systeme entwickelt, die durch Sprachdaten die Stimmungslage der Patient:innen erfassen können (DeSouza et al., 2021; Huang et al., 2020). Wie in Abbildung 1 dargestellt, werden bei diesen Systemen kurze Sprachaufnahmen von Patient:innen über ein Smartphone aufgenommen, die im Anschluss mithilfe von Algorithmen auf bestimmte Merkmale analysiert werden. Diese Merkmale wie Tonlage, Sprechtempo, Wortwahl und Pausen können als Indikatoren für emotionale Zustände dienen. Anhand dieser Analyse wird ein Stimmungswert erzeugt, der an den Behandelnden zurückgemeldet wird. Auf Basis dessen können behandlungsbezogene Entscheidungen getroffen werden (Sonde Health, 2023).

1.2 Therapeut:innenorientierte Tools in der psychischen Gesundheitsversorgung

Neben patient:innenorientierten Anwendungen existieren auch Systeme, die sich auf die Behandelnden konzentrieren. Dazu gehören Anwendungen im Be-

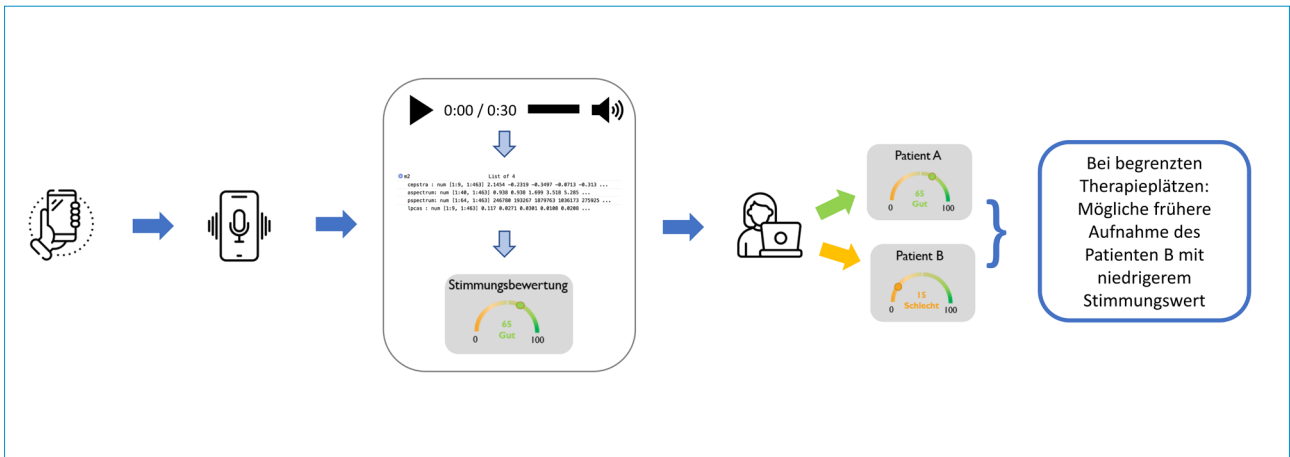


Abbildung 1: Schematische Darstellung der Funktionsweise eines KI-gestützten Behandlungstools in der psychischen Gesundheitsversorgung.

reich des Praxismanagements und Feedback-Tools. Erstere können klinische und administrative Arbeitsabläufe durch die Transkription von Therapiesitzungen und Integration dieser in Krankenakten automatisieren und so den Verwaltungsaufwand verringern (Sadeh-Sharvit et al., 2023). Feedback-Tools zielen darauf ab, den Praktizierenden Feedback über die Qualität ihrer Interaktion mit den Patient:innen zu geben, indem Sprachdaten aus den Gesprächen zwischen Therapeut:innen und Patient:innen analysiert werden, siehe Abbildung 2 (Cummins et al, 2019; Hirsch et al., 2018; Sadeh-Shavitt, 2023). Die gewonnenen Audioaufnahmen werden auf bestimmte Merkmale hin analysiert, zum Beispiel, um festzustellen, ob Komponenten einer effektiven psychotherapeutischen Gesprächsmethode, wie dem Motivational Interviewing (MI), zum Einsatz kamen. Die Praktizierenden erhalten so wichtige Hinweise zur Verbesserung ihrer Therapie (Cummins et al., 2019).

2. Herausforderungen und Limitationen der Nutzung KI-gestützter Systeme

Trotz der vielversprechenden Chancen KI-gestützter Technologien gibt es immer noch Herausforderungen und offene Fragen, die einer sicheren Implementierung entgegenstehen. Der Einsatz der Technologien ist weniger fortgeschritten als in anderen medizinischen Fachbereichen, u.a. durch die Abhängigkeit der Behandlungsergebnisse von der Beziehung zwischen Praktizierenden und Patient:innen (Jin et al., 2023). Offene Fragen bezüglich ethischer und regulatorischer Bedenken (Wer haftet bei Fehlern?), Datenqualität (Sind die Daten, mit denen das System trainiert wurde, repräsentativ für die klinische Gesamtstichprobe?), Datenschutz (Wie wird mit sensiblen Patient:innendaten umgegangen?) und Übertragbarkeit (Sind die Systeme in den alltäglichen klinischen Alltag integrierbar?) sind noch zu klären (Aung et al., 2021). Neben dieser allgemein

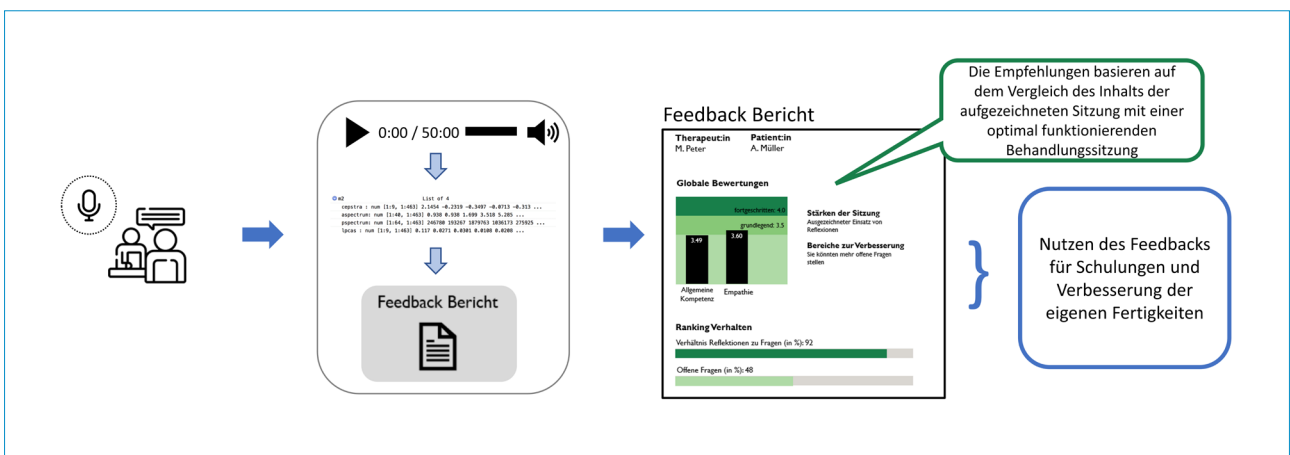


Abbildung 2: Schematische Darstellung der Funktionsweise eines KI-gestützten Feedback-Tools in der psychischen Gesundheitsversorgung.

geringeren Verfügbarkeit und den verschiedenen Systemfaktoren ist auch die fehlende allgemeine Akzeptanz der KI-Systeme aufgrund der Skepsis der Anwender:innen eine Herausforderung (Blease et al., 2020). Hier spielen individuelle Faktoren auf Seiten der Anwender:innen eine wichtige Rolle.

3. Einflussfaktoren auf die Nutzung KI-gestützter Systeme

Die Nutzungsbereitschaft von Anwender:innen kann durch verschiedene individuelle Einflussfaktoren erklärt werden, u.a. durch die folgenden vier Hauptfaktoren der „Unified Theory of Acceptance and Use of Technology“ (UTAUT), die aus verschiedene Theorien integriert wurden (Venkatesh et al., 2003).

1. Die Leistungserwartung beschreibt dabei die Erwartung einer Person, dass die Nutzung der Technologie ihre Leistung oder Produktivität verbessern wird. Konkret könnte das bedeuten, dass ein:e Psychotherapeut:in erwartet, dass die Verwendung eines Tools zur Patient:innendokumentation die eigene Produktivität steigert, indem die Technologie die Sitzung automatisch dokumentiert und so das zeitaufwendige Übertragen des Sitzungsinhaltes in die Patientenakten abgenommen wird.
2. Die Aufwandserwartung beschreibt die Wahrnehmung einer Person über die Leichtigkeit oder Schwierigkeit der Nutzung einer neuen Technologie. Wenn eine neue Technologie leicht zu bedienen ist, ist die Wahrscheinlichkeit auch höher, dass sie angenommen und genutzt wird. Der:Die Psychotherapeut:in schätzt das neue System zur Patientendokumentation als benutzerfreundlich ein und erwartet daher, dass diese intuitive Plattform die Arbeitsbelastung verringert und somit mehr Zeit für direktere Interaktionen mit den Patient:innen schafft.
3. Der Soziale Einfluss bezieht sich auf den Einfluss, den die Meinungen und Unterstützung von anderen, wie beispielsweise Kolleg:innen, auf die Akzeptanz und Nutzung der Technologie haben. Die positiven Rückmeldungen der Kolleg:innen zur Nutzung dieser spezifischen Anwendung haben einen Einfluss darauf, wie der:Die Psychotherapeut:in die Technologie wahrnimmt und tragen maßgeblich zur Entscheidung bei, diese in der eigenen Praxis zu nutzen.

4. Erleichternde Bedingungen meint die wahrgenommene Unterstützung und Ressourcen, die verfügbar sind, um die Nutzung der Technologie zu erleichtern. Dazu zählen beispielsweise technische Unterstützung, Schulungen oder die Verfügbarkeit von Ressourcen. Der:Die Psychotherapeut:in erkennt die Verfügbarkeit von Schulungen und Unterstützung durch den technischen Support als erleichternde Bedingungen an. Diese Ressourcen werden als bedeutend angesehen, da sie das Vertrauen in eine erfolgreiche Nutzung der Technologie und den Erhalt von Hilfestellungen bei Problemen stärken.

4. Einflussfaktoren auf die Nutzungsabsicht der Praktizierenden am Beispiel KI-gestützter Behandlungs- und Feedback-Tools

Das UTAUT wird im Bereich des Gesundheitswesens häufig verwendet (Rouidi et al., 2022). Im Bereich der psychischen Gesundheitsversorgung untersuchte eine von uns durchgeführte online-Studie auf Basis des UTAUTs die Nutzungsabsichten von insgesamt 206 deutsch- und englischsprachigen Psychologiestudierenden und Psychotherapeut:innen in Ausbildung (Kleine, Kokje, Lerner, et al., 2023). Die Teilnehmenden erhielten eine kurze theoretische Einführung in das Behandlungs- und Feedback-Tool und wurden im Anschluss gebeten, für beide Tools Fragen aus dem UTAUT zu beantworten. Zusätzlich zu den vier Hauptfaktoren wurden in die Studie die beiden Faktoren Vertrauen und Wissen über KI und als Kontrollvariablen Datenschutzbedenken und Ängstlichkeit gegenüber KI-Systemen erhoben. Ergebnisse der Studie zeigten, dass die wahrgenommene Nützlichkeit und der soziale Einfluss mit den Nutzungsabsichten für beide Tools in einem positiven Zusammenhang stand. Die Leistungserwartung hatte keinen Einfluss auf die Nutzungsabsicht des Feedback-Tools, sie war jedoch mit der Nutzungsabsicht des Behandlungstools negativ assoziiert. Die Ängstlichkeit gegenüber KI-Systemen war negativ mit den Nutzungsabsichten für beide Tools assoziiert. Vertrauen und Datenschutzbedenken hatten bei beiden Tools keinen Einfluss auf die Nutzungsabsichten.

Die Ergebnisse der Studie zeigen, dass die Akzeptanz und Nutzungsabsichten KI-gestützter Systeme stark

von den Einflüssen des Arbeitsumfelds abhängig sind. Vor allem Berufseinsteiger:innen sind in ihren berufsbezogenen Entscheidungen möglicherweise noch nicht gefestigt und scheinen diese eher an den Erwartungen und Entscheidungen anderer auszurichten (Gado et al., 2022; Tran et al., 2021). Der negative bzw. fehlende Einfluss der Aufwandserwartung auf die Nutzungsabsicht der KI-gestützten Systeme lässt den Schluss zu, dass die wahrgenommene Benutzerfreundlichkeit in der theoretischen Interaktion eine eher untergeordnete Rolle spielt und erst in der direkten Verwendung eine tragende Rolle einnimmt. Auch der fehlende Zusammenhang zwischen Vertrauen und den Nutzungsabsichten und der negative Zusammenhang mit Ängstlichkeit gegenüber KI-Systemen lässt sich ähnlich erklären: Vertrauen in die Systeme erfordert ein tiefergehendes Verständnis, die Teilnehmenden haben den Einsatz des Tools jedoch nur in der Theorie und nicht in der Praxis kennengelernt. Daher könnte die intuitive, emotionale Reaktion der Ängstlichkeit eine größere Rolle in der Vorhersage der Nutzungsabsichten als das Vertrauen gespielt haben (Kwak et al., 2022).

Datenschutzbedenken korrelierten sowohl negativ mit dem Vertrauen als auch der wahrgenommenen Benutzerfreundlichkeit der beiden Tools, der direkte Effekt auf die Nutzungsabsichten blieb jedoch aus. Im Hinblick auf weitere Forschung, die zeigte, dass das Vertrauen in den Datenschutz bei KI-basierter Therapie im Vergleich zu Therapeuten geringer war (Aktan et al., 2022), wird deutlich, dass Datenschutzbedenken nach wie vor zu den größten Herausforderungen beim Einsatz neuer Technologien gehören (Grande et al., 2020). Zusätzlich spielt auch das Wissen über KI eine Rolle bei den Nutzungsabsichten. Ein höheres spezifisches Verständnis des jeweiligen Tools war bei beiden Tools mit einer höheren wahrgenommenen Benutzerfreundlichkeit assoziiert. Zudem gab es einen positiven Zusammenhang zwischen dem allgemeinen Wissen über KI und der Nutzungsabsicht des Feedback-Tools. Hier wurden durch das grundlegende KI-Wissen die Grenzen und die Vorteile des Tools möglicherweise ausreichend erkannt. Für Anwendungen mit direktem Patient:innenkontakt, wie dem Behandlungs-Tool, scheint hingegen ein tiefgreifenderes Verständnis oder die direkte Interaktion mit der Anwendung erforderlich zu sein.

5. Implikationen und Fazit

Grundsätzlich lassen sich folgende Schlussfolgerungen zur Akzeptanz und Nutzung KI-gestützter Systeme in der psychischen Gesundheitsversorgung ziehen:

1. Vielfältige Einsatzbereiche, andauernde Herausforderungen: Die Anwendungsmöglichkeiten von KI-gestützten Systemen sind breit gefächert und bieten das Potenzial, um eine Verbesserung in der psychischen Gesundheitsversorgung von mehreren Seiten bewirken zu können. Dennoch stehen der Nutzung und der sicheren Implementierung dieser KI-Technologien noch viele Herausforderungen entgegen. Dazu zählen eine allgemeine fehlende Akzeptanz, ethische Bedenken und Datensicherheit.
2. Unterschiedliche Technologien, unterschiedliche Akzeptanz: Es ist wichtig zu erkennen, dass sich der Einsatz verschiedener Technologien unterschiedlich auf Patient:innen und Anwender:innen auswirkt und mit unterschiedlichen Voraussetzungen einhergeht. Daher ist es ratsam, nicht pauschal von einer Anwendung auf eine andere zu schließen, sondern sich differenziert mit den Merkmalen und Anwendungsbereichen verschiedener Tools auseinanderzusetzen, um ein umfassendes Verständnis und eine effektive und sichere Integration von KI-gestützten Technologien in der psychischen Gesundheitsversorgung gewährleisten zu können.
3. Zielgruppen genau betrachten: Unterschiedliche Zielgruppen könnten unterschiedlich auf KI-gestützte Systeme reagieren. Es ist daher wichtig, die Bedürfnisse und Erwartungen verschiedener Gruppen (z.B. Berufseinsteiger:innen vs. erfahrenes Fachpersonal) zu berücksichtigen, um gezielte Einführungs- und Schulungsstrategien zu entwickeln.
4. Erfahrung und Interaktion als Grundlage für Vertrauen: Theoretische Kenntnisse über KI-gestützte Tools bilden zwar die Grundlage, aber das tatsächliche Vertrauen und Verständnis entwickelt sich durch direkte Erfahrungen im Umgang mit KI-gestützten Technologien. Das Erleben und Ausprobieren der Tools in sicherer Umgebung kann Ängste mindern und ein tieferes Verständnis für deren Wirksamkeit schaffen.

