kom.sd
Kompetenzzentrum Soziale Dienste

Künstliche Intelligenz in der Wohlfahrtspflege

Eine empirische Studie im Rahmen der wissenschaftlichen Begleitung des Bundesförderprogramms „Zukunftssicherung der Freien Wohlfahrtspflege durch Digitalisierung“

des Bundesministeriums für Bildung, Familie, Senioren, Frauen und Jugend (BMBFSFJ)

Hannes Breder
Dr. Konstantin Rink
Prof. Dr. Udo Seelmeyer

31. Dezember 2025

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 International Lizenz
http://creativecommons.org/licenses/by-nc-nd/4.0/

Die Veröffentlichung stellt keine Meinungsäußerung des BMBFSFJ dar. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei der Autorin/dem Autor.



Gefördert vom:

Bundesministerium
für Bildung, Familie, Senioren,
Frauen und Jugend

EXECUTIVE SUMMARY **1**

1 Einleitung 4
2 Grundlagen und historische Entwicklung von KI 6
3 Anwendungsbereiche und Potentiale von KI in der Sozialen Arbeit 11
3.1 Dialogsysteme *11*
3.2 Entscheidungsunterstützung *14*
3.3 Unterstützung im Alltag *16*
3.4 Unterstützung der alltäglichen Dokumentation und Information *18*
4 Herausforderungen und Risiken 20
4.1 Verzerrungen und fehlerhafte KI-Berechnungen *20*
4.2 Irrtümer über Fähigkeiten und Eigenschaften von KI *21*
4.3 Verantwortung und Transparenz in KI-Systemen *23*
4.4 Kompetenzen & Qualifizierung *25*
5 Aktuelle Überblicksstudien zum KI-Einsatz im Dritten Sektor in Deutschland 26
5.1 Forschungsbericht: Civic Coding *26*
5.2 Arbeitspapiere: KI für Nichtregierungsorganisationen *27*
5.3 Forschungsbericht: Künstliche Intelligenz in der Sozialwirtschaft *28*
6 Fragestellung und Erhebungsdesign der Studie 30
6.1 Forschungslücken und Fragestellung *30*
6.2 Methodisches Vorgehen *30*
7 Auswertung der Interviews 35
7.1 Übersicht der erhobenen Interviews und Vorstellung der folgenden Kategorien- & typengeleiteten Auswertung *36*
7.2 Innovationssemantiken *38*
7.3 Soziale Welten *43*
7.4 Innovationsprozess *49*
8 Inhaltsanalytische Auswertung der Fokusgruppen 54
8.1 Formelle und informelle Nutzung von KI *54*
8.2 Zukunftsbilder in der Fokusgruppe: Zwischen organisational Neuem und professioneller Bedrohung *55*
9 Die Freie Wohlfahrtspflege als spezifischer Kontext für KI-Innovationen 60
9.1 Stellung der Freien Wohlfahrtspflege im deutschen Sozialstaat *60*
9.2 Föderale Organisationsstruktur und ihre Implikationen für KI-Innovationen *61*
9.3 Zur Frage der Feldspezifik und Übertragbarkeit *62*
10 Gelingensbedingungen und Herausforderungen von KI-Innovationen in der Wohlfahrtspflege 63
10.1 Herausforderungen *63*
10.1 Gelingensbedingungen *67*
11 Fazit 71
11.1 Limitationen der Studie *71*
11.2 Einordnung Ergebnisse vor dem Hintergrund des Diskurses *72*
11.3 Ausblick und Ableitungen *74*

ABBILDUNGSVERZEICHNIS **78**

TABELLENVERZEICHNIS **78**

ABKÜRZUNGSVERZEICHNIS **78**

LITERATURVERZEICHNIS **79**

Executive Summary

Seit 2019 unterstützt das BMBFSFJ (bis 2025: BMFSFJ) die Spitzenverbände der Freien Wohlfahrtspflege im Bundesförderprogramm „Zukunftssicherung der Freien Wohlfahrtspflege durch Digitalisierung" bei der Bewältigung der Herausforderungen der Digitalisierung. Ziel des Programms ist es, innovative digitale Ansätze zu entwickeln und zu erproben, um die Zukunftsfähigkeit der Freien Wohlfahrtspflege zu sichern. Die wissenschaftliche Begleitung durch das Kompetenzzentrum Soziale Dienste (kom.sd) zielt einerseits darauf ab, Maßnahmen und Ergebnisse der Verbände zu evaluieren, die Fortschritte im Gesamtprogramm zu bilanzieren und Handlungsempfehlungen abzuleiten. Darüber hinaus wurde im Förderjahr 2025 eine Studie zu KI in der Wohlfahrtspflege durchgeführt, deren Ergebnisse mit diesem Bericht vorgelegt werden.

Dieser Bericht gibt Einblicke in die wissenschaftliche Auseinandersetzung mit Künstlicher Intelligenz in der Sozialen Arbeit und untersucht den aktuellen Status quo von konkreten KI-Innovationen in der deutschen Wohlfahrtspflege. Dazu wurde eine qualitative Studie durchgeführt, die Expert:inneninterviews (n=12) sowie eine Fokusgruppendiskussion mit Fachkräften umfasst. Die Analyse identifiziert Innovationsstrategien auf Projekt- und Leitungsebenen sowie Nutzungsbarrieren und Risikowahrnehmungen auf der operativen Ebene. Ziel ist die Ableitung evidenzbasierter Steuerungsimpulse für Politik und Verbände.

Zentrale Befunde

- **Drei Typen der Innovation:** Die Wohlfahrtspflege agiert zwischen *Pragmatischer Modernisierung* (Einkauf von Standardsoftware, mit Fokus auf Effizienz), *Reflexiver Sozialinnovation* (Partnerschaften mit Wissenschaft/ Startups) und *Disruptiver Eigenentwicklung* (Aufbau technischer Souveränität durch „Pioniere“).
- **Weniger Projekte als gedacht:** Der Feldzugang über die Spitzenverbände der Freien Wohlfahrtspflege bot eine gute Feldübersicht, dabei wurden jedoch weniger strategische Implementierungsprojekte vorgefunden, als es die aktuelle Popularität des Themas vermuten lässt. Es ist zu erwarten, dass in den nächsten Jahren zahlreiche neue Projekte in Forschung und Praxis der Sozialen Arbeit umgesetzt werden.
- **Die „Projekt-Falle“:** Innovationsvorhaben scheitern häufig nicht an technischer Machbarkeit, sondern an der Struktur der Förderlandschaft. Nach

Ablauf der Anschubfinanzierung fehlt es an Modellen für den Dauerbetrieb und die Wartung von Software (Transferproblem).

- **„Schatten-KI" durch Fachkräfte:** Während formelle KI-Strategien in Einrichtungen oft fehlen oder durch strenge Datenschutzvorgaben blockiert sind, deutet vieles darauf hin, dass Fachkräfte privat längst Tools wie ChatGPT nutzen. Dies führt zu unregulierten, datenschutzrechtlich riskanten „informellen Praktiken".
- **Akzeptanz durch Entlastung:** Fachkräfte lehnen KI in der direkten Interaktion (Beziehungsarbeit) eher ab, fordern aber vehement eine Entlastung von administrativen „Zeitfressern".

Strategische Handlungsempfehlungen

i. Für die Zuwendungsgeber (Ministerien auf Bund- und Länderebene)

- **Paradigmenwechsel in der Förderung:** Weg von der reinen Projektförderung (Aufbau) hin zur Infrastrukturförderung (Betrieb). Software ist kein statisches Produkt, sondern ein Prozess. Förderrichtlinien müssen Wartung, Updates und Weiterentwicklung als refinanzierbare Kosten anerkennen, um Investitionsruinen zu vermeiden.
- **Förderung technologischer Souveränität:** Um Abhängigkeiten von wenigen globalen Tech-Akteuren zu mindern, sollten Open-Source-Ansätze und verbandsübergreifende Plattformlösungen („Wohlfahrts-Cloud") gezielt incentiviert werden.

ii. Für Spitzenverbände und Träger der Freien Wohlfahrtspflege

- **Formalisierung statt Tabuisierung:** Die informelle Nutzung ist faktisch vorhanden. Träger müssen diesen „Graubereich" durch bereitgestellte, datenschutzkonforme Experimentierräume (Sandbox-Umgebungen) in formelle Bahnen lenken.
- **Re-Framing der Kommunikation:** Ein akzeptanzförderndes Narrativ könnte in den Fokus rücken, dass KI einerseits bei der Bewältigung administrativer Aufgaben unterstützt, damit mehr Zeit für den Menschen bleibt, zugleich aber auch in fachlichen Fragen und bei der Sozialen Arbeit mit Adressat:innen unterstützen kann.
- **Synergien bündeln:** Die Studie zeigt, dass Eigenentwicklungen teilweise an Ressourcengrenzen stoßen. Es empfiehlt sich die Bildung von Entwicklungskonsortien, um die Innovationskraft einzelner Pioniere für den Gesamtverband nutzbar zu machen und Skaleneffekte zu erzielen.

Die Soziale Arbeit steht unter hohem Innovationsdruck. Derzeit droht eine Schere zwischen einer unregulierten Nutzung an der Basis und stagnierenden Pilotprojekten auf Leitungsebene. Eine erfolgreiche digitale Transformation erfordert *nachhaltige* Finanzierungsstrukturen seitens der Politik und eine *mutige, regulierende* Gestaltungskraft seitens der Verbände, um KI vom Risiko zum Werkzeug der Professionalität zu wandeln.

1 Einleitung

Seit 2019 fördert das Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend (BMBFSFJ, bis 2025 BMFSFJ) die sechs Spitzenverbände der Wohlfahrtspflege und die Bundesarbeitsgemeinschaft der Freien Wohlfahrtspflege (BAGFW) durch das Förderprogramm „Zukunftssicherung der Freien Wohlfahrtspflege durch Digitalisierung“. Ziel des Programms ist es, entlang thematischer Leitlinien innovative Lösungsansätze im Kontext der Digitalisierung zu entwickeln und zu erproben.

Im Rahmen der wissenschaftlichen Begleitung erfolgt eine fortlaufende formative Evaluation, um die Maßnahmen der Verbände zu analysieren und die Digitalisierung der Freien Wohlfahrtspflege zu unterstützen. Während in den Berichten aus den Vorjahren das Förderprogramm und dessen Outputs wie Outcomes primär adressiert wurden, legt der nachfolgende Bericht für das Förderjahr 2025 einen expliziten Schwerpunkt auf das Thema Künstliche Intelligenz (KI), dessen fachliche Relevanz für das Feld der Wohlfahrtspflege insbesondere seit der Veröffentlichung generativer Sprachmodelle massiv an Bedeutung gewonnen hat.

Die zentrale Zielsetzung dieses Berichts geht über eine reine Bestandsaufnahme und Deskription bestehender KI-Innovationsprojekte hinaus. Es geht darum, die spezifischen Innovationslogiken zu rekonstruieren, die sich aktuell in der Wohlfahrtspflege herausbilden. Im Kern der Untersuchung steht die Identifikation und Analyse von drei distinkten Innovationstypen, die unterschiedliche Rollen im digitalen Wandel des Sektors einnehmen. Der Bericht zielt darauf ab, die jeweiligen Logiken, beteiligte soziale Welten und entsprechende Innovationssemantiken zu beschreiben und zu verstehen. Im Anschluss soll aufgezeigt und kontrastiert werden, wie diese Typen mit den realen Nutzungsbedingungen an der Basis – insbesondere der informellen Nutzung von KI durch Fachkräfte – kollidieren oder korrespondieren.

Der Aufbau des Berichts folgt dieser analytischen Logik und nähert sich dem Thema schrittweise an. Zunächst werden in den Kapiteln 2 und 3 die technologischen Grundlagen von KI sowie ihre potenziellen Anwendungsfelder in der Sozialen Arbeit, etwa in Form von Dialogsystemen oder zur Entscheidungsunterstützung, dargelegt. Daran anschließend widmet sich Kapitel 4 den ethischen und praktischen Herausforderungen, wie etwa algorithmischen Verzerrungen und Transparenzfragen, gefolgt von einer Einordnung des Forschungsstands in Kapitel 5, die auf einer Auswertung zentraler Überblicksarbeiten basiert.

Den empirischen Kern der Arbeit bilden die Kapitel 6 bis 9. Nach der Vorstellung des methodischen Designs in Kapitel 6 erfolgt in Kapitel 7 die detaillierte Herleitung der

oben genannten Innovationstypen auf Basis von Expert:inneninterviews. Ergänzend dazu analysiert Kapitel 8 die Perspektiven von Fachkräften, die in einer Fokusgruppendiskussion erhoben wurden, wobei vor allem die informelle Nutzung und die Zukunftsbilder beleuchtet werden. In Kapitel 9 werden die Befunde aus den Expert:inneninterviews und der Fokusgruppe diskursiv eingeordnet.

2 Grundlagen und historische Entwicklung von KI

Wer sich, unabhängig von Branche oder Fachgebiet, mit Fragen der Digitalisierung auseinandersetzt, kommt um den Sammelbegriff der Künstlichen Intelligenz nicht herum. Seit dem Meilenstein der Einführung von ChatGPT im November 2022 sind in nahezu allen gesellschaftlichen Bereichen rasante Entwicklungen zu beobachten, die zu einer weiten Verbreitung der Technologie in Alltagsanwendungen und zu vielfältigen Einsatzmöglichkeiten geführt haben. Anders als frühere technologische Fortschritte, die sich oft über Jahre oder gar Jahrzehnte vollzogen, erfolgen gegenwärtige Innovationen und Leistungssteigerungen im Bereich der KI mitunter im Monatstakt. Die größten Technologiekonzerne, darunter Microsoft, Meta und Google, übertreffen sich nicht nur gegenseitig in der Weiterentwicklung der Fähigkeiten ihrer KI-Modelle, sondern investieren auch in bislang beispiellosem Umfang in den Ausbau von Rechenkapazitäten[1] sowie in die Rekrutierung hochqualifizierter Entwickler:innen[2].

Obwohl die jüngsten Fortschritte oft als radikal neu wahrgenommen werden, ist die Idee einer „intelligenten Maschine“ keineswegs ein Produkt des 21. Jahrhunderts, sondern Teil einer langen ideengeschichtlichen Entwicklung. Wie Macsenaere (2024) nachzeichnet, reicht das Konzept bis in die Antike zurück, wo erste Vorstellungen von selbstständigen, menschenähnlichen Apparaten entstanden. In der Frühen Neuzeit wurden diese Konzepte in Form aufwendiger mechanischer Automaten weiterentwickelt[3], die allerdings eher technische Illusionen als tatsächlich intelligente Maschinen darstellten (vgl. ebd. S. 15f.). Einen entscheidenden theoretischen Impuls lieferte Alan Turing 1950 mit seinem Aufsatz „Computing Machinery and Intelligence“ und der Einführung des sogenannten Turing-Tests als Kriterium für maschinelle Intelligenz und setzte damit einen Maßstab, der bis heute in der Diskussion computerbasierter Dialogsysteme häufig zitiert wird (vgl. Hatch et al. 2025, S. 13). In den 1960er-Jahren demonstrierte Joseph Weizenbaums Computerprogramm „ELIZA“ dann, dass Computer bereits damals in der Lage waren, in textbasierten Dialogen die Illusion eines menschlichen Gesprächspartners zu

[1] Allein das Projekt „Stargate“ in Partnerschaft von Investoren, Softwareanbietern und Regierung plant Ausgaben iHv. 500 Milliarden Dollar in den nächsten vier Jahren vgl. Eidam und Bauer 2025, S. 8.

[2] Tech-Konzern Meta bietet ausgewählten Spitzenforscher:innen bis zu 300 Millionen Dollar für vier Jahre (t3n 07.2025)

[3] Wolfgang von Kempelens „Mechanischer Schachspieler“ erweckte den Eindruck, das Gerät spiele Schach ohne menschliche Intervention

erzeugen[4]. Erste Demonstratoren, wie Weizenbaums „ELIZA“ oder Simon & Newells „General Problem Solver“, führten jedoch zu einer euphorischen Überschätzung der Potenziale im Bereich der Künstlichen Intelligenz (Hatch et al. 2025; vgl. Buxmann und Schmidt 2021, S. 5). Dennoch: Die damals noch auf simplen und regelbasierten Algorithmen basierenden Erfolge bildeten die Grundlage für die theoretische Auseinandersetzung mit dem Ziel, Computern menschenähnliche Intelligenz zu verleihen.

Zur Umsetzung dieser großen Ideen fehlten jedoch sowohl Rechen- als auch Speicherkapazitäten (vgl. Kreidenweis 2024, S. 12.). Enttäuschung über ausbleibende technologische Durchbrüche, gepaart mit überzogenen Erwartungen und anschließender Kritik, führte dazu, dass das junge Forschungsfeld in den 1970er und frühen 1990er Jahren mehrere sogenannte „KI-Winter“ durchlebte. Diese Phasen waren von einer Skepsis innerhalb der entwickelnden Community sowie gekürzten Förderungen aufgrund abnehmenden öffentlichen Interesses gekennzeichnet (vgl. Connolly und Gill 2015, S. 108ff.). In dieser Zeit wurde zudem die begriffliche Unterscheidung zwischen „starker KI“, verstanden als hypothetische, menschenähnliche Intelligenz, und „schwacher KI“, die auf spezifische, eng umgrenzte Aufgabenfelder beschränkt ist, eingeführt (vgl. Searle 1980, S. 417). Gleichwohl kam die Forschung im Bereich der Künstlichen Intelligenz nicht vollständig zum Erliegen. Vielmehr verschob sich der Fokus auf eben jene Formen schwacher KI, etwa Expertensysteme, wissensbasierte Diagnosetools oder industrielle Optimierungsverfahren, die trotz ihrer eingeschränkten Flexibilität erste praktische Erfolge vorweisen konnten (vgl. WBGU 2019, S. 65). Kurz vor der Jahrtausendwende erzeugte der Wettstreit zwischen IBMs Deep Blue und Schachweltmeister Gari Kasparow, der im Sieg der Maschine endete, weltweit mediale Aufmerksamkeit und leitete eine neue Phase des allgemeinen Interesses an KI ein (vgl. Macsenaere 2024, S. 17). Mit dem Übergang ins 21. Jahrhundert setzte ein rasanter Fortschritt in der Hardware-Entwicklung ein. Die Zahl der Transistoren auf Prozessoren verdoppelte sich weiterhin in regelmäßigen Abständen, entsprechend dem sogenannten Mooreschen Gesetz. Parallel dazu entwickelte sich die Grafikprozessor-Technologie (GPU), die durch ihre Fähigkeit zur massiv parallelen Datenverarbeitung eine Schlüsselrolle für das spätere Training komplexer KI-Anwendungen übernahm (vgl. WBGU 2019, S. 77). Ab den späten 2000er-Jahren stieß die Miniaturisierung der Transistoren jedoch an physikalische Grenzen: Die nur noch wenige Nanometer großen Bauteile ließen sich kaum weiter verkleinern.

[4] ELIZA reagierte nach Vorbild der klientenzentrierten Gesprächsführung auf bestimmte Schlüsselwörter mit passenden, oft weiterführenden Fragen (z. B. „Ich habe ein Problem mit meinem Vater.“ → „Erzählen Sie mir mehr über Ihre Familie!“).

Seitdem wird nennenswerte Leistungssteigerung im Bereich des Computing verstärkt durch horizontale Skalierung erreicht, insbesondere durch Cloud Computing und verteilte Systeme, die eine umfassende Parallelisierung von Rechenprozessen ermöglichen. Mindestens ebenso bedeutsam wie der Hardware-Fortschritt war die zunehmende Datafizierung seit den frühen 2000er-Jahren. Gemeint ist die Repräsentation der physischen Welt in maschinenlesbaren und quantifizierbaren Datenpunkten (vgl. WBGU 2019, S. 73). Ein zentraler Wendepunkt war dabei der Zeitpunkt, ab dem mehr digitale als analoge Daten existierten (vgl. Hilbert und López 2011, S. 60). Plattformen wie soziale Netzwerke, Suchmaschinen und Videodienste erzeugten enorme Mengen an Text-, Bild- und Videodaten, die sich als Trainingsmaterial für maschinelle Lernverfahren eigneten. Hinzu kam das „Internet der Dinge“ (Internet of Things - IoT), das kontinuierlich Echtzeitdaten aus Industrie, Verkehr und privaten Anwendungen generierte. Sensoren, vernetzte Geräte und intelligente Maschinen liefern dabei Informationen über Prozesse, Umweltbedingungen und Nutzerverhalten, wodurch die digitale Abbildung der physischen Welt immer vollständiger und granularer wird (vgl. WBGU 2019, S. 70f.). In diesem Zusammenhang prägte Humby (2006) den Begriff „Daten als das Öl des 21. Jahrhunderts“, um zu verdeutlichen, dass Daten zwar in großem Umfang vorliegen, jedoch erst durch gezielte Aufbereitung und Analyse, in der Analogie eine Art „Raffinierung“, ihren Wert entfalten. Aufgrund der schier unendlichen Mengen an zugänglichen Daten, erfolgt diese Aufbereitung in der Regel mittels algorithmischer Verfahren, die als mathematisch beschreibbare, endliche Handlungsanweisungen zur Lösung spezifischer Probleme die Grundlage datenverarbeitender Technologien bilden (vgl. Gapski 2020, S. 160). Die Kombination aus gesteigerter Rechenkapazität, leistungsfähiger Hardware, einer tiefen Alltagsdurchdringung digitaler Technologien sowie bislang beispielloser Datenmengen erzeugte einen hohen Bedarf an neuartigen Verfahren der Datenverarbeitung (vgl. WBGU 2019, S. 77). Um in den stetig wachsenden Datenmengen Muster zu erkennen und nutzbar zu machen, wurden bereits seit den 1990er-Jahren statistische Verfahren entwickelt, die kontinuierlich weiter ausgebaut wurden und unter dem Oberbegriff des Maschinellen Lernens (Machine Learning) zusammengefasst werden können. Unter maschinellem Lernen versteht man Methoden, bei denen Modelle ihre Parameter anhand von strukturierten Daten iterativ anpassen, sodass sich statistische Vorhersagen oder Klassifikationen verbessern (vgl. Linnemann et al. 2023, S. 199). Grundlage dieser Verfahren ist, dass Maschinen oder Softwareprogramme durch „Erfahrungen“ in Form von Daten „trainiert“ werden (vgl. Kreidenweis 2024, S. 14). Dabei lassen sich unterschiedliche Lernparadigmen unterscheiden: Beim überwachten Lernen wird ein Modell mit bereits gelabelten Daten trainiert, beim unüberwachten Lernen sucht der Algorithmus selbstständig Strukturen oder Cluster in den Daten, und beim

bestärkenden Lernen erfolgt das Training durch Rückmeldungen in Form von “Belohnungen” oder “Strafen” (ebd. S. 18f.). Mit dem Aufkommen des Deep Learning, einer Unterkategorie des maschinellen Lernens, wurden diese Ansätze erheblich erweitert. Deep-Learning-Modelle basieren ebenfalls auf statistischen Prinzipien, nutzen jedoch mehrschichtige künstliche neuronale Netze, die es ermöglichen, hochkomplexe Muster in großen und unstrukturierten Datensätzen (z. B. Sprache, Bilder, Videos) zu erkennen (ebd. S. 15).
Mit der Veröffentlichung des Papers „Attention Is All You Need“ (Vaswani et al. 2017) durch ein Team von Google-Forschenden setzte ein entscheidender Durchbruch für die moderne Entwicklung generativer KI ein. Es markierte den Beginn einer neuen Generation algorithmischer Modelle, die auf der sogenannten Transformer-Architektur basieren. Diese Architektur nutzt einen Mechanismus, der es erlaubt, jedes Element einer Sprachsequenz in Relation zu allen anderen zu setzen und so das Ausgangselement in seinem statistischen Kontext zu „verstehen“, um so die Anfrage in natürlicher Sprache zu verarbeiten. Entscheidend war dabei weniger ein einzelnes technisches Detail als vielmehr die Tatsache, dass Sprachverarbeitung (und in der Folge auch Bild- und Videobearbeitung) in bislang unerreichter Genauigkeit und Geschwindigkeit möglich wurde.
Abschließend ist zu betonen, dass die anderen beschriebenen Architekturen der Künstlichen Intelligenz keineswegs an Relevanz verloren haben, sondern als eigenständige Entwicklungslinien weiterhin erfolgreich fortgeführt werden. Gleichwohl erfahren diese etablierten Ansätze durch die technologische Dynamik generativer Systeme eine signifikante Erweiterung bzw. Neubewertung. Ein Großteil der im Folgenden behandelten Projekte lässt sich in diesen konventionellen Strängen verorten und grenzt sich damit von den aktuellen Paradigmen der Generativen KI ab.

Im Anschluss wird aufgezeigt, wie die zuvor beschriebenen Vorstufen der KI in Publikationen der Sozialen Arbeit thematisiert werden und wie sich der Begriff mit dem Auftreten der großen Sprachmodelle verändert hat. Es ist jedoch hervorzuheben, dass die in den vergangenen drei Jahren stark beschleunigte und in Teilen exponentielle Entwicklung im Bereich der Künstlichen Intelligenz zu einer signifikanten Zunahme wissenschaftlicher Veröffentlichungen geführt hat. Diese Veröffentlichungen werden jedoch mitunter bereits durch die Einführung nachfolgender Modelle obsolet, sei es in Bezug auf den Funktionsumfang, die aktuelle gesellschaftliche Nutzung oder die allgemeine Verfügbarkeit. Die in dieser Arbeit dargestellten Projekte entsprechen dem Stand der wissenschaftlichen Publikationen bis zum Sommer 2025. Es ist jedoch davon auszugehen, dass eine erhebliche Zahl weiterer, bislang nicht dokumentierter Erprobungsprojekte in der Praxis existiert. Der daraus resultierende Forschungsbedarf kann mit der Dynamik der technologischen

Entwicklungen derzeit nicht Schritt halten, sodass die wissenschaftliche Aufarbeitung strukturell hinter der ubiquitären Verbreitung zurückbleibt. Gleichwohl ist eine deutliche Zunahme der Publikationsaktivität zu verzeichnen, die insbesondere in den vergangenen beiden Jahren in mehreren Sammelbänden ihren Ausdruck gefunden hat.

3 Anwendungsbereiche und Potentiale von KI in der Sozialen Arbeit

In den folgenden Unterkapiteln werden konkrete Anwendungsfälle von Künstlicher Intelligenz in Handlungsfeldern der Freien Wohlfahrtspflege vorgestellt. Jedes Unterkapitel beginnt mit einer allgemeinen Darstellung der jeweiligen technologischen Grundlagen und allgemeiner Nutzungsformen. Anschließend werden literaturbasiert spezifische Anwendungsfälle in Sozialer Arbeit und Wohlfahrtspflege präsentiert und in Bezug auf ihre Potenziale und gesellschaftlichen Implikationen diskutiert.

3.1 Dialogsysteme

Chatbots im Sinne text- und sprachbasierter Dialogsysteme, die automatisiert Nutzeranfragen beantworten können, sind kein neues Phänomen. Vielmehr werden sie besonders als Werkzeug des Customer-Relationship-Management (CRM) bereits seit den frühen 2000er Jahren als Service-Chats kommerziell genutzt und sind schließlich vor gut 10 Jahren, basierend auf Fortschritten im Bereich des Natural Language Processing (NLP), in Alltagstechnologien wie Amazons Alexa (2014) und Google Assistant (2016) gemündet. In sozialen Handlungsfeldern werden Chatbots wie etwa die psychotherapeutische App „Deprexis“ (BfArM 2021), die auch im Verzeichnis für digitale Gesundheitsanwendungen geführt wird, oder das Pilotprojekt „safety-for-refugees“, in dem ein mehrsprachiger Chatbot zur Unterstützung von Schutzsuchenden in der Schweiz entwickelt wurde (vgl. Bühler und Reiser 2023, S. 16f.), bereits angewendet. Diese Dialogsysteme nutzen jedoch keine künstliche Intelligenz, sondern arbeiten auf Basis regelbasierter Software, indem vordefinierte Anfragen durch festgelegte Antworten beantwortet werden. Hieraus folgt ein wesentliches Problem: Selbst technisch ausgereifte Chatbots konnten logische Dialogbrüche nicht vollständig vermeiden und die Kohärenz einer Interaktion dauerhaft aufrechterhalten (vgl. Lotze 2018, S. 42). Ihre Antwortfähigkeit ist durch die vordefinierten Auswahlbäume beschränkt. Werden Anfragen gestellt, die von den geskripteten Konversationen abweichen, etwa durch Rechtschreibfehler im Text, muss der Chatbot die Beantwortung verweigern, oder auf eine allgemeine Information bzw. einen menschlichen Kontakt verweisen, was zu signifikanten Abbruchraten auf Seiten der Hilfesuchenden führt. So zeigten die Untersuchungen von Bühler & Reiser, dass lediglich 10% der Nutzer:innen das angebotene

Hilfsangebot in Anspruch nahmen, nachdem der Chatbot keine adäquate Antwort liefern konnte (vgl. Bühler et al. 2023, S. 9).
Unter anderen um solche Kommunikationsbrüche zu vermeiden, wird LLM-gestützten Chatbots ein großes Potential für den fachlichen Einsatz (in der Beratung) zugerechnet (vgl. Krings 2024, S. 26). Lehmann bestätigt diese Einschätzung und beschreibt KI-gestützte Beratung als konsequente Weiterentwicklung digitaler Beratungsformate, die sich in den vergangenen Jahren bereits von der E-Mail- über Foren- und Chatberatung bis hin zu Messenger- und Videoberatung kontinuierlich ausdifferenziert haben (vgl. Lehmann 2025, S. 118ff.). Er skizziert dabei die grundlegenden Potenziale von Künstlicher Intelligenz in verschiedenen Phasen des Beratungsgeschehens: Vor der Beratung – etwa im Rahmen der Beratungsausbildung –, während der Beratung als „Copilot“ zur Unterstützung von Fachkräften sowie nach der Beratung als Reflexionsinstrument. Während diese Ansätze KI primär als Werkzeug für Fachkräfte verstehen, wird zugleich ihr Potenzial betont, Ratsuchende unmittelbar zu unterstützen – etwa durch die Förderung der Kommunikation, die niederschwellige Bereitstellung von Informationen oder in Form einer eigenständigen Beratungsinstanz (ebd. S. 120ff.). Fachlich zeichnen sich erste integrierte Ansätze ab, in denen traditionelle, „menschliche“ Beratungsprozesse durch den Einsatz künstlicher Intelligenz ergänzt werden. Im 2024 gestarteten Modellprojekt „KIESA – Inklusiv beraten“, einer Kooperation mehrerer Wohlfahrtsakteure mit der Technischen Hochschule Nürnberg, wird eine künstliche Intelligenz entwickelt, die Beratungsprozesse unterstützt, beispielsweise durch automatische Übersetzungen in einfache Sprache oder die Zusammenfassung des bisherigen Gesprächsverlaufs (vgl. Kieslinger 2024, 338f.). Erste vollständig KI-gestützte Beratungsangebote, wie der Bürgergeld-Chatbot[5], sind 2025 in den Praxiseinsatz übergegangen und bieten eine fundierte, KI-basierte Erstberatung (sogenannte 0-Level-Beratung), die zuvor von Fachberater:innen durchgeführt worden wäre. Darüber hinaus ist der Einsatz von Chatbots zur emotional-affektiven Unterstützung ein hochaktuelles gesellschaftliches Thema. Nicht ausreichende Beratungskapazitäten in analogen Settings, uhrzeit-unabhängige Verfügbarkeit und die ubiquitäre und niederschwellige Zugänglichkeit von Chatbots, in Verbindung mit einer hohen sprachlich-empathischen und inhaltlichen Qualität der von ihnen generierten Antworten führen zu einer vermehrten informellen Nutzung zu Beratungszwecken (vgl. Linnemann et al. 2024a, S. 13f.). Aktuelle Modelle sind in der Lage, sprachliche Reaktionen oder Verhaltensweisen in beachtlicher Qualität zu interpretieren und auch selbst zu generieren, die den Anschein empathischer Sorge erwecken und den Eindruck einer intentionalen Unterstützung vermitteln (vgl.

