

Heft 1/2026

Soziale Arbeit und KI

Mit Beiträgen von

Fabian Beloch, Angelika Beranek, Edeltraud Botzum, Cordula Endter, Franziska Erdtmann, Max Haberstroh, Andrea Hahn, Flemming Hansen, André Heinz, Josephine Jahn, Sabine Klinger, Dennis Klinkhammer, Pascal Krimmer, Sebastian Krumm, Anne Martin, Jonathan May, Florian Müller, Marcel Pietsch, Thomas Pudelko, Julia Rehn, Adrian Roeske, Susanne Sackl-Sharif, Philipp Seitz, Daniela Cornelia Stix, Tina Wenz, Dennis Wittmer und Marc Witzel

in den Rubriken

Editorial

Das Thema

Aktuelle Entwicklungen in der Praxis

Lehre und Digitalisierung

Forschungsnotizen

Diskursraum

Abschlussarbeiten

Heft 1, Jahrgang 2, 2026

Soziale Arbeit und KI

Herausgegeben von Adrian Roeske, Marc Witzel, Angelika Beranek,
Cordula Endter, Max Haberstroh, Thomas Pudelko und Daniela
Cornelia Stix

This work is licensed under a Creative Commons
Attribution 4.0 International License
<http://creativecommons.org/licenses/by/4.0/>



Roeske, A., Witzel, M., Beranek, A., Endter, C., Haberstroh, M., Pudelko, T., & Stix, D.C. (Hrsg.). (2026). *Soziale Arbeit und KI: 2(1)*. OAPublishing Collective. <https://doi.org/10.21240/dimesoz/2026/1>

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Titel: Soziale Arbeit und KI
Herausgebende: Adrian Roeske, Marc Witzel, Angelika Beranek, Cordula Endter, Max Haberstroh,
Thomas Pudelko und Daniela Cornelia Stix
Produktion: Klaus Rummeler
Satz: büro mn, Bielefeld
Verlag: OAPublishing Collective Genossenschaft
Herstellung: Books on Demand GmbH, Norderstedt, Deutschland
Nummer: 1
Jahr: 2026
Jahrgang: 2



ISBN (print): 978-3-03978-181-2
ISBN (online): 978-3-03978-182-9
DOI-URL: <https://doi.org/10.21240/dimesoz/2026/1>
ISSN (online): 2813-947X
ISSN (print): 2813-9488

© Zürich, 31. Jan 2026. Creative Commons Attribution 4.0 International (CC BY 4.0), alle Rechte liegen
bei den Autor:innen

Das Werk und jeder seiner Beiträge, sind urheberrechtlich geschützt. Sie dürfen das Material in jed-
wedem Format oder Medium vervielfältigen und weiterverbreiten, das Material remixen, verändern
und darauf aufbauen und zwar für beliebige Zwecke. Unter folgenden Bedingungen: Namensnennung
– Sie müssen angemessene Urheber- und Rechteangaben machen, einen Link zur Lizenz einschl.
Original-DOI beifügen und angeben, ob Änderungen vorgenommen wurden. Diese Angaben müssen
den üblichen wissenschaftlichen Zitierformaten folgen.

Inhalt

Editorial

Adrian Roeske, Marc Witzel, Angelika Beranek, Cordula Endter,
Max Haberstroh, Thomas Pudelko und Daniela Cornelia Stix

Editorial. Soziale Arbeit und KI **i**

Das Thema

Fabian Beloch und Anne Martin

**Generative KI in der Beratung. Empirische Befunde
mit Blick auf Nutzungsbedingungen von Chatbots** **1**

Dennis Wittmer

**Ethische und rechtliche Dimensionen des Einsatzes
von KI-Technologie in der Sozialen Arbeit und
Leitlinien zur informierten Einwilligung der Klientel** **21**

Andrea Hahn und Flemming Hansen

**Algorithmen haben keine Empathie. KI als
Kommunikationsbrücke und Reflexionshilfe in der
Sozialen Arbeit** **45**

Dennis Klinkhammer

Erklärbare Künstliche Intelligenz für die Soziale Arbeit **55**

Aktuelle Entwicklungen in der Praxis

Tina Wenz

**Künstliche Intelligenz und digitale
Teilhabe. Potenziale, Spannungsfelder und
Gestaltungsanforderungen für Menschen mit
Beeinträchtigungen in der Sozialen Arbeit** **75**

Julia Rehn

**Mensch vs. Chatbot – Resonanzraum oder
Echokammer des Selbst? Eine qualitative
Einzelfallstudie** **91**

Lehre und Digitalisierung

Philipp Seitz

**Künstliche Intelligenz und kritisch-reflexive Haltung
im Studium der Sozialen Arbeit** 113

Florian Müller und Edeltraud Botzum

**Künstliche Intelligenz im Peer-Mentoring.
Reflexionen zur Förderung digitaler Future Skills
im Studiengang Soziale Arbeit** 131

Forschungsnotizen

Marcel Pietsch, Pascal Krimmer und André Heinz

**KI als Antwort auf den sozialprofessionellen
Fachkräftemangel? Eine Verbindung
professionstheoretischer Überlegungen mit
empirischen Perspektiven
aus dem Caritaspanel 2024** 139

Sabine Klinger, Susanne Sackl-Sharif und Josephine Jahn

**Künstliche Intelligenz in der Sozialen Arbeit mit
Kindern und Jugendlichen. Eine komparative Analyse
der Sicht- und Nutzungsweisen von Fachkräften aus
Österreich und Deutschland** 159

Diskursraum

André Heinz und Marcel Pietsch

**„Ceci n'est pas une réalité“ – Zur Dekonstruktion
von Wirklichkeit in einer magrittesken Realität von
Filterblasen und KI-Algorithmen** 185

Abschlussarbeiten

Jonathan May

Maschinelle Übersetzung in der Kommunikation mit unbegleiteten Minderjährigen mit Fluchterfahrung. Zwischen Inklusion, Exklusion und Teilhabe 201

Franziska Erdtmann

Die Bedeutung von Peer Support bei der Bewältigung belastender Online-Erfahrungen von Mädchen* 207

Sebastian Krumm

KI und Soziale Arbeit. Eine Studie zur Akzeptanz des Technologieeinsatzes unter Fachkräften der Sozialen Arbeit 215

Adrian Roeske¹, Marc Witzel², Angelika Beranek³, Cordula Endter⁴,
Max Haberstroh⁵, Thomas Pudelko⁶ und Daniela Cornelia Stix⁵

Editorial

Soziale Arbeit und KI

- 1 Institut für Informationsmanagement Bremen
- 2 Evangelische Hochschule Dresden
- 3 Hochschule München
- 4 Katholische Hochschule für Sozialwesen Berlin
- 5 Brandenburgische Technische Universität Cottbus-Senftenberg
- 6 Hochschule für Soziale Arbeit und Pädagogik, Berlin

Sogenannte Künstliche Intelligenz¹ (KI) ist aktuell eines der relevantesten Themen mit Bezug zur Digitalisierung, weshalb sich zahlreiche Publikationen finden lassen. Im Kontext der Sozialen Arbeit seien an dieser Stelle beispielhaft folgende Veröffentlichungen genannt: Linnemann et al., 2025; Feist-Ortmanns & Macsenaere, 2025; Plafky & Badertscher, 2025; Kieslinger, 2024; Beranek, 2025. Die Beiträge zum Themenfeld Künstliche Intelligenz weisen mehrere Aspekte auf, die in den aufgezählten Publikationen fokussiert werden. Dabei wird der Einsatz algorithmischer Systeme diskutiert zur (Unterstützung von) Entscheidungsfindung, z. B. durch Predictive Risk Modelling in der Kinder- und Jugendhilfe (vgl. Plafky, 2025; Burghardt et al., 2024), in der Beratung(sausbildung) (vgl. Engelhardt, 2024; Lehmann et al., 2021) oder im Hinblick auf die Lehre in Studiengängen der Sozialen Arbeit (vgl. Engelhardt & Ley, 2025; Witter et al., 2024).

Mit der Veröffentlichung von ChatGPT 3.5 im November 2022 hat die Diskussion um die sogenannte Künstliche Intelligenz gesellschaftlich und auch innerhalb der Sozialen Arbeit in Disziplin und Profession eine starke Zunahme erfahren. Neben den öffentlichen Diskursen, die zumindest in Teilen zwischen Heilsversprechen und apokalyptischen Befürchtungen

1 Mit der Formulierung „sogenannte Künstliche Intelligenz“ weisen wir auf die Spannungsfelder um den Intelligenzbegriff hin, die im Hinblick auf KI immer wieder thematisiert werden. Beispielhaft sei auf die begriffliche Differenz zwischen „natürlicher“ Intelligenz und KI verwiesen.

Roeske, A., Witzel, M., Beranek, A., Endter, C., Haberstroh, M., Pudelko, T., & Stix, D.C. (2026). Editorial. Soziale Arbeit und KI. *DiMe-Sozial*, 2(1), i–viii. <https://doi.org/10.21240/dimesoz/2026/1/138>

schwanken, findet ein fachlicher Diskurs statt, der sich im Wesentlichen auf das transformatorische Potenzial von sogenannter KI für das fachliche und organisationale Handeln der Sozialen Arbeit bezieht.

Auch in der Redaktion der *DiMe-Sozial* haben wir im Rahmen des Gründungsprozesses das Thema Künstliche Intelligenz bereits diskutiert und als relevant identifiziert. Zugleich zeichnete sich ab, dass sich aktuell viele Periodika diesem Thema zuwenden. Aus diesem Grund haben wir uns dazu entschlossen, erst mit Heft 3 das Thema „Soziale Arbeit und KI“ als Schwerpunkt festzulegen. Trotz der Wartezeit zeigte sich anhand der Menge der eingereichten Beitragsvorschläge ein reges Interesse an diesem Schwerpunkt.

Der Diskurs um das Verhältnis von Sozialer Arbeit und sogenannter Künstlicher Intelligenz erhält eine besondere Relevanz durch zwei spezifische Aspekte: Zum einen scheinen interaktive und kommunikative Prozesse – und damit spezifische Erbringungsformen personenbezogener Dienstleistungen – technologisierbar. Zum anderen wird statistischen Entscheidungsprozessen, die unter Einbezug komplexer Daten und umfassender Datenmengen stattfinden, eine hohe Relevanz zugestanden. Beides verweist auf zentrale Kriterien professionellen Handelns in der Sozialen Arbeit, nämlich insbesondere dort, wo sie – wenn auch in Form zu kritisierender Utopien – in konkreten Erbringungsprozessen bspw. in Form von digitalen Assistenten gedacht werden (vgl. Schulze, 2022).

Soziale Arbeit und KI – Die Frage nach der Technologisierbarkeit des Subjekts

Beide genannten Aspekte des Diskurses um Soziale Arbeit und sogenannter Künstlicher Intelligenz verweisen darauf, dass bisher herrschende Gewissheiten innerhalb der Profession herausgefordert werden.

Sowohl mit Berücksichtigung von Daten in der Entscheidungsfindung, in der kommunikativen Arbeit *am* Fall – im Sinne eines Zusammenwirkens von Fachkräften in kasuistischen Verstehensprozessen – als auch in der kommunikativen Arbeit *im* Fall – im Sinne der Arbeit mit den Adressat:innen und weiteren beteiligten Akteur:innen – stellen die verfügbaren Werkzeuge, ebenso wie der Diskurs über diese, eine Infragestellung bisheriger Sicherheiten dar.

Würden die reflexiven und kommunikativen Anteile als zentral für das professionelle Handeln (vgl. Dewe & Otto, 2012; von Spiegel, 2021) verstanden, so scheint auch dies im Angesicht neuer Möglichkeiten der Verarbeitung großer Datenmengen sowie des *Verstehens* menschlicher Sprache nun technologisierbar.

Vor diesem Hintergrund empfiehlt sich die Relektüre des zentralen Aufsatzes zum Technologiedefizit in der Sozialen Arbeit von Luhmann und Schorr (1979). Beide argumentieren mit der kaum zu bewältigenden Flut an Informationen innerhalb pädagogischer Prozesse. Zugleich stellen sie jedoch das Subjekt sowie dessen Veränderung im Sinne des People Processing (Hasenfeld, 1972) klar in den Mittelpunkt ihrer Argumentation der Nicht-Technologisierung pädagogischer Prozesse. Diese entsteht gerade durch die Selbstreferenzialität des Subjekts und nicht durch die Überforderung der vorliegenden Informationen in der Kommunikation. Subjekte selbst stehen im Zentrum des professionellen Handelns (vgl. Winkler, 2021). Als solche zeichnen sie sich dadurch aus, dass sie in selbsttätiger Auseinandersetzung mit der sie umgebenden Welt agieren. Folglich sind Subjekte im Sinne rationaler Prozesse kausal nicht zu beeinflussen bzw. nicht in technologisierbaren Akten veränderbar.

Zugleich lassen sich die Wirkungen und die Bedeutung dieser technologischen Entwicklungen bisher tatsächlich kaum überblicken. Es ist unklar, in welchem Maße Adressat:innen Sozialer Arbeit Unterstützung mithilfe von auf Sprachmodellen basierenden Chatbots suchen und finden (Linnemann et al., 2024). Weiterhin gilt es zu untersuchen, inwiefern es gelingt, professionelles Handeln stärker zu quantifizieren und auf diese Kontroll-, Automatisierungs- und letztlich Rationalisierungsmomente zuzuführen, die die Unterwerfung unter betriebswirtschaftliche Logiken in der Sozialen Arbeit vorantreiben (vgl. hierzu grundlegend Pasquinelli, 2024).

Vor dem Hintergrund dieser grob skizzierten Überlegungen möchte diese Ausgabe der *DiMe-Sozial* einen Beitrag leisten zur Aufklärung über das Verhältnis Soziale Arbeit und sogenannter Künstlicher Intelligenz. Dazu lagen den Herausgeber:innen nach Ende der Einreichungsfrist insgesamt 30 Beitragsvorschläge vor. Aufgrund dieser hohen Zahl sowie deren spannenden und vielfältigen Bearbeitungen umfasst das Heft 1/2026 der *DiMe-Sozial* in der Rubrik „Das Thema“ sowie in den weiteren Rubriken einen Beitrag mehr als geplant. Das Heftthema ist zudem durchgehend am Titel orientiert, da die Einreichungen in allen Rubriken das Themenfeld Soziale Arbeit und Künstliche Intelligenz aufgegriffen haben.

Zu den Beiträgen des Hefts

Die Beiträge dieser Ausgabe fokussieren nahezu durchgehend das Verhältnis von sogenannter Künstlicher Intelligenz und Sozialer Arbeit. In der Rubrik „Das Thema“ erfolgt die Thematisierung dieses Verhältnisses systematisch in vier unterschiedlichen Beiträgen. Diese greifen in empirischer, ethischer und rechtlicher, systematischer sowie konzeptioneller Weise unterschiedliche Dimensionen des breiten Diskurses um sogenannte Künstliche Intelligenz auf.

Fabian Beloch und Anne Martin (2026) präsentieren in ihrem Beitrag empirische Befunde mit Blick auf Nutzungsbedingungen von Chatbots in der Sozial- und Schuldenberatung. Dabei zeigt sich, dass sowohl auf Seiten der Adressat:innen als auch auf Seiten der Fachkräfte Kompetenzen im Umgang mit den neuen Technologien ausgebildet werden müssen, um diese in spezifischen Beratungsangeboten einzusetzen.

Dennis Wittmer (2026) diskutiert die ethischen und rechtlichen Aspekte des Einsatzes sogenannter Künstlicher Intelligenz unter Berücksichtigung des EU AI-Acts sowie der DSGVO. Dabei werden die hohen ethischen Anforderungen beim Technologieeinsatz in der Sozialen Arbeit hervorgehoben und in Relation zu fachlichen Anforderungen gestellt.

Andrea Hahn und Flemming Hansen (2026) setzen sich in ihrem Beitrag systematisch mit der Frage des Einsatzes von Chatbots als *Kommunikationsbrücke* auseinander, die sowohl im Kontakt mit Adressat:innen wie in der Fallbearbeitung zum Einsatz kommen kann. Mitzudenken ist allerdings, dass sie als Werkzeuge zu betrachten sind, da ihnen Empathie, implizites Wissen und ethische Urteilskraft fehlen.

Dennis Klinkhammer (2026) befasst sich in seinem Beitrag mit der Notwendigkeit professioneller KI-Kompetenzen in der Sozialen Arbeit. Gerade um KI als Werkzeug in seiner Reichweite und bezüglich seiner Einsatzmöglichkeiten einschätzen zu können, bedarf es einer dezidierten Beschäftigung mit den statistischen Grundlagen ihrer Funktionsweisen, um schließlich den Anspruch erklärbarer KI zu erfüllen.

In der Rubrik „Aktuelle Entwicklungen in der Praxis“ finden sich in dieser Ausgabe zwei Beiträge. Zum einen diskutiert **Tina Wenz** (2026) die Potenziale und Herausforderungen sogenannter Künstlicher Intelligenz im Kontext der Teilhabe von Menschen mit Beeinträchtigungen und erarbeitet hierzu zehn Thesen. Zum anderen stellt **Julia Rehn** (2026) die Ergebnisse einer Einzelfallstudie vor, welche die parallele Inanspruchnahme von professioneller Beratung und Chatbots untersucht. Dabei wird vor allem das Verständnis von Kommunikation als Resonanzraum zentral, das sich in erster Linie als ein in zwischenmenschlicher Kommunikation erfahrbares Phänomen verstehen lässt.

In der Rubrik „Lehre und Digitalisierung“ untersucht der erste Beitrag von **Philipp Seitz** (2026) die Ausbildung einer kritisch-reflexiven Haltung im Studium und greift dafür die Figur der *Denkbewegung* auf. Seitz' Überlegungen thematisieren die Möglichkeit, generative Künstliche Intelligenz im Rahmen solcher Denkbewegungen zu nutzen. Im zweiten Beitrag dieser Rubrik diskutieren **Florian Müller und Edeltraud Botzum** (2026), inwiefern KI-gestütztes Arbeiten zur Förderung digitaler Schlüsselkompetenzen bei Studierenden der Sozialen Arbeit beitragen kann. Dabei steht insbesondere der Einsatz von sogenannter Künstlicher Intelligenz in

Verbindung mit Peer-Mentoring im Fokus. Ziel ist hier die Vorbereitung der Studierenden auf eine technologisierte Arbeits- und Bildungswelt, auch in der Sozialen Arbeit.

Der erste Beitrag in der Rubrik „Forschungsnotizen“ von **Marcel Pietsch, Pascal Krimmer und André Heinz (2026)** geht vor dem Hintergrund der Gleichzeitigkeit von Fachkräftemangel und zunehmendem Einsatz sogenannter Künstlicher Intelligenz der Frage nach, inwieweit der betriebliche Einsatz dieser Technologien eine Reaktionsmöglichkeit bei wachsenden Personallücken darstellt. Empirisch wird diese Diskussion durch die Ergebnisse des Caritaspanels (2024) gestützt, in welchem auch nach den Möglichkeiten des Einsatzes sogenannter generativer Künstlicher Intelligenz gefragt wurde. Im zweiten Beitrag dieser Rubrik präsentieren **Sabine Klinger, Susanne Sackl-Sharif und Josephine Jahn (2026)** die Ergebnisse einer internationalen Vergleichsstudie zu den Sicht- und Nutzungsweisen von Fachkräften in der Kinder- und Jugendarbeit sowie der Schulsozialarbeit in Deutschland und Österreich. Die Ergebnisse verweisen auf die notwendige Beschäftigung im Hinblick auf den Kompetenzerwerb der Fachkräfte sowie organisationaler Rahmenbedingungen.

In der Rubrik „Diskursraum“ stellen **André Heinz und Marcel Pietsch (2026)** Überlegungen zur Diskussion, die das Verhältnis von realen Erfahrungen und digitalen Repräsentationen aufgreifen. Dabei werden insbesondere die medialen Repräsentationen zum Gegenstand genommen, um sich mittels ästhetischer Bildung mit Fragen des Verhältnisses von Realität und Darstellung auseinanderzusetzen.

In der Rubrik „Abschlussarbeiten“ stellen **Franziska Erdtmann (2026), Jonathan May (2026)** und **Sebastian Krumm (2026)** ihre Arbeiten in drei Beiträgen vor. Die Bachelorarbeit von May diskutiert den Einsatz maschineller Übersetzung im Rahmen professioneller Arbeit mit unbegleiteten Minderjährigen mit Fluchterfahrung. Dabei wird einerseits die Ermöglichung von Kommunikation diskutiert und andererseits die problematischen Effekte maschineller Übersetzungen. Erdtmann erarbeitet in ihrer Masterarbeit die Bedeutung von Peer-Support für junge Frauen vor dem Hintergrund sexualisierter Belästigung, Hate Speech und Cybermobbing. Fokussiert werden dabei die Kompetenzen von Peers als Ansprechpartner:innen. Sebastian Krumm präsentiert in seiner Masterarbeit Ergebnisse aus einer empirischen Erhebung zur Akzeptanz des Technologieeinsatz bei Fachkräften in der Sozialen Arbeit.

Aus der Redaktion wünschen wir eine anregende Lektüre!

Literatur

- Beloch, F., & Martin, A. (2026). Generative KI in der Beratung: Empirische Befunde mit Blick auf Nutzungsbedingungen von Chatbots. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 1–20. <https://doi.org/10.21240/dimesoz/2026/1/57>
- Beranek, A. (2025). Der Einsatz digitaler Tools und Künstlicher Intelligenz in der Sozialen Arbeit. *Archiv für Wissenschaft und Praxis der Sozialen Arbeit*, 3, 4–14.
- Burghardt, J., Lehmann, R., Reder, M., Koska, C., Kraus, M., & Müller, N. (2024). Kann Künstliche Intelligenz sozialarbeiterische Entscheidungsprozesse unterstützen? Ethik und digitale Operationalisierung im Feld der Kindeswohlgefährdung. *unsere jugend*, 76(7+8), 300–310. <https://doi.org/10.2378/uj2024.art40d>
- Dewe, B., & Otto, H.-U. (2012). Reflexive Sozialpädagogik. In W. Thole (Hrsg.), *Grundriss Soziale Arbeit* (S. 197–217). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-94311-4_8
- Engelhardt, E., & Ley, T. (2025). Voll (dia)logisch? Ein Werkstattbericht über den Einsatz von generativer KI in der Hochschulbildung für Soziale Arbeit. Curriculare Überlegungen und veränderte Akteurskonstellationen. In M. Wunder & A. Giercke-Ungermann (Hrsg.), *Digitalisierung in der Hochschulbildung für Soziale Arbeit* (S. 251–264). Verlag Julius Klinkhardt. <https://doi.org/10.35468/6166-16>
- Engelhardt, E. M. (2024). Einsatz generativer KI in der Beratungsausbildung Lehrmethoden und Praxisbeispiele aus dem Hochschulkontext. In S. Neumaier, M. Dörr, & E. Botzum (Hrsg.), *Praxishandbuch Digitale Projekte in der Sozialen Arbeit* (S. 63–79). Beltz Juventa. <https://doi.org/10.3262/978-3-7799-7769-8>
- Erdtmann, F. (2026). Die Bedeutung von Peer Support bei der Bewältigung belastender Online-Erfahrungen von Mädchen*. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 207–213. <https://doi.org/10.21240/dimesoz/2026/1/96>
- Feist-Ortmanns, M., & Macsenaere, M. (2025). KI in der Kinder- und Jugendhilfe. In G. Linnemann, J. Löhe, & B. Rottkemper (Hrsg.), *Künstliche Intelligenz in der Sozialen Arbeit: Grundlagen für Theorie und Praxis* (S. 90–101). Beltz Juventa. <https://doi.org/10.3262/978-3-7799-8562-4>
- Hahn, A., & Hansen, F. (2026). Algorithmen haben keine Empathie: KI als Kommunikationsbrücke und Reflexionshilfe in der Sozialen Arbeit. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 45–54. <https://doi.org/10.21240/dimesoz/2026/1/78>
- Hasenfeld, Y. (1972). People Processing Organizations: An Exchange Approach. *American Sociological Review*, 37(3), 256. <https://doi.org/10.2307/2093466>

- Heinz, A., & Pietsch, M. (2026). „Ceci n'est pas une réalité“ – Zur Dekonstruktion von Wirklichkeit in einer magrittesken Realität von Filterblasen und KI-Algorithmen. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 185–200. <https://doi.org/10.21240/dimesoz/2026/1/65>
- Kieslinger, D. (2024). Mit KI-Unterstützung zu einer inklusiven Kinder- und Jugendhilfe. *unsere jugend*, 76(7+8), 338–339. <https://doi.org/10.2378/uj2024.art43d>
- Klinger, S., Sackl-Sharif, S., & Jahn, J. (2026). Künstliche Intelligenz in der Sozialen Arbeit mit Kindern und Jugendlichen: Eine komparative Analyse der Sicht- und Nutzungsweisen von Fachkräften aus Österreich und Deutschland. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 159–183. <https://doi.org/10.21240/dimesoz/2026/1/59>
- Klinkhammer, D. (2026). Erklärbare Künstliche Intelligenz für die Soziale Arbeit. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 55–74. <https://doi.org/10.21240/dimesoz/2026/1/58>
- Krumm, S. (2026). KI und Soziale Arbeit: Eine Studie zur Akzeptanz des Technologieeinsatzes unter Fachkräften der Sozialen Arbeit. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 215–200. <https://doi.org/10.21240/dimesoz/2026/1/74>
- Lehmann, R., Albrecht, J., Domes, M., Petrlic, R., Bradl, M., Burghardt, J., Kiener, D., Stieler, M., Widerhold, J.-P., & Zauter, S. (2021). *Gutachten über die Einsatzmöglichkeiten von Künstlicher-Intelligenz-Software in aufsuchenden, digitalen Angeboten der Migrationsberatung*. TH Nürnberg; Institut für E-Beratung.
- Linnemann, G., Löhe, J., & Rottkemper, B. (2024). Bedeutung von Selbstoffenbarungseffekten in quasisozialen Beziehungen mit auf generativer KI basierten Systemen in Settings von Onlineberatung und -therapie. *e-beratungsjournal.net*, 20(1), 1–21. <https://doi.org/10.48341/9x1s-5y11>
- Linnemann, G., Löhe, J., & Rottkemper, B. (Hrsg.). (2025). *Künstliche Intelligenz in der Sozialen Arbeit: Grundlagen für Theorie und Praxis*. Beltz Juventa. <https://doi.org/10.3262/978-3-7799-8562-4>
- Luhmann, N., & Schorr, K. E. (1979). Das Technologiedefizit der Erziehung und die Pädagogik. *Zeitschrift für Pädagogik*, 25(3), 345–365.
- May, J. (2026). Maschinelle Übersetzung in der Kommunikation mit unbegleiteten Minderjährigen mit Fluchterfahrung: Zwischen Inklusion, Exklusion und Teilhabe. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 201–206. <https://doi.org/10.21240/dimesoz/2026/1/83>
- Müller, F., & Botzum, E. (2026). Künstliche Intelligenz im Peer-Mentoring: Reflexionen zur Förderung digitaler Future Skills im Studiengang Soziale Arbeit. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 131–138. <https://doi.org/10.21240/dimesoz/2026/1/85>

- Pasquinelli, M. (2024). *Das Auge des Meisters: Eine Sozialgeschichte Künstlicher Intelligenz* (K. Hermes, Übers.; 1. Auflage). UNRAST.
- Pietsch, M., Krimmer, P., & Heinz, A. (2026). KI als Antwort auf den sozialprofessionellen Fachkräftemangel? Eine Verbindung professionstheoretischer Überlegungen mit empirischen Perspektiven aus dem Caritaspanel 2024. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 139–158. <https://doi.org/10.21240/dimesoz/2026/1/70>
- Plafky, C. S. (2025). Entscheidungsunterstützung in der Sozialen Arbeit durch prädiktive Analytik. *Archiv für Wissenschaft und Praxis der Sozialen Arbeit*, 3, 36–39. <https://doi.org/10.1093/bjsw/bcv031>
- Plafky, C. S., & Badertscher, H. (2025). *Künstliche Intelligenz in der Sozialen Arbeit: Potenziale, Herausforderungen und Ethik im digitalen Zeitalter*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-49009-6>
- Rehn, J. (2026). Mensch vs. Chatbot – Resonanzraum oder Echokammer des Selbst? Eine qualitative Einzelfallstudie. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 91–112. <https://doi.org/10.21240/dimesoz/2026/1/137>
- Schulze, S. (2022). Künstliche Intelligenz – Warum sollte sie nur von Autobauern genutzt werden? *unsere jugend*, 74(3), 130–132. <https://doi.org/10.2378/uj2022.art18d>
- Seitz, P. (2026). Künstliche Intelligenz und kritisch-reflexive Haltung im Studium der Sozialen Arbeit. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 113–129. <https://doi.org/10.21240/dimesoz/2026/1/68>
- Spiegel, H. von (with Sturzenhecker, B.). (2021). *Methodisches Handeln in der Sozialen Arbeit: Grundlagen und Arbeitshilfen für die Praxis* (7., durchgesehene Auflage). Ernst Reinhardt Verlag.
- Wenz, T. (2026). Künstliche Intelligenz und digitale Teilhabe: Potenziale, Spannungsfelder und Gestaltungsanforderungen für Menschen mit Beeinträchtigungen in der Sozialen Arbeit. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 75–89. <https://doi.org/10.21240/dimesoz/2026/1/63>
- Winkler, M. (2021). *Eine Theorie der Sozialpädagogik: Neuausgabe mit einem neuen Nachwort*. Herausgegeben von Gaby Flösser und Marc Witzel (G. Flösser & M. Witzel, Hrsg.). Beltz Juventa.
- Witter, S., Meinhardt-Injac, B., Siemer, L., Späte, J., & Fachhochschule Potsdam. (2024). *ChatGPT im Studium der Sozialen Arbeit: Eine quantitative Studie zur Nutzung, Bewertung und Thematisierung in der Hochschule aus Studierendensicht*. <https://doi.org/10.34678/OPUS4-3382>
- Wittmer, D. (2026). Ethische und rechtliche Dimensionen des Einsatzes von KI-Technologie in der Sozialen Arbeit und Leitlinien zur informierten Einwilligung der Klientel. *DiMe-Sozial | Soziale Arbeit, Medien und Digitalisierung*, 2(1), 21–43. <https://doi.org/10.21240/dimesoz/2026/1/60>

Fabian Beloch¹ und Anne Martin²

Generative KI in der Beratung

Empirische Befunde mit Blick auf Nutzungsbedingungen von Chatbots

¹ Universität zu Köln

² TH Nürnberg, Institut für E-Beratung

Zusammenfassung

Die zunehmende Digitalisierung sozialer Dienstleistungen zeigt zentrale Fragen zur Teilhabe von Klient:innen sowie zur Professionalisierung von Fachkräften auf. Am Beispiel einer Untersuchung zur Sozial- und Schuldenberatung werden Unterstützungsbedarfe sowie Potenziale und Herausforderungen KI-gestützter Beratungsformate analysiert. Eine Erkenntnis ist, dass es individuelle Konzepte für die Gestaltung von Beratungen braucht, die sich an den Ressourcen und Unterstützungsbedarfen der Klient:innen orientieren. Gleichzeitig stehen Fachkräfte vor einer Rollen-Neudefinition im Sinne einer Professionalisierung vor dem Hintergrund der Digitalisierung der Beratung.

Schlüsselwörter: Generative KI, Chatbots, Beratung, digitale Medien, digitale Teilhabe, Unterstützungsbedarfe

Generative AI in Counselling. Empirical Findings With Regards to the Conditions of Use of Chatbots

Abstract

Despite the increasing digitization of social services raises key questions regarding the participation of clients and the professionalisation of specialists. Using the example of a study on social and debt counselling, the need for support as well as the potential and challenges of AI-supported counselling formats are analysed. One finding is that individual concepts are needed for the design of counselling sessions that are geared towards resources and needs of clients. At the same time, professionals are facing a redefinition of their roles in terms of professionalisation against the backdrop of the digitization of counselling.

Keywords: Generative AI, chatbots, counselling, digital media, digital participation, support needs

Beloch, F., & Martin A. (2026). Generative KI in der Beratung: empirische Befunde mit Blick auf Nutzungsbedingungen von Chatbots. *DiMe-Sozial*, 2(1), 1–20. <https://doi.org/10.21240/dimesoz/2026/1/57>

1. Einleitung

Generative Künstliche Intelligenz (KI) in Form von Chatbots findet seit einigen Jahren in der Sozialen Beratung zunehmend Verwendung, parallel dazu befassen sich verschiedene Forschungsprojekte mit dieser Thematik. Es folgen einige Beispiele: Das Potenzial von ChatGPT als eigenständiger Schwangerschaftsberater wurde getestet; dabei zeigten sich die Antworten als sachlich korrekt, werden aber oft als zu oberflächlich empfunden, was den Schluss nahelegt, dass menschliche Beratende besser seien (Kubitza, 2024). Mit dem Einsatz eines Chatbots als virtuelle:r Testklient:in wurden Chancen für die Beratungsausbildung aufgezeigt (Lippert et al., 2024). In der Studie von Martin & Pengel (2024) wurde der Einsatz von Chatbots in der Unterstützung Studierender im Selbststudium untersucht, mit dem Vorteil, dass generative KI eine menschenähnliche Interaktion gestalten kann und einen großen Möglichkeitsraum bietet.

Der aktuelle Diskurs zur Nutzung von KI in der Beratung ist von der fortschreitenden Digitalisierung und Mediatisierung sozialer Dienstleistungen geprägt. Zentrale Diskussionslinien betreffen dabei Fragen der fachlichen Kontrolle, ethischen Verantwortung und professionellen Einbettung solcher Systeme. Erste empirische Befunde zeigen zwar eine hohe Akzeptanz und potenzielle Entlastung von Fachkräften, betonen aber die Notwendigkeit einer reflektierten, sozialarbeiterisch fundierten Implementierung. Insgesamt zeigt sich ein sich dynamisch entwickelnder Forschungs- und Diskursstand, der zwischen technologischem Fortschritt und professionsethischer Verantwortung vermittelt (Lehmann, 2025). In ihrem kürzlich erschienenen Band *Künstliche Intelligenz in der Beratung* beleuchten Engelhardt & Kühne (2025) Forschungsdesiderata unter anderem hinsichtlich der Zugangsbedingungen marginalisierter Adressat:innen sozialer Beratung einerseits sowie der Auswirkungen auf die Rolle der Beratenden andererseits (S. 160f.). Kurzum gilt es daher, folgende Frage zu beantworten: Welche Bedingungen der Beratungspraxis sind beim Einsatz generativer KI in Form von Chatbots zu berücksichtigen?