Abschließend ist anzumerken, dass trotz der wachsenden Forschung zu KI-gestützten Technologien eine stärkere Verknüpfung von Theorie und Praxis notwendig ist, um Bedarfe zu erkennen und Hindernissen bei der Im-

plementierung in die klinische Praxis aktiv begegnen zu können. Erst dann kann der angestrebte Nutzen einer effektiven und sicheren Integration dieser Technologien in der psychische Gesundheitsversorgung erreicht werden.

Literaturangaben:

- Aafjes-van Doorn, K.; Kamsteeg, C.; Bate, J.; Aafjes, M.: A scoping review of machine learning in psychotherapy research. *Psychotherapy Research*, 31(1). 2021, S. 92–116. Online: <https://doi.org/10.1080/10503307.2020.1808729>.
- Aktan, M. E.; Turhan, Z.; Dolu, İ.: Attitudes and perspectives towards the preferences for artificial intelligence in psychotherapy. *Computers in Human Behavior*, 133, 107273. 2022. Online: <https://doi.org/10.1016/j.chb.2022.107273>.
- Aung, Y. Y. M.; Wong, D. C. S.; Ting, D. S. W.: The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare. *British Medical Bulletin*, 139(1). 2021, S. 4–15. Online: <https://doi.org/10.1093/bmb/ldab016>.
- Blease, C.; Locher, C.; Leon-Carlyle, M.; Doraiswamy, M.: Artificial intelligence and the future of psychiatry: Qualitative findings from a global physician survey. *Digital Health*, 6, 2055207620968355. 2020. Online: <https://doi.org/10.1177/2055207620968355>.
- Cummins, R.; Ewbank, M. P.; Martin, A.; Tablan, V.; Catarino, A.; Blackwell, A. D.: TIM: A Tool for Gaining Insights into Psychotherapy. [Conference presentation]. WWW ,19: The Web Conference. San Francisco, CA, USA 2019, May 13-17. Online: <https://doi.org/10.1145/3308558.3314128>.
- DeSouza, D. D.; Robin, J.; Gumus, M.; Yeung, A.: Natural Language Processing as an Emerging Tool to Detect Late-Life Depression. *Frontiers in Psychiatry*, 12, 719125. 2021. Online: <https://doi.org/10.3389/fpsy.2021.719125>.
- Deutscher Bundestag: Wartezeiten auf eine Psychotherapie—Studien und Umfragen. 2022. Online: <https://www.bundestag.de/resource/blob/916578/53724d526490deea69f736b1fda83e76/WD-9-059-22-pdf-data.pdf>. [abgerufen am 04.12.2023].
- Gado, S.; Kempen, R.; Lingelbach, K.; Bipp, T.: Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence? *Psychology Learning & Teaching*, 21(1). 2022, S. 37–56. Online: <https://doi.org/10.1177/14757257211037149>.
- Grande, D.; Luna Marti, X.; Feuerstein-Simon, R.; Merchant, R. M.; Asch, D. A.; Lewson, A.; Cannuscio, C. C.: Health Policy and Privacy Challenges Associated With Digital Technology. *JAMA Network Open*, 3(7), e208285. 2020. Online: <https://doi.org/10.1001/jamanetworkopen.2020.8285>.
- Hirsch, T.; Soma, C.; Merced, K.; Kuo, P.; Dembe, A.; Caperton, D. D.; ... & Imel, Z. E.: „It's hard to argue with a computer“: Investigating Psychotherapists' Attitudes towards Automated Evaluation. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 2018, June, S. 559-571.
- Huang, Z.; Epps, J.; Joachim, D.; Sethu, V.: Natural Language Processing Methods for Acoustic and Landmark Event-Based Features in Speech-Based Depression Detection. *IEEE Journal of Selected Topics in Signal Processing*, 14(2). 2020, S. 435–448. Online: <https://doi.org/10.1109/JSTSP.2019.2949419>.
- Jin, K. W.; Li, Q.; Xie, Y.; Xiao, G.: Artificial intelligence in mental healthcare: An overview and future perspectives. *The British Journal of Radiology*, 96(1150), 20230213. 2023. Online: <https://doi.org/10.1259/bjr.20230213>.
- Kellogg, K. C.; Sadeh-Sharvit, S.: Pragmatic AI-augmentation in mental healthcare: Key technologies, potential benefits, and real-world challenges and solutions for frontline clinicians. *Frontiers in Psychiatry*, 13. 2022. Online: <https://www.frontiersin.org/articles/10.3389/fpsy.2022.990370>.
- Kleine, A.-K.; Kokje, E.; Lerner, E.; Gaube, S.: Attitudes Toward the Adoption of 2 Artificial Intelligence–Enabled Mental Health Tools Among Prospective Psychotherapists: Cross-sectional Study. *JMIR Human Factors*, 10, e46859. 2023. Online: <https://doi.org/10.2196/46859>.
- Kleine, A.-K.; Kokje, E.; Hummelsberger, P.; Lerner, E.; Gaube, S.: AI-Enabled Clinical Decision Support Tools for Mental Healthcare: A Product Review. Preprint. 2023. Online: <https://osf.io/ez43g>.
- Kleine, A.-K.; Lerner, E.; Cecil, J.; Heinrich, A.; Gaube, S.: Advancing mental health care with AI-enabled precision psychiatry tools: A patent review. *Computers in Human Behavior Reports*, 12, 100322. 2023. Online: <https://doi.org/10.1016/j.chbr.2023.100322>.
- Kwak, Y.; Ahn, J.-W.; Seo, Y. H.: Influence of AI ethics awareness, attitude, anxiety, and self-efficacy on nursing students' behavioral intentions. *BMC Nursing*, 21(1), 267. 2022. Online: <https://doi.org/10.1186/s12912-022-01048-0>.
- Rouidi, M.; Elouadi, A. E.; Hamdoune, A.; Choujtani, K.; Chati, A.: TAM-UTAUT and the acceptance of remote healthcare technologies by healthcare professionals: A systematic review. *Informatics in Medicine Unlocked*, 32, 101008. 2022. Online: <https://doi.org/10.1016/j.jimu.2022.101008>.
- Sadeh-Sharvit, S.; Camp, T. D.; Horton, S. E.; Hefner, J. D.; Berry, J. M.; Grossman, E.; Hollon, S. D.: Effects of an Artificial Intelligence Platform for Behavioral Interventions on Depression and Anxiety Symptoms: Randomized Clinical Trial. *Journal of Medical Internet Research*, 25(1), e46781. 2023. Online: <https://doi.org/10.2196/46781>.
- Sonde Health: Sonde Health homepage. 2023. Online: <https://www.sondehealth.com/>. [abgerufen am 04.12.2023].
- Tran, A. Q.; Nguyen, L. H.; Nguyen, H. S. A.; Nguyen, C. T.; Vu, L. G.; Zhang, M.; Vu, T. M. T.; Nguyen, S. H.; Tran, B. X.; Latkin, C. A.; Ho, R. C. M.; Ho, C. S. H.: Determinants of Intention to Use Artificial Intelligence-Based Diagnosis Support System Among Prospective Physicians. *Frontiers in Public Health*, 9. 2021. Online: <https://www.frontiersin.org/articles/10.3389/fpubh.2021.755644>.
- Venkatesh, V.; Morris, M. G.; Davis, G. B.; Davis, F. D.: User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425. 2003. Online: <https://doi.org/10.2307/30036540>.

Julia Cecil

LMU Center for Leadership and People Management

Projekt: Human-AI-Interaction in Healthcare

julia.cecil@psy.lmu.de

Lebenslauf in Kürze

Julia Cecil ist Doktorandin am Center for Leadership and People Management (CLPM) der LMU München in dem von der Volkswagen-Stiftung geförderten Projekt „Human-AI-Interaction in Healthcare: Identifying Factors Contributing to Clinical Utility“. Sie studierte Psychologie an der Universität Regensburg (B.Sc. und M.Sc.). Während ihres Studiums war Julia Cecil als wissenschaftliche Hilfskraft am Lehrstuhl für Sozial-, Arbeits-, Organisations- und Wirtschaftspsychologie an der Universität Regensburg und am CLPM tätig.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegt darin, die psychologischen Prozesse zu erforschen, die der Interaktion zwischen Mensch und Technologie (im Gesundheitswesen) zugrunde liegen, um ein tieferes Verständnis über Entscheidungsfindungen mit KI-Systemen zu erlangen.

Christiane Saure

Generative KI im Krankenhauswesen: Eine wertschöpfungskettenorientierte Betrachtung

Künstliche Intelligenz („KI“) hat seit dem Beginn ihrer Entwicklung einen tiefgreifenden Wandel durchlaufen. Während in den Anfängen fortgeschrittene Analytik auf der Basis von Algorithmen im Vordergrund stand, lag der Fokus im Anschluss daran auf maschinellem Lernen und schließlich auf Deep Learning. Generative KI ist das neueste und bisher interessanteste Tool in der Entwicklung von KI und erweitert die bisherigen Nutzungspotenziale von KI im Sinne eines Paradigmenwechsels. Generative KI ist ein Sammelbegriff für diejenigen KI-basierten Systeme, welche menschliche Fertigkeiten simulieren und damit die Ergebnisse menschlicher Intelligenz mindestens erreichen, wenn nicht sogar übertreffen. Dabei handelt es sich um die Fähigkeit, textliche, bildliche, visuelle oder sprachliche Inhalte selbständig zu generieren. Die Basis Generativer KI-Systeme bilden sog. Foundation Models, welche erhebliche Datenmengen unter Nutzung von Tiefen Neuronalen Netzen verarbeiten. Als konkrete Anwendungsbeispiele Generativer KI können derzeit beispielsweise ChatGPT, Med-PaLM 2, BERT, LaMDA, DALL-E genannt werden.

Herkömmliche KI-Lösungen kommen bereits seit einiger Zeit im Krankenhauswesen zum Einsatz. Generative KI-Systeme im Gesundheitswesen sind jedoch vergleichsweise neu und befinden sich noch ganz überwiegend im Anfangsstadium der Anwendung bzw. sogar in der Entwicklung. Branchenübergreifende Studien weisen darauf hin, dass durch die Anwendung Generativer KI erhebliche Automatisierungspotenziale erschlossen werden können (Chui, et al., S. 8 ff.; Daugherty et al., S. 11). Vor diesem Hintergrund besteht das Ziel des vorliegenden Beitrags darin, potenzielle Auswirkungen Generativer KI auf ausgewählte medizinische und nicht-medizinische Prozesse im Krankenhaus darzustellen.

Als Systematisierungs- und Strukturierungsansatz dient dabei das von Porter entwickelte Konzept der Wertschöpfungskette (Porter, S. 65 ff.). Danach beinhaltet die Wertschöpfungskette des Krankenhauses eine Abfolge von Versorgungsstufen, die unter Berücksichtigung der krankheitsindividuellen, persönlichen sowie sozialen Situation des Patienten zur fallabschließenden

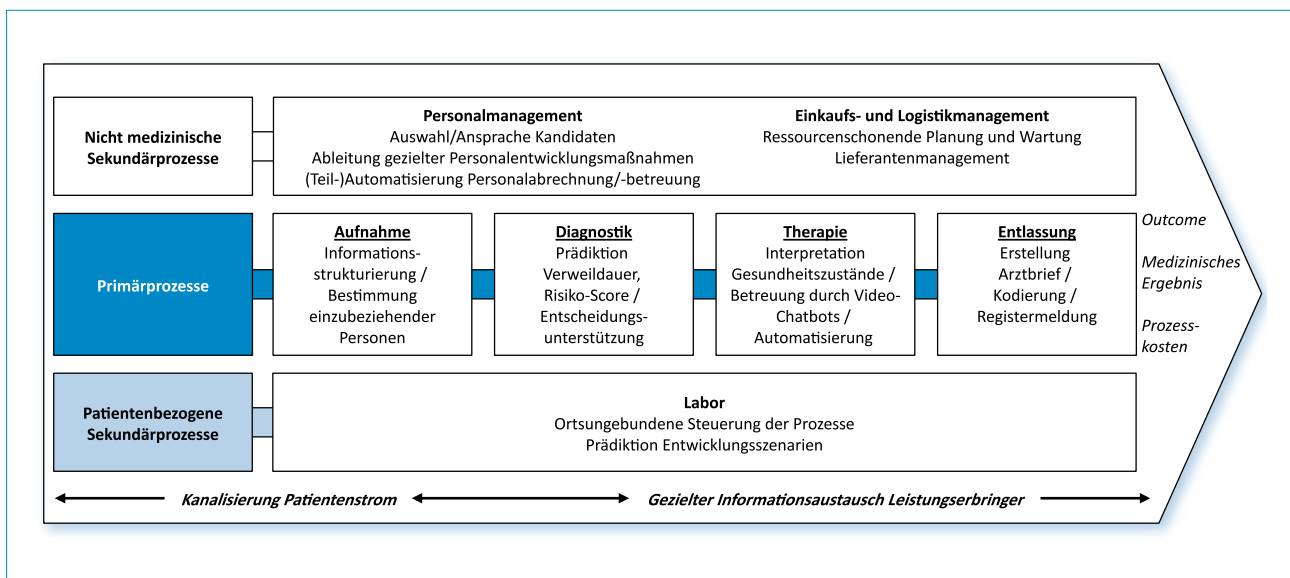


Abbildung 1: Einsatzmöglichkeiten Generativer KI entlang der Krankenhaus-Wertschöpfungskette

Behandlung durchlaufen werden müssen. Die Darstellung visualisiert diese Wertschöpfungskette sowie die im Folgenden beschriebenen Einsatzmöglichkeiten Generativer KI im Krankenhaus.

Primärprozesse

Der Prozess der Aufnahme eines Patienten in das Krankenhaus wird durch den Einsatz Generativer KI maßgeblich weiterentwickelt werden können. Wesentliches Kennzeichen dieser Weiterentwicklung bildet die Zurverfügungstellung und Sammlung sowie die Systematisierung relevanter patientenbezogener Informationen aus unterschiedlichen Datenquellen. Neben Informationen aus der zukünftig für freigeschaltete Leistungserbringer zugänglichen digitalen Patientenakte werden Eintragungen in Gesundheits-Apps sowie Aufzeichnungen von Wearables auf relevante Symptome und Beschwerden hin analysiert. Der prädiktive Charakter der KI-Systeme macht es in der Folge möglich, diejenigen ärztlichen sowie pflegerischen Fachdisziplinen im Krankenhaus frühzeitig und zielsicher zu bestimmen und automatisiert systemseitig zu kontaktieren, welche im Rahmen der Diagnostik aus fachlichen Erwägungen einzubeziehen sind.

Generative KI-Modelle ermöglichen eine spezifischere, zeitnähere und damit qualitativ hochwertigere Diagnostik. Durch die Zusammenführung von Ergebnissen

verschiedener Informationsquellen, wie z.B. bisheriger Befunde, vor Ort im Krankenhaus erstellter Aufnahmen und Ergebnisse wissenschaftlicher Studien generieren KI-Systeme Aussagen zu Behandlungsszenarien und den damit verbundenen Folgen für die Verweildauer, die Mortalität sowie patientenspezifische Risiko-Scores (Gutierrez, S. 6). Somit nimmt die KI zukünftig eine große Rolle im Rahmen der Prädiktion zukünftiger Ereignisse und Ergebnisse ein, welche mit bestimmten Behandlungsmustern verbunden sind. Ärzte und Gesundheitsdienstleister werden insoweit in die Lage versetzt, zielgenaue Diagnosen zu stellen und im Sinne einer bestmöglichen Patientenversorgung Therapieentscheidungen zu treffen. (Antweiler et al., S. 4).

Im Rahmen der Therapie werden punktuell heute schon Medizin- wie auch Pflegeroboter eingesetzt. Diese unterstützen Ärzte bei chirurgischen Eingriffen und Pflegekräfte im Zusammenhang mit der Verrichtung körperlich schwerer Tätigkeiten. Generative KI wird zukünftig die Virtualisierung der therapeutisch-medizinischen wie auch der pflegerischen Prozesse vorantreiben. Denkbar ist zum einen, dass Gesundheitszustände von Patienten während der Therapiephase durch Generative KI interpretiert werden. Die KI könnte sodann unterstützend und richtungweisend bei der Entscheidung für weitere therapeutische und pflegerische Maßnahmen wirken.