[5] Ein Angebot der Caritas NRW unter https://ki-chat.das-steht-dir-zu.de/

Macsenaere 2024, S. 13). Empirische Untersuchungen zeigen, dass generative KI-Modelle hinsichtlich Qualität und Empathie mitunter mit professionellen (klinischen) Fachkräften vergleichbar sind und in einzelnen Studien sogar vorteilhafter bewertet wurden (in der Psychotherapie: Hatch et al. 2025; in einem Onlineforum zu medizinischen Fragen: Ayers et al. 2023; in der Beziehungsberatung: Vowels et al. 2024). Dabei sind Menschen in der Kommunikation mit Chatbots offenbar bereit, auch sensible Informationen zu teilen. Bühler und Reiser beschreiben beispielsweise, dass Nutzer:innen eines Chatbots für Migrant:innen von Streitigkeiten und Unrechtserfahrungen bis zu physischer Gewalt im Wohnumfeld berichten (vgl. Bühler und Reiser 2023, S. 17). Umfassend beschreiben Linnemann, Löhe und Rottkemper (2024a) diese Selbstoffenbarungseffekte im Kontext von „quasisozialen Beziehungen“[6]. Ein entscheidender Faktor hierfür ist die reduzierte Bewertungsangst: Nutzende offenbaren sich einer KI leichter, wenn sie annehmen, dass diese kein soziales Urteil fällt (vgl. ebd. S. 11). In diesem Kontext sind zuletzt besonders spezialisierte „Companion-Apps“ (zu Deutsch: Begleitung) wie „replika.ai“ (vgl. Linnemann 2025, S. 41) oder „friend.com“ Gegenstand der wissenschaftlichen und öffentlichen Diskussion (zu Chancen und Risiken weiterführend Hollanek und Sobey 2025). Mit gezielter anthromorpher Sprache werben Anbieter wie „replika.ai“ („Der AI-Begleiter, der sich um dich sorgt. Immer da, um zuzuhören und zu reden. Immer an deiner Seite.“) oder „friend.com“ („Dein neuer Mitbewohner wartet auf dich“) um Nutzer:innen. Auch die Sprachmodelle der großen Anbieter, wie ChatGPT usw., werden zu ähnlichen Beratungen herangezogen. Dieses Phänomen wurde jedoch bisher erst wenig wissenschaftlich beschrieben, wobei die bisher umfangreichsten Studien von den Anbietern selbst stammen. Mit jeweils gut 4 Millionen untersuchten Konversationen beschäftigten sich OpenAI (Phang et al. 2025) und Anthropic (McCain et al. 2025) mit der emotional-affektiven Nutzung ihrer Modelle und fanden dabei jeweils etwa 2-3% entsprechende Konversationen. Es deutet sich an, dass die technologischen Entwicklungen und die tatsächliche Nutzungspraxis der potentiellen Ratsuchenden den fachlichen Diskurs überholen, der besonders aufgrund der gravierenden ethischen und datenschutzrechtlichen Bedenken (vgl. Kap. 3) berechtigterweise kritisch geführt wird. Bislang gelten für den Einsatz von KI im Bereich professioneller Beratung und Gesundheitsversorgung allerdings strenge Auflagen und in bestimmten Fällen sogar Ausschlüsse (vgl. „Hochrisikotechnologien“ im AI-Act). Dass der Markt trotzdem

[6] „Eine quasisoziale Beziehung beschreibt die Beziehung zwischen einem Menschen und einem künstlichen Agenten, bei der die Merkmale zwischenmenschlicher Beziehungsbildung (1) Soziale Präsenz, (2) Vertrauen, (3) emotionale Bindung und (4) gegenseitige Beeinflussung gleichzeitig vorhanden sind. Quasisoziale Beziehungen weisen soziale Elemente auf, sind aber dennoch von echten zwischenmenschlichen Beziehungen zu unterscheiden.“ (Rottkemper et al. 2023a, S.11).

scheinbar unaufhaltsam auf die Nutzung von großen Sprachmodellen in digitalen Chatbots zusteuert, zeigt das Beispiel des vielzitierten amerikanischen Branchenpioniers WoeBot Health (vgl. Bertsche und Como-Zipfel 2023, S. 160f.; vgl. Linnemann et al. 2024a, S. 2): Das Unternehmen nahm im Juni 2025 seine bis dahin aufwändig zugelassene, regelbasierte Anwendung vom amerikanischen Markt, um einen Relaunch auf Grundlage generativer KI vorzubereiten. Die vorgestellten Befunde werfen nicht nur Fragen nach der Rolle von KI in zwischenmenschlicher Kommunikation auf, sondern berühren auch grundlegende Herausforderungen für professionelle Beratungsinstitutionen der Freien Wohlfahrtspflege. Diese stehen vor der Herausforderung, Konzepte und technische Lösungen zu entwickeln, wie sich generative KI in qualitätsgesicherte Beratungs- und Unterstützungsprozesse integrieren lässt. Ohne eine aktive Gestaltung droht die Gefahr, dass institutionalisierte Angebote in Konkurrenz zu frei verfügbaren KI-Anwendungen geraten und an Bedeutung verlieren.

3.2 Entscheidungsunterstützung

Neben der Unterstützung von Kommunikation und Dokumentation wird KI auch für die Entscheidungsunterstützung in Wohlfahrtskontexten eingesetzt. Erste Expertensysteme wurden bereits in den 1980er-Jahren insbesondere im medizinischen Bereich zur Entscheidungsunterstützung entwickelt und haben seitdem eine stetige Weiterentwicklung erfahren (vgl. Sutton et al. 2020). Heutige Systeme gelten insbesondere bei Aufgaben der visuellen Klassifizierung, etwa in der Tumorerkennung, generell als leistungsfähiger gegenüber der menschlichen Einschätzung und sind ein etabliertes medizinisches Instrument (exemplarisch Fujita 2020). In der Sozialen Arbeit werden vergleichbare Ansätze als Algorithmische Entscheidungssysteme (ADM; auch decision support systems (DSS)) diskutiert. Seit den 2010er-Jahren sind insbesondere im angloamerikanischen Raum Projekte umgesetzt worden, die sich vorrangig auf den Kinderschutz konzentrieren (Neuseeland: Gillingham 2017; Großbritannien: Broadhurst et al. 2015; USA: La Valle et al. 2016). Im Rahmen von Verfahren der Kinder- und Jugendhilfe dienen diese Systeme der präventiven Risikoeinschätzung mit dem Ziel, eine fundierte Prognose über potenzielle Gefährdungslagen zu ermöglichen und dadurch zukünftigen Schaden zu verhindern (Predictive Risk Modelling). Die frühen prädiktiven Systeme wiesen jedoch mehrere zentrale Einschränkungen auf. Die Datengrundlage war oftmals unzureichend, da sie entweder nicht repräsentativ war, nur eine begrenzte Stichprobengröße umfasste oder wichtige Merkmale nicht erfasste. Hinzu kam eine mangelnde Erklärbarkeit der Ergebnisse, da viele der eingesetzten Modelle aufgrund

ihrer technischen Funktionsweise keine Einblicke in die Entscheidungsfindung ermöglichen. Evaluationen zeigten, dass Verzerrungen in den Daten mitunter zu sexistischen und rassistischen Entscheidungen der Systeme führten. Zusammenfassend besteht die grundlegende Schwierigkeit darin, menschliches Verhalten ausschließlich auf Basis quantitativer Daten vorherzusagen, da dieses von vielfältigen qualitativen, situativen und schwer messbaren Faktoren beeinflusst wird. Die Geschichte der Entwicklung entsprechender Systeme ist daher eng begleitet von kritischer Rezeption und eingestellten Projekten (vgl. Gillingham 2016, S. 34; vgl. Chouldechova et al. 2018, S. 3f.; vgl. Bühler und Reiser 2023, S. 16; vgl. Lehmann und Burghardt 2024, S. 68). Ein Beispiel für Ansätze, die diese Defizite adressieren, bietet die Studie von Amrit et al. (2017). Dort wurde ein prädiktives System für den Kinderschutz entwickelt, das strukturierte Sozialdaten mit unstrukturierten Textinformationen aus Fallnotizen kombinierte. Die Ergebnisse zeigten, dass durch die Verbindung beider Datentypen eine deutlich höhere Prognosegenauigkeit erreicht werden konnte. Am Beispiel der Stadt Amsterdam wurde berichtet, dass die kombinierte Nutzung strukturierter und unstrukturierter Daten eine deutlich höhere Vorhersagegenauigkeit ermöglichte, wobei insbesondere die schriftlichen Fallnotizen aus verpflichtenden Kinder- und Jugenduntersuchungen einen erheblichen Mehrwert gegenüber rein strukturierten Datensätzen lieferten (vgl. Amrit et al. 2017, S. 23f.). Mit dem Aufkommen großer multimodaler KI-Systeme erfahren solche integrativen Ansätze eine neue Relevanz. Aktuelle Forschungsvorhaben binden zunehmend unstrukturierte schriftliche Falldaten in die Modellierung ein und nutzen die Fähigkeit moderner Modelle, verschiedene Datentypen unabhängig von ihrer Strukturiertheit zu verarbeiten (etwa Plafky et al. 2025; Stoll et al. 2025). Dabei wird die moderne generative Komponente eingesetzt, etwa um Reflexionsfragen zu Falldaten zu formulieren oder Hypothesen zum Fallgeschehen zu entwickeln (vgl. Lehmann und Burghardt 2024, S. 69). Die gezielte Integration domänenspezifischen Fachwissens reagiert zugleich auf Kritik, wonach bisherigen Systemen eine hinreichende fachliche Fundierung fehlte. Auffällig ist, dass die Systeme damit nicht mehr ausschließlich übergeordnete Controlling-Funktionen übernehmen, sondern zunehmend in die pädagogische Fallgestaltung selbst eingreifen. Mit dieser neuartigen intensiveren Eingriffstiefe verbindet sich die Forderung nach einer diskursiven Auseinandersetzung sowie nach der Sicherstellung einer letztverantwortlichen menschlichen Entscheidungsinstanz (vgl. Linnemann et al. 2024b, S. 92f.). Diese entspricht zugleich einer Mindestanforderung der europäischen KI-Verordnung (AI Act, Artikel 4). Vor diesem Hintergrund ist den Autoren zum Zeitpunkt der Studienerstellung kein System bekannt, das in Deutschland bereits produktiv im Einsatz wäre.

3.3 Unterstützung im Alltag

Mit Assistenzsystemen und Ambient Assisted Living (AAL) sind technologische Systeme gemeint, die Unterstützung in der selbstständigen Lebensführung bieten, zum Beispiel für Senior:innen oder Menschen mit einer Behinderung (vgl. Krings 2024, S. 26). Bereits in den frühen 2010er Jahren wurden solche Systeme KI-gestützt entwickelt, um ältere und pflegebedürftige Menschen in ihrem häuslichen Umfeld zu unterstützen. Dazu gehören etwa intelligente Sensoren zur Überwachung von Aktivitäten oder erste Roboter. Dabei ist die Abgrenzung zwischen regelbasierter Programmierung und künstlicher Intelligenz nicht immer trivial. Ein exemplarisches Anwendungsfeld hierfür sind Pflegeroboter: Während Sprach- und Bilderkennung oder die Anpassung an Nutzerpräferenzen auf maschinellem Lernen beruhen und als KI-Komponenten klassifiziert werden können, folgen grundlegende Bewegungsabläufe oder Routineaufgaben häufig festgelegten Regeln, und können damit nicht als autonome oder „intelligente“ Pflegeunterstützung gezählt werden (vgl. Graf 2020, S. 610). Aktuelle Entwicklungen im Bereich des maschinellen Lernens und der kontextsensitiven KI erweitern das Anwendungsspektrum dieser Systeme erheblich. Moderne AAL-Lösungen sind in der Lage, individuelle Routinen zu erlernen und sich adaptiv an die Bedürfnisse der Nutzer:innen anzupassen. Beispielsweise können intelligente Sprachassistenzsysteme Erinnerungsfunktionen übernehmen, etwa bei der Einnahme von Medikamenten, oder bei kognitiven Einschränkungen unterstützende Hinweise geben. Für den Bereich der Teilhabe von Menschen mit Behinderung am Arbeitsleben hat zuletzt hat das Verbundprojekt KI.Assist eine systematische Erhebung von KI-gestützten Assistenzsystemen durchgeführt. Die Studie zeigt, dass sich die meisten identifizierten Projekte noch auf dem Umsetzungsstand wissenschaftlicher Forschungsarbeiten befinden (vgl. Lippa 2024, S. 71). Insgesamt untersuchten die Autor:innen knapp 160 Technologien, die Unterstützung in den sieben Hauptkategorien Wahrnehmung und Kommunikation (29%), Interaktion (13%), Navigation und Orientierung (7%), konkrete Arbeit (7%), Lernen (12%), physische Aspekte (22%) sowie psychische Aspekte (6%) bieten und eine breite Palette technischer Hardware als Ausgangspunkt für den Einsatz von KI nutzen (vgl. Blanc und Beudt 2022, S. 16). Die Bandbreite reicht von sensorbasierten Erkennungssystemen über tragbare Brillen, die Inhalte per Augmented Reality projizieren, bis zu Orthesen, die Bewegungsabläufe unterstützen (ebd. S.22). Dabei heben die Autor:innen hervor, dass KI-basierte Systeme in diesem Handlungsfeld vor allem Sinnes- und Körperbehinderungen adressieren, seltener jedoch psychische oder Lernbehinderungen (ebd. S.19).

Jovanovic et al. (2022) führten ein Scoping Review zum Einsatz von Künstlicher Intelligenz im Bereich des Ambient Assisted Living durch. Mit Fokus auf die Versorgung von alten Menschen fanden Sie in über 100 Artikeln überwiegend Lösungen zur Unterstützung von Aktivitätsassistenz (Mobilitätsassistenz), Aktivitätserkennung (Sensorsysteme) und Kommunikation. Kennzeichnend für den Einsatz im häuslichen Kontext ist das stetige Monitoring von Alltagsdaten (Bewegungssensoren, Gesundheitsdaten) mit dem Ziel, Notsituationen, wie Sturzgeschehen frühzeitig zu erkennen. Hier werden neben verbauten Sensoren und mobilen Geräten, wie Smartphones oder Smartwatches auch assistive Roboter benannt (vgl. Jovanovic et al. 2022, S. 17). Anthropomorphe bzw. zoomorphe Roboter wie NAO oder die Robbe PARO wurden gezielt für die Interaktion mit Menschen entwickelt und sind seit rund zwei Jahrzehnten auf dem Markt. Sie nutzen Techniken der Künstlichen Intelligenz im Bereich der Sprach- und Gesichtserkennung, um eine möglichst natürliche Kommunikation zu ermöglichen und speichern das Interaktionsverhalten. Kirchner et al. (2019) heben darüber hinaus Exoskelette als besondere Form der Robotik hervor. Erst der Einsatz von Künstlicher Intelligenz ermöglicht die präzise Steuerung komplexer mechanischer Exoskelette, wie sie zur Nachbildung der Morphologie der menschlichen Extremitäten notwendig sind (vgl. Kirchner et al. 2019, S. 419). Modernste, multimodale KI-Modelle sind im Anwendungsbereich der Alltagsunterstützung für Menschen mit einer Behinderung besonders hervorzuheben. In einer eindrucksvollen Demonstration zeigte OpenAI bei der Veröffentlichung des Modells GPT-4o (das „o“ steht für omni, lateinisch für „alles“, um auf die Multimodalität des Modells zu verweisen, das neben Text auch auditive und visuelle Daten verarbeiten kann), wie eine KI mithilfe einer Smartphone-Kamera die Umgebung von sehbehinderten Personen präzise beschreiben und dadurch bei der Navigation unterstützen kann (OpenAI, 2024). Eine vergleichbare Unterstützung im Alltag wurde bislang durch zivilgesellschaftliche Initiativen wie die NGO „Be My Eyes“ ermöglicht, bei der Freiwillige per Videoanruf visuelle Informationen in Sprache übertragen (Avila et al. 2016). Technische Hilfsmittel wie die OrCam MyEye geben geschriebenen Text als Sprache aus, wobei sie als zugelassene medizinische Hilfsmittel ein aufwändiges Prüfverfahren durchlaufen müssen, um ihre Refinanzierung im Gesundheitssystem zu gewährleisten. Deutlich wird daran, wie gravierend die jüngsten Durchbrüche im Bereich der KI sind: Funktionen, die bislang als Alleinstellungsmerkmale spezialisierter Produkte galten, können durch allgemein verfügbare multimodale Modelle zunehmend ersetzt werden.

3.4 Unterstützung der alltäglichen Dokumentation und Information

Kaum eine Tätigkeit nimmt in der sozial- bzw. wohlfahrtspraktischen Arbeit einen so hohen zeitlichen Anteil ein wie die (pädagogische) Dokumentation. Die Auseinandersetzung mit dem Lesen und Schreiben fallrelevanter Informationen erfüllt dabei vielfältige Funktionen, die sowohl der internen als auch der externen Absicherung professionellen Handelns dienen (vgl. Reichmann 2016, S. 38ff.). Plafky (2025, S. 191) bezeichnet die Dokumentation sogar als „unerlässlich“, um Soziale Arbeit ethisch, rechtlich und zielgerichtet zu gestalten. Trotz der hohen Relevanz existieren nur wenige belastbare empirische Untersuchungen zum tatsächlichen Zeitaufwand der Dokumentationstätigkeiten. Dieser variiert je nach Handlungsfeld und definitorischen Einschlusskriterien erheblich. Ein Anteil von rund 30 % der Arbeitszeit, der für das Lesen und Verfassen von Dokumentationen aufgewendet wird, gilt jedoch als verbreitet akzeptierter Richtwert (vgl. Halfar 2019, S. 16; vgl. Schmidberger 2024, S. 147). In Sitzungsprotokollen, Hilfeplanberichten, Wunddokumentationen und all den weiteren Dokumentationsformen wird professionelles und domänenspezifisches Handlungswissen verbalisiert und verstetigt. Problematisch bleibt jedoch, dass dieses fixierte Wissen bislang kaum systematisch in die fachliche Arbeit zurückfließt. Holz et al. sprechen daher von einem „Wissensschatz“, den es zu heben gilt (ebd. S. 174). Sowohl in der Erstellung als auch in der Nutzung von Dokumentationseinträgen scheinen also Optimierungspotentiale zu bestehen. Werfen wir zunächst einen Blick auf die Eingabe von Dokumentation, dann finden wir, wie auch in den anderen Einsatzfeldern im Gesundheitswesen bereits erste KI-gestützte Produkte, die am Markt etabliert sind. Anbieter – beispielsweise „Voize“ - bieten eine Teilautomatisierung der Dokumentationsaufgaben durch Spracheingaben, die von der Anwendung über Schnittstellen unmittelbar in einheitliche, maschinenlesbare Datenpunkte umgewandelt und in die digitale Fallakte der jeweiligen Fachsoftware integriert werden (vgl. Schmidberger 2024, S. 150ff.). Weitere Beispiele für entsprechende Lösungen finden sich bei Plafky et al. (2025, S. 191). Im Bereich der Sozialen Dienste gestaltet sich die vereinfachte Eingabe von Dokumentationen aufgrund des hohen Anteils an Freitexten im Unterschied zu den stärker quantifizierbaren Daten etwa im Pflegebereich, wie zum Beispiel Blutdruckwerten, deutlich komplexer. Die meisten Anbieter von Fachsoftware stellen jedoch bereits Diktierfunktionen bereit oder planen deren zeitnahe Implementierung (vgl. Kreidenweis und Diepold 2024, S. 36). Es ist davon auszugehen, dass in naher Zukunft weitere KI-Funktionalitäten ihren Weg in die Fachsoftware der Sozialen Arbeit finden werden. Über 70 Prozent der

befragten Softwarehersteller bewerten das Potenzial von KI für ihr Produkt als hoch bis sehr hoch (ebd. S. 35), und erste Funktionen wurden bereits eingeführt, beispielsweise die KI-gestützte Dienstplanung in Medifox DAN oder das Modul VIVA in Connext Vivendi, welches einen Chat mit der digitalen Patientenakte ermöglicht. Die neuen Möglichkeiten generativer KI erfordern eine vertiefte Debatte über den Sinn und Zweck von Dokumentation. „Darf" die KI aus den zuvor lose diktierten Gedanken einen wohlformulierten und fachsprachlichen Absatz für meinen nächsten Bericht generieren? Oder ist gerade die reflexive Auseinandersetzung mit den fachlichen Wahrnehmungen und Einschätzungen ein entscheidender Schlüsselmoment professionellen Handelns? Darüber hinaus werden derzeit auch Systeme erprobt, die unmittelbar in die fachliche Arbeit und Fallreflexion eingebunden sind. Das Assistenzsystem KAIMo beispielsweise extrahiert aus der Freitextdokumentation jene Informationen, die für eine Gefährdungseinschätzung relevant sind, und unterstützt Fachkräfte anschließend bei der Hypothesenbildung zum Fall (vgl. Lehmann und Burghardt 2025, S. 69).

Nicht zuletzt dienen KI-gestützte Wissensdatenbanken der effizienten und gezielten Auffindbarkeit fallrelevanter Informationen. Besonders vielversprechend sind hier Ansätze der Retrieval-Augmented Generation (RAG), bei denen ein etabliertes Foundation Model Zugriff auf unternehmensinterne Daten erhält und diese mit besonderer Gewichtung in seine Antwortgenerierung einbezieht. Auf diese Weise können Informationen aus verschiedenen Datenquellen effektiv zusammengeführt und kontextualisiert werden. Entsprechende Anwendungen sind auch außerhalb geschlossener Systeme denkbar. Ein Beispiel dafür ist die Plattform Jugendhilfe.ai[7], die auf Grundlage einschlägiger Fachliteratur Fragen zu gesetzlichen Rahmenbedingungen, Finanzierungsmöglichkeiten und komplexen Rechtsfragen mittels generativer KI beantwortet. In jüngster Zeit wird vermehrt die Verwendung sogenannter „KI-Agenten" diskutiert, die viele der oben genannten Funktionen theoretisch erfüllen könnten. KI-Agenten meinen Instanzen eines Foundational Model, denen ein initialer System- oder Basisprompt vor jeder Interaktion zugewiesen wird. Durch diese Vorgehensweise lässt sich der Agent auf unterschiedliche Aufgaben vordefinieren (z. B. die automatische Formulierung einer formellen E-Mail an externe Akteure). Im Kontext der Wohlfahrt zeichnen sich hier erste innovative Geschäftsmodelle ab: So hat der AWO Landesverband Thüringen e.V. mit seiner Tochtergesellschaft K-KISS GmbH („Kompetenzzentrum KI für Soziale Services") eine entsprechende KI-Plattform[8] etabliert und bietet diese Lösung als Dienstleister auch anderen sozialen Trägern an.

[7] https://jugendhilfe.ai/

[8] https://awo-ki-spielplatz.de/

4 Herausforderungen und Risiken

Trotz der vielfältigen Potenziale und neuen Anwendungsmöglichkeiten von Künstlicher Intelligenz bestehen zugleich zahlreiche Herausforderungen und Risiken. Eine vertiefte Auseinandersetzung mit diesen Aspekten kann an dieser Stelle nicht erfolgen; sie sollen jedoch zumindest benannt und durch weiterführende Quellen für interessierte Leser:innen zugänglich gemacht werden.

4.1 Verzerrungen und fehlerhafte KI-Berechnungen

Datenbias

Ein Bias der KI kann zu einer Überrepräsentation bestimmter Bevölkerungsgruppen bzw. Merkmale führen und mangelnde Transparenz im Entscheidungsprozess kann Benachteiligungen und Ungleichheiten verstärken (Chouldechova et al. 2018). Dies kann auch durch die Tendenz von Menschen, „in der Zusammenarbeit mit Maschinen entweder nachlässig, übermäßig abhängig oder unangemessen zurückhaltend zu werden“ (Zerilli et al. 2019, S. 555), begünstigt werden.

Weiterführend: Plafky (2025): Entscheidungsunterstützung in der Sozialen Arbeit durch prädiktive Analytik

Halluzinationen

Zwar können Large Language Models grundlegende Informationsanfragen beantworten, doch dabei besteht das zentrale Problem der Halluzinationen. Da Sprachmodelle primär die Wahrscheinlichkeiten von Wortsequenzen innerhalb eines semantischen Ähnlichkeitsraums berechnen, sind sie nicht in der Lage, zwischen Wahrheit und Falschheit von Aussagen zu unterscheiden. Sie verfügen zudem nicht über ein metakognitives Bewusstsein dafür, ob sie zu einem bestimmten Thema valide Informationen besitzen. Infolgedessen kann es zu sogenannten „Halluzinationen“ kommen, bei denen das Modell Informationen aus seinem neuronalen Netzwerk generiert, die sprachlich plausibel erscheinen, jedoch nicht mit der Realität übereinstimmen (vgl. Kreidenweis 2024, S. 17). Im Kontext von rechtlicher oder psychosozialer Beratung ist diese faktische Korrektheit jedoch essenziell, ja in manchen Fällen sogar lebenswichtig, sodass die Adaption von generativer KI hier nicht unmittelbar erfolgt ist (vgl. Plafky 2025, S. 422f.). In der technischen Entwicklung der Large Language Models konnten hier jedoch zuletzt Fortschritte gemacht werden durch 1) die Anbindung der Modelle an das Internet und

2) „Retrieval-augmented Generation“ (RAG), bei der die Wissensbasis für generative Inhalte durch eigene Dokumente ergänzt wird (vgl. Löhe 2024, S. 105f.).

Weiterführend: Gapski, Harald (2020): Digitale Transformation: Datafizierung und Algorithmisierung von Lebens- und Arbeitswelten.

4.2 Irrtümer über Fähigkeiten und Eigenschaften von KI

Automation Bias

Darüber hinaus sind Organisationen der Sozialen Arbeit dem sogenannten Automation Bias ausgesetzt. Dieser beschreibt die Tendenz von Menschen, automatisierten Systemen übermäßiges Vertrauen zu schenken, selbst wenn die eigenen fachlichen Einschätzungen oder empirischen Hinweise widersprechen. Studien zeigen, dass insbesondere in komplexen Entscheidungssituationen Fachkräfte dazu neigen, Vorschläge von Software oder KI als verbindlich anzusehen, auch wenn diese fehlerhaft oder unvollständig sind (Parasuraman und Manzey 2010). Im Kontext der Sozialen Arbeit kann dies bedeuten, dass kritische fachliche Bewertungen, die individuelle Lebenslagen und soziale Dynamiken berücksichtigen, zugunsten standardisierter, algorithmisch erzeugter Empfehlungen vernachlässigt werden. Der Automation Bias kann sowohl zu einer Überakzeptanz falscher maschineller Vorschläge als auch zu einer Abwertung der eigenen Expertise führen. Um dem entgegenzuwirken, sind gezielte Schulungen, die Reflexion algorithmischer Entscheidungsprozesse und die Einführung von Mechanismen zur Überprüfung von Systemvorschlägen notwendig.