Der vorliegende Beitrag bietet darauf Antworten anhand von drei Themenkomplexen: Zuerst werden die potenzielle Zielgruppe von Beratung und die professionell Beratenden vor dem Hintergrund ihrer Ausgangsbedingungen betrachtet. An der (Nicht-)Zusammenarbeit von professionellen Beratenden und ehrenamtlichen Mitarbeitenden, die selbst anteilige Beratungsaufgaben übernehmen (sollen), wird daraufhin erläutert, welche Reibungspunkte und professionelle Strategien bei der *Mitarbeit* von Chatbots als dritte Akteure entstehen können. Abschließend werden weitere Fragen abgeleitet, die Anknüpfungspunkte für zukünftige Forschungsarbeiten und die Erweiterung des Diskurses um KI in der Sozialen Arbeit ermöglichen sollen. Diskussionsgrundlage sind empirische

Erkenntnisse aus einem kürzlich abgeschlossenen Praxisforschungsprojekt¹ zur digitalen Teilhabe im Kontext hybrider Beratungsstrukturen (Beloch et al., 2025).

2. Ausgangsbedingungen der potenziellen Zielgruppe von Beratung

Inwieweit sich Chatbots zum Einsatz in der Praxis tatsächlich eignen, muss zunächst an den Ausgangsbedingungen derjenigen bemessen werden, die mit Beratungsangeboten erreicht werden sollen: Ratsuchende. Empirische Bevölkerungsstudien geben hier erste Hinweise. Laut der ARD/ZDF-Medienstudie nutzen 95 % der deutschen Bevölkerung das Internet generell; somit gelten 5 % als *Offliner:innen*, unter den Älteren sind dies sogar rund 25 % (ARD/ZDF, 2024, S. 3). Der D21-Digital-Index gibt parallel dazu an, dass nur die Hälfte der Bevölkerung über alle nötigen digitalen Basiskompetenzen verfügt, wobei Menschen mit hoher formaler Bildung besser abschneiden als Menschen mit niedriger formaler Bildung (Initiative D21, 2025, S. 21). Ein ähnliches Muster zeigt sich in Bezug auf KI-Anwendungen: Diese werden generell von 39 % der Bevölkerung genutzt, vergleichsweise seltener von Menschen mit niedriger formaler Bildung oder höheren Alters (Initiative D21, 2025, S. 18). Deutlich werden damit Tendenzen digitaler Ungleichheit, die mit Bezug auf Chatbots in der Beratung von besonderer Bedeutung sind, da die Adressierten Sozialer Arbeit als mehrfach vulnerabel und auch unter Digitalisierungsaspekten benachteiligt gelten (Iske & Kutscher, 2020).

Um empirisches Wissen über die Ausgangsbedingungen der potenziellen Zielgruppe von Beratung im Hinblick auf digitale Angebotsformate zu generieren, wurde eine quantitative Erhebung mit Nutzenden der Sozial- und Schuldenberatung durchgeführt. Diese Beratungsfelder eignen sich in besonderer Weise als Betrachtungsgegenstand im Kontext digitaler Ungleichheit, da beide im Kern die Lebenslage finanzieller Armut bearbeiten (van Rießen & Bleck, 2023). Vor allem in Armut lebende Menschen verfügen häufig nicht über ausreichenden Zugang zu digitalen Möglichkeiten (Bertsche & Como-Zipfel, 2023, S. 132). Mithilfe eines Pencil-and-Paper-Fragebogens wurden in den Beratungseinrichtungen Ende des Jahres 2022 insgesamt 165 Angebotsnutzende befragt. Die Stichprobe beinhaltet Menschen verschiedenen Geschlechts, Alters und sozialer sowie kultureller Herkunft und spiegelt die Diversität der Zielgruppe von Beratungsangeboten in vielen Punkten wider. Einschränkend ist anzumerken, dass in der Stichprobe nicht diejenigen Adressierten vorkommen, die ausschließlich die digitalen Angebotsformate der Einrichtungen nutzen. Der Fragebogen enthielt unter anderem Fragen zu Medienbesitz, Mediennutzung, medienbezogenen Fähigkeiten und Präferenzen in Bezug auf (digitale) Beratungsformate; die Ergebnisse werden im Folgenden dargestellt.

¹ Das Projekt wurde ermöglicht durch die Förderung der Sozialstiftung NRW.

Von Interesse sind mit Blick auf den Medienbesitz (Abbildung 1) hauptsächlich Endgeräte wie Desktop-PC, Laptop, Tablet und Smartphone, da durch sie der Zugriff auf und die Interaktion mit textbasierten KI-Anwendungen überhaupt erst möglich ist.

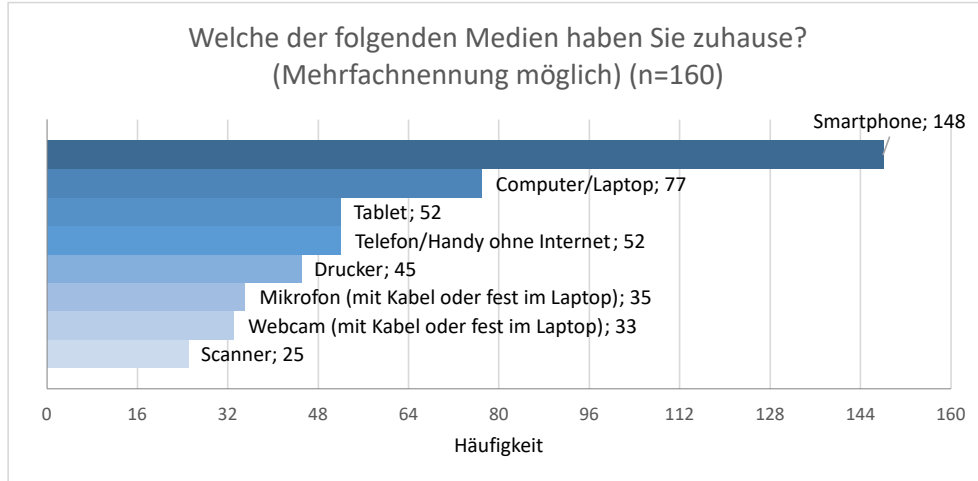


Abbildung 1: Medienbesitz der Ratsuchenden (Quelle: Beloch et al., 2025, S. 32).

Den Angaben der befragten Ratsuchenden zufolge besitzen 148 von 160 (etwa 93%) von ihnen ein Smartphone, fast die Hälfte Computer/Laptop und etwa ein Drittel ein Tablet. Keines der abgefragten digitalen Endgeräte besitzen dagegen 12 von 160 Befragte (7%). Diesbezüglich zeigt sich auch, dass der Besitz digitaler Medien allgemein und insbesondere des Smartphones von den älteren Befragten tendenziell seltener angegeben wird. Prinzipiell ist die Möglichkeit der Nutzung von Chatbots mit dem weitverbreiteten Vorhandensein des Smartphones also für fast alle Ratsuchenden gegeben. Allerdings müssen auch diejenigen Personengruppen mitgedacht werden, die keinen selbstbestimmten Zugang zu digitalen Medien haben.

Da der Medienbesitz allein noch wenig zum tatsächlichen Nutzungsverhalten aussagt, konnten die Ratsuchenden zusätzlich angeben, wie häufig sie bestimmte Geräte und Software-Anwendungen nutzen. Insgesamt zeigt sich bei ihnen eine vor allem auf private Kommunikation und Unterhaltung ausgerichtete Kommunikation. Die prävalentesten Medien und medialen Anwendungen sind demnach: Smartphone, E-Mail, Suchmaschinen wie Google/Bing, Messenger wie WhatsApp, Social Media-Anwendungen wie Facebook und Instagram. Höher-schwellige Software wie Office-Anwendungen oder Clouddienste sind kaum vertreten. Ebenfalls randständig sind datenschutz-sensible Alternativen wie beispielsweise Startpage/DuckDuckGo/Ecosia als Suchmaschinen und Telegram/Signal/Threema als Messenger. Da die Nutzung von KI-Anwendungen zum Zeitpunkt der Erhebung in der breiten Bevölkerung noch unüblich war, wurden dazu keine Daten erhoben. Ausgehend von den übrigen Angaben und unter Hinzunahme der oben genannten Studie der Initiative

D21 (2025) ist dennoch davon auszugehen, dass Chatbots unter den befragten Ratsuchenden allenfalls eine geringe bis keine Rolle gespielt haben dürften.

Anhand von Daten dazu, wie fähig Ratsuchende im Umgang mit digitalen Technologien allgemein sind, lässt sich zumindest hypothetisch darauf schließen, wie wahrscheinlich es ist, dass sie Chatbots für ihre Zwecke nutzen können. Digitale Anforderungen, mit denen Ratsuchende nicht zurechtkommen, lassen sich zudem als Unterstützungsbedarfe verstehen, für deren Deckung Chatbots womöglich (weiter-)entwickelt werden könnten. Im Fragebogen konnten die Ratsuchenden eine Selbsteinschätzung vornehmen, ob sie bestimmte Aufgaben mit digitalen Medien bewältigen können (Abbildung 2).

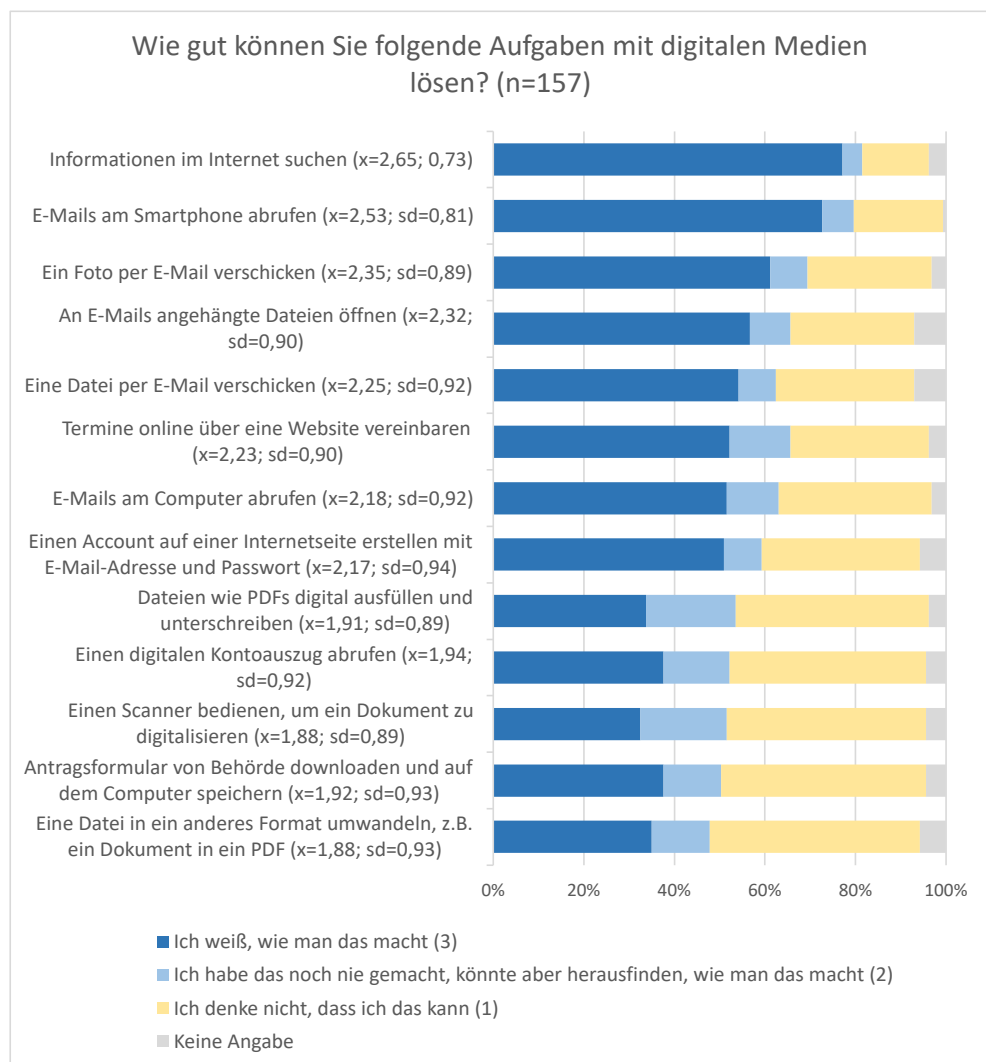


Abbildung 2: Medienbezogene Fähigkeiten der Ratsuchenden
(Quelle: Beloch et al., 2025, S. 33).

Zunächst ist positiv anzuerkennen, dass fast die Hälfte der Befragten sich zutraut, jede der abgefragten Aufgaben mit digitalen Medien entweder selbst zu lösen oder herausfinden zu können, wie diese zu lösen ist. Da-

gegen geben etwa 15 % an, keine der Aufgaben lösen zu können. Auch hier sind Ungleichheitsdimensionen aktuell: Menschen mit niedriger formaler Bildung sowie ältere Menschen geben geringere medienbezogene Fähigkeiten an. Prägnant ist, dass vergleichsweise wenige Ratsuchende Vorgänge wie das Beschaffen digitaler Dokumente von Behörden, das Ausfüllen und Unterschreiben von PDF-Dateien oder das Umwandeln von Dateien in ein anderes Format wie PDF beherrschen – alles Anforderungen, die die Arbeit mit digitalen Formularen und Anträgen auszeichnen. In Bezug auf die Frage, was ein Unterstützungsangebot für sie beinhalten sollte, erachten es parallel dazu 93 % der Befragten als wichtig oder sehr wichtig, „Dass mir jemand bei den Formularen hilft“. Auch das Vereinbaren von Terminen über eine Website – eine immer häufiger auftretende Anforderung im Alltag – stellt für fast ein Drittel der Befragten ein Problem dar.

Daraus lässt sich ableiten, dass die Nutzung von digitalen Möglichkeiten für viele Adressierte eine Hürde darstellt, da ihnen erforderliche medienbezogene Fähigkeiten fehlen. Dies muss nicht automatisch bedeuten, dass ihnen auch die Nutzung eines Chatbots für Beratungszwecke schwerfallen würde, denn dieser ließe sich in seinen Grundzügen in ähnlicher Weise nutzen wie Messenger-Anwendungen, die bereits von der überwiegenden Mehrheit der Ratsuchenden genutzt werden. Erschwerend kommt dabei hinzu, dass die Kommunikation per Schriftsprache Literarität und verbale Reflexivität voraussetzt (Engel, 2008); nur wenn eine Beratungsanfrage verständlich eingegeben werden kann, kann ein Chatbot entsprechend hilfreiche Reaktionen generieren.

Zum anderen deuten die Daten daraufhin, an welchen Stellen Chatbots womöglich unterstützend wirken könnten: Gesteuert durch alltagssprachliche Schriftkommunikation könnten diese die Online-Vereinbarung von Terminen niedrigschwelliger gestalten oder Aufgaben übernehmen, wie die Recherche und das Herunterladen von Antragsdokumenten sowie deren Ausfüllen. Praktisch vorstellbar wäre beispielsweise, dass sie die für das jeweilige Formular erforderlichen Angaben in einfacher Sprache bei den Adressierten abfragen, und anschließend die womöglich umgangssprachlich formulierte Antwort auslesen, um die benötigten Informationen an der richtigen Stelle zu platzieren. Dieser Anwendungsfall bespielt zugegebenermaßen einen eher komplizierten Workaround für ein ohnehin zu hinterfragendes Mittel wie digitale Formulare, die letztlich nur eine computerisierte Version von Papierformularen darstellen. Womöglich wäre es angesichts der neuen technologischen Potenziale angebracht, nicht nach KI-gestützten Lösungen für die Erleichterung des Umgangs mit solchen hochschwelligem althergebrachten Mitteln zu suchen, sondern behördliche Kommunikationswege generell neu zu denken. Mit Blick auf den aktuellen Stand des bürokratischen Systems scheint der benannte Workaround allerdings ein denkbarer Anwendungsfall für Beratungsnutzende zu sein. Sodann stellen sich allerdings auch Fragen rund um die Speicherung und

Verarbeitung personenbezogener Daten. Wie oben erwähnt, finden sich in der vorgestellten Befragung Hinweise darauf, dass die wenigsten Ratsuchenden in der Mediennutzung einen sensiblen Umgang mit den eigenen Daten pflegen.

Unabhängig davon, inwieweit die Ratsuchenden zur Nutzung von Chatbots für Beratung in der Lage wären, stellt sich die Frage, ob sie diese überhaupt nutzen möchten. Schließlich sind sie es, denen das Angebot nutzen soll; entsprechend sollten bei der Wahl der Kommunikationsformate die Bedürfnisse und Präferenzen der Adressierten berücksichtigt werden. Im Fragebogen konnten die Ratsuchenden angeben, wie gerne sie die fünf gängigsten Beratungsformate in den ihnen vertrauten Einrichtungen nutzen würden (Abbildung 3).

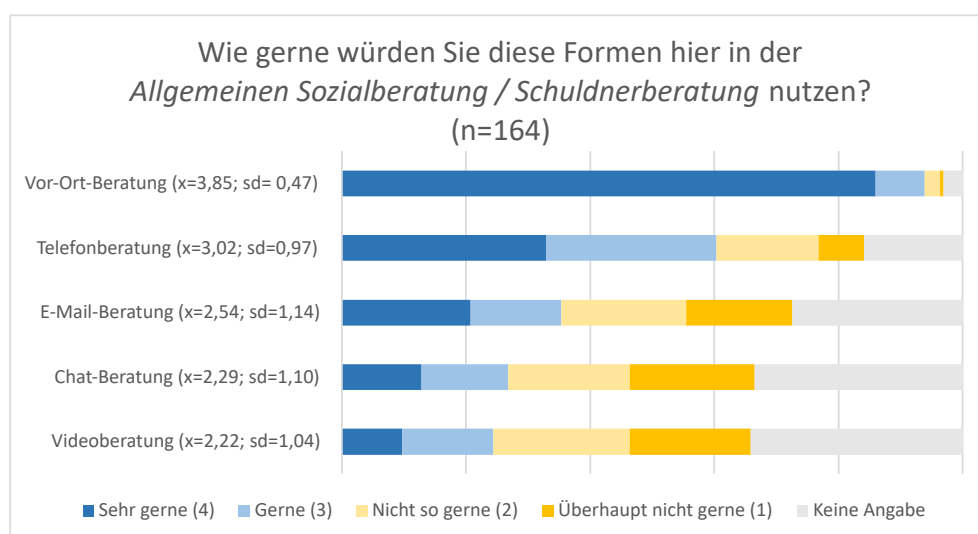


Abbildung 3: Präferenzen der Ratsuchenden zu Beratungsformaten (Quelle: Beloch et al, 2025, S. 28).

Das Ergebnis ist dahingehend sehr deutlich, dass Vor-Ort-Beratung von fast allen Befragten mit „sehr gerne“ oder „gerne“ bewertet wird; nur ein kleiner Teil von etwa 6% gibt dagegen „nicht so gerne“ oder „überhaupt nicht gerne“ an oder macht keine Angabe. Telefonberatung wird von immerhin mehr als der Hälfte wohlwollend bewertet. Die drei digitalen Beratungsformate der E-Mail-, Chat- und Videoberatung bewerten nur ein Drittel oder weniger Personen wohlwollend; ein anderes Drittel etwa macht hier keinerlei Angaben. Wie oben angemerkt, wurde nur die Meinung von Menschen in Einrichtungen vor Ort erhoben, es fehlt also die Meinung derjenigen, die ausschließlich technologisch vermittelte Beratungsformate nutzen. Dennoch ergeben die Werte eine Tendenz, die Schlussfolgerungen auf den Einsatz von Chatbots in der Beratung zulassen: Gemessen an der Bewertung des Formats „Chat-Beratung“ ist davon auszugehen, dass ein erheblicher Teil der Ratsuchenden einer Nutzung von Chatbots zur Beratung als Alternative zur Vor-Ort-Beratung mit menschlichen Beratenden zumindest skeptisch gegenüberstehen dürfte.

Darüber hinaus zeigen die Daten auch einen Zusammenhang zwischen den medienbezogenen Fähigkeiten und der Präferenz digitaler Beratungsformate: Befragte mit höheren medienbezogenen Fähigkeiten bewerten die Chat-Beratung wohlwollender. Was auch festzustellen ist: Jedes Beratungsformat hat seine Befürwortenden. Was für die eine Person gut passt, mag für die andere Person eher problematisch sein. In diesem Sinne lassen sich mit Blick auf (digitale) Beratungsformate keine Pauschalaussagen im Sinne einer *One-size-fits-all-Lösung* treffen.

Ausgehend von den empirischen Daten zu Ausgangsbedingungen der Zielgruppe lässt sich zusammenfassend festhalten, dass zwar vielen der technologische Zugang zu Chatbots durch den Besitz von Smartphones gegeben ist, jedoch nach wie vor bedacht werden muss, dass ein gewisser Teil von der Inanspruchnahme abgeschieden bleibt, sei es aufgrund fehlender Hardware oder mangelnder medienbezogener Fähigkeiten. Zudem variieren die Präferenzen zu bestimmten Beratungsformaten. Die Verlagerung von Beratungsangeboten ins Digitale kann für einen Teil der Zielgruppe also durchaus hilfreich und von ihnen gewünscht sein, keineswegs jedoch für alle gleichermaßen. Um Exklusionsrisiken nicht zu verstärken und damit die Zielgruppe tatsächlich vom Chatbot-Einsatz profitieren kann, müssten deshalb zusätzlich Brückenstrukturen implementiert werden, zum Beispiel in Form von Qualifizierung und Begleitung für marginalisierte Adressierte. Inwieweit Chatbots also ungeachtet ihrer möglichen Alltagsnähe ein tatsächlich niedrigschwelliges Beratungsmittel darstellen (können), ist kritisch zu prüfen.

3. Ausgangsbedingungen professioneller Beratender hinsichtlich Digitalisierung und digitaler Beratungsformate

Wenn Chatbots professionelle Beratung ergänzen sollen, fungiert eine menschliche Beratungsperson weiterhin als hauptsächlich Fallverantwortliche und ein Chatbot als technologisches Hilfsmittel. Damit diese Konstellation einen Mehrwert bietet, müssen die Beratenden das Hilfsmittel im Beratungsprozess an den entsprechenden Stellen – zum Beispiel zur Unterstützung bei digitalen Aufgaben einbringen und dessen Implikationen berücksichtigen können. Obgleich digitale Medien seit vielen Jahren für verschiedene Zwecke in der Beratungsarbeit genutzt werden, zeichnen sich Beratungseinrichtungen durch eine hohe Diversität aus in Bezug auf Mediennutzung, Nutzungskompetenzen und Haltungsfragen. Zunächst ist das empirische Wissen aus den im vorherigen Abschnitt erwähnten Studien (ARD/ZDF, 2024; Initiative D21, 2025) in weiten Teilen auf die Personengruppe der Beratenden übertragbar; auch hier sind Ungleichheitsdimensionen aktuell. Darüber hinaus liegen wenige einschlägige Studien vor, in denen spezifisch Beratende befragt wurden.

Engelhardt et al. (2019) haben auf Basis einer quantitativen Befragung von 157 Beratenden herausgefunden, dass längerfristige Onlineberatung oder *Blended Counseling* nur sehr geringfügig genutzt werden. Praktisch erfüllt internetbasierte Beratung für die Befragten überwiegend die Funktion einer Überbrückung für die Präsenzberatung oder sie wird für besondere Zwecke (z. B. Erstkontakt, Krisenintervention) eingesetzt. Als mögliche Ursache verweisen die Autor:innen auf ältere Arbeiten über die notorische Skepsis von Fachkräften im sozialpädagogischen Bereich, wenn es um technologische Mittelbarkeit von Face-to-Face-Begegnungen geht. Kasten & Lampert (2021) diversifizieren diese Perspektive, indem sie auf Basis einer Befragung von 100 Beratenden drei unterschiedliche Einstellungen respektive Zugänge zum Format der Onlineberatung rekonstruieren: Den konservativ geprägten Zugang zeichnet die Annahme aus, dass Onlineberatung der klassischen Face-to-Face-Beratung unterlegen ist und allenfalls als Einstieg in die *richtige* Beratung oder zur Ermutigung der weiteren Inanspruchnahme dienen kann. Dagegen nimmt der innovative Zugang keine solche Hierarchisierung vor, sondern sieht Vorteile in beiden Zugängen. Letztlich beinhaltet der professionelle Zugang unter anderem einen kritischen Blick auf Fragen des Datenschutzes und der Datenspeicherung, der verschiedenen Beratungsmethoden sowie der Beziehungsgestaltung.

Mit den Erkenntnissen aus der durchgeführten Erhebung mit Beratungsfachkräften können weitere Bedingungen hinzugefügt werden. Zu zwei Zeitpunkten wurden je ein Team der Sozial- und ein Team der Schuldenberatung im Rahmen von Gruppendiskussionen nach ihren Erfahrungen mit digitalen Technologien in der Beratungsarbeit gefragt. Die ersten beiden Diskussionen fanden Mitte 2022 statt, bevor organisatorische Maßnahmen zur Hybridisierung der jeweiligen Beratungsdienste umgesetzt wurden. Die nächsten beiden fanden Mitte 2024 statt, nachdem die Beratungsbüros mit zusätzlichen digitalen Medien ausgestattet wurden, Schulungen durchgeführt und Nutzungserfahrungen mit neuen hybriden Strukturen gesammelt werden konnten. Die Datenauswertung wurde in Anlehnung an die Qualitative Inhaltsanalyse nach Mayring (2022) vorgenommen. Im Folgenden werden Erkenntnisse aus den vier Gruppendiskussionen zusammenführend diskutiert. Diese Betrachtungsweise eignet sich besonders mit Blick auf die einleitend formulierte Fragestellung, da die Daten sowohl Aufschluss geben über Ausgangsbedingungen vor der Implementierung zusätzlicher digitaler Technologien in der Praxis als auch über Erfahrungswerte im Übergangsprozess und mögliche fachliche Hindernisse. Zudem sind diese Daten übertragbar auf den Fall der Implementierung von Chatbots.

Grundsätzlich herrscht bei den befragten Fachkräften die Wahrnehmung vor, dass digitale Kommunikation relevanter wird. Damit ist die Annahme verknüpft, dass sich Beratungsprozesse in den nächsten Jahren zwangsläufig verändern werden, da jüngere Generationen mit einem selbstver-

ständlichen Bezug zu digitalen Medien aufwachsen und entsprechend affin für die Nutzung seien und Beratungsdienste zunehmend online einfordern würden. Die Beratenden zeigen aber auch eine Sensibilität dafür, dass digitale Beratungsformate nicht für alle Ratsuchenden gleichermaßen geeignet sind, womit sie thematisieren, was im vorigen Abschnitt anhand der Ratsuchendenbefragung beschrieben wurde. Ungeachtet des Wissens darüber, wie die Adressierten zu digitalen Beratungsformaten stehen, bilden die befragten Teams sehr diverse Haltungen zur eigenen Nutzung digitaler Medien ab. Anders als die oben angeführte Einteilung von Kasten & Lampert (2021) legen die Daten eine Einteilung in folgende Gruppen nahe, womit mehr auf den Antrieb zur Nutzung neuer Technologien angespielt wird: die Motivierten, die Skeptischen, die Schwankenden.

Beratende der Gruppe der *Motivierten* geben ein persönliches Interesse am Einsatz digitaler Medien an und verknüpfen damit überwiegend Vorteile:

„Ich bin hier gestartet, und mir wurde gesagt: „Okay, es gibt dieses Digitalisierungsprojekt.“ Und ich habe gesagt, dass ich das interessant finde und dass ich das spannend finde und bin jetzt insofern eingestiegen.“ (Sozialberater:in A1)

„Ich bin ja affin und mir macht das ganze Spaß und ich nutze, ja, digitale Medien, solange es sinnvoll ist, sie zu nutzen, mit den Ratsuchenden.“ (Schuldenberater:in B1)

Insofern zeigen sie sich offen und probieren bereitwillig unterschiedliche technische Lösungen aus. Ausgehend vom Vergleich ihrer Aussagen vor und nach der Implementierung der hybriden Beratungsstrukturen hat sich die Nutzung digitaler Möglichkeiten bei ihnen verstetigt. Nutzten sie vorher beispielsweise noch überwiegend Papierakten und Fax, arbeiteten sie später ausschließlich mit digitalen Akten und dem Versenden von Dokumenten per Mail. Gleichermäßen integrieren sie digitale Beratungswege (im Projekt: Videoberatung, Chatberatung) nachhaltig in ihre Arbeitsabläufe. Hier findet sich auch die einzige explizite Erwähnung von KI:

„Ich persönlich habe auch schon künstliche Intelligenz ausprobiert, jetzt Sachen in Texte reinkopiert, die jetzt datenschutzrechtlich kein Ding sind, um die mir da nochmal anders zusammenfassen zu lassen.“ (Schuldenberater:in B1)

Im beforschten Projekt war die Etablierung von KI-Anwendungen nicht als Maßnahme vorgesehen. Das hier zitierte Zusammenfassen von Texten (wahrscheinlich mithilfe eines Chatbots) ist deshalb als individuelles, innovatives Nutzungsverhalten zu interpretieren. Deutlich wird in der Aussage ferner ein Bewusstsein für Datenschutz und die Datenverarbeitung durch KI-Anbieter.

Im klaren Gegensatz dazu steht die Gruppe der *Skeptischen*:

„Ich bin technisch einfach relativ langsam. Ich brauche sehr viel Zeit, um mich in solche Dinge einzufinden, ähm muss das auch immer wiederholen, um überhaupt da fit zu werden.“
(Sozialberater:in A2, *vor* der Implementierungsphase)

„Also ich habe mich nicht darum bemüht, weil [...] das ist auch nicht meine Welt, diese digitale Welt. Und dann denke ich, nee, es ist mir zu aufwendig, zu wenig effektiv.“ (Sozialberater:in A2, *nach* der Implementierungsphase)

Sie zeigen eine eher zurückhaltende Nutzung digitaler Medien in der Beratung, die sich im Verlauf der Implementierungsphase ausgehend von den Schilderungen nicht verändert hat. Sie benennen verstärkt Vorbehalte gegen den Einsatz neuer Technologien, empfinden diese für Beratungskommunikation als nicht zweckmäßig und verzeichnen eine klare Bevorzugung der klassischen Face-to-Face-Beratung. Entsprechend wird sich „nicht darum bemüht“, neue Technologien in gewohnte Arbeitsprozesse einzufügen.

„Für die Zukunft muss man sicherlich gucken, gibt es da / hat das Potenzial? Will ich auch gar nicht in Abrede stellen. Aber hier war einfach so viel auch an anderen Dingen noch zu tun.“
(Sozialberaterin A2, *nach* der Implementierungsphase)

Sie verkennen nicht pauschal, dass digitale Kommunikationsmittel Potenzial haben können; bis auf Weiteres verharren sie jedoch in gewohnten und für sie funktionierenden Mustern. Ihr Verharren scheint zu großen Teilen in mangelnden Nutzungserfahrungen und -fähigkeiten begründet zu sein, was die Vermutung nahelegt, dass zusätzliches Wissen über die Vorteile digitaler Möglichkeiten und ausreichend Raum zum Üben zum Umdenken bewegen könnte (Beloch, 2022).

Werden die Motivierten und die Skeptischen als zwei Enden eines Spektrums begriffen, sind die Gruppe der *Schwankenden* in dessen Mitte einzuordnen:

„Es ist natürlich auch zeitgemäß, was so passiert und es ist für mich ja, würde ich mir jetzt wünschen, auch 'ne Bereicherung überhaupt so digital zu arbeiten.“ (Schuldenberater:in B2, *vor* der Implementierungsphase)

„... was mein Engagement, den Leuten dieses Tool nahezubringen, dann auch einigermaßen gekillt hat.“ (Sozialberater:in A3, *nach* der Implementierungsphase)

Zu Beginn des Projekts hatten die dieser Gruppe Zugehörigen einen Antrieb, sich in neue Beratungsformate einzuarbeiten und verschiedene Medien zur Ergänzung ihrer Arbeitsabläufe zu nutzen. Explizit damit verknüpft steht eine intrinsische Motivation nach persönlicher Weiterentwicklung. Was in der Umsetzung ohne große Hürden gelang, wurde nachhaltig übernommen; so etwa die vermehrt digitale Aktenführung im Team der Schuldenberatung. In anderen Bereichen führten anfängliche Schwierigkeiten während des Medieneinsatzes dazu, dass davon wieder Abstand genommen und zu alten Arbeitsweisen zurückgekehrt wurde. Ein solches Beispiel war der Einsatz von Tablets, um mit den Ratsuchenden gemeinsam Formulare zu bearbeiten; nachdem Ratsuchende mehrfach ihre Zugangsdaten für Online-Konten nicht mitgebracht hatten, verzichtete oben zitierte:r Sozialberater:in irgendwann gänzlich auf den Tablet-Einsatz. Dieselbe Beratungsperson berichtete auch davon, Videoberatung während der COVID-19-Pandemie gezwungenermaßen, aber durchaus erfolgreich genutzt zu haben, nach der Aufhebung der Kontaktbeschränkungen dann jedoch gänzlich zur Face-to-Face-Beratung zurückgekehrt zu sein, da dieses Format als einfacher empfunden wurde. Insgesamt steht die Gruppe der Schwankenden digitalen Medien also offen gegenüber, ob es jedoch zur langfristigen Nutzung kommt, ist stärker von äußeren Bedingungen abhängig.

Spezifisch mit Blick auf den möglichen Einsatz von Chatbots ist in allen drei Gruppen ungeachtet der individuellen Ausgangsbedingungen eine Tendenz weg von textbasierter Onlineberatung hin zu Face-to-Face-Beratung festzustellen. Insbesondere bei psychosozialen Beratungsanliegen ist Chatberatung aus Sicht der Befragten aufgrund der Abwesenheit non-verbaler Elemente weniger geeignet.