Use Cases: Diagnostik

Aktuell befinden sich verschiedene Anwendungen Generativer KI zur Diagnostik und Befundung in der Entwicklung. Dabei ist festzustellen, dass die meisten Produkteigenschaften konzeptionell und technisch getestet werden, allerdings noch nicht kommerziell verfügbar sind.

Siemens Healthineers entwickelt derzeit in Zusammenarbeit mit dem Universitätsklinikum Essen einen Prototyp zur Evaluation und Weiterentwicklung eines Softwareassistenten für die radiologische Befundung. Ziel ist es dabei, ein Chatsystem bereit zu stellen, welches Antworten auf medizinische Fragestellungen, Reports und Bilder miteinander verknüpft. So wird perspektivisch der passende Bereich innerhalb eines Reports durch Klicks auf ein medizinisches Bild hervorgehoben werden können. Darüber hinaus sollen Reports bei der Durchsicht diagnostischer Bilder erstellt und in Echtzeit priorisiert werden (Ehealthcom, 2023).

Ein weiteres Anwendungsbeispiel sind die von Paige.AI entwickelten Produkte zur Entdeckung und Diagnostik von Prostata- sowie Brustkrebs. Paige.AI ist das erste Unternehmen, welches eine Zulassung der U.S Food and Drug Administration (FDA) für den Einsatz von Generativer KI in der Pathologie erhalten hat. Aktuell arbeitet Paige.AI mit Microsoft zusammen, um ein bildbasiertes KI-Modell zur Erkennung von bösartigen Neubildungen zu entwickeln (Paige.AI, 2023).

Zum anderen ist zu erwarten, dass Generative KI zukünftig auch emotionale Betreuungsleistungen übernimmt. Diese Leistungen wurden in den bisherigen wissenschaftlichen Diskussionen zumeist als nicht substituierbar qualifiziert. Video-Chatbots oder auch humanoide virtuelle Assistenten in 3D, welche direkt auf die Sprache von Patienten reagieren, geben medizinisch fundierte Antworten und bieten darüber hinaus emotionale Unterstützung für Patienten. Auf diese Weise erfolgt eine Automatisierung ausgewählter Arbeitsschritte und das medizinische wie auch das pflegerische Personal wird entlastet.

Diese Entwicklungen ermöglichen es, dass der potenziell weiter anwachsende Fach- und Pflegekräftemangel reduziert wird. So gehen Daugherty et al. davon aus, dass Generative KI 37% der Tätigkeiten von Ärzten und 35% der Tätigkeiten des unterstützenden Personals maßgeblich beeinflussen wird und diese Arbeitsschritte potenziell automatisierbar sind (Daugherty et al., S. 15). Wesentliches Kennzeichen hierbei ist, dass im Gegensatz zu bisherigen KI-Modellen

und KI-Systemen auf der anderen Seite führt potenziell zu einer Verbesserung der Prozess- wie auch der Ergebnisqualität.

Die Erstellung von Artribriefen im Rahmen der Entlassung des Patienten aus dem Krankenhaus bindet Ressourcen und gehört zu den weniger geschätzten administrativen Tätigkeiten des ärztlichen Personals. Mithilfe Generativer KI werden strukturierte wie auch unstrukturierte Informationen aus verschiedenen Datenquellen, wie der elektronischen Patientenakte, Sprachaufnahmen o. ä., genutzt und zu einem Textvorschlag für den Arztbrief zusammengeführt. Lediglich die finale Bearbeitung des Textes erfolgt in diesem Kontext durch das ärztliche Personal (Antweilder et al., S. 9).

Automatisierbar ist darüber hinaus auch die Meldung von Erkrankungen an medizinische Register. Die Kodierung der erbrachten medizinischen Leistungen erfolgt aktuell durch speziell geschulte Medizincontroller, könnte jedoch aus in Zukunft von Generative KI übernommen werden (Antweilder et al., S. 9).

Use Case Medizinische Dokumentation

HCA Healthcare, Nashville, betreibt 180 Krankenhäuser, ca. 2.300 ambulante Pflegedienste sowie weitere Einrichtungen in 20 Bundesstaaten der USA sowie in Großbritannien. Im Rahmen einer strategischen Partnerschaft mit Augmedix, einem Anbieter von Apps für medizinisches Personal und Google Cloud testet HCA Healthcare seit April 2023 eine Anwendung zur automatischen Sprache-zu-Text-Verarbeitung. Auf der Basis des Generativen und auf den Gesundheitssektor zugeschnittene KI-Systeme Med-PaLM 2 werden Informationen aus Gesprächen zwischen Arzt und Patient aufgezeichnet, extrahiert und in medizinische Notizen umgewandelt. Diese Notizen müssen in einem nächsten Schritt durch die Ärzte kontrolliert und ggf. ergänzt werden, bevor sie direkt in die elektronische Patientenakte (EPA) des Krankenhauses übertragen werden. Auf diese Weise erfolgt eine Teilautomatisierung des Dokumentationsprozesses und zeitliche wie auch personelle Ressourcen werden geschont (vgl. Google Cloud).

nicht lediglich einzelne ausgewählte Tätigkeiten (z.B. Umbetten und Lagern des Patienten, Anreichen von OP-Besteck, Setzen einer Naht, Intra- und Postoperative Überwachung) substituiert werden können. Die Besonderheit besteht vielmehr darin, die gesamten Behandlungs- und Pflegeprozesse durch den Einsatz von KI so zu gestalten, dass Arbeitsschritte verlagert, beschleunigt, zusammengefasst und parallelisiert werden. Die ständige Interaktion zwischen medizinischem und pflegerischem Personal auf der einen Seite

Patientenbezogene Sekundärprozesse

In Laboren wird täglich eine große Menge an Daten erzeugt. Schon heute sind digitale Werkzeuge unabdingbar für die Erfassung, Auswertung und Speicherung von Messresultaten. Mithilfe von Generativer KI-Systeme wird der Automatisierungsgrad in Laboren zukünftig noch weiter zunehmen. Dabei werden Prozesse über das Smartphone ortsungebunden gesteuert. Während derzeit die Analyse mikroskopischer Aufnahmen am Bildschirm erfolgt, wird Generative KI dazu beitragen,

dass Labormediziner virtuell durch das Innere der Untersuchungsobjekte wandern, um Unregelmäßigkeiten festzustellen. Dies führt dazu, dass Labormedizinern sehr viel stärker als bisher eine prädiktive Rolle bei der Beschreibung von Szenarien der Weiterentwicklung von Zellen (z.B. Entartungsrisiken) zukommt.

Nicht-Medizinische Sekundärprozesse

Zu den nicht-medizinischen Sekundärprozessen zählen diejenigen Prozesse, die die Verwaltung des Krankenhauses, das Controlling, das Personalmanagement oder die Steuerung der logistischen Ressourcen betreffen. Die Anwendung von KI-Systemen erstreckt sich hier auf die Berechnung optimaler Allokationen, die Verbesserung von Abrechnungen oder die Steuerung und Beschleunigung von Lieferketten.

In Zeiten des Fachkräftemangels ist das Personalmanagement eine wichtige Kernfunktion innerhalb der Verwaltungsprozesse. Insbesondere für das Recruiting aber auch die Personalentwicklung kann Generative KI dazu beitragen, Mitarbeiter zu gewinnen und langfristig an die Organisation zu binden. Die Erstellung von Wunschprofilen für Mitarbeiter durch das Krankenhaus dient als Grundlage für die internetbasierte Suche und zielgerichtete Ansprache von Wunschkandidaten und die Auswertung von Bewerbungen. Ein weiterer Anwendungsbereich erstreckt sich auf die Analyse von Daten der Mitarbeiter, wie z.B. vorzuziehenswerte Einsatzbereiche sowie Arbeitszeiten und die anschließende Entwicklung von personalisierten Schulungs- und Entwicklungsplänen. Auch in diesem Bereich erschließt sich erhebliches Automatisierungspotenzial. Hierbei handelt es sich um die Erstellung und Anpassung von Gehaltsabrechnungen, die Planung von Leistungsbeurteilungen oder die Beantwortung von individuellen Fragen der Mitarbeiter. In der Gesamtschau kann Generative KI somit als ein Instrument zur Herstellung eines höheren Mitarbeiterbindung angesehen werden.

Im Einkaufs- und Logistikmanagement von Krankenhäusern trägt Generative KI zukünftig dazu bei, den Ausfall von medizinischen Geräten im Voraus zu bestimmen. In der Folge können die Wartung und Reparatur dieser Geräte ressourcenschonender geplant und Ausfallzeiten minimiert werden. Ein großer Teil der deutschen Krankenhäuser ist in Einkaufsgemeinschaften organisiert.

Generative KI hat dabei das Potenzial, durch die Analyse großer Datenmengen spezifische und passgenaue Lieferanten zu identifizieren, so dass die Verfügbarkeit von Produkten, die Qualität, der Preis und auch der Servicelevel verbessert werden (Chui, et al., S. 8 ff.).

Schnittstellenmanagement zwischen Krankenhäusern sowie vor- und nachgelagerten Versorgungsstufen Wesentliches Kennzeichnen qualitativ hochwertiger Patientenversorgung ist die schnittstellenübergreifende und zielorientierte Zusammenarbeit des Krankenhauses mit allen dem Krankenhaus vor- und nachgelagerten Versorgungsstufen. Generative KI wird zukünftig dazu beitragen, den Patientenstrom im Gesundheitswesen zu kanalisieren und patientenspezifische Informationen zu dem Gesundheitszustand und potenziellen Therapieansätzen zwischen relevanten Leistungserbringern entlang der Versorgungskette zu kommunizieren. Auf diese Weise können Verweildauern gesenkt, das medizinische und pflegerische Personal entlastet und die Zusammenarbeit der an der Versorgung beteiligten Einrichtungen wie niedergelassenen Ärzte, Pflegeeinrichtungen, Krankenhäusern oder Rehabilitationseinrichtungen verbessert werden.

Anforderungen an die Umsetzung Generativer KI im Krankenhauswesen

An die Nutzung und den Einsatz von Generativer KI sind sehr hohe Anforderungen zu stellen, die sich aus den nachfolgend dargestellten Aspekten ergeben:

- | Inhalte, welche durch Generative KI erzeugt wurden, sind schwer von echten Inhalten zu unterscheiden.
- | Die durch Generative KI generierten Informationen zu Diagnose und Therapie sind unter Umständen inkorrekt, ohne dass dies für die Anwender erkennbar wird.
- | Die Verwendung medizinischer Patientendaten durch Generative KI entspricht möglicherweise nicht den datenschutzrechtlichen Rahmenbedingungen.
- | Generative KI erarbeitet unter Umständen wenig spezifische oder auch diskriminierende Inhalte.

Wichtige Kriterien zur Evaluation von KI-Systemen sind somit Datenschutz, Sicherheit, Vertrauenswürdigkeit und Transparenz. Unter Berücksichtigung dieser Anforderungen hat die Europäische Kommission im Dezember 2023 den EU AI Act verabschiedet. Dieser ist als

eines der ersten umfassenden KI-Gesetze weltweit zu qualifizieren. Die darin enthaltenen Vorschriften legen fest, welche Pflichten Anbieter und Nutzer unter Berücksichtigung unterschiedlicher Risikograde Generativer KI erfüllen müssen.

Empfehlungen für das Krankenhausmanagement
Vor dem Hintergrund der dargelegten Chancen und Verbesserungspotenziale, die durch den Einsatz Generativer KI erzielbar erscheinen, sind folgende Empfehlungen für das Krankenhausmanagement abzuleiten. Krankenhäuser sollten zunächst eine spezifische Strategie für den Einsatz Generativer KI implemen-

tieren. In der Folge wird es notwendig sein, Investitionen in IT-Infrastruktur, in Datenmanagement- wie auch Analysetools zu tätigen und die Intraoperabilität von Daten und Systemen zu forcieren. Eine große Herausforderung für das Krankenhausmanagement wird darin bestehen, Generative KI-Modelle mit einer großen Menge von krankenhausspezifischen Daten und Informationen zu trainieren. In der Gesamtschau werden alle Beteiligten gefordert sein, nicht nur die technischen Voraussetzungen für Generative KI im Krankenhauswesen zu schaffen, sondern auch konkrete Use Cases durch anwendungsorientierte Forschung zu erschließen.

Literaturangaben:

- Antweiler, D.; Beckh, K.; Chakraborty, N.; Giesselbach, S.; Klug, K.; Rüping, S.: Language Processing in der Medizin. Von der automatisierten Befundanalyse bis zum Arztbrief-generator: Künstliche Intelligenz für dokumentenbasierte Prozesse im Krankenhaus, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS (Hrsg.). 2023. Online: <https://publica-rest.fraunhofer.de/server/api/core/bitstreams/2c49467b-4350-4099-9040-f539cb4302f9/content> [abgerufen am 11.12.2023].
- Chui, M.; Hazan, E.; Roberts, R.; Singla, A.; Smaje, K.; Sukharevsky, A.; Yee, L.; Zimmel, R.: The economic potential of generative AI The next productivity, McKinsey & Company (Hrsg.). 2023. Online: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier/> [abgerufen am 11.12.2023].
- Daugherty, P.; Ghosh, B.; Narain, K.; Guan, L.; Wilson, J.: A new era of generative AI for everyone. The technology underpinning ChatGPT will transform work and reinvent business, Accenture (Hrsg.). 2023. Online: <https://www.accenture.com/content/dam/accenture/final/accenture-com/document/Accenture-A-New-Era-of-Generative-AI-for-Everyone.pdf> [abgerufen am 11.12.2023].
- Ehealthcom, Siemens Healthineers zeigt die Möglichkeiten von generativer KI in der Medizintechnik. 2023. Online: <https://e-health-com.de/details-unternehmensnews/siemens-healthineers-zeigt-die-moeglichkeiten-von-generativer-ki-in-der-medizintechnik/> [abgerufen am 11.12.2023].
- Google Cloud, HCA Healthcare Collaborates with Google Cloud to Bring Generative AI to Hospitals. 2023. Online: <https://www.googlecloudpresscorner.com/2023-08-29-HCA-Healthcare-Collaborates-with-Google-Cloud-to-Bring-Generative-AI-to-Hospitals> [abgerufen am 11.12.2023].
- Gutierrez, G.: Artificial Intelligence in the Intensive Care Unit, in: Critical care. 2020, Online: <https://doi.org/10.1186/s13054-020-2785-y> [abgerufen am 11.12.2023].
- Paige.AI, U.S. FDA Grants Paige Breakthrough Device Designation for Cancer Detection in Breast Lymph Nodes. 2023. Online: <https://paige.ai/u-s-fda-grants-paige-breakthrough-device-designation-for-cancer-detection-in-breast-lymph-nodes> [abgerufen am 11.12.2023].
- Porter, M.E.: Wettbewerbsvorteile: Spitzenleistungen erreichen und behaupten, 8. Aufl.. Frankfurt/New York 2014.

Prof. Dr. Christiane Saure

Frankfurt University of Applied Sciences, Frankfurt am Main
Fachbereich 4: Soziale Arbeit und Gesundheit
saure.christiane@fb4.fra-uas.de

Lebenslauf in Kürze

Frau Saure wurde im Jahr 2016 zur Professorin für Change Management und Organisationsentwicklung in Gesundheitseinrichtungen an die Frankfurt University of Applied Sciences berufen. Seitdem hat sie sich in verschiedenen Funktionen, u.a. als geschäftsführende Kanzlerin der Hochschule sowie als Prodekanin des Fachbereichs Soziale Arbeit und Gesundheit, engagiert. Frau Saure blickt auf eine langjährige Tätigkeit als Unternehmensberaterin zurück. So hat sie von 1998 bis 2016 zahlreiche Krankenhäuser, Uniklinika, Rehabilitationskliniken und Krankenkassen zu finanzwirtschaftlichen wie auch organisatorischen Fragestellungen u.a. bei PricewaterhouseCoopers (PwC) beraten. Frau Saure ist ausgebildete Bankkauffrau und studierte Betriebswirtschaftslehre sowie Anglistik an der Westfälischen Wilhelms-Universität Münster.



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Meine Leidenschaft in diesem Bereich ist die Beantwortung der Frage, auf welche Weise KI den Pflegekräftemangel zukünftig beseitigen kann.