Weiterführend: Parasuraman und Manzey (2010): Complacency and bias in human use of automation: an attentional integration

Erkennen von Werbeversprechen

Kritisch weisen Kreidenweis und Diepold (2024) darauf hin, dass zahlreiche „branchenspezifische Produkte mit dem Label „KI-gestützt“ versehen werden, obwohl es sich nicht immer um „echte“ KI im Sinne selbstlernender Systeme, sondern teils nur um Optimierungsalgorithmen handelt.“ (Kreidenweis und Diepold 2024, S. 11). Wenngleich dieser Aussage eine enge Definition von KI zugrunde liegt, lassen sich doch solche Beispiele für „KI-Washing“ finden. Beispielhaft kann hier die Dienstplanungs-KI eines Fachsoftware-Herstellers benannt werden. Tatsächlich

handelt es sich jedoch nicht um ein selbstlernendes Deep-Learning-System, sondern um eine parametrisierte Optimierungs- und Priorisierungs-Engine. Sämtliche Planvorschläge werden durch ein regelbasiertes System erzeugt, das vordefinierte Bedingungen und gewichtete Kriterien systematisch abgleicht - im Gegensatz zu adaptiven, neuronalen Netzen. Der Begriff „KI-gestützt“ dient in diesem Kontext vor allem der Vermarktung und verweist auf die Automatisierung durch Software-Intelligenz, nicht auf Methoden des Machine Learning. Organisationen der Sozialen Arbeit sollten daher zunächst klären, für welche Aufgaben der Einsatz von KI sinnvoll und notwendig ist und in welchen Fällen regelbasierte Systeme effizienter und zuverlässiger arbeiten, um auf dieser Grundlage fundierte Einschätzungen zur Sinnhaftigkeit angebotener Produkte treffen zu können.

Weiterführend: Gutwald et al. (2021): Soziale Konflikte und Digitalisierung - Chancen und Risiken digitaler Technologien bei der Einschätzung von Kindeswohlgefährdungen.

Zuschreibung menschlicher Eigenschaften

In der öffentlichen Auseinandersetzung wie auch in privaten Nutzungskontexten werden der KI Eigenschaften zugesprochen, die über ihre tatsächliche technische Funktionsweise hinausgehen. Dabei zeigt sich zum einen die Tendenz zur Anthropomorphisierung, also der Zuschreibung menschlicher Fähigkeiten, Intentionen und Emotionen an ein nicht-menschliches System. Während diese „Vermenschlichung“ kein neues Phänomen ist, sondern bereits bei früheren Technologien wie dem Computer zu beobachten war (Echterhoff et al. 2006), erhält sie im Kontext aktueller KI-Entwicklungen eine neue Relevanz und Intensität. Mitunter sind es auch bewusste Designentscheidungen der Entwickler:innen bzw. des Marketings, wie zum Beispiel die Gestaltung von Avataren, die gezielte Simulation von Emotionalität oder entsprechende Werbeslogans. Dabei impliziert eine anthropomorphe Wahrnehmung nicht ausschließlich negative Folgen: Die Vermenschlichung bis hin zur quasisozialen Beziehung bedeuten ein hohes Maß an Vertrauen und Akzeptanz gegenüber der Technologie.

Weiterführend: Linnemann et al. (2024b): Mensch-KI-Interaktion in der Kinder- und Jugendhilfe.

Zuschreibung magischer Fähigkeiten

Zum anderen lassen sich gerade im Hype um generative KI thaumaturgische Deutungsmuster[9] feststellen: KI wird nicht nur als technisches Werkzeug verstanden, sondern durch ein magisches Narrativ überhöht. Dieses Narrativ wird gezielt von Herstellern und Anbietern geschürt, wenn sie in ihrer Bildsprache etwa Zauberstab-Icons, Feenstaub-Effekte oder vergleichbare Symbole in der Gestaltung ihrer KI-Anwendung verwenden. Solche Gestaltungsentscheidungen verschleiern die tatsächliche Funktionsweise der Systeme und verstärken den Eindruck, es handle sich bei KI um eine Art übernatürliche Kraft, was sie einer technischen Steuerbarkeit zu entziehen scheint. Auch hier fördert die „Aufgabe" der eigenen Handlungsfähigkeit das Vertrauen in die Technologie, kann aber zu einer unkritischen Nutzung führen.

Weiterführend: Lupetti und Murray-Rust (2024): (Un)making AI Magic: A Design Taxonomy.

4.3 Verantwortung und Transparenz in KI-Systemen

Ethik

Die Soziale Arbeit kann sich auf die sieben allgemein anerkannten Kriterien für einen ethisch gerechten KI-Einsatz berufen, die eine von der EU-Kommission eingesetzte Expert:innengruppe erarbeitet hat (vgl. Independent High-Level Expert Group on Artificial Intelligence 2020, S.3).

Sie lauten:

1. menschliche Entscheidungsfreiheit und Kontrolle,
2. technische Robustheit und Sicherheit,
3. Datenschutz,
4. Transparenz,
5. Diversität, Nichtdiskriminierung sowie Fairness,
6. gesellschaftliches und ökologisches Wohlergehen und
7. Zurechenbarkeit

Für die Soziale Arbeit fungieren diese Leitlinien nicht als bloße technische Checkliste, sondern als essenzielles normatives Korrektiv, um das professionelle Mandat der Hilfeleistung auch im digitalen Raum zu wahren. Insbesondere vor dem Hintergrund vulnerabler Adressat:innengruppen erhalten Aspekte wie Nichtdiskriminierung und

[9] In der historischen Thaumaturgie wurden mechanische Artefakte im 16. Jahrhundert aufgrund mangelnden technischen Verständnisses für magisch gehalten.

Transparenz eine gesteigerte Relevanz, da intransparente algorithmische Entscheidungen (etwa bei der Ressourcenvergabe) existenzielle Ungleichheiten verstärken könnten. Somit dient dieser Rahmen dazu, den Einsatz von KI konsequent an den Zielgruppen der Wohlfahrt auszurichten und sicherzustellen, dass technologische Effizienz niemals die menschliche Urteilskraft und professionelle Verantwortung substituiert.

Weiterführend: Görder (2020): Wann ist KI ok?

Blackbox

Die maschinellen Lernverfahren, die die Grundlage künstlicher Intelligenz sind, funktionieren innerhalb einer menschlich nicht nachvollziehbaren Blackbox. Die künstlichen neuronalen Netze entwickeln im Laufe des Trainings eine Struktur aus statistischen Verknüpfungen, die es dem System ermöglichen, Muster und Zusammenhänge in großen Datenmengen zu erkennen und daraus Vorhersagen abzuleiten. Diese statistischen Verknüpfungen entstehen jedoch ohne explizit nachvollziehbare Regeln, wodurch die internen Entscheidungsprozesse der Modelle weitgehend undurchsichtig bleiben. Dieses Phänomen wird als Blackbox-Charakter der künstlichen neuronalen Netze bezeichnet. Es erschwert die Nachvollziehbarkeit und Erklärbarkeit der Ergebnisse und stellt somit eine zentrale Herausforderung für die verantwortungsvolle Nutzung und Weiterentwicklung Künstlicher Intelligenz dar.

Weiterführend: Gillingham (2016): Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the 'Black Box' of Machine Learning

Datenschutz

Eine weitere zentrale Herausforderung bildet der Datenschutz. Da in der Sozialen Arbeit regelmäßig hochsensible personenbezogene Daten verarbeitet werden, etwa zu finanziellen Verhältnissen oder dem Gesundheitsstatus, unterliegen diese Informationen einem besonders hohen Schutzbedarf. Die Nutzung von Large Language Models (LLM) birgt hier spezifische Risiken, da eingegebene Informationen auf externen Servern verarbeitet und potenziell für das Training der Modelle weiterverwendet werden können. Zudem bestehen Gefahren durch sogenannte „Privacy Attacks“, bei denen Angreifer durch gezielte Abfragen versuchen, sensible Informationen aus den Trainingsdaten der Modelle zu rekonstruieren. Datenschutzkonformität erfordert daher neben technischen Maßnahmen wie

Verschlüsselung zwingend auch eine entsprechende KI-Kompetenz der Fachkräfte, um ungewollte Datenabflüsse zu verhindern.

Weiterführend: Pelzl (2025): IT-Sicherheit und Datenschutz im Kontext von KI-Sprachmodellen

4.4 Kompetenzen & Qualifizierung

Für den professionellen Einsatz von KI und datenbasierten Systemen in der Wohlfahrtspflege müssen Fachkräfte über vier miteinander verzahnte Kompetenzdimensionen verfügen:

Wissen: Es bedarf nicht nur instrumenteller Kenntnisse zu Bedienung und Systemfunktionen, sondern auch eines Verständnisses der zugrundeliegenden Algorithmen, möglicher nicht-intendierter Effekte sowie Datenschutz- und rechtlicher Rahmenbedingungen.

Können: Digitale Handlungsfähigkeit entsteht erst durch erprobtes, erfahrungsbasiertes Tun. Ansätze wie Peer Learning können helfen, die Unterschiede in der technischen Expertise in der Mitarbeitendenschaft auszugleichen, müssen jedoch durch strukturierte Lernräume und Anreizsysteme ergänzt werden.

Haltung: Eine professionelle Grundhaltung erfordert die reflexive Auseinandersetzung mit ethischen Leitlinien, Ziel- und Wertekonflikten, beispielsweise zwischen digitaler Teilhabe und dem Schutz vulnerabler Gruppen.

Autonomie: Fachliche Urteilsbildung setzt Gestaltungs- und Entscheidungsspielräume in der Entwicklung und Anwendung von Systemen voraus. Nur so wird gewährleistet, dass digitale Anwendungen unterstützend wirken und nicht als zusätzliche Belastung empfunden werden (vgl. Seelmeyer 2020, S. 96f.).

Die Ausbildung dieser Kompetenzen befähigt Fachkräfte dazu, sich aktiv und reflektiert in die Gestaltung, Einführung und Nutzung von KI-Systemen einzubringen. Sie schafft damit die Grundlage, um für eine verantwortungsvolle, ethisch fundierte und sozial verträgliche KI-Entwicklung einzutreten (vgl. Steiner und Tschopp 2022, S. 470). Kompetenz bedeutet in diesem Zusammenhang nicht nur die sachgerechte Nutzung technischer Systeme, sondern auch die Fähigkeit, deren Einsatz kritisch zu hinterfragen, Position zu beziehen und entsprechende Rahmenbedingungen mitzugestalten.

Weiterführend: Seelmeyer, Udo (2020): Big Data & Künstliche Intelligenz – Neue Anforderungen an die Fachlichkeit in sozialen Berufen.

5 Aktuelle Überblicksstudien zum KI-Einsatz im Dritten Sektor in Deutschland

Nachdem in den vorangegangenen Kapiteln Einsatzmöglichkeiten und fachliche Aspekte zur Nutzung Künstlicher Intelligenz in Sozialen Organisationen dargestellt wurden, soll im Folgenden ein Überblick über drei empirische Studien gegeben werden, die für den Gegenstandbereich der KI-Nutzung in der Wohlfahrtspflege relevant erscheinen. Die Arbeiten beleuchten jeweils in Teilen auch den Einsatz von KI im Forschungsfeld, unterscheiden sich jedoch hinsichtlich des Feldzugangs und der inhaltlichen Fokussierung. Im Anschluss werden zentrale Ergebnisse vorgestellt und für die Weiterarbeit der wissenschaftlichen Begleitung eingeordnet.

	Züger et al. 2022 **Forschungsbericht**: „Civic Coding“	Schultz et al. 2023; 2024; 2025 **Studie:** „KINiro – KI in Nichtregierungsorganisationen“	Kreidenweis & Diepold 2024 **Forschungsbericht**: „Künstliche Intelligenz in der Sozialwirtschaft“
Sample	**Zivilgesellschaftliche KI-Expert:innen** für gemeinwohlorientierte Entwicklung	**Interessierte Personen; Digitalverantwortliche** aus Nichtregierungsorganisationen (NRO)	**Führungskräfte** aus der Sozialwirtschaft; **KI-Expert:innen** aus dem sozialwirtschaftlichen Umfeld
Methode	Expert:inneninterviews *(n=17)*; Fallstudien *(n=10)*	Expert:inneninterviews *(n=15)*; Quantitative Onlinebefragung *(n=335)*	Expert:inneninterviews *(n=9)*; Quantitative Onlinebefragung *(n=185)*
Fokus	▪ Überzeugungen und Erfahrungen der Praxis ▪ zentrale Probleme und Bedarfe; mögliche Lösungsansätze	▪ Grundsätzliche Akzeptanz ▪ Bedarfe, Erwartungen& Bedenken ▪ Erste Nutzung	▪ Kenntnisstand zum Thema KI ▪ Grundsätzliche Akzeptanz ▪ Erste Nutzung ▪ Künftig vorstellbare Anwendungsfälle

Tabelle 1: Überblick der untersuchten Studien

5.1 Forschungsbericht: Civic Coding

Der Forschungsbericht „Civic Coding“ von Züger und Kolleg:innen untersucht, wie KI-Entwicklung im Sinne des Gemeinwohls definiert, organisiert und umgesetzt werden kann. Er wurde im Auftrag des Bundesministeriums für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV) in Kooperation mit dem Bundesministerium für Arbeit und Soziales (BMAS) und dem Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend (BMBFSFJ) als Teil der Initiative „Civic Coding – Innovationsnetz KI für das Gemeinwohl” erstellt. Die Studie stellt einen frühen Überblick über Grundlagen, Herausforderungen und Chancen gemeinwohlorientierter KI dar und greift dazu auf Expert:inneninterviews (n=17) sowie die Vorstellung konkreter KI-Projekte aus dem zivilgesellschaftlichen Sektor anhand von Fallstudien (n=10) zurück (vgl. Züger et al. 2022, S. 12f.). Zu den

umsetzenden Organisationen in der Stichprobe zählen vor allem Hochschulen und kleinere zivilgesellschaftliche Initiativen, die sich in ihrer Verfasstheit von den stark formalisierten und hierarchischen Strukturen der Freien Wohlfahrtspflege unterscheiden. Die Unterschiede in Struktur, Finanzierung, vorhandenen Dateninfrastrukturen und gesellschaftlicher Stellung erschweren direkte Vergleiche und eine Übertragung von Erkenntnissen auf die hier vorliegende Studie. Dennoch sind die breit ausgeführten definitorischen Grundlagen, Herausforderungen, Erfolgsfaktoren, Bedarfe und Empfehlungen relevante Ansatzpunkte auch für KI-Projekte aus der Wohlfahrt. Die Studie identifiziert zentrale Potenziale und Risiken gemeinwohlorientierter KI und entwickelt 16 konkrete Empfehlungen zur nachhaltigen Förderung solcher Projekte. Ein zentrales Ankerprojekt, welches in Folge der Studie initiiert wurde, ist das Civic Data Lab. Dieses wird von einem Konsortium aus der Gesellschaft für Informatik, CorrelAid und dem Deutschen Caritasverband getragen und vom BMBFSFJ gefördert. Das Civic Data Lab unterstützt die Entwicklung gemeinwohlorientierter Datenvorhaben von der Konzeption bis zur Umsetzung sowie die allgemeine Entwicklung von Datenstrategien in zivilgesellschaftlichen Organisationen.

5.2 Arbeitspapiere: KI für Nichtregierungsorganisationen

Das Forschungsprojekt KINiro (Künstliche Intelligenz für Nichtregierungsorganisationen) der Ostbayerischen Technischen Hochschule Regensburg untersuchte im Rahmen der BMBFSFJ-Förderlinie „Künstliche Intelligenz für das Gemeinwohl“ zwischen 2023 und 2025 systematisch den Bedarf, die Akzeptanz und die Umsetzungsmöglichkeiten von Künstlicher Intelligenz in deutschen Nichtregierungsorganisationen. Das Projekt verfolgte einen sequenziellen Mixed-Methods-Ansatz, der qualitative und quantitative Erhebungsmethoden systematisch kombinierte. Die Forschungsarbeit gliederte sich in vier aufeinander aufbauende Phasen, die jeweils in separaten Arbeitspapieren dokumentiert wurden (vertiefend Scharf et al. 2023; Schultz et al. 2024; Hauer et al. 2024; Schultz et al. 2025). Zunächst erfolgte eine systematische Literaturrecherche, gefolgt von explorativen Experteninterviews, einer quantitativen Online-Befragung mit interessierten Mitarbeitenden und abschließenden vertiefenden Interviews. Über alle Erhebungsformate zeigt sich, dass sich der Einsatz von Künstlicher Intelligenz in deutschen Nichtregierungsorganisationen noch in einem frühen Entwicklungs-stadium befindet. Gut 15% der befragten Nichtregierungs-organisationen nutzen KI bereits regelmäßig (vgl. Hauer et al. 2024, S. 28), größtenteils jedoch erst seit den Durchbrüchen im Bereich der großen Sprachmodelle im Jahr 2022 (ebd. S.33). Der

Einsatz wird dabei eher als experimentell charakterisiert und setzt besonders auf leicht zugängliche Systeme großer Anbieter (z.B. ChatGPT, CoPilot, DeepL), vor Allem für die Verarbeitung von Text (z.B. Korrektur, Übersetzung, Recherche).

5.3 Forschungsbericht: Künstliche Intelligenz in der Sozialwirtschaft

Die Studie von Kreidenweis und Diepold (2024) liefert empirisch fundierte Einblicke in den aktuellen Kenntnisstand, die Akzeptanz sowie zu möglichen zukünftigen Nutzungsperspektiven von KI-Technologien in der Sozialwirtschaft. Damit steht sie in engem Bezug zum Forschungsinteresse dieser Studie. Methodisch folgt die Untersuchung einem Multi-Stakeholder-Ansatz im Mixed-Methods-Design, das qualitative und quantitative Erhebungen kombiniert. Neben leitfadengestützten Expert:inneninterviews (n = 9) wurde ein standardisierter Fragebogen an Führungskräfte von Organisationen der Sozialwirtschaft eingesetzt (n = 185). Darüber hinaus wurden Anbieter von Branchensoftware (n=27) als tertiäre Zielgruppe mittels eines weiteren quantitativen Fragebogens befragt, dessen Ergebnisse jedoch in dieser Publikation nicht vertiefend dargestellt werden (vgl. Kreidenweis & Diepold, S. 5 f.). Für die Ansprache der Teilnehmenden konnte die Studie auf den Adresspool des jährlich erscheinenden IT-Report für die Sozialwirtschaft zurückgreifen, einem Branchenreport zur Softwarenutzung in sozialen Organisationen. Die Ergebnisse zeigen, dass der tatsächliche Einsatz von KI in Verwaltungs- und Organisationsprozessen bislang gering ist: Verschiedene Nutzungsformen von KI zu Verwaltungs- und Organisationszwecken werden im Mittelwert von lediglich 13% der Befragten genutzt. Ausnahmen bilden hier KI-basierte Chatbots (43% tatsächliche Nutzung) und KI-Unterstützung im Rechnungseingang (30% tatsächliche Nutzung). Deutlich höher fällt hingegen die künftige Nutzungsbereitschaft aus: 75 % der Befragten können sich den Einsatz von KI im Verwaltungsbereich vorstellen. In der direkten Arbeit mit Klient:innen ist der tatsächliche Einsatz derzeit noch marginal (8 %), doch auch hier geben 80 % der Befragten an, sich einen künftigen Nutzen zu fachlichen Zwecken vorstellen zu können (ebd. S. 22f.). Diese Befunde sind jedoch mit Vorsicht zu interpretieren. Erstens erlauben sie keine Rückschlüsse auf die organisationsweite Nutzung, sondern spiegeln ausschließlich die Perspektive der befragten Führungskräfte wider. Zweitens bleibt unklar, welche Art und welcher Umfang der Technologien konkret gemeint sind (z. B. ob bereits eine einzelne Nutzung von maschineller Übersetzung als „Einsatz von KI“ gewertet wird). Drittens können Verzerrungen durch die Auswahl und Formulierung der abgefragten Items entstehen. So würde etwa eine differenzierte Erhebung zwischen „generativer KI zur

Texterstellung und -zusammenfassung“ und „eigenentwickelten KI-basierten Chatbots zur automatisierten Beantwortung von Anfragen“ vermutlich zu veränderten Prozentwerten führen. Grundsätzlich ist die Akzeptanz dort am höchsten, wo bekannte Anwendungsformen mit klar erkennbaren Vorteilen für Personal oder Klient:innen verbunden sind (vgl. ebd. S. 24). Abschließend benennen die Befragten insbesondere die eigenen Verbands- und Trägerstrukturen als zentrale Ressource beim Einsatz von KI, sowohl in Bezug auf die Auswahl geeigneter Lösungen als auch hinsichtlich der notwendigen Kompetenzentwicklung (vgl. Kreidenweis & Diepold 2024, S. 28f.).

6 Fragestellung und Erhebungsdesign der Studie

6.1 Forschungslücken und Fragestellung

Die vorliegende empirische Studie knüpft an die zuvor beschriebenen Entwicklungen und Studien im Bereich der Künstlichen Intelligenz an und berücksichtigt dabei die spezifische Nutzung in sozialen Diensten. In Abgrenzung zu bestehenden Untersuchungen, die sich vorwiegend auf medizinisch-pflegerische Kontexte (z. B. Wundversorgung) den Katastrophenschutz (z. B. KI-gestützte Auswertung von Baumbeständen) oder den Nichtregierungssektor im Allgemeinen konzentrieren, liegt der Fokus dieser Studie auf Organisationen der Freien Wohlfahrtspflege und deren Zielgruppen. Die jüngsten Fortschritte im Bereich der generativen KI haben eine verstärkte Auseinandersetzung mit der Implementierung und Nutzung von KI sowohl in der Wissenschaft als auch in der Praxis angestoßen. Aktuelle Publikationen zur Anwendung gemeinwohlorientierter KI in der Sozialen Arbeit zeigen erste mögliche Ansätze zur Nutzung auf, beleuchten jedoch eher prospektiv künftige Nutzungsszenarien, wobei bislang eine systematische Untersuchung der praktischen Nutzung und des konkreten Innovationsgeschehens innerhalb wohlfahrtsspezifischer Kontexte weitgehend fehlt. Treffend spitzen Plafky et al. zu: „Was bislang fehlt, sind verlässliche und robuste Forschungsergebnisse bezüglich tatsächlicher Einsatzmöglichkeiten in Bezug von KI im Bereich der Arbeit in sozialen Einrichtungen unterschiedlicher Handlungsfelder“ (Plafky et al. 2022, S. 117). Die vorliegende Studie hat sich deshalb zum Ziel gesetzt, den aktuellen Status quo von KI-Innovationen in der deutschen Wohlfahrtspflege und die formelle und informelle Nutzung von KI in der Freien Wohlfahrtspflege zu untersuchen. Ein besonderer Fokus liegt dabei auf der Analyse von Innovationsstrategien in KI-Projekten (Expert:inneninterviews) sowie den Wahrnehmungen zu Nutzungsbarrieren und möglichen Risiken auf der Ebene von Fachkräften (Fokusgruppe).

6.2 Methodisches Vorgehen

Aufbauend auf den theoretischen Vorarbeiten wurde ein Studiendesign entwickelt, das mittels leitfadengestützter Expert:inneninterviews und einer Fokusgruppendiskussion das Innovationsgeschehen und die konkrete Nutzung von KI in Organisationen der Freien Wohlfahrtspflege empirisch untersucht. Zunächst wurden in einer ersten Phase modellhafte KI-Projekte identifiziert und befragt, in

einer zweiten Phase sollte dann untersucht werden, wie sich die KI-Nutzung auf lokaler Ebene jenseits modellhafter Vorhaben zeigt.

Beschreibung der Interviewstudie

Für die erste Erhebung wurde ein qualitatives Studiendesign mit leitfadengestützten Expert:inneninterviews umgesetzt. Dazu wurden Personen aus den Wohlfahrtsverbänden befragt, die Auskunft über konkrete Umsetzungsprojekte mit KI-Technologien innerhalb der Wohlfahrt geben konnten. Solche Expert:inneninterviews sind besonders gut geeignet, um kontextspezifisches Wissen, hier insbesondere zur Arbeitspraxis in KI-Projekten in Strukturen der Wohlfahrtspflege, zu rekonstruieren. Die Expert:innen liefern einen zielgerichteten Zugang zum kontextspezifischen Wissen ihrer Arbeitspraxis und (vgl. Gläser und Laudel 2010, S. 11ff.). Es wurde ein Leitfaden entwickelt, der entlang von vier Frageblöcken strukturiert war, die von einleitenden Fragen über das allgemeine organisationale KI-Klima bis hin zur Darstellung eines spezifischen KI-Projekts beim jeweiligen Träger reichten und mit Überlegungen zum zukünftigen Einsatz von KI abschlossen (Abb. 1).

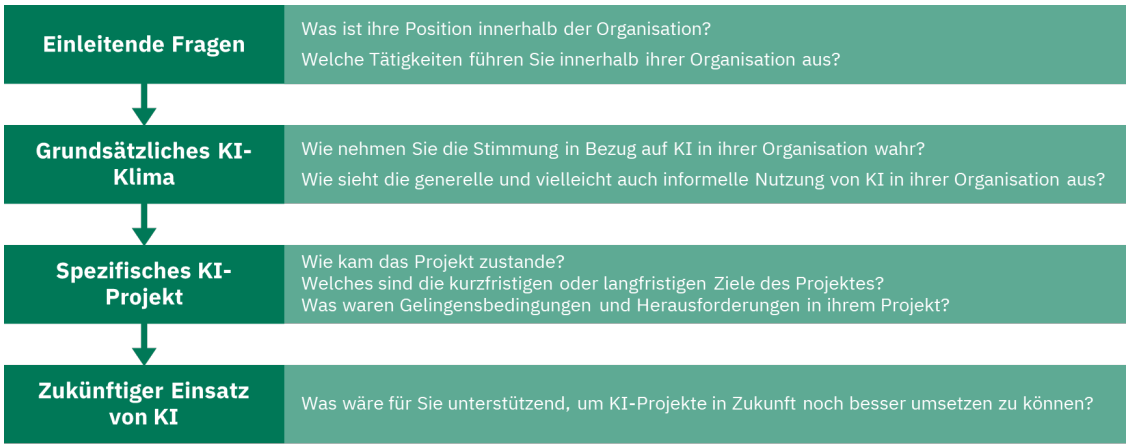


Abbildung 1: Exemplarische Ausschnitte aus dem Interviewleitfaden

Die Ansprache potentieller Interviewpartner:innen erfolgte mittels eines Schnellballverfahrens, ausgehend von den Projektleitungen des begleiteten Förderprogramms zur „Zukunftssicherung der Freien Wohlfahrtspflege durch Digitalisierung“ in den sechs Spitzenverbänden. Ergänzend wurde eine Desktop-Recherche durchgeführt, um potenzielle weitere Teilnehmer:innen zu identifizieren. Auf dieser Grundlage wurden 19 Projekte identifiziert, in denen KI strategisch eingesetzt wird und die von Organisationen der Freien Wohlfahrtspflege hauptverantwortlich getragen werden. Die Verantwortlichen dieser Projekte wurden per E-Mail zur Teilnahme an der Studie eingeladen. Bei sieben Projekten erfolgte

jedoch entweder keine Rückmeldung oder der Hinweis, dass sich das Vorhaben noch in einer sehr frühen Phase befinde.
Im Zeitraum von März bis Mai 2025 wurden insgesamt zwölf leitfadengestützte Expert:inneninterviews durchgeführt. In der Stichprobe (n=12) sind fünf von sechs Wohlfahrtsverbänden vertreten. Ausgehend von Hinweisen auf der Bundesebene und weiteren Recherchen wurden auch Projekte auf Landes- und kommunaler Ebene identifiziert. Insgesamt zeigte sich, dass entgegen der starken Thematisierung von KI weniger konkrete Umsetzungsprojekte identifiziert werden konnten, als zunächst gedacht. Die befragten Personen kamen vor allem aus dem mittleren Management und hatten in ihrer Funktion zumeist mit Themen der Digitalisierung zu tun.

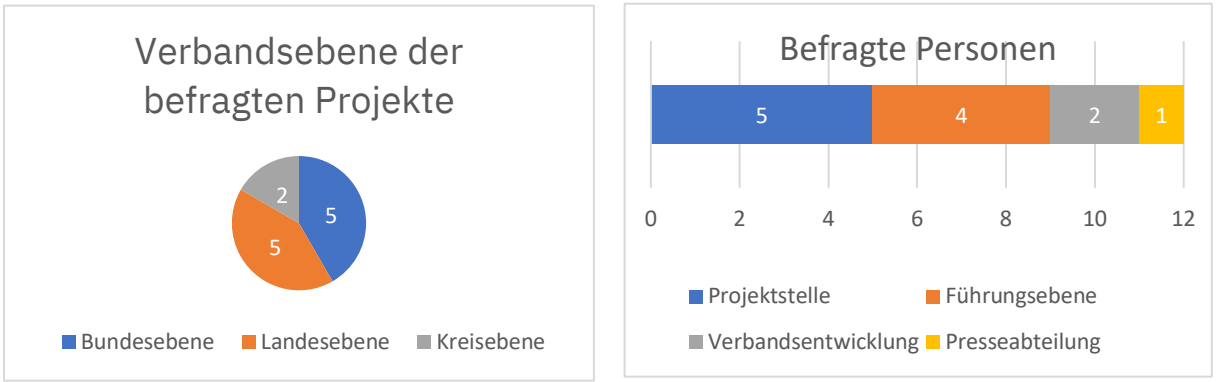


Abbildung 2: Visualisierung der Stichprobe

Es ist zu betonen, dass die so erhobene Stichprobe jedoch nur einen Teil der KI-Projekte in der Freien Wohlfahrtspflege umfasst. Es ist davon auszugehen, dass gerade in den letzten beiden Jahren bedeutende Projektvorhaben angestoßen wurden, die 1) (noch) nicht publiziert wurden (und deshalb nicht bei der Desktopresearch aufgefunden wurden), oder 2) nicht von den Projektleitungen benannt wurden oder 3) die im Rahmen der Ansprache im Schneeballverfahren erreicht wurden, sich aber nicht zurückgemeldet haben.