„Die schreiben was und das ist... da sind so viele Fragezeichen für den, der es liest, dass man dann ja eigentlich entweder 'nen Standardtext zurückschreibt: „Rufen Sie uns am besten mal an“, oder nachfragt. Und dann hat man so 'ne Art Chat und der ist dann auch echt anstrengend.“ (Schuldenberater:in B1)

Die scheinbar elementare Wichtigkeit der körperlichen Anwesenheit für bestimmte Themen und die generelle Schwierigkeit, dass schriftliche Anfragen von Ratsuchenden missverständlich sein können und klärende Rückfragen nach sich ziehen, wirft Fragen auf: Was können Chatbots an welcher Stelle in Ergänzung zur menschlichen Vor-Ort-Beratung leisten? Was sollten sie leisten können – bei jedem Thema eigenständig *beraten* oder thematisch beschränkt sein und bei bestimmten Schlagwörtern an die menschlichen Beratenden verweisen? Wie kann sichergestellt werden, dass Chatbots nicht *falsch* reagieren, weil sie unklar formulierte Anfragen der Ratsuchenden anders *verstehen* als intendiert?

Ein anderer Vorbehalt wiederum deutet auf ein besonderes Potenzial von Chatbots hin:

„Also die Online-Anfragen [...] die trudeln ja irgendwie rein. Wenn's dir gerade gar nicht passt und dann haben wir 24 Stunden, die zu beantworten – das war so 'ne Vorgabe. Das macht halt so 'n bisschen Druck und passt nicht wirklich, also kann's gar nicht steuern.“
(Schuldenberater:in B1)

Die theoretisch immerwährende Verfügbarkeit der Onlineberatenden am einen Ende der Beratungsplattform bei gleichzeitig fehlender Planbarkeit des Aufkommens von schriftlichen Anfragen kann Beratende unter Druck setzen. Chatbots könnten hier eine Entlastung darstellen, da sie tatsächlich ununterbrochen verfügbar sind und reagieren können. Anfragen in der Onlineberatung könnten auf diesem Wege kanalisiert werden. Die oben genannten Gestaltungsfragen bleiben hier aber gleichsam relevant.

Es wurde deutlich, dass sich die Affinität für digitale Medien innerhalb von Beratungsteams deutlich unterscheiden kann und dass in der Praxis ganz unterschiedliche Erfahrungen in der digitalen Kommunikation mit Adressat:innen gemacht werden. Deshalb ergeben sich auch von fachlicher Seite Klärungsbedarfe, wenn es um die Etablierung beratungsergänzender KI-Anwendungen geht.

4. Fachkräfte im Spannungsfeld zwischen eigener Professionalität und dritten Akteur:innen im Beratungsprozess

In den vorigen Abschnitten wurden Chatbots aus einer Nutzungsperspektive als digitales Medium verhandelt. Diese Konzeptualisierung legt nahe, dass Chatbots lediglich technologische Werkzeuge ohne relevante soziale Implikationen wären. Unter Hinzunahme von Studien zur Interaktion von Menschen und Chatbots muss diese Perspektive allerdings erweitert werden. Bereits 1966 konnte Joseph Weizenbaum durch Experimente mit dem prototypischen Chatbot *ELIZA* empirisch belegen, dass die Mensch-Maschine-Interaktion ähnliche soziale Effekte wie die Mensch-Mensch-Interaktion entfalten kann: Das Computerprogramm analysierte Eingaben von Menschen auf bestimmte Schlüsselwörter und gab dann Textantworten, mit denen es die Gesprächsführung durch Psychotherapeut:innen imitierte. Einige Nutzende waren fest davon überzeugt, mit einem Menschen zu chatten und attribuierten *ELIZA* verschiedene Eigenschaften, wie z. B. Hintergrundwissen und echtes Verstehen (Weizenbaum, 1966, S. 42). Diese Antropomorphisierung technologischer Systeme und das daraus resultierende Eingehen einer „quasisozialen Beziehung“ (Linnemann et al., 2024) zwischen Chatbot und Nutzer:in scheint vor allem angesichts der mittlerweile ausgereiften Transformer-Technologie naheliegend und hat deshalb Konsequenzen für

den Einsatz von Chatbots in der Beratung. So wurde unter dem Stichwort „Eliza effect“ (Turkle, 2007) beispielsweise nachgezeichnet, dass Menschen unter Umständen die *soziale* Interaktion mit einer Maschine gegenüber der Interaktion mit einem Menschen bevorzugen können. Insofern treten Beratungs-Chatbots nicht einfach als weiteres Werkzeug im Methodenkoffer der Beratenden auf, sondern als dritte Akteure im Beratungsprozess – und somit als potenzielle Konkurrenz oder zumindest Mitstreitende, wenn es um die Gelegenheit zur Erfüllung von beratungsbezogenen Aufgaben geht.

Eine ähnliche Konstellation des Nebeneinanders stellt die Verfügbarkeit von Fachkräften und ehrenamtlichen Mitarbeitenden in Beratungseinrichtungen dar. Dieser Praxisbereich ist häufig von wechselseitigen Abgrenzungstendenzen geprägt und so stellte schon Bauer (1998) mit Bezug auf die Soziale Arbeit die Frage, ob ehrenamtlich Mitarbeitende die hauptamtlichen Fachkräfte als *Hilfstruppen* ergänzen oder sie als *Jobkiller* im Zuge kostensparender Substitutionsstrategien ersetzen würden. Diese Grundsatzfrage hat bis heute nichts an Aktualität verloren (Jakob, 2023). Am Beispiel der Zusammenarbeit zwischen Fachkräften und Ehrenamtlichen lassen sich vor diesem Hintergrund weitere Bedingungen für die Implementierung von Chatbots als Ergänzung menschlicher Fachberatung elaborieren. Dafür werden erneut Daten aus den im letzten Abschnitt eingeführten Gruppendiskussionen mit Sozial- und Schuldenberatenden herangezogen. Ehrenamtlich Mitarbeitende werden von den Befragten darin stellenweise als hilfreiche Ressource im Gesamtprozess der Beratungsarbeit thematisiert, vielfach werden mit der Zusammenarbeit allerdings auch Nachteile und problematische Erfahrungen assoziiert.

Das Hinzuziehen von Ehrenamtlichen bedeutet für manche Beratende Umständlichkeit, weshalb der Nutzen in Frage gestellt wird:

„Mache ich es selbst innerhalb von fünf Stunden oder mache ich Folgetermine, führe ein Gespräch mit einem Ehrenamtlichen, mit dem Klienten, gestalte das so transparent, dass das alle verstehen [...] also es ist sehr sehr viel umständliches Arbeiten teilweise.“
(Sozialberater:in A4)

Ein kooperatives Beratungsformat würde erfordern, dass der:die Berater:in zusätzliche Vorkehrungen trifft: Die ehrenamtliche Kraft müsste mit Informationen über die adressierte Person und die geplanten Aufgaben in Kenntnis gesetzt werden, um an die Arbeit des:der Berater:in anschließen zu können. Die adressierte Person müsste über das Hinzuziehen des:der Ehrenamtlichen informiert und der Kontakt zwischen ihnen hergestellt werden. All dies würde weitere Gespräche erfordern, später müsste zudem eine Nachbesprechung erfolgen, damit der:die Berater:in weiß, was inwieweit erledigt wurde und wie fortgefahren werden kann. Demgegenüber steht die Option, die anstehenden Aufgaben nicht zu delegieren, sondern

unter hohem Zeitaufwand selbst zu erledigen. Für andere Beratende spricht gegen das Delegieren an Ehrenamtliche nicht vorrangig der Mehraufwand, sondern die damit einhergehende temporäre Distanzierung:

„Aber wenn ich dann selber nicht da bin, und es kommt eine Rückfrage, das ist dann für mich auch so eine Belastung, dass ich die schon gar nicht dahin schicke.“ (Sozialberater:in A5)

Aufgaben zu delegieren bedeutet, die Fallarbeit für eine Zeit abzugeben und während des darauffolgenden Prozesses nicht anwesend zu sein, um bei Bedarf Rückfragen oder Ähnliches zu beantworten. Dieser Umstand scheint bei dem:der zitierten Berater:in emotionale Belastung zu verursachen, die vom kooperativen Beratungsformat abhält. In ähnlicher Weise wird an anderer Stelle die erfahrungsbasierte Sorge davor beschrieben, dass ehrenamtlich Mitarbeitende ihre Zuständigkeitsgrenzen überschreiten und dabei Fehler machen können:

„Man musste dann auch die Ehrenamtlichen immer wieder daran erinnern, dass der Ablauf so ist, dass die nicht zuerst mit den Leuten eine halbe Beratung machen und uns erst dann dazuschalten. So, das ist, glaube ich, das Beste [...] wenn wir mitkriegen, worum es geht und bevor irgendwelche falschen Sachen gesagt werden, wir dann, ja, unseren Senf dazugeben können.“ (Schuldenberater:in B1)

Was in den Daten an verschiedenen Stellen aktualisiert wird, lässt sich mit einer zentralen Erkenntnis zusammenführen: Die befragten Beratenden sehen den Beratungsfall als *ihr Projekt* und sich selbst als verantwortlich für dessen Erfolg. Ehrenamtliche Mitarbeit stellt in diesem Gefüge ein schwer kalkulierbares Risikopotenzial – gewissermaßen eine Blackbox – dar, weshalb die Befragten sich entweder gegen die Zusammenarbeit oder wenn, dann nur für eine Zusammenarbeit unter enger Begleitung und klar begrenztem Aufgabenprofil der Ehrenamtlichen entscheiden. Äquivalent dazu zeichnen sich Positivbeispiele für ehrenamtliche Mitarbeit in den Befragungen dadurch aus, dass die Ehrenamtlichen in ihrem vereinbarten, engen Aufgabenbereich bleiben und sich bei Unklarheiten direkt an die Beratenden wenden.

Aus diesen fachlichen Perspektiven lassen sich Klärungsfragen für die *Zusammenarbeit* mit beratungsergänzenden Chatbots ableiten. Wenn Chatbots als dritte Akteure beratungsbezogene Aufgaben übernehmen sollen, inwieweit würden die menschlichen Beratenden diese dann überhaupt als unterstützendes Mittel annehmen und in die eigenen Arbeitsprozesse integrieren? Denkbar wäre, dass es hier zu ähnlichen Hemmnissen kommt, wie oben thematisiert: Die Beratungsarbeit würde temporär an Dritte abgegeben werden, deren Handlungen von den Beratungsfachkräften kaum kontrollierbar sind und entgegen deren Intention oder gar fehlerhaft ausfallen könnten. Sicherlich können entsprechende Algorithmen und

umfassende Trainingsdatensätze zumindest die Fehleranfälligkeit eines Chatbots minimieren; das Problem der mangelnden Einsicht in dessen Arbeitsprozesse bleibt aber bestehen. Wie könnte also eine Kommunikation zwischen den menschlichen Beratenden und dem Chatbot ermöglicht werden, damit ähnlich wie in Fallübergabegesprächen mit einer ehrenamtlichen Kraft das Weitergeben von erforderlichen Informationen und Arbeitsständen und somit ein tatsächlich kooperatives Arbeiten erfolgen kann? Damit wird auch die Ebene der Falldokumentation berührt.

Eine transparente Dokumentation wird von den Beratenden als zwingende Grundlage für das kooperative Arbeiten mehrerer Beteiligter gewertet. Mit Blick auf Chatbots stellen sich Fragen des Umgangs mit personenbezogenen Daten in der Beratungsarbeit deshalb neu: Inwieweit kann bzw. muss eine Datenspeicherung bei der Nutzung von Chatbots durch Ratsuchende erfolgen? Sollten die Chat-Daten, bspw. in Form von Gesprächszusammenfassungen, den Beratungsfachkräften übermittelt werden? Sollten die Dokumentationstexte der Beratungsfachkräfte von dem Chatbot ausgelesen werden? Welche datenschutzrechtlichen Faktoren müssen dabei bedacht werden? Und vor allem: Wäre eine solche Schnittstelle überhaupt im Sinne der Adressierten?

5. Ausblick und Limitationen

5.1 *Digitale Teilhabe und Unterstützungsbedarfe aufseiten der Klient:innen*

Mit Blick auf die Adressat:innen von Beratung am Beispiel der vorgestellten Untersuchung der Sozial- und Schuldenberatung konnte festgehalten werden, dass der Besitz digitaler Medien allgemein und insbesondere von Smartphones von älteren und weniger gebildeten Befragten tendenziell seltener angegeben wird. Prinzipiell erlauben die vorhandenen technischen Mittel zwar die Möglichkeit der Nutzung von digitalen Beratungsangeboten, allerdings trifft dies nicht auf diejenigen zu, die keinen Zugang zu digitalen Infrastrukturen haben. Auf Basis der Daten kann nur hypothetisch auf Kompetenzen im Umgang mit KI-Anwendungen geschlossen werden. Trotzdem ist anzunehmen, dass die Erhöhung von Kompetenzen, die einen sicheren und reflektierten Umgang erlauben, dazu beitragen kann, die Nutzungsabsicht zu erhöhen – darunter würden auch Mehrwert generierende und gut nutzbare Anwendungen fallen. Bei Klient:innen könnte dies zu einer verstärkten Hinwendung zu KI-Anwendungen und digitalen Beratungsangeboten führen.

In der durchgeführten Untersuchung wurden Unterstützungsbedarfe in Bezug auf alltagsrelevante Aspekte der befragten Zielgruppe sichtbar, bspw. das Ausfüllen von Formularen als konkreten Anwendungsfall. Auch hierbei spielen Kompetenzen eine Rolle, die, ähnlich wie beim Mediennutzungsverhalten, von Älteren und formal weniger Gebildeten

als niedriger eingestuft werden. Hier finden sich Hinweise auf zielgruppenspezifische Konzeptionen von Beratungsangeboten und ebenfalls auf gezielten Fortbildungsbedarf für Fachkräfte und Begleitung für marginalisierte Adressierte. Außerdem bestehen hier Ansatzpunkte für den Einsatz von Chatbots zur Entlastung der Klient:innen bzw. zur gezielten Unterstützung im Alltag bei organisatorischen Fragen, etwa bei Terminvergabe oder Kommunikation mit Behörden u. Ä. Auch wird deutlich, dass es individuelle Konzepte für Klient:innen braucht, gemessen an deren Ressourcen und Unterstützungsbedarfen sowie an den individuellen Umständen.

5.2 *Beratungsformate und Professionalisierung aus Sicht der Fachkräfte*

In Bezug auf die Gruppe der Beratenden konnte festgehalten werden, dass seit Jahren eine Steigerung der Nutzung digitaler Medien für verschiedene Zwecke in der Beratungsarbeit konstatiert werden kann und Beratungseinrichtungen eine hohe Diversität in Bezug auf Mediennutzung, Nutzungskompetenzen und Haltungsfragen zeigen. Zugleich kann festgehalten werden, dass sich Chatbots prinzipiell dazu eignen, ergänzend zu professionellen Beratenden im direkten Kontakt mit Ratsuchenden zu deren Unterstützung eingesetzt zu werden. Es ist denkbar, dass der Chatbot selbst als Berater unmittelbar für Ratsuchende zum Einsatz kommen kann oder dass er nur mittelbar eingesetzt wird, beispielsweise zur Risiko-diagnostik oder als Dokumentationshilfe und Fallzusammenfassung, worin Entlastungspotential für die Fachkräfte liegen kann.

Der in der Untersuchung konstatierte Fortbildungsbedarf der Akteur:innen zeigt den Bedarf der Fachkräfte (Beloch et al., 2025): Es braucht genügend freigestellte Zeit zur Einarbeitung in die jeweils neue Technologie, um sich in einem geschützten Rahmen damit vertraut machen zu können. Dies ist im Beratungsalltag *nebenher* und ohne Anleitung kaum realisierbar. Es braucht ferner Methoden bzw. Verknüpfung digitaler und nicht-digitaler Beratungsformate, zumal sich gezeigt hat, dass Präsenzberatung aktuell nach wie vor als relevante Form für die Klient:innen angesehen wird. Schulungsformate könnten sein: Theorie-Inputs, aber vor allem praktische Anwendung, d. h. Ausprobieren, Testen, Wiederholen. Hier bieten sich Workshops und praxisnahe Fortbildungen an, die bspw. auf Rollenspiele und vor allem Reflexion setzen, um den Umgang mit KI-Technologien anwendungsnah zu trainieren und ebenso kritisch zu hinterfragen und dabei gleichzeitig Vorbehalte zu verringern.

Für Fachkräfte kann dies ein Autonomieerleben und ein Schritt in Richtung Professionalisierung bedeuten, trotz und gerade wegen eines dritten Players – der KI. Von dieser können sich die Fachkräfte emanzipieren, indem sie ihre Rolle als menschlicher Partner in der Beratung fokussieren und stärken sowie indem sie in der Kommunikation mit Klient:innen

und der Planung des Prozesses das hervorheben, was die Interaktion aus menschlicher Sicht ausmacht und unersetzlich macht: das empathische Hinwenden zum Menschen.

5.3 Methodische Limitationen und offene Fragen

Daten zu Chatbots im Konkreten wurden in der vorliegenden Studie nicht erhoben, da die Nutzung von KI-Anwendungen zum Erhebungszeitpunkt noch nicht flächendeckend vorhanden war. Trotz dieser Limitation eröffnen sich gleichzeitig weiterführende Fragestellungen: Wie kann die Bedeutung von Chatbots in der Beratung sowohl für Adressat:innen als auch für Fachkräfte gesteigert werden mit Blick auf ihr entlastendes Potential? Inwiefern können Adressat:innen sowie Fachkräfte Sozialer Arbeit für KI-Anwendungen sensibilisiert und für einen reflektierten Einsatz aller Beteiligten geschult werden mit Blick auf Datenschutz, Datenethik und -autonomie, Selbstbestimmung und kritischer Reflexion? Auch in Bezug auf die technische Realisierung stellen sich Fragen: Inwiefern bedarf es der Umsetzung und Gestaltung von Chatbots in Anpassung an mobile Geräte nach dem Prinzip des Responsive Design, um so die niedrigschwellige und individuelle Nutzung zu gewährleisten?

Bezugnehmend auf die einleitend formulierte Frage wurde deutlich, dass sich die Affinität für digitale Medien innerhalb von Beratungsteams deutlich unterscheiden kann und dass in der Praxis ganz unterschiedliche Erfahrungen in der digitalen Kommunikation mit Adressat:innen gemacht werden, insbesondere im Fall der Zusammenarbeit mit dritten Akteur:innen. Deshalb ergeben sich auch von fachlicher Seite Klärungsbedarfe, wenn es um die Etablierung beratungsergänzender KI-Anwendungen geht. Die Implementierung von beratungsergänzenden Chatbots ist somit keine triviale Angelegenheit, sondern zieht Klärungsbedarfe in Bezug auf die Rollen menschlicher und maschineller Beratung und deren Kombination nach sich. Zudem sind bestimmte infrastrukturelle Bedingungen erforderlich, die unter anderem davon abhängig sind, was jene Chatbots leisten sollen.

Literatur

- ARD/ZDF. (2024). *ARD/ZDF-Medienstudie 2024 Grundlagenstudie im Auftrag der ARD, ZDF und Forschungskommission*. https://www.ard-zdf-medienstudie.de/files/Download-Archiv/Medienstudie_2024/Basispraesentation_ARD-ZDF-Medienstudie_2024.pdf
- Bauer, R. (1998). Macht das Ehrenamt arbeitslos? Oder hilft das Ehrenamt neue Arbeitsplätze zu schaffen? Freiwillige in der Sozialen Arbeit: Pioniere, Hilfstruppen, Jobkiller. *Sozial Extra*, 10, 2–3.
- Beloch, F. (2022). Alles nur eine Frage der Kompetenzen? Faktoren für eine gelingende Digitalisierung der sozialarbeiterischen Praxis. *FORUM sozialarbeit + gesundheit*, 2/2022.

- Beloch, F., Franken, S., & Kutscher, N. (2025). Dimensionen digitaler Teilhabe im Kontext hybrider Beratung: Forschungsbericht der Wissenschaftlichen Begleitung zum Modellprojekt „Telefonzelle 4.0“. Universität zu Köln. <https://doi.org/10.18716/ESA/UTZ4>
- Bertsche, O., & Como-Zipfel, F. (2023). *Digitalisierung. Herausforderungen und Handlungsansätze für die Soziale Arbeit*. Verlag W. Kohlhammer.
- Engel, F. (2008). Erziehungsberatung online. In G. Hörmann & W. Körner (Hrsg.), *Einführung in die Erziehungsberatung* (S. 280). Kohlhammer.
- Engelhardt, E., Henrich, M., Reindl, R., Weinhardt, M., Zauter, S., & Dietrich, C. (2019). Beratungsbedingte Internetnutzung. Welche Dienste und Angebote nutzen Beratungsfachkräfte privat und beruflich? *e-beratungsjournal*, 15(1), 39–55.
- Engelhardt, E., & Kühne, S. (2025). *Künstliche Intelligenz in der Beratung: Ein Kompass für die systemische Praxis*. Vandenhoeck & Ruprecht.
- Initiative D21. (2025). *D21-Digital-Index 2024/25. Jährliches Lagebild zur Digitalen Gesellschaft*. https://initiatived21.de/uploads/03_Studien-Publikationen/D21-Digital-Index/2024-25/D21_Digital_Index_2024_2025_final.pdf
- Iske, S., & Kutscher, N. (2020). Digitale Ungleichheiten im Kontext Sozialer Arbeit. In N. Kutscher, T. Ley, U. Seelmeyer, F. Siller, A. Tillmann & I. Zorn (Hrsg.), *Handbuch Soziale Arbeit und Digitalisierung* (S. 115–124). Beltz Juventa.
- Jakob, G. (2023). Freiwilliges soziales Engagement und Soziale Arbeit: eine Kooperation mit Herausforderungen. Begriffsvielfalt, Indienstnahme und Konflikte, Wohlfahrtsmix. *FORUM sozial*, 04/2023, 11–15.
- Kasten, A., & Lampert, A. (2021). Blickpunkt Online-Beratung. Erwartungen, Themen und Zugänge aus der Sicht von Berater_innen. In S. Erbring & J. Fischer (Hrsg.), *Zukunft der Beratung: Bd. 5. Sonderband Sozialmagazin* (S. 175–186). Beltz Juventa.
- Kubitza, E. (2024). ChatGPT als digitale Anlaufstelle für Beratungsanliegen zum Thema Schwangerschaftsabbruch. Eine explorative Untersuchung. *e-beratungsjournal*, 20(1), 79–104. <https://doi.org/10.48341/jpaa-3p25>
- Kühne, S., & Engel, F. (2025, Mai 18). *ELIZA und die Falle der Anthropomorphisierung. Kommunikation mit Künstlicher „Intelligenz“: Mutige Gedanken zu KI und Beratung*. <https://kuehne-engel.blogspot.com/2025/05/eliza-und-die-falle-der.html>
- Lehmann, R. (2025). KI in der Beratung. In G. Linnemann, J. Löhe & B. Rottkemper (Hrsg.), *Künstliche Intelligenz in der Sozialen Arbeit: Grundlagen für Theorie und Praxis* (S. 117–127). Beltz Juventa.

- Linnemann, G., Löhe, J., & Rottkemper, B. (2024). Bedeutung von Selbstoffenbarungseffekten in quasisozialen Beziehungen mit auf generativer KI basierten Systemen in Settings von Onlineberatung und -therapie. *e-beratungsjournal*, 20(1), 1–21. <https://doi.org/10.48341/9x1s-5y11>
- Lippert, C., Rudolph, E., Poltermann, A., Engert, N., Lehmann, R., & Albrecht, J. (2024). Generative KI in der beraterischen Ausbildung. Der Einsatz eines*r virtuellen Klient*in als Übungstool für angehende Onlineberater*innen. *e-beratungsjournal*, 20(1), 41–60. <https://doi.org/10.48341/tcgc-st69>
- Martin, A., & Pengel, N. (2024). *Beratung via Chatbot? Möglichkeiten und Anforderungen beim Einsatz Generativer KI in einem bildungswissenschaftlichen Lehramts-Modul*. <https://doi.org/10.48341/RFBQ-T940>
- Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13. Aufl.). Beltz.
- Turkle, S. (2007). Authenticity in the age of digital companions. *Interaction Studies*, 8(3), 501–517.
- van Rießen, A., & Bleck, C. (Hrsg.). (2023). *Handlungsfelder und Adressierungen der Sozialen Arbeit*. Kohlhammer.
- Weizenbaum, J. (1966). ELIZA – A Computer Program For the Study of Natural Language Communication Between Man And Machine. *Communications of the AMC*, 9(1), 36–45.

Dennis Wittmer¹

Ethische und rechtliche Dimensionen des Einsatzes von KI-Technologie in der Sozialen Arbeit und Leitlinien zur informierten Einwilligung der Klientel

¹ Katholische Hochschule Nordrhein-Westfalen

Zusammenfassung

Dieser Beitrag setzt sich mit den ethischen und rechtlichen Implikationen des KI-Einsatzes in der Sozialen Arbeit auseinander. Ethische Bedenken, wie algorithmische Voreingenommenheit, die Wahrung der professionellen Grundsätze und die unklare Verteilung von Verantwortung werden ebenso diskutiert wie die rechtlichen Rahmenbedingungen durch die DSGVO und die KI-Verordnung. Ziel ist es, einen verantwortungsvollen und transparenten Umgang mit KI in der Sozialen Arbeit zu fördern und die Profession zur reflektierten Positionierung zu befähigen, während die Rechte der Klientel in den Mittelpunkt gestellt werden. Aufbauend auf dieser Analyse werden praxisorientierte Leitlinien für die informierte Einwilligung von Klient:innen entwickelt.

Schlüsselwörter: Soziale Arbeit, Digitalisierung, Künstliche Intelligenz (KI), Ethik, Recht, DSGVO, KI-VO, informierte Einwilligung

Ethical and Legal Dimensions of AI Use in Social Work and Guidelines for Clients' Informed Consent

Abstract

This paper examines the ethical and legal implications of the use of AI in social work. Ethical concerns, such as algorithmic bias, upholding professional principles, and the unclear distribution of responsibility, are discussed alongside the legal frameworks of the GDPR and the AI Act. The aim is to promote a responsible and transparent approach to AI in social work and to enable the profession to adopt a reflective stance, while centering clients' rights. Based on this analysis, practice-oriented guidelines for the informed consent of clients are developed.

Keywords: Social Work, Digitalization, Artificial Intelligence (AI), Ethics, Law, GDPR, AI Act, Informed Consent

Wittmer, D. (2026). Ethische und rechtliche Dimensionen des Einsatzes von KI-Technologie in der Sozialen Arbeit und Leitlinien zur informierten Einwilligung der Klientel. *DiMe-Sozial*, 2(1), 21–43. <https://doi.org/10.21240/dimesoz/2026/1/60>

1. Einleitung

Die rasante und stetig fortschreitende Entwicklung von Künstlicher Intelligenz (KI) durchdringt mittlerweile nahezu alle Lebensbereiche und bringt somit auch Implikationen für die Soziale Arbeit mit sich (Linnemann et al., 2023; Stieler et al., 2024). Dabei ist die Soziale Arbeit womöglich kein Bereich, welcher bei Überlegungen bezüglich Veränderungen durch den KI-Fortschritt als erstes in den Sinn kommt. Dennoch zeigt sich das Veränderungspotenzial auch hier zunehmend (Linnemann et al., 2023). Die Frage, die sich hierbei stellt, ist nicht, *ob* KI Einzug in die Soziale Arbeit finden wird, sondern vielmehr *wie* (Stieler, 2021). Dabei besteht die Gefahr, dass KI weitgehend unbemerkt und unreflektiert in den Alltag der Sozialen Arbeit integriert wird (Görder, 2021). Die Profession steht somit vor der dringenden Herausforderung, sich angesichts der rasanten Entwicklungen zu positionieren und Leitlinien für den Umgang mit KI mitsamt Chancen und Risiken zu entwickeln (Reamer, 2023). Der vorliegende Beitrag greift diesen Auftrag auf und widmet sich der Auseinandersetzung mit ethischen und rechtlichen Implikationen des KI-Einsatzes in der Sozialen Arbeit, während fachliche Grundsätze und relevante Regularien in den Blick genommen werden. Auf dieser Grundlage werden zudem Leitlinien für die informierte Einwilligung von Klient:innen entwickelt, die zur Anwendung kommen, wenn in deren Fallkontext KI-Systeme personenbezogene Daten verarbeiten. Diese Leitlinien sollen dazu dienen, die rechtlichen Anforderungen einer informierten Einwilligung zu verstehen und einzuhalten. Zudem sollen sie anleiten, wie die Klientel dazu befähigt werden kann, diese Einwilligung zu erteilen.

Zunächst folgt ein Überblick des aktuellen KI-Einsatzes in der Sozialen Arbeit, welcher Potenziale und Anwendungsbereiche behandelt, die für die Kontextualisierung der weiteren Inhalte entscheidend sind. Darauf folgend werden die ethischen Implikationen dieser Entwicklung aufgegriffen und diskutiert, während Themenschwerpunkte wie Herausforderungen, Verantwortlichkeit, Autonomie und fachlich gebotene Grundsätze, dargestellt werden. Auf diese Behandlung der ethischen Dimension folgt die der rechtlichen. An dieser Stelle werden einschlägige Gesetze und Regularien dargestellt und diskutiert. Durch die Betrachtung dieser beiden Dimensionen ergeben sich die Rahmenbedingungen, auf denen die Leitlinien zur informierten Einwilligung entwickelt werden.

2. Aktueller KI-Einsatz in der Sozialen Arbeit

Die Soziale Arbeit verändert sich seit dem frühen 20. Jahrhundert ständig durch die Integration von Technologie; begonnen mit Schreibmaschinen und Telefonen, gefolgt durch das Internet und die sozialen Medien. Jede dieser Entwicklungen steht für eine eigene Revolution im Hinblick auf Arbeitsprozesse, Fallmanagement sowie die Interaktion zwischen Sozial-

arbeitenden und ihrer Klientel (Nuwasiima et al., 2024). Aktuell ist davon auszugehen, dass KI-Technologie in der Sozialen Arbeit, trotz des diskutierten Spannungsverhältnisses zwischen Chancen und Risiken, ebenfalls das Wirkungspotenzial solcher vergangenen Entwicklungen besitzt (Plafky et al., 2022; Stieler et al., 2024). KI gewinnt auch zunehmend an allgemeiner Relevanz in der Gesellschaft, weshalb sie nicht als isoliertes Phänomen, sondern als integraler Bestandteil und Einflussfaktor betrachtet werden sollte. Sie wird künftig explizit sowie implizit Auswirkungen auf Arbeitsprozesse, Lebensgestaltung, Alltagskommunikation und Entscheidungsprozesse von Sozialarbeitenden sowie derer Klient:innen haben (Stieler et al., 2024).

Auch wenn der Einsatz von KI-Technologie im Bereich der Sozialen Arbeit noch am Anfang steht und häufig erst in Pilotprojekten erprobt wird, werden bereits Potenziale und Anwendungsbereiche deutlich. Die Chancen des KI-Einsatzes unterscheiden sich dabei je nach konkretem Einsatzkontext. Üblicherweise zielt dieser auf eine Steigerung der Effizienz, Qualität oder Personalisierung ab. So kann KI-Technologie eingesetzt werden, um Routinetätigkeiten zu automatisieren, administrative Aufgaben in der Fallverwaltung zu übernehmen und Entscheidungsprozesse zu verbessern, wodurch sich die Sozialarbeitenden intensiver auf die direkte Interaktion mit ihrer Klientel konzentrieren können (Fernando & Ranasinghe, 2023; Nuwasiima et al., 2024). Prädiktive KI-Analysen können eingesetzt werden, um Fälle zu analysieren, beispielsweise um Risikogruppen im Bereich des Kinderschutzes oder der psychischen Gesundheit zu identifizieren, bei denen die Notwendigkeit sofortiger Interventionen vorliegt. So lassen sich Fälle nach der Wahrscheinlichkeit negativer Ergebnisse priorisieren, damit die Unterstützung bei den Klient:innen ankommt, die diese am dringendsten benötigen (Nuwasiima et al., 2024). Solche KI-gestützten Analysen ermöglichen auch eine schnellere und umfassendere Rekonstruktion der Lebenswelten, während Ähnlichkeiten und Unterschiede zu anderen Fällen berücksichtigt werden. Dadurch ergeben sich für alle Beteiligten förderliche Perspektiven, die zuvor unzugänglich oder schwer erreichbar waren (Linnemann et al., 2023; Stieler et al., 2024). Eine weitere Anwendungsmöglichkeit im direkten Fallkontext sind KI-gestützte Chatbots, welche zur emotionalen Unterstützung, Psychoedukation, zur Symptomverfolgung und für Selbsthilfeübungen eingesetzt werden können. Ziel solcher Chatbots ist es, besonders in Zeiten hoher Nachfrage und geringer Kapazitäten, den Zugang zu Ressourcen zu erweitern und auch zwischen den Sitzungen eine kontinuierliche Unterstützung zu bieten (Fernando & Ranasinghe, 2023; Nuwasiima et al., 2024; Stieler et al., 2024). Dabei kommt den Chatbots gerade bei der Unterstützung von sozial isolierten Klient:innen eine wichtige Bedeutung zu (Alotaibi & Alshahre, 2024). Auch außerhalb des direkten Fallkontexts bestehen erwähnenswerte Potenziale. So birgt die KI-Integration für die Ausbildung von Sozialarbeitenden einzigartige Gelegenheiten, komplexe Herausforderungen zu üben. Durch

Simulationen können Studierende an sensible Themen wie Kinderschutz und häusliche Gewalt herangeführt und darauf vorbereitet werden, was zuvor auf einer derart immersiven Ebene nur schwer möglich war. Untersuchungen zeigen, dass simulationsbasierte Trainings in der Lage sind, die Empathie, das Selbstbewusstsein und eine positivere Einstellung gegenüber Klient:innen zu fördern (Dalziel et al., 2024). KI-Systeme sind auch in der sozial(arbeits)wissenschaftlichen Forschung vielseitig anwendbar, beispielsweise bei Analysen und Simulationen. Darüber hinaus können soziale Organisationen KI einsetzen, um ihre Ressourcen effizienter zu nutzen. Dies gelingt etwa durch die Analyse von Trends und Mustern bei der Inanspruchnahme von Dienstleistungen (Kuckartz & Rädiker, 2024; Nuwasiima et al., 2024; Xu et al., 2024). Für eine vertiefende Auseinandersetzung mit den vielseitigen Einsatzfeldern von KI-Systemen in der Sozialen Arbeit bieten Linnemann et al. (2025) in ihrem Sammelband eine umfassende Darstellung.