Sibel Altin

Relevanz von KI-Anwendungen aus der Perspektive von Krankenversicherungen

Einführung und Relevanz

Die steigenden Anforderungen an das Gesundheitswesen erfordern innovative Lösungen, um die Effizienz zu steigern und gleichzeitig eine qualitativ hochwertige Versorgung zu gewährleisten. Künstliche Intelligenz (KI) bietet in diesem Kontext vielfältige Möglichkeiten, insbesondere für Krankenversicherungen, um die Herausforderungen der modernen Gesundheitslandschaft anzugehen. Dieser Beitrag untersucht die Relevanz von KI-Anwendungen in der Krankenversicherungsbranche und stellt Einsatzbereiche sowie interessante Fallbeispiele aus dem deutschen Gesundheitswesen dar, um die Potenziale von KI-Anwendungen für die Steigerung der Innovationsfähigkeit und Effizienz im Krankenversicherungswesen aufzuzeigen.

KI-Anwendungen können in vielfältigen Aufgabenbereichen des Krankenversicherungswesens ihre positiven Potenziale entfalten. Die Einsatzbereiche lassen sich dabei in drei maßgebliche Bereiche einteilen. So können KI-Anwendungen in der Verwaltung, in der Leistungssteuerung sowie in der Kundenkommunikation von Krankenversicherungen eingesetzt werden.

In der Verwaltung können durch KI-basierte Anwendungen, einfache bis komplexere Routineaufgaben in der Rechnungsprüfung sowie die Bearbeitung von Mitgliedsanfragen vereinfacht und automatisiert werden. In beiden Einsatzbereichen können maßgebliche Effizienzreserven gehoben werden, da Rechnungsprüfungsaufgaben sowie die Mitgliedsbetreuung im Hinblick auf die Prüfung von Anspruchsvoraussetzungen für Leistungsansprüche in einem großen Umfang personelle Kapazitäten binden. Beide Bereiche eignen sich zudem auch deswegen für KI-basierte Lösungen, weil eine Standardisierung der Fallarten möglich ist und große Testdatensätze generierbar sind, um jeweilige KI-Modelle zu trainieren. Ergänzend lassen sich KI-basierte Ansätze ebenfalls in der rechnungsprüfungsbasierten

Betrugserkennung mittels komplexer Mustererkennung einsetzen, wobei auch hier maßgebliche Einsparungen generiert werden können. Zuletzt sind KI-basierte Ansätze ebenfalls in der Prognose des Beitragsaufkommens sowie der Leistungsanspruchnahme einsetzbar, um den Haushaltsplanungsprozess von Krankenversicherungen zu unterstützen. Die benannten Einsatzmöglichkeiten können Krankenversicherungen maßgebliche Effizienzvorteile und hierdurch eine Reduktion der Verwaltungsausgaben ermöglichen. Im Hinblick auf die aufwandsarme Integration von KI-Anwendungen in die Rechen- und Verarbeitungssysteme kommt den Krankenversicherungen der heute bereits hohe Digitalisierungsgrad im Versicherungswesen entgegen, so dass die Grundlage zur digitalen Verarbeitung großer Datenmengen grundsätzlich gegeben ist.

In der Gestaltung und Steuerung der Gesundheitsversorgung der Versicherten stecken weitere maßgebliche Potenziale, die durch KI-Anwendungen genutzt werden können. So können mittels KI-basierter Tools die Analysemöglichkeiten der Abrechnungsdaten erweitert und auf diese Weise die Versorgungsplanung für den individuellen Versicherten sowie für gesamte Versorgungsregionen verbessert werden. In solche Modelle können individualisierte Diagnose-, Behandlungs- und Arzneimitteldaten als auch Daten zum Inanspruchnahmeverhalten von Versicherten im Hinblick auf Leistungserbringer (Versorgungspfade) einfließen. Ergänzend können die verbesserten Datenanalysemöglichkeiten und die zusätzliche Einbeziehung von Patientendaten (selbst erhobene Vital- und Fitnessdaten des Patienten) dazu genutzt werden, die individuelle Gesundheitsversorgung in der Form zu verbessern, als dass mittels Prädiktionsanalysen personalisierte Präventions- und Vorsorgeempfehlungen entwickelt und den Versicherten offeriert werden könnten. So würden beispielweise Patienten mit Risikofaktoren (z.B. Bluthochdruck, Adipositas) gezielt Präventionsangebote unterbreitet. Eine

solche Umgestaltung hin zu einer datengetriebenen Vorsorge- und Versorgungsplanung für Versicherte würde zum einen die Versorgung des Versicherten verbessern und zum anderen durch den verbesserten Service die Versichertenbindung erhöhen. Hierdurch würde ein Qualitätswettbewerb zwischen Krankenversicherungen eingeleitet und der derzeit bestehende alleinige Preiswettbewerb sinnvoll ergänzt. Zudem könnten durch eine stärkere Patientensteuerung und eine intensivierete, präventive Versorgung langfristige Kosteneinsparungen generiert werden.

Im Hinblick auf die Kundenkommunikation könnten KI-basierte Spracherkennungssysteme dazu genutzt werden, die Service- und Telefonie Einheiten von Krankenversicherungen zu unterstützen oder Mitgliedsanfragen standardisiert mittels Chatbots zu verarbeiten. Im Hinblick auf die historisierte Auswertung von Mitgliedsanfragen könnte auch ein Kundentargeting durchgeführt und gezielt ergänzende Versicherungs- oder Leistungsangebote offeriert werden. Diese Maßnahmen würden für Krankenversicherungen kostensenkende Effekte bei gleichzeitiger Serviceverbesserung ermöglichen.

Interessante Fallbeispiele

1) Versorgungsverbesserung durch Gesundheitsplattformen für eine individualisierte Gesundheitsversorgung

Zur Umsetzung der KI-basierten individualisierten Vorsorge- und Gesundheitsplanung für Versicherte können eigens von Krankenversicherungen angebotene appbasierte Gesundheitsplattformen genutzt werden. Auf diesen können den Versicherten sodann Analyseergebnisse zu ihrem jeweiligen Gesundheitsprofil dargestellt und darauf aufbauende Leistungsempfehlungen unterbreitet werden. Ergänzend können auch Termine mit Leistungserbringern (Terminservice) vereinbart werden, um die Wahrscheinlichkeit der Inanspruchnahme der empfohlenen Vorsorge- und Gesundheitsleistungen zu erhöhen. Erste größere Gesetzliche Krankenversicherungen bieten bereits Apps mit ersten KI-basierten Anwendungen an. So bietet die AOK Plus in Sachsen mit der NAVIDA App ihren Versicherten eine Gesundheitsplattform an, die als persönlicher Gesundheitsassistent dient [AOK PLUS]. Die Versicherten können die App zum einen nutzen, um mit ihrer Versicherung zu kommunizieren (z.B. digitalisierte Anfrage- und Antragsmöglichkeiten).

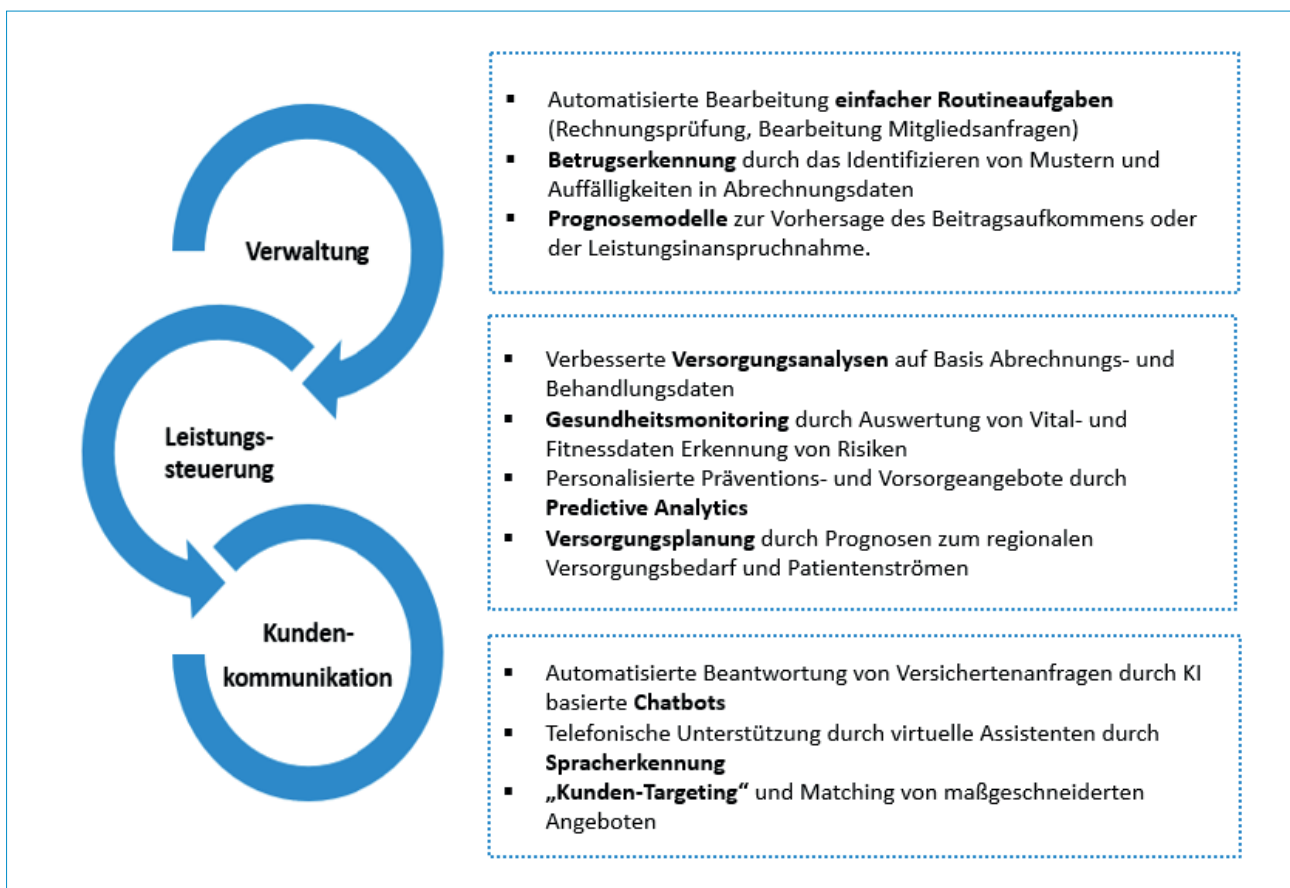


Abbildung 1: Einsatzbereiche von KI-Anwendungen in Krankenversicherungen (Eigene Darstellung)

Daneben unterbreitet die App individualisierte Vorsorgeempfehlungen und stellt Kursangebote zur Verfügung, so dass eine Inanspruchnahme empfohlener Leistungen durch die Niedrigschwelligkeit vereinfacht wird. Ergänzend ist in die App ein KI-basierter Symptomchecker integriert, der den Versicherten hilft, die Bedeutung akuter Gesundheitssymptome richtig einzuordnen und gleichzeitig Empfehlungen darüber gibt, welche Handlungsoptionen (z.B. wait and see, Inanspruchnahme Hausarzt oder Facharztbesuch, Bedarf Notfallversorgung) ihnen offenstehen. Gesundheitsplattformen dieser Art sind derzeit auf dem Vormarsch und könnten die Vorteile KI-basierter Analysemöglichkeiten für die Steuerung der Versorgung der Versicherten zeitnah in die Versorgung tragen, so dass auf der einen Seite Versicherte von einer verbesserten Versorgungssteuerung profitieren und auf der anderen Seite Krankenversicherungen eine bedarfs- und zielgerichtete Leistungsanspruchnahme erreichen können.

2) Versorgungsverbesserung durch KI-bezogene Forschungsförderung

Mit dem Ziel einer Beschleunigung der Forschung an KI-basierten Analysetools mit medizinischen Versorgungsdaten aus Deutschland und der damit verbundenen Hoffnung auf sich hieraus ergebenden Erkenntnissen und daraus ableitbaren Versorgungsverbesserungen arbeitet das Bundesministerium für Gesundheit (BMG) derzeit an dem Gesundheitsdatennutzungsgesetz (GDNG) [BMG]. Dieses sieht den Aufbau eines Forschungsdatenzentrums (FDZ) vor, welches beim Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) angesiedelt werden soll. In dem Forschungsdatenzentrum sollen die Abrechnungsdaten der Kranken- und Pflegekassen sowie die Daten aus der elektronischen Patientenakte (ePA) und klinische Registerdaten (z.B. Krebsregister) pseudonymisiert zusammengefasst und verknüpft werden. Auf diese Weise würde der weltweit erste umfassende Gesundheitsdatensatz dieser Art für eine gesamte Bevölkerung entstehen. Durch die intelligente Zusammenführung und Nutzung dieser Daten zum Training von KI-Anwendungen könnten enorme Entwicklungsfortschritte bei KI-Modellen und darauf aufbauend auch für die medizinische Forschung entstehen und Deutschland könnte zum Zentrum der KI-

basierten Medizinforschung werden. Der Zugang zu den Daten wird dabei explizit für Wissenschaftler ermöglicht werden, die vor dem Zugang ein Prüfungsverfahren für ihr Forschungsvorhaben mitsamt eines Antrags beim BfArM durchlaufen müssen, um unter anderem zu gewährleisten, dass die Forschung an den Daten im Sinne des Gemeinwohls und nicht mit einem Gewinninteresse erfolgt.

Ergänzend zur Etablierung des Forschungsdatenzentrums sieht das GDNG vor, die Datennutzungsrechte und Informationspflichten der Gesetzlichen Krankenkassen zu erweitern. Die gesetzlichen Krankenkassen haben ihren Versicherten grundsätzlich durch Aufklärung, Beratung und Leistungen zu helfen (§1 SGB V). Diese Pflicht soll nun durch das GDNG erweitert werden, um die Informierung der Versicherten über ihr individuelles Risikoprofil und die Empfehlung möglicher Gesundheitsmaßnahmen (§25b SGB V). Zur Erkennung des individuellen Risikoprofils dürfen Krankenkassen nun Abrechnungsdaten auf Individualebene verarbeiten und bei festgestellten Gesundheitsrisiken Versicherte informieren. Diese Erweiterung der Datenanalysemöglichkeiten wird es Krankenversicherungen zukünftig ermöglichen, Risikodaten der Versicherten auf Individualebene auszuwerten. Aus diesen Auswertungen könnte beispielsweise eine Empfehlung für eine risikoadaptierte Früherkennung bei erhöhtem Krebsrisiko abgeleitet werden. Zudem könnten nicht-leitlinienkonforme Behandlungsregime oder Wechselwirkungen bei Medikamenten leichter identifiziert und den Versicherten als Risiko angezeigt werden bei gleichzeitiger Darstellung von Empfehlungen zu Handlungsoptionen.

Insbesondere die Erweiterung des Analyse- und Informationsrechts der Krankenversicherungen birgt ein enormes Potential für eine zukünftig individualisierte Versorgungssteuerung. Hierbei könnten in Verbindung mit krankenkasseneigenen Gesundheitsplattformen direkte Steuerungsinstrumente eingesetzt werden, um Versicherte zu informieren und deren Versorgung effektiv zu steuern.

Fazit

Die Integration von Künstlicher Intelligenz in die Krankenversicherungsbranche eröffnet eine Vielzahl

von Chancen zur Steigerung der Innovationsfähigkeit und Wirtschaftlichkeit. Der verantwortungsbewusste Einsatz von KI wird nicht nur die betriebliche Effizienz verbessern, sondern auch die Qualität der Versorgung für Versicherte erhöhen. Die kontinuierliche

Zusammenarbeit zwischen Krankenversicherungen, Gesundheits-einrichtungen und Regulierungsbehörden ist entscheidend, um die positive Entwicklung dieser Technologien sicherzustellen und gleichzeitig ethische Standards zu wahren.

Literaturangaben:

AOK PLUS. AOK NAVIDA-Ihre persönliche Gesundheitsassistentin. Online: <https://www.aok.de/pk/plus/app-navida-gesundheitsassistentin/> [abgerufen am 08.12.2023].

Bundesministerium für Gesundheit (BMG). Kabinettsentwurf eines Gesetzes zur verbesserten Nutzung von Gesundheitsdaten (Gesundheitsdatennutzungsgesetz – GDNG) vom 30.08.2023. Online: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/G/GDNG_Kabinettd.pdf [abgerufen am 08.12.2023]

Prof. Dr. Sibel Altin

Frankfurt University of Applied Sciences
Fachbereich 4: Soziale Arbeit und Gesundheit
sibel.altin@fb4.fra-uas.de

Lebenslauf in Kürze

Frau Prof. Dr. Altin studierte Gesundheitsökonomie mit dem Schwerpunkt Management im Gesundheitswesen an der Universität zu Köln. Anschließend war sie zwischen 2013 und 2016 als wissenschaftliche Mitarbeiterin am Institut für Gesundheitsökonomie und klinische Epidemiologie (IGKE) an der Universitätsklinik Köln beschäftigt und promovierte 2016 zum Thema „Patientenorientierung in Gesundheitssystemen“. Zwischen 2016 und 2018 arbeitete sie als Post-Doc am Lehrstuhl für Medizinmanagement an der Universität Duisburg-Essen und gehörte dem Arbeitsbereich „Gesundheitssystem, Gesundheitspolitik und Arzneimittelsteuerung“ an. Anschließend war sie zwischen 2018 und 2023 in verschiedenen Positionen bei der AOK Rheinland/Hamburg tätig. Im Rahmen ihrer Tätigkeit leitete sie die Abteilung Versorgungsinnovationen und verantwortete die Entwicklung und Umsetzung innovativer, digitaler Versorgungsangebote für die Versicherten der AOK Rheinland/Hamburg.