Auswertungsstrategie

Die Auswertung der so zusammengesetzten Stichprobe erfolgte methodisch orientiert an der Qualitativen Inhaltsanalyse nach Kuckartz (2018). Da das Forschungsinteresse explorativ ausgerichtet war, wurde im ersten Analyseschritt ein induktives Vorgehen gewählt. Das bedeutet, das Kategoriensystem wurde nicht a priori festgelegt, sondern direkt aus dem Datenmaterial der 12 Interviews entwickelt.

Hierfür wurden die Transkripte[10] sequenziell durchgearbeitet, relevante Textstellen paraphrasiert und schrittweise zu abstrakteren Kategorien zusammengefasst. Dieses Vorgehen stellte sicher, dass die spezifischen Charakteristika und Nuancen der untersuchten KI-Projekte unvoreingenommen erfasst wurden. Das finalisierte Kategoriensystem bildete einen geeigneten Ausgangspunkt, um die Einzelfälle miteinander zu vergleichen und Muster zu identifizieren. Auf der Basis dieser so entwickelten Datengrundlage kam zusätzlich zur Inhaltsanalyse das Verfahren der Typenbildung nach Kuckartz (2010) zur Anwendung. Die Typenbildung dient der Strukturierung des Forschungsfeldes, um die Singularität von Einzelfällen zu überwinden. Dabei erfolgt eine strikte Abgrenzung zur intuitiven Alltagstypisierung: Das Vorgehen ist wissenschaftlich regelgeleitet und methodisch kontrolliert. Orientiert haben wir uns an qualitativen Prinzipien, wobei die kontinuierliche Reflexion und lückenlose Dokumentation aller Arbeitsschritte im Zentrum standen. Bei der Typenbildung werden „aufgrund von Ähnlichkeiten in ausgewählten Merkmalsausprägungen [...] Objekte zu Typen zusammengefasst. Dabei sollen die Objekte desselben Typs einander möglichst ähnlich, die verschiedenen Typen hingegen möglichst unähnlich sein." (Kuckartz 2010, S. 555f.). Die Typenbildung dient als zentrale Strategie, um subjektive Vielfalt auf übergeordnete Muster zu verdichten. Über den analytischen Wert hinaus besitzt die Typenbildung hohe praktische Relevanz, da sie eine fundierte Basis für die Entwicklung zielgerichteter Interventionsmaßnahmen bildet (ebd.). „Mit dem Begriff Typus werden die gebildeten Teil- oder Untergruppen bezeichnet, die gemeinsame Eigenschaften aufweisen und anhand der spezifischen Konstellation dieser Eigenschaften beschrieben und charakterisiert werden können" (Kelle und Kluge 2010, S. 85).
Der Prozess der Typenbildung gliedert sich in vier Phasen (vgl. Kelle und Kluge 2010, S. 91f.):

1. **Erarbeitung von Vergleichsdimensionen:** Festlegung relevanter Merkmale und Kategorien, um Ähnlichkeiten und Unterschiede zwischen den Fällen zu erfassen.
2. **Gruppierung und Analyse:** Zusammenfassung der Fälle basierend auf empirischen Regelmäßigkeiten. Ziel ist eine hohe **interne Homogenität** innerhalb eines Typs bei gleichzeitiger **externer Heterogenität** zwischen den Typen.
3. **Analyse von Sinnzusammenhängen:** Inhaltliche Deutung der Gruppen, um den Gegenstand nicht nur zu beschreiben, sondern durch Reduktion auf wenige Typen erklärend zu verstehen.

[10] Die Transkription ist angelehnt an die Regeln nach Dresing und Pehl 2016.

4. **Charakterisierung:** Abschließende Beschreibung der konstruierten Typen anhand ihrer spezifischen Merkmalskombinationen.

Ein Typus entsteht somit aus der Verschränkung von empirischen Regelmäßigkeiten und inhaltlichen Sinnzusammenhängen.

Beschreibung der Fokusgruppe

Für die zweite Erhebung wurde die Methode der moderierten Fokusgruppe gewählt. Während Einzelinterviews vorrangig individuelle Sichtweisen abbilden, ermöglichen Fokusgruppen eine Erhebung kollektiver und geteilter Perspektiven auf den KI-Einsatz in sozialen Einrichtungen. Im Gruppendiskurs werden nicht nur persönliche Einschätzungen, sondern auch gemeinsame Deutungsmuster und Spannungsverhältnisse sichtbar, die für das Verständnis der aufkommenden Nutzung zentral sind. Fokusgruppen fördern einen distanzierteren, vergleichenden Blick auf die eigene Einrichtung und das Handlungsfeld insgesamt. So kann beispielsweise erörtert werden, inwiefern KI-Alltagstechnologien bereits in der Praxis verankert sind bzw. erprobt werden, ohne dass dies institutionell-strategisch gesteuert wird. Gerade das Spannungsfeld zwischen Managementebene und fachlicher Praxis, also zwischen strategischer Planung und konkreter Anwendung, lässt sich in der Diskussion mit mehreren Teilnehmenden differenzierter beleuchten. Zugleich ist zu berücksichtigen, dass in Fokusgruppen bestimmte Erfahrungen möglicherweise nicht offen geteilt werden, etwa aufgrund von sozialer Erwünschtheit oder der Zurückhaltung, interne Sachverhalte der eigenen Organisation vor Vertreter:innen anderer Verbände preiszugeben. Dies kann dazu führen, dass sensible oder kritische Aspekte unterrepräsentiert bleiben.

In der Fokusgruppe wurde auf eine verbandsübergreifende Zusammensetzung der Stichprobe geachtet. Dieses Vorgehen sollte sicherstellen, dass die Analyse nicht durch eine einseitige Zusammensetzung des Samples verzerrt wird. Die Fokusgruppendiskussion wurde im Oktober 2025 mit 5 Teilnehmenden aus Wohlfahrtsorganisationen in Ostwestfalen-Lippe (OWL) durchgeführt, die jeweils auf mittleren Managementebenen tätig waren. Diese Gruppe wurde adressiert, da davon ausgegangen wurde, dass auf dieser Ebene noch ein hinreichend guter Einblick in die ganz konkrete Arbeitspraxis vor Ort besteht und sich dieser Einblick zugleich aber nicht nur auf ein Team bezieht, sondern einen etwas größeren Bereich überblickt.

7 Auswertung der Interviews

In nahezu allen gesellschaftlichen Bereichen ist gegenwärtig von **Innovation** die Rede: Sie gilt nicht nur in der Wirtschaft als Schlüssel der Wettbewerbsfähigkeit, sondern wird auch in Kunst, Wissenschaft und Politik positiv bewertet und gewünscht. Innovationen lassen sich auf struktureller Ebene als Evolutionsprozess verstehen, in dem sich Praktiken und Ordnungen grundlegend erneuern. Um als Innovation zu gelten, muss „eine Idee, eine Praktik oder ein Objekt als etwas Neues und Besseres wahrgenommen, wertgeschätzt und weiterkommuniziert werden muss" (Rammert 2010, S. 34). Innovation ist stets eine Zweiheit: „Sie umfasst sowohl eine Strukturänderung, die in einer spezifischen Art und Weise geschieht, als auch eine entsprechende Form der semantischen Begleitung dieses Wandlungsprozesses" (Jungmann und Besio 2018, S. 14f.).

Genau diese **spezifische Art des Wandels** markiert die Grenze zur Revolution oder zur Reform. Der Unterschied zwischen Innovationen und Revolutionen (Brunkhorst 2003) bezieht sich auf die Radikalität der Veränderung. Während Innovationen evolutionär verlaufen, sind sie „keine Revolutionen, die, im Sinne Mertons, implizieren würden, dass [sich] sowohl die Mittel als auch die Ziele einer Gesellschaft, d. h. eine etablierte Sozialordnung komplett verändern" (Zapf 1989, S. 177). Was sich im Falle einer Revolution ändert, wird folglich nicht als einzelnes Artefakt oder Programm, sondern als umfassender Strukturkomplex verstanden.

Eine weitere Differenzierung zeigt sich im Vergleich zur **Reform**. Zwar rekurrieren beide Konzepte auf die Idee der Verbesserung, doch setzen sie unterschiedliche Akzente. Während es sich bei Innovationen um die Nutzung des Zufalls für die Zukunft handelt – oft bewerkstelligt durch eine:n Entrepreneur:in und durch das Einfügen neuer Artefakte (Jungmann und Besio 2018) –, betont die Reform die Identifikation problematischer Aspekte der Vergangenheit (Baecker 2005). Hier agiert eher der:die Analyst:in, der Problemlagen in reparierender Absicht bearbeitet (Jungmann und Besio 2018). Reform bezeichnet somit eine Modifikation bestehender Programme, um Defizite der Vergangenheit zu beheben, wohingegen Innovation das Neue in die Zukunft projiziert.

Im folgenden Kapitel beziehen wir uns auf dieses Verständnis von Innovation und stellen die zentralen Ergebnisse unserer Auswertung vor, wobei die Identifikation und Charakterisierung spezifischer **Innovationstypen** im Mittelpunkt steht.

7.1 Übersicht der erhobenen Interviews und Vorstellung der folgenden Kategorien- & typengeleiteten Auswertung

Die durchgeführte Interviewstudie fängt exemplarische Momentaufnahmen aktueller KI-Projekte in der Wohlfahrt ein. Dies wird daran deutlich, dass elf der zwölf Vorhaben sich in der aktiven Umsetzung befinden und lediglich ein Projekt mangels einer Anschlussfinanzierung ruht. Eine knappe Mehrheit der Projekte wurde zum Auftakt durch Eigenmittel finanziert, was auf eine hohe interne Relevanz der Themen und damit verbunden eine Investitionsbereitschaft innerhalb der Verbände hindeutet. Technologisch konzentrieren sich alle Projekte auf die Verarbeitung natürlicher Sprache. Zehn Vorhaben nutzen dabei Ansätze generativer Künstlicher Intelligenz, während zwei Projekte auf nicht-generative Lösungen wie Übersetzungs- oder Dokumentationshilfen setzen. In der Beschaffung zeigt sich eine Tendenz zu vollständigen Eigenentwicklungen oder Entwicklungskooperationen, gegenüber dem Einkauf von „schlüsselfertigen“ Standardlösungen (Abb. 2).

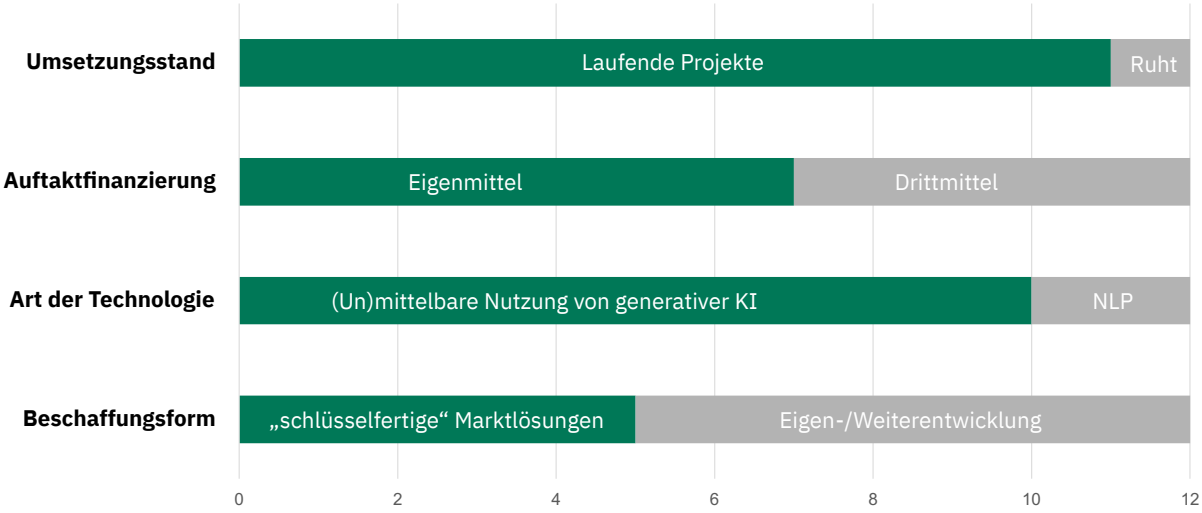


Abbildung 3: Fallübersicht der durchgeführten Interviewstudie

Die folgenden Kapitel widmen sich der detaillierten Betrachtung dreier unterschiedlicher Innovationstypen in KI-Projekten im Feld der Wohlfahrt. Die aus dem Interviewmaterial gewonnen Oberkategorien der Innovationssemantiken, der Sozialen Welten sowie des Innovationsprozesses dienen dabei als strukturierende Klammer. Um ein grundlegendes Verständnis für diese analytischen Perspektiven zu schaffen, wird jede der drei Kategorien zu Beginn des jeweiligen Kapitels kurz theoretisch und inhaltlich eingeleitet.
Entlang dieser Dimensionen entfaltet sich das Profil der drei Innovationstypen. Der Typus der Pragmatischen Modernisierung (Typ A) wird als Modus vorgestellt, der primär auf die Stabilisierung des Status Quo zielt und Sicherheit durch etablierte

Verfahren sucht. Demgegenüber wird mit der Reflexiven Sozialinnovation (Typ B) ein Ansatz beschrieben, der auf strategische Weiterentwicklung in Netzwerken und eine multiperspektivische Abschätzung setzt. Als dritte Ausprägung wird der Typus der Disruptiven Innovation (Typ C) eingeführt, dessen Fokus auf einer Emanzipation der Wohlfahrt und dem Streben nach technologischer Souveränität liegt. Die nachstehende Tabelle bietet einen ersten Überblick über diese Zusammenhänge und die Merkmalsausprägungen, bevor sie im Anschluss tiefgehend erörtert werden.

	Typ A:	Typ B:	Typ C:
	Pragmatische Modernisierung	**Reflexive Sozialinnovation**	**Experimentell-disruptive Innovation**
Innovationssemantik			
Zukunftsentwurf	Stabilisierung des Status Quo	Strategische Entwicklung	Emanzipation der Wohlfahrt
Risiko- und Begrenzungsstrategien	Vorsichtsprinzip	Multiperspektive Abschätzung	Risiko als kalkulierter Bestandteil des Innovativen
Organisationswandel	Exogene Anpassung	Prozessöffnung	Strukturelle Disruption
Soziale Welten			
Akteurskonstellation	Formale Interdependenz	Partnernetzwerke	Entkoppelte Avantgardist:innen
Ressourcen	Nutzung institutioneller Ressourcen	Projekt- und Netzwerkbasierte Ressourcen	Individuelle Ressourcen
Dynamiken & Konflikte	Sicherheit durch Verfahren	Zusammenarbeit mit Externen	Legitimation durch Funktion
Innovationsprozess			
Ideengenerierung	Reaktion auf ein wahrgenommenes Funktionsdefizit	Chancenbasierte Ideenfindung	Neugierige Pionier:innen
Ideenakzeptierung	Formalisierter Auswahlprozess	Kooperative Aushandlung	Diffusion durch Vorreiter:innen
Ideenrealisierung	Return on Investment	Projekt-Falle	Technologische Souveränität

Tabelle 2: Überblick über Typen und Merkmalsdimensionen

7.2 Innovationssemantiken

Unter **Semantik** ist allgemein zu verstehen, dass ein Sinn sich durch Wiederholung verfestigt und entsprechend generalisiert zur Verfügung steht. "Semantiken umfassen Begriffe und Unterscheidungen, die Beschreibungen in einem bestimmten sozialen Kontext anleiten. Sie werden nicht in Frage gestellt, sondern sind für bestimmte kommunikative Kontexte schlicht einleuchtend und plausibel" (Jungmann und Besio 2018, S. 16). Strukturen und Semantiken können dabei durchaus auseinanderlaufen.

Innovationssemantiken lassen sich am besten verstehen, wenn sie gegenüber anderen Formen des menschengemachten Neuen abgegrenzt werden, beispielsweise gegenüber dem Revolutionärem oder der Reform. Innovation gilt als eine artifizielle Abweichung – eine „mindful deviation" (Garud und Karnøe 2001) –, die sich um eine bewusst hergestellte, klar abgrenzbare symbolische oder materielle Einheit zentriert. Typischerweise wird dann die Innovation semantisch begleitet (vgl. Jungmann und Besio 2018, S. 28). Das heißt, auf Basis eines Artefaktes wird eine positive Zukunft prospektiv antizipiert, die sich von dem Vergangenen abhebt. Innovationen folgen demnach der Logik eines "Vorher-Nachher-Vergleichs" (ebd. S. 31).

Zusammengenommen zeichnet sich die Innovationssemantik als Form der Kommunikation dadurch aus, dass sie als eine Beschreibung vorliegt, die nicht in Frage gestellt wird. Innovativ ist eine Semantik dann, wenn sie auf ein Artefakt, wie eben die KI, zentriert ist, sie mit einer zukünftigen Positivität verbunden ist und das integrierte Artefakt auf seine Folgen hin kommunikativ begleitet wird.

Entsprechend dieses Verständnisses lassen sich sehr unterschiedliche Innovationssemantiken bei den von uns Interviewten feststellen. Sie unterscheiden sich nicht so sehr in der Artefaktzentrierung, hier gehen alle von einer recht homogenen Technologie der Large Language Models aus. Vielmehr unterscheiden sich die Semantiken hinsichtlich der inhaltlichen Konturierung der zukünftigen Positivität, der Deutung und des proklamierten Umgangs mit Risiken und Begrenzungsbedarfen sowie der erwarteten organisationalen Veränderungen. Diese Aspekte werden im Folgenden herausgearbeitet. Wichtig zu beachten ist dabei, dass es sich bei allen nachfolgenden Beschreibungen und insbesondere bei den Innovationssemantiken gerade nicht um beobachtete Praktiken handelt, sondern vielmehr um Erzählungen, über deren Deckung mit real vorfindbaren sozialen Praktiken hier keine Aussagen getroffen werden können.

Theoretisch kann der hier verwendete Begriff der „Semantik" an das Konzept der „soziotechnischen Imaginäre" anschließen, wie es von Jasanoff und Kim (2015) in

den Science and Technology Studies (STS) entwickelt wurde. Soziotechnische Imaginäre beschreiben kollektiv geteilte, institutionell stabilisierte und öffentlich performte Zukunftsvisionen, die durch geteilte Vorstellungen wünschenswerter sozialer Ordnungen geprägt sind und durch wissenschaftlich-technologischen Fortschritt ermöglicht werden sollen (Jasanoff 2009). Vor dem Hintergrund des Konzeptes wird erkennbar, dass KI-Innovationssemantiken nicht lediglich individuelle Meinungen darstellen, sondern kollektive Deutungsmuster, die organisationales Handeln, Entscheidungen und institutionelle Prioritäten maßgeblich beeinflussen (Jasanoff & Kim, 2015). Übertragen auf den Kontext Sozialer Dienste zeigt sich, dass die jeweiligen Imaginäre nicht nur abbilden, wie Akteure über Technologie sprechen, sondern auch, welche Zukünfte sie als erstrebenswert oder vermeidbar rahmen und welche sozialen Ordnungsvorstellungen dabei transportiert werden (Bareis & Katzenbach 2022). Gerade im Feld von KI ist die Wirkmächtigkeit solcher Imaginäre besonders ausgeprägt, da KI als transformative Technologie mit weitreichenden gesellschaftlichen Implikationen verhandelt wird und dabei konkurrierende Narrative von Fortschritt, Risiko und Kontrolle hervorruft (Sartori & Bocca 2023). Die im Folgenden rekonstruierten Innovationssemantiken der Befragten lassen sich vor diesem Hintergrund als organisational situierte Varianten soziotechnischer Imaginäre lesen, die das Verhältnis von KI, Wohlfahrtspflege und organisationalem Wandel auf je spezifische Weise konfigurieren.

Zukunftsentwürfe

Der Zukunftsentwurf beim Typus der **Pragmatischen Modernisierung** (Typ A) zielt primär auf die **Stabilisierung des Status Quo**. Angesichts eines als bedrohlich wahrgenommenen externen Ressourcendrucks wird KI als notwendiges Kompensationsinstrument gerahmt. Sie dient dazu, das bestehende System unter Mangelbedingungen funktionsfähig zu halten: „Die Alternative [zu KI] ist, dass eben da niemand ist.“ (ID5, Pos.52). Die positive Nutzungsperspektive liegt in der Entlastung von administrativen Routinetätigkeiten um „Freiräume“ für die analoge, zwischenmenschliche Kerntätigkeit, zu bewahren (ID4, Pos.33). KI als Innovation wird dabei beschrieben als die nächste Evolutionsstufe der Digitalisierung, als der „normale technische Fortschritt [...] der normale Prozess, dem folgt man“ (ID7, Pos.44). Die Innovation erscheint in dieser Semantik als exogen induziert und als reaktiver Prozess in Form einer Implementierung marktfähiger, kommerziell erworbener Softwareprodukte in der Organisation, oder der Erweiterung bestehender Lösungen um KI-Komponenten.

Demgegenüber hebt die Semantik beim Typ der **Reflexiven Sozialinnovation** (Typ B) die **Strategische Entwicklung** im Zusammenhang mit Künstlicher Intelligenz hervor.

Das Ziel ist eine qualitative Weiterentwicklung des „Dienstes am Menschen“, der als „Kernauftrag“ der Wohlfahrt unterstrichen wird (ID4, Pos.109). In diesem Entwurf wird KI verstanden als „Potenzial einen besseren Service anzubieten, statt unsere Mitarbeiter zu reduzieren und Geld zu sparen“ (ID11, Pos.151) und langfristig als Möglichkeit einer datengestützten Steuerung („besser lenken und leiten“, ID4, Pos. 69).

Am radikalsten ist die Innovationssemantik im Typ der **Experimentell-disruptiven Innovation** (Typ C) mit der Zukunftsvorstellung einer **Emanzipation der Wohlfahrt**. Hier wird KI als Instrument der Befreiung von strukturellen Abhängigkeiten im Kontext der Digitalisierung von Organisationen gedeutet. KI gilt hier als „größte digitale Revolution seit dem Internet“ (ID8, Pos.12), die es der Wohlfahrt ermöglicht, maßgeschneiderte Werkzeuge „inhouse“ zu produzieren und damit die eigene Handlungsfähigkeit fundamental zu erweitern.

Das Narrativ bricht mit der Rolle der Wohlfahrt als passiver Konsumentin von Softwarelösungen und der damit einhergehenden Abhängigkeit von Dienstleistern. Dem entgegen gesetzt wird die positive Vision der Rückgewinnung der technologischen Souveränität: Soziale Arbeit soll sich „selber auf den Weg machen“ (ID1, Pos.90) und mit Hilfe von KI eigene Lösungen entwickeln, um so ihre Autonomie gegenüber profitorientierten Agenturen zurückzugewinnen.

Risiko- & Begrenzungsstrategien

Die Art und Weise, wie der Einsatz der KI-Innovation beobachtet und bewertet wird, offenbart spezifische Risikokulturen, die bei den drei identifizierten Typen sehr unterschiedlich ausgeprägt sind.

Die **Pragmatische Modernisierung** (Typ A) folgt im Umgang mit Risiken dem **Vorsichtsprinzip**. Die umfangreichen Funktionen, die die marktfähigen Lösungen bieten, werden zwar als nützlich, aber auch als potenziell fehleranfällig beschrieben. KI liefere „nie 100% Output“ und erfordere daher ständige Kontrolle, man müsse „immer nachbessern“ (ID2, Pos. 32). Entsprechende Limitationen werden darum betont: „KI darf keine Autorität ersetzen, da müssen wir klare Grenzen setzen“ (ID2, Pos. 50). Die Innovation gilt in diesem Narrativ als reversibel und muss sich an ihrer Effizienz messen lassen, wie die Befragten im Kontext eingekaufter Marktlösungen formulieren: „Wenn wir merken, wir sparen nicht damit Zeit ein, dann werden wir es auch definitiv wieder abschaffen [...]“ (ID4, Pos.43).

Die Semantik der **Reflexiven Sozialinnovation** (Typ B) betont eine **multiperspektivische Abschätzung** des Risikos. Hier werden Risiken frühzeitig und systemisch abgewogen. Insbesondere wird die Gefahr der Arbeitsverdichtung thematisiert, da durch KI freiwerdende Zeitfenster sofort neu gefüllt werden könnten:

„[Der KI-Einsatz] muss moderiert werden. Die Leute brauchen Sicherheit, sonst verlieren wir sie.“ (ID3, Pos. 20). Zudem werden datenschutzrechtliche Grenzen bei der Arbeit mit sensiblen Daten als absolute Schranke anerkannt. Die KI „soll nicht auf Klientendaten zugreifen, damit fallen manche KI-Anwendungen einfach weg“ (ID4, Pos. 70), was eine bewusste Begrenzung des Innovationsraums zum Schutz der Adressat:innen darstellt.

Hingegen deutet die dominierende Semantik der **Experimentell-disruptiven Innovation** (Typ C) **Risiko als kalkulierten Bestandteil des Innovativen**. Die Semantik verschiebt sich hier vom Risikovermeiden zum Risikomanagement durch Agilität. Projekte werden beispielsweise „bewusst einfach angesetzt, um keinen Datenschutzärger zu haben“ (ID8, Pos. 55). Um in der Entwicklung und Testung schnell voran zu kommen, wird eine „Koalition der Willigen“ gebildet, die bereit und fähig ist, die Veränderungen mitzutragen (ID1, Pos. 21). Das größte Risiko wird hier nicht in der Anwendung der Technik gesehen, sondern im Unterlassen der Innovation. Fehlende Ressourcen wie Entwickler:innen werden dabei nicht als Stopp-Signal gewertet, sondern agil kompensiert („wir haben es dann anders gelöst“, ID1, Pos. 36).

Organisationale Veränderungen

Die drei Typen implizieren in ihren Semantiken unterschiedlich tiefgreifende Eingriffe in organisatorische Strukturen.

Die **Pragmatische Modernisierung** (Typ A) charakterisiert die KI-bedingten Wandlungsprozesse vor allem als **exogene Anpassung**. In diesem Modus bleibt die institutionelle Struktur weitgehend unangetastet, während neue KI-Werkzeuge in bestehende organisationale Strukturen integriert werden. Die Organisation agiert hierbei als reaktiver Anwender des „normalen technischen Fortschritts“ (ID7, Pos. 42), der Marktstandards wie den Microsoft Copilot übernimmt. Dabei wird die Verantwortung für den Wandel teilweise externalisiert, indem ein „KI-Schub [...] vonseiten der Krankenkassen [oder] Ministerien“ (ID6, Pos. 59) eingefordert wird, um die Implementierung zu finanzieren und so überhaupt erst zu ermöglichen.

Im Kontrast dazu fordert die **Reflexive Sozialinnovation** (Typ B) eine **Prozessöffnung** der Organisationsstruktur. Der Wandel vollzieht sich hier prozessorientiert, indem die KI der Organisation einen „Spiegel“ vorhält, der aufdeckt, wie „tradiert“ manche Prozesse sind und erzwingt so eine Diskussion über deren Sinnhaftigkeit (ID3, Pos.14). Durch diese Reflexion sollen im Zuge der KI-Innovation interne Abläufe aufgebrochen und neu verhandelt werden.

In der Semantik der **Experimentell-disruptiven Innovation** (Typ C) wird die Einführung von KI als **strukturelle Disruption** etablierter organisationaler

Digitalisierungslogiken gerahmt. Die Befragten beschreiben diesen Wandel als ein „Herauslösen“ von Innovationen aus starren Projektlogiken, wodurch eine agile Projektorganisation entsteht. Bezeichnend für diese Semantik ist der Vergleich mit der Startup-Kultur, in welcher „Startup-like viele, viele kleine Produkte“ (ID1, Pos. 14) selbst entwickelt werden. Innerhalb dieser Argumentationslogik erfährt auch die Rolle der Entwickler:innen eine Neudeutung, indem sie nicht mehr primär als Ersteller, sondern als Instanz zur „Absicherung“ und „Implementierung“ (ID1, Pos. 30) semantisiert werden.

7.3 Soziale Welten

Anselm Strauss (1978), auf den das Konzept „sozialer Welten" zurückgeht, versteht unter sozialen Welten Gruppen, die durch mindestens eine „primary activity" (Strauss 1978, S. 122) gekennzeichnet sind. Die in ihnen Handelnden haben ein gemeinsames Anliegen, gemeinsame Ziele und teilen dieselben Ansichten. Das heißt, die Akteure einer sozialen Welt haben eine geteilte Perspektive auf ihr Handlungsfeld. Soziale Welten sind keine geschlossenen Gemeinschaften, sie sind veränderbar. Sie entwickeln mitunter formelle Organisationen und können eventuell auch »Subwelten« haben (Strauss 1998, S. 293).
Neben den Aktivitäten teilen die Mitglieder Ressourcen verschiedenster Art, um ihre Ziele zu erreichen (Clarke 1991, S. 131). Dadurch können sich soziale Welten in ihrer Größe, ihrer Geschichte, ihrer Veränderungsrate, ihren Ressourcen oder auch ihrem Verhältnis zu Technologien sowie Artefakten voneinander unterscheiden (Strauss 1993, S. 213). „Some worlds are quite local, others regional or national, while an increasing number are international in scope of activities and membership" (ebd.). Soziale Welten können Teil einer Organisation sein, genauso wie mehrere Organisationen sich unter ihnen subsumieren oder wie mehrere Welten sich an mehreren Organisationen beteiligen können.
Eine weitere Eigenschaft sozialer Welten ist, dass sie fluide Grenzen besitzen. „It is a universe where fragmentation, splintering, and disappearance are the mirror images of appearance, emergence, and coalescence. This is a universe where nothing is strictly determined" (Strauss 1978, S. 123). Aus zwei konkurrierenden Welten kann eine neue Welt entstehen. Eine Welt kann in viele kleine Segmente zerfallen oder eine Welt verschwindet ganz von der Oberfläche (vgl. Clarke 1991, S. 133). In den jeweiligen Innovationstypen können verschiedene Organisationen und soziale Welten rekonstruiert werden, die in den Innovationen involviert sind.