Der Einsatz von KI-Technologie ist daher insgesamt eine vielversprechende und möglicherweise notwendige Entwicklung für die Praxis der Sozialen Arbeit, welche aktuell vor mehreren Herausforderungen steht, die die Effektivität beeinträchtigen. So steigt die Nachfrage nach sozialen Dienstleistungen, während die Ressourcen begrenzt sind, weshalb Sozialarbeitende oft mit zu hohen Fallzahlen konfrontiert sind. Dieser Umstand begünstigt bekannterweise Burnouts und die Abnahme der Betreuungsqualität. Die Integration von KI-Technologien in die Soziale Arbeit birgt das Potenzial, diese Herausforderungen bewältigen zu können und sollte daher entsprechend berücksichtigt werden (Nuwasiima et al., 2024).

3. Ethische Dimension für die Soziale Arbeit

In der Politik, Wirtschaft und Wissenschaft werden zahlreiche ethische Herausforderungen beschrieben, welche die aktuelle Entwicklung von KI-Technologie aufwerfen (Görder, 2021). Die meisten dieser Überlegungen sind dabei nicht ausschließlich auf den Diskurs in der Sozialen Arbeit beschränkt, sondern von allgemeiner Relevanz. Jedoch stellen sich diese ethischen Herausforderungen besonders für die Soziale Arbeit, denn ethische Handlungskompetenz gilt als wichtige Säule dieser Profession (Graumann & Maaser, 2023). In Anbetracht der geschilderten Annahme, dass KI-Technologie nicht nur Einzug in die Soziale Arbeit finden wird, sondern auch als gesellschaftlicher Bestandteil gesehen werden sollte, welcher so unvermeidbar als Einflussfaktor auf die Soziale Arbeit wirkt, erscheint die ethische Auseinandersetzung mit dieser Entwicklung umso bedeutender (Nuwasiima et al., 2024; Stieler et al., 2024). Aktuell sieht sich die Soziale Arbeit durch den damit einhergehenden Paradigmenwechsel überwältigt, da bisher die ethische Reflexion sowie konkrete Strategien fehlen (Fernando & Ranasinghe, 2023; Reamer, 2023).

3.1 Bedenken und Herausforderungen

Ein zentraler in der Literatur diskutierter Aspekt ist die Vertraulichkeit und der damit einhergehende Schutz von Informationen. KI-Systeme erfordern große Datenmengen, darunter häufig auch private und sensible Informationen, um trainiert zu werden und effizient funktionieren zu können. Dies wirft Bedenken hinsichtlich der Privatsphäre, des Datenschutzes und der Überwachung auf (Dutta, 2024; Nuwasiima et al., 2024). Die digitale Speicherung personenbezogener Daten ist immer mit einem Risiko für den Datenschutz verbunden, denn digitale Systeme und Server können zwar vielseitig geschützt werden, trotzdem sollte eine Kompromittierung nie ganz ausgeschlossen werden. Solche Daten sollten daher, falls situativ möglich, stets pseudonymisiert oder anonymisiert gespeichert werden (Alpers et al., 2024; LfDI BW, 2024). Auch wird mögliche algorithmische Voreingenommenheit (Bias) in den Trainingsdaten kritisiert. Diese kann dazu führen, dass KI-Systeme aufgrund von Faktoren wie Ethnie, Geschlecht oder sozioökonomischem Status ungleich handeln, womit das Risiko einer systematischen Benachteiligung bestimmter Gruppen besteht (Nuwasiima et al., 2024; Stieler et al., 2024). Neben solch einer ungerechtfertigten Diskriminierung besteht allerdings auch die Frage nach gerechtfertigter Diskriminierung, also bei Fällen, in denen Ungleichbehandlungen ethisch legitim oder gar geboten sind. Wenn der klassische aristotelische Gerechtigkeitsgrundsatz befolgt wird, kann es aus Gründen der Gerechtigkeit sogar gefordert sein, *Gleiches gleich und Ungleiches ungleich* zu behandeln. Denn zum Ausschöpfen des Potenzials von KI in der Sozialen Arbeit gehört auch die Möglichkeit, Kombinationen von Merkmalen, wie Risikofaktoren, zu identifizieren und aufgrund dieser Personen anders zu behandeln. Es ist also zu klären, welche Unterschiede in Bezug auf Eigenschafts- oder Verhaltensmerkmale, welche Formen der Ungleichbehandlung legitimieren oder gar verlangen (Görder, 2021). Ein weiterer Faktor, welcher im Kontext des KI-Einsatzes berücksichtigt werden sollte, ist die Fehleranfälligkeit der KI-Ausgaben. So können KI-Systeme Fachkräften Fehlinformationen mitteilen, welche zu Fehlentscheidungen führen (Fernando & Ranasinghe, 2023; Nuwasiima et al., 2024; Stieler et al., 2024). In diesem Zusammenhang besteht der *Automation Bias*. Dieser beschreibt die Tendenz von Menschen, automatisierte Vorschläge unkritisch zu übernehmen (Stieler et al., 2024). Die Verlässlichkeit und Korrektheit von KI-Ausgaben sollten daher stets bewusst kritisch hinterfragt werden. Denn durch KI-Technologie lässt sich zwar die Kindeswohlgefährdung in einem Fallkontext statistisch berechnen, doch es sollte nicht ausschließlich auf der Grundlage dieser Berechnungen gehandelt werden. Besonders im Rahmen des Kinderschutzes können Fehleinschätzungen fatale Folgen haben (Fernando & Ranasinghe, 2023).

In der Vergangenheit waren digitale Technologien lediglich Werkzeuge in den Händen der Sozialarbeitenden. Sie hatten den komplexen Anforderungen der Sozialen Arbeit wenig entgegenzusetzen, da sie nicht

selbstständig aktiv werden konnten (Stieler et al., 2024). Durch KI erhalten diese Werkzeuge zunehmend neue Autonomien und Einsatzmöglichkeiten, so können sich beispielsweise Chatbots, ohne das Zutun von Sozialarbeitenden, mit den Klient:innen unterhalten. Hierbei stellt sich allerdings die Frage, ob solche Entwicklung in Zukunft auch vollends genutzt werden sollten. Die eigene Begrenzung des KI-Einsatzes, um ethische und professionelle Grundsätze zu wahren, sollte daher einen zentralen Diskussionspunkt darstellen. Denn Soziale Arbeit benötigt vertrauensvolle menschliche Interaktionen, die Empathie und Feingefühl für Nuancen in komplexen sozialen Situationen erfordern – was KI-Systeme nicht leisten können (Fernando & Ranasinghe, 2023).

3.2 *Autonomie und Verantwortlichkeit*

Neben Begrenzungen und Risiken des KI-Einsatzes stellt sich die zentrale ethische Frage, wem die Verantwortlichkeit für das Wirken von KI-Systemen zukommt. Besonders die zunehmende Autonomie von KI-Systemen führt aus ethischer Sicht zu unklaren Verantwortlichkeiten (aus rechtlicher Sicht wird dieser Umstand in Kapitel 4.5 aufgegriffen). Mit steigender Komplexität der Technologie wird die Zuweisung von Verantwortung für die Handlungen von KI-Systemen stetig herausfordernder (Dutta, 2024; Nuwasiima et al., 2024). Leistungsfähige Algorithmen werden von manchen fast als moralische Akteure betrachtet, in der ethischen Diskussion wird jedoch mehrheitlich die Position vertreten, dass Verantwortung nur personalen Subjekten zugeschrieben werden kann (Görder, 2021). Dem Deutschen Ethikrat (2023, S. 143) zufolge, können nur natürliche Personen moralische Verantwortung übernehmen, wenn sie über Handlungsfähigkeit verfügen, also in der Lage sind aktiv, zweckgerichtet und kontrolliert auf die Umwelt einzuwirken und dadurch Veränderungen zu verursachen. Wenn dies auf KI-Systeme zutreffen würde, könnten diese als verantwortungsfähig verstanden werden, doch müsste ihnen dann auch der Personenstatus zugeschrieben werden. Zu den Eigenschaften, die diesen Status begründen, gehören unter anderem Intelligenz, Bewusstsein, Selbstbewusstsein, Selbsterkenntnis, Emotionen, freier Wille und Autonomie. Dabei müssen diese nicht kumulativ vorliegen, so wird beispielsweise auch einem Embryo der Personenstatus zugesprochen. Daher stellt sich die Frage, ob KI-Systemen, die möglicherweise in der Zukunft diese Eigenschaften aufweisen könnten, der Personenstatus zugesprochen werden kann. Nach vorherrschender Auffassung wird der Personenstatus für KI-Systeme auch in Zukunft nicht erreichbar sein, da Maschinenhandeln bislang nur die Imitation einer Handlung darstellt und grundlegende Elemente menschlicher Intelligenz und moralischer Urteilsfähigkeit fehlen (Dt. Ethikrat, 2023, S. 143; Jafari-Kammann, 2025, S. 6, 19). Somit kann Verantwortung nicht direkt von maschinellen Systemen übernommen werden. Diese fällt entsprechend den Menschen zu, die in unterschiedlichen Funktionen hinter solchen Systemen stehen. Die Festlegung klarer Verantwortungen ist daher unerlässlich, um den Einsatz von

KI-Systemen verantwortungsvoll zu gestalten (Dutta, 2024; Nuwasiima et al., 2024). Dies wirft die Frage auf, wem die (Teil-)Verantwortung für die Wirkungen von KI-Systemen zukommt oder zukommen sollte, beispielsweise wenn diese Fehler machen. Infrage kommen dabei jene Personen bzw. Organisationen, die KI-Systeme entwickeln, verkaufen, installieren oder benutzen. Die Entscheidungen werden letztlich aber immer von Menschen getroffen, bei denen auch die Verantwortlichkeit für diese liegt (Görder, 2021). Wenn ein KI-System im Rahmen automatisierter Datenauswertung Schlussfolgerungen erzeugen soll, liegt die Verantwortung bei den Menschen, ein solches System in einer ethisch vertretbaren Weise zu entwickeln und einzusetzen (Dt. Ethikrat, 2023, S. 143–144). Grundlage zur Realisierung dieser Perspektive bildet die Annahme, KI-Systeme als unterstützendes Werkzeug zu verstehen, nicht als autonome Kraft. Besonders im Bereich der Entscheidungsfindung sollten KI-Systeme nur zur Entscheidungsunterstützung eingesetzt werden und nicht als Entscheidungersatz. Andernfalls kann nicht verhindert werden, dass Verantwortlichkeit unklar wird und wirksame Kontrollmöglichkeiten verloren gehen. Letztendlich muss immer eine entscheidungsbefugte Person sichtbar bleiben, welche die Verantwortung übernimmt (Dt. Ethikrat, 2023, S. 350). Solche eine Zuschreibung der Verantwortlichkeit ist bisher weder ethisch noch rechtlich (Kapitel 4.5) eindeutig geregelt. Nichtsdestotrotz ist eine klare Verantwortlichkeit unabdingbar, damit der KI-Einsatz ethisch vertretbar erfolgt.

3.3 Ethische Entwicklung und Nachvollziehbarkeit

Grundsätzlich stellen die ethische Gestaltung, Umsetzung und Evaluation von KI-Systemen zentrale Herausforderungen dar, die schon bei der Entwicklung solcher Systeme bedacht werden müssen. Damit die ethische Gestaltung, Umsetzung und Evaluation von KI-Systemen überhaupt gelingen können, müssen ethische Bedenken bereits in die Entwicklung und Planung einfließen. Das bedeutet in der Konsequenz, dass entwickelnde Personen, Unternehmen und Mitwirkende dies trotz Interessenkonflikten, sicherstellen sollten (Dutta, 2024; Nuwasiima et al., 2024; Stieler et al., 2024). Ethische Verantwortung lässt sich nicht auf KI-Systeme auslagern, auch wenn in einen KI-Algorithmus ethische Regeln einprogrammiert werden können, wodurch er in einem gewissen Maß *moralisiert* wird. Es wäre allerdings ein Trugschluss deswegen anzunehmen, dass die KI dadurch selbst ethisch verantwortungsvolle Entscheidungen treffen kann. Denn nicht alles, was eine ethische Entscheidung ausmacht, lässt sich in einem Algorithmus übersetzen. Für streng deontologische oder utilitaristische Ethiken mag dies zwar denkbar sein, aber gerade die Ethik der Sozialen Arbeit bezieht sich aus guten Gründen auf andere Ethiktraditionen. Darunter fallen tugend- und situationsethisch orientierte Ansätze, welche die Haltungen, Motive, Beziehungsqualität, Gefühle oder moralische Intuitionen berücksichtigen. So setzen Werte wie Empathie, Loyalität, Zugewandtheit oder Nächstenliebe zwischenmenschliche Begegnung voraus

und lassen sich nicht in Algorithmen überführen. Die Absicht, eine moralische Maschine zu entwickeln, ist daher möglicherweise paradox, denn dies würde die menschlichen Qualitäten entfernen, die ethisch qualifizierte Entscheidungen erst ermöglichen (Görder, 2021).

Ein weiterer zentraler Aspekt des ethischen KI-Einsatzes ist die Kontrolle von eingesetzten KI-Systemen. Dabei sind Transparenz und Nachvollziehbarkeit algorithmischer Systeme für die Kontrolle und Verantwortung weder zwingend notwendig noch ausreichend (Dt. Ethikrat, 2023, S. 377). Kontrolle und Verantwortung sind auch ohne vollständige Transparenz möglich. Den herstellenden oder anwendenden Personen von KI-Systemen kann die volle Verantwortung für den Einsatz von KI-Systemen zugewiesen werden, auch wenn ihnen die Details der Verarbeitung, die auf Deep-Learning-Ansätzen beruhen, unbekannt sind. Zentral hierbei ist die getragene Verantwortung derartige Systeme zum Einsatz gebracht zu haben und dementsprechend auch begründen zu können, weshalb diese Intransparenz akzeptabel ist. Das kann der Fall sein, wenn der mögliche Schaden der Intransparenz gering ist oder der zusätzliche Nutzen die Nachteile der Intransparenz überwiegt. Ein anschauliches Beispiel hierfür ist, dass im medizinischen Kontext ein KI-System in der Krebsdiagnostik eingesetzt wird, dessen Prognosen eine hohe Genauigkeit aufweisen und somit die Nachteile überwiegt. Auch wenn diese nicht mehr nachvollziehbar von den verantwortungstragenden Personen erklärt werden können (Dt. Ethikrat, 2023, S. 377–378). Dabei betont der Deutsche Ethikrat (2023, S. 350, 378), dass dies in anderen Kontexten und sensiblen Entscheidungen der öffentlichen Verwaltung bedeuten kann, dass intransparente Systeme nicht eingesetzt werden dürfen, wenn die Notwendigkeit der vollständig nachvollziehbaren Begründung nicht gewährleistet ist. Denn Betroffene müssen, insbesondere bei tiefgreifenden Entscheidungen, nachvollziehen können, auf welcher Grundlage der Algorithmus zu seinem Ergebnis kam. Für die Soziale Arbeit ist dies insofern relevant, weil diese in diversen Behörden der öffentlichen Verwaltung agiert, wie beispielsweise im Jugend-, Sozial- oder Gesundheitsamt. Allerdings sei dabei angemerkt, dass sich diese Gedanken auf KI-Systeme im Kontext der (automatisierten) Entscheidungsfindung und -unterstützung beziehen. Einsatzbereiche von KI-Systemen, wie in der Dokumentation (z. B. Transkription von Gesprächen oder dem Erstellen von Gesprächsvermerken) und Organisation (z. B. Terminmanagement), sind in diesem Sinne weniger problematisch und durchaus auch interessant für öffentliche Verwaltungen.

3.4 Ethisch und fachlich gebotene Grundsätze

Um den zuvor dargestellten Überlegungen gerecht zu werden, ist die Einhaltung gewisser Grundsätze notwendig, damit KI-Systeme in der Sozialen Arbeit ethisch vertretbar eingesetzt werden können, während Fach-/Qualitätsstandards weiterhin gewahrt werden. Auf der bereits vertieften

Grundlage der zuvor dargestellten Literatur, lassen sich sechs zentrale ethische Grundsätze herausstellen, die unter anderem aus den Empfehlungen des Deutschen Ethikrats (2023) für den KI-Einsatz in der öffentlichen Verwaltung abgeleitet wurden:

- 1) *Klientelzentrierung*: Die Bedürfnisse, der Wille sowie die Rechte der Klient:innen sollten stets im Zentrum des KI-Einsatzes stehen. Wenn in der Sozialen Arbeit KI-Technologie zum Einsatz kommt, zählen diese zu den direkt Betroffenen. Die Klientel muss daher dazu ermächtigt werden, den KI-Einsatz in ihrem Fallkontext zu verstehen, um diesem zuzustimmen oder ablehnen zu können (Molala & Mbaya, 2023). Hierbei kommt der informierten Einwilligung eine zentrale Bedeutung zu, welche in Kapitel 5 näher dargestellt wird.
- 2) *Verantwortlichkeit*: Da moralische Verantwortung grundsätzlich nur von Menschen übernommen werden kann, muss für jeden KI-Einsatz eine sichtbare, entscheidungsbefugte Person die finale Verantwortung übernehmen. Wer solche Systeme zum Einsatz bringt, trägt die Verantwortung und muss diesen sowie die Inkaufnahme möglicher Risiken begründen können (Dt. Ethikrat, 2023, S. 143, 150, 377–378; Dutta, 2024; Görder, 2021; Nuwasiima et al., 2024).
- 3) *Professionelle Werte und Fachlichkeit*: Die Ethik der Sozialen Arbeit stützt sich primär auf Ethiktraditionen, bei denen tugend- und situationsethische Ansätze im Vordergrund stehen. Werte wie Empathie, Loyalität oder Zugewandtheit sind untrennbar mit der direkten zwischenmenschlichen Begegnung verbunden und lassen sich nicht in Algorithmen übersetzen. Es muss folglich sichergestellt werden, dass die entsprechenden Systeme die grundlegenden fachlichen Standards der professionellen und menschlichen Interaktion nicht ersetzen oder untergraben. Dies erfordert die Wahrung der vertrauensvollen menschlichen Interaktion, die Empathie und Feingefühl für Nuancen in komplexen sozialen Situationen sichert – was KI-Systeme nicht leisten können. Die Beibehaltung dieser essenziellen menschlichen Werte und der Schutz der professionellen Beziehungsqualität sind zentral. Daher sollte die bewusste Begrenzung des KI-Einsatzes zur Wahrung ethischer und professioneller Grundsätze stets reflektiert werden. Eine schleichende Aushöhlung der sozialrechtlich gebotenen, individuellen Bedarfsermittlung zugunsten von Kostenersparnis oder Standardisierung muss verhindert werden (Dt. Ethikrat, 2023, S. 331–332; Fernando & Ranasinghe, 2023; Görder, 2021).
- 4) *Transparenz und Nachvollziehbarkeit*: Für einen verantwortungsvollen KI-Einsatz müssen ausgewogene Standards für Transparenz, Erklärbarkeit und Nachvollziehbarkeit entwickelt werden, die sich flexibel an die jeweilige Aufgabe und den Kontext anpassen. Dabei ist es entscheidend, einen angemessenen Ausgleich zu finden, der auch Aspekte wie Sicherheit, Schutz vor Missbrauch oder Datenschutz berücksichtigt. Diese Standards müssen durch verbindliche technische und organisatorische Maßnahmen (TOM) umgesetzt werden, deren Grundlagen bereits bei der Entwicklung der Software durch klare Dokumen-

tations- und Offenlegungspflichten zu schaffen sind. Generell ist einzu- fordern, dass entwickelnde Personen, Unternehmen und Mitwirkende, den verantwortungsvollen und ethischen KI-Einsatz trotz Interessen- konflikten sicherstellen müssen (Dt. Ethikrat, 2023, S. 378–379; Dutta, 2024; Nuwasiima et al., 2024; Stieler et al., 2024).

- 5) *Risikobewusster und kompetenter Umgang*: Um den Risiken algorithmischer Voreingenommenheit (Bias) wirksam zu begegnen, müssen sowohl die Technologie selbst als auch die anwendenden Personen in die Pflicht genommen werden. Einerseits sollten KI-Systeme und ihre Trainingsdaten regelmäßig auf Verzerrungen überprüft werden. Damit kann die systematische Benachteiligung, etwa aufgrund von Ethnie oder Geschlecht, verhindert werden, wobei die komplexe Abgrenzung zu einer ethisch gebotenen Ungleichbehandlung zu berücksichtigen ist. Andererseits muss diese technische Absicherung durch eine umfassende Schulung der anwendenden Personen ergänzt werden. Diese Schulungen müssen über die reine Bedienung hinausgehen und schwer- punktmäßig ein kritisches Verständnis für die Grenzen, potenziellen Verzerrungen und Risiken der KI-Technologie fördern, um deren verantwortungsvollen Einsatz zu gewährleisten (Dt. Ethikrat, 2023, S. 330–331; Görder, 2021; Nuwasiima et al., 2024; Stieler et al., 2024).
- 6) *Qualitätssicherung*: Es sollten (qualitäts-)sichernde Maßnahmen etab- liert und evaluiert werden, darunter beispielsweise das mindestens stich- probenartige Befolgen des Vier-Augen-Prinzips bei KI-Ausgaben, die regelmäßige Reflexion des KI-Einsatzes oder die Sicherstellung, dass KI-Technologie keine Entscheidungen über menschliche Schicksale treffen darf. Auch Ansätze, wie die Umkehr der Begründungspflicht – also nicht die Abweichung von einem KI-Vorschlag – muss gerecht- fertigt werden, sondern die Entscheidung, einem solchen Vorschlag zu folgen, soll geprüft werden (Dt. Ethikrat, 2023, S. 330, 349–350; Dutta, 2024; Fernando & Ranasinghe 2023; Görder, 2021; Molala & Mbaya, 2023; Nuwasiima et al., 2024; Stieler et al., 2024).

Die Berücksichtigung dieser allgemeinen Grundsätze ist dafür zentral, dass der KI-Einsatz in der Sozialen Arbeit ethisch und im Sinne der Fach-/Qualitätsstandards erfolgen kann. Allerdings sei dazu ergänzt, dass je nach konkretem Anwendungskontext, weitere Vorkehrungen not- wendig sein können und neben diesen allgemeinen Grundsätzen diverse spezifischere geboten sein können. Darüber hinaus muss neben dieser ethischen Dimension, die moralisch gebotene Aspekte fordert, noch die verpflichtende rechtliche Dimension bedacht werden, die selbstverständ- lich stets berücksichtigt werden muss.

4. Rechtliche Dimension für die Soziale Arbeit

Neben der ethischen Dimension, die es bei dem KI-Einsatz (in der Sozia- len Arbeit) zu berücksichtigen gilt und welche aufzeigt, was moralisch ge- boten ist, besteht die verbindliche rechtliche Dimension, die berücksichtigt

werden *muss*. Der Einsatz von KI-Technologie wird direkt und indirekt durch ein Zusammenspiel nationaler Gesetze, internationaler Abkommen und Richtlinien geprägt (Pelzl, 2025, S. 217). Es sind grundsätzlich Datenschutzgesetze, KI-Regularien, aber auch sonstige einschlägige Gesetze wie das Urheberrecht, einzuhalten (Europäische Union, 2024; Industrie- und Handelskammer für München und Oberbayern [IHK München], 2025b). Dieses Kapitel soll besonders im Kontext der Sozialen Arbeit einen Überblick über alle zentralen Regularien bieten und gleichzeitig ungeklärte Dynamiken in der aktuellen Rechtslage darstellen. Dies ist nicht nur wichtig, um eine generelle Übersicht zu geben, sondern auch, um zu verdeutlichen, welche Rechte den Betroffenen, also letztendlich den Klient:innen in diesem Kontext zustehen. Die Soziale Arbeit hat den Auftrag, sich für die Interessen und Rechte der Klient:innen einzusetzen (Deutscher Berufsverband für Soziale Arbeit e. V., 2025). Es ist daher unabdingbar, diesen Auftrag sicherzustellen und Klient:innen dazu zu ermächtigen und sie nicht zu übergehen.

4.1 Europäische Datenschutzgrundverordnung (DSGVO)

Die Europäische Datenschutz-Grundverordnung (DSGVO) garantiert umfassende Betroffenenrechte und beeinflusst die datenschutzfreundliche Gestaltung soziotechnischer Systeme weltweit. Diese Regelungen müssen von den verantwortlichen Stellen durch (technisch-)organisatorische Maßnahmen umgesetzt werden. Mit der Verordnung werden daher die Betroffenenrechte und Pflichten für die Verarbeitung personenbezogener Daten hinsichtlich Transparenz und Auskunft zentral reguliert. Die Bereitstellung langer Datenschutzerklärungen und andere traditionelle Vorgehensweisen werden dem Anspruch der informationellen Selbstbestimmung im Sinne der DSGVO allein nicht mehr gerecht (Grünewald & Pallas, 2021). Im Zentrum der DSGVO steht der Begriff der *personenbezogenen Daten* (Art. 4 DSGVO). Dieser bezieht sich auf alle Informationen, die eine identifizierbare natürliche Person betreffen. Eine Identifizierbarkeit kann dabei auch durch die Zuordnung von (indirekten) Merkmalen (Stimme, Name, Aussehen, biografische Daten, ...) gegeben sein, auch wenn für die Zuordnung gewisses Zusatzwissen erforderlich ist, wie bei IP-Adressen oder pseudonymisierten (Gesundheits-)Daten. Jeglicher Umgang mit diesen Daten wird nach der DSGVO (Art. 4 Abs. 2) als Verarbeitung verstanden, welche Vorgänge wie das Erheben, Speichern, Verändern oder Löschen von Daten umfasst, unabhängig davon, ob diese automatisiert erfolgen. Die Verarbeitung von personenbezogenen Daten ist entsprechend DSGVO (Art. 6 Abs. 1) nur unter mindestens einer von sechs Bedingungen zulässig. Dies ist beispielsweise der Fall, wenn die Verarbeitung dem Schutz in lebensbedrohlichen Situationen dient, sie in Ausübung öffentlicher Gewalt bzw. zur Erfüllung rechtlicher Verpflichtungen im öffentlichen Interesse erfolgt (manchmal zutreffend für die Soziale Arbeit) oder wenn eine informierte Einwilligung der betroffenen Person vorliegt. Besonders letztere ist für diesen Beitrag zentral

(Näheres in Kapitel 5). Dabei stellt eine besonders streng geschützte Kategorie, die sogenannten *besonderen Kategorien personenbezogener Daten* dar (Art. 9 Abs. 1 DSGVO), deren Verarbeitung initial grundsätzlich untersagt ist. Hierzu zählen sensible Informationen, wie die ethnische Herkunft, politische Meinungen, religiöse/weltanschauliche Überzeugungen, Daten zur sexuellen Orientierung, genetische, biometrische oder gesundheitliche Daten. Ihre Verarbeitung ist nur unter bestimmten Ausnahmefällen erlaubt (Art. 9 Abs. 2 DSGVO), wie nach einer informierten und ausdrücklichen Einwilligung der betroffenen Person. Für den KI-Kontext sind diese Regularien insofern relevant, weil KI-Systeme auf der Sammlung, Analyse und Verarbeitung großer Datenmengen basieren, welche häufig personenbezogene Daten enthalten, die entsprechend reguliert werden. Besonders in der Sozialen Arbeit werden in vielen Bereichen sehr sensible Daten verarbeitet (Pelzl, 2025, S. 205, 217). Allerdings sei betont, dass diese dargestellten Regularien der DSGVO nur für personenbezogene Daten gelten, nicht für anonymisierte Daten (DSGVO, 2016, Erwägungsgrund 26). Die Anonymisierung ist ein Verfahren, bei dem personenbezogene Daten so verändert werden, dass sie permanent und unwiderruflich, auch nicht indirekt, keiner natürlichen Person mehr zugeordnet werden können. Im Gegensatz dazu, ist die Pseudonymisierung ein Verfahren, das den Personenbezug nicht aufhebt, sondern lediglich verschleiert. Hierbei werden die ursprünglich identifizierenden Daten durch Pseudonyme, bspw. eine zufällig generierte ID-Nummer, ersetzt. Die Informationen, die eine Re-Identifizierung der Person ermöglichen würden, werden getrennt vom pseudonymisierten Datensatz und unter strengen Sicherheitsmaßnahmen aufbewahrt. Dies sind gemäß DSGVO (Art. 4 Abs. 5) pseudonymisierte Daten und daher weiterhin personenbezogene Daten, da die Möglichkeit der Re-Identifizierung besteht. Hervorzuheben ist daher, dass die DSGVO auf pseudonymisierte Daten Anwendung findet, auf anonymisierte Daten jedoch nicht. Die entsprechenden Konsequenzen für die Praxis müssen im Einzelfall geprüft werden. So kann bei KI-gestützten Fallanalysen gut reguliert werden, welche Art der Daten verarbeitet wird. Wenn Klient:innen dagegen eigenständig mit einem Chatbot interagieren, eher weniger. Auch wenn die zentral relevanten Artikel der DSGVO dargestellt wurden, bestehen darüber hinaus weitere, die beim KI-Einsatz relevant sein können. Darunter Artikel 44, der die Übermittlung personenbezogener Daten in Drittländer reguliert und welcher zu berücksichtigen ist, wenn KI-Systeme im Ausland gehostet werden. Grundsätzlich besteht auf der Ebene der IT-Sicherheit und des Datenschutzes im Kontext von KI-Systemen viel Beachtenswertes, weshalb zur weiteren Vertiefung Pelzl (2025) empfohlen wird.

4.2 KI-Verordnung (KI-VO)

Im Jahr 2024 verabschiedete das Europäische Parlament die KI-Verordnung (KI-VO), auch als AI-Act bekannt, die neben der DSGVO eine weitere zentrale Verordnung darstellt, die es beim KI-Einsatz

zu berücksichtigen gilt. Ziel dieser Verordnung ist die Vorschreibung eines menschenzentrierten und risikobasierten Ansatzes bei KI-Anwendungen. Je höher das Risiko eines KI-Systems ist, desto strenger sind die Anforderungen. Um solche Risiken zu kategorisieren, wird zwischen verbotenen KI-Praktiken, Hochrisiko-KI-Systemen, KI-Systemen mit begrenztem Risiko sowie minimalen Risiko unterschieden. Jede Kategorie unterliegt eigenen Regulationen; so gelten strenge Anforderungen an Technik und Organisation bei Hochrisiko-KI-Systemen. KI-Anwendungen mit einem inakzeptablen Risiko, werden durch die Verordnung gänzlich verboten (Dötterl, 2025; Europäische Union, 2024; IHK München, 2025b). Die KI-VO ist am 1. August 2024 in Kraft getreten und die Verbote und Verpflichtungen erlangen nun schrittweise Gültigkeit (Europäische Union, 2024; Kompetenzplattform Künstliche Intelligenz Nordrhein-Westfalen, 2024). Auch wenn die Konkretisierung der rechtlichen Anforderungen aktuell noch an diversen Stellen erfolgt und die Umsetzung in nationales Recht aussteht, sollte sich bereits den Anforderungen der KI-VO angenommen werden. Die KI-VO nimmt jegliche beteiligte Person entlang der Wertschöpfungskette in unterschiedlichem Maß in die Verantwortung und kategorisiert dabei nach Rollen wie Anbietern, Herstellern oder Betreibern (Europäische Union, 2024; IHK München, 2025a; 2025b). In der Sozialen Arbeit ist der potenzielle Einsatz von KI aus fast allen Risikokategorien denkbar. Die umsetzenden Stellen können nach der KI-VO als *Betreiber* verstanden werden. Entsprechend muss beim KI-Einsatz geprüft werden, um welche Risikokategorie es sich handelt und welche Maßnahmen entsprechend ergriffen werden müssen. Das gilt besonders für Hochrisiko-KIs, für die unter anderem Pflichten bestehen wie eine KI-Kompetenzvermittlung der mitarbeitenden Personen, entsprechende technisch und organisatorische Maßnahmen oder menschliche Aufsichtspflichten (Europäische Union, 2024; IHK München, 2025a). Für eine nähere Vertiefung empfiehlt sich der Beitrag von Dötterl (2025).

4.3 DSGVO und KI-VO im (Spannungs-)Verhältnis

In der Umsetzung scheinen die DSGVO mit ihrem Fokus Datenschutz sowie die KI-VO mit der Regulierung von KI-Systemen in einem Spannungsverhältnis zu stehen. Dieses wird ausgelöst durch die aktuell gängige Praxis des Trainings von KI-Modellen mithilfe von wahllos aus dem Internet gesammelten Datensätzen. Der Europäische Datenschutzausschuss (EDSA, 2024) veröffentlichte hierzu eine rechtlich bedeutsame Klarstellung mit potenziell großer zukünftiger Auswirkung. Der EDSA betont, dass genau diese aktuelle Praxis in der Regel nicht mit der DSGVO vereinbar ist. Das massenhafte Sammeln von personenbezogenen Daten aus dem Internet zum Anlernen von Algorithmen ist in den meisten Fällen nicht rechtskonform. Denn die Verarbeitung der personenbezogenen Daten geschieht auf einer fehlenden Rechtsgrundlage und verstößt gegen die Zweckbindung (EDSA, 2024). Zudem mangelt es an

Transparenz, womit die Rechte der betroffenen Personen nicht umsetzbar sind. Unternehmen, die KI entwickeln, müssen ihre Trainingsmethoden also grundlegend überdenken und werden zukünftig auf ordnungsgemäß lizenzierte Datensätze, synthetische Daten oder Daten mit einer echten Einwilligung zurückgreifen müssen. KI-Modelle, die auf rechtswidrig gesammelten Daten trainiert wurden, könnten sogar mit empfindlichen Bußgeldern und potenziellen Verboten belegt werden. Die Botschaft ist daher: Die Verantwortung für die Einhaltung der DSGVO liegt bereits in der Entwicklungs- und Trainingsphase bei den KI-entwickelnden Personen und Unternehmen. Die konkreten Auswirkungen dessen stehen jedoch noch aus.