Ihre Forschungsschwerpunkte liegen in den Bereichen Versorgungsforschung sowie Gesundheitssystemforschung.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Implementierung und Evaluation KI-basierter Digital Health Ansätze für die Patientenversorgung der Zukunft.



4. Die juristische und ethische Perspektive

Olivia Alig

KI im Konfliktmanagement – rechtliche & ethische Aspekte

Prolog

„In der Konfliktlösung bleibt trotz der unterstützenden Rolle von KI die menschliche Empathie, Intuition und Urteilsbildung unverzichtbar. Künstliche Intelligenz sollte die Expertise beim Konfliktmanagement erweitern, nicht jedoch ersetzen. Es ist die menschliche Einsicht, die den entscheidenden Unterschied macht und durch KI-Tools lediglich bereichert wird.“

<https://chat.openai.com>, Prompt: KI & Konfliktmanagement, ChatGPT 4, 12.11.23

1. Einleitung

Die Themen Künstliche Intelligenz und Konfliktmanagement erscheinen auf den ersten Blick nicht zusammen zu passen. Dies ergibt sich zum einen daraus, dass die Branchen der alternativen Streitklärungsmethoden, u.a. Schlichtung, Mediation und Cooperative Praxis, per se nicht technikaffin sind. Zudem anderen handelt es sich um eine Thematik, die zutiefst menschlich geprägt ist. Sie verlangt von den professionell Beteiligten Einfühlungsvermögen, das Erkennen von Gefühlen, Stimmungen und Empathie. – Alles Qualitäten, die KI nicht leisten kann oder KI nicht überlassen werden sollte. Arbeitsweisen und Strukturen von Konfliktmanagementsystemen könnten sich jedoch für die Anwendung von KI als Assistenz und Werkzeug eignen. Dieser mögliche Einsatz von KI soll hiermit einer ersten Analyse unterzogen werden.

2. These

Konfliktklärung und -lösung und damit auch das Konfliktmanagement in Unternehmen, Institutionen und Einrichtungen werden sich in Zukunft durch den Einsatz von KI verändern (vgl. Steffek 2023: 121ff).

3. Untersuchungsgegenstand

Die Fragestellung ist, wenn KI (sinnvoll) im Konfliktmanagement eingesetzt werden könnte, ob dies rechtlich und ethisch zulässig und/oder auch sinnvoll wäre. Da-

raus resultiert: Was sind die rechtlichen und ethischen Herausforderungen des Einsatzes von KI in Konfliktmanagementsystemen, besonders im Healthcare-Sektor?

4. Konfliktmanagement

Unter Konfliktmanagement werden Maßnahmen verstanden, die in Institutionen, Organisationen, Unternehmen mit dem Ziel des Erkennens und der Bewältigung von Konflikten sowie für die Deeskalation und zur Prävention eingesetzt werden (vgl. Schienle u.a. 2019: 3). Konfliktmanagement befasst sich dabei hauptsächlich mit den Konflikt-Prozessen, jedoch auch mit Konfliktpotential und -folgen (vgl. Glasl 2020: 21). Dabei werden beispielsweise Konflikte erfasst, die im Rahmen von Strategie, Planung und Umsetzung von Arbeitsprozessen sowie durch den Betrieb und Umstrukturierungen entstehen können. Zum Konfliktmanagement gehört auch die Mustererkennung von Strukturen, Prozessen, etc. – Hier könnte die Arbeitsweise von KI ansetzen. Gerade im Healthcare-Sektor, in dem Hierarchien und ritualisierte Arbeitsabläufe vorherrschen, kann es zu wiederkehrenden Konflikten kommen. Diese Streitigkeiten können intern (z.B. Konflikte im Team) oder extern (z.B. zwischen Ärzt:innen und Patient:innen) sein. Denkbar sind u.a. Konflikte über Ziele, Rollen, Methoden, Werte, u.ä. – Diese Streitigkeiten stören Arbeitsabläufe, vermindern die Professionalität, machen ggf. krank, usw.. Sogenannte ADR-Verfahren („Alternative Dispute Resolution“), wie Schlichtung oder Mediation, können hier das Mittel der Wahl sein, um die Konflikte zu klären und zu lösen (vgl. zu den verschiedenen Methoden Glasl 2020: 419ff).

5. Künstliche Intelligenz

Wenn es um den Einsatz von KI im Konfliktmanagement geht, liegt zunächst der Nutzen von generativen Sprachmodellen bzw. Large-Language-Modellen nahe (siehe Heetkamp u.a. 2023: 80ff). Um im Prozess der Konfliktbearbeitung Thesen, Strukturen, Muster, Ver-

strickungen, Themen, Interessen o.ä. zu visualisieren, können jedoch auch KI-Tools in Betracht kommen, die Bilder aus Textbeschreibungen aufgrund maschinellen Lernens erstellen (sogenannte Image-Creatoren).

6. Konfliktmanagement & KI

Im Rahmen des Konfliktmanagements besteht die Möglichkeit, KI als Assistenz und/oder Werkzeug der Konfliktmanager:in einzusetzen: z.B. in der Konfliktanalyse, für das Bilden von Arbeitshypothesen und den Perspektivwechsel. Dies gilt auch für die Mustererkennung von typischen Konflikten im Unternehmen, z.B. von Arbeitsprozessen, Machtungleichgewichten und Strukturen. Existieren aufgrund bestimmter Aufgabenverteilungen und Abläufe wiederkehrende Konflikte? Dies könnte bei Aufgaben relevant sein, die eine Zusammenarbeit von Ärzt:innen und Pfleger:innen voraussetzen, die Hierarchie-Ebenen überschreiten oder beim Ärzt:innen- und Patient:innengespräch. Hier kann KI präventiv zum Erkennen von typischen Konfliktpotential eingesetzt werden. Ergebnisse des Konfliktmanagementsystems könnten auch zu Schulungszwecken genutzt werden, um die Konflikt- und Kommunikationsfähigkeiten des Fachpersonals zu trainieren. Weitere Einsatzmöglichkeiten sind denkbar, um Prozesse hinsichtlich der Kommunikation und des Patientenmanagements zu optimieren.

Die Nutzung von KI kann nur im Rahmen von Recht und Ethik erfolgen. Daraus ergeben sich komplexe Herausforderungen und Perspektiven (siehe z.B. Hoeren u.a. 2022: 1ff). Der konkrete Anwendungsfall von KI bedarf daher einer individuellen Analyse und Prüfung.

7. Rechtliche Sicht

Bei der Untersuchung der rechtlichen Sicht, sind mindestens vier Ebenen zu unterscheiden:

- | allgemeine rechtliche Beurteilung des Einsatzes von KI,
- | Rahmen und Rolle des Rechts im Konfliktmanagement selbst,
- | rechtliche Zulässigkeit der Anwendung von KI im Konfliktmanagement,
- | Sondernormen des Healthcare-Sektors.

Zahlreiche Gesetze und Normen sind bei der allgemeinen rechtlichen Beurteilung von KI im deutschem und

europäischem Rechtsraum zu beachten: Von der Verfassung (GG) und den dort geregelten Grund- und Freiheitsrechten, über z.B. das BGB, UrhG, IT-SiG, TMG, Datenschutzrecht (BDSG, DSGVO), den Digital Markets Act (DMA), Digital Services Act (DAS), dem Daten-Governance-Gesetz (DGA) bis hin zum noch nicht in Kraft getretenen AI-ACT der EU (Artificial Intelligence Act). Zudem ist die EU-Grundrechte-Charta zu beachten.

Der AI-Act-Entwurf hat einen risikobasierenden Ansatz hinsichtlich der Entwicklung und des Einsatzes von KI. Er sieht vor, dass das KI-System in Risikogruppen einzuordnen ist, je nach Gefahr für Gesundheit, Sicherheit und Grundrechte. Dadurch sollen jeweils die Anforderungen an das KI-Tool und damit einhergehenden Pflichten geregelt werden (z.B. Folgenabschätzung für Grundrechte, Informations- und Transparenzpflichten). Im Trilog wurde vorläufig vereinbart, dass „Allzweck-KI“ (General Purpose AI), die für viele Zwecke verwendet werden kann, und „Basismodelle“ (Foundation Models), auf denen andere KI-Anwendungen aufbauen, unter den AI-Act fallen sollen (Rat der EU, 09.12.23). – Dies wird für Konfliktmanagementsysteme, die sich solcher KI-Modelle bedienen, relevant werden. – Die vorläufige Einigung zum AI-Entwurf sieht zudem das Verbot vor, KI zur kognitiven Verhaltensmanipulation oder zum Erkennen von Emotionen am Arbeitsplatz einzusetzen (Rat der EU, 09.12.23).

Die EU geht bei KI generell von der Betroffenheit gewichtiger Grundrechte der EU-Bürger:innen aus: „Durch ihre besonderen Merkmale (z. B. Undurchsichtigkeit, Komplexität, Datenabhängigkeit, autonomes Verhalten) kann die Verwendung von KI dazu führen, dass einige der in der EU-Grundrechtecharta ... verankerten Grundrechte verletzt werden. Der Vorschlag zielt darauf ab, diese Grundrechte in hohem Maße zu schützen und durch einen klar festgelegten risikobasierten Ansatz verschiedene Ursachen für Risiken anzugehen.“ (Europäische Kommission, 2021: Begründung, 3.5 Grundrechte). Durch den AI-Act soll ein Gleichgewicht zwischen der Förderung der KI-Technik und dem Schutz der Grundrechte geschaffen werden. – Dem Verhaltenskodex der G7-Staaten zu KI, dem sogenannten Code-of-Conduct (2023), liegen diese Werte und der Schutz der Sicherheit, Gesundheit und demokratischer Werte als Selbstverpflichtung ebenfalls zugrunde.

Im Hinblick auf die Trainingsdaten der KI sind außerdem Regelungen zum Data-Mining, also urheberrechtliche Überlegungen und sogenannten Fair-Use-Fragen von Bedeutung (vgl. §§ 44b, 60d UrhG). Bei der Anwendung von KI spielen auch haftungsrechtliche Fragen eine Rolle, insbesondere im Healthcare-Sektor.

Beim Konfliktmanagement muss sogenanntes „zwingendes“ Recht beachtet werden. Von sogenannten dispositiven Recht kann hingegen im Rahmen von alternativen Konfliktlösungen einvernehmlich seitens der Beteiligten abgewichen werden. Diese Unterscheidung des anzuwendenden Rechts ist je nach betroffenem Rechtsgebiet, z.B. Arbeitsrecht, Berufsrecht, etc., genau zu analysieren.

Bei der Untersuchung der rechtlichen Zulässigkeit der Anwendung von KI im Rahmen von Konfliktmanagementsystemen sind die Grund- und Freiheitsrechte der Betroffenen, auf Würde, Schutz des Allgemeinen Persönlichkeitsrechtes, des Rechts auf Informationelle Selbstbestimmung, der Datenschutz, das Gleichheitsgebot bzw. Diskriminierungsverbot, Transparenz- und Informationsgebote, das Verbot von automatisierten Entscheidungen, einschließlich des Profilings (siehe § 22 DSGVO) und menschliche Überprüfungsspflichten zu beachten. Sogenanntes „Social Scoring“, also die Bewertung und/oder Einstufung sozialen Verhaltens, soll nach dem Entwurf des AI-Acts unzulässig sein (vgl. bereits jetzt § 31 BDSG).

Je nach Branche und Konfliktthema müssen unter Umständen Besonderheiten des Arbeitsrechts, weitergehende Datenschutzbestimmungen, spezifische Branchengesetze, etc. beachtet werden. Im Gesundheitswesen gelten z.B. erhöhte Anforderungen an den Schutz von Patienten- bzw. Gesundheitsdaten, besonderes Berufsrecht sowie Regulierungen zur kritischen Infrastruktur und zum Produkthaftungsrecht.

8. Ethische Sicht

Eine KI-Ethik muss auch deshalb diskutiert werden, da bei der Entwicklung und Programmierung häufig die Chancen und Funktionalitäten der KI-Technik im Vordergrund stehen, ohne deren Wirkung angemessen

in einen sozialen Kontext zu stellen (vgl. Hoeren, u.a. 2023: 4). Um die ethischen Fragestellungen in Bezug auf Konfliktmanagement untersuchen zu können, ist zuerst ein Perspektivwechsel hin zum Menschen im „System“ zu vollziehen. Wie ist die (psychische) Verfasstheit der vom Konflikt Betroffenen? – Sie befinden sich in Streit und Stress, also in einer Belastungs- und Ausnahmesituation. Damit müssen die Verantwortlichen für das Konfliktmanagement äußerst verantwortungsbewusst umgehen. Dies kann nicht einem KI-System überlassen werden.

Folgende Fragen sind ebenfalls relevant: Zu welchem konkreten Zweck soll KI verwendet werden? Welche Trainingsdaten und damit (Fach-)Inhalte liegen zugrunde? Wäre der Einsatz tatsächlich gewinnbringend für die Anwender:innen und Betroffenen? Führt er zu Arbeitserleichterungen, so dass die Verantwortlichen des Konfliktmanagements, beispielsweise Coaches, Schlichter, Supervisor:innen und Mediator:innen entlastet werden und die jeweilige Expertise ergänzt? Würde damit das Verfahren und die Ergebnisse des Konfliktmanagements insgesamt verbessert werden? Steht der Mensch im Mittelpunkt, mit seinen Interessen, Bedürfnissen und Gefühlen? – Häufig werden bei Überlegungen zu KI und Ethik Begriffe wie Vertrauenswürdigkeit, Transparenz, Informiertheit (über die Nutzung von KI), Nachvollziehbarkeit, Überprüfbarkeit, die Vermeidung von Diskriminierung und Voreingenommenheit angeführt (vgl. Grimm u.a. 2022: 167). Teilweise werden diese bereits im Recht abgebildet. Zudem sind Gleichwertigkeit, Gerechtigkeit und Fairness Werte, die dem KI-System zugrunde liegen und vom Menschen kontrolliert werden müssen. – Diese „Verankerung“ könnte in der Folge dann auch die Tätigkeit der Konfliktmanager:in erleichtern bzw. verbessern. Der Dialog mit einem KI-gestützten Sprachmodell könnte beispielsweise die eigene, mögliche Voreingenommenheit der Konfliktberater:innen sichtbar machen, den zu vollziehenden Perspektivwechsel und die Neutralität erleichtern, das Verstehen und Verständnis für die Konfliktparteien, z.B. im Dialog mit einem Sprachmodell verbessern, beim Bilden von Arbeits-hypothesen helfen usw.. Dies würde jedoch voraussetzen, dass ethische Werte, wie z.B. Diskriminierungsfreiheit im KI-Tool und dessen Programmierung

unabänderbar festgeschrieben werden (=„Ethics by Design“ – Grimm u.a. 2022: 167). Zudem sind ethische und politisch begründete Entscheidungen notwendig, was KI darf und was nicht, wo und wie sie eingesetzt wird und welche Bedingungen sie dafür erfüllen muss. Es gehört dabei auch zu den ethischen Überlegungen dazu, sich mit den Interessen der Anbieter des KI-Systems (z.B. bzgl. Daten und Rechte) angemessen auseinanderzusetzen und dies entsprechend zu regulieren.

9. Fazit

KI sollte im Rahmen von Konfliktmanagementsystemen ausschließlich als „erweiterte“ Intelligenz eingesetzt werden und ein humanistisches Menschenbild mit dessen Werten unveränderbar im Zentrum stehen. Als Assistenz-Tool, Werkzeug zum Dialog, zur Reflexion und ggf. als Entscheidungshilfe kann KI in Konfliktmanagementsystemen genutzt werden, dies gilt grundsätzlich auch im Healthcare-Sektor.

Der Einsatz von KI darf den Menschen jedoch nicht ersetzen, sondern sollte ihm „dienen“. Der Bedarf und der rechtliche sowie ethische Rahmen muss für die jeweilige Institution und für die konkrete Anwendung genau analysiert, geprüft und im KI-System implementiert werden. Im Gesundheitswesen sind dabei die rechtlichen und ethischen Anforderungen höher und sensibler einzuschätzen als in anderen Branchen.