Akteurskonstellation

Die Analyse zeigt über alle Typen hinweg heterogene beteiligte Akteursgruppen. Analytisch bedeutsamer als welche Akteure teilnehmen, ist jedoch die temporale Dynamik: Der Zeitpunkt und der Modus, in dem spezifische Akteursgruppen in das Innovationsgeschehen eintreten, variieren systematisch.
Die Akteurskonstellation der **Pragmatischen Modernisierung** (Typ A) ist durch eine **formale Interdependenz** gekennzeichnet, bei der Innovation durch etablierte Kanäle und organisationsinterne Strukturen prozessiert wird. Zentral sind hier in der Regel ständige Einrichtungen wie das „[Name des Mitarbeiterforums]" (ID9, Pos.41), in

dem fachbereichsübergreifend Mitarbeitende zusammenkommen, um über digitale Themen zu lernen und Bedarfe der Praxis zu benennen, der „Leitungskreis“ aus Geschäftsführung und Vorstand (ID4, Pos.27), oder institutionalisierte Austauschplattformen wie der „Arbeitskreis zu KI“ (ID2, Pos.108), der einen informellen Austausch zu aktuellen Entwicklungen ermöglicht. In diesen Beispielen sind es also bestehende Strukturen, die als Nukleus des Innovationsvorhabens betrachtet werden können. Ausgehend davon müssen für die Innovation andere innerorganisationale Gatekeeper gewonnen werden, die ein Veto-Recht geltend machen können. Geschäftsführung und Vorstand müssen Investitionen „abnicken“ (ID4, Pos.27), die Einführung ist „sehr stark [...] abhängig“ vom Betriebsrat (ID9, Pos.25), Datenschutzstellen und interne IT müssen rechtliche und technische Details prüfen und die Unbedenklichkeit bestätigen (vgl. ID6, Pos.14), bevor die Erprobung beginnen kann.

Anders stellt sich dies im Falle der **Reflexiven Sozialinnovation** (Typ B) dar: Die zentralen Strukturen existieren selten a priori, sondern werden im Prozess der Zusammenarbeit mit Partnern für das Projekt neu konstituiert. Die befragten Personen sind hier oft Teil von spezialisierten Teams (vgl. ID10, Pos.24), die qua Amt Ansatzpunkt für Innovation sind (durch die Einwerbung von Fördermitteln; durch direkte Ansprache) und die als Broker zwischen der Organisation und externen Akteuren (z.B. Hochschulen, vgl. ID11, Pos.73; Institute, ID7, Pos.65 oder Dienstleister, ID10, Pos.24) agieren. Die Akteurskonstellation ist hybrid und überschreitet Organisationsgrenzen. Exemplarisch für diese prozessuale Strukturbildung steht ein Gremium, welches spezifisch für das Innovationsvorhaben eingesetzt wurde und in dem „mehr als 30 Leute“ aus Haupt- und Ehrenamt versammelt sind (ebd.). Es ist kein vorbestehendes Entscheidungsgremium, sondern ein temporärer Resonanzraum, der sich erst durch das Innovationsvorhaben mit Leben füllt. Die Konstellation ist fragil, da sie von der Kooperation externer Dienstleister abhängt, die teils tief in interne Prozesse eingreifen (vgl. ID11, Pos.93) und deren Ausfall („Partner verweigert die Arbeit“, ID1, 1 Pos.121) die gesamte Konstellation gefährdet. Es wird diesbezüglich der Bedarf nach „vertrauenswürdigen Partnern, die nicht einfach nur Dienstleister sind und bei denen wir nicht in eine Abhängigkeit geraten“ (ID 12, Pos.98) betont. Optimalerweise sollen diese Entwicklungspartner ein Interesse daran haben, wie die Wohlfahrt „soziale Herausforderungen anzugehen“ (ebd.).

Typ C (**Experimentell-disruptive Innovation)** wiederum bewegt sich vorwiegend in einem strukturellen Vakuum. Hier gibt es keine vorgeschalteten Gremien, die die Innovation steuern. Die „**eigenmächtigen Avantgardist:innen**“ agieren weitgehend entkoppelt: „Ich habe den Entscheider zwar direkt über mir [aber] ist schon ganz gut, dass ich nicht noch andere Gremien dazwischen habe“ (ID8, Pos.78). Innerhalb ihres

Verantwortungsbereichs haben die Avantgardist:innen einen hohen Gestaltungsspielraum, der es ihnen erlaubt, ihre „Traumprojekte“ (ID8, Pos.74) weitestgehend eigenmächtig voranzutreiben, wobei organisationale ‘Watchdogs’ wie die IT-Abteilung nur in geringem Maße in den Innovationsprozess eingebunden sind: „mit der habe ich nichts zu tun [...] nur wenn ich zum Beispiel wissen muss, was haben wir für eine technische Infrastruktur“ (ID8, Pos.10). In ihrer Selbstbeschreibung definiert eine projektverantwortliche Person ihren Auftrag so: „[...] unser Ziel ist es, Zukunft zu schaffen.“ Sie hebt KI als ihr persönliches „Treiberthema“ hervor und konstatiert, dass das Projekt „im Prinzip mit [ihrem] Interesse gematcht“ habe (ID1, Pos. 7). Flankiert wird dies durch eine „Koalition der Willigen“ (ID1, Pos.21), die sich ad-hoc aus „Early Adoptern, die ganz viel machen und auch Sachen machen, die vielleicht gar nicht geregelt sind, weil noch gar nichts geregelt ist“ (ID10, Pos.20)“ bildet und die einen Proof-of-Concept mitverantworten. Erst nachdem diese experimentelle Innovation („eigenes KI-System gebaut“, ID1, Pos.14) erfolgreich war, beginnt die Strukturbildung, über die das Produkt an einen größeren Nutzer:innenkreis ausgerollt wird. Die formale Akteurskonstellation steht hier also am Ende, nicht am Anfang des Prozesses.

Ressourcen

Über die ausgewerteten Interviews hinweg konnte keine typenbezogene Unterschiede in Fragen der Finanzierungsmodelle für die Innovationsvorhaben festgestellt werden. Vielmehr kann von einem heterogenen Finanzierungsmix aus Förder- und Eigenmitteln für die KI-Projekte gesprochen werden. Knapp die Hälfte der Vorhaben ist zumindest in der Anschubfinanzierung über Eigenmittel finanziert, was für eine grundsätzliche Investitionsbereitschaft in KI-Technologien steht. Jedoch divergieren die Typen im Hinblick auf den Legitimationsdruck bezüglich des Mitteleinsatzes.

Was die eingesetzten organisationseigenen Personalressourcen angeht, werden kleinere Projekte häufig in Personalunion und integriert in das alltägliche Aufgabenspektrum der Befragten umgesetzt, während sich umfangreichere Vorhaben durch die Arbeit in multiprofessionellen Teams auszeichnen, die auch nicht ausschließlich die Aufgabe des KI-Projekts haben. Diese Teams bestehen beispielsweise aus Governance-, Finanzierungs- und Projektmanagement-Expert:innen. Entwickler:innen im Sinne von Personen, die ein KI-System technisch programmieren, sind in keinem der Projekte angestellt.

Im Sinne einer **Pragmatischen Modernisierung** greift Typ A in der Adaption marktgängiger KI-Anwendungen auf **organisationsinterne Ressourcen** zurück. KI-Vorhaben werden durch „interne Förderung“ angeschoben (ID5, Pos.38) oder es

werden Stellen geschaffen, die „nicht über Projektmittel finanziert sind“ (ID6, Pos.26). Doch nicht nur finanzielle und personelle Ressourcen werden eingesetzt. Bestehende Gremien werden als Erprobungsräume für die Innovation genutzt, datenschutzrechtliche und technische Prüfverfahren werden von den bestehenden Strukturen umgesetzt. Diese institutionelle Einbettung schafft langfristige Planungssicherheit und begünstigt die Beschaffung standardisierter Marktlösungen (ID7, Pos.44; ID12, Pos.20). Damit ist die Innovation eng mit der Gesamtorganisation verbunden. In diesem Kontext betonen die Befragten zudem die Kopplung der KI-Projekte an interne Weiterbildungsprogramme, durch welche den Mitarbeitenden begleitend Basiskompetenzen für den KI-Einsatz vermittelt werden (ID7, Pos.44).

Die für die Umsetzung der **Reflexiven Sozialinnovation** (Typ B) notwendigen **projekt- und netzwerkbasierten Ressourcen** setzen sich aus einer komplementären Verbindung von verbandlichem Domänenwissen und externer technologischer Expertise zusammen. Während die finanzielle Grundausstattung primär durch Drittmittel und seltener durch Eigenmittel gedeckt wird (ID8, Pos. 42), bringen die Wohlfahrtsverbände vor allem fachlich-inhaltliche Ressourcen ein. Hierzu zählt beispielsweise der Zugriff auf spezialisierte Fachreferate, welche die Qualität der Trainingsdaten durch eine inhaltliche Validierung sicherstellen (ID10, Pos. 14). Die Relevanz von guten Trainingsdaten wird dabei besonders benannt: „[...] wenn wir das irgendwann mal maschinell auswerten wollen, dann brauchen wir gute Daten“ (ID12, Pos. 27). Viele Wohlfahrtsorganisationen könnten die „geforderte Menge und Qualität nicht bereitstellen“, da zentrale Datensysteme über Jahrzehnte nicht aufgebaut wurden (ID12, Pos. 16). Auch andere Projekte benennen die Aktualität der Datengrundlage als kritische Ressource („wie geht man mit veralteten Daten um“, ID11, Pos.117). Die technischen Produktionsressourcen werden hingegen weitgehend von den Entwicklungspartnern bereitgestellt. Dies umfasst die notwendige Programmier-Expertise sowie die physische Infrastruktur wie Serverkapazitäten und Plattformen (ID10, Pos. 24).

In den **Individuell geprägten Sozialen Welten** der **Experimentell-disruptiven Innovation** (Typ C) nehmen persönliche Ressourcen eine zentrale Funktion ein. Die Projektverantwortlichen bringen eine hohe intrinsische Motivation mit und beschäftigen sich auch über ihre Tätigkeit hinaus vertiefend mit aktuellen Entwicklungen im Bereich KI: „[...] das ist nicht im Projekt entstanden, sondern aus der persönlichen Leidenschaft...“ (ID1, Pos.32). Wenn externe Partner zu langsam arbeiten, übernehmen sie technische Aufgaben selbst („dann mache ich die selbst“, ID8, Pos. 8), und sichern damit kontinuierliche Fortschritte. Dabei wird KI selbst als ermächtigende Ressource hervorgehoben, indem sie als Werkzeug beschrieben wird, welches fehlende technische Kenntnisse kompensiert: „[...] ich bin kein

Programmierer, also ich kenne HTML, CSS [...] aber nicht in der Tiefe [...] aber mit KI kann ich das“ (ID1, Pos.38).

Dynamik & Konflikte

In der Auseinandersetzung mit Dynamiken und Konflikten in den Sozialen Welten wird eine „Konkurrenz“ innerhalb der Wohlfahrt hervorgehoben, die eine synergetische Zusammenarbeit und das Teilen von Best Practices grundsätzlich erschwert. Kooperation erfolgt vielfach auf „Goodwill“, weil Konkurrenzlogiken die Zusammenarbeit beeinträchtigen (ID12, Pos. 38; ID10 Pos.10). Dieser als „Koopkurrenz“ (Schönig 2015) bekannte Zustand beschreibt das wohlfahrtsspezifische Spannungsfeld zwischen einem geteilten normativen Ziel und dem faktischen Wettbewerb um Ressourcen, Reputation und abrechenbare Fälle.

In den **organisationsinternen Strukturen** der **pragmatischen Modernisierung** (Typ A) entsteht ein Spannungsfeld zwischen der Geschwindigkeit technologischer Entwicklungen und der Dauer interner Abstimmungsprozesse. Die notwendige Einbindung von Gremien wie dem Betriebsrat (ID9, Pos.25) oder Leitungskreisen bei der Einführung einer eingekauften Lösung führt zu verlängerten Entscheidungszyklen, da Beschlüsse hierarchisch legitimiert werden müssen. Durch die Berücksichtigung der formellen organisationsinternen Abstimmungswege wird „Sicherheit durch Verfahren“ gewährleistet. Gleichzeitig erzeugen die laufenden Kosten für standardisierte Softwarelösungen einen Rechtfertigungsdruck gegenüber der Geschäftsführung („Warum ist das so teuer?“).

Charakteristisch für die **Reflexive Sozialinnovation** (Typ B) ist die Zusammenarbeit mit externen Akteuren zur Umsetzung des Innovationsvorhabens und damit einhergehende Rollenkonflikte. Die gemeinnützige Logik der Wohlfahrtsverbände trifft auf die kommerzielle „Businessorientierung“ der entwickelnden Dienstleister, die primär „Geld machen“ wollen (ID1, Pos. 30). Dies manifestiert sich in Reibungsverlusten, etwa durch intransparente Kostenstrukturen, die erst spät offenbar werden (ID8, Pos. 52) sowie durch eine herausfordernde Passung von Geschäftsmodell und Wohlfahrtstruktur: Während kommerzielle Anbieter Lizenzen an Einzelnutzer:innen verkaufen wollen, möchte der Verband eine plattformisierte Zentrallösung für seine Mitglieder bereitstellen, was „überhaupt nicht der Business Case von dem Anbieter“ ist (ID3, Pos.47). Dem gegenüber wird die Zusammenarbeit mit wissenschaftlichen Akteuren eher als vertrauensbildend beschrieben, da diese die technische Machbarkeit unabhängig validieren (vgl. ID4, Pos. 55; ID1, 1 Pos.139). Unabhängig vom Entwicklungspartner besteht die Gefahr einer Projektfalle: Wurde im Projekt keine Refinanzierungsmöglichkeit gefunden, kann dies dazu führen, dass erfolgreiche Pilotprojekte stranden, weil der Transfer in den Regelbetrieb nicht in die

gängigen „Förderschablonen“ passt, da Wartung und Betrieb darüber nicht abgebildet werden. Insofern stünden sich die „Förderkulissen selbst im Weg“ (ID7, Pos. 8), wie ein Befragter hervorhebt. Das Resultat ist, dass vielversprechende Ansätze nach der Projektlaufzeit ruhen müssen: „Das ist leider vorbei. […] konnte nicht umgesetzt werden“ (ID7, Pos. 22). Oder noch drastischer, wie es eine andere Person beschreibt: „Wenn die Förderung ausläuft, ist die Idee tot.“ (ID1, Pos.27)

Bei der **Experimentell-disruptiven Innovation** (Typ C) folgt die Projektdynamik dem Prinzip der **Legitimation durch Funktion**. Neue Funktionen werden ohne langfristige Problemdefinition entwickelt, ihre sich im Betrieb erweisende Nützlichkeit schafft Fakten, die eine nachträgliche Akzeptanz begünstigen. Ist die Nützlichkeit nicht gegeben, wird auf der Grundlage von Feedback der „Early Adopter“ iterativ eine andere, passende Lösung erarbeitet. Durch ihren Erfolg erfahren die Anwendungen durch interne Nachfrage eine Ausweitung, die der Akteur jedoch nicht als Belastung empfindet. Vielmehr führen die steigenden Anforderungen und Nachfragen zu Überlegungen, die Idee „mit einer Vision dahinter“ (ID1, Pos.32) zu skalieren und in ein eigenes Geschäftsmodell zu überführen. Der Avantgardist befindet sich in seinen Vorhaben in einem anhaltenden Ringen nach Handlungs- und Gestaltungsspielräumen „nach dem Motto: Hier hast du eine Summe, überschreite die nicht und das sind die Sachen, die du alle selber gestalten kannst“ (ID8, Pos.74).

7.4 Innovationsprozess

Innovationsprozesse werden in der Fachliteratur häufig als Abfolge mehrerer Phasen dargestellt, um den Ablauf von der Ideenfindung bis zur Markteinführung strukturiert zu beschreiben. Ein etabliertes Phasenmodell stammt von Trommsdorff (1995), der insgesamt sechs idealtypische Prozessphasen (in drei Blöcken) unterscheidet. Das Modell soll als Heuristik dienen, um komplexe Innovationsaufgaben zu strukturieren, wobei kontinuierliche Analyse- und Testschleifen über alle Phasen hinweg frühzeitig Entscheidungshilfen für Fortführung oder Abbruch von Innovationsprojekten liefern sollen.

Der *erste Block, Ideengenerierung*, beginnt mit der *Phase der Situationsanalyse*, in der das Umfeld und mögliche Problemfelder untersucht werden. Daran schließt die *Phase der Ideengewinnung* an, bei der systematisch neue Ideen entwickelt werden.

Der *zweite Block, Ideenakzeptierung*, umfasst die *Phase der Ideenselektion und -bewertung* sowie die anschließende *Phase der Strategiefindung* (Rüggeberg und Burmeister 2008). In dieser Phase werden vielversprechende Ideen mittels Screenings identifiziert und evaluiert, gefolgt von strategischen Entscheidungen zur Ausrichtung der Innovation (z. B. Festlegung von Zielmärkten, Produktkonzept und benötigten Ressourcen).

Der *dritte Block, Ideenrealisierung*, beinhaltet Entwicklung und Markteinführung der Innovation. In der *Entwicklungsphase* wird das Konzept technisch umgesetzt, begleitet von laufenden Tests (Prototypen- und Markttests), um Machbarkeit und Marktakzeptanz zu prüfen. Zum Abschluss erfolgt die *Phase der Einführung/Durchsetzung* des neuen Produkts am Markt, also die Kommerzialisierung der Innovation

Der als Phasenmodell dargestellte Innovationsprozess bildet stets einen idealtypischen Verlauf ab, der sich in der Praxis aufgrund hoher Komplexität, Variabilität sowie ausgeprägter Dynamik und Unsicherheit nur selten in dieser idealtypischen Form zeigt. In realen Innovationsprozessen überschneiden sich Phasen häufig, verlaufen iterativ und sind nicht eindeutig voneinander trennbar, was zu erheblichen Einschränkungen bei der Übertragbarkeit linearer Modelle führt. Zudem wird die in der Praxis angestrebte Parallelisierung von Prozessschritten zur Effizienzsteigerung in gängigen Phasenmodellen nicht adäquat berücksichtigt (vgl. Vahs und Brem 2015, S. 234).

Ungeachtet der Schwierigkeit, die linearen Phasenmodelle eins-zu-eins auf die Empirie zu übertragen, wollen wir die Phasen als Heuristik verwenden. Wie die einzelnen Phasen sich in den drei Typiken darstellen, werden wir im Folgenden darstellen.

Ideengenerierung

Im Falle der **Pragmatischen Modernisierung** (Typ A) entsteht die Idee zur Innovation als **Reaktion auf ein wahrgenommenes Funktionsdefizit**. Beispiele hierfür sind Prozesse (z. B. Papier-Handbücher), die als „anachronistisch" empfunden werden (ID6, Pos. 4) oder stark repetitive Tätigkeiten (z.B. Dokumentation, ID5, Pos.38). Dabei wirken die neuen Möglichkeiten von KI als externer Schock, also als ein unvorhergesehenes Ereignis von außen, das die etablierten Strukturen schlagartig aufbricht und an dem die Ideengenerierung ansetzt. Bisherige Ideenfindungsprozesse waren z.B. wegen hoher Kosten einer digitalen Lösung ins Stocken geraten („Wir hatten auch keine Idee mehr [...] zum Glück kam dann [LLM-Modell] und wir haben doch noch eine Idee entwickelt", ID6, Pos. 30).

Beim Typus der **Reflexiven Sozialinnovation** (Typ B) lassen sich bei den Entwicklungspartnerschaften zwei grundlegende Modi der **chancenbasierten Ideengenerierung** identifizieren. Zum einen werden Projektideen von außen an die Wohlfahrtsverbände herangetragen. Dies geschieht beispielsweise durch Beratungsunternehmen („Wir werden ein Projekt umsetzen, was uns [IT-Systemhaus] angeboten hat", ID9, Pos.75) oder durch wissenschaftliche Akteure, wobei oft vorangegangene Kooperationen als Auslöser fungieren („Wir hatten eine Studie mit [Hochschule] gehabt und so ist der Kontakt entstanden", ID7, Pos. 16). Zum anderen erfolgt der Anstoß endogen durch die Organisationen selbst, sofern sie über eigene Expertise-Strukturen verfügen oder dies als ihren expliziten Auftrag begreifen: „Unsere Aufgabe ist es, digitale Projekte zu finden [...] wo wir sagen, die haben Potenzial" (ID10, Pos. 6). Ideen können auch als Eigenentwicklungen beginnen („Ich war schon dabei, eine Plattform selber zu bauen", ID3, Pos.26) und durch das Auftreten des richtigen Entwicklungspartners in eine Entwicklungspartnerschaft übergehen. Ungeachtet des Ursprungs werden die Konzepte strategisch auf ihre Förderfähigkeit hin optimiert. Die Orientierung an Förderlinien ist zentral, wie ein Befragter betont: „unsere ganze Projektidee war eigentlich, ein Konzept zu erschaffen für weitere Förderungen", ID1, 1, Pos. 95).

Wie im Verlauf des Berichts bereits deutlich geworden ist, sind bei der **Experimentell-disruptiven Innovation** (Typ C) einzelne **neugierige Pionier:innen** wesentlich für die Ideengenerierung verantwortlich („Vision meiner Person", ID1, Pos.30). Sie sind „ChatGPT-Nutzer der ersten Stunde" (ID8, Pos.10), die in Form von „Experimenten" private Erfahrungen mit KI-Technologien gemacht haben und dabei „im Grunde genommen unendlich viele Möglichkeiten [gesehen]" (ID8, Pos.19) haben. Eigenmächtig oder in kleinen Allianzen ([Der Geschäftsführer] wollte unbedingt ein erstes KI-Projekt haben", ID8, Pos.39) entwickeln sie diese mit hoher intrinsischer Motivation.

Ideenakzeptierung

Typenunabhängig ist für die Akzeptanz von Innovationen entscheidend “wer der Datenschutzbeauftragte ist [und] wer da in der Geschäftsführung ist, wie vorsichtig die sind“ (ID11 Pos.49). Damit werden diese Entscheider:innen als zentrale Voraussetzung für einen erfolgreichen Projektstart hervorgehoben.

Die Akzeptanz der generierten Idee wird im Typus der **Pragmatischen Modernisierung** (Typ A) maßgeblich durch die niederschwellige Verfügbarkeit von KI-Tools gefördert. Die Integration kostenloser KI-Funktionen in bereits bestehende Software-Suiten, beispielsweise die Erweiterung von MS365 um Microsofts Copilot, ermöglicht den Mitarbeitenden eine erste Exploration der Technologie. Während einige Organisationen zu deren Einhegung zunächst Compliance-Richtlinien erarbeiten (vgl. ID4, Pos.10), gehen andere einen liberaleren Weg und „haben erstmal beschlossen, dass jeder Mitarbeitende prinzipiell KI nutzen kann“ (ID9, Pos.41). Nach dieser ersten Erprobungsphase setzt Typ A auf die Adaption von „Stangenlösungen“, „individuelle Appentwicklung möchte man nicht, das hält man für zu risikoreich, zu kostenintensiv“ (ID6, Pos. 32). Auch werden hier große technologische Fortschritte erwartet, denen eine Eigenentwicklung möglicherweise nicht standhält: „Wenn wir was entwickeln und dann haben wir drei Jahre lang was Schickes und dann nichts mehr“ (ID6, Pos. 30). Ausgehend von dieser Entscheidung folgt ein **formalisierter Auswahlprozess**, bei dem verschiedene Lösungen anhand von Kriterien wie Datenschutz, Kosten, Skalierbarkeit und technischer Stabilität ausgesucht, erprobt und abgewogen werden. Der Typus A verfolgt damit eine regulierte, risikoabsichernde Strategie, in die Organisationseinheiten wie IT, Datenschutz und Betriebsrat als Voraussetzung für Akzeptanz innerhalb der Organisation früh eingebunden werden.

Die Innovationsvorhaben der **Reflexive Sozialinnovation** (Typ B) werden als auf Zeit angelegte Projekte in den Organisationen verankert. Die **kooperative Aushandlung** von technischen Details ermöglicht eine höhere Steuerbarkeit, die sich beispielsweise in einer kontrollierten Datenverarbeitung äußert. Die Nutzung einer abgeschirmten Cloud bzw. Datenbank stellt sicher, dass „alle Daten in dieser Cloud bleiben“ und nicht zu Trainingszwecken verwendet werden (ID10, Pos. 24). Besonders hervorgehoben wird die Möglichkeit, organisationsinterne Daten aus verschiedenen Cloudumgebungen zu verknüpfen, sodass KI „dort sucht statt im ganzen Web“ (ID4, Pos. 57). Dies schafft Vertrauen und erlaubt die Arbeit mit sensiblen personenbezogenen Daten, was als strategischer „Hebel“ beschrieben wird (ID3, Pos. 18).

Bei der **Experimentell-disruptiven Innovation** (Typ C) entsteht die Akzeptanz der Innovation nicht durch langwierige Planung, sondern durch **Diffusion durch Vorreiter:innen**: Erfolgreiche Experimente einzelner Akteure erzeugen einen viralen

Sog, der neue Bedarfe konstituiert („Dann kamen im Prinzip die Kollegen auf uns zu und haben gesagt, Das und Das brauchen wir auch“, ID1, Pos. 16).

Ideenrealisierung

Die untersuchten Projekte befinden sich, mit einer Ausnahme, derzeit noch in einer frühen Implementierungsphase. Zur langfristigen Akzeptanz und Diffusion dieser Innovationen in der Wohlfahrt lassen sich daher keine Aussagen treffen. Dennoch müssen die involvierten Akteure langfristige strategische und inhaltliche Fragestellungen klären, um die Weiterentwicklung zu gewährleisten.

Die Phase der Realisierung ist im Modus der **Pragmatischen Modernisierung** (Typ A) im Gegensatz zu projektförmigen oder experimentellen Vorhaben auf eine langfristige Erweiterung des Softwareportfolios ausgerichtet. Die Anwendung muss daher spezifische Qualitätsanforderungen mit „Blick auf Nachhaltigkeit“ (ID6, Pos.30) auch im laufenden Betrieb erfüllen und so einen **Return on Investment** bieten. Die Finanzierung aus Eigenmitteln erfolgt langfristig nur, wenn die Innovation „betriebswirtschaftlich Sinn macht“ (ID5, Pos.50). Um diese Passung zu sichern, wird auch der Wechsel zwischen vergleichbaren Anbietern benannt. Dass Projekte auf diesem Weg stocken, liegt auch an einer Infrastruktur-Lücke: Die langfristig angedachte Software trifft auf eine noch weitgehend ‘analoge’ Realität, in der Einrichtungen „drei Schritte in einem machen müssen“ (ID5, Pos. 62), da eine grundlegende Digitalisierung noch nicht flächendeckend erfolgt ist: „[...] wir haben da immer noch die Kollegen, die mit Faxgerät und Stift und Zettel da sitzen“ (ID1, 2, Pos. 22).

In der langfristigen Realisierung der Idee stehen Projekte der **Reflexiven Sozialinnovation** (Typ B) in stetiger Gefahr einer **Projekt-Falle**. Da die Entwicklung der Innovation stark an die Förderlogik und ihr Betrieb an die Förderlaufzeit gebunden ist, ist der Transfer in den Regelbetrieb („Dauerbetrieb“) ein kompliziertes Unterfangen. Projekte stagnieren im Status des „Testmodells“ (ID1, 1, Pos. 93) oder ruhen vollständig, sobald die externe Finanzierung endet („seit anderthalb Jahren ungefähr ruht das aber“, ID7, Pos. 22). Dies führt dazu, dass Projekte nach Ablauf der Förderung oft „im Keller verschwinden“ (ID11, Pos. 139) oder „eingestampft“ werden (ID10, Pos. 48).

Die **Experimentell-disruptive Innovation (Typ C)** arbeitet in der Realisierungsphase an der Genese einer Lösung, die sich von dem zunächst experimentellen Status emanzipiert. Die bisherigen Eigenentwicklungen sollen als robuste Stand-Alone-Lösung ausgebaut werden. Die Befragten artikulieren den Wunsch nach einer **technologischen Souveränität**: Es soll etwas sein, das „spezieller gebaut [ist] und nicht das Fertigding, was ich mir in die Cloud setze“ (ID8, Pos. 74). Realisierung bedeutet hier den Aufbau einer eigenen Infrastruktur („unternehmensinterne KI“),

die dauerhafte Kontrolle ermöglicht („Ich kann das Ding abschalten“, ID8, Pos. 93) und Datenschutzstandards jenseits der großen Anbieter gewährleistet.

8 Inhaltsanalytische Auswertung der Fokusgruppen

Neben den Interviews mit den Protagonisten der Innovationsprojekte haben wir im Zuge der wissenschaftlichen Evaluation auch mit Fachkräften der Sozialen Arbeit im Rahmen einer Fokusgruppe gesprochen. Zentrale Forschungsfrage der Fokusgruppe war, wie KI semantisch aufgeladen wird, welche Neuheit der KI prospektiv zugeschrieben wird, und wo Grenzen der KI-Nutzung gezogen werden. Zudem haben wir nach der formellen und informellen Nutzung bei den jeweiligen Einrichtungen und Trägern gefragt.

Zum ersten zeigt sich in dem transkribierten Fokusgruppenmaterial, dass die Innovationssemantik eine dichotome Struktur aufweist: Die KI wird in administrativen Belangen als Entlastung konstruiert, während es in der direkten Dienstleistungserbringung als Entkernung der Professionalität gedeutet wird. Zum zweiten zeigt sich in dem Material eine deutliche Diskrepanz zwischen dem theoretischen Diskurs und den Deutungen der Fachkräfte über die formelle wie informelle Nutzung in der sozialarbeiterischen Praxis. Zunächst soll auf die formelle und informelle Nutzung eingegangen werden, bevor auf die Zukunftsbilder der KI-Nutzung innerhalb der Fokusgruppe eingegangen wird.