4.4 Besonderheiten bei schutzbedürftigen Personen

Wenn KI-Systeme personenbezogene Daten verarbeiten, sind die Vorgaben der DSGVO zu beachten. Es ist darauf hinzuweisen, dass sensible Daten (Art. 9 DSGVO) wie Gesundheitsdaten, besonderen Einschränkungen unterliegen (Dötterl, 2025). Hierbei kristallisieren sich Datenschutzprobleme für die Soziale Arbeit heraus, wenn beispielsweise (schutzbedürftige) Personen mit KI-gestützten Chatbots interagieren und dabei Einblicke in ihr persönliches und intimes Leben preisgeben. Dies wirft die Frage auf, wie diese Daten übermittelt, gespeichert und ausgewertet werden dürfen (Steiner & Tschopp, 2022). Besonders die personenbezogenen Daten von schutzbedürftigen Personen wie Minderjährigen, erhalten nach der DSGVO einen besonderen Schutzstatus (Art. 9 DSGVO). Nach der KI-VO sind KI-Systeme verboten, welche die Vulnerabilität oder Schutzbedürftigkeit ausnutzen. Besagte Vulnerabilität kann sich dabei aus Alter, Behinderung oder auch der sozialen und wirtschaftlichen Situation ergeben. Ein konkretes Beispiel hierfür wäre ein *vermenschlichtes* KI-System, das bei Kindern zu einer emotionalen Abhängigkeit oder zu einer Behinderung ihrer sozialen und emotionalen Entwicklung führen könnte (Dötterl, 2025). An dieser Stelle sollte daher grundsätzlich betont werden, dass im Bereich der Schutzbedürftigkeit von Klient:innen, sei es in Kontexten wie der psychischen Gesundheit oder des Kinderschutzes, strengere Regularien gelten, die stets berücksichtigt werden müssen.

4.5 Ungeklärte rechtliche Verantwortlichkeit

Die KI-VO nimmt zwar verschiedene beteiligte Personen entlang der Wertschöpfungskette auf unterschiedliche Weise in die Verantwortung, doch ähnlich wie in der ethischen Betrachtung, bestehen im rechtlichen Kontext nicht klar verteilte Verantwortlichkeiten. Die Frage nach entsprechenden Verantwortlichkeiten stellt sich dabei auf verschiedenen Ebenen von der KI-Entwicklung bis hin zu ihrem Einsatz. Für diesen Beitrag ist besonders die Ebene des praktischen Einsatzes im Alltag der Sozialen Arbeit relevant, auf der auch die meisten Unklarheiten bei der

Verantwortungszuschreibung und Haftung bestehen. Eine rein rechtliche Betrachtung zeigt, dass nach derzeitigem Stand keine eindeutige Antwort möglich ist. Die Unklarheiten bezüglich der Haftungsfragen sind dafür der ausschlaggebende Grund, weshalb auch entsprechende Unklarheiten in der Verantwortlichkeit bestehen. Momentan können für Unternehmen die Fragen der Haftung für den KI-Einsatz nicht beantwortet werden (Kunitz, 2023). Grund dafür ist die in den Rechtswissenschaften ungeklärte Anwendbarkeit der Maßstäbe aktueller Rechtsprechung sowie die fehlenden Präzedenzfälle. Es ist daher rechtlich ungeklärt, wer bei Fehlern, die auf KI-Systeme zurückgehen, haftet. Das gilt nicht nur für automatisierte Entscheidungssysteme, sondern auch bei menschlichen Entscheidungen, die von KI-Prädikation abweichen oder bei informelleren Nutzungsweisen, wie beispielsweise Chatbots in Beratungssituationen. In all diesen Fällen ist die Forschung defizitär und die Rechtsgrundlage ungeklärt (Steiner & Tschopp, 2022). Solche Haftungsfragen können sich außerdem je nach sozialer Einrichtung bzw. Trägertyp grundsätzlich unterscheiden. Dies lässt die KI-anwendenden Personen in der Sozialen Arbeit aktuell entsprechend ernüchert mit Unsicherheiten zurück. An dieser Stelle gilt es, die stetigen rechtlichen Entwicklungen zu beobachten und sich auf die zuvor aufgezeigten ethisch gebotenen Grundsätze bezüglich der Verantwortlichkeit zu besinnen.

4.6 Umsetzung der rechtlichen Anforderungen in der Praxis

Um den bestehenden vielseitigen und komplexen rechtlichen Anforderungen gerecht zu werden, haben unterschiedliche offizielle Stellen begonnen, entsprechende Diskussionspapiere, Prüfkriterien oder Leitlinien zu veröffentlichen. Darunter beispielsweise das Bayerische Landesamt für Datenschutzaufsicht (BayLDA, 2024), der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit (HmbBfDI, 2023), der Landesbeauftragte für den Datenschutz und die Informationsfreiheit Baden-Württemberg (LfDI BW, 2024) und das Kompetenzzentrum Chemnitz (Kunitz, 2021). Diese vielfältigen Beiträge nehmen unterschiedliche Überlegungen in den Fokus, die an dieser Stelle nur exemplarisch dargestellt werden. So wird herausgestellt, dass je nach Einsatzkontext beachtenswerte Unterschiede in der Rechtsgrundlage gelten, beispielsweise wenn es sich um eine Notfallsituation handelt oder wenn öffentliche oder nicht-öffentliche Stellen für den KI-Einsatz verantwortlich sind (LfDI BW, 2024). Auch wird auf die Notwendigkeit eines Risikomodells verwiesen (gemäß Art. 35 DSGVO). Durch dieses soll dokumentiert werden, ob Schutzziele wie Fairness, Autonomie, Kontrolle, Transparenz, Verlässlichkeit, Sicherheit und Datenschutz sichergestellt werden (BayLDA, 2024). Zentral für alle Beiträge ist die informierte und zweckgebundene Einwilligung der betroffenen Personen, deren personenbezogene Daten verarbeitet werden.

Im Rahmen dieses Beitrags können die vielschichtigen und relevanten rechtlichen Aspekte, welche zu beachten sind, lediglich grob thematisiert werden. Im Folgenden wird spezifisch für die Klientel der Sozialen Arbeit die ethische und rechtliche Erfüllung der informierten Einwilligung (gemäß Art. 4 Nr. 11 DSGVO) in den Blick genommen in deren Fallkontext KI-Technologie zum Einsatz kommt und personenbezogene Daten verarbeitet werden. Daher ist es wichtig zu erwähnen, dass die zuvor aufgeführten Quellen (o. Ä.) zur intensiveren Einarbeitung genutzt werden sollten, damit alle rechtlichen Pflichten erfüllt werden und nicht nur Teile davon. Zum Gesamtbild zählen unter anderem die Regularien zur Erfassung und Speicherung von personenbezogenen Daten, dem Training von KI-Systemen mit solchen Daten, der Pseudo-/Anonymisierung von Daten, dem Erstellen von Risikomodellen, Regularien bei automatisierten Entscheidungen (Art. 22 DSGVO) sowie die Berücksichtigung der KI-VO (BayLDA, 2024; Dötterl, 2025; Europäische Union, 2024; Grünewald & Pallas, 2021; HmbBfDI, 2023; LfDI BW, 2024; Pelzl, 2025).

5. Leitlinien zur informierten Einwilligung

Wie zuvor ausgeführt, ist die Erteilung einer informierten Einwilligung der Klient:innen aufgrund rechtlicher Vorgaben und ethischer Überlegungen gegebenenfalls notwendig, wenn KI-Technologie im Fallkontext eingesetzt wird und in diesem Zuge personenbezogene Daten verarbeitet werden. Diese Art der Einwilligung soll Klient:innen ermöglichen, die Implikationen der KI-Nutzung in ihrem Fallkontext zu verstehen und die KI-Nutzung gegebenenfalls abzulehnen (Molala & Mbaya, 2023). Im Folgenden wird das Konzept der informierten Einwilligung näher beschrieben, gefolgt von Leitlinien, die aufzeigen, wie diese umgesetzt werden sollte.

Jedoch sollte dabei rechtlich differenziert werden, denn eine informierte Einwilligung (gemäß Art. 6 Abs. 1 DSGVO) ist nicht zwingend erforderlich, wenn personenbezogene Daten verarbeitet werden und dies in der Ausübung eines gesetzlichen Auftrags im öffentlichen Interesse geschieht, was im Kontext der Sozialen Arbeit durchaus vorkommt. Dennoch wird empfohlen, die Prinzipien der folgenden Leitlinien auch in solchen Fällen zu berücksichtigen.

5.1 Was ist eine informierte Einwilligung?

Die Berechtigung zur Verarbeitung von personenbezogenen Daten im Kontext von (KI-Einsatz) innerhalb der Sozialen Arbeit wird (gemäß Art. 6 Abs. 1 DSGVO) unter anderem zulässig, wenn die betroffene Person in informierter Weise eine Einwilligung bekundet. Bei personenbezogenen Daten der besonderen Kategorien (Art. 9 DSGVO, siehe Kapitel 4.1) muss diese Einwilligung (gemäß Art. 9 Abs. 2 DSGVO) in ausdrücklicher Weise, also unmissverständlich erfolgen. Die informierte

Einwilligung muss aktiv, freiwillig und ohne Druck bekundet werden und setzt voraus, dass die betroffene Person über die Verwendung der Daten aufgeklärt wurde. Dies muss in einfacher und verständlicher Sprache erfolgen, sodass die betroffene Person den konkreten Sachverhalt sowie alle vorgesehenen Verarbeitungsvorgänge verstanden hat. Außerdem kann die betroffene Person ihre Einwilligung jederzeit grundlos widerrufen (Kunitz, 2021). Während die informierte Einwilligung nach klassischen Autonomietheorien die Bereitstellung wesentlicher Informationen und die anschließende Nicht-Einmischung in die Entscheidung der Klientel erfordert, sollte der Prozess darüber hinaus so gestaltet werden, dass ihre Würde nicht untergraben und stattdessen die fantasievolle Reflexion verschiedener Optionen gefördert wird. Dies kann beispielsweise gelingen, indem gemeinsam verschiedene Szenarien unterschiedlicher Optionen, mit den entsprechenden Chancen und Risiken, durchgespielt werden. So können sich Klient:innen schließlich ermächtigt fühlen und für sich selbst sprechen (Stoljar, 2021).

5.2 Aufklärung über Chancen, Risiken und Alternativen

Damit Klient:innen eine informierte Einwilligung oder Ablehnung geben können, müssen sie über die Chancen, Risiken und Alternativen des KI-Einsatzes aufgeklärt werden. Diese Aufklärung muss in einfacher und leichter Sprache geschehen, sodass diese auch Personen ohne technisches Vorwissen verstehen. Dabei müssen alle vorgesehenen Verarbeitungsvorgänge transparent offengelegt werden sowie welche Entscheidungen und Prozesse konkret durch den KI-Einsatz beeinflusst werden oder werden könnten (Kunitz, 2021). Die Chancen und Vorteile des KI-Einsatzes unterscheiden sich je nach Einsatzkontext. Üblicherweise zielt dieser auf eine Steigerung der Effizienz, Qualität oder Personalisierung ab (siehe Kapitel 2). Klient:innen müssen auch darüber aufgeklärt werden, dass KI-Systeme möglichen Bias aufweisen oder dass die Verlässlichkeit und die Korrektheit der KI-Ausgaben fehlerhaft sein können (siehe Kapitel 3.1). Es müssen Strategien bestehen, um diesen Risiken entgegenzuwirken, die ebenfalls transparent dargestellt werden sollten, jedoch ohne die Risiken zu verharmlosen. Solche Strategien sind beispielsweise Strategien zum Datenschutz, klare Verteilung von Verantwortung oder regelmäßige Reflexion.

Neben der Erwähnung der Chancen und Risiken vom KI-Einsatz, sollten auch die Alternativen genannt werden. Den Klient:innen muss es möglich sein, den KI-Einsatz abzulehnen. Demensprechend sollte im Rahmen der informierten Einwilligung ebenfalls erläutert werden, was die Konsequenzen einer Ablehnung konkret bedeuten. Die besagten Konsequenzen können je nach Kontext der Situation unterschiedlich aussehen, sollten aber stets möglichst geringgehalten werden. Im Fall einer Ablehnung müssen die Klient:innen die Risiken des KI-Einsatzes nicht in Kauf nehmen,

profitieren aber demensprechend auch nicht von dessen Vorteilen und Chancen. Eine Alternative könnten hybride Lösungen sein, bei denen abgesprochen wird, dass der KI-Einsatz minimiert und eingeschränkt wird.

5.3 Einverständniserklärung und Widerruf

Wenn eine Einverständniserklärung (gemäß Art. 6 Abs. 1, Art. 7 DSGVO) für die Verarbeitung personenbezogener Daten vorausgesetzt wird, kann diese auf verschiedene Weisen erteilt werden. So ist neben der schriftlichen Einwilligung, beispielsweise auch eine elektronische oder ggf. sogar mündliche zulässig. Diese Einwilligung kann von der betroffenen Person nur gegeben werden, wenn die Willensbekundung in einer unmissverständlichen Form abgegeben wird, während die Person in informierter Kenntnis der Sachlage die personenbezogenen Daten für einen eindeutig festgelegten Zweck (Zweckbindung gemäß Art. 5 Abs. 1 DSGVO) freigibt (LfDI BW, 2024). Wichtig ist auch, dass der eigenen Nachweispflicht nachgekommen werden kann, weshalb in diesem Kontext ausdrücklich schriftliche oder elektronische Einverständniserklärungen empfohlen werden. Der konkrete Inhalt der Einverständniserklärungen muss an den spezifischen Anwendungskontext angepasst werden, weshalb an dieser Stelle keine allgemeingültigen Vorlagen möglich sind. Schon die Tatsache, ob KI-Systeme lokal ausgeführt werden oder über externe Dienstleister, hat eine Bedeutung für die Informations- und Sorgfaltspflichten.

Darüber hinaus ist es essenziell, dass den Klient:innen ihre datenschutzrechtlichen Rechte mitgeteilt und beachtet werden. Dazu gehört, dass sie ihre Einwilligung jederzeit widerrufen können (Art. 7 Abs. 3 DSGVO), was bedeutet, dass die personenbezogenen Daten nicht weiterverarbeitet werden dürfen. Außerdem haben sie ein Auskunftsrecht (Art. 15 DSGVO), durch welches sie jederzeit Auskunft über die verarbeiteten personenbezogenen Daten sowie die möglichen Empfänger dieser Daten verlangen können. Ein weiteres Recht ist die Datenübertragbarkeit (Art. 20 DSGVO), durch das gefordert werden kann, dass ihnen (oder anderen verantwortlichen Personen) die gespeicherten personenbezogenen Daten in einem maschinenlesbaren Format übermittelt werden müssen. Zudem besteht stets das Recht auf Berichtigung, Löschung und Einschränkung (Art. 16–18 DSGVO) der personenbezogenen Daten.

5.4 Zusammenfassung der Leitlinien

Die Klient:innen müssen über den Grund, die Chancen und Risiken des KI-Einsatzes sowie deren Alternativen informiert werden. Dies umfasst auch die konkrete Darlegung aller vorgesehenen Verarbeitungsabläufe, Entscheidungen und Prozesse, die KI involvieren und die durch diese beeinflusst werden könnten (siehe Kapitel 5.2). Ein besonderer Fokus liegt dabei auf einer grundsätzlich transparenten und neutralen Darstellung

aller relevanten Informationen. Die Kommunikation der Aufklärungsinhalte ist dabei so zu gestalten, dass sie in einfacher und leichter Sprache erfolgt und Sprachbarrieren berücksichtigt werden. Während dieses Prozesses sollte die Würde der Klient:innen stets gewahrt sowie deren fantasievolle Reflexion verschiedener Entscheidungen aktiv gefördert werden (siehe Kapitel 5.1). Die formale Einwilligung erfolgt durch eine Einverständniserklärung, deren Inhalt klar dargestellt werden muss. Die diesbezüglichen Rechte der Klient:innen, wie das Widerspruchsrecht, müssen dabei explizit vermittelt werden (siehe Kapitel 5.3). Zusätzlich sollten im Gesamtkontext des jeweiligen KI-Einsatzes stets die zuvor dargestellten ethischen, rechtlichen und fachlichen Rahmenbedingungen eingehalten (siehe Kapitel 3 und 4) werden. Insbesondere die in Kapitel 3.4 definierten Grundsätze sollten zwingend beachtet werden. Darüber hinaus müssen stets zu berücksichtigende Änderungen innerhalb der rechtlichen Grundlagen geprüft werden, da sich die Rechtsgrundlage im KI-Kontext aktuell im stetigen Wandel befindet.

6. Ausblick

Der KI-Einsatz in der Sozialen Arbeit unterliegt diversen Regularien und fordert ethische sowie fachliche Reflexion. Bemühungen seitens der Politik, wie die neue KI-VO, sind wichtige Schritte, um rechtliche Handlungssicherheit herzustellen und einen Rahmen für den KI-Einsatz zu definieren, in dem Betroffenenrechte gewahrt werden. Allerdings bleiben die konkreten Auswirkungen solcher Bemühungen zunächst in großen Teilen aus, weshalb zukünftige Entwicklungen abzuwarten sind. Darunter fallen besonders ungeklärte Haftungsfragen zur Verantwortlichkeit und die Auswirkungen von rechtlichen Entwicklungen, wie die Klarstellungen der EDSA (2024) (Kapitel 4.3) oder einschlägigen Präzedenzurteilen, die aufgrund des begrenzten Rahmens dieses Beitrags nicht näher beleuchtet werden konnten. Diese Umstände sorgen aktuell dafür, dass der KI-Einsatz in der Sozialen Arbeit ein heikles Themenfeld darstellt. Besonders KI-Systeme, die nach der KI-VO als Hochrisiko-KI verstanden werden, sollten daher zunächst noch nicht außerhalb von Pilotprojekten eingesetzt werden. Trotzdem ist das Potenzial der KI-Integration eindeutig gegeben, selbst wenn es sich nur um die KI-gestützte Transkription von Gesprächen und deren anschließende Dokumentation handelt. Zwar ist viel mehr als das möglich, jedoch bleibt es eine überwältigende Aufgabe, allen Regularien und fachlichen Grundsätzen gerecht zu werden. Dennoch ist dies eine Aufgabe, deren Bewältigung sich für die Soziale Arbeit lohnen dürfte.

Dabei sollten jedoch unbedingt die Fach-/Qualitätsstandards der Sozialen Arbeit berücksichtigt und gewahrt werden. Denn das Tripelmandat führt die Soziale Arbeit in ein Spannungsfeld zwischen Ökonomie und Qualität. KI-Systeme locken mit Prozessoptimierung, wobei die Gefahr besteht, dass die Qualitätsstandards der Profession oder sogar die Rechte und der Wille der Klientel zugunsten von Effizienz übergangen werden. So sollte

der Grundsatz gelten, dass vertrauensvolle menschliche Interaktionen, die Empathie, das Feingefühl für Nuancen und Werte wie Loyalität oder Zugewandtheit benötigen, nicht von KI-Systemen geleistet werden können (AvenirSocial, 2015; Fernando & Ranasinghe, Görder, 2021). Außerdem sollte die Klientel stets dazu ermächtigt werden, ihre Rechte und Interessen in diesem Kontext wahren zu können. Denn auch hier besteht die Gefahr, diese zugunsten einer besseren Ökonomie zu übergehen (AvenirSocial, 2015; Deutscher Berufsverband für Soziale Arbeit e. V., 2025; Molala & Mbaya, 2023). Zusammenfassend ist es also unabdingbar, dass sich Personen der Sozialen Arbeit bei der KI-Integration sowie dem KI-Einsatz für die Einhaltung der professionellen Fach-/Qualitätsstandards einsetzen und sich dagegen wehren, wenn diese übergangen werden.

Als Ausblick bleibt der dringliche Auftrag an die Soziale Arbeit, weitere ethische Reflexionen und konkrete Strategien zu entwickeln, damit sie den rasanten Entwicklungen der KI-Technologie entsprechen kann (Fernando & Ranasinghe 2023; Reamer, 2023). Ein konkretes Beispiel hierfür wäre die flüchtig angeführte Thematik der (freiwilligen) Einsatzbegrenzung von KI-Technologie. Denn auch wenn KI-Systeme zunehmend autonom agieren und in weiteren Aufgabenfeldern tätig sein können, sollten Einsatzmaxime und Grenzen definiert werden, um die Fach-/Qualitätsstandards der Profession zu wahren.

Literatur

- Alotaibi, J. O. & Alshahre, A. S. (2024). The role of conversational AI agents in providing support and social care for isolated individuals. *Alexandria Engineering Journal*, 108, 273–284. <https://doi.org/10.1016/j.aej.2024.07.098>
- Alpers, S., Krings, B. J., Becker, C., Rill, M. & Weinreuter, M. (2024). Ethische, rechtliche und soziale Aspekte (ELSA) der Gestaltung von KI-Systemen: Systematisierung der Betrachtung durch Vorgehensmodelle und Leitfäden. *Zeitschrift für Arbeitswissenschaft*, 78, 160–169. <https://doi.org/10.1007/s41449-024-00422-9>
- AvenirSocial. (2015). *Was ist gute Soziale Arbeit? Diskussionspapier von Avenir Social Schweiz zur Qualität in der Sozialen Arbeit*. https://avenirsocial.ch/wp-content/uploads/2018/12/AG_Qualitaet_DEF_D_1.pdf
- Bayerisches Landesamt für Datenschutzaufsicht (BayLDA). (2024). *Checkliste zur Bewertung von KI-Systemen gemäß Art. 22 DSGVO*. Abgerufen am 28. Juli 2025, von https://www.lda.bayern.de/media/ki_checkliste.pdf
- Dalziel, M. V., Schaffer, K. & Martin, N. (2024). *Navigating AI in social work and beyond: A multidisciplinary review*. arXiv. <https://doi.org/10.48550/arXiv.2411.07245>

- Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit (HmbBfDI). (2023). *Checkliste zur Bewertung von LLM-Chatbots gemäß DSGVO*. Abgerufen am 28. Juli 2025, von https://datenschutz hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/20231113_Checkliste_LLM_Chatbots_DE.pdf
- Der Landesbeauftragte für den Datenschutz und die Informationsfreiheit Baden-Württemberg (LfDI BW). (2024). *Rechtsgrundlagen Datenschutz & KI*. Abgerufen am 28. Juli 2025, von <https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/>
- Deutscher Berufsverband für Soziale Arbeit e. V. (2025). *Berufsethische Prinzipien und Leitlinien des DBSH* (2. Fassung). https://www.dbsh.de/media/public/dbsh-bund/Profession/2025-05-21_DBSH_Berufsethik_2._Fassung.pdf
- Deutscher Ethikrat. (2023). *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz*. <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>
- Dötterl, S. (2025). EU AI Act und Soziale Arbeit: Die KI-Verordnung und ihre Auswirkungen. In G. Linnemann, J. Löhe & B. Rottkemper (Hrsg.), *Künstliche Intelligenz in der Sozialen Arbeit: Grundlagen für Theorie und Praxis* (S. 221–240). Beltz Juventa. <https://doi.org/10.3262/978-3-7799-8562-4>
- Dutta, D. (2024). Descartes' legacy in AI ethics: Reconsidering Cartesian dualism in conceptualizing ethics of AI. *The Interdisciplinary Journal of Human and Social Studies*, 3(2).
- Europäischer Datenschutzausschuss (EDSA). (2024). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf
- Europäische Union. (2024). *Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) (Text von Bedeutung für den EWR)*. Amtsblatt der Europäischen Union, L 2024/1689. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32024R1689>
- Fernando, N. & Ranasinghe, P. (2023). Integration of artificial intelligence in social work: Opportunities, challenges, and considerations. *Journal of Computational Social Dynamics*, 8(9), 13–24. <https://vectorial.org/index.php/JCSD/article/view/35>

- Görder, B. (2021). Die Macht der Muster: Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch ‚künstliche Intelligenz‘. *EthikJournal*, 7(2), 1–22. https://www.ethikjournal.de/fileadmin/user_upload/ethikjournal/Texte_Ausgabe_2021_2/Goerder_Ethikjournal_2.2021.pdf
- Graumann, S. & Maaser, W. (2023). Ethik in der Sozialen Arbeit. In C. Neuhäuser, M. L. Raters & R. Stoecker (Hrsg.), *Handbuch Angewandte Ethik* (S. 557–561). J. B. Metzler. https://doi.org/10.1007/978-3-476-05869-0_77
- Grünewald, E. & Pallas, F. (2021). Datensouveränität für Verbraucher:innen: Technische Ansätze durch KI-basierte Transparenz und Auskunft im Kontext der DSGVO. In A. Boden, T. Jakobi, G. Stevens & C. Bala (Hrsg.), *Verbraucherdatenschutz – Technik und Regulation zur Unterstützung des Individuums* (Schriften der Verbraucherinformatik, Band 1). https://doi.org/10.18418/978-3-96043-095-7_02
- Industrie- und Handelskammer für München und Oberbayern (IHK München). (2025a, September). *Artificial Intelligence Act – das müssen Unternehmen wissen*. Abgerufen am 24. September 2025, von <https://www.ihk-muenchen.de/ratgeber/digitalisierung/kuenstliche-intelligenz/ai-act/>
- Industrie- und Handelskammer für München und Oberbayern (IHK München). (2025b, März). *KI und Datenschutz*. Abgerufen am 24. September 2025, von <https://www.ihk-muenchen.de/de/Service/Recht-und-Steuern/Datenschutz/ki/>
- Jafari-Kammann, H. (2025). *Künstliche Intelligenz und Menschenwürde: Untersuchungen im Umfeld von Mensch-Maschine-Interaktionen* [Dissertation, FernUniversität in Hagen]. deposit_hagen. https://ub-deposit.fernuni-hagen.de/servlets/MCRFileNodeServlet/mir_derivate_00002902/Diss_Jafari-Kammann_K%C3%BCnstliche_Intelligenz_und_Menschenw%C3%BCrde_2025.pdf
- Kompetenzplattform Künstliche Intelligenz Nordrhein-Westfalen. (2024). *Der AI Act – die KI-Verordnung der EU*. Abgerufen am 24. September 2025, von <https://www.ki.nrw/ai-act/>
- Kuckartz, U., & Rädiker, S. (2024). *Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und künstlicher Intelligenz* (6. Auflage). Juventa Verlag.
- Kunitz, S. (2023). *Haftung für künstliche Intelligenz? Der Rat des Algorithmus*. Herausgegeben von Mittelstand-Digital Zentrum Chemnitz. Abgerufen am 24. September 2025, von https://digitalzentrum-chemnitz.de/downloads/ng_ki-und-dsgvo/
- Kunitz, S. (2021). *KI-Anwendungen & DSGVO*. Herausgegeben von Mittelstand 4.0-Kompetenzzentrum Chemnitz. Abgerufen am 24. September 2025, von https://digitalzentrum-chemnitz.de/downloads/ng_ki-und-dsgvo/

- Linnemann, G., Löhe, J., & Rottkemper, B. (Hrsg.). (2025). *Künstliche Intelligenz in der Sozialen Arbeit: Grundlagen für Theorie und Praxis*. Beltz Juventa. <https://doi.org/10.3262/978-3-7799-8562-4>
- Linnemann, G. A., Löhe, J., & Rottkemper, B. (2023). Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit: Eine exemplarische arbeitsfeldübergreifende Betrachtung des Natural Language Processing (NLP). *Soziale Passagen*, 15, 197–211. <https://doi.org/10.1007/s12592-023-00455-7>
- Molala, T. S., & Mbaya, T. W. (2023). Social work and artificial intelligence: Towards the electronic social work field of specialisation. *International Journal of Social Science Research and Review*, 6(4), 613–621. <https://doi.org/10.47814/ijssrr.v6i4.1206>
- Nuwasiima, M., Ahonon, M. P. & Kadiri, C. (2024). The role of artificial intelligence (AI) and machine learning in social work practice. *World Journal of Advanced Research and Reviews*, 24(1), 80–97. <https://doi.org/10.30574/wjarr.2024.24.1.2998>
- Plafky, C., Kratz, N., Kuck, A. & Frischhut, H. (2022). KI-basierte Entscheidungsunterstützung in der Praxis Sozialer Arbeit. *Unsere Jugend*, 74, 115–121. <https://doi.org/10.2378/uj2022.art16d>
- Pelzl, J. (2025). IT-Sicherheit und Datenschutz im Kontext von KI-Sprachmodellen. In G. Linnemann, J. Löhe & B. Rottkemper (Hrsg.), *Künstliche Intelligenz in der Sozialen Arbeit: Grundlagen für Theorie und Praxis* (S. 204–220). Beltz Juventa. <https://doi.org/10.3262/978-3-7799-8562-4>
- Reamer, F. G. (2023). Artificial intelligence in social work: Emerging ethical issues. *International Journal of Social Work Values and Ethics*, 20(2), 52–71. <https://doi.org/10.55521/10-020-205>
- Steiner, O., & Tschopp, D. (2022). Künstliche Intelligenz in der Sozialen Arbeit. *Sozial Extra*, 46, 466–471. <https://doi.org/10.1007/s12054-022-00546-4>
- Stieler, M. (2021). Künstliche Intelligenz als Sozialarbeiterin. *neue caritas*. Deutscher Caritasverband e. V. <https://www.caritas.de/neue-caritas/heftarchiv/jahrgang2021/artikel/kuenstliche-intelligenz-als-sozialarbeiterin>
- Stieler, M., Domes, M., Burghardt, J. & Lehmann, R. (2024). KI: Innovation oder alter Wein in neuen Schläuchen? Herausforderungen und (An-)Fragen an die Soziale Arbeit als Beziehungsfprofession. *SLR*, (89).
- Stoljar, N. (2021). Informierte Einwilligung und relationale Konzepte von Autonomie. In N. Biller-Andorno, S. Monteverde, T. Krones & T. Eichinger (Hrsg.), *Medizinethik. Grundlagentexte zur angewandten Ethik* (S. 175–187). Springer VS. https://doi.org/10.1007/978-3-658-27696-6_9
- Xu, R., Sun, Y., Ren, M., Guo, S., Pan, R., Lin, H., Sun, L. & Han, X. (2024). AI for social science and social science of AI: A survey. *Information Processing & Management*, 61(3). <https://doi.org/10.1016/j.ipm.2024.103665>

Andrea Hahn¹ und Flemming Hansen¹

Algorithmen haben keine Empathie

KI als Kommunikationsbrücke und Reflexionshilfe in der Sozialen Arbeit

¹ Hochschule für Angewandte Wissenschaften Kiel

Zusammenfassung

Die fortschreitende Digitalisierung stellt die Soziale Arbeit vor neue Herausforderungen hinsichtlich Standardisierung, professioneller Handlungskompetenz und hinsichtlich des Einsatzes von Künstlicher Intelligenz (KI). Vor dem Hintergrund der Sozialen Arbeit als „Menschenrechtsprofession“ (Staub-Bernasconi, 2007) wird argumentiert, dass Empathie, Kontextsensibilität und geschulte Intuition zentrale Elemente professionellen Handelns bleiben, die nicht durch Technik ersetzt werden können. Zugleich eröffnet KI neue Möglichkeiten, Kommunikation zu erleichtern, Fachkräfte bei Routineaufgaben zu entlasten und Reflexion zu fördern. Der Beitrag diskutiert KI als Kommunikationsbrücke (1), als Werkzeug der Standardisierung (2) und als Reflexionshilfe (3). Abschließend wird betont, dass ihr Potenzial nur bei ethischer Rahmung, Datenschutz, Transparenz und Beteiligung der Adressat:innen verantwortungsvoll ausgeschöpft werden kann. KI kann Empathie unterstützen, aber nicht ersetzen.

Schlüsselwörter: Soziale Arbeit, Künstliche Intelligenz (KI), Kommunikationsbrücke, Reflexionshilfe, professionelle Urteilskraft, Empathie

Algorithms Lack Empathy. The Role of AI as a Communication Bridge and Reflective Tool in Social Work Practice

Abstract

Ongoing digitalisation presents Social Work with new challenges regarding professional competence, standardisation, and the use of Artificial Intelligence (AI). Building on the understanding of Social Work as a “human rights profession” (Staub-Bernasconi, 2007), this paper argues that empathy,

contextual sensitivity, and trained intuition remain core elements of professional practice that cannot be replaced by technology. At the same time, AI opens up new possibilities to facilitate communication, relieve practitioners from repetitive tasks, and support structured reflection. The paper discusses three central functions of AI in Social Work: as a communication bridge (1), as a tool for standardisation (2) and decision support, and as a reflective aid (3). It concludes that the potential of AI can only be realised through ethical frameworks ensuring data protection, transparency, and participation of service users. AI can assist professionals in promoting empathy and inclusion—but it cannot replace genuine human connection.

Keywords: Social Work, Artificial Intelligence (AI), communication bridge, reflective aid, professional judgement, empathy

1. Einleitung

Die Digitalisierung und der Einsatz von Künstlicher Intelligenz (KI) werfen grundlegende Fragen für die Soziale Arbeit auf. Diese sind jedoch nicht völlig neu, sondern knüpfen an bestehende Debatten darüber an, was professionelles Handeln ausmacht und inwieweit es standardisierbar ist. Aus professionstheoretischer Perspektive gilt die Soziale Arbeit als „Menschenrechtsprofession“ (Staub-Bernasconi, 2007b, S. 20) mit Fokus auf Beziehungsgestaltung und Kontextsensibilität (Staub-Bernasconi, 2018). Luhmann und Schorr (1982) beschrieben den Begriff des „strukturellen Technologiedefizits“ als grundlegende Herausforderung pädagogischer und sozialer Systeme: Anders als in technisch-funktionalen Systemen stehen in der Sozialen Arbeit keine zuverlässigen Mittel zur Verfügung, mit denen definierte Ziele sicher erreicht werden können. Der Hintergrund ist, dass sich Bewusstsein und Verhalten von Menschen nicht durch Kommunikation oder Interventionen sozialer Fachkräfte intentional determinieren lassen. Die Autoren sprechen in diesem Zusammenhang von „rudimentären Kausalplänen“ bzw. „subjektiven Technologien“ (Luhmann & Schorr, 1982, S. 19), die zwar handlungsleitend wirken, aber keine planbare Steuerung garantieren. Soziale Probleme sind daher nicht im technischen Sinne lösbar. Sie erfordern individuell ausgehandelte, relationale und kontextsensible Antworten im Zusammenspiel aller Beteiligten.

Einfühlsame Kommunikation und menschliche Empathie sind in der Sozialen Arbeit grundlegend (Schröder, 2022), denn Beziehungsgestaltung lebt von Resonanz, Perspektivenübernahme und Aushandlung. Künstliche Intelligenz kann dies – wenn überhaupt – nur simulieren. Dennoch hält sie zunehmend Einzug in die Soziale Arbeit, beispielsweise in Form algorithmischer Entscheidungsunterstützung (Bastian et al., 2020), digitaler Sprachverarbeitung (Linnemann et al., 2023) oder durch automatisierte Kommunikationssysteme wie Chatbots (Falck et al., 2023).