Die Nutzung von KI muss stets von den Verantwortungsträger:innen (an-)geleitet und kontrolliert werden.

Menschliche Interaktion ist also notwendig („human in the loop“). Z.B. muss auch eine Überprüfung sogenannter „Halluzinationen“ stattfinden. Das die KI „bedienende“ Fachpersonal sollte entsprechend geschult werden, z.B. auch hinsichtlich des „Promptings“, der Kontrollkompetenz, des Erkennens von Bias und falschen Ergebnissen, etc. – Dann könnte die Verwendung von KI dazu beitragen, durch die Verbindung von „Mensch und Maschine“ Konflikte, auch in Institutionen, Unternehmen und Organisationen, in Zukunft besser zu managen und diesen ggf. vorzubeugen. – Frei nach Habermas erfordert der Einsatz von hoch komplexen Systemen, wie KI, den Einsatz von professionalisiertem Personal, von sogenannten „Gatekeepern“ (Habermas 2022: 39 zu digitalen Medien). Die Diskussion verlangt meines Erachtens jedoch ebenfalls, dass sich nicht nur Expert:innen, sondern auch Generalist:innen interdisziplinär miteinander auseinandersetzen. Dabei bedarf es eines gesellschaftspolitischen Diskurses, in dem wir uns zurzeit befinden. Über das „Ob“ (Verwendung von KI) wurde scheinbar bereits entschieden. Jetzt geht es um das „Wie“ und die Beantwortung der umfassenden Frage: „In welcher Welt wollen wir leben?“ – Für diesen gesellschaftlichen Aushandlungsprozess eignen sich hervorragend Methoden aus dem Konfliktmanagement (ohne oder mit KI?) ...

Literaturangaben:

Europäische Kommission, Vorschlag AI-Act, COM(2021) 206 final, 2021/0106(COD) Brüssel 21.04.21, Online: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206> [abgerufen am 08.12.2023].

Glasl, F.; Konfliktmanagement, Ein Handbuch für Führung, Beratung und Mediation, 12. Aufl., Stuttgart 2020.

Grimm, P.; Keber, T. O.; Zöllner, O.: Digitale Ethik, Leben in vernetzten Welten, Bonn 2022.

Habermas, J.: Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik, Berlin 2022.

Heetkamp, S. J.; Piroutek, C.: ChatGPT in Mediation und Schlichtung, in: zkm 3/23, S. 80ff.

Hoeren, T., (Hrsg.) Pinelli, S.: Künstliche Intelligenz – Ethik und Recht, München 2022.

Rat der EU, Presse Mitteilung 09.12.23, 01:27, Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world, Online: <https://www.consilium.europa.eu/de/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> [abgerufen am 09.12.2023].

Steffek, F.: Die Veränderung der Konfliktlösung durch künstliche Intelligenz – Teil 2, in: zkm 4/23, S. 121ff.

Olivia Alig

Frankfurt University of Applied Sciences
Rechtsanwältin & Mediatorin
Fachbereich 4: Soziale Arbeit und Gesundheit
olivia.alig@medienanwaeltin.de
alig@fb4.fra-uas.de

Lebenslauf in Kürze

Olivia Alig arbeitet als Rechtsanwältin und Mediatorin in Frankfurt/M. mit den Schwerpunkten Medien, Vertrag, Familie, Wirtschaft & Digitales. Zudem ist sie an der Frankfurt University of Applied Sciences, Fachbereich 4, als Lehrbeauftragte tätig (Medienrecht, Kinder- & Jugendmedienschutz). Sie ist verheiratet und hat zwei Kinder.



Nach ihrem Abitur 1988 studierte Olivia Alig zunächst Germanistik und Kunstgeschichte an der J.W.-Goethe Universität Frankfurt/M. Im Sommersemester 1990 nahm sie dann ihr Studium der Rechtswissenschaften an der J.W.-Goethe Universität Frankfurt/M. auf. Während ihres Studiums war sie als wissenschaftliche Hilfskraft und Tutorin im Dekanat und am Lehrstuhl Prof. Dr. Ingolf Pernice (Öffentliches Recht, Völker- und Europarecht) sowie nach dem 1. Staatsexamen 1996 bei Prof. Dr. Rudolf Wiethölter (Wirtschaftsrecht) an der J.W.-Goethe Universität tätig. Im Sommer 1996 verbrachte sie einen Studienaufenthalt mit Stipendium am Europäischen Hochschulinstitut in Florenz (Europarecht). Ihr anschließendes Referendariat am Landgericht Frankfurt/M. ab 1997, mit Stationen u.a. an der HfG Offenbach (Kanzlerin) und bei Sony Music, beendete sie 1999 mit ihrem 2. Juristischen Staatsexamen. Im selben Jahr wurde sie als Rechtsanwältin zugelassen. Von 1999 – 2002 war sie in der Kanzlei Kornmeier Rechtsanwälte Frankfurt/M. tätig (Schwerpunkt Medien- & Musikrecht). Seit März 2002 ist Olivia Alig Inhaberin ihrer eigenen (Medien-) Kanzlei in Frankfurt am Main, u.a. im Bereich Medien-, Urheber-, Internet- und Vertragsrecht sowie Recht der Digitalisierung und KI. Von 2010 – 2016 war sie zudem Geschäftsführerin (Business – und Legal Affairs) der Footage Online GmbH, Frankfurt/M. Sie absolvierte eine Weiterbildung als Verfahrensbeiständige (2016 – 2017, Zertifikat Paritätisches Bildungswerk Bundesverband, anerkannt vom BVEB e.V) und eine Mediationsausbildung mit anschließender Zertifizierung bei IKOM Frankfurt/M. (2018 – 2019, Familie & Wirtschaft). Seit 2019 ist sie als zertifizierte Mediatorin tätig. Inzwischen ist sie lizenzierte Mediatorin BAFM, Mediatorin BM® und erhielt 2023 das Gütesiegel als QVM®-Mediatorin. 2020/21 nahm sie u.a. an Fortbildungen in Online-Mediation und Kooperativer Praxis (DVCP®) teil. Seit 2020 ist Olivia Alig ehrenamtliches Mitglied im Bundesfachausschuss „Digitales Leben“ des DKSB Bundesverband e.V. (Beratung u.a. zum JuSchG, UN-KRK, Comment No. 25) und seit WS 18/19 als Lehrbeauftragte an der Frankfurt University of Applied Sciences, FB 04 für Medienrecht & Kinder-/Jugendmedienschutz tätig.

Publikationen u.a.: Mediation – eine Chance für die Medienbranche? in: Die Mediation, IV/2022, S. 66ff und Sharenting, Mama-Blogger, Kinderinfluencer & Co. – Eine rechtliche Betrachtung, in: BPJMAktuell 4/2021, S. 9ff.

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

Seit Jahren befasse ich mich als Rechtsanwältin mit Rechtsfragen im Zusammenhang mit Medien, Internet, der Digitalisierung, dem Jugendmedienschutz und mit aktuellen technischen Entwicklungen. Als Mediatorin beschäftige ich mich mit Konfliktmanagement sowie der Klärung und Lösung von (digitalen) Konflikten, auch im Online-Format, in den Bereichen Vertrag, Medien, Familie und Wirtschaft. Zudem interessieren mich die Chancen der alternativen Konfliktlösung bei Streitigkeiten im Urheberrecht. Mein Fokus gilt neuen Technologien, auch in Bezug auf das Thema Künstliche Intelligenz. Hierzu setze ich mich jedoch nicht nur kritisch mit dem Anwendungs-

potential, sondern ebenfalls mit den damit einhergehenden Risiken und Herausforderungen auseinander. Dazu gehört auch, den rechtlichen und ethischen Rahmen beim Einsatz von KI zu analysieren. Während meiner Lehrtätigkeit an der Frankfurt University of Applied Sciences, FB 04 (Medienrecht, Kinder- & Jugendmedienschutz) habe ich Gelegenheit, in den Austausch mit Studierenden zu treten und dabei auch das Thema KI wissenschaftlich zu vertiefen.

Meiner Meinung nach geht es bei der Diskussion über KI, wie im Konfliktmanagement, um verschiedene individuelle und gesellschaftliche Interessen und Bedürfnisse. Sie müssen im Rahmen von Recht und Ethik miteinander in Einklang gebracht werden, um den Nutzen, die Chancen und Risiken sinnvoll miteinander abzuwägen. Diesen Diskurs möchte ich gerne erweitern und mich darüber interdisziplinär austauschen. Erfreulicherweise ist das aufgrund der Teilnahme, u.a. an den Projekten von Frau Prof. Dr. Barbara Klein, FUAS, FB 04, möglich (Woman KI Days, digitale Buchveröffentlichung: KI im Healthcare-Sektor).

Titel Beitrag: „KI im Konfliktmanagement – rechtliche & ethische Aspekte“

Name(n): Olivia Alig, Rechtsanwältin & Mediatorin, Lehrbeauftragte FUAS

Institution(en): Medien- & Mediationskanzlei Ffm und FUAS, FB 04

Susanne Beck

Rechtlicher Rahmen des Einsatzes von KI in der Medizin

In der modernen Medizin spielt Künstliche Intelligenz (KI) eine immer größere Rolle. Im Folgenden sei ein Überblick über rechtliche Herausforderungen und Rahmenbedingungen gegeben, die mit dem Einsatz von KI in der Medizin einhergehen. Dabei wird ein Fokus auf den Ausgleich zwischen innovativen medizinischen Technologien und dem Schutz der Patientenrechte gelegt.

Dabei ist keinesfalls zu vernachlässigen, dass die Einführung von KI in der Medizin die Diagnostik und Behandlung revolutioniert und wichtige Vorteile mit sich bringt. KI-Systeme können Muster in großen Datenmengen erkennen, was Ärzt:innen hilft, präzisere Diagnosen zu stellen. Dennoch ergeben sich hieraus rechtliche Fragen, unter anderem bezüglich der Verantwortlichkeit für Diagnoseentscheidungen und dem rechtlichen Umgang mit den Risiken von KI-Systemen. Diese vorab zu diskutieren, ggf. gesetzgeberisch zu reagieren und einige Antworten auf die Fragen schon jetzt zu finden, ist dabei nicht zwingend nachteilig für die technologische Entwicklung, im Gegenteil, oft entstehen Nachteile eher durch zu spät diskutierte rechtliche Herausforderungen.

I. Überblick über einige rechtliche Aspekte

Im Folgenden sei auf einige rechtliche Aspekte hingewiesen, die sich bei der Nutzung von KI im medizinischen Bereich als problematisch darstellen können.

1. Aufklärung und Einwilligung

Die informierte Einwilligung ist bekanntermaßen ein Eckpfeiler im medizinischen Behandlungsprozess. Insofern wird diskutiert, ob Patienten auch über die Integration von KI in den Diagnose- und Behandlungsprozess aufgeklärt werden müssen. Vieles spricht dafür, zumindest derzeit aufgrund der Neuartigkeit der Technologie eine solche Aufklärung, auch über mögliche Risiken und die Grenzen der Technologie, durchzuführen.

Dabei besteht eine Herausforderung in der Vermittlung komplexer KI-Prozesse in einer für Laien verständlichen Form. Auch ist fraglich, ob die bloße Aufklärung über die Tatsache, dass KI genutzt wird, ausreichend ist. Stattdessen sollte jedenfalls dann, wenn die Ärztin die KI bei der Entscheidungsfindung befragt, auch die Patientenautonomie durch Einbeziehung in diesen Prozess gestärkt werden. Dabei ist darauf zu achten, dass die KI die Entscheidung des Patienten nicht unangemessen beeinträchtigt, vgl. dazu sogleich unter II.

2. Zivilrechtliche Haftung

Bei Fehlern und Schäden, die durch KI-Systeme in der Medizin verursacht werden, stellt sich die Frage nach der zivilrechtlichen Haftung der Beteiligten. Dabei wird diskutiert, ob traditionelle Haftungskonzepte, wie das Produkthaftungsrecht, auf KI anwendbar sind, oder ob jedenfalls zusätzlich neue rechtliche Rahmenbedingungen geschaffen werden müssen. Die Klärung der Haftungsverteilung zwischen Herstellern, Anwendern (z.B. Ärzten) und möglicherweise den Entwicklern ist schwierig, zugleich ist sie entscheidend, um Vertrauen in die Technologie zu schaffen und Patientenschutz zu gewährleisten. Hierzu gibt es derzeit viele rechtswissenschaftliche und gesetzgeberische Debatten, um angemessene Lösungen und Konzepte zu entwickeln, wobei unter anderem Gefährdungshaftungen oder spezifische Versicherungslösungen diskutiert werden. Wichtig ist an dieser Stelle jedenfalls, dass die Haftung nicht auf unangemessene Weise auf die Letztentscheiderin abgewiesen werden sollte, d.h. etwa die Ärztin, die KI ohne hinreichende Einführung oder Kenntnis des KI-Systems nutzt und sogar nutzen muss, weil die Klinik dieses System angeschafft hat, dann aber ohne Einschränkung für fehlerhafte Entscheidungen, die sie gemeinsam mit der KI getroffen hat, persönlich haften soll. Die Angemessenheit einer derart weitgehenden Haftung lässt sich durchaus plausibel hinterfragen. Gerade zivilrechtlich scheint es plausibler, wenn die materielle Absicherung

bzw. die Kompensation von Schäden durch die Beteiligten erfolgt, die tatsächlich von der Herstellung und dem Einsatz der KI profitieren – hier kann das Zivilrecht plausible Lösungen finden, da es in diesem Rechtsgebiet primär um materiellen Ausgleich geht.

3. Strafrechtliche Verantwortlichkeit

Anders ist es mit Blick auf die strafrechtliche Verantwortlichkeit – das Strafrecht blickt jeweils auf die konkret handelnde Person und fragt ausschließlich danach, ob diese Person sich strafbar gemacht hat. Dabei schließt die Strafbarkeit einer Person die Strafbarkeit anderer Beteiligten nicht aus. Es ist also konkret zu diskutieren, wie sich die verschiedenen Beteiligten im KI-Entwicklungsprozess, aber auch bei der Diagnose- oder Behandlungsentscheidung, individuell strafbar gemacht haben. Bezüglich der an der Entwicklung Beteiligten gilt das in anderen Kontexten entwickelte Produktstrafrecht, wobei zu beachten ist, dass die Anforderungen bei der Entwicklung eines als riskant angesehenen Produkts nicht zu hoch angesetzt werden sollten.

Darüber hinaus stellt sich im Kontext des KI-Einsatzes in der Medizin die Frage auf, inwiefern Ärzte für gemeinsam mit KI getroffene Entscheidungen zur Verantwortung gezogen werden können, also ob und in welchem Maße Ärzte für die Überprüfung der von KI-Systemen gelieferten Empfehlungen verantwortlich sind. Dabei ist zu beachten, dass die KI aufgrund der geringen Transparenz ihrer Funktionsweise, der zugrundeliegenden Daten oder der schweren Nachvollziehbarkeit von Fehlfunktionen ein schwer zugängliches Werkzeug ist. Insofern spricht vieles dafür, jedenfalls dann, wenn die KI ohne weitere Erklärung nur eine „Ja-oder-Nein“-Alternative anbietet und ggf. sogar noch unter Zeitdruck zu entscheiden ist, eine Verantwortlichkeit in Frage zu stellen. Strafrechtsdogmatisch ist hier etwa die Frage aufzuwerfen, ob die Schädigung tatsächlich als „Werk“ der handelnden Ärztin anzusehen ist. Eine umfassende individuelle Verantwortung für jede gemeinsam mit der KI getroffene Entscheidung lässt sich also bezweifeln, genaueres hierzu sogleich unter II.

4. Datenschutzrechtliche Aspekte

Der Einsatz von KI in der Medizin erfordert die Verarbeitung großer Mengen sensibler Patientendaten, was erhebliche datenschutzrechtliche Implikationen

mit sich bringt. Die Herausforderung für das Recht, aber auch für die Praxis, liegt insofern darin, Datenschutz und Datenverfügbarkeit für das Training und die Anwendung von KI-Systemen in Einklang zu bringen. Besondere Aufmerksamkeit erfordert insofern Ausgestaltung der Einwilligung zur Datenverarbeitung und -speicherung, aber auch die Sicherstellung von Transparenz und Kontrolle über die künftige Datennutzung. Zudem ist problematisch, dass die Einwilligung eigentlich jederzeit widerrufen werden kann, zugleich nach einer Einspeisung in ein Lernendes System eine Isolierung und „Herauslösung“ spezifischer persönlicher Daten kaum realisierbar erscheint. Insofern wäre eine Anpassung des Datenschutzrechts an die Herausforderungen durch die verstärkte Nutzung von KI in vielerlei Hinsicht sicherlich wünschenswert.