8.1 Formelle und informelle Nutzung von KI

Die Analyse des Datenmaterials aus der Fokusgruppe zeigt ein signifikantes Spannungsfeld in der Nutzung von KI im Kontext der Sozialen Arbeit. Unter dem Analysefokus der formellen versus informellen Nutzung lässt sich eine Dichotomisierung rekonstruieren: Während die formelle Implementierung von KI durch strukturelle Defizite und datenschutzrechtliche Hürden aus Sicht der Diskutanten stagniert, diffundieren informelle Nutzungspraktiken aus dem privaten Alltag in den professionellen Kontext, welcher wiederrum vor den Gefahren des KI-Einflusses geschützt werden soll.

Auf der formellen Ebene ist die Nutzung von KI weitgehend abwesend. Die Diskutanten verorten ihr Arbeitsfeld metaphorisch am „Ende der Nahrungskette“ (Pos. 16). Es herrscht ein Konsens darüber, dass die Digitalisierung und der Einsatz von KI im eigenen Träger hinter den notwendigen Standards zurückbleibt („nicht so weit, wie wir eigentlich hätten sein müssen“, Pos. 16). Formelle Nutzungsszenarien – also durch die Organisation bereitgestellte, rechtssichere KI-Tools und klare Handlungsanweisungen – werden von den Diskutanten mit Blick auf die Praxis in den eigenen Einrichtungen und Trägern nicht identifiziert. Stattdessen wird der Status quo durch ein kollektives „Noch-Nicht“ (Pos. 14-15) beschrieben. Von allen Diskutanten

wird geteilt, dass die KI nicht einmal thematisch im Träger in Erscheinung getreten ist.
Die Leerstelle der formellen Nutzung wird diskursiv primär durch den Datenschutz legitimiert. Der Datenschutz fungiert aus Sicht der Diskutanten als „Gatekeeper“, der Innovationen ausbremst. Die Unsicherheit über Serverstandorte („Firmen aus dem Ausland“, Pos. 16) und die Sorge um die Integrität sensibler Adressat:innendaten führen zu einer organisationalen Paralyse. Sobald rechtliche Fragen auftauchen, bricht der Diskurs über Möglichkeiten ab („Dann verstummen die halt so“, Pos. 37). Aus Sicht der Diskutanten ist diese Defensivhaltung der Träger durchaus rational begründet, denn die Daten der Adressat:innen sind äußerst sensibel, was gleichzeitig aber dazu führt, dass keine offiziellen Experimentierräume entstehen.
In dieses organisationale Vakuum dringen informelle Praktiken ein. Die Analyse zeigt, dass KI-Nutzung nicht top-down implementiert, sondern bottom-up durch individuelle Akteure in die Organisation getragen wird. Dies wird von den Diskutanten als generationaler Habitus gerahmt: Für die „junge Generation“ (Pos. 46) ist die Nutzung von Tools wie ChatGPT eine selbstverständliche Alltagspraxis, die nahtlos auf berufliche Aufgaben übertragen wird. Die Übertragung aus dem Alltag auf die berufliche Tätigkeit wird von Diskutanten kritisch bewertet. Die informelle Nutzung geschieht im Affekt einer „Euphorie“ und schafft vollendete Tatsachen (Pos. 47). Erst nachträglich entsteht ein Bewusstsein für den Regelverstoß („Oh Gott, jetzt hast du da personenbezogene Daten [...] eingegeben“). Die informelle Nutzung von KI in der Berufspraxis stößt bei den Diskutanten auf Ablehnung, da zu schnelle Umsetzungen den Datenschutz aus ihrer Sicht zu wenig beachten.
In diesem Sinne verhindert die formelle Abstinenz von KI-Strategien die Nutzung nicht, sondern verlagert sie in den informellen, unregulierten Bereich. Dies birgt erhebliche Risiken. Solange Soziale Dienste keine formellen Angebote machen, bleibt die KI-Nutzung den einzelnen Fachkräften überlassen, bei der Datenschutzverstöße eher akzidentiell passieren als mutwillig herbeigeführt werden.

8.2 Zukunftsbilder in der Fokusgruppe: Zwischen organisational Neuem und professioneller Bedrohung

Die vorliegende Analyse widmet sich der Rekonstruktion von Innovationssemantiken innerhalb einer Fokusgruppe von Fachkräften der Sozialen Arbeit. Als theoretische Folie dient hierbei das eingangs entwickelte Verständnis von Innovation (s.o.). Innovation wird hier, in Anlehnung an Garud und Karnøe (2001), als „mindful deviation“, also eine achtsame, bewusste Abweichung, verstanden, welche sich nicht abstrakt vollzieht, sondern sich um ein konkretes Artefakt, in unserem Fall im weiten Sinne die Künstliche Intelligenz, zentriert. Wie genau diese Neuheit gedeutet wird,

dazu gibt es zwischen der Fokusgruppe und den Interviews teils deutliche Differenzen.

i. KI als administrative Entlastung

In der Fokusgruppe haben die Teilnehmenden die KI zunächst an der organisationalen Peripherie der Sozialen Arbeit verortet. Die Fachkräfte konstruieren eine klare Trennung zwischen ihrer eigentlichen, professionellen Tätigkeit (Arbeit am Menschen) und den bürokratischen Notwendigkeiten, die als Störung empfunden werden. KI dient in diesen bürokratischen Nebentätigkeiten als Mittel zur Effizienzsteigerung. Die Innovation wird semantisch nicht als Umbruch verstanden, sondern als funktionales Werkzeug, das sich in bestehende Abläufe integriert, um diese zu beschleunigen.

> *„[...] das nimmt wirklich auch einen Teil Arbeitszeit und kann eine Entlastung sein [...] Ja, so die Zeitfresser im Alltag. Na ja, und da hänge ich aber auch bei den Mitarbeitern mit der Tagungsdokumentation zum Beispiel so Ich sehe, das ist viel Arbeit auch oder das frisst viel Zeit.“ (Pos. 14/17).*

Die Metapher des „Zeitfressers“ ist hier zentral. Das Bestehende wird als defizitär beschrieben, weil es Ressourcen bindet, die an anderer Stelle fehlen. KI wird als Lösung antizipiert, die diese Ressourcen freisetzt. Bereits hier lassen sich bemerkenswerte Parallelen zum frühen Digitalisierungsdiskurs der 1990er Jahre ziehen, wo der Personal Computer auf „einfachste Schreib-, Dokumentations-, Informations- und Entscheidungsarbeiten“ (Ehlert 1991, S. 41) reduziert wurde. Besonders prägnant wird dies im Material in der Vorstellung eines automatisierten Anfragemanagements, bei dem KI als *Gatekeeper* fungiert:

> *„[...] dass die Jugendämter auf diese Fragen dann quasi Antworten geben und dass dann eine Vorauswahl schon passiert, also dass die Anfragen schon direkt in einer Maske ist und dass man dann sagt Zack, zack, zack, zack“ (Pos. 39).*

Die Wendung *„Zack, zack, zack, zack“* verdeutlicht die Geschwindigkeit und Reibungslosigkeit, die der KI zugeschrieben wird. Hier zeigt sich die prospektive Neuheit: Statt die Anfragen manuell zu sichten, kann durch die KI ein Filter sein, durch den Fragen schneller beantwortet werden können. Interessant ist hierbei der Verweis auch auf andere Sektoren:

> *„[...] und die halt in der in der anderen Wirtschaft, in der Produktions und Produktionsbetrieben auch schon teilweise so betrieben wird“ (Pos. 39).*

Der Einsatz von KI wird legitimiert, indem eine Analogie zur Produktionswirtschaft hergestellt wird. Im Kontext der Administration wird die „Produktionslogik“ von KI

antizipiert und sogar herbeigesehnt, um *„lästige Randthemen“* (Pos. 55) loszuwerden. Die Innovationssemantik verspricht hier eine *Rückgewinnung von Spielräumen für die Profession* durch die Auslagerung des Unliebsamen an die KI. Der professionelle Tätigkeitsbereich lässt sich, so schreibt schon Ehlert 1991 in Bezug auf Computer im Allgemeinen und so wiederholen es die Fachkräfte in Hinblick auf KI fast nahtlos, hingegen nicht in die Logik von Computern zwängen (Ehlert 1991).

ii. Die Unersetzbarkeit der menschlichen Interaktion

Sobald die Innovationssemantik nicht mehr um den Bereich der Verwaltung kreist und sich auf die Interaktion mit Adressat:innen bezieht, verändert sich die Bewertung der Fachkräfte. Das prospektiv Neue stellt sich als Bedrohung und Verfall dar, nicht als Fortschritt. Hier wird KI nicht mehr als nützliche Weiterentwicklung gesehen, sondern als defizitäre *Simulation*. Die Fachkräfte betonen, dass Soziale Arbeit auf einer spezifischen Qualität der Beziehung beruht, die durch KI nicht reproduziert werden kann. Es wird eine ontologische Differenz zwischen Mensch und KI aufgemacht. Der Kern dieser Differenz liegt in der Authentizität der Emotion und der Resonanz.

> *„Wenn ich jetzt aber in der Beratung bin und da mit Humor und Menschen habe, die mitkriegen, der mir ein Lächeln schenkt und quasi kein generiertes Lächeln, weil der Algorithmus das sagt, sondern weil der Mensch gerade genau entscheidet. Ich lächel, weil ich das gerade mache. Das ist nicht zu. Das ist auch nicht zu ersetzen.“ (Pos. 73).*

Ein „generiertes Lächeln“ ist zwar durch die KI vorhanden, aber es fehlt die intentionale Tiefe des Menschlichen. KI wird als *„emotional [...] eingeschränkt“* (Pos. 84) beschrieben. Besonders deutlich wird die Abgrenzung gegenüber einer therapeutischen Innovation durch KI. Die Komplexität menschlicher Kommunikation wird als unzugänglich für Algorithmen markiert:

> *„Und Körpersprache, Ton und wenn einer redet und einmal überlegt [...] Da müsste dann müsste das kann ja nicht einfach so ein Chatbot machen“ (Pos. 63).*

Die Innovation wird hier als Gefahr gedeutet, vulnerable Zielgruppen (z.B. Menschen mit Bindungsstörungen) mit einem technologischen Artefakt zu konfrontieren, das per definitionem keine Bindung eingehen kann (*„die KI stellt keine Bindung her“*, Pos. 84). Hier fungiert die Innovationssemantik als Schutzmechanismus: Um die Qualität der Arbeit zu sichern, soll das Neue aus dem Kernbereich ferngehalten werden. Es wird betont: *„Sozialarbeit ist ja mit Menschen, das ist nichts Technisches“* (Pos. 50).

Sofern die KI aber in dem Kernbereich der Professionalität zum Einsatz kommt, konstruieren die Interviewten ein Bedrohungs- wenn nicht gar Verfallsszenario.

iii. De-Professionalisierung durch KI

Ein weiterer Strang der Analyse in Hinblick auf die KI als Innovation betrifft die Kompetenzen der Fachkräfte. Hier wird die Einführung von KI nicht nur als ethisch problematisch, sondern als Angriff auf die eigene Kompetenzstruktur gedeutet. Die Innovationssemantik wird mit einer *Verlustrechnung* verknüpft. Diese Verlustrechnung wird am Beispiel der Dokumentationspraxis durch die Konstruktion eines Vorher-Nachher-Vergleichs aufgemacht. Das ist insofern bemerkenswert, als dass die Dokumentation in den Interviews mit den Protagonisten der konkreten KI-Innovationen tendenziell als notwendiges Übel, welches es an die KI auszulagern gilt, gedeutet wurde. Aus Sicht der Fachkräfte in der Fokusgruppe wird das Schreiben von Berichten, Hilfeplänen oder Tagesdokumentationen jedoch nicht als bloße, organisationale Verpflichtung verstanden, die es auszulagern gilt, sondern als reflexive Praxis – als ein Prozess des Verstehens.

> *„Also ich finde, wenn man jetzt einen Hilfeplan vorbereitet oder eine Tagesdoku macht, wenn man das von KI machen lässt, dann nimmt man sich die Reflexionsmöglichkeiten des Schreibens und des Überlegens“ (Pos. 37).*

Wenn die KI diese Aufgabe übernimmt, so die Befürchtung, entfällt der Moment der reflexiven Durchdringung des sozialarbeiterischen Falles. Innovation führt hier paradoxerweise zu einer Verflachung der Dokumentation und, eo ipso, der Reflexivität (da es *„zu oberflächlich bleibt“*, Pos. 17). Das Konzept der „erlernten Hilflosigkeit“ (Pos. 17) wird explizit genannt. Dies deutet darauf hin, dass die Innovation als eine Form der Entmündigung wahrgenommen wird. Die Sorge ist, dass junge Fachkräfte oder auch erfahrene Mitarbeiter:innen verlernen, methodisch zu denken, wenn die KI ihnen die Ergebnisse liefert:

> *„Ja, also du hast ja irgendwie doch methodisches Handwerkszeug mitbekommen [...] Wo man dann manchmal gar nicht mehr so richtig drankommt.“ (Pos. 17).*

Hier grenzt sich die Innovationssemantik deutlich von der „Reform“ ab. Eine Reform würde Prozesse anpassen, aber die Innovation durch KI droht, das Sozialarbeiter:innen überflüssig zu machen oder zu degradieren. Wenn die Fachkraft sich selbst nur noch als *Bedienerin eines Algorithmus erlebt, verliert sie die Fähigkeit, Klientinnen* zur Autonomie zu befähigen. Dieser Aspekt unterscheidet die Soziale Arbeit von anderen Bereichen. Während ein Fehler in der Logistik („LKW falsch

beladen“) lediglich ökonomische Schäden verursacht, geht es in der Sozialen Arbeit um Existentielles.

> *„Sag mal so, wenn der LKW falsch beladen ist am Ende des Tages ärgerlich, aber da wird jetzt keiner von uns umkommen [...] Das ist bei uns anders, finde ich.“ (Pos. 45).*

Aufgrund dieser möglichen, existentiellen Folgen betonen die Befragten in der Fokusgruppe, dass die KI-Innovation verlangsamt werden sollte. Die Geschwindigkeit der Innovation („Hype“) wird kritisch gesehen. Es wird eine Verlangsamung gefordert, bei der der Mensch als letzte Instanz der „Vernunft“ fungiert:

> *„Machen wir aber nicht, weil wir, weil da dann der, der der Verstand und die Vernunft kommt.“ (Pos. 75).*

Hierin liegt ein zentrales Element: Innovation wird nicht per se abgelehnt, aber sie muss unter das Primat menschlicher Vernunft gestellt werden. KI sollte aus Sicht der Befragten niemals autonom agieren.

Die Argumentationsmuster der Fachkräfte gleichen frappierend jenen auf die Einführung des Personal Computers in den 1980er-Jahren. Eine zentrale Parallele ist die strikte Trennung von Administration und Beziehungsarbeit. Ehlert (1991: 41) konstatierte bereits damals, dass zwar „Zahlen und Berechnungen“ automatisierbar seien, sich der „professionelle Tätigkeitsbereich“ jedoch „nicht in die Logik von Computern zwängen“ lasse. Dies spiegelt die Haltung der Fokusgruppe wider, KI für administrative „Randthemen“ zu nutzen, sie aber vom sozialarbeiterischen Kern fernzuhalten.

Auch die Sorge vor De-Professionalisierung steht in einer historischen Tradition. Die heutige Angst vor dem Verlust der Reflexionsfähigkeit durch KI findet in den damaligen Warnungen vor „Berichten von der Stange“ (Lumma 1991, S. 92) und einer „Computerisierung des Denkens“ (Schurz 1991, S. 22) ihr historisches Äquivalent. Bezüglich der Interaktion betonte beispielsweise Mehlich (1997, S. 17) die „begrenzte Algorithmisierbarkeit“ sozialer Situationen, während Brinckmann (1993: 263) davor warnte, dass „der Kontakt zum Klienten“ erschwert werde.

Die KI-Skepsis ist somit keine neue Reaktion, sondern die Reaktivierung einer traditionsreichen Abwehrhaltung: Der Kern des professionellen Handelns wird erneut gegen eine drohende technologische Rationalisierung verteidigt.

9 Die Freie Wohlfahrtspflege als spezifischer Kontext für KI-Innovationen

Um die im vorliegenden Bericht dargestellten empirischen Befunde zum Einsatz Künstlicher Intelligenz in der Wohlfahrtspflege angemessen einordnen zu können, bedarf es einer Berücksichtigung der spezifischen Merkmale des untersuchten Feldes. Organisationen der Freien Wohlfahrtspflege verfügen über eine historisch gewachsene, sozialstaatlich eingebettete Sonderstellung, die in spezifischer Weise den Rahmen für technologische Innovationsprozesse absteckt. Die in den nachfolgenden Kapiteln herausgearbeiteten Studien, Ergebnisse und Innovationstypen müssen vor diesem Hintergrund gelesen werden.

9.1 Stellung der Freien Wohlfahrtspflege im deutschen Sozialstaat

Die Freie Wohlfahrtspflege nimmt im deutschen Sozialstaat eine Stellung ein, die sie von den meisten anderen Organisationstypen unterscheidet. Ihre Rolle ist historisch eng mit dem Subsidiaritätsprinzip verknüpft, das den Vorrang freier Träger gegenüber öffentlichen Trägern bei der Erbringung sozialer Dienstleistungen festschreibt. Die Verbände betreiben sehr heterogene Einrichtungen und Dienste, darunter u.a. Kindertagesstätten, Beratungsstellen, besondere Wohnformen, Pflegeeinrichtungen und Werkstätten für behinderte Menschen. Dieses breite Aufgabenspektrum bedingt eine besondere Komplexität. Anders als privatwirtschaftliche Unternehmen vereinen die Wohlfahrtsverbände mehrere Funktionen und Zielsetzungen in sich. Sie sind Erbringer sozialer Dienstleistungen, sozialpolitische Interessenvertretungen, Mitgliederorganisationen und wirtschaftliche Akteure zugleich. Grohs (2017, S. 26) weist darauf hin, dass insbesondere die Rollen als politische Interessenvertretung, als Anwalt vulnerabler Gruppen und als ökonomisch denkender Dienstleister in Widerspruch zueinander geraten (können). Evers und Ewert (2010, S. 106) beschreiben die Wohlfahrtsverbände daher als hybride Organisationen, die als intermediäre Sphäre eine enorme Pluralität an Orientierungen, organisatorischen Formen und Einfluss nehmenden Prinzipien vereinen.

Für den Gegenstand dieses Berichts ist diese Hybridität aus mehreren Gründen relevant. Erstens müssen KI-Innovationen in der Wohlfahrtspflege gleichzeitig unterschiedlichen Rationalitäten gerecht werden – ökonomischen Effizienzerwartungen ebenso wie fachlichen und ethischen Anforderungen der Sozialen Arbeit. Zweitens führt die Vielzahl der Aufgabenbereiche dazu, dass KI-Anwendungen in unterschiedlichen Kontexten zum Einsatz kommen, von der administrativen Dokumentation über die Beratung bis hin zur strategischen

Steuerung und dies wiederum jeweils in den unterschiedlichen Handlungsfeldern, wie etwa der Kinder- und Jugendhilfe, der Behindertenhilfe oder der Altenhilfe. Drittens bewegen sich die Verbände in einem Spannungsfeld zwischen der Arbeit mit hochsensiblen personenbezogenen Daten und dem Druck zur Effizienzsteigerung, was die Einführung datengetriebener Technologien besonders voraussetzungsreich macht.

9.2 Föderale Organisationsstruktur und ihre Implikationen für KI-Innovationen

Die Wohlfahrtsverbände sind föderal organisiert und in Bundes-, Landes- und Ortsverbände sowie Träger und Einrichtungen auf kommunaler Ebene gegliedert. Aus Perspektive der Bundesverbände, auf deren Ebene die Projekte des hier vorliegenden Förderkontextes angesiedelt sind, handelt es sich bei den Verbänden auf Landes- und Regionalebene um überwiegend unabhängige Organisationen. Das deutsche Wohlfahrtssystem funktioniert daher nicht nach einem einfachen Top-Down-Prinzip, sondern als Kombination aus Top-Down- und Bottom-Up-Dynamiken auf der Basis unterschiedlich stark ausgeprägter Kopplungen (vgl. Weick 1995; Ahrne & Brunsson 2008). Innerhalb der kommunalen Ebene, zwischen Trägern und ihren Einrichtungen, besteht eine vergleichsweise enge, weisungsgebundene Kopplung. Zwischen den föderalen Ebenen hingegen ist der Einfluss der Spitze auf die Mitgliedsorganisationen strukturell begrenzt (vgl. Nowoczyn 2017, S. 12).

Trotz struktureller Ähnlichkeiten unterscheiden sich die einzelnen Wohlfahrtsverbände bei näherer Betrachtung voneinander. Die Spitzenverbände divergieren untereinander ebenso wie ihre Mitgliedsorganisationen – in Größe, im Verhältnis der verschiedenen Ebenen zueinander und im Grad der bereits erreichten digitalen Reife. Aus ihrer Funktion als übergeordnete Organe ergeben sich für die Bundesverbände besondere Rollen, darunter die Qualitäts- und Strukturentwicklung der Organisationen, die Bereitstellung zentraler Dienstleistungen wie Daten- und Wissensmanagement und die Wahrnehmung sozialpolitischer Interessenvertretung (vgl. Grohs 2017).

Für KI-Innovationen hat diese Organisationsstruktur weitreichende Konsequenzen. Die lose Kopplung der Verbandsebenen erschwert die Durchsetzung einheitlicher KI-Strategien, verbindlicher Datenschutzstandards oder gemeinsamer Dateninfrastrukturen erheblich. Eine Bundesebene, die ein KI-gestütztes Dokumentationssystem entwickelt oder beschafft, kann dessen Nutzung auf lokaler Ebene nicht anordnen, sondern muss auf Überzeugung, Anreize und freiwillige Übernahme setzen. Die in Kapitel 9 dieses Berichts herausgearbeiteten Innovationstypen spiegeln diese Dynamik wider. Die Pragmatische Modernisierung

(Typ A) setzt auf marktverfügbare Standardlösungen, die sich vergleichsweise niedrigschwellig in bestehende Strukturen integrieren lassen. Die Reflexive Sozialinnovation (Typ B) erfordert demgegenüber kooperative Netzwerke, die über die Grenzen einzelner Verbandsebenen hinausreichen. Die Experimentell-disruptive Innovation (Typ C) gelingt oft nur dort, wo einzelne Akteure genügend Autonomie besitzen, um unabhängig von formellen Abstimmungsprozessen agieren zu können.

Zugleich begünstigt die föderale Struktur das Entstehen informeller Nutzungspraktiken. Wo formelle KI-Strategien fehlen oder sich in Abstimmungsprozessen zwischen den Ebenen befinden, nutzen Fachkräfte auf lokaler Ebene eigenständig frei verfügbare KI-Werkzeuge, häufig ohne Wissen der übergeordneten Organisationseinheiten. Die in Kapitel 10 dokumentierten Phänomene der informellen KI-Nutzung sind daher nicht nur Ausdruck individuellen Handelns, sondern auch ein strukturelles Merkmal lose gekoppelter Organisationen, in denen Steuerungsimpulse der Spitze nur begrenzt an der Basis ankommen.

9.3 Zur Frage der Feldspezifik und Übertragbarkeit

Angesichts der beschriebenen Merkmale stellt sich die Frage, inwieweit die in diesem Bericht generierten Erkenntnisse wohlfahrtsspezifisch sind oder auch für andere Organisationstypen Geltung beanspruchen können. Manche der identifizierten Dynamiken, etwa die Herausforderung, KI-Innovationen von der Projekt- in die Regelstruktur zu überführen, oder das Spannungsverhältnis zwischen Datenschutzanforderungen und Innovationsdruck, dürften auch in anderen Sektoren anzutreffen sein. Andere Befunde hingegen sind eng an die spezifische Verfasstheit der Wohlfahrtspflege gebunden. Die Gleichzeitigkeit von sozialarbeiterischem Auftrag und ökonomischer Rationalität, die sozialstaatliche Einbettung über das Subsidiaritätsprinzip, die föderale Organisationslogik mit ihren losen Kopplungen sowie die besondere Sensibilität der verarbeiteten Daten konstituieren ein Bedingungsgefüge, das in dieser Kombination anderen Feldern nicht ohne Weiteres eigen ist.

Gleichwohl lassen sich mögliche Parallelen zu verwandten Organisationstypen benennen, etwa zu föderal strukturierten Verbänden in anderen Bereichen des Dritten Sektors, zu Organisationen im Bildungs- oder Gesundheitswesen mit ähnlichen Hybriditätsmerkmalen oder zu öffentlichen Verwaltungen, die vergleichbare Steuerungsprobleme in Mehrebenensystemen aufweisen.

10 Gelingensbedingungen und Herausforderungen von KI-Innovationen in der Wohlfahrtspflege

Die empirischen Befunde aus den Interviews (Kap. 7) und der Fokusgruppendiskussion (Kap. 8) lassen sich entlang wiederkehrender Muster zu zentralen Herausforderungen und Gelingensbedingungen verdichten. Die folgende Zusammenstellung basiert auf dem erhobenen Datenmaterial und den daraus gewonnenen analytischen Kategorien. Sie bündelt die über die drei Innovationstypen hinweg identifizierten förderlichen und hemmenden Faktoren für eine KI-Innovation in der Wohlfahrtspflege. Diese Ergebnisse können als erste Anhaltspunkte für eine gelingende KI-Innovation gelesen werden. Allerdings muss einschränkend darauf verwiesen werden, dass die befragten Projekte zum Zeitpunkt der Interviews mehrheitlich noch nicht abgeschlossen waren und daher keine valide Bewertung für oder gegen eine erfolgreiche Implementierung vorgenommen werden kann.

10.1 Herausforderungen

i. Die „Projekt-Falle" und bürokratische Hürden

Die wohl markanteste Herausforderung, die sich quer durch die Innovationstypen zog, war die sogenannte „Projekt-Falle". Erfolgreiche Innovationen können nach Ablauf der Förderlaufzeit nicht in den Regelbetrieb überführt werden, da die bestehenden Förderstrukturen keine Finanzierung von Wartung, Betrieb und Weiterentwicklung vorsehen. Die Konsequenzen wurden drastisch formuliert: „Wenn die Förderung ausläuft, ist die Idee tot" (ID1, Pos. 27), Projekte „verschwinden im Keller" (ID11, Pos. 139) oder werden „eingestampft" (ID10, Pos. 48). In den Interviews wurde das Problem primär strukturell verortet. Gemeinnützige Organisationen dürften per Satzung keine unternehmerischen Risiken eingehen und könnten deshalb nach Projektende keine Geschäftsmodelle entwickeln. Ohne eine Überführung in die Regelfinanzierung drohen die erarbeiteten Lösungen wieder zu verschwinden. Verschärft wird dies dadurch, dass gängige Förderprogramme vorrangig Prozessbegleitung finanzieren, nicht jedoch die für eine nachhaltige Implementierung erforderlichen technischen Fachstellen und Infrastrukturen. Einzelne Projekte reagierten auf diese strukturelle Hürde mit innovativen Lösungsansätzen, etwa der Gründung einer eigenen GmbH, um Entwicklungsergebnisse in ein tragfähiges Geschäftsmodell überführen zu können.