Der Begriff „Künstliche Intelligenz“ umfasst kein einheitliches technisches Verfahren, sondern eine Vielzahl unterschiedlicher Ansätze – von regelbasierten Entscheidungsbäumen bis hin zu lernenden Systemen wie neuronalen Netzen oder Verfahren der Sprachverarbeitung (Deutscher Ethikrat, 2021). Für die Soziale Arbeit ist die Unterscheidung zwischen einfachen und lernenden KI-Systemen besonders relevant, da sich daraus unterschiedliche Anforderungen an Transparenz, Nachvollziehbarkeit und ethische Kontrolle ergeben. Während regelbasierte Systeme in der Regel nachvollziehbar arbeiten, stellen lernende Modelle wie neuronale Netze insbesondere hinsichtlich möglicher Verzerrungen (Bias) und fehlender Erklärbarkeit (*Explainability*) spezifische Herausforderungen dar (Bastian et al., 2020; Deutscher Ethikrat, 2021).

Dieser Beitrag erörtert, welche Chancen KI in der Kommunikation der Sozialen Arbeit bietet, wo ihre Grenzen liegen und wie sie als Werkzeug zur Unterstützung professioneller Urteilsbildung, Teilhabe und Reflexion genutzt werden kann.

2. KI als Kommunikationsbrücke zur Nutzer:innenkommunikation

KI-gestützte Anwendungen, wie beispielsweise Chatbots, können als Kommunikationsbrücke dienen, indem sie Sprachbarrieren überwinden oder rund um die Uhr ansprechbar sind. Beispielsweise können Sprachtools im Kontakt mit geflüchteten Adressat:innen hilfreich sein, etwa durch Echtzeit-Übersetzungen in Beratungsgesprächen. Studien zeigen, dass solche Technologien Zugänge erleichtern, sprachliche Barrieren reduzieren und zur inklusiven Kommunikation beitragen können (Linnemann et al., 2023; Erdwiens, 2024; UNESCO, 2021).

Chatbots kommen bereits stellenweise zum Einsatz, um niederschwellige Informationen bereitzustellen. Bei diesen handelt es sich um Programme, die text- oder sprachbasierte, natürliche Interaktionen zwischen Mensch und Maschine simulieren (Steiner & Tschopp, 2022). Sie haben die Aufgabe, Menschen unterstützend zur Seite zu stehen, sei es beim Recherchieren von Inhalten oder in anderen Anwendungsfeldern. Sowohl im Berufs- als auch im Privatleben fungieren sie als digitale Assistent:innen und können durch ihre permanente Erreichbarkeit (24/7) überzeugen (Bruns & Kowald, 2023).

KI-gestützte Chatbots wurden bereits für den Erstkontakt mit Ratsuchenden erprobt, um niederschwellig Informationen bereitzustellen oder passende Ansprechpersonen zu empfehlen (Linnemann et al., 2023). Wenn KI-Systeme adressat:innengerecht gestaltet werden, können sie z. B. Entscheidungsfindungsprozesse unterstützen und zur Förderung digitaler Selbstbestimmung, Teilhabe und Empowerment beitragen – etwa durch barrierearme Zugänge, mehrsprachige Unterstützung oder individualisierte

Informationsbereitstellung (vgl. Breuer-Nyhse, 2023; Erdwiens et al., 2024; Kaminsky, 2021). Voraussetzung ist dabei, dass Chatbots auf umfangreiche Datenmengen zugreifen, natürliche Sprachverarbeitung nutzen und auf maschinellem Lernen basieren, sodass sie in unterschiedlichen Bereichen wirksame Unterstützung und Anleitung bieten (vgl. Meng et al., 2023). Dennoch bleibt festzuhalten: So hilfreich diese Systeme auch sind – zwischenmenschliche Empathie können sie nicht kompensieren. Chatbots mögen menschliche Kommunikation imitieren, doch sie ersetzen keine Vertrauensbeziehung. Als zusätzlicher Kommunikationskanal, etwa außerhalb üblicher Sprechzeiten oder für Menschen mit Hemmschwellen, können sie jedoch hilfreich sein (Linnemann et al., 2023).

3. KI als Unterstützung bei Fallbearbeitung und Standardisierung

KI-gestützte Anwendungen können Fachkräfte in der Sozialen Arbeit von repetitiven Aufgaben entlasten. Dazu zählen etwa die automatische Erstellung von Gesprächsprotokollen per Spracherkennung oder die Auswertung umfangreicher Fallakten durch KI-gestützte Textanalyse (Falck et al., 2023; Erdwiens et al., 2024; Breuer-Nyhse, 2023). Wenn KI im Hintergrund unterstützende Aufgaben wie Dokumentation oder Informationsbeschaffung übernimmt, kann dies Fachkräfte entlasten und mehr Zeit für Beziehungsarbeit schaffen. Studien stellen heraus, dass sich dadurch Effizienz und Zielgenauigkeit erhöhen lassen, ohne den menschlichen Kontakt zu verlieren, sofern KI als ergänzendes Werkzeug verstanden wird (Falck et al., 2023; Erdwiens et al., 2024; Mollick & Mollick, 2023).

Vor diesem Hintergrund werden international derzeit verschiedene KI-gestützte Systeme in der Sozialen Arbeit erprobt, insbesondere in Feldern, in denen Entscheidungen unter Unsicherheit getroffen werden müssen. Ein Beispiel ist das *Predictive Risk Modeling* (PRM) im Kinderschutz, bei dem vorhandene Falldaten genutzt werden, um Risiken für Kindeswohlgefährdungen algorithmisch zu prognostizieren. Studien deuten darauf hin, dass solche Tools Fachkräfte unterstützen können, indem sie individuelle Verzerrungen reduzieren und Entscheidungen nachvollziehbarer machen (Oak, 2016).

Dem gegenüber stehen bekannte Risiken wie algorithmische Verzerrungen oder sogenannte „Feedback Loops“ (Oak, 2016, S. 1215), die bestehende Benachteiligungen unbeabsichtigt reproduzieren können. KI in der Sozialen Arbeit kann je nach Kontext, Datenbasis und Steuerung sowohl befähigend als auch diskriminierend wirken (Deutscher Ethikrat, 2021). Ergänzend dazu bildet der EU AI Act (Europäische Union, 2024) einen verbindlichen Rechtsrahmen für den sicheren und vertrauenswürdigen Einsatz von KI in Europa. Er legt risikobasierte Anforderungen, Transparenzpflichten und das Prinzip der menschlichen Aufsicht fest, Aspekte,

die auch für den sozialarbeiterischen Kontext zentrale Bedeutung haben, um Diskriminierung und Machtasymmetrien vorzubeugen. Gleichzeitig zeigen sich klare Grenzen: PRM-Systeme arbeiten mit vereinfachenden Kategorien und können dazu verleiten, komplexe Lebenslagen auf schematische Risikowerte zu reduzieren. Sie dürfen deshalb nicht als Ersatz professionellen Urteils verstanden werden (Schneider & Seelmeyer, 2019). Die Verantwortung bleibt bei der Fachkraft, die Ergebnisse kritisch reflektieren und mit dem Fallkontext abgleichen muss.

4. KI als Reflexionshilfe zur Förderung geschulter Intuition

Der eigentliche Mehrwert von KI liegt vor diesem Hintergrund weniger in automatisierten Entscheidungen als in der Unterstützung von Kommunikation, Reflexion und mehrperspektivischem Arbeiten. Soziale Arbeit lebt von Aushandlungsprozessen, Beziehungsarbeit und Kontextsensibilität. Hier können insbesondere KI-Systeme aus dem Bereich der Sprachverarbeitung hilfreich sein, denn auch zwischen Fachkräften und anderen Beteiligten eines Falls kann KI als Brücke dienen. Große Sprachmodelle wie ChatGPT können darüber hinaus helfen, Informationsbarrieren in interdisziplinären Hilfesystemen und -netzwerken zu überwinden – etwa durch Übersetzungen oder die Übertragung komplexer Inhalte in unterschiedliche sprachliche Formate. Generative KI agiert dabei als Vermittlerin zwischen verschiedenen Fach- und Sprachkulturen, wodurch Zusammenarbeit erleichtert wird.

Neben der operativen und kommunikativen Unterstützung kann generative KI – insbesondere dialogorientierte Systeme wie ChatGPT – auch eine reflexive Unterstützungsfunktion übernehmen. Sabine Ader (2021) beschreibt professionelle Urteilsbildung in der Sozialen Arbeit als Zusammenspiel von planvollem, analytischem Denken und „geschulter Intuition“. Während sich analytisches Denken auf explizit verfügbares Wissen stützt, entsteht Intuition u. a. durch Erfahrung, Reflexion und kollegialen Austausch. Gigerenzer (2007) definiert Intuition als eine Form von erfahrungsbasierter Heuristik. Dies bedeutet, dass Intuition auf gesammelten Erfahrungen beruht und oft in der Lage ist, in komplexen und unsicheren Situationen schneller und effizienter zu einer guten Entscheidung zu kommen als eine ausführliche, rationale Analyse. Intuition greift auf unbewusste Mustererkennung und frühere Erlebnisse zurück, um rasch Handlungsoptionen zu bewerten. Diese Überlegung lässt sich durch den von Nørretranders (1998) eingeführten Begriff der „Exformation“ ergänzen. Exformation bezeichnet das implizite Wissen, das in unserem Denken und Handeln eine zentrale Rolle spielt. Es handelt sich um Informationen, die zwar verarbeitet wurden und unsere Entscheidungen und unsere Kommunikation beeinflussen, uns jedoch nicht direkt bewusst sind. Dieses implizite Wissen bleibt oft unausgesprochen, bildet aber die Grundlage für intuitive Urteile.

Aus dieser Perspektive betrachtet, können KI-Systeme eine wichtige strukturierende Funktion übernehmen. Sie können Fachkräfte dabei unterstützen, ihr implizites Erfahrungswissen zu identifizieren und in expliziter Form darzustellen. Dies kann beispielsweise durch datenbasierte Analysen oder Mustererkennung geschehen, die unbewusste Entscheidungsgrundlagen sichtbar machen. Zudem können KI-Systeme – wie z. B. für Fallarbeit spezifizierte Chatbots – das Überprüfen alternativer Sichtweisen und das Durchspielen von hypothetischen Szenarien ermöglichen, um Entscheidungsprozesse zu optimieren. So tragen sie dazu bei, sowohl intuitive als auch rationale Entscheidungsstrategien zu verbessern.

Generative KI kann solche impliziten Prozesse durch dialogische Rückfragen, hypothetisches Probehandeln oder Perspektivwechsel sichtbar und reflektierbar machen. Sie ersetzt keine Kolleg:innen oder Supervision – kann aber als ergänzendes Werkzeug zur Schulung professioneller Urteilskraft beitragen. In kasuistischen Lernprozessen, auch zur Unterstützung „analoger“ kollegialer Beratung oder im Studium und in der Ausbildung, kann sie als Reflexionshilfe wertvoll sein.

Solche reflexiven KI-Anwendungen könnten künftig gezielt für die Soziale Arbeit entwickelt werden, etwa in Form von spezialisierten Chatbots zur Fallreflexion, die Fachkräften strukturierte Rückfragen stellen und methodische Perspektiven vorschlagen. „Practice Panels“ in Sinne digitaler Reflexionsumgebungen, die beispielweise auf Fachwissen, Falltypen und Theorien der Sozialen Arbeit trainiert sind, könnten Fachkräfte bei der Analyse komplexer Entscheidungssituationen unterstützen. Damit ließen sich kasuistische, intuitive und analytische Perspektiven produktiv miteinander verbinden – stets im Bewusstsein, dass KI keine normative Letztinstanz darstellt, sondern ein Werkzeug zur Anregung professioneller Urteilskraft ist.

Gleichwohl bleibt festzuhalten, dass KI keine genuine Empathie besitzt. Sie erzeugt sprachliche Muster, ohne Bedeutung tatsächlich zu verstehen. Bender et al. (2021) beschreiben große Sprachmodelle als „stochastic parrots“, die zwar sprachlich kohärent agieren, aber ohne Bewusstsein oder Intentionalität funktionieren. In diesem Sinne können KI-Systeme empathische Kommunikation nur *simulieren*, nicht erleben. Turkle (2017) bezeichnet dies als „synthetische Nähe“, die emotionale Resonanz technisch nachbildet, ohne echtes Mitgefühl zu enthalten.

5. Rahmenbedingungen für den ethischen und professionellen Einsatz

Die Wirkung von Künstlicher Intelligenz (KI) in der Sozialen Arbeit ist wesentlich vom organisationalen Kontext abhängig. Ob ein KI-System befähigend oder restriktiv wirkt, entscheidet sich nicht allein anhand der Technologie selbst, sondern an der jeweiligen institutionellen Struktur

(Breuer-Nyhsen, 2023). In kontrollorientierten Einrichtungen besteht die Gefahr, dass KI primär zur Überwachung und Standardisierung eingesetzt wird – während in lernoffenen, dialogisch ausgerichteten Organisationen Potenziale für kooperative Aushandlungsprozesse und reflexives Arbeiten stärker genutzt werden können (Kaminsky, 2021). Eine ethisch vertretbare und fachlich sinnvolle Implementierung von KI erfordert daher die aktive Mitwirkung von Fachkräften in der Entwicklung und Einführung entsprechender Systeme (Breuer-Nyhsen, 2023). Zugleich müssen grundlegende Prinzipien wie Datenschutz, Transparenz, Partizipation sowie der Schutz der Rechte der Adressat:innen gewährleistet sein (Deutscher Ethikrat, 2021). Nur wenn die Anwendung kontextsensibel erfolgt – also auf die Lebenslagen, Bedürfnisse und Machtverhältnisse vor Ort abgestimmt ist – kann KI in der Sozialen Arbeit professionell und verantwortungsvoll eingesetzt werden (Kaminsky, 2021; Breuer-Nyhsen, 2023).

Ein verantwortungsbewusster Umgang mit KI-Tools ist insofern für die weitere Entwicklung der Profession und Disziplin Soziale Arbeit von entscheidender Bedeutung. Diese sind zweifellos in der Lage, fachliche Standards zu setzen oder zu verändern. Wie Hansen (2010) betont, sind Standards in der Sozialen Arbeit „weder Himmelsgabe noch Teufelswerk, sondern konstitutionelle Steuerungsinstrumente in sozialen Systemen“ (S. 163). Allerdings können – je nach organisationsbezogener Rahmung – grundsätzlich Risiken im Sinne von sozialtechnischen Handlungsoptionen zum Tragen kommen. Fachliche Standards können aber auch kommunikativ-orientierende Funktionen erfüllen, wenn sie vorrangig als „Kommunikationsbrücken“, beispielsweise zwischen Fachlichkeit, Organisation und Politik, verstanden und eingesetzt werden. Diese Funktion von Standards zielt auf institutionelle Verständigung ab und erfordert, dass sie normativ verhandelt und legitimiert sind. KI-Systeme operieren jedoch auf der Grundlage von Datenstrukturen, u. U. ohne (professions-)ethisch begründete Maßstäbe. Der Einsatz von KI verlangt daher eine reflexive Rahmung, eine klare professionsbezogene Grundlegung sowie eine Anschlussfähigkeit an Theorien und Methoden der Sozialen Arbeit, um keine Technologisierung des Sozialen zu bewirken.

Fazit

KI kann in der Sozialen Arbeit eine Brückenfunktion erfüllen, indem sie Kommunikation erleichtert, Informationen zugänglicher macht und Fachkräfte von Routineaufgaben entlastet. Insbesondere in einem von Vielfalt geprägten Feld bietet der gezielte Einsatz von KI die Chance, inklusiver zu arbeiten – zum Beispiel mehrsprachige Adressat:innen besser einzu beziehen oder rund um die Uhr erreichbar zu sein. Diese Chancen gilt es zu nutzen, um Angebote niedrigschwelliger und effizienter zu gestalten. Gleichzeitig sind Empathie und ethisches Handeln nicht automatisierbar, und die Technik darf nicht zum Selbstzweck werden. Sozialarbeiter:innen tragen die Verantwortung, mit KI zu arbeiten, ohne sich von ihren Vor-

gaben blind leiten zu lassen. Wo KI vorgeschaltet ist – etwa in Form eines Chatbots als erste Anlaufstelle –, muss im Bedarfsfall schnell eine menschliche Fachkraft ansprechbar sein. Menschlichkeit und Zuwendung bleiben der nicht-technische Kern der Sozialen Arbeit, und KI sollte sie unterstützen, nicht verdrängen.

Zugleich zeigen aktuelle Studien, dass Nutzer:innen Interaktionen mit KI-Systemen, insbesondere mit Chatbots, teilweise als empathisch wahrnehmen, obwohl keine genuine Empathie vorhanden ist (Meng et al., 2023; Liu et al., 2024). Diese sogenannte „wahrgenommene Empathie“ kann Vertrauen und Zugänglichkeit fördern, etwa in Situationen, in denen Hemmschwellen oder Sprachbarrieren bestehen. Für die Soziale Arbeit stellt sich damit die Frage, wie solche techno-empathischen Zuschreibungen professionell eingeordnet werden können. Entscheidend bleibt, dass die Beziehung zwischen Menschen nicht durch eine vermeintlich empathische Maschine ersetzt, sondern durch deren gezielten Einsatz reflektiert unterstützt wird.

Schließlich sind Algorithmen Werkzeuge, keine Kolleg:innen. Sie haben keine Empathie, kein implizites Wissen und keine wertebasierte Urteilskraft. Daher sollte KI in der Sozialen Arbeit stets mit Augenmaß, Kompetenz und Reflexion eingebunden werden. Gelingt dies, kann sie viele Menschen besser mit den Hilfen verbinden, die sie benötigen – ohne die Beziehung zwischen Menschen zu ersetzen.

KI kann in der Sozialen Arbeit zur Entscheidungsunterstützung und Standardisierung beitragen. Besonders wertvoll ist sie jedoch als Kommunikationshilfe und Reflexionsinstrument. Ob sie die Profession ergänzt oder substituiert, hängt weniger von der technischen Entwicklung ab, sondern vielmehr von der Professionalität der Fachkräfte und den institutionellen Rahmenbedingungen. Algorithmen haben keine Empathie – sie können Resonanz nur simulieren, nicht erleben. Sie können jedoch helfen, diese in der Fallarbeit (weiter-) zu entwickeln.

Literatur

- Ader, S. (2021). Analytischer „Scharfsinn“ und geschulte Intuition im Dialog. *Sozial Extra*, 45(4), 245–250. <https://doi.org/10.1007/s12054-021-00394-8>
- Bastian, P., Schrödter, M., & Taylor, B. (2020). Risikodiagnostik und Big Data Analytics in der Sozialen Arbeit. In N. Kutscher, T. Ley, U. Seelmeyer, F. Siller, A. Tillmann & I. Zorn (Hrsg.), *Handbuch Soziale Arbeit und Digitalisierung* (S. 255–264). Beltz Juventa.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). ACM. <https://doi.org/10.1145/3442188.3445922>

- Breuer-Nyhsen, J. (2023). Digitale Transformation in der Sozialen Arbeit – Potenziale und Grenzen des Technologieeinsatzes aus professionsethischer Sicht [Dissertation]. Universität Vechta.
- Bruns, J., & Kowald, L. (2023). Praxisleitfaden Chatbots. Springer Fachmedien Wiesbaden.
- Deutscher Ethikrat. (2021). Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>
- Erdwiens, M., & Institut für Sozialstrategie. (2024). KI in der Sozialwirtschaft – Einsatz, Haltung, Entwicklung (Studie im Auftrag der Diakonie Deutschland). Institut für Sozialstrategie. <https://www.institut-fuer-sozialstrategie.de>
- Europäische Union. (2024). *Verordnung (EU) 2024/1689 über Künstliche Intelligenz (AI Act)*. Amtsblatt der Europäischen Union. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32024R1689>
- Falck, N., Alles, L., Flick, U., & Schulz, K. (2023). *KI-Kompetenzmodell für die Soziale Arbeit (Projektbericht im Verbundvorhaben KIWA)*. Hochschule Esslingen. https://kiwa-projekt.de/wp-content/uploads/2023/12/KI-Kompetenzmodell_VK-KIWA_Falck_Alles_Flick_Schulz_v1_1.pdf
- Gigerenzer, G. (2007). *Bauchentscheidungen: Die Intelligenz des Unbewussten und die Macht der Intuition*. Goldmann.
- Hansen, F. (2010). *Standards in der Sozialen Arbeit*. Deutscher Verein für öffentliche und private Fürsorge e. V.
- Kaminsky, C. (2021). Digitale Transformation Sozialer Arbeit? – Ethische Orientierungen auf neuem Terrain. *EthikJournal*, 7(2), 1–21.
- Linnemann, G. A., Löhe, J., & Rottkemper, B. (2023). Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit: Eine exemplarische arbeitsfeldübergreifende Betrachtung des Natural Language Processing (NLP). *Soziale Passagen*, 15, 197–211. <https://doi.org/10.1007/s12592-023-00455-7>
- Luhmann, N., & Schorr, K. E. (1982). Das Technologiedefizit in der Erziehung und die Pädagogik. In N. Luhmann & K. E. Schorr (Hrsg.), *Zwischen Technologie und Selbstreferenz: Fragen an die Pädagogik* (S. 11–40). Suhrkamp.
- Liu, T., Giorgi, S., Aich, A., et al. (2024). *The Illusion of Empathy: How AI Chatbots Shape Conversation Perception*. arXiv preprint arXiv:2411.12877. <https://arxiv.org/abs/2411.12877>
- Meng, J., Rheu, M., Zhang, Y., Dai, Y., & Peng, W. (2023). Mediated social support for distress reduction: AI chatbots vs. human. *Proceedings of the ACM on Human-Computer Interaction*, 7, 1–25. <https://doi.org/10.1145/3579505>
- Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. *SSRN*. <https://doi.org/10.2139/ssrn.4377900>

- Nørretranders, T. (1998). *Spüre die Welt: Die Wissenschaft des Bewusstseins*. Rowohlt.
- Oak, E. (2016). A minority report for social work? The predictive risk model (PRM) and the Tuituia assessment framework in addressing the needs of New Zealand's vulnerable children. *British Journal of Social Work*, 46(5), 1208–1223. <https://doi.org/10.1093/bjsw/bev028>
- Schneider, D., & Seelmeyer, U. (2018). Der Einfluss der Algorithmen. *Sozial Extra*, 42(3), 21–24. <https://doi.org/10.1007/s12054-018-0046-y>
- Schröder, C. (2022). Wunderressource Empathie? *Sozial Extra*, 46, 350–355. <https://doi.org/10.1007/s12054-022-00520-0>
- Staub-Bernasconi, S. (2018). *Soziale Arbeit als Handlungswissenschaft: Auf dem Weg zu kritischer Professionalität* (2., überarb. u. erw. Aufl.). Barbara Budrich / UTB.
- Staub-Bernasconi, S. (2007b). Soziale Arbeit: Dienstleistung oder Menschenrechtsprofession? Zum Selbstverständnis Sozialer Arbeit in Deutschland mit einem Seitenblick auf die internationale Diskussionslandschaft. In A. Lob-Hüdepohl & W. Lesch (Hrsg.), *Ethik Sozialer Arbeit – Ein Handbuch* (S. 17–53). Schäffer.
- Steiner, O., & Tschopp, D. (2022). Künstliche Intelligenz in der Sozialen Arbeit. *Sozial Extra*, 46(6), 466–471. <https://doi.org/10.1007/s12054-022-00546-4>
- Turkle, S. (2017). *Reclaiming Conversation: The Power of Talk in a Digital Age*. Penguin.
- UNESCO (2021). *AI and education: Guidance for policy-makers*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf00000376705>

Dennis Klinkhammer¹

Erklärbare Künstliche Intelligenz für die Soziale Arbeit

¹ FOM – Hochschule für Oekonomie und Management

Zusammenfassung

Künstliche Intelligenz (KI) verspricht durch eine automatisierte Entscheidungsunterstützung, präzise Vorhersagemodelle und personalisierte Interventionsangebote Effizienzsteigerungen und eine verbesserte Versorgung von Klient:innen im Kontext der Sozialen Arbeit. Gleichzeitig gilt es, die möglichen Herausforderungen und Risiken adäquat zu berücksichtigen. Bisherige Beiträge zur KI in der Sozialen Arbeit weisen dabei stellenweise ein wenig differenziertes Verständnis von KI-Systemen sowie deren statistischen Grundlagen auf. So bspw. in Form von nicht trennscharfen Definitionen, die für eine versierte Anwendung von KI-Systemen sowie für eine kritische Auswahl unter bestehenden KI-Systemen und zur Entwicklung zielgruppenspezifischer KI-Systeme essenziell wären. Im bisherigen wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit wird allerdings bereits explizit auf die Erfordernis statistischer Grundlagen verwiesen, um eine erklärbare KI in der Sozialen Arbeit etablieren zu können, ohne diese jedoch im Detail zu explizieren. Daran anknüpfend erfolgt in diesem Beitrag eine Ausdifferenzierung des maschinellen Lernens in die statistischen Grundlagen, bestehend aus den Regressions-, Klassifikations-, Clustering- sowie Dimensionsreduktionsverfahren sowie Einblicke in die Funktionsweise von neuronalen Netzwerken für das sogenannte Deep Learning mit einem einschlägigen Anwendungsbeispiel zu erklärbarer KI. Dabei wird die zentrale Rolle von Algorithmen sowie der Weights und Biases geklärt, welche gemeinsam mit der Transformer-Architektur als Grundlage von Large Language Models, wie bspw. ChatGPT zu verstehen sind. Insbesondere die statistischen Grundlagen erweisen sich dabei als essenziell für einen gleichberechtigten Dialog von Sozialarbeiter:innen mit Stakeholdern der Digitalisierung und KI und können zur Entwicklung zielgruppenspezifischer KI-Systeme für die Ausbildung und Praxis der Sozialen Arbeit beitragen.

Schlüsselwörter: Künstliche Intelligenz, Machine Learning, Deep Learning, Algorithmen, Soziale Arbeit

Klinkhammer, D. (2026). Erklärbare Künstliche Intelligenz für die Soziale Arbeit. *DiMe-Sozial*, 2(1), 55–74. <https://doi.org/10.21240/dimesoz/2026/1/58>

Explainable Artificial Intelligence for Social Work

Abstract

Artificial intelligence (AI) promises increased efficiency and improved client care in the context of social work through automated decision support, precise predictive models, and personalized interventions. At the same time, it is crucial to adequately consider the potential challenges and risks. Previous contributions to AI in social work sometimes exhibit a somewhat vague understanding of AI systems and their statistical foundations. This is evident, for example, in the lack of precise definitions, which would be essential for the proficient application of AI systems, for critically selecting from existing systems, and for developing target group-specific AI systems. However, the existing academic discourse on AI in social work explicitly acknowledges the need for statistical foundations to establish explainable AI in social work, without, however, elaborating on these foundations in detail. Building on this, this article differentiates machine learning into its statistical foundations, consisting of regression, classification, clustering and dimensionality reduction methods, as well as providing insights into the functioning of neural networks for so-called deep learning with a relevant application example of explainable AI. This clarifies the central role of algorithms, weights, and biases, which, together with the transformer architecture, form the basis of large language models such as ChatGPT. In particular, the statistical foundations prove essential for an equitable dialogue between social workers and stakeholders in digitalization and AI, and can contribute to the development of target group-specific AI systems for the education and practice of social work.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Algorithms, Social Work

1. Einleitung

1.1 Künstliche Intelligenz in der Sozialen Arbeit

Die Integration von Künstlicher Intelligenz (KI) in die Soziale Arbeit gewinnt nicht zuletzt durch die zunehmende Digitalisierung an Bedeutung und eröffnet sowohl Chancen als auch Herausforderungen für die Profession (Klinger et al., 2022; Kutscher et al., 2020). Aktuelle Diskussionen fokussieren bspw. KI-Systeme, um zukünftige Ereignisse vorherzusagen und präventiv einzugreifen zu können. Allerdings zeigen vorläufige Forschungsergebnisse, dass der Einsatz von KI-Systemen in der Sozialen Arbeit nicht zwangsläufig zu einer fachlichen Aufwertung führt (Stummbaum & Rusert, 2021). Im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit wird u. a. das Black-Box-Problem einiger KI-Systeme als Herausforderung für eine transparente und zielgruppenspezifische Soziale Arbeit gesehen (Görder, 2021). Die in manchen Fällen mit KI-Systemen einhergehenden Intransparenzen hinsichtlich der zugrunde-

liegenden Entscheidungsprozesse erschweren es demnach professionellen Akteur:innen aus dem Kontext der Sozialen Arbeit, die Ergebnisse nachzuvollziehen und kritisch zu hinterfragen. Dies steht im Widerspruch zu den ethischen Grundsätzen der Sozialen Arbeit, die Transparenz und Nachvollziehbarkeit in Entscheidungsprozessen fordern (Görder, 2021). Zudem können KI-Systeme bestehende gesellschaftliche Ungleichheiten reproduzieren oder sogar verstärken, da sie in der Regel auf historischen Daten basieren, die bereits entsprechende Verzerrungen enthalten können. Dies stellt eine erhebliche ethische Herausforderung dar, insbesondere, wenn vulnerable Gruppen betroffen sind.

Dabei zeichnet sich im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit bisweilen eine Unschärfe ab, die sowohl das Verständnis von KI betrifft, als auch deren Funktionsweise. In einigen Beiträgen wird KI-Systemen bspw. Autonomie zugeschrieben:

„Algorithmen, die gegenwärtig zum Einsatz kommen, sind in der Lage, sich selbst zu verändern, indem sie sich dynamisch und adaptiv auf eigene Operationen sowie die Umwelt beziehen. So können Algorithmen im Grunde Algorithmen schreiben, ohne dass hierbei Programmierer:innen von außen in den Rechenablauf eingreifen.“ (Wunder, 2021, S. 44).

Diese Einschätzung wird der Funktionsweise von Algorithmen dahingehend nicht gerecht, dass diese sich ausschließlich innerhalb eines vorgegebenen Rahmens verändern können. Algorithmen haben in der Regel eine vorgegebene Funktionsweise, bspw. als Aktivierungsfunktionen in neuronalen Netzwerken beim sogenannten *Deep Learning* (LeCun et al. 2015), wobei die Anpassung der KI-Systeme durch statistisch vorgegebene Optimierungsprozesse, wie bspw. dem Gradientenabstiegsverfahren, auch bekannt als *Gradient Descent*, basiert. Eine kreative Neukonstruktion der Algorithmen liegt dabei nicht vor, lässt sich aber stellenweise dem wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit entnehmen (Steiner & Tschopp, 2022; Schneider & Seelmeyer, 2018). Selbst bei komplexen KI-Systemen müssen menschliche Expert:innen in der Regel die Problemstellung und die Leistungskennzahlen definieren und die Algorithmen überwachen. Auch werden in den Beiträgen zu KI-Systemen stellenweise negative Aspekte gehoben, ohne dass die positiven Aspekte gleichermaßen benannt werden. Bezugnehmend auf gesellschaftliche Ungleichheiten lässt sich dem wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit bspw. folgende Aussage entnehmen: „Diese werden durch KI-Algorithmen reproduziert und verstärkt, sodass es zu ungerechtfertigter Diskriminierung kommt“ (Görder, 2021, S. 7) – obschon bereits mehrere algorithmenbasierte Ansätze in der Praxis dazu verwendet werden, gesellschaftliche Ungleichheiten zu erkennen und diesen entgegenzuwirken (Burton, 2023). Eine gewisse Skepsis scheint sich entsprechend mit Beiträgen wie „Zukünfte Sozialer Arbeit – digital und wie bei Ikea.

Szenarien Sozialer Arbeit in der Digitalisierung“ (Stummbaum & Rusert, 2021) in den wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit eingeschlichen zu haben.

Diese und weitere Unschärfen können teilweise auf eine fehlende Berücksichtigung der methodischen Grundlagen der KI, insbesondere der Statistik, zurückgeführt werden. Viele Algorithmen, insbesondere im Bereich des maschinellen Lernens, basieren dabei auf statistischen Verfahren (Trabs et al., 2021), bspw. dem Regressionsverfahren – dieses gilt sogar als einer der am häufigsten verwendeten Algorithmen im maschinellen Lernen. Ergänzend können Klassifikationsverfahren, Clusteringverfahren und Verfahren zur Dimensionsreduktion genannt werden, die nicht nur in vielen akademischen Ausbildungen im Curriculum verankert sind (Ärlebäck & Kawakami, 2023; Friedrich et al., 2022), sondern auch eine entsprechende Grundlage für ein Verständnis der methodischen Grundlagen der KI bilden (Klinkhammer & Keller, 2024). Insbesondere die Statistik ermöglicht dabei ein fundiertes Verständnis von Datenanalysen und -interpretationen, welche für den verantwortungsvollen Einsatz von KI-Systemen unerlässlich sind. So schult Statistik bspw. die Fähigkeit, Unsicherheiten zu quantifizieren und die Gültigkeit von Modellen zu beurteilen (Ärlebäck & Kawakami, 2023). Dieses kritische Denken ist somit entscheidend, um die Möglichkeiten und Grenzen von KI-Systemen im Kontext der Sozialen Arbeit realistisch einzuschätzen und ethische sowie gesellschaftliche Implikationen zu berücksichtigen. Im wissenschaftlichen Diskurs zu KI – nicht nur in der Sozialen Arbeit – scheinen diese Grundlagen allerdings unterrepräsentiert und wenig ausdifferenziert.

Angesichts dieser Herausforderungen scheint es unerlässlich, dass die Soziale Arbeit ein fundiertes Verständnis der methodischen Grundlagen von KI entwickelt. Nur so kann sichergestellt werden, dass der Einsatz von KI-Systemen die professionelle Verantwortung und die ethischen Standards der Sozialen Arbeit unterstützt (Kutscher et al., 2020), anstatt sie zu untergraben. Die kritische Auseinandersetzung mit den methodischen Grundlagen von KI scheint daher nicht nur eine akademische Notwendigkeit, sondern auch eine praktische Voraussetzung für eine zielführende und gerechte Soziale Arbeit im digitalen Zeitalter.