5. Zulassungsrecht

Das Zulassungsverfahren für Medizinprodukte garantiert Sicherheit und Wirksamkeit der Produkte und ist somit von entscheidender Bedeutung. Zugleich ist das derzeitige Zulassungsverfahren mit Blick auf KI problematisch, geht es doch davon aus, dass ein Medizinprodukt zum Zulassungszeitpunkt „fertig“ ist und nicht mehr verändert wird. Ein Weiterlernen der Systeme ist damit nicht vereinbar. Das wiederum steht jedoch im Widerspruch zu der Funktion der Systeme, die gerade von ihren Fehlern lernen und sich weiterentwickeln sollen. Deshalb besteht hier die Herausforderung darin, geeignete Zulassungsverfahren zu entwickeln, die sowohl die Besonderheiten von KI, also kontinuierliches Lernen und Anpassung, als auch die Notwendigkeit einer robusten klinischen Bewertung berücksichtigen. Dabei geht es auch um die Schaffung von Standards und Richtlinien, die die Qualität und Sicherheit von KI-Systemen sicherstellen.

II. Das Konzept der „Meaningful Human Control“

Wie sich im Überblick über die rechtlichen Probleme gezeigt hat, bestehen wichtige Herausforderungen durch die Nutzung von KI in der Medizin mit Blick auf die (individuelle) Verantwortlichkeit der Beteiligten. Das ist nicht nur dann der Fall, wenn die KI autonom agiert, sondern bereits dann, wenn die Ärztin gemeinsam mit der KI Diagnose- oder Therapie-Entscheidungen trifft. Denn auch dann ist kaum noch nachvollziehbar, warum eine Entscheidung ggf. fehlerhaft

war. Gerade wenn die Ärztin nicht selbst entschieden hat, die KI einzusetzen, sondern dies von der Klinikleitung entschieden wurde, erscheint eine umfassende Haftung in jedem Fall, in dem die Ärztin mit der KI entscheidet, unangemessen. Denn es ist durchaus denkbar, dass sich der Fehler daraus ergibt, dass das System fehlerhaft trainiert wurde oder im weiteren Verlauf dazu gelernt hat. Oder es ist eben überhaupt nicht erklärbar, warum eine falsche Diagnose oder Therapie vorgeschlagen wurde.

Nun könnte man argumentieren, die Ärztin habe ja die „Letztentscheidung“, ja dies wird sogar oft gefordert in dem Sinne, dass eben letztlich ein Mensch die Entscheidung treffen solle. Dabei muss jedoch eine pauschale Zurechnung dieser Entscheidung zum menschlichen Entscheider kritisch gesehen werden. Kriterium für die Zuschreibung muss vielmehr die Kontrolle, die Herrschaft über das System sein. Um eine Fehlentscheidung der Ärztin vorwerfen zu können, bedarf es eines gewissen Grads an Entscheidungsgewalt über das System. Mit Blick hierauf wurde das Konzept der Meaningful Human Control (bedeutsame menschliche Kontrolle) entwickelt, ursprünglich im Kontext Autonomer Waffensysteme, inzwischen aber auch in anderen Kontexten von Bedeutung. Dabei geht es normativ also darum, dass zum einen Verantwortung nur dann zugeschrieben werden soll, wenn der jeweilige Mensch eine solche Form der Kontrolle über das Lernende System hatte. Zum anderen ist grundsätzlich sicherzustellen, dass Menschen in zentralen Lebensbereichen generell eine solche Kontrolle über die KI behalten. Die konkrete Ausgestaltung einer solchen bedeutsamen Kontrolle ist letztlich von den Umständen des Einzelfalls abhängig. Dazu gehört Transparenz, die Möglichkeit, die Gründe für die Entscheidungsvorschläge zu

verstehen oder Nachfragen zu stellen. Somit ist die Ermöglichung dieser Kontrolle bereits bei der technischen Entwicklung der Systeme mitzudenken (MHC by design). Auch im Zulassungsverfahren kann solche Kontrolle ausgeübt werden, wenn etwa die Betroffenen für die Frage nach der Zulassung eine Stimme erhalten – das gilt auch für Patient:innen.

Letztlich bedarf es eines interdisziplinären, wissenschaftlichen Diskurses über die Kriterien, die einer bedeutsamen menschlichen Kontrolle im Einzelfall zugrunde liegen. Diese Kontrolle kann für jeden Lebensbereich, für jedes lernende System anders ausgestaltet sein. Das Konzept dient als Orientierung, als Rahmen für das Design, für die Zulassung, aber auch für den Gesetzgeber.

Deshalb sollte auch de lege ferenda über Lösungsmöglichkeiten diskutiert werden, wie man bedeutsame menschliche Kontrolle auch gesetzgeberisch durchsetzen kann. Nur so können Entscheidungen tatsächlich noch als „menschliche Entscheidungen“ angesehen werden, nur so bleibt individuelle Verantwortlichkeit in diesen Kontexten möglich.

III. Abschließende Überlegungen

Abschließend sei noch einmal betont, dass der dringend nötige rechtliche Rahmen für den Einsatz von KI in der Medizin dynamisch und anpassungsfähig sein muss, um sowohl den technologischen Fortschritt als auch die Patientensicherheit zu gewährleisten. Interdisziplinäre Forschung und entsprechende Debatten, die rechtliche, ethische und medizinische Expertise vereinen, sind für die Entwicklung von Standards, Richtlinien sowie Gesetzen, die den Einsatz von KI in der Medizin einhegen, in Zukunft unerlässlich.

Literaturangaben:

- Ballestrem, J. Graf u.a. (Hrsg.): Künstliche Intelligenz – Rechtsgrundlagen und Strategien in der Praxis, Wiesbaden 2020.
Beck, S.: Zum Einsatz von Robotern im Palliativ- und Hospizbereich, *Medizinrecht*, 2018, S. 772-778.
Ebers, M.; Heinze, C.; Krügel, T.; Steinrötter, B. (Hrsg.) Künstliche Intelligenz und Robotik, München 2020.
Katzenmeier, Ch.: Big Data, E-Health, M-Health, KI und Robotik in der Medizin, *Medizinrecht*, 2019, S. 259-271.
Münch, F.: Autonome Systeme im Krankenhaus, Baden-Baden 2017.

Prof. Dr. Susanne Beck, LL.M. (LSE)

Leibniz Universität Hannover
Kriminalwissenschaftliches Institut
susanne.beck@jura.uni-hannover.de

Lebenslauf in Kürze

Susanne Beck studierte Rechtswissenschaften in Würzburg und an der London School of Economics. Ihr juristisches Referendariat absolvierte sie in Würzburg und Sydney (AUS). 2006 promovierte sie zu „Stammzellforschung und Strafrecht“. Im Anschluss arbeitete sie am United International College in Zhuhai (China). Von 2008 bis 2012 war sie Wissenschaftliche Mitarbeiterin in Würzburg, wo sie sich 2013 zu „Strafrecht und Kollektive“ habilitierte. Seit 2013 hat sie den Lehrstuhl für Strafrecht, Strafprozessrecht, Strafrechtsvergleichung und Rechtsphilosophie an der Leibniz Universität Hannover inne. Susanne Beck forscht u.a. zu „Moderne Technologien und Recht“, Medizin(straf)recht und Rechtstheorie. Sie ist u.a. Mitglied von acatech und im Vorstand der Akademie für Ethik in der Medizin. Ein interdisziplinäres Forschungsprojekt zur Nutzung von KI in der Medizin (vALID) wurde gerade erfolgreich abgeschlossen, derzeit ist sie u.a. an einem BMBF-Projekt zu Meaningful Human Control im Kontext Autonomer Waffensysteme beteiligt (MeHuCo).



Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... ist es zu beobachten, wie sich durch diese Entwicklung neue Fragen der individuellen Verantwortung und der Haftung stellen. Dabei ist mir besonders wichtig, die Zurechnung gerecht zu verorten, insbesondere nicht die Letztentscheiderin in unangemessener Weise zu benachteiligen – ein besonders wichtiges Konzept hierfür ist die „Meaningful Human Control“. Hiernach ist individuelle Verantwortung nur dann angemessen, wenn die entsprechende Person bedeutsame Kontrolle über das Geschehen hatte. Weiterhin ist mir besonders wichtig, ein Bewusstsein für mögliche Diskriminierungen durch KI hervorzurufen und so solche Diskriminierungen weitgehend zu verringern – soweit ich das als Wissenschaftlerin kann.

Janina Loh

Zur ethischen Evaluation von Technologien am Beispiel von ChatGPT

In den anwendungsorientierten, alltagsnahen und praxisbezogenen Bereichsethiken wie etwa der Technik-, Medizin- und Pflegeethik steht die ethische Reflexion und kritische Einschätzung von Sachverhalten, Handlungsoptionen und -mitteln auf der Tagesordnung. So hat etwa insbesondere in der Ethikberatung sowie im Rahmen der Durchführung und Moderation ethischer Fallbesprechungen im Gesundheitssektor und in der Sozialen Arbeit das Modell von Tom L. Beauchamp und James F. Childress (2009), das auf den vier Prinzipien der Autonomie, Fürsorge, Nichtschaden und Gerechtigkeit baut, Tradition. Mit Blick auf eine ethische Bewertung von Technologien v.a. für die Medizin, Therapie und Pflege erlangte das MEESTAR-Modell (Manzeschke et al. 2013; Weber 2015) eine gewisse Bekanntheit. Schließlich lässt sich noch das am Institut Mensch, Ethik und Wissenschaft (IMEW) entwickelte und vom Bundesministerium für Bildung und Forschung geförderte Projekt FreTiP – Fragen zur ethischen Reflexion von digitalen Technologien in der Pflegepraxis (Grüber et al. 2023) als aktuelles Beispiel für ein solches Evaluationsmodell anführen.

Das im Folgenden vorgestellte Modell zur ethischen Evaluation von Technologien, das ich entwickelt habe (zuletzt in Loh/Grote 2023), ist mehr als Ergänzung zu diesen und weiteren bestehenden Modellen zu verstehen und nicht als eine Alternative für eben jene. In der Tat ist es überdies weder auf eine spezifische ethische Schule wie etwa die deontologische, konsequenzialistische oder Tugendethik beschränkt noch überhaupt auf eine ethische Evaluation. Mit demselben Modell lässt sich bei Bedarf ebenso bspw. eine politische, ökonomische oder rechtliche Bewertung einer Technologie vornehmen.

Im Folgenden stelle ich mein Vier-Dimensionen-Modell am Beispiel einer KI-Technologie vor, die auch im Gesundheitssektor von wachsender Präsenz ist, nämlich generative KI in Form von ChatGPT – etwa mit Blick auf

Möglichkeiten der personalisierten Medizin, der Kommunikation mit Patient:innen und Klient:innen oder der Interpretation unstrukturierter medizinischer Daten.

Im Rahmen der ersten Dimension ergeben sich ethische Fragen, die Herstellung und Design einer jeweiligen Technologie thematisieren. Darunter fallen z.B. Fragen, die darauf zielen, unter welchen Herstellungs-, etwa Arbeits- und Umweltbedingungen, sie entwickelt wurde. Hierzu gehören auch Fragen mit Blick auf die Bedingungen für die involvierten Lebewesen (Menschen, Tiere, Pflanzen). Außerdem zählen dazu Fragen, die sich auf das Design und damit auf die äußere und innere Gestaltung, einer Technologie konzentrieren – welche Werte etwa in Form von Stereotypen in diese (bewusst oder unbewusst) eingegangen sind.

Hinsichtlich ChatGPT ist im Rahmen dieser Kategorie etwa zu fragen, in welcher Weise Menschen und Nichtmenschen wie Tiere und Pflanzen in die Herstellung dieser Technologie involviert waren? Kam es zu Ausbeutung, Unterdrückung und Diskriminierung? (Stichwort »Clickwork«) Welche Menschen waren in die Entwicklung von ChatGPT bei OpenAI, Sama etc., involviert? War Diversität gewährleistet? Hier ergibt sich überdies ein Brückenschlag zur vierten Dimension. Wird ChatGPT nachhaltig hergestellt? Woher kommen die nötigen Ressourcen? Wie viel Energie »kostet« etwa eine Anfrage bzw. das Training von ChatGPT (bzw. von LLMs generell)? Und schließlich soll ChatGPT offensichtlich nicht nach irgendwelchen Geschlechterkriterien entworfen sein. Doch ist diese Technologie tatsächlich frei von jeglichen (diskriminierenden/problematischen) Geschlechtszuschreibungen? Sind sonstige Stereotypisierungen in ChatGPT verborgen? An dieser Stelle lässt sich eine Verbindung zur dritten Dimension herstellen.

In der zweiten Dimension geht es um die (ethischen) Fragen, die sich mit Blick auf das Thema Autonomie

und Aufgabenbereich stellen. Diese Dimension dreht sich etwa um Fragen, die die autonome Wirkungsweise einer Technologie betreffen, die sich also darauf richten, was die fragliche Technologie quasi allein machen können sollte. Darüber hinaus geht es auch um Fragen, die darauf abzielen, was der Wirkungsbereich bzw. die Aufgabe und damit der (primäre) Sinn und Zweck einer jeweiligen Technologie ist, wo und wie sie also agieren können sollte. Mit Blick auf ChatGPT ist im Rahmen dieser Dimension etwa zu fragen, wie (sofern das überhaupt möglich ist) sichergestellt werden kann, dass ChatGPT ausschließlich »die Wahrheit« erzählt (Stichwort »Halluzination«). Sollte ChatGPTs Autonomie nicht in diesem Zusammenhang eingeschränkt werden und inwiefern würde das die ethische Evaluation von ChatGPT beeinflussen? Weiterhin ist zu fragen, ob ChatGPT auf bestimmte Anfragen gar nicht oder lediglich in bestimmter Weise antworten können sollte? Was wissen wir darüber, inwiefern ChatGPTs Antwortmöglichkeiten (also ChatGPTs Autonomie in diesem Zusammenhang) durch OpenAI beschränkt sind? Mit dieser Frage ergibt sich ein Brückenschlag zur dritten Dimension. Schließlich sollte im Rahmen dieser Dimension in den Blick genommen werden, ob ChatGPT in bestimmten Institutionen oder Branchen gar nicht oder lediglich in bestimmter Weise genutzt werden sollte? Und damit ergibt sich ein Brückenschlag zur vierten Dimension.

Die dritte Dimension dreht sich um die (ethischen) Fragen, die sich innerhalb des Themenbereichs Daten und Sicherheit stellen. Damit sind zunächst einmal jene Fragen angesprochen, die die im Einsatz einer jeweiligen Technologie erhobenen Daten betreffen. Es geht aber auch um Fragen, die die Sicherheit einer Technologie im Rahmen ihres Einsatzes fokussieren – einerseits wiederum mit Blick auf die Daten, andererseits aber auch hinsichtlich der sonstigen Soft- aber auch Hardware der fraglichen Technologie. Hinsichtlich ChatGPT ist innerhalb dieser Dimension etwa zu fragen, welche Daten über jene, die ChatGPT nutzen, erhoben werden, wer davon in welcher Weise profitiert bzw. wer darauf zugreifen kann? Außerdem stellt sich die Frage – da ja das Training einer bestimmten Version von ChatGPT zu einem konkreten Zeitpunkt abgeschlossen war –, welche Daten im Rahmen des Trainings genutzt wurden bzw. welche Daten in eine jeweilige Version von ChatG-

PT (nicht mehr) eingehen. Darüber hinaus stellt sich die Frage, wie sicher die Daten Dritter, auf die OpenAI (für das Training von ChatGPT etwa) zugegriffen hat, sind (siehe bspw. den Fall einer kurzzeitigen Sperrung von ChatGPT in Italien aufgrund der Sorge um die Sicherheit der personenbezogenen Daten der italienischen Bevölkerung, die für das Training von ChatGPT genutzt wurden). Und schließlich bleibt die Frage nach der Sicherheit der Hardware, auf der ChatGPT jeweils installiert ist bzw. die für den Einsatz von ChatGPT gebraucht wird (wie etwa Server etc.).