Im Zusammenhang mit der „Projekt-Falle" wurden auch die bürokratischen Anforderungen der Förderprogramme selbst als Hindernis beschrieben. Die

Komplexität der Antragstellung und Dokumentation wurde als abschreckend wahrgenommen. Besonders pointiert wurde der Widerspruch zwischen Innovation und Verwaltungsanforderungen benannt: „Wir sind hier das Innovationslabor und digitales Projekt. Und dann komme ich mit diesen Dokumenten, wo man die Leute onboarden muss, die Selbsterklärung — aber es ist nicht digital" (ID1, Pos. 75). Darüber hinaus wurde kritisch angemerkt, dass der Verwaltungsaufwand für die Fördermittelbewirtschaftung selbst zusätzliche personelle Ressourcen binde, die angesichts begrenzter Förderbudgets an anderer Stelle sinnvoller eingesetzt werden könnten.

ii. Datenschutz: Zwischen Schutz und organisationaler Paralyse

Der Datenschutz erwies sich als ambivalenter Faktor. Einerseits stellte er eine notwendige Schutzfunktion dar, insbesondere angesichts hochsensibler personenbezogener Daten in der Sozialen Arbeit. Die datenschutzrechtliche Unsicherheit, etwa bezüglich der Frage, ob und wohin Daten fließen — „natürlich gehen auch Daten in die USA" (ID1, Pos. 57) — und die noch nicht im Recht verankerte KI-Richtlinie der EU erzeugten eine „Grauzone" (ID3, Pos. 66). Andererseits fungierte der Datenschutz in der Praxis häufig als Barriere, der Innovationen ausbremste. In der Fokusgruppe wurde beispielsweise berichtet, dass Diskurse abbrachen, sobald rechtliche Fragen aufkamen. Insbesondere beim Umgang mit klient:innenbezogenen Daten war die Haltung bewusst restriktiv. Die daraus resultierende Zurückhaltung führte in Organisationen dazu, dass KI-Anwendungen auf Bereiche ohne Personenbezug beschränkt blieben und potenziell sinnvolle Einsatzszenarien, etwa in der Dokumentation oder der Fallarbeit, von vornherein ausgeschlossen wurden.

iii. Informelle Nutzung

Das Fehlen formeller KI-Strategien führte nicht unbedingt zur Nicht-Nutzung, sondern zur Verlagerung in den informellen Bereich. Die Analyse zeigte eine „informelle, dann wahrscheinlich auch viel über private Accounts" stattfindende Nutzung (ID3, Pos. 18). Die Fokusgruppe verdeutlichte, dass Fachkräfte privat längst Tools wie ChatGPT nutzen und diese Praxis nahtlos auf berufliche Aufgaben übertragen. Die informelle Nutzung geschehe dabei im Affekt einer „Euphorie" und schaffe vollendete Tatsachen (Fokusgruppe, Pos. 47). Die informelle Dynamik wurde auch auf der Leitungsebene wahrgenommen. Mit dem schrittweisen Inkrafttreten der europäischen KI-Verordnung entstand ein regulatorischer Druck, der insbesondere Führungskräfte dazu veranlasste, sich mit der Frage einer rechtskonformen und

formellen Nutzung von KI-Anwendungen auseinanderzusetzen (ID3, Pos. 18). Die Herausforderung bestand dabei nicht nur in der Einhaltung datenschutzrechtlicher Vorgaben, sondern auch in der Entwicklung organisationaler Rahmenbedingungen, die eine kontrollierte und zugleich innovationsoffene Nutzung ermöglichen. Für viele Organisationen stellt sich damit die Frage, wie der Übergang von einer informellen, individuell verantworteten Nutzung hin zu einer institutionell abgesicherten KI-Strategie gelingen kann.

iv. Defizitäre digitale Infrastruktur als Grundlagenproblem

Eine wesentliche Hürde lag in der noch weitgehend analogen Realität vieler Einrichtungen. Die langfristig angedachte KI-gestützte Software trifft auf eine grundlegende Digitalisierungslücke. In zahlreichen Einrichtungen fehlt es an basaler digitaler Infrastruktur, sodass die Einführung von KI nicht als einzelner Innovationsschritt, sondern als Mehrfachanforderung erlebt wird, bei der grundlegende Digitalisierung und KI-Implementierung gleichzeitig bewältigt werden müssen (ID5, Pos. 62; ID1, 2, Pos. 22). Hinzu kommt, dass viele Wohlfahrtsorganisationen nicht über die erforderliche Datengrundlage verfügen, da zentrale Datensysteme über Jahrzehnte nicht systematisch aufgebaut wurden (ID12, Pos. 16). Der Aufbau eigener KI-Infrastrukturen übersteigt dabei die finanziellen und personellen Kapazitäten der meisten Träger und bleibt allenfalls für sehr große Organisationen realisierbar.

v. Fehlende eigene Entwicklungskapazitäten und Spannungsfeld mit externen Dienstleistern

In keinem der untersuchten Projekte waren Entwickler:innen im Sinne von Personen, die ein KI-System technisch programmieren, angestellt. Die Befragten beschrieben eine strukturelle Lücke zwischen vorhandener fachlicher Konzeptionsstärke und fehlender technischer Umsetzungskompetenz, die eine Abhängigkeit von externen Dienstleistern begründe. Die Förderbedingungen verschärften dieses Problem, da über Projektmittel in der Regel allein Workshopreferent:innen, nicht jedoch weitere technische Fachstellen finanziert werden konnten (ID1, Pos. 67). Die daraus resultierende Empfehlung mehrerer Befragter war, dass eine nachhaltige KI-Implementierung die Schaffung eigener Entwicklungsstellen voraussetze. Einzelne Akteure kompensierten dieses Defizit, indem sie KI-Werkzeuge selbst als Entwicklungsressource nutzten und so trotz fehlender Programmierkenntnisse

eigene Anwendungen erstellten. Eine tragfähige Lösung für die Gesamtorganisation stellt dies jedoch nicht dar.

Die daraus folgende Zusammenarbeit mit externen Entwicklungspartnern war dabei von grundlegenden Spannungen geprägt. Die gemeinnützige Logik der Wohlfahrtsverbände und die kommerzielle Orientierung der Dienstleister erzeugten wiederholt Reibungen, etwa durch intransparente Kostenstrukturen oder durch eine mangelnde Passung der Geschäftsmodelle. Während Verbände plattformisierte Zentrallösungen für ihre Mitgliedsorganisationen anstrebten, orientierten sich kommerzielle Anbieter häufig an Einzellizenzmodellen, die mit der verbandlichen Nutzungslogik nicht vereinbar waren. Zudem unterlagen die Preismodelle der zumeist jungen Unternehmen einer eigenen Dynamik, die eine langfristige Kalkulierbarkeit erschwerte.

vi. Unterschätzte Prozessanpassung

Eine vielfach benannte, aber häufig unterschätzte Herausforderung lag in der Notwendigkeit, organisationale Prozesse an die neuen technologischen Möglichkeiten anzupassen. Die Befragten betonten, dass das Potenzial von KI nur dann ausgeschöpft werden könne, wenn bestehende Arbeitsabläufe entsprechend neu ausgerichtet würden. Die bloße Bereitstellung von KI-Zugängen ohne begleitende Prozessberatung und Schulung wurde als unzureichend bewertet. Vielmehr erfordere eine gelingende Implementierung einen systemischen Ansatz, der technische Einführung, Kompetenzentwicklung und organisationale Prozessanpassung als zusammenhängende Aufgaben begreift. Ohne diese begleitende Organisationsentwicklung drohe die Innovation wirkungslos zu bleiben, da neue Werkzeuge lediglich auf veraltete Strukturen aufgesetzt würden. In mehreren Interviews wurde in diesem Zusammenhang beschrieben, wie die Auseinandersetzung mit KI gleichsam als Katalysator wirkte, der tradierte Prozesse sichtbar machte und eine überfällige Diskussion über deren Sinnhaftigkeit anstieß.

vii. Konkurrenz innerhalb der Wohlfahrt und fehlende Kooperation

Eine übergreifende Herausforderung stellte die Konkurrenz innerhalb der Wohlfahrt dar, die eine synergetische Zusammenarbeit und das Teilen von Best Practices grundsätzlich erschwerte. Kooperation erfolgte vielfach auf „Goodwill", weil Konkurrenzlogiken die Zusammenarbeit beeinträchtigten (ID12, Pos. 38; ID10, Pos.

10). Dieser als „Koopkurrenz" (Schönig 2015) bekannte Zustand beschreibt das wohlfahrtsspezifische Spannungsfeld zwischen einem geteilten normativen Ziel und dem faktischen Wettbewerb um Ressourcen, Reputation und Fälle (Kap. 7.3). Zugleich zeigten sich Ansätze, diese Dynamik zu überwinden: „Ich glaube auch, dass wir in der Wohlfahrt ganz gut beraten sind, wenn wir uns da gemeinschaftlich ein bisschen aufstellen" (ID3, Pos. 66). Die Forderung nach verbandsübergreifender Zusammenarbeit wurde von mehreren Befragten geteilt.

10.1 Gelingensbedingungen

i. Rückhalt und Offenheit der Leitungsebene

Als grundlegende Voraussetzung für die Initiierung und Umsetzung von KI-Innovationen erwies sich in den Interviews die aktive Unterstützung durch die Leitungsebene. Entscheidend war dabei nicht nur die Freigabe von Ressourcen, sondern eine persönliche Offenheit für das Thema. So wird etwa die direkte Anbindung an die Geschäftsführung als durchaus ambivalent, aber letztlich förderlicher Faktor beschrieben: Die Nähe zur Entscheidungsebene ermöglichte schnelles Handeln, wenn der Entscheider „auch sehr fasziniert von diesem Thema KI [war] und unbedingt ein erstes KI Projekt haben" wollte (ID8, Pos. 39). Befunde korrespondieren mit den in Kapitel 7.4 beschriebenen typübergreifenden Dynamiken, in denen hervorgehoben wird, dass für die Akzeptanz von Innovationen entscheidend sei, „wer der Datenschutzbeauftragte ist [und] wer da in der Geschäftsführung ist, wie vorsichtig die sind" (ID11, Pos. 49). In einem anderen Fall wird die Offenheit der gesamten Leitungsebene als Voraussetzung hervorgehoben, wobei explizit betont wird, dass es sich dabei nicht um „blinde Naivität", sondern um eine informierte Aufgeschlossenheit handle (ID4, Pos. 27).

ii. Engagierte Schlüsselpersonen und organisationale Freiräume

Über alle drei Innovationstypen hinweg erwies sich das Vorhandensein engagierter Schlüsselpersonen als eine entscheidende Gelingensbedingung. Die Befragten betonten, dass die meisten Einrichtungen „gar nicht die Kapazität [haben], das dann zusätzlich zum täglichen Business umzusetzen" und es deshalb „Pioniere im Verband" brauche, die sagen, „wir machen das jetzt" (ID5, Pos. 40). Diese Pionier:innen zeichneten sich durch eine hohe intrinsische Motivation aus, die auch über die reguläre Arbeitszeit hinausging: „Das ist nicht im Projekt entstanden, sondern aus der persönlichen Leidenschaft" (ID1, Pos. 32). Ebenso wichtig war das

Engagement der operativen Ebene, die neben den Schlüsselpersonen die Innovationen von Anfang an mittragen, „die von Anfang an sehr interessiert an dieser Technologie gewesen“ sind (ID5, Pos. 40). Besonders ausgeprägt zeigte sich dies bei dem Innovationstypus C, in der aus Early Adoptern „Koalitionen der Willigen" (ID1, Pos. 21) gebildet wurden. Die Diffusion der Innovation erfolgte hier über den Nachweis praktischer Nützlichkeit. Erfolgreiche Experimente erzeugten einen Sog, der neue Bedarfe in der Organisation konstituierte. Mitarbeitende kamen auf die Innovator:innen zu und fragten: „Das und das brauchen wir auch" (ID1, Pos. 16). An diesem Beispiel zeigt sich, dass die Schlüsselpersonen auch zur Diffusion von Wissen und als zentrale Ansprechpersonen dienen.

Neben dem Vorhandensein von Schlüsselpersonen hoben die Befragten hervor, dass gelingende KI-Innovation nicht allein durch engagierte Einzelpersonen, sondern erst im Zusammenhang mit organisationalen Freiräumen ermöglicht wurde. Als entscheidend wurde das Vorhandensein von „Overheads" beschrieben, also Personen und Teams, „die nicht im täglichen Geschäft drin sind, die solche Technologien aussuchen können, die die Implementierung organisieren können" (ID5, Pos. 54). Diese Kapazitäten ermöglichten es, Marktrecherche zu betreiben, Produkte zu evaluieren und die Einführung zu steuern. Auch die Einrichtung physischer und organisatorischer Innovationsräume wurde als förderlich benannt; beispielsweise Räumlichkeiten, die eine „Start-up Atmosphäre" schufen und Mitarbeitende „aus ihrem Alltag" herausziehen konnten (ID1, Pos. 73).

iii. Datenschutzkonforme Infrastrukturen und technologische Grundsatzentscheidungen

In den Interviews erwies sich die frühzeitige und bewusste Auseinandersetzung mit datenschutzrechtlichen Anforderungen als förderlich, sofern sie nicht defensiv als Blockade, sondern konstruktiv als Gestaltungsrahmen angelegt wurde. Die kooperative Aushandlung datenschutzkonformer technischer Lösungen stellte insbesondere im Typ B einen Erfolgsfaktor dar. Die Nutzung abgeschirmter Cloud-Umgebungen stellte sicher, dass „alle Daten in dieser Cloud bleiben" und nicht zu Trainingszwecken verwendet werden (ID10, Pos. 24). Besonders hervorgehoben wurde die Möglichkeit, organisationsinterne Daten zu verknüpfen, sodass KI „dort sucht statt im ganzen Web" (ID4, Pos. 57). Als konkrete Anforderung an eine tragfähige technische und datenschutzsichere Infrastruktur wurde formuliert, dass auf der einen Seite ein „sauberes System", in dem nichts „in die Trainingsdaten rein [geht] und nicht nach Übersee" läuft (ID3, Pos. 26), geschaffen werden soll, auf der anderen Seite wurde die notwendige Technologieoffenheit eines solchen Systems

betont, also die Fähigkeit, verschiedene Sprachmodelle anbinden zu können, um die Abhängigkeit von nur einem Modell zu reduzieren.

iv. Betriebsvereinbarungen und Compliance-Rahmen

Die frühzeitige Erarbeitung von Betriebsvereinbarungen und Compliance-Richtlinien wurde als strukturelle Absicherung der Innovation beschrieben. Es sei „wichtig, mit den Einrichtungen und mit den Organisationen von vornherein darüber zu sprechen, dass der Einsatz dieser Technologie nicht zu einer Arbeitsverdichtung führt" (ID5, Pos. 64). Dabei wurden sowohl regionale als auch landesweite Betriebsvereinbarungen als Referenz genannt. Ergänzend dazu erarbeiteten Verbände „Leitlinien zum KI-Einsatz" (ID5, Pos. 34) und umfassende Compliance-Rahmen „ethischer Art und Weise, aber eben auch datenschutzbezogen und [bezüglich] der Sicherheit der eigenen Unternehmensdaten" (ID4, Pos. 10). Die Entwicklung von KI-Richtlinien im Tandem mit Datenschutzbeauftragten wurde dabei als bewährte Praxis hervorgehoben.

v. Systematischer Kompetenzaufbau und Multiplikator:innen-Ansatz

Dass KI-Projekte nur dann gelingen, wenn sie von gezielten Qualifizierungsmaßnahmen begleitet werden, zog sich als Erkenntnis durch nahezu alle Interviews. Mehrfach wurde dabei ein mehrstufiger Ansatz beschrieben. Zunächst wurden ausgewählte Mitarbeitende als Multiplikator:innen „in einer Intensivschulung" qualifiziert, „um sich selber mit diesem Thema zu befassen und sodann auch als Multiplikatoren aktiv zu werden" (ID2, Pos. 20-21). Anschließend wurde eine strukturierte Skalierungsstrategie verfolgt. Das heißt, „wir haben 100 Multiplikatorinnen [...] die haben jetzt natürlich in dem halben Jahr eine sehr hohe Kompetenz aufgebaut [...] Und jetzt skalieren wir das quasi" (ID3, Pos. 26). Inhaltlich umfasste die Qualifizierung nicht nur technische Bedienungskompetenz, sondern ein grundsätzliches Verständnis der Technologie. Die Schulungen wurden zudem als Gelegenheit genutzt, „den Leuten erstmal die Berührungsängste" zu nehmen, wozu eine „vertrauensvolle Arbeitsoberfläche" und eine „gewisse Fehlerkultur" von Anfang an als „A und O für den Einstieg" beschrieben wurden (ID3, Pos. 53). Darüber hinaus etablierten sich Formate des Peer Learnings, in denen Teilnehmende eigenständig Prompt-Vorlagen und KI-Assistenten entwickelten und diese mit der Gemeinschaft teilten.

vi. Partizipation, iteratives Vorgehen und Wissensteilung

Ein durchgängiges Muster erfolgreicher Innovationsprojekte war die konsequente Beteiligung der Nutzer:innen an der Entwicklung und Weiterentwicklung der KI-Anwendungen. In einem Projekt wurden sogenannte Clubs gegründet, die sich „alle sechs Wochen in verschiedenen Zusammensetzungen" trafen (ID1, Pos. 16), um gemeinsam zu definieren, was als nächstes konstruiert werden soll. Dieses iterative Vorgehen schuf eine geteilte Verantwortung für das Ergebnis. Als weitere Gelingensfaktoren wurden die Transparenz über die Möglichkeiten und Grenzen der Technologie benannt — „falsche Versprechungen [sind] Gift für das ganze Thema" (ID3, Pos. 53) — sowie die Messbarkeit der Ergebnisse. Darüber hinaus wurde die aktive Wissensteilung innerhalb der Projekte als förderlich beschrieben. Teilnehmende entwickelten eigenständig Prompt-Vorlagen und stellten diese auf gemeinsamen Plattformen bereit.

Auch in der Zusammenarbeit mit externen Partnern stellten die Zusammenarbeit und das iterative Vorgehen eine wesentliche Gelingensbedingung dar. Kooperationen mit wissenschaftlichen Akteuren, etwa dem Fraunhofer-Institut, wurden als vertrauensbildend beschrieben, da diese die technische Machbarkeit unabhängig validieren konnten. Bei der Auswahl kommerzieller Entwicklungspartner erwies sich die Bewertung der unternehmerischen Solidität und Produktentwicklung als relevant. Die intensive nutzer:innenzentrierte Produktentwicklung durch Start-ups, die „monatelang in Pflegeeinrichtungen gesessen und analysiert [haben], wie die Arbeitsabläufe sind" (ID5, Pos. 54), wurde als qualitätssicherndes Merkmal hervorgehoben. Zugleich wurde ein Bedarf nach „vertrauenswürdigen Partnern, die nicht einfach nur Dienstleister sind und bei denen wir nicht in eine Abhängigkeit geraten" formuliert (ID12, Pos. 98).

11 Fazit

11.1 Limitationen der Studie

Die in dieser Studie vorgenommene Rekonstruktion von Falltypen auf Grundlage eines Korpus von zwölf Interviews unterliegt mehreren wissenschaftlichen Limitationen.

Erstens ist die Begrenztheit der **Stichprobengröße** hervorzuheben: der vorliegende Datensatz ermöglicht zwar dichte, kontextualisierte Einblicke, schränkt jedoch die theoretische Sättigung und die Verallgemeinerbarkeit der Ergebnisse ein, da aus forschungspraktischen Gründen die Anzahl an Interviews beschränkt war und nicht bis zu einer Sättigung im Sinne des Theoretical Samplings im Sinne der Grounded Theory erweitert werden konnte. Insofern ist nicht auszuschließen, dass sich bei weiterer Forschung auch noch andere, hier nicht beschriebene Innovationstypen rekonstruieren lassen. Die herausgearbeiteten Innovationstypen stellen idealtypische Konstruktionen dar, die sich in praktischen Innovationsprojekten durchaus auch im zeitlichen Verlauf auch abwechseln oder in Mischformen auftreten können.

Zweitens besteht eine **Kontextabhängigkeit der Daten**. Unsere Interviewpartner in der Interviewstudie gehören zu einer Gruppe von Personen, die besonders affin gegenüber technologischen Neuerungen sind. Zudem handelt es sich um Personen, die nicht direkt in der Interaktionsarbeit mit Adressat:innen arbeiten, sondern in Verwaltungs- und Leitungspositionen. Sie sind besonders offen gegenüber den technologischen Innovationsprozessen und sind selbst in die Entwicklung oder Implementierung verwickelt. Dadurch fehlt zuweilen eine Distanz gegenüben den Innovationen. Auf der anderen Seite finden sich bei der Zusammensetzung der Fokusgruppe Fachkräfte aus der Praxis und der unmittelbaren Interaktionsarbeit. Ihre Aussagen zur Einführung von KI in der jeweiligen Organisation müssen ebenfalls im Kontext gesehen werden.

Drittens besteht eine Limitation bei der empirischen Samplingstrategie. Viele der von uns untersuchten Projekten wurden über das Schneeballverfahren ermittelt. Dadurch bleiben Innovationsprojekte, die nicht über die existierenden Kontakte, andere Protagonisten im Feld oder im Internet aufzufinden sind, verschlossen und konnten nicht mit in das Sample aufgenommen wurden. Hierdurch entsteht ein sogenannter **Selection Bias**. Für unsere Studie wurden primär solche Projekte erreicht, die eine gute Anbindung an ihren Bundesverband hatten, oder eine gewisse Öffentlichkeitsarbeit für ihr Projekt gemacht haben. Dadurch ist es sehr wahrscheinlich, dass wir mit der Studie vor allem aktive, erfolgreiche Projekte ermittelt haben und weniger Projekte, die gescheitert sind oder solche, die gerade

erst gestartet haben. Dies führt potenziell zu einer Homogenisierung des Samples, da Empfehlungen häufig innerhalb ähnlicher Zirkel ausgesprochen werden.
Viertens muss auf die **retrospektive Verzerrung** und **soziale Erwünschtheit** hingewiesen werden. Da ein Teil der Daten auf Interviews beruhen, besteht die Möglichkeit, dass die Befragten die oft chaotischen Innovationsprozesse im Rückblick linearer und rationaler darstellen, als diese tatsächlich verlaufen sind. Des Weiteren ist das Thema Innovation derzeit ein positiv besetzter Begriff. Dies könnte eine soziale Erwünschtheit bei der Beantwortung der Fragen mit sich gebracht haben. Befragte könnten Probleme heruntergespielt haben, um das Projekt, ihre Organisation oder sich selbst als erfolgreich darzustellen.
Schließlich liegt eine Limitierung mit Blick auf die Fokusgruppe in der Zusammensetzung der Teilnehmenden, die alle auf mittleren Leitungspositionen tätig waren. Hier besteht die Möglichkeit, dass vor allem aus einer Organisationsperspektive heraus argumentiert wurde, die geprägt ist von einer Leitungsrolle, die vor allem eine rechtskonforme und fachlich angemessene Nutzung von KI sicherzustellen hat (z.B. das KI-Anwendungen datenschutzkonform sein müssten). Ergebnisse dieser Fokusgruppe könnten in anschließenden Forschungen kontrastiert werden durch Fokusgruppen mit Personen, die über keine Leitungspositionen verfügen. Unsere Hypothese wäre, dass hier noch einmal stärker informelle Nutzungspraktiken berichtet werden könnten, die etwa mit datenschutzrechtlichen Anforderungen konfligieren.

11.2 Einordnung Ergebnisse vor dem Hintergrund des Diskurses

Die vorliegende Untersuchung verfolgte das Ziel, den aktuellen Status der digitalen Transformation – spezifisch den Einsatz Künstlicher Intelligenz – in der Freien Wohlfahrtspflege nicht nur deskriptiv zu erfassen, sondern in seinen Innovationsdynamiken zu verstehen und Regelmäßigkeiten in Form von Typen zu rekonstruieren. Aufbauend auf der Erläuterung technologischer Grundlagen in Kapitel 2 haben die in Kapitel 3 entfaltete Rekonstruktion zentraler Anwendungsfälle von KI in der Wohlfahrtspflege - von Dialogsystemen über komplexe Entscheidungs-unterstützungs- und Assistenzsysteme bis hin zur Optimierung der administrativen Prozesse und der Dokumentation – und die in Kapitel 4 aufgezeigten Risiken und Herausforderungen von KI-Technologien, ein grundlegendes Verständnis zu möglichen Szenarien einer Nutzung von KI in der Wohlfahrtspflege geschaffen. Führt man diese aus dem fachlichen Diskurs entwickelte Rahmung mit den in den Kapiteln 7 und 8 gewonnenen empirischen Einblicken zusammen, offenbaren sich Diskrepanzen zwischen der in der Literatur proklamierten Potentialität und den

untersuchten konkreten Anwendungsfällen. Die Realität der Sozialen Arbeit, so zeigen die Befunde, nutzt das sozio-technische Potenzial von Künstlicher Intelligenz nicht in seiner Breite.

Ein zentraler Befund dieser Studie stellt die **faktische Absenz** bestimmter theoretischer Diskursstränge im Praxisfeld und in den Innovationstypen dar. Die im fach-wissenschaftlichen Diskurs beschriebenen Szenarien etwa einer KI-gestützten Diagnostik finden sich in der von den Befragten beschriebenen Praxis des KI-Einsatzes nicht wieder. Mehr noch: In der Fokusgruppe (Kapitel 8.1) kippt die Innovationssemantik ins Negative, sobald sie sich nicht mehr rein auf administrative Aspekte bezieht und den Kern der pädagogischen Interaktion berührt.

Was theoretisch als Erweiterung professioneller Handlungsfähigkeit gedacht ist, wird in der Fokusgruppe praktisch als „Bedrohung“ oder „defizitäre Simulation“ menschlicher Zuwendung gedeutet. Die Furcht vor existentiellen Fehlentscheidungen – prägnant illustriert am Vergleich des „falsch beladenen LKWs“ mit der Arbeit am Menschen – führt dazu, dass KI-Systeme zur Entscheidungsunterstützung in der Praxis weitgehend ausgeklammert bleiben. Hier zeigt sich keine bloße Implementierungslücke, sondern eine manifeste kulturelle Barriere: Es zeigt sich eine protektive Haltung des professionellen Habitus, der die Unmittelbarkeit der Beziehungsarbeit gegen eine als systemfremd empfundene Technologisierung in Stellung bringt. Die in Kapitel 4 theoretisch diskutierten Risiken der „Verzerrung“ und „Fehlannahmen“ werden in der Praxis nicht technisch, sondern ethisch verhandelt und führen zu einer präventiven Abwehrhaltung gegenüber fallbezogenen KI-Anwendungen.
Demgegenüber steht eine bemerkenswerte Konvergenz, die sich wie ein roter Faden durch die Äußerungen der Befragten zieht: Die Dominanz der Verwaltungsperspektive. Beide Analyseebenen – Interviews und Fokusgruppe - identifizieren die „Unterstützung im Alltag der Sozialen Arbeit“ (Kapitel 3.3) als weitgehendes Desiderat. Zwar kommen im theoretischen Diskurs die Möglichkeiten von KI im Bereich von *Ambient Assisted Living* oder in Klienten-Tools vor, in der Empirie spielen diese jedoch eine marginale Rolle. Die „Technik am Klienten“ bleibt eine Randerscheinung, die im Schatten der „Technik für die Akten“ verschwindet.

Dies korrespondiert mit dem zweiten, überaus deutlichen Konvergenzbefund: Der Fokussierung auf die Dokumentation. Bereits in der theoretischen Herleitung (Kapitel 3.4) wurde die Dokumentationsarbeit als immenser zeitlicher Ressourcenaufwand identifiziert. Im Gegensatz zu den vorgenannten Visionen findet *diese* theoretische Position in der Empirie ihre faktische Entsprechung. Die Dokumentation fungiert als das „Einfallstor der geringsten Gegenwehr“. In diesem Kontext wird KI nicht als Gefährdung der professionellen Identität wahrgenommen, sondern als

willkommenes Entlastungsinstrument zur Bewältigung administrativer Routinen. Es zeigt sich eine bemerkenswerte Konvergenz zwischen der theoretisch diagnostizierten Überbürokratisierung und der pragmatischen Lösungsstrategie, KI als Schreibassistenz zu nutzen. Innovation wird in der Sozialen Arbeit an dieser Stelle somit als *Entlastung von fachfremden Tätigkeiten* zugelassen, um den *eigentlichen* Kern der Arbeit – die Interaktion – frei von Technologie zu halten.

Diese „Schieflage" – hier die überbordende Nutzung für Textproduktion, dort die defensive Vermeidung in der Fallarbeit – birgt ein Risiko, das in Kapitel 8 als „informelle Praxis" identifiziert wurde. Weil offizielle Strategien fehlen oder sich auf „sichere" Verwaltungsaspekte beschränken, entsteht ein „organisationales Vakuum". In diesen leeren Raum dringen informelle Praktiken ein: Fachkräfte nutzen Tools wie ChatGPT auf eigene Faust, oft ohne Datenschutzkonformität, getrieben von einer „euphorischen" Alltagspraxis.

Die Soziale Arbeit steht damit an einem Scheideweg. Verbleibt sie in der aktuellen Haltung, droht sie die Gestaltungsmacht über die Technologie zu verlieren. Wenn KI nur dort legitim ist, wo sie Akten schreibt, aber dort tabuisiert wird, wo sie Entscheidungen vorbereitet oder Klienten unterstützt, überlässt die Profession das Feld der fallbezogenen KI externen, tech-getriebenen Akteuren oder der unregulierten Schatten-IT.

Es bedarf daher dringend einer offensiven Strategie, die über die bloße Unterstützung von Dokumentationsaufgaben hinausgeht. Die in Kapitel 4 angemahnte „Kompetenzentwicklung" darf sich nicht in der Bedienung von Prompts erschöpfen, sondern muss die Fähigkeit zur kritischen Reflexion hybrider Entscheidungsarchitekturen umfassen. Nur wenn die Soziale Arbeit beginnt, KI nicht nur als „Bürokratiekiller", sondern als potenzielles (und kritisch zu prüfendes) Fachinstrument zu begreifen, kann sie die in diesem Bericht aufgezeigte Lücke zwischen theoretischem Potenzial und praktischer Nutzung im Sinne ihrer Adressat:innen schließen.

11.3 Ausblick und Ableitungen

Die vorliegende Untersuchung offenbart eine große Heterogenität in der Auseinandersetzung mit Künstlicher Intelligenz (KI) in der Sozialen Arbeit. Während sich auf der Ebene von KI-Projekten drei distinkte Innovationstypen – von der *Pragmatischen Modernisierung* (Typ A) über die *Reflexive Sozialinnovation* (Typ B) bis zur *Disruptiven Innovation* (Typ C) – ausdifferenzieren, zeigt die operative Ebene der Fachkräfte eine Ambivalenz zwischen professionsethischer Skepsis und informeller Nutzung.

Im Sinne des Konzepts der soziotechnischen Imaginäre (Jasanoff & Kim, 2015) lassen sich diese Innovationstypen als kollektiv geteilte Zukunftsvisionen verstehen, die jeweils spezifische Vorstellungen wünschenswerter sozialer Ordnung mit konkreten technologischen Entwicklungspfaden verknüpfen. Auffällig ist dabei jedoch, dass trotz der Unterschiede zwischen den drei Typen eine gemeinsame Grundfigur dominiert, die als Rationalisierungsimaginäres bezeichnet werden kann. Über alle drei Typen hinweg wird KI primär als Instrument zur Effizienzsteigerung, Entlastung von administrativen Tätigkeiten und Optimierung bestehender Prozesse legitimiert.

Diese Dominanz des Rationalisierungsnarrativs ist aus analytischer Perspektive folgenreich. Soziotechnische Imaginäre wirken, wie Jasanoff und Kim (2015) betonen, nicht nur ermöglichend, sondern auch begrenzend. Indem sie bestimmte Zukünfte als wünschenswert und erreichbar rahmen, schließen sie zugleich alternative Entwicklungspfade aus dem kollektiven Vorstellungsraum aus. Bezogen auf die Wohlfahrtspflege bedeutet dies, dass Einsatzszenarien von KI, die über Rationalisierung hinausgehen, systematisch unterbelichtet bleiben. Denkbar wären etwa partizipative Ansätze, in denen Adressatinnen und Adressaten aktiv an der Gestaltung KI-gestützter Ansätze beteiligt werden, oder explorative Nutzungsformen, die KI als Werkzeug fachlicher Reflexion und professioneller Wissensgenese einsetzen, statt sie lediglich als Mittel zur Entlastung von Routinetätigkeiten zu begreifen. Ebenso kommen gemeinwohlorientierte Anwendungen, etwa im Bereich der sozialräumlichen Analyse oder der datengestützten Bedarfsplanung, kaum in den Blick, wenn das vorherrschende Imaginäre auf individuelle Arbeitsentlastung verengt bleibt. Die in den Interviews rekonstruierten Innovationssemantiken spiegeln damit nicht nur unterschiedliche Strategien im Umgang mit KI wider, sondern auch die Grenzen des gegenwärtig geteilten Vorstellungsraums darüber, was KI in der Sozialen Arbeit sein und leisten könnte.

Aus dieser Diskrepanz zwischen organisationaler Innovationssemantik und gelebter Praxis lassen sich spezifische Handlungsbedarfe für die politische Steuerung und die verbandliche Strategiebildung ableiten.

Für die staatliche Förderpolitik ergibt sich die Notwendigkeit einer Anpassung der Förderlogiken, um die empirisch belegte „Projekt-Falle“ (vgl. Typ B) zu überwinden.

- **Verschiebung von Projekt- zu Infrastrukturförderung:** Die Ergebnisse zeigen, dass Innovationen (Typ B und C) häufig nach Auslaufen der Anschubfinanzierung stagnieren, da Wartung und Betrieb in der

Regelprojektlogik nicht abgebildet werden. Um Innovationen nachhaltig zu implementieren, bedarf es Förderinstrumente, die den Transfer vom Experimentierraum in den Regelbetrieb („Dauerbetrieb“) finanziell absichern.

- **Förderung technologischer Souveränität:** Angesichts der Abhängigkeit von kommerziellen „Stangenlösungen“ (Typ A) und dem gleichzeitigen Wunsch nach Emanzipation (Typ C), sollte die Entwicklung von *Open-Source*-Lösungen und verbandsübergreifenden Plattformen gezielt incentiviert werden. Dies stärkt die Unabhängigkeit der Wohlfahrtspflege von volatilen Marktlogiken großer Tech-Konzerne.

Für die Spitzenverbände und Träger der Freien Wohlfahrtspflege resultieren aus der Studie drei zentrale strategische Imperative:

- **Formelle fachliche Alternativen zu informellen Praktiken (Shadow-IT):** Die Auswertung der Fokusgruppendiskussion legt nahe, dass eine formelle Abstinenz von KI-Strategien nicht zur Nicht-Nutzung führt, sondern eine risikobehaftete, informelle Nutzungspraxis (Nutzung privater Accounts) fördern könnte. Träger müssen diesen „Graubereich“ durch die Bereitstellung datenschutzkonformer Sandbox-Umgebungen und klarer Richtlinien in formelle Bahnen lenken. Das „Noch-Nicht“ der Organisation darf nicht zum Risiko der Fachkraft werden.
- **Differenzierte Innovationskommunikation:** Die Sorge der Fachkräfte vor einer De-Professionalisierung und dem Verlust reflexiver Kompetenzen muss ernst genommen werden. Es gilt einerseits, das Narrativ der „Re-Professionalisierung“ zu stärken: KI übernimmt administrative „Zeitfresser“, um den humanzentrierten Kern der Sozialen Arbeit zu schützen. Die Kommunikation darf andererseits KI aber auch nicht pauschal als Effizienzsteigerung rahmen (Typ A), sondern muss – analog zu Typ B – ebenso den mögliche Nutzungsszenarien und Mehrwerte für Anwendungen auf der Ebene der *Beziehungsarbeit* herausstellen.
- **Allianzen statt Einzelkämpfertum:** Während Typ C („Disruptive Innovatoren“) beweist, dass technologische Souveränität durch *Inhouse*-Entwicklung möglich ist, übersteigt dies oft die Ressourcen einzelner Träger. Es empfiehlt sich die Bildung von Entwicklungskonsortien oder Genossenschaftsmodellen. Hierbei können die Agilität und Visionskraft der „Pioniere“ (Typ C) mit den strukturellen Ressourcen der etablierten Organisationen (Typ A/B) synergetisch verknüpft werden.

Die Soziale Arbeit steht an einem Scheideweg zwischen exogener Anpassung an Marktstandards und einer emanzipatorischen Gestaltung eigener digitaler

Werkzeuge. Welcher Entwicklungspfad eingeschlagen wird, hängt nicht zuletzt davon ab, welche soziotechnischen Imaginäre (Jasanoff & Kim, 2015) sich innerhalb der Wohlfahrtspflege durchsetzen und institutionell stabilisieren. Gelingt es, die Innovationskraft der „Koalition der Willigen“ strukturell zu verstetigen und gleichzeitig die berechtigte Skepsis der Basis durch partizipative Prozessöffnungen aufzugreifen, kann KI vom bloßen Rationalisierungsinstrument zu einem Ermöglichungsmedium professionellen Handelns avancieren.

Abbildungsverzeichnis

Abbildung 1: Exemplarische Ausschnitte aus dem Interviewleitfaden 31
Abbildung 2: Visualisierung der Stichprobe 32
Abbildung 3: Fallübersicht der durchgeführten Interviewstudie 36

Tabellenverzeichnis

Tabelle 1: Überblick der untersuchten Studien 26
Tabelle 2: Überblick über Typen und Merkmalsdimensionen 37

Abkürzungsverzeichnis

Abb.	Abbildung
AAL	Ambient Assisted Living
AWO	Arbeiterwohlfahrt Bundesverband e. V.
BAGFW	Bundesarbeitsgemeinschaft der Freien Wohlfahrtspflege e. V.
BMAS	Bundesministerium für Arbeit und Soziales
BMBFSFJ	Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend
BMUV	Bundesministeriums für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz
DCV	Deutscher Caritasverband e. V.
Diakonie	Diakonie Deutschland - Evangelisches Werk für Diakonie und Entwicklung e. V.
DRK	Deutsches Rotes Kreuz e. V.
Kap.	Kapitel
Kom.sd	Kompetenzzentrum Soziale Dienste
Tab.	Tabelle
ZWST	Zentralwohlfahrtsstelle der Juden in Deutschland e. V.

Literaturverzeichnis

Amrit, Chintan; Paauw, Tim; Aly, Robin; Lavric, Miha (2017): Identifying child abuse through text mining and machine learning. In: *Expert Systems with Applications* 88, S. 402–418. DOI: 10.1016/j.eswa.2017.06.035.

Avila, Mauro; Wolf, Katrin; Brock, Anke; Henze, Niels (2016): Remote Assistance for Blind Users in Daily Life. In: Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments. PETRA '16: 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments. Corfu Island Greece, 29 06 2016 01 07 2016. New York, NY, USA: ACM, S. 1–2.

Ayers, John W.; Poliak, Adam; Dredze, Mark; Leas, Eric C.; Zhu, Zechariah; Kelley, Jessica B. et al. (2023): Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. In: *JAMA internal medicine* 183 (6), S. 589–596. DOI: 10.1001/jamainternmed.2023.1838.

Baecker, Dirk (2005): Die Reform der Gesellschaft. In: *Corsi, G./Esposito, E. (Hrsg.): Reform und Innovation in einer unstabilen Gesellschaft.* Stuttgart: Lucius und Lucius.

Bareis, Jascha & Katzenbach, Christian (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. Science, Technology, & Human Values, 47(5), 855-881.

Bertsche, Oliver; Como-Zipfel, Frank (2023): Digitalisierung. Herausforderungen und Handlungsansätze für die Soziale Arbeit. 1. Auflage. Stuttgart: Verlag W. Kohlhammer (Soziale Arbeit in der Gesellschaft).

BfArM (2021): deprexis | DiGA-Verzeichnis. Online verfügbar unter https://diga.bfarm.de/de/verzeichnis/00450.

Blanc, B.; Beudt, S. (2022): Monitoring KI-gestützter Assistenztechnologien für Menschen mit Behinderungen. Ergebnisbericht des Projekts KI.ASSIST.: Bundesverband Deutscher Berufsförderungswerke e. V.

Broadhurst, Karen; Alrouh, Bachar; Yeend, Emily; Harwin, Judith; Shaw, Mike; Pilling, Mark et al. (2015): Connecting Events in Time to Identify a Hidden Population: Birth Mothers and Their Children in Recurrent Care Proceedings in England. In: *The British Journal of Social Work* 45 (8), S. 2241–2260. DOI: 10.1093/bjsw/bcv130.

Brunkhorst, Hauke (2003): Evolution und Revolution — Hat die Evolution des politischen Systems eine normative Seite? In: Kai-Uwe Hellmann (Hg.): Das System

der Politik. Niklas Luhmanns politische Theorie. 1. Aufl. Wiesbaden: Westdt. Verl., S. 326–335.

Bühler, Judith; Hartmann, Andrea; Tuggener, Don; Cieliebak, Mark; Baier, Dirk (2023): Safety-for-refugees.ch: eine Informations- und Meldewebseite für Ukraine-Flüchtlinge : Schlussbericht.

Bühler, Judith; Reiser, Meret (2023): Sind unsere neuen Arbeitskolleg*innen bald auch Chatbots?

Buxmann, Peter; Schmidt, Holger (2021): Künstliche Intelligenz. Berlin, Heidelberg: Springer Berlin Heidelberg.

Chouldechova, Alexandra; Benavides-Prado, Diana; Fialko, Oleksandr; Vaithianathan, Rhema (2018): A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In: *Conference on Fairness, Accountability and Transparency*, S. 134–148. Online verfügbar unter https://proceedings.mlr.press/v81/chouldechova18a.html.

Clarke, Adele (1991): Social worlds/arenas theory as organizational theory. In: *D. R. Maines (Hrsg.) Social Organization and Social Process: Essays in Honor of Anselm Strauss*, S. 119–158.

Connolly, Amy J.; Gill, Grandon T. (2015): Debates in Information Technology. A Capstone Class Project: Informing Science.

Dresing, Thorsten; Pehl, Thorsten (Hg.) (2016): Praxisbuch Transkription. Regelsysteme, Software und praktische Anleitungen für qualitative ForscherInnen. 6. Aufl. Marburg: Dr. Dresing und Pehl GmbH.

Echterhoff, Gerald; Bohner, Gerd; Siebler, Frank (2006): “Social Robotics” und Mensch-Maschine-Interaktion. In: *Zeitschrift für Sozialpsychologie* 37 (4), S. 219–231. DOI: 10.1024/0044-3514.37.4.219.

Ehlert, Wiking (1991): Rechner verändern institutionelle Strukturen! Computereinsatz, Organisationsentwicklung und Arbeit. In: Bernhard Meyer (Hg.): Hilfe vom Bildschirm: Computer in der sozialen Arbeit. Freiburg: Bernhard Meyer Verlag, S. 35–54.

Eidam, Benjamin; Bauer, Dan (2025): Einleitung: Warum jetzt jeder auf KI setzen muss. In: Benjamin Eidam und Dan Bauer (Hg.): Der KI-Schlüssel für Unternehmen. Wiesbaden: Springer Fachmedien Wiesbaden, S. 1–18.

Fujita, Hiroshi (2020): AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. In: *Radiological physics and technology* 13 (1), S. 6–19. DOI: 10.1007/s12194-019-00552-4.

Gapski, Harald (2020): Digitale Transformation: Datafizierung und Algorithmisierung von Lebens- und Arbeitswelten. In: Nadia Kutscher, Thomas Ley, Udo Seelmeyer, Friederike Siller, Angela Tillmann und Isabel Zorn (Hg.): Handbuch Soziale Arbeit und Digitalisierung. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 156–166.

Garud, Raghu; Karnøe, Peter (2001): Path Creation as a Process of Mindful Deviation. In: Raghu Garud und Peter Karnøe (Hg.): Path dependence and creation. Online-Ausg. New York: Psychology Press (Lawrence Erlbaum Associates), S. 1–40.

Gillingham, Philip (2016): Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the 'Black Box' of Machine Learning. In: *The British Journal of Social Work* 46 (4), S. 1044–1058. DOI: 10.1093/bjsw/bcv031.

Gillingham, Philip (2017): Predictive Risk Modelling to Prevent Child Maltreatment: Insights and Implications from Aotearoa/New Zealand. In: *Journal of Public Child Welfare* 11 (2), S. 150–165. DOI: 10.1080/15548732.2016.1255697.

Gläser, Jochen; Laudel, Grit (2010): Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen. 4. Auflage. Wiesbaden: VS Verlag. Online verfügbar unter http://d-nb.info/1002141753/04.

Görder, Björn (2020): Wann ist KI ok? In: *SozA* 69 (1), S. 16–22. DOI: 10.5771/0490-1606-2020-1-16.

Graf, Birgit (2020): Assistenzroboter für die Pflege : Verfügbare Produkte und Forschungsfelder. In: *Zeitschrift fur Gerontologie und Geriatrie* 53 (7), S. 608–614. DOI: 10.1007/s00391-020-01782-7.

Halfar, Bernd (2019): Fachkräftemangel: ein Blick aus dem Elfenbeinkeller. Jahresbericht 2019. Hg. v. Regens-Wagner-Stiftung. Dillingen.

Hatch, S. Gabe; Goodman, Zachary T.; Vowels, Laura; Hatch, H. Dorian; Brown, Alyssa L.; Guttman, Shayna et al. (2025): When ELIZA meets therapists: A Turing test for the heart and mind. In: *PLOS Ment Health* 2 (2), e0000145. DOI: 10.1371/journal.pmen.0000145.

Hauer, Franziska; Haug, Sonja; Schultz, Maximilian; Scharf, Anna; Weber, Karsten (2024): Künstliche Intelligenz für Nichtregierungsorganisationen - Quantitative Erforschung der Umsetzung in deutschen NROs. 3.Arbeitspapier.

Hilbert, Martin; López, Priscila (2011): The world's technological capacity to store, communicate, and compute information. In: *Science (New York, N.Y.)* 332 (6025), S. 60–65. DOI: 10.1126/science.1200970.

Hollanek, Tomasz; Sobey, Aisha (2025): AI Companions for Health and Mental Wellbeing: Opportunities, Risks and Policy Implications. Unter Mitarbeit von Apollo - University of Cambridge Repository und University of Cambridge: Leverhulme Centre for the Future of Intelligence. Online verfügbar unter https://www.repository.cam.ac.uk/items/d3229fe5-db87-42ff-869b-11e0538014d8.

Humby, Clive (2006): Data is the new oil. ANA Senior Marketer's Summit. Kellogg School of Management. Evanston, USA, 2006.

Jasanoff, Sheila (2015): Future imperfect: Science, technology, and the imaginations of modernity. In Sheila Jasanoff & Sang-Hyun Kim (Hrsg.), Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power (S. 1–33). University of Chicago Press.

Jasanoff, Sheila & Kim, Sang-Hyun (2009). Containing the atom: Sociotechnical imaginaries and nuclear power in the United States and South Korea. Minerva, 47(2), 119–146. https://doi.org/10.1007/s11024-009-9124-4

Jovanovic, Mladjan; Mitrov, Goran; Zdravevski, Eftim; Lameski, Petre; Colantonio, Sara; Kampel, Martin et al. (2022): Ambient Assisted Living: Scoping Review of Artificial Intelligence Models, Domains, Technology, and Concerns. In: *Journal of medical Internet research* 24 (11), e36553. DOI: 10.2196/36553.

Jungmann, Robert; Besio, Cristina (2018): Semantiken des sozialen Wandels: zur diskursiven Gestaltung von Innovation. In: *Zeitschrift für Diskursforschung / Journal for Discourse Studies* 6 (2), S. 11–42.

Kelle, Udo; Kluge, Susann (2010): Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung. 2., überarbeitete Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften (SpringerLink Bücher).

Kieslinger, Daniel (2024): Mit KI-Unterstützung zu einer inklusiven Kinder- und Jugendhilfe. In: *unsere jugend*. Online verfügbar unter https://elibrary.utb.de/doi/pdf/10.2378/uj20240707.

Kirchner, Elsa A.; Will, Niels; Simnofske, Marc; Kampmann, Peter; Benitez, Luis Manuel Vaca; Gea Fernández, José de; Kirchner, Frank (2019): Exoskelette und künstliche Intelligenz in der klinischen Rehabilitation. In: Mario A. Pfannstiel, Patrick Da-Cruz und Harald Mehlich (Hg.): Digitale Transformation von Dienstleistungen im Gesundheitswesen. Wiesbaden, Heidelberg: Springer Gabler, S. 413–435.

Kreidenweis, Helmut; Diepold, Maria (2024): Künstliche Intelligenz in der Sozialwirtschaft. Forschungsbericht.

Krings, Markus (2024): Künstliche Intelligenz und Soziale Arbeit. In: *BdW* 171 (1), S. 25–27. DOI: 10.5771/0340-8574-2024-1-25.

Kuckartz, Udo (2010): Typenbildung. In: Günter Mey und Katja Mruck (Hg.): Handbuch qualitative Forschung in der Psychologie. 1. Auflage. Wiesbaden: VS Verl. für Sozialwiss, S. 553–568.

La Valle, Ivana; Holmes, Lisa; Gill, Chloe, Brown, Rebecca, Hart, Di; Barnard, Matt (2016): Improving children's social care services: results of a feasability study. London: CAMHS Press.

Lehmann, Robert (2025): KI in der Beratung. In: Gesa Linnemann, Julian Löhe und Beate Rottkemper (Hg.): Künstliche Intelligenz in der Sozialen Arbeit: Beltz Juventa.

Lehmann, Robert; Burghardt, Jennifer (2024): Die Akzeptanz von Künstlicher Intelligenz bei der Entscheidungsunterstützung in der Kinder- und Jugendhilfe. In: Jens Albrecht, Martin Brinkmann, Jennifer Burghardt, Monika Feist-Ortmanns, Hans Jürgen Frischhut, Daniel Hahn et al.: Künstliche Intelligenz in der Kinder- und Jugendhilfe. 1. Auflage. Hg. v. Michael Macsenaere. München: Ernst Reinhardt Verlag, S. 67–73.

Lehmann, Robert; Burghardt, Jennifer (2025): KI-Kompetenz für Fachkräfte der Sozialen Arbeit. Empfehlungen für die Fort- und Weiterbildung. In: Deutscher Verein für öffentliche und private Fürsorge e.V. (Hg.): Digitalisierung und KI in der Sozialen Arbeit. Ausgabe 3/2025 - Archiv für Wissenschaft und Praxis der Sozialen Arbeit: Lambertus-Verlag, S. 76–84.

Linnemann, Gesa (2025): Grundlagen der „Mensch-KI“-Interaktion – Auswirkungen auf den Einsatz im Kontext- der Sozialen Arbeit. In: Gesa Linnemann, Julian Löhe und Beate Rottkemper (Hg.): Künstliche Intelligenz in der Sozialen Arbeit: Beltz Juventa, S. 35–46.

Linnemann, Gesa; Löhe, Julian; Rottkemper, Beate (2024a): Bedeutung von Selbstoffenbarungseffekten in quasisozialen Beziehungen mit auf generativer KI basierten Systemen in Settings von Onlineberatung und-therapie. In: *E-Beratungsjournal* 20 (1).

Linnemann, Gesa; Löhe, Julian; Rottkemper, Beate (2024b): Mensch-KI-Interaktion in der Kinder- und Jugendhilfe. In: Jens Albrecht, Martin Brinkmann, Jennifer Burghardt, Monika Feist-Ortmanns, Hans Jürgen Frischhut, Daniel Hahn et al.: Künstliche Intelligenz in der Kinder- und Jugendhilfe. 1. Auflage. Hg. v. Michael Macsenaere. München: Ernst Reinhardt Verlag, S. 87–93.

Linnemann, Gesa Alena; Löhe, Julian; Rottkemper, Beate (2023): Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit. In: *Soz Passagen* 15 (1), S. 197–211. DOI: 10.1007/s12592-023-00455-7.

Lippa, Barbara (2024): KI-gestützte Assistenzsysteme für die Teilhabe von Menschen mit Behinderungen am Arbeitsleben. In: Helmut Kreidenweis (Hg.): KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis. 1. Auflage. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG (Edition Sozialwirtschaft, 47), S. 65–80.

Löhe, Julian (2024): KI-Anwendung in der Kinder- und Jugendhilfe. In: Helmut Kreidenweis (Hg.): KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis. 1. Auflage. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG (Edition Sozialwirtschaft, 47).

Lotze, Netaya (2018): Zur sprachlichen Interaktion mit Chatbots – Eine linguistische Perspektive. Unter Mitarbeit von mediarep.org, Theo Hug und Günther Pallaver: Innsbruck University Press.

Lumma, W. (1991): Berichte von der Stange - in der Bewährungshilfe. In: Bernhard Meyer (Hg.): Hilfe vom Bildschirm: Computer in der sozialen Arbeit. Freiburg: Bernhard Meyer Verlag, S. 92–105.

Lupetti, Maria Luce; Murray-Rust, Dave (2024): (Un)making AI Magic: A Design Taxonomy. In: Florian Floyd Mueller, Penny Kyburz, Julie R. Williamson, Corina Sas, Max L. Wilson, Phoebe Toups Dugas und Irina Shklovski (Hg.): Proceedings of the CHI Conference on Human Factors in Computing Systems. CHI '24: CHI Conference on Human Factors in Computing Systems. Honolulu HI USA, 11 05 2024 16 05 2024. New York, NY, USA: ACM, S. 1–21.

Sartori, Laura, Bocca, Giulia (2023): Minding the gap(s): Public perceptions of AI and socio-technical imaginaries. AI & Society, 38, 443-458

Macsenaere, Michael (2024): Grundlagen Künstlicher Intelligenz (KI) in der Kinder- und Jugendhilfe. In: Jens Albrecht, Martin Brinkmann, Jennifer Burghardt, Monika Feist-Ortmanns, Hans Jürgen Frischhut, Daniel Hahn et al.: Künstliche Intelligenz in der Kinder- und Jugendhilfe. 1. Auflage. Hg. v. Michael Macsenaere. München: Ernst Reinhardt Verlag, S. 11–48.

McCain, Miles; Linthicum, Ryn; Lubinski, Chloe; Tamkin, Alex; Huang, Saffron; Stern, Michael et al. (2025): How People Use Claude for Support, Advice, and Companionship. Online verfügbar unter https://www.anthropic.com/news/how-people-use-claude-for-support-advice-and-companionship.

Mehlich, Harald (1997): Software im Sozialwesen: Anwendungsschwerpunkte und Auswirkungen. In: *Theorie und Praxis der sozialen Arbeit* (3), S. 10–18.

Parasuraman, Raja; Manzey, Dietrich H. (2010): Complacency and bias in human use of automation: an attentional integration. In: *Human factors* 52 (3), S. 381–410. DOI: 10.1177/0018720810376055.

Pelzl, Jan (2025): IT-Sicherheit und Datenschutz im Kontext von KI-Sprachmodellen. In: Gesa Linnemann, Julian Löhe und Beate Rottkemper (Hg.): Künstliche Intelligenz in der Sozialen Arbeit: Beltz Juventa, 204-220.

Phang, Jason; Lampe, Michael; Ahmad, Lama; Agarwal, Sandhini; Fang, Cathy Mengying; Liu, Auren R. et al. (2025): Investigating Affective Use and Emotional Well-being on ChatGPT.

Plafky, Christina; Kratz, Norbert; Kuck, André; Frischhut, Hans (2022): KI-basierte Entscheidungsunterstützung in der Praxis Sozialer Arbeit. In: *4* 74 (3). DOI: 10.2378/uj2022.art16d.

Plafky, Christina; Purandare, Mitra; Plattner, Benjamin (2025): Aktennotizerstellung in der Sozialen Arbeit durch Künstliche Intelligenz –Erkenntnisse aus einem Mixed-Method-Forschungsprojekt. In: Gesa Linnemann, Julian Löhe und Beate Rottkemper (Hg.): Künstliche Intelligenz in der Sozialen Arbeit: Beltz Juventa, S. 190–203.

Plafky, Christina S. (2025): Entscheidungsunterstützung in der Sozialen Arbeit durch prädiktive Analytik. In: Deutscher Verein für öffentliche und private Fürsorge e.V. (Hg.): Digitalisierung und KI in der Sozialen Arbeit. Ausgabe 3/2025 - Archiv für Wissenschaft und Praxis der Sozialen Arbeit: Lambertus-Verlag, S. 36–39.

Rammert, Werner (2010): Die Innovationen der Gesellschaft. In: Soziale Innovation. Auf dem Weg zu einem postindustriellen Innovationsparadigma. 1. Aufl. Wiesbaden: VS Verl. für Sozialwiss (Dortmunder Beiträge zur Sozialforschung), S. 21–51.

Reichmann, Ute (2016): Schreiben und Dokumentieren in der Sozialen Arbeit. Struktur, Orientierung und Reflexion für die berufliche Praxis. Opladen, Berlin, Toronto: Verlag Barbara Budrich (utb-studi-e-book, 4579).

Rüggeberg, Harald; Burmeister, Kjell (2008): Innovationsprozesse in kleinen und mittleren Unternehmen. 06. Aufl. Berlin (Working Papers of the Institute of Management Berlin at the Berlin School of Economics, Paper No. 41).

Scharf, Anna; Hauer, Franziska; Schultz, Maximilian; Haug, Sonja; Weber, Karsten (2023): Künstliche Intelligenz in Nichtregierungsorganisationen: Ein Scoping Review über ein junges Forschungsfeld. 1.Arbeitspapier.

Schmidberger, M. (2024): KI in der Dokumentation: Sprachsteuerung und Spracherfassung. In: Helmut Kreidenweis (Hg.): KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis. 1. Auflage. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG (Edition Sozialwirtschaft, 47).

Schultz, Maximilian; Hauer, Franziska; Haug, Sonja; Weber, Karsten (2025): Künstliche Intelligenz in Nichtregierungsorganisationen: In-depth-Interviews. 4.Arbeitspapier.

Schultz, Maximilian; Scharf, Anna; Hauer, Franziska; Haug, Sonja; Weber, Karsten (2024): Künstliche Intelligenz in Nichtregierungsorganisationen: Explorative Studie in einem neuen Forschungsfeld. 2.Arbeitspapier.

Schurz, R. (1991): Computerisierung des Denkens? Zuwendung, Individualität und Funktionalität. In: Bernhard Meyer (Hg.): Hilfe vom Bildschirm: Computer in der sozialen Arbeit. Freiburg: Bernhard Meyer Verlag, S. 22–34.

Searle, John R. (1980): Minds, brains, and programs. In: *Behav Brain Sci* 3 (3), S. 417–424. DOI: 10.1017/S0140525X00005756.

Seelmeyer, Udo (2020): Big Data & Künstliche Intelligenz – Neue Anforderungen an die Fachlichkeit in sozialen Berufen. In: *BdW* 167 (3), S. 95–98. DOI: 10.5771/0340-8574-2020-3-95.

Steiner, Olivier; Tschopp, Dominik (2022): Künstliche Intelligenz in der Sozialen Arbeit. In: *Sozial Extra* 46 (6), S. 466–471. DOI: 10.1007/s12054-022-00546-4.

Stoll, Dragan; Wehrli, Samuel; Lätsch, David (2025): Case reports unlocked: Harnessing large language models to advance research on child maltreatment. In: *Child abuse & neglect* 160, S. 107202. DOI: 10.1016/j.chiabu.2024.107202.

Strauss, Anselm (1978): A Social World Perspective. In: Norman K. Denzin (Hg.): Studies in Symbolic Interaction (Vol. 1). An Annual Compilation of Research. Greenwich, Connecticut, S. 119–128.

Strauss, Anselm L. (1993): Continual permutations of action. Hawthorne, NY: Aldine De Gruyter (Communication and social order).

Sutton, Reed T.; Pincock, David; Baumgart, Daniel C.; Sadowski, Daniel C.; Fedorak, Richard N.; Kroeker, Karen I. (2020): An overview of clinical decision support systems: benefits, risks, and strategies for success. In: *npj Digit. Med.* 3 (1), S. 17. DOI: 10.1038/s41746-020-0221-y.

Trommsdorff, Volker; Gizycki, Vittoria von (Hg.) (1995): Fallstudien zum Innovationsmarketing. München: Vahlen.

Vahs, Dietmar; Brem, Alexander (2015): Innovationsmanagement. Von der Idee zur erfolgreichen Vermarktung. 5. überarbeitete Auflage 2015. Stuttgart: Schäffer-Poeschel.

Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gomez, Aidan N. et al. (2017): Attention Is All You Need.

Vowels, Laura M.; Francois-Walcott, Rachel R.R.; Darwiche, Joëlle (2024): AI in relationship counselling: Evaluating ChatGPT's therapeutic capabilities in providing relationship advice. In: *Computers in Human Behavior: Artificial Humans* 2 (2), S. 100078. DOI: 10.1016/j.chbah.2024.100078.

WBGU (2019): Unsere gemeinsame digitale Zukunft. Online verfügbar unter https://digital.zlb.de/viewer/fulltext/34277218/1/.

Zapf, Wolfgang (1989): Über soziale Innovationen. In: *Soziale Welt* 40 Jg. (Nr. 1/2), S. 170–183.

Zerilli, John; Knott, Alistair; Maclaurin, James; Gavaghan, Colin (2019): Algorithmic Decision-Making and the Control Problem. In: *Minds & Machines* 29 (4), S. 555–578. DOI: 10.1007/s11023-019-09513-7.

Züger, Theresa; Faßbender, Judith; Kuper, Freia; Nenno, Sami; Katzy-Reinshagen, Anna; Kühnlein, Irina (2022): Civic coding. Grundlagen und empirische Einblicke zur Unterstützung gemeinwohlorientierter KI. Berlin.