1.2 Aufbau und Fragestellung

Ausgehend von der in der Einleitung bereits angedeuteten Auseinandersetzung der Sozialen Arbeit als Profession mit KI-Systemen sowie deren zugrundeliegende Funktionsweise ist das vornehmliche Ziel dieses Beitrags eine Einführung in die statistischen Grundlagen der KI (Rashidi et al., 2023; Trabs et al., 2021; Wennker, 2020), um zu den Entwicklungen auf dem Gebiet der erklärbaren KI (Haque et al., 2023; Meske et al., 2020) überleiten zu können und in ein entsprechendes Anwendungsbeispiel zu überführen. Zu diesem Zweck wird zunächst in den

theoretischen Grundlagen die methodische Anschlussfähigkeit des wissenschaftlichen Diskurses zu KI in der Sozialen Arbeit herausgearbeitet. Ergänzend werden die im Rahmen einer kumulativen Habilitation systematisch eruierten Fortbildungsmöglichkeiten zu den statistischen Grundlagen der KI vorgestellt, die für den wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit sowie einer curricularen Verankerung zweckdienlich sein können. In den zugrundeliegenden Fragestellungen des Beitrags geht es somit insbesondere um die ausdifferenzierte Darstellung der statistischen Grundlagen für ein grundlegendes Verständnis von KI-Systemen sowie um die damit verbundenen Chancen, die sich im Sinne erklärbarer KI-Systeme für die Soziale Arbeit ergeben.

2. Entwicklung eines professionsbezogenen KI-Verständnisses

Die Auseinandersetzung mit KI in der Sozialen Arbeit hat sich in den letzten Jahren gewandelt – von eher allgemeinen, teils noch unscharf gefassten Definitionen hin zu differenzierteren Darstellungen, die statistische Grundlagen und methodische Aspekte stärker einbeziehen. Dabei zeigt sich, dass das Verständnis von KI-Systemen an Präzision gewonnen hat, auch wenn einzelne Formulierungen gelegentlich vereinfachend oder verallgemeinernd wirken. Um diese Entwicklung zumindest in ihren Grundzügen nachzeichnen zu können, erfolgt zunächst ein exemplarischer Abriss des wissenschaftlichen Diskurses zu KI in der Sozialen Arbeit zwischen den Jahren 2018 bis 2025. Das Jahr 2018 wurde deshalb als Anfangszeitpunkt der Betrachtungen gewählt, weil kurz zuvor mit der Vorstellung der sogenannten Transformer-Architektur (Vaswani et al., 2017) ein wesentlicher Grundstein für Large Language Models – wie bspw. ChatGPT – gelegt worden ist. Dadurch wurde ein scheinbar niedrigschwelliger Zugang zu KI geschaffen, in dessen Folge die interdisziplinäre Wahrnehmung von KI-Systemen und deren Verbreitung zugenommen hat, ohne dass sich dabei die statistischen Grundlagen verändert haben.

Bereits die frühen Ansätze, wie bspw. bei Schneider und Seelmeyer (2018), verstehen KI als Technologie, die auf maschinellem Lernen basiert und neuronale Netzwerke nutzen kann, „um unstrukturierte Daten zu analysieren, Muster zu erkennen und eigenständige Klassifikationen zu entwickeln“ (Schneider & Seelmeyer, 2018, S. 22). Dieser Beitrag differenziert bereits zwischen regelbasierten Algorithmen und datengetriebenen Outputs und verweist entsprechend früh im wissenschaftlichen Diskurs über KI in der Sozialen Arbeit auf die zugrundeliegenden „Statistischen Berechnungen“ (Schneider & Seelmeyer, 2018, S. 21) entsprechender KI-Systeme. Der Schwerpunkt liegt dabei allerdings auf den sogenannten Klassifikationsverfahren, während andere Anwendungsbereiche – etwa Regressions- oder Clusteringverfahren – weniger im Blickfeld stehen. Die ethische Dimension, wie mögliche Stigmatisierung durch datenbasierte Prognosen, wird ebenfalls reflektiert und setzt sich in nachfolgenden Beiträgen entsprechend fort (Remke & Schiffhauer, 2025; Kutscher et al.,

2020). Exemplarisch für die weitere Entwicklung des wissenschaftlichen Diskurses zu KI in der Sozialen Arbeit knüpft Görder (2021) an diesem Verständnis von KI an und verweist darüber hinaus auf „Algorithmen für spezifische, abgegrenzte Problemstellungen, die zu diesem Zweck lernfähig programmiert worden sind“ und „auf der Grundlage großer Datensätze z. B. mit den Methoden des maschinellen Lernens trainiert werden und so selbst Verfahren der Mustererkennung entwickeln“ (Görder, 2021, S. 2). Hier rückt vermehrt die Fähigkeit zur Mustererkennung in den Vordergrund: „Künstliche Intelligenz identifiziert Muster, also Korrelationen und Regelmäßigkeiten in unserer Gesellschaft in einem bisher nicht gekannten Ausmaß“ (Görder, 2021, S. 16). Die Betonung der Korrelation macht allerdings deutlich, dass die Differenz zwischen beschreibenden und prognostischen Verfahren zu diesem Zeitpunkt noch nicht trennscharf gezogen wurde, was sich entsprechend in anderen Beiträgen widerspiegelt (Telkmann, 2023; Wunder, 2021). Mit dem sogenannten Black-Box-Problem neuronaler Netzwerke wird zudem ein Thema aufgegriffen, das methodische Fragen der Nachvollziehbarkeit von KI-Entscheidungen berührt, auch wenn dessen technische Hintergründe noch nicht im Detail entfaltet werden. Bereits ein Jahr später stellen Steiner und Tschopp (2022) KI als breites, interdisziplinäres Teilgebiet der Informatik dar. Sie integrieren zentrale Konzepte wie Algorithmen, maschinelles Lernen, neuronale Netzwerke und *Deep Learning* und illustrieren dies am Beispiel des Predictive Risk Modelling (PRM). PRM basiert allerdings vornehmlich auf rein statistischen Methoden, wird im entsprechenden Beitrag allerdings als Beispiel für KI diskutiert. Dies mag ein Hinweis darauf sein, dass die Abgrenzung zwischen statistischen Methoden und KI-Systemen im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit weiterhin in Bewegung ist (Telkmann, 2023). Die Aussage, „Deep Learning-Verfahren haben daher den Vorteil, dass Trainingsdaten vorgängig nicht mehr manuell gekennzeichnet (gelabelt) werden müssen“ (Steiner & Tschopp, 2022, S. 467) verweist darüber hinaus auf eine Perspektive, welche die Vielfalt praktischer Umsetzungen nur teilweise widerspiegelt, da leistungsfähige Systeme oft weiterhin auf umfangreiche manuell annotierte Datensätze zurückgreifen. Dennoch deutet sich hier eine wachsende technische Präzision verbunden mit einem breiteren methodischen Horizont an, da der Aspekt des Trainings von KI-Systemen in den Vordergrund gerückt wird. Dies führen Linnemann et al. (2023) mit einer Definition von KI als Oberbegriff für Verfahren fort, „die aus Daten neue Informationen generieren, insbesondere in Form von Vorhersagen, Klassifikationen oder Ordnungen“. In diesem Beitrag wird explizit auf die statistischen Grundlagen des maschinellen Lernens verwiesen: „Maschinelles Lernen umfasst Verfahren, bei denen Systeme lernen, was bedeutet, dass sich ihre auf statistischen Modellen beruhenden Vorhersagen durch Rückmeldungen in Form von Feedbackdaten verbessern“ (Linnemann et al., 2023, S. 199). Fortgeschrittenes Natural Language Processing wird in diesem Beitrag ebenso einbezogen, wobei der Fokus in der Darstellung stellenweise auf spezifische Anwendungsformen wie Chatbots verengt bleibt. Besonders hervor-

zuheben ist der erstmalige Verweis auf das Forschungsfeld der *Explainable Artificial Intelligence* (XAI), „das der Fragestellung gewidmet [ist], wie diese Informationen transparent gemacht werden können“ (Linnemann et al., 2023, S. 200). Damit tritt neben die technische und statistische Fundierung auch eine methodische Perspektive, die für den zukünftigen verantwortungsvollen Einsatz von KI in der Sozialen Arbeit wegweisend sein dürfte. Entsprechend kongruent schließt sich die Forderung nach einer „Kompetenz im Bereich von Informationsanalyse und Datenerhebung“ (Remke & Schiffhauer, 2025, S. 82) im Kontext der Sozialen Arbeit inklusive einer entsprechenden curricularen Anpassung (Erdwiens & Seidel, 2022) an. Kritisch zu sehen sind die stellenweise explizierten und methodisch nicht immer zutreffenden Verweise auf die Korrelation (Telkmann, 2023; Görder, 2021) als zugrundeliegende Funktionsweise von KI-Systemen, die nicht umfassende Berücksichtigung der statistischen Grundlagen (Steiner & Tschopp, 2022; Schneider & Seelmeyer, 2018) sowie die den Algorithmen unterstellte Eigenständigkeit (Wunder, 2021). Dies ist besonders deshalb fragwürdig, da die statistischen Grundlagen auch unabhängig der Einführung der Transformer-Architektur (Vaswani et al., 2017) nicht nur seit einem Jahrzehnt mit entscheidenden *Deep-Learning*-Grundlagen (LeCun et al., 2015) gleichgeblieben sind, sondern seit längerem zu den klassischen Datenanalyseverfahren der Statistik zählen.

Zusammenfassend lässt sich sagen, dass die Soziale Arbeit als Profession bereits frühzeitig auf die Notwendigkeit eines Verständnisses der statistischen Grundlagen von KI-Systemen hingewiesen hat (Linnemann et al., 2023; Telkmann, 2023; Schneider & Seelmeyer, 2018). Grund hierfür war die Reflexion sowohl theoretischer (Görder 2021; Kutscher et al. 2020) als auch methodischer (Telkmann, 2023) Bedenken im Umgang mit diesen KI-Systemen, um so bspw. eine algorithmenbasierte Diskriminierung zu verhindern (Remke & Schiffhauer, 2025). Insbesondere der Verweis auf XAI verdeutlicht die Bedeutung erklärbarer KI-Systeme im Kontext der Sozialen Arbeit und eröffnet eine neue Perspektive (Linnemann et al., 2023), die ohne diese statistischen Grundlagen nicht möglich wäre. Der vorliegende Beitrag widmet sich nachfolgend nicht nur detailliert den statistischen Grundlagen von KI-Systemen mit dem Ziel einer erklärbaren KI für die Soziale Arbeit, sondern schließt sich auch bereits explizierten Handlungsempfehlungen an (Remke & Schiffbauer, 2025; Linnemann et al., 2023; Erdwiens & Seidel, 2022) mit einem entsprechenden Anwendungsbeispiel.

3. Methodik

3.1 Forschungs- und Weiterbildungsderivat

Ausgehend von der Einführung der Transformer-Architektur (Vaswani et al., 2017) wurden zum Zweck eines besseren Verständnisses der zugrundeliegenden Funktionsweise einschlägige Weiterbildungsangebote recherchiert, absolviert und systematisch kontrastiert. Dies geschah vor dem Hintergrund, die statistischen Grundlagen von KI zu identifizieren und im Rahmen einer kumulativen Habilitation an der Justus-Liebig-Universität Gießen zur *Implementierung von Methoden der computergestützten Datenanalyse in sozialwissenschaftlicher Lehre und Forschung* zusammenzutragen, um die Disziplinen *Social Science* und *Data Science* als *Social Data Science* bzw. *Computational Social Science* zusammenzuführen – eine Zusammenführung, wie sie sich bspw. bei GESIS als Leibniz-Institut für Sozialwissenschaften etabliert hat. Entsprechend handelt es sich bei dem vorliegenden Beitrag um ein Forschungs- und Weiterbildungsderivat.

3.2 Übersicht über einschlägige Weiterbildungsangebote

Die Weiterbildungsangebote lassen sich zunächst in die Mathematik hinter den statistischen Grundlagen, den statistischen Grundlagen des maschinellen Lernens und *Deep Learnings* selbst sowie weiteren Vertiefungen zum *Natural Language Processing* unterteilen. Nachfolgend ausgewiesen sind jeweils die absolvierten Weiterbildungsstunden und die dazugehörigen Jahreszahlen. Die Mathematik vertieft dabei die Funktionsweise der statistischen Grundlagen im Rahmen der Anwendung im maschinellen Lernen und beim *Deep Learning*. Zu den absolvierten Weiterbildungsangeboten zählten daher zunächst die allgemeine Spezialisierung *Data Science* der Johns Hopkins University, basierend auf zehn sukzessiven Selbstlernkursen (ca. 280 Stunden zwischen 2017 und 2018) und die Spezialisierung *Mathematics for Machine Learning* des Imperial College London, welche auf drei sukzessiven Selbstlernkursen (ca. 40 Stunden in 2018) basiert. *Machine Learning* von Stanford Online wurde als Selbstlernkurs (ca. 20 Stunden in 2019) durch *Data Science Math Skills* der Duke University als Selbstlernkurs (ca. 10 Stunden in 2020) ergänzt, um auf die Spezialisierung *Deep Learning* von DeepLearning.AI vorzubereiten, die auf fünf sukzessiven Selbstlernkursen (ca. 120 Stunden zwischen 2020 und 2021) basiert. Die Spezialisierung *Natural Language Processing* von DeepLearning.AI, basierend auf fünf sukzessiven Selbstlernkursen (ca. 120 Stunden zwischen 2022 und 2023) ist eine Konsequenz der Markteinführung von ChatGPT gewesen – das u. a. auf der Transformer-Architektur basiert – und wurde ergänzt um die Selbstlernkurse *Generative AI with Large Language Models* von DeepLearning.AI (ca. 20 Stunden in 2024) und schließlich um *Probabilistic Deep Learning with TensorFlow 2* vom Imperial College London (ca. 50 Stunden in 2025).

Die insgesamt 690 im Rahmen der kumulativen Habilitation absolvierten und zusammengetragenen Weiterbildungsstunden greifen dabei jeweils auf die Programmiersprachen R oder Python zurück. In allen Weiterbildungsangeboten werden die statistischen Grundlagen von KI entweder ausdifferenziert dargestellt, oder aber entsprechend vertieft, worauf sich die nachfolgenden Ausführungen des vorliegenden Beitrags beziehen werden. Einschlägige Fachpublikationen rahmen das Forschungs- und Weiterbildungsderivat, wobei Autor:innen wie LeCun et al. (2015) und Ng et al. (2001) aufgrund ihrer zentralen Rolle in der Entwicklung von KI-Systemen, deren statistischen Grundlagen oder einschlägigen Weiterbildungsangeboten keine weitere Vorstellung erfordern.

4. Auswertung

4.1 Grundlagen der Künstlichen Intelligenz

In einschlägigen Definitionen wird KI vornehmlich als ein symbolischer Überbegriff verwendet (Trabs et al., 2021), der insbesondere in die Grundlagen des maschinellen Lernens im Allgemeinen und *Deep Learning* als Teilbereich des maschinellen Lernens im Besonderen ausdifferenziert werden kann (LeCun et al., 2015). Dadurch ist zunächst das maschinelle Lernen als wesentlicher Teilbereich der Künstlichen Intelligenz zu definieren, das darauf abzielt, Computern die Fähigkeit zu verleihen, Muster in Daten zu erkennen (Wennker, 2020). Das maschinelle Lernen lässt sich grundsätzlich in überwachtes und unüberwachtes Lernen unterteilen (Rashidi et al., 2023). Beim überwachten Lernen werden Modelle mit zuvor von Menschen gelabelten Daten trainiert, um Vorhersagen oder Zuordnungen vorzunehmen. Hierzu zählen u. a. die aus der Statistik bekannten Regressionsverfahren (Trabs et al., 2021), die kontinuierliche Werte vorhersagen, sowie Klassifikationsverfahren, die Daten in vorgegebene Kategorien einteilen (Friedrich et al., 2022). Beim unüberwachten Lernen hingegen arbeitet das Modell mit zunächst nicht gelabelten Daten und versucht inhärente Strukturen zu identifizieren. In diesem Bereich finden bspw. Clusteringverfahren Anwendung (Ng et al., 2001), die Datenpunkte auf Basis ihrer Ähnlichkeit gruppieren, sowie Verfahren zur Dimensionsreduktion, die hochdimensionale Daten auf eine kompaktere Darstellung abbilden (Rashidi et al., 2023; Trabs et al., 2021). Sowohl Clusteringverfahren als auch Verfahren zur Dimensionsreduktion sind ebenfalls vornehmlich statistische Verfahren (Friedrich et al., 2022). Für alle vier explizierten Verfahren existiert entsprechend eine Vielzahl an Algorithmen (Klinkhammer & Keller, 2024), die je nach Problemstellung und Beschaffenheit der Daten eingesetzt werden können. Einige dieser Algorithmen, darunter die Sigmoid-Funktion der logistischen Regression, können darüber hinaus als sogenannte Aktivierungsfunktionen in neuronalen Netzwerken verwendet werden (siehe Abbildung 1). Ein Algorithmus ist dabei eine endliche, wohl definierte Folge von Anweisungen zur Lösung eines Problems oder zur Durchführung einer Berechnung. Im

Bereich der KI und insbesondere im maschinellen Lernen bestimmen Algorithmen die Art und Weise, wie Modelle aus Daten lernen, Muster erkennen und Vorhersagen treffen können (Rashidi et al., 2023).

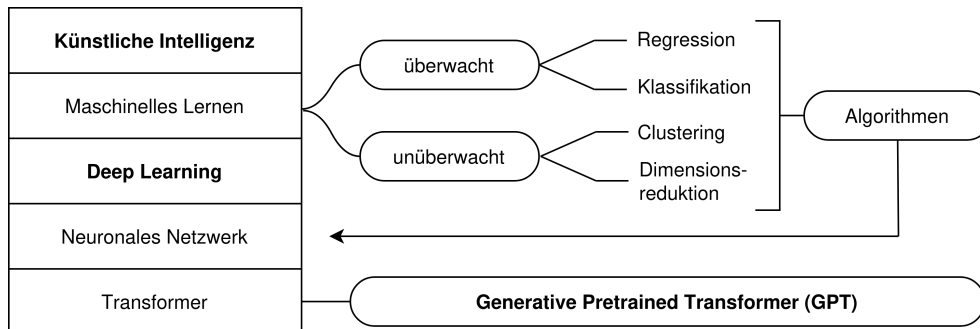


Abbildung 1: Statistische Grundlagen der Künstlichen Intelligenz, eigene Darstellung.

Neuronale Netzwerke stellen demnach eine fundamentale Klasse von Modellen innerhalb des *Deep Learnings* dar, einem Teilbereich des maschinellen Lernens, der auf einer mehrschichtigen Architektur basiert (Rashidi et al., 2023; LeCun et al., 2015). Sie bestehen aus mehreren Schichten mit einer vorab definierten Anzahl an Neuronen, die mittels Aktivierungsfunktionen nichtlineare Zusammenhänge zwischen Eingangs- und Ausgangsdaten modellieren sollen. *Deep Learning* hat insbesondere in den letzten Jahren erhebliche Fortschritte in verschiedenen Anwendungsgebieten, darunter Bild- und Sprachverarbeitung, erzielt (Rashidi et al., 2023). Eine spezialisierte Architektur innerhalb des *Deep Learning* sind die sogenannten Transformer, die in neuronalen Netzwerken zur sequenziellen Datenverarbeitung eingesetzt werden (Vaswani et al., 2017). Die Transformer-Architektur basiert dabei auf der Selbstaufmerksamkeit, auch Self-Attention genannt, sowie Mechanismen zur Gewichtung von Sequenzbestandteilen, wodurch sie besonders effizient in der Verarbeitung von Sprach- und Textdaten sind (Vaswani et al., 2017). Ein bekanntes Beispiel für ein auf der Transformer-Architektur basierendes KI-System ist der *Generative Pretrained Transformer* (GPT), der durch vortrainierte Wissensrepräsentationen und Feinabstimmung auf spezifische Aufgaben leistungsfähige Ergebnisse in der natürlichen Sprachverarbeitung (*Natural Language Processing*) erzielen kann. Die Transformer-Architektur hat sich dadurch als wegweisend für zahlreiche moderne KI-Systeme, wie bspw. ChatGPT, erwiesen und ist ein Schlüsselkonzept im aktuellen Stand der Forschung zur KI.

4.2 Grundlagen der Erklärbaren Künstlichen Intelligenz

Erklärbare KI, auch *Explainable Artificial Intelligence* (XAI) genannt (Haque et al., 2023; Meske et al., 2020), befasst sich vor diesem Hintergrund mit der Transparenz und Nachvollziehbarkeit von KI-Systemen, insbesondere in komplexen maschinellen Lernverfahren wie dem *Deep Learning* (Rashidi et al., 2023). Während die leistungsstarken KI-Sys-

teme im wissenschaftlichen Diskurs manchmal auch als Black Box bezeichnet werden, da ihre internen Entscheidungsprozesse für Menschen schwer verständlich sein können, zielt XAI darauf ab, deren Funktionsweise für Nutzer:innen transparent darzulegen und interpretierbar zu machen (Haque et al., 2023). Einblicke in die zugrundeliegenden Algorithmen, deren Umsetzung in einer verständlichen Programmiersprache sowie die transparente Ausweisung (und ggfls. manuelle Veränderbarkeit) von relevanten Parametern wie Weights und Biases, die maßgeblich für die Informationsverarbeitung und -weitergabe in neuronalen Netzwerken verantwortlich sind, gelten als Möglichkeiten zur Umsetzung von XAI-Systemen (Meske et al., 2020).

Im Bereich des überwachten Lernens ermöglichen XAI-Systeme die Erklärung von Klassifikations- und Regressionsentscheidungen, bspw. durch Techniken wie *Feature-Importance-Analysen* oder *Local Interpretable Model-Agnostic Explanations* (LIME) (Saarela & Jauhiainen, 2021). Solche Ansätze helfen dabei, zu verstehen, welche Merkmale eines Datensatzes maßgeblich zu einer bestimmten Entscheidung beitragen (Haque et al., 2023). Im unüberwachten Lernen ist die Erklärbarkeit eine besondere Herausforderung, da Clusteringverfahren und Verfahren zur Dimensionsreduktion oft keine direkten Entscheidungsregeln liefern. Methoden wie die Visualisierung von Clusterstrukturen oder die Untersuchung latenter Merkmalsräume tragen allerdings dazu bei, Muster und Beziehungen in den Daten verständlicher zu machen. Neuronale Netzwerke, insbesondere tiefgehende Architekturen mit einer Vielzahl an Neuronen und Schichten, erfordern einen spezialisierten XAI-Ansatz, um ihre Vorhersagen nachvollziehbar zu gestalten (Haque et al., 2023). Gradientenbasierte Methoden, Aktivierungsvisualisierungen oder sogenannte *Attention Maps* für die Transformer-Architektur sind Beispiele für Ansätze, die zur Interpretierbarkeit beitragen können. Gerade bei Transformer-Architekturen wie GPT sind Mechanismen zur Analyse der Selbstaufmerksamkeit essenziell, um zu verstehen, welche Textbestandteile maßgeblich für eine Antwort sind.

Die Integration von XAI in KI-Systeme gewinnt vor diesem Hintergrund zunehmend an Bedeutung (Haque et al., 2023), insbesondere in kritischen Anwendungen, in denen Entscheidungen überprüfbar und vertrauenswürdig sein müssen (Meske et al., 2020). Durch erklärbare KI-Systeme können nicht nur bessere Einsichten in Modellentscheidungen gewonnen, sondern auch ethische und regulatorische Anforderungen adressiert werden.

4.3 *Passung zum professionsbezogenen KI-Verständnis in der Sozialen Arbeit*

Während im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit oftmals auf die sogenannten Klassifikationsverfahren verwiesen wird, konnten darüber hinaus Regressionsverfahren, Clusteringverfahren sowie Verfahren zur Dimensionsreduktion als statistische Grundlagen für KI-Systeme identifiziert werden. Alle vier Verfahren greifen dabei auf verschiedene Algorithmen im Kontext des maschinellen Lernens zurück. Exemplarisch sei an den Klassifikationsverfahren nochmals hervorgehoben, dass eine Vielzahl an Algorithmen zur Auswahl steht, um Muster in den Daten zu erkennen; Gleiches gilt für die anderen erwähnten Verfahren.

Während sich im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit stellenweise noch ein synonyme Gebrauch von Algorithmen, KI und KI-Systemen abzeichnet, geht aus den einschlägigen Weiterbildungsangeboten hervor, dass Algorithmen als eine systematische, sequenzielle Anweisung zur Lösung eines spezifischen Problems definiert werden können. KI umfasst dabei ein breiteres Spektrum an Technologien, die in der Lage sind, aus Daten zu lernen sowie Muster und Zusammenhänge zu erkennen. Algorithmen dienen im Kontext von KI als Grundlage, um KI-Systeme zu trainieren, etwa in Form von Optimierungsverfahren oder Lernmethoden. Als solche sind sie jedoch nicht ausreichend, um als KI bezeichnet zu werden oder eine solche zu definieren.

Ethische Fragen, wie Datenschutz und Exklusion werden bereits differenziert im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit besprochen, inklusive einschlägigen Verweisen auf das mögliche Black-Box-Problem und auf das mögliche Bias-Problem von KI-Systemen. Ein vertiefender Blick auf die oftmals zugrundeliegenden Verzerrungen in den Trainingsdaten und deren Konsequenzen für das maschinelle Lernen im Allgemeinen sowie *Deep Learning* im Besonderen wäre allerdings notwendig, um mögliche Schwachstellen hinsichtlich der Interpretierbarkeit einschätzen, erkennen und diesen – bspw. über die Parameter neuronaler Netzwerke – entgegenwirken zu können. Eine Kenntnis der statistischen Grundlagen erweist sich hier als Voraussetzung, um einem Bias in den Trainingsdaten, den Algorithmen oder den zugrundeliegenden Annahmen, die beim Modellaufbau getroffen werden, entgegenwirken zu können. Den einschlägigen Weiterbildungsangeboten und dem explizierten Anwendungsbeispiel können entsprechende Lösungsvorschläge entnommen werden und es wird deutlich, dass ein solcher Bias in neuronalen Netzwerken die generalisierende Fähigkeit der Modelle zwar beeinflussen kann, diese deshalb aber auch eine versierte Supervision menschlicher Akteur:innen erfordern.

Insbesondere um eigene und zielgruppenspezifische KI-Systeme zu entwickeln, die bspw. durch die Expertise von Sozialarbeiter:innen auf die Bedarfe vulnerabler Gruppen abgestimmt sind, bedarf es eines versierten Rückgriffs auf die statistischen Grundlagen. Eine systematische Unterscheidung zwischen statistischen Verfahren, bzw. dessen Einordnung im Kontext des maschinellen Lernens, kann daher als zielführend für den wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit erachtet werden.

5. Erklärbare KI: Anwendungsbeispiel

Das nachfolgende Anwendungsbeispiel stellt eine Einführung in ein diskriminierungsfreies *Natural Language Processing* dar, welches auf Grundlage der Spezialisierung von DeepLearning.AI eigenständig erstellt werden kann. Es wird entsprechend ergänzt um die zugrundeliegende Mathematik aus der Spezialisierung *Mathematics for Machine Learning* des Imperial College London sowie aus dem Selbstlernkurs *Data Science Math Skills* der Duke University. In dem Anwendungsbeispiel sind zwei unterschiedliche Personen A und B in einem dreidimensionalen Koordinatensystem hinterlegt. Das dreidimensionale Koordinatensystem basiert entsprechend auf einer X-Achse, einer Y-Achse und einer Z-Achse (siehe Abbildung 2), woraus sich die Koordinaten der Person A mit (X_A, Y_A, Z_A) sowie der Person B mit (X_B, Y_B, Z_B) ableiten lassen. Da sämtliche Wörter in großen Sprachmodellen wie bspw. ChatGPT als Koordinaten hinterlegt und somit über Vektoren und Matrizen berechnet werden können, werden Vektoren und Matrizen in den Weiterbildungsangeboten zur zugrundeliegenden Mathematik eingehend behandelt. Zusätzlich greift das *Natural Language Processing* großer Sprachmodelle – wie bspw. ChatGPT – auf neuronale Netzwerke und somit auf *Deep Learning* zurück, welches in der Spezialisierung von DeepLearning.AI vermittelt wird. In einem neuronalen Netzwerk können die Koordinaten der Personen A und B als sogenannte Features X, Y und Z mit vorab hinterlegten Values 0 oder 1 verarbeitet und somit das neuronale Netzwerk trainiert werden. Die Values können dabei für tatsächliche oder zugeschriebene Eigenschaften der Personen A und B stehen, bspw. negativ (0) oder positiv (1). Sowohl die Features als auch die Values beeinflussen das Lernverhalten eines neuronalen Netzwerks. Die erste Möglichkeit einer Diskriminierungsvermeidung ist dabei die manuelle Überschreibung der Features oder Values, welche bspw. in einer Datentabelle hinterlegt sein können. Somit kann Person A in dem dreidimensionalen Koordinatensystem näher an Person B heranrücken oder Person A eine positive Eigenschaft (1) zugeschrieben werden.

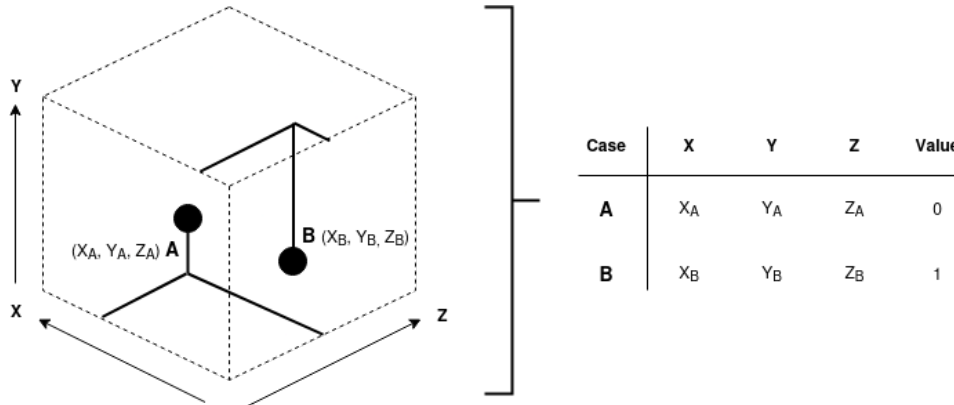


Abbildung 2: *Natural Language Processing* ohne Diskriminierung, eigene Darstellung.

Die zweite Möglichkeit einer Diskriminierungsvermeidung setzt an den Parametern neuronaler Netzwerke an. Wie in den Selbstlernkursen *Generative AI with Large Language Models* von DeepLearning.AI und *Probabilistic Deep Learning with TensorFlow 2* des Imperial College London verdeutlicht, lassen sich die Weights und Biases als Parameter bereits trainierter neuronaler Netzwerke nachträglich manuell anpassen. Neuronale Netzwerke basieren zunächst auf Neuronen, die in Schichten angeordnet sind. Die Neuronen führen Berechnungen auf der Grundlage sogenannter Aktivierungsfunktionen durch, wie bspw. der logistischen Regression für binäre Klassifikationen (hier: Value = 0 bzw. Value = 1), wie in der Spezialisierung *Data Science* der Johns Hopkins University oder dem Selbstlernkurs *Machine Learning* von Stanford Online verdeutlicht. Erreicht die Aktivierungsfunktion einen bestimmten Schwellenwert, gibt diese die entsprechenden Values an die Neuronen der nachfolgenden Schicht weiter. Entsprechend werden die Weights und Biases beim Training zunächst so angepasst, dass die Vorhersagegenauigkeit des neuronalen Netzwerks optimiert wird. Die Weights würden in dem vorliegenden Beispiel die jeweilige Relevanz der Features X, Y und Z bestimmen, mit der diese auf die Aktivierungsfunktion einwirken. Die Biases hingegen wirken auf den Schwellenwert der Aktivierungsfunktion ein, ohne die Aktivierungsfunktion selbst zu verändern. Dies kann mit der Höhe einer Hürde bei einem Hindernislauf verglichen werden. Sowohl die Weights als auch die Biases können nach dem Training manuell angepasst werden, um nachträglich das Ergebnis der binären Klassifikation zu beeinflussen. Vereinfacht ausgedrückt ermöglicht dies die Korrektur einer fälschlicherweise von dem neuronalen Netzwerk mit einer negativen Eigenschaft versehenen Person hin zu einer positiven Eigenschaft, sodass das angepasste neuronale Netzwerk die erkannte Form der Diskriminierung bei zukünftigen Vorhersagen unter ähnlichen Voraussetzungen weniger wahrscheinlich als Ergebnis hervorbringen wird. Durch den Blick auf die Parameter ist das neuronale Netzwerk nicht nur transparent, sondern auch erklärbar und anpassbar für entsprechend versierte Nutzer:innen.

Um dieses und weitere Anwendungsbeispiele aus dem Kontext der Sozialen Arbeit auch ohne die Kenntnis einer Programmiersprache bspw. im Rahmen der Hochschulausbildung auszuprobieren, wurde das *Open-Source-Framework* KI-ENNA *academic* entwickelt (Klinkhammer, 2025), für das u. a. als Anwendungsbeispiel die Beurteilung einer Kindeswohlgefährdung im Kontext der Sozialen Arbeit direkt in den Demo-Daten zur Verfügung steht. Damit lernen Studierende nicht nur autodidaktisch selbst ein neuronales Netzwerk zu trainieren, sondern am Anwendungsbeispiel der Kindeswohlgefährdung auch dessen Lernverhalten transparent einzusehen und die Fälle in der zugrundeliegenden Datentabelle zu identifizieren, die für ein solches KI-System herausfordernd sein können, um die Parameter manuell zu überarbeiten. Das *Open-Source-Framework* gilt deshalb als besonders niedrigschwellig, da es sowohl das Training als auch die Ausweisung der Weights und Biases direkt im Browser und somit ohne Hard- und Softwarevoraussetzungen ermöglicht. Neben der kostenlosen Browser-App (<https://www.statistical-thinking.de/ki-enna-academic.html>), welche kompatibel zu vollständigen KI-Frameworks wie TensorFlow in der Programmiersprache Python ist, stehen bei Bedarf entsprechende Tutorials entsprechende Tutorials (<https://statistical-thinking.de/ki-tutorial-soziale-arbeit.html>) zur Verfügung. Das Codebuch zur Beschreibung der Demo-Daten sowie weitere Anwendungsbeispiele stehen auf der Homepage der Browser-App zur Verfügung.

6. Zusammenfassung und Ausblick

6.1 Erklärbare KI als Handlungsempfehlung

Ausgehend von den zuvor dargestellten Ausführungen zu den statistischen Grundlagen der KI scheinen curriculare Anpassungen in der akademischen Ausbildung sowie einschlägige Weiterbildungsangebote erforderlich, um Sozialarbeiter:innen die Möglichkeit zu eröffnen, KI-Systeme und deren Funktionsweise fundiert einschätzen zu können. Hierfür bietet sich eine interdisziplinäre Zusammenarbeit mit Informatiker:innen und Expert:innen aus dem Bereich *Data Science* an, bspw. in Form von gemeinsamen Lehrveranstaltungen wie „KI und Soziale Arbeit“ von OPEN vhb – was für *virtuelle hochschule bayern* steht. Die Einführung eines konsekutiven Masterprogramms oder Zertifikatskurses zu KI-Systemen in der Sozialen Arbeit würde anstelle dieser allgemeinen Handlungsempfehlungen eine besonders versierte Gruppe von Expert:innen in der Profession der Sozialen Arbeit hervorbringen, die als Multiplikator:innen und an der Schnittstelle zu KI-Systemen zur Qualitätssicherung beitragen könnten. Dabei sind Vertiefungen in den statistischen Grundlagen genauso denkbar wie niedrigschwellige Anwendungsbeispiele zur Vermittlung der Funktionsweise von KI-Systemen. Insgesamt scheint es wichtig, die kritische Reflexion und Praxisorientierung zu stärken (Remke & Schiffbauer, 2025; Erdwiens & Seidel, 2022), indem Studierende und bereits ausgebildete Sozialarbeiter:innen in der Erkennung von Bias und

ethischen Herausforderungen geschult werden und dabei eigene, praxisnahe Fallstudien entwickeln können, um KI-Systeme kritisch zu hinterfragen und zielgruppenspezifisch auszugestalten. GitHub könnte eine entsprechende Plattform für einschlägige Vorhaben sein. Dabei ist davon auszugehen, dass für die genannten Regressions-, Klassifikations-, Clustering- und Dimensionsreduktionsverfahren bereits Expert:innen im Kontext der Sozialen Arbeit vertreten sind, da diese statistischen Grundlagen oftmals ein fester Bestandteil der akademischen Ausbildung professionsübergreifender Fachdisziplinen sind – wie bspw. der Erziehungswissenschaften, der Soziologie und der Psychologie. Diese Fachdisziplinen tragen nicht nur zur interdisziplinären Profession der Sozialen Arbeit bei, sondern sind gleichermaßen Bestandteil des professoralen Kollegiums und akademischen Mittelbaus. Unschärfen im Verständnis von KI sowie auch die entsprechende Vermarktung von KI-Systemen außerhalb des Fachdiskurses mit Buzzwords könnten von diesem professionsinternen Vorteil abgelenkt haben.

Wie im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit bereits expliziert, könnte darüber hinaus ein Schwerpunkt auf erklärbaren KI-Systemen liegen, die als XAI-Ansatz vorgestellt worden sind (Haque et al., 2023). Diese erhöhen die Transparenz von KI-Systemen und ermöglichen u. a. über die statistischen Grundlagen die versierte Auseinandersetzung mit diesen (Linnemann et al., 2023; Trabs et al., 2021). Bestrebungen auf dem Gebiet der XAI setzen sich zu diesem Zweck mit der Herausforderung auseinander, KI-Systeme und deren Entscheidungen für den Menschen verständlicher und nachvollziehbarer zu gestalten (Meske et al., 2020). Dabei geht es nicht nur darum, dass die Entscheidungen von KI-Systemen adäquat interpretiert werden können, sondern auch darum, dass die Entscheidungsprozesse und die zugrunde liegenden Modelle transparent und nachvollziehbar umgesetzt werden. Ein KI-System, das dem XAI-Ansatz folgt, könnte beispielsweise transparent aufzeigen, welche Daten und Faktoren in den Entscheidungsprozess eingeflossen sind und wie sie miteinander in Beziehung stehen (Haque et al., 2023). Ein solches Verständnis schafft nicht nur Vertrauen bei den professionellen Akteur:innen aus dem Kontext der Sozialen Arbeit (Klinger et al., 2022), welche KI-Systeme nutzen, sondern auch bei Klient:innen, die von diesen Entscheidungen betroffen sein können. Zudem könnten gezielt Forschungsprojekte und Open-Source-Initiativen gefördert werden, um zielgruppenspezifische KI-Systeme für die Soziale Arbeit zu entwickeln und Sozialarbeiter:innen die Möglichkeit zu geben, eigene KI-Systeme zu gestalten.

6.2 Anschlussfähigkeit und Zukunftsperspektive

Der Beitrag verdeutlicht, dass KI in der Sozialen Arbeit stellenweise als technologische Einheit betrachtet wird, ohne die zugrundeliegenden Algorithmen entsprechend in Regressions-, Klassifikations-, Clustering- und

Dimensionsreduktionsverfahren ausdifferenzieren. Dies kann zu einer unzureichenden Einschätzung der Chancen und Risiken führen, insbesondere hinsichtlich der ethischen Herausforderungen, die sich aus Bias und dem sogenannten Black-Box-Problem ergeben können. Diese Herausforderungen können insbesondere bei neuronalen Netzwerken und dem *Deep Learning* einen nicht unerheblichen Einfluss auf KI-Systeme nehmen, wobei die Ursachen vornehmlich über fundierte Statistikkenntnisse adressiert werden können. Aktuell zeichnet sich ab, dass bei KI-Systemen im wissenschaftlichen Diskurs zu KI in der Sozialen Arbeit vornehmlich auf ChatGPT verwiesen wird, ohne dabei alternative und zum Teil für den Kontext der Sozialen Arbeit besser geeignete KI-Systeme zu reflektieren oder die zugrundeliegende Funktionsweise, bspw. in Form der Transformer-Architektur, zu berücksichtigen.

Erklärbare KI stellt vor diesem Hintergrund eine zentrale Zukunftsperspektive für die Soziale Arbeit dar. Sie bietet Möglichkeiten, die Transparenz und Nachvollziehbarkeit von KI-Systemen zu verbessern und deren Einsatz verantwortungsbewusst zu gestalten. Mit gezielten Einblicken in die Algorithmen sowie deren Umsetzung in einer verständlichen Programmiersprache kann bspw. der Prozess der Informationsverarbeitung und -weitergabe in neuronalen Netzwerken veranschaulicht werden. Um jedoch von diesen Potenzialen zu profitieren, bedarf es einer intensiveren methodischen Schulung von professionellen Akteur:innen der Sozialen Arbeit, insbesondere hinsichtlich der statistischen Grundlagen und der Mechanismen maschinellen Lernens. Entsprechende Weiterbildungsangebote wurden in diesem Beitrag hervorgehoben und es empfiehlt sich eine Ausgestaltung ähnlicher Weiterbildungsangebote mit spezifischem Fokus auf die Soziale Arbeit als Profession. Curriculare Anpassungen in der Ausbildung sowie interdisziplinäre Kooperationen mit Informatiker:innen und *Data-Science*-Expert:innen können in diesem Sinne dazu beitragen, das Verständnis für KI-Systeme zu vertiefen und eigene, praxisnahe Anwendungen zu entwickeln.

Der zukünftige wissenschaftliche Diskurs zu KI in der Sozialen Arbeit könnte sich dadurch verstärkt der Entwicklung erklärbarer und ethisch reflektierter KI-Systeme widmen, die spezifisch auf die Bedarfe der Klient:innen sowie der Profession der Sozialen Arbeit zugeschnitten sind. Hierbei könnten auch Open-Source-Initiativen eine zentrale Rolle spielen, um partizipative Ansätze zu fördern und Sozialarbeiter:innen eine aktive Mitgestaltung an KI-Systemen zu ermöglichen. Durch eine solche strategische Ausrichtung könnte KI nicht nur als technologische Innovation, sondern auch als integraler Bestandteil einer professionellen und ethisch verantwortungsvollen Sozialen Arbeit etabliert werden.

Literatur

- Ärleböck, J. B., & Kawakami, T. (2023). *The relationships between statistics, statistical modelling and mathematical modelling. Advancing and consolidating mathematical modelling: International perspectives on the teaching and learning of mathematical modelling*. Springer. https://doi.org/10.1007/978-3-031-27115-1_18
- Burton, J. (2023). Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence. *Technology in Society*, 75. <https://doi.org/10.1016/j.techsoc.2023.102262>
- Erdwiens, D., Seidel, A. (2022). Zur Verankerung von Themen der Digitalisierung in Modulhandbüchern der Studiengänge Sozialer Arbeit. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung, (Occasional Papers)*, 22–42. <https://doi.org/10.21240/mpaed/00/2022.06.13.X>
- Friedrich, S., Antes, G., Behr, S., et al. (2022). Is there a role for statistics in artificial intelligence? *Advances in Data Analysis and Classification*, 16. <https://doi.org/10.1007/s11634-021-00455-6>
- Görder, B. (2021). Die Macht der Muster: Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch Künstliche Intelligenz. *EthikJournal*, 7(2).
- Haque, A. B., Islam, A. K. M. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change*, 186, 122120. <https://doi.org/10.1016/j.techfore.2022.122120>
- Klinger, S., Mayr, A., & Sackl-Sharif, S. (2022). Digitalisierung der Handlungspraxis in der Sozialen Arbeit. Kontrastierungen nach Angebotsstrukturen und Zielgruppen. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 22(2). <https://doi.org/10.17169/FQS-22.2.3851>
- Klinkhammer, D. (2025). AI-ANNE: (A) (N)eural (N)et for (E) xploration. *Journal of Open Source Software*, 10(110). <https://doi.org/10.21105/joss.08039>
- Klinkhammer, D., & Keller, K. (2024). *Grundlagenwissen Künstliche Intelligenz*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-44959-9_21
- Kutscher, N., Ley, Z., Seelmeyer, U., Siller, F., Tillmann, A. & Zorn, I. (Hrsg.) (2020). *Handbuch Soziale Arbeit und Digitalisierung*. Beltz Juventa.

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. Springer Science and Business Media LLC. <https://doi.org/10.1038/nature14539>
- Linnemann G. A., Löhe, J. & Rottkemper, B. (2023). Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit. Eine exemplarische arbeitsfeldübergreifende Betrachtung des Natural Language Processing (NLP). *Soziale Passagen*, 15. <https://doi.org/10.1007/s12592-023-00455-7>
- Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2020). *Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities*. Informa UK Limited. <https://doi.org/10.1080/10580530.2020.1849465>
- Ng, A., Jordan, M., & Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 2, 849–856. <https://ai.stanford.edu/~ang/papers/nips01-spectral.pdf>
- Rashidi, H. H., Albahra, S., Robertson, S., Tran, N. K., & Hu, B. (2023). Common statistical concepts in the supervised Machine Learning arena. *Frontiers in Oncology*, 13. <https://doi.org/10.3389/fonc.2023.1130229>
- Remke, S., Schiffbauer, B. (2025). Digitalisierung im Curriculum kompetenzorientiert umsetzen: Erste Impulse in Orientierung an Lehrerfahrungen und am Deutschen Qualifikationsrahmen Soziale Arbeit. (2025). *DiMe-Sozial | Soziale Arbeit, Medien Und Digitalisierung*, 1(1), 79–88. <https://doi.org/10.21240/dimesoz/2025/1/15>
- Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Appl. Sci.* 3, 272. <https://doi.org/10.1007/s42452-021-04148-9>
- Schneider, D., & Seelmeyer, U. (2018). *Der Einfluss der Algorithmen*. Springer Science and Business Media LLC. <https://doi.org/10.1007/s12054-018-0046-y>
- Steiner, O., & Tschopp, D. (2022). Künstliche Intelligenz in der Sozialen Arbeit. *Sozial Extra* 46, 466–471. <https://doi.org/10.1007/s12054-022-00546-4>
- Stummbaum, M., & Rusert, K. (2021). *Zukünfte Sozialer Arbeit – digital und wie bei Ikea. Szenarien Sozialer Arbeit in der Digitalisierung*. Verlag Julius Klinkhardt. <https://doi.org/10.35468/5909-13>
- Telkmann, D. (2023). *Reboot: social work. Inwieweit verändert sich die Funktionsbestimmung Sozialer Arbeit durch den Einsatz künstlicher Intelligenz?* Bremer Schriften zur Sozialen Arbeit, 3. Hochschule Bremen. <https://doi.org/10.26092/elib/2765>

- Trabs, M., Jirak, M., Krenz, K., & Reiß, M. (2021). *Statistik und maschinelles Lernen: Eine mathematische Einführung in klassische und moderne Methoden*. Springer Nature. <https://doi.org/10.1007/978-3-662-62938-3>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N. & Polosukhin, I. (2017). Attention Is All You Need (Version 7). arXiv. <https://doi.org/10.48550/ARXIV.1706.03762>
- Wennker, P. (2020). *Künstliche Intelligenz in der Praxis*. Springer Gabler. <https://doi.org/10.1007/978-3-658-30480-5>
- Wunder, M. (2021). *Digitalisierung und Soziale Arbeit. Transformationen und Herausforderungen*. Julius Klinkhardt. <https://doi.org/10.25656/01:23157>

Tina Wenz¹

Künstliche Intelligenz und digitale Teilhabe

Potenziale, Spannungsfelder und Gestaltungsanforderungen für Menschen mit Beeinträchtigungen in der Sozialen Arbeit

¹ Landesverein für Innere Mission

Zusammenfassung

Künstliche Intelligenz (KI) transformiert die Soziale Arbeit und bietet neue Chancen für die Teilhabe von Menschen mit Beeinträchtigungen. Zugleich drohen neue Risiken: algorithmische Diskriminierung, Black-Box-Systeme, technikgetriebene Normierung und der Ausschluss nicht digital kompetenter Nutzer:innen. Der Beitrag analysiert Potenziale und Herausforderungen entlang von zehn Thesen zu ethischen, technischen und organisationalen Anforderungen. Er zeigt: Teilhabe erfordert partizipative Technikgestaltung, diskriminierungssensible Datenpraxis, Barrierefreiheit, adaptive Systeme und Governance. Die Finanzierung inklusiver KI stellt gerade kleinere Träger vor große Herausforderungen. Der Beitrag schließt mit konkreten Empfehlungen für Praxis, Politik und Forschung: ethische Standards operationalisieren, Ressourcen sichern, Peer-Learning fördern, analoge Alternativen erhalten und partizipative Prozesse institutionalisieren.

Schlüsselwörter: Digitale Teilhabe, Künstliche Intelligenz, Behinderung, Soziale Arbeit, Ethik, Barrierefreiheit, Governance, Digitalisierung

Artificial Intelligence and Digital Inclusion. Potentials, Tensions, and Design Imperatives for People with Disabilities in Social Work

Abstract

Artificial Intelligence (AI) is transforming social work and opening new opportunities for the inclusion of people with disabilities. At the same time, it poses new risks: algorithmic discrimination, black-box systems, technologically driven standardization, and the exclusion of digitally under-skilled users.

Wenz, T. (2026). Künstliche Intelligenz und digitale Teilhabe. Potenziale, Spannungsfelder und Gestaltungsanforderungen für Menschen mit Beeinträchtigungen in der Sozialen Arbeit. *DiMe-Sozial*, 2(1), 75–89. <https://doi.org/10.21240/dimesoz/2026/1/63>

This article analyzes the potentials and challenges of AI along ten theses addressing ethical, technical, and organizational requirements. It argues that meaningful inclusion requires participatory technology development, discrimination-sensitive data practices, accessibility, adaptive systems, and robust governance. The financing of inclusive AI solutions poses significant challenges, especially for smaller organizations. The article concludes with concrete recommendations for practice, policy, and research: operationalize ethical standards, secure funding, promote peer learning, maintain analog alternatives, and institutionalize participatory processes.

Keywords: Digital Inclusion, Artificial Intelligence, Disability, Social Work, Ethics, Accessibility, Governance, Digitalization

1. Einleitung

Die Digitalisierung und insbesondere Künstliche Intelligenz (KI) transformieren die Soziale Arbeit und verändern Prozesse der Unterstützung, Kommunikation und Entscheidungsfindung. Für Menschen mit Beeinträchtigungen entstehen dadurch neue Chancen zur Teilhabe – etwa durch assistive Technologien, barrierefreie Kommunikation und datenbasierte Unterstützungsangebote. Gleichzeitig zeigen Technik- und Disability Studies, dass algorithmische Diskriminierung, Intransparenz und Digital Divide neue Formen sozialer Ungleichheit erzeugen können (Benjamin, 2019; Eubanks, 2018; OECD, 2024).

Im Zentrum des Beitrags steht die Frage, unter welchen Bedingungen KI-Systeme Selbstbestimmung fördern und wie Ausschlussmechanismen vermieden werden können (Lee & Zhang, 2023; Mittelstadt et al., 2016). Der Beitrag diskutiert diese Fragestellung anhand von zehn Thesen, die ethische, technische und organisationale Dimensionen beleuchten. KI wird dabei nicht als rein technologische Innovation verstanden, sondern als sozial wirksame Technologie, deren Gestaltung entscheidend für die Qualität von Teilhabe ist.

2. Analytischer und begrifflicher Rahmen: Von Empowerment zu Inklusion

Zentrale Konzepte wie Teilhabe, Empowerment und digitale Souveränität bilden den normativen Rahmen dieses Beitrags. Teilhabe meint die aktive, selbstbestimmte Mitwirkung an gesellschaftlichen Prozessen; Empowerment zielt auf die Befähigung zur selbstbestimmten Nutzung digitaler Angebote; digitale Souveränität beschreibt den reflektierten, sicheren und kritischen Umgang mit Technologien (Ladau, 2021; Bennett & Rosner, 2019). Diese Begriffe stehen im Spannungsfeld von Equality (Gleichbehandlung) und Equity (Chancengerechtigkeit): Gleichbehandlung bie-

tet gleiche Mittel, Chancengerechtigkeit hingegen adressiert strukturelle Barrieren und unterschiedliche Bedarfe (Benjamin, 2019; Wachter et al., 2021).

Techniksoziologische und Disability-Studies-Ansätze betonen, dass Technik nie neutral ist, sondern Machtverhältnisse widerspiegelt und Diskriminierungen fortschreiben kann, wenn sie nicht partizipativ und adaptiv gestaltet wird (Benjamin, 2019; Eubanks, 2018). Der analytische Rahmen verknüpft diese Perspektive mit Konzepten der intersektionalen Inklusionsforschung, die betont, dass digitale Exklusion durch das Zusammenspiel mehrerer sozialer Kategorien (z. B. Behinderung, Gender, Herkunft, Klasse) verstärkt werden kann.

Empowerment und Inklusion werden dabei nicht als additive Begriffe verstanden, sondern als aufeinander bezogene Prinzipien: Empowerment als Voraussetzung individueller Handlungsfähigkeit, Inklusion als strukturelle Ermöglichung sozialer Gleichberechtigung. Die theoretische Rahmung dient zugleich als Grundlage für die empirisch gestützten Thesen, die im Folgenden zentrale Anforderungen und Gestaltungsprinzipien KI-gestützter Teilhabe zusammenfassen.

3. Zehn Thesen zu KI und digitaler Teilhabe: Chancen, Risiken, Grenzen und Anforderungen

Die Potenziale und Risiken von Künstlicher Intelligenz (KI) für digitale Teilhabe in der Sozialen Arbeit lassen sich systematisch entlang von zehn Thesen analysieren. Diese strukturieren zentrale ethische, technische und organisationale Anforderungen und berücksichtigen unterschiedliche Nutzer:innengruppen sowie Kontexte wie Alltag, Bildung und Arbeit. Grundlage bilden empirische Befunde, kritische Theorien und Good-Practice-Beispiele (Benjamin, 2019; Eubanks, 2018; OECD, 2024; Bennett & Rosner, 2019).

These 1: Partizipation ist Voraussetzung für gelingende digitale Teilhabe

Die konsequente Einbindung von Menschen mit Beeinträchtigungen in Entwicklung, Auswahl und Evaluation von KI-Systemen ist essenziell, um Teilhabe im digitalen Raum tatsächlich zu verwirklichen. Partizipative Technikentwicklung ermöglicht es, die Bedarfe unterschiedlicher Gruppen frühzeitig zu berücksichtigen und realitätsnahe Lösungen zu schaffen. Studien zeigen jedoch, dass Beteiligungsprozesse häufig symbolisch bleiben und bestehende Machtasymmetrien reproduzieren, wenn sie nicht strukturell abgesichert sind (Bennett & Rosner, 2019; Shinohara & Wobbrock, 2016).

Fehlt diese Verankerung, werden KI-Systeme häufig von Personen ohne eigene Erfahrung mit Beeinträchtigungen konzipiert – ein Umstand, der dazu führt, dass Anwendungen an tatsächlichen Bedürfnissen vorbeigehen oder normierend wirken. Ethisch problematisch ist Partizipation ohne Entscheidungsmacht; sie verletzt das Prinzip der Achtung der Menschenwürde und der Selbstbestimmung.

Good Practice umfasst institutionell verankertes Co-Design, Nutzer:innenbeiräte, kontinuierliche Feedbackschleifen und barrierefreie Beteiligungsformate. Ethisch geboten sind transparente Mandate, reale Mitentscheidung, Ressourcen für Beteiligte und Rechenschaft über die Umsetzung von Empfehlungen.

These 2: Chancengerechtigkeit erfordert diskriminierungssensible Daten und Algorithmen

Chancengerechtigkeit in KI-gestützten Verfahren setzt eine diskriminierungssensible Daten- und Modellpraxis voraus. Werden marginalisierte Gruppen – insbesondere Menschen mit atypischen, mehrfachen oder seltenen Beeinträchtigungen – in Trainings- und Evaluationsdaten unterrepräsentiert, entstehen systematische Verzerrungen, die bestehende Machtverhältnisse fortschreiben (Benjamin, 2019; Lee & Zhang, 2023). Intersektionale Überschneidungen mit Armut, Herkunft oder Gender verstärken diese Effekte.

Ohne entsprechende Sensibilität werden Benachteiligungen automatisiert: Standardisierte Bewertungs- und Leistungsentscheidungen orientieren sich an Durchschnittsprofilen und verfehlen individuelle Lebenslagen. Ethisch steht hier das Gebot der Nicht-Diskriminierung im Zentrum; intransparente Datenherkünfte und unklare Fairnesskriterien verletzen dieses Prinzip.

Bewährte Praxis sind repräsentative Datenerhebung, partizipative Datengovernance, dokumentierte Feature-Auswahl, systematische Bias-Analysen über den gesamten Entwicklungszyklus sowie externe Audits (Jobin et al., 2019). Ergänzend sind klare Fairnessziele, Datenminimierung, Monitoring im Betrieb und Widerspruchsrechte zu verankern.

These 3: Empowerment setzt Nachvollziehbarkeit und Kontrolle voraus

Digitale Souveränität entsteht nur, wenn Entscheidungen von KI-Systemen nachvollziehbar sind. Erklärbarkeit ermöglicht es Nutzer:innen und Fachkräften, Funktionslogiken zu verstehen, Risiken einzuschätzen und Entscheidungen zu hinterfragen. Empirische Analysen wie bei Geiger & Halfaker (2017) verdeutlichen zudem, dass Kooperation zwischen automatisierten und menschlichen Entscheidungen nur funktioniert, wenn

Verantwortlichkeiten und Transparenzmechanismen klar geregelt sind. Gerade für Menschen mit kognitiven Einschränkungen ist Transparenz zentral, um Kontrolle über die eigene Nutzung zu behalten (Mittelstadt et al., 2016).

Fehlt Nachvollziehbarkeit, werden fehlerhafte oder voreingenommene Entscheidungen schwer erkennbar; Black-Box-Modelle unterminieren Selbstbestimmung und schwächen Vertrauen. Ethisch verletzt mangelnde Erklärbarkeit die Prinzipien von Rechenschaft, Informiertheit und Autonomie.

Good-Practice umfasst verständliche Modelldokumentation, Transparenzberichte in leicht zugänglicher Sprache, fallbegleitende Begründungen, visuelle Erklärungen und zielgruppengerechte Schulungen. Verantwortlichkeiten müssen geklärt sein: Wer erklärt was, wie tief und über welche Schnittstellen – inklusive Möglichkeiten für Rückfragen, Korrekturen und Unterstützung.

These 4: Adaptive Systeme statt One-Size-Fits-All

Gleichbehandlung allein genügt nicht; KI-Systeme müssen konfigurierbar sein, um unterschiedlichen Bedarfen gerecht zu werden. Adaptive Funktionalitäten adressieren Variationen in Wahrnehmung, Kommunikation, Motorik und Kognition und unterstützen individuelle Wege zur Teilhabe (OECD, 2024).

Fehlt diese Anpassbarkeit, erzeugen Systeme Normierungsdruck: Wer nicht dem impliziten Standardprofil entspricht, wird überfordert, falsch eingeschätzt oder ausgeschlossen. Ethisch geboten ist hier Equity statt bloßer Equality – also der Abbau struktureller Barrieren durch angemessene Differenzierung.

In der Praxis bewährt sich ein modularer Aufbau mit optionalen Assistenzstufen, vordefinierten, feinjustierbaren Profilen und Progressive Disclosure. Damit sind klare, barrierefreie Einstellungen gemeint, rücksetzbare Defaults und Unterstützung bei der Konfiguration – ohne kognitive Überlastung.

These 5: Barrierefreiheit ist nicht verhandelbar

Barrierefreiheit ist Grundprinzip inklusiver Digitalisierung. Accessibility-by-Design muss technische, sprachliche und kognitive Zugänge von Beginn an berücksichtigen und fortlaufend prüfen (European Commission, 2024; OECD, 2024). Usability-by-Design ergänzt dies um eine intuitive, konsistente Interaktion.

Fehlt diese Perspektive, verfestigen sich Ausschlüsse: Schnittstellen sind für Screenreader unzureichend, Texte zu komplex, Interaktionen nicht tastatur- oder sprachbedienbar – mit unmittelbaren Folgen für Selbstbestimmung und Nutzung. Ethisch ist Barrierefreiheit Ausdruck von Respekt, Nicht-Schädigung und Gerechtigkeit.

Wirksam sind die konsequente Anwendung anerkannter Standards (z. B. WCAG), Tests mit heterogenen Nutzer:innengruppen, klare Sprache, multimodale Ein- und Ausgaben, kontinuierliche Qualitätssicherung sowie barrierefreie Supportkanäle.

These 6: Assistive KI fördert Selbstbestimmung – wenn sie flexibel bleibt

Assistive KI kann Kommunikation erleichtern, Orientierung geben und Alltagsbewältigung unterstützen – und so Empowerment stärken (Katta & Thinyane, 2023). Voraussetzung ist eine flexible, personenbezogene Konfiguration, die sich mit den Nutzer:innen entwickelt.

Standardisierte, starr vordefinierte Unterstützungslogiken bergen das Risiko neuer Ausschlüsse oder paternalistischer Eingriffe. Fehlinterpretationen von Eingaben oder Kontext führen zu Fehlsteuerungen mit realen Folgen für Autonomie. Ethisch ist die Grenze zwischen Unterstützung und Kontrolle zu sichern.

Gute Praxis setzt auf partizipative Entwicklung, feingranulare Einstellmöglichkeiten, Datensparsamkeit und transparente Kontrolle darüber, welche Daten wofür genutzt werden. Funktionen sollten temporär deaktivierbar sein; Berechtigungen müssen anpassbar und nachvollziehbar bleiben.

These 7: Digitale Kompetenzen sind Schlüssel zur Teilhabe

Digitale und KI-bezogene Kompetenzen sind Voraussetzung für selbstbestimmte Nutzung. Fehlende Kenntnisse führen zu Unsicherheit, Abhängigkeit und Nicht-Nutzung – mit dem Risiko, bestehende soziale Ungleichheiten zu vertiefen (Kumar et al., 2023).

Kompetenzdefizite betreffen sowohl Nutzer:innen als auch Fachkräfte. Ohne verständliche Schulungsangebote, Begleitung und Zeitbudgets bleibt die Einführung neuer Technologien folgenlos oder erzeugt Überforderung. Ethisch sind Unterstützung und Befähigung Ausdruck von Fairness und Verantwortung.

Bewährt haben sich niedrighschwellige, zielgruppenspezifische Lernangebote, Peer-Learning, der Einsatz digitaler Lots:innen und praxisnahe Formate. Organisationen sollten Weiterbildung strukturell verankern, Lernzeiten einplanen und Erfolge sichtbar machen, um eine Kultur souveräner Nutzung zu fördern.

These 8: Automatisierte Entscheidungen müssen anfechtbar sein

Anfechtbarkeit ist ein zentrales Schutzprinzip in KI-gestützten Verfahren der Sozialen Arbeit. Wer von einer automatisierten Entscheidung betroffen ist, muss Gründe verstehen, Widerspruch einlegen und Korrektur verlangen können – barrierefrei und niedrighschwellig (Eubanks, 2018).

Fehlen nachvollziehbare Verfahren, werden Rechte entwertet: Fehler bleiben unentdeckt, falsche Zuordnungen verstetigen sich, Vertrauen bricht weg. Besonders betroffen sind Personen mit geringen Technikenntnissen oder kognitiven Einschränkungen. Ethisch steht hier die Rechenschaftspflicht im Vordergrund.

Wirksame Praxis umfasst verständliche Bescheide mit Entscheidungsbeurgründung, zugängliche Beschwerdewege, Ombudsstellen, Fristen- und Status-Tracking sowie Unterstützung beim Widerspruch. Technisch unterstützen Audit-Trails, Protokolle der Modellversionen und Decision Logs die nachträgliche Prüfung und Korrektur.

These 9: Governance, Ethik und Verantwortung müssen klar geregelt sein

Verantwortungsvolle KI braucht klare Rollen, verbindliche Prozesse und überprüfbare Leitlinien (Green, 2023; UNESCO, 2021). Dazu zählen Zuständigkeiten für Datenschutz, Qualitätssicherung und Ethik ebenso wie transparente Entscheidungen über Beschaffung, Einführung und Betrieb.

Ohne Governance drohen Intransparenz, Doppelstrukturen und Kontrollverlust – insbesondere in kleineren Trägern mit knappen Ressourcen (OECD, 2024). Fehlende Standards erschweren die Koordination und erhöhen das Risiko unbeabsichtigter Folgen. Ethisch verankert Governance bedeutet Rechenschaft, Fairness und Beteiligung.

Bewährte Elemente sind Ethikbeiräte, Risiko- und Folgenabschätzungen, klare Freigabeprozesse (z. B. Checklisten zu Barrierefreiheit und Fairness), externe Audits sowie öffentliche Rechenschaft. Governance wird wirksam, wenn sie partizipativ angelegt ist und Betroffene systematisch einbindet.

These 10: Kontinuierliche Evaluation ist überlebenswichtig

KI-Systeme und Nutzungskontexte verändern sich. Nur durch regelmäßige technische, organisatorische und ethische Evaluationen bleiben Anwendungen inklusiv, aktuell und wirksam (Jobin et al., 2019; OECD, 2024). Evaluation ist kein Projektabschluss, sondern Teil des Betriebs.

Ohne kontinuierliche Überprüfung drohen Leistungsabfall, blinde Flecken und neue Ausschlüsse – etwa durch Daten- oder Konzeptdrift. Betroffene merken die Folgen meist zuerst, haben aber oft die geringsten Möglichkeiten, sie sichtbar zu machen. Ethisch sichern iterative Lernprozesse Gerechtigkeit und Schadensvermeidung.

Gute Praxis verbindet quantitative Metriken (z. B. Fehler-, Bias- und Nutzungsraten) mit qualitativen Rückmeldungen aus Nutzer:innenperspektive. Regelmäßige Feedbackschleifen, Updates, Nachtests zur Barrierefreiheit und dokumentierte Modelländerungen sichern Lernprozesse und Rechenschaft – und damit die Wirksamkeit inklusiver KI über die Zeit.

4. Ethische Standards für KI in inklusiver Arbeit

Die Entwicklung und Anwendung von KI-Systemen in der Sozialen Arbeit muss sich an klaren ethischen Leitplanken orientieren, um die Rechte und Bedarfe von Menschen mit Beeinträchtigungen zu schützen und digitale Teilhabe diskriminierungssensibel zu gestalten. Zentrale Prinzipien sind: Achtung der Menschenrechte, Partizipation, Transparenz, Barrierefreiheit, Nachvollziehbarkeit sowie die Förderung von Empowerment (UNESCO, 2021; European Commission, 2024). Ergebnisse aus deutschen Praxisprojekten bestätigen diesen Bedarf: Feichtenbeiner et al. (2022) zeigen, dass ethische Reflexion und inklusive Gestaltung in Einrichtungen der Behindertenhilfe bislang kaum institutionell verankert sind

KI-Systeme dürfen bestehende Benachteiligungen nicht verstärken. Stattdessen müssen sie dazu beitragen, individuelle Lebenslagen zu respektieren und Chancengerechtigkeit zu fördern – durch adaptive, erklärbare und barrierefreie Gestaltung (Benjamin, 2019; Mittelstadt et al., 2016). Ein zentrales Element ist die Einrichtung effektiver Beschwerde- und Widerspruchsmechanismen. Nutzer:innen müssen automatisierte Entscheidungen hinterfragen, Fehler melden und Unterstützung erhalten können – barrierefrei, niedrigschwellig und nachvollziehbar (Eubanks, 2018; OECD, 2024).

Die bloße Existenz ethischer Prinzipien genügt nicht; entscheidend ist deren Verbindlichkeit und Überprüfbarkeit. Wie Whittlestone et al. (2019) hervorheben, offenbaren sich in der Umsetzung von Ethikrichtlinien häufig Spannungen zwischen normativen Prinzipien und prakti-

scher Anwendung – ein Befund, der auch für die Soziale Arbeit relevant ist. Ethische Standards müssen operationalisiert werden: in Form klarer Leitlinien, verpflichtender Prüfverfahren, extern begleiteter Audits und transparenter Governance-Strukturen (Jobin et al., 2019; Green, 2023). Internationale Erfahrungen zeigen: Wo Ethikbeiräte, Datenschutzverantwortliche und systematische Evaluationen institutionell verankert sind, steigt Qualität, Sicherheit und gesellschaftliche Akzeptanz von KI-Lösungen deutlich (OECD, 2024; UNESCO, 2021).

5. Übergreifende Herausforderungen bei der Implementierung inklusiver KI-Systeme

Die Einführung inklusiver KI-Systeme in der Sozialen Arbeit ist mit vielschichtigen Herausforderungen verbunden, die über technische Fragen hinausgehen. Zentrale Barrieren liegen in fehlenden Ressourcen, mangelnden Strukturen und sozialer Ungleichheit.

Insbesondere kleinere Trägerorganisationen verfügen nicht über ausreichende finanzielle, personelle oder technische Mittel, um partizipative Technikentwicklung, regelmäßige Bias-