Die vierte Dimension fokussiert nun die (ethischen) Fragen, die sich mit Blick auf Kontext und Einsatzbereich einer jeweiligen Technologie stellen. Das sind insbesondere Fragen bzgl. der Arbeit und Lebensumstände von im Einsatzbereich einer Technologie tätigen Menschen und Tieren. Denn ChatGPT verändert unsere Arbeit – etwa im Dienstleistungssektor, im Journalismus, in der Bildung und im Gesundheitswesen. Welche gesellschaftlichen Auswirkungen hat die Herstellung und serienmäßige Produktion von LLMs wie ChatGPT auf die in diesen Bereichen tätigen Menschen? Welche langfristigen gesellschaftlichen Auswirkungen könnte der Einsatz von LLMs wie ChatGPT haben und sind die Risiken und negativen Folgen durchweg (moralisch) gerechtfertigt? Unsere aktuelle Sicht auf digitale und etwa generative KI wie z.B. ChatGPT ist zumeist dualistisch und dadurch unangemessen einseitig. Es ist voreilig, ChatGPT ausschließlich entweder als reine Erleichterung für die Menschen, die diese Technologie nutzen, zu sehen oder als Gefahr für Arbeitsplätze und Menschlichkeit.

Wollen wir dieses Vier-Dimensionen-Modell zur Evaluation einer Technologie nun anwenden, ist es sinnvoll, folgendermaßen vorzugehen: Zunächst einmal ist die zu bewertende Technologie selbst so konkret wie möglich festzulegen. Handelt es sich bspw. um eine (bestimmte) generative KI, ein LLM, eine App oder einen Roboter? Darüber hinaus ist zu klären, wo der Einsatz der fraglichen Technologie erfolgen soll – z.B. in einem konkreten Unternehmen, einer Schule, einem Altenpflegeheim oder geht es um die private Nutzung? Schließlich muss die eigentliche Frage, die im Rahmen der Beurteilung beantwortet werden soll, definiert werden. Geht es vielleicht darum zu entscheiden, ob sich eine Einrichtung im Gesundheitswesen die jewei-

lige Technologie zulegen sollte oder soll eine Technologie, die bereits im Einsatz ist, zwischenevaluieren werden?

Sind diese Fragen beantwortet, sollte geklärt werden, welche Evaluation konkret durchgeführt werden soll – etwa eine ethische, eine ökonomische oder politische – und welche Maßstäbe und Prinzipien für die Bewertung innerhalb der Einrichtung oder dem Unternehmen bereits zur Verfügung stehen. Verfügt die Einrichtung vielleicht über Haltungssätze oder ethische Leitlinien, gibt es in dem Unternehmen eine Unternehmensphilosophie oder ethische Richtlinien? Auch die DSGVO enthält (zumindest implizit) ethische Kategorien, die mit bedacht werden müssen. Außerdem bleibt zu klären, inwiefern das individuelle moralische Gewissen jener, die die Evaluation durchführen, für die Beurteilung eine Rolle spielt? Gibt es schließlich spezifische moralische Werte wie z.B. Autonomie, Sicherheit oder Fürsorglichkeit, denen ein besonderer Stellenwert zukommt?

Sind alle diese Fragen beantwortet, ›durchläuft‹ die fragliche Technologie die vier Dimensionen. Ist das geschafft, bleibt zu definieren, in welcher Weise die Ergebnisse gesichert und die Entscheidung zur Beantwortung der Ausgangsfrage der Evaluation getroffen werden soll. Es wäre z.B. denkbar, dass Antworten auf die Fragen innerhalb der vier Dimensionen (und evtl. sogar die Gesamtentscheidung mit Blick auf eine jeweilige Dimension) mit Symbolen bzw. Zeichen veranschaulicht werden – etwa ein + für Ja, ein – für Nein, ein ? dann, wenn die Antwort (moralisch) uneindeutig ausfällt –, um den Abwägungs- und Entscheidungsfindungsprozess zu erleichtern. Auch wäre vorstellbar, dass hierfür Zahlen bzw. Werte genutzt werden, wie bspw. ein Wert von 1 oder 2, wenn die Antwort auf eine Frage (moralisch) gut oder sehr gut ausfällt, eine 0, wenn die Antwort (moralisch)

uneindeutig ist sowie ein Wert von -1 oder -2, je nachdem ob die Antwort auf eine Frage etwa (moralisch) problematisch oder nicht akzeptabel ausfällt.

Schließlich bleibt darüber nachzudenken, welche Form das Ergebnis der Evaluation haben soll. Handelt es sich etwa um einen Zahlenwert? Wie wird dieses Ergebnis in eine durchführbare Entscheidung übersetzt – etwa im Sinne von: Ja, wir wollen in unserer Einrichtung die fragliche Technologie nutzen, aber unter folgenden Bedingungen...? Wie sieht der Entscheidungsfindungsprozess konkret aus bzw. wer entscheidet und wie? Gibt es etwa eine unabhängige Moderation, wenn es sich um einen demokratischen Gruppenentscheidungsfindungsprozess handelt? Was ist zu tun, wenn viele der Antworten (moralisch) uneindeutig sind oder wenn aus anderen Gründen die Entscheidung zum aktuellen Zeitpunkt noch nicht getroffen werden kann?

Abschließend bleibt festzuhalten, dass die vier Dimensionen meines Evaluationsmodells zwar sicherlich für alle Technologien relevant und dass sie auch miteinander verzahnt sind. Allerdings sind individuelle Gewichtungen der vier Dimensionen mit Blick auf konkrete Technologien festzustellen, die Sensibilität und ein gesteigertes ethisches Bewusstsein erfordern. Das Modell lässt sich in unterschiedlichen Kontexten (zur privaten Evaluation, in einer Einrichtung in der sozialen Arbeit etc.) und mit unterschiedlicher disziplinärer Ausrichtung (ethisch, ökonomisch, politisch etc.) anwenden. Auch sind verschiedene Vorgehensweisen, Entscheidungsfindungsprozesse und Formen der Ergebnissicherungen möglich. Das Ziel ist, mit diesem Bewertungsmodell ein pragmatisches und möglichst intuitiv anwendbares Werkzeug für die Evaluation von Technologien bereitzustellen. Ich habe die Vorlage für einen Bewertungsbogen entworfen, den ich gerne auf Anfrage zur Verfügung stelle.

Literaturangaben:

Beauchamp, T. L.; Childress, J. F.: Principles of Biomedical Ethics. 6. Auflage. Oxford/New York 2009.

Grüber, K.; Loevsckaya, E.; Ossino, V.: Fragen zur ethischen Reflexion von digitalen Technologien in der Pflegepraxis (FreTiP). Institut Mensch, Ethik und Wissenschaft gGmbH. 2023. Online: https://www.google.com/urlsa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKewiji6Czt7GCAxUwg_OHHVgEC2wQFnoECAOQAQ&url=https%3A%2F%2Fwww.imew.de%2Ffileadmin%2FDokumente%2FVolltexte%2FIMEW_FreTiP-2023_barrierefreie.pdf&usq=A0vVawOWXKNGO2488BjntxniHy&opi=89978449 [abgerufen am 7.11.2023].

Loh, J.; Grote, G.: Einleitung: Medizin – Technik – Ethik. Spannungsfelder zwischen Theorie und Praxis. In: Loh, J. u. a.: Medizin – Technik – Ethik. Spannungsfelder zwischen Theorie und Praxis. Techno:Phil – Aktuelle Herausforderungen der Technikphilosophie. Band 5. Berlin 2023, S. 1–14.

Manzeschke, A.; Weber, K.; Rother, E.; Fangerau, H.: Ergebnisse der Studie "Ethische Fragen im Bereich Altersgerechter Assistenzsysteme". VDI/VDE Innovation + Technik GmbH. 2013.

Weber, K.: MEESTAR: Ein Modell zur ethischen Evaluierung sozio-technischer Arrangements in der Pflege- und Gesundheitsversorgung. In: Weber, K. u. a.: Technisierung des Alltags - Beitrag für ein gutes Leben?. Stuttgart 2015, S. 247–262.

Evaluationsbogen

Prof. Dr. Janina Loh
 Stiftung Leibniz, Subjektive Ethik
 Max-Planck-Ethik der Technik und ihrer sozialen Kontexte
 in BSC, Zentrum für Ethik und Verantwortung (ZEV)
janina.loh@uni-wuerzburg.de

EVALUATIONSBOGEN

Datum: _____

Protokollant*in: _____

Weitere Beteiligte an der Evaluation: _____

Technologie: _____

Ort: _____

Form der Evaluierung (ethisch, ökonomisch, rechtlich, politisch...): _____

Bewertungsmaßstab: _____

Relevante Werte und Prinzipien: _____

Frage: _____

Seite 3 von 6

Prof. Dr. Janina Loh
 Stiftung Leibniz, Subjektive Ethik
 Max-Planck-Ethik der Technik und ihrer sozialen Kontexte
 in BSC, Zentrum für Ethik und Verantwortung (ZEV)
janina.loh@uni-wuerzburg.de

1. DIMENSION: HERSTELLUNG UND DESIGN

Frage	-2.000	-1.000	0.000	+1.000	+2.000
In welcher Weise waren Menschen und Nichtmenschen in die Herstellung involviert? Kam es zu Ausbeutung, Unterdrückung und Diskriminierung?					
Welche Menschen waren in die Entwicklung involviert? Wer Diversität gewährleistet?					
Wird die Technologie nachhaltig hergestellt? Woher kommen die nötigen Ressourcen?					
Sind in die Technologie Standards implementiert (z.B. hinsichtlich Gesundheit, Alter etc.)?					

BEWERTUNGEN MIT BLICK AUF HERSTELLUNG UND DESIGN, DIE FÜR DIE EVALUATION RELEVANT SIND:

Seite 2 von 6

Prof. Dr. Janina Loh
 Stiftung Leibniz, Subjektive Ethik
 Max-Planck-Ethik der Technik und ihrer sozialen Kontexte
 in BSC, Zentrum für Ethik und Verantwortung (ZEV)
janina.loh@uni-wuerzburg.de

2. DIMENSION: AUTONOMIE UND AUFGABENBEREICH

Frage	-2.000	-1.000	0.000	+1.000	+2.000
Ist die Funktion und Aufgabe der Technologie eindeutig und klar definiert?					
Sollte der Einsatz der Technologie in spezieller Form eingeschränkt werden? Sollte sie bestimmte Aufgaben nicht ausführen dürfen?					
Wird durch die Autonomie der Technologie die Autonomie der involvierten Menschen gestützt?					
Wird durch die Autonomie der Technologie die Autonomie der involvierten Menschen beeinträchtigt?					
Ist die Technologie intuitiv verständlich und erlernbar für die involvierten Menschen?					

BEWERTUNGEN MIT BLICK AUF AUTONOMIE UND AUFGABENBEREICH, DIE FÜR DIE EVALUATION RELEVANT SIND:

Seite 3 von 6

Prof. Dr. Janina Loh
 Stiftung Leibniz, Subjektive Ethik
 Max-Planck-Ethik der Technik und ihrer sozialen Kontexte
 in BSC, Zentrum für Ethik und Verantwortung (ZEV)
janina.loh@uni-wuerzburg.de

3. DIMENSION: DATEN UND SICHERHEIT

Frage	-2.000	-1.000	0.000	+1.000	+2.000
Ist der Nutzer der im Einsatz der Technologie erhobenen Daten (schongest) klar informiert von den Daten?					
Wird Missbrauch angemessen vorgebeugt?					
Wirden Persönlichkeitsrechte (z.B. mit Blick auf personenbezogene private Daten und Informationen) im Einsatz der Technologie geschützt?					
Ist die Technologie hinsichtlich der in ihrem Einsatz erhobenen Daten transparent designed?					
Ist die Technologie sicher mit Blick auf die Soft- aber auch die Hardware?					

BEWERTUNGEN MIT BLICK AUF DATEN UND SICHERHEIT, DIE FÜR DIE EVALUATION RELEVANT SIND:

Seite 4 von 6

Prof. Dr. Janina Loh
 Stiftung Leibniz, Subjektive Ethik
 Max-Planck-Ethik der Technik und ihrer sozialen Kontexte
 in BSC, Zentrum für Ethik und Verantwortung (ZEV)
janina.loh@uni-wuerzburg.de

4. DIMENSION: KONTEXT UND EINSATZBEREICH

Frage	-2.000	-1.000	0.000	+1.000	+2.000
Ist der Einsatz der Technologie mit Blick auf die involvierten Menschen (z.B. ihre Arbeit) positiv zu bewerten?					
Fördert meine/unsere Nutzung dieser Technologie die Gesellschaft? Unterstützt sie mich/uns bei der Lösung der Technologie in positiver Weise gesellschaftliche Situationen?					
Welchen Prozessrisiken durch den Einsatz der Technologie dokumentiert (auch außerhalb des Rahmens, in dem ich/wir sie einsetze)?					
Sind die (auch langfristig zu erwartenden) Risiken und negativen Folgen im Einsatz dieser Technologie (moralisch/ökonomisch/politisch) rechtlich gerechtfertigt?					

BEWERTUNGEN MIT BLICK AUF KONTEXT UND EINSATZBEREICH, DIE FÜR DIE EVALUATION RELEVANT SIND:

Seite 5 von 6

Prof. Dr. Janina Loh
 Stiftung Leibniz, Subjektive Ethik
 Max-Planck-Ethik der Technik und ihrer sozialen Kontexte
 in BSC, Zentrum für Ethik und Verantwortung (ZEV)
janina.loh@uni-wuerzburg.de

5. DIMENSION: GEWISSEN UND VERANTWORTUNG

Die nächsten Schritte

Wie ist wofür zuständig?

Seite 6 von 6

Prof. Dr. Janina Loh

Stiftung Liebenau

Janina.loh@stiftung-liebenau.de

Hochschule Bonn-Rhein-Sieg

janina.loh@h-brs.de

Lebenslauf in Kürze

Janina Loh (geb. Sombetzki) ist Ethiker:in auf einer Stabsstelle Ethik bei der Stiftung Liebenau in Meckenbeuren am Bodensee und hat eine Honorarprofessur an der Hochschule Bonn-Rhein-Sieg am Zentrum für Ethik und Verantwortung (ZEV) für Ethik der Technik und ihrer sozialen Kontexte inne. 2018 erschien von Loh die erste deutschsprachige Einführung in den Trans- und Posthumanismus (Junius, 4. Auflage 2023, koreanische Übersetzung 2021). Loh publizierte 2019 eine Einführung in die Roboterethik (Suhrkamp). Lohs viertes Buch / Habilitationsprojekt entwirft eine Kritisch-posthumanistische Ethik der Companionship für die Wissensräume (Arbeitstitel).

Meine Leidenschaft im Bereich der künstlichen Intelligenz ...

... liegen insbesondere in der Roboterethik, in der feministischen Technikphilosophie sowie der Ethik der künstlichen Intelligenz und Digitalisierung. Ich interessiere mich für die Weisen einer Evaluation von Technologien und für die ethischen Fragen, die sich v.a. mit Blick auf Robotik, KI und Digitalisierung im Gesundheitswesen, in der Sozialen Arbeit, im autonomen Fahren und hinsichtlich der Bildung emotionaler Bindungen mit Technologien stellen.



Impressum

Herausgeberin: Prof. Dr. Barbara Klein,
Frankfurt University of Applied Sciences

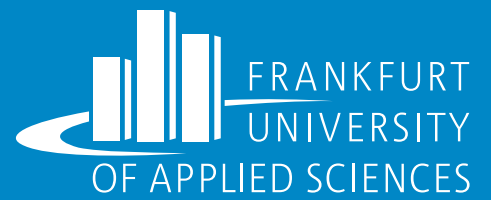
Redaktion: Prof. Dr. Barbara Klein, Paulin Krause.
Für inhaltliche Korrektheit und Vollständigkeit der
Angaben übernimmt die Redaktion keine Gewähr.

Gestaltung: Frank Muckenheim, Frankfurt am Main

Bildnachweis: Titel: © xartproduction | Stock.Adobe.
com; Seite 9: © Frankfurt UAS | Kevin Rupp; Seite 11: ©
Frankfurt UAS | Nils Urbach; Seite 13: © Frankfurt UAS
| Kevin Rupp, © Frankfurt UAS; Seite 29: © Frankfurt
UAS | Kevin Rupp; Seite 34: © Patrick Bal; Seite 44:
© Peter Witkop; Seite 53: © Frankfurt UAS | Kevin Rupp,
© Frankfurt UAS; Seite 73: © Frankfurt UAS | Kevin
Rupp; Seite 99: © Balsereit; Seite 105: © Olivia Alig;
Seite 110: © Franz Fender; Seite 115: © Lukas Beck
Alle anderen Bilder: Quelle privat bzw. © Frankfurt UAS

1. Auflage: 01/2024

Doi: 10.48718/1cw9-3c06



Frankfurt University of Applied Sciences

Nibelungenplatz 1

D-60318 Frankfurt am Main

Tel. 0 69 15 33-2131

www.frankfurt-university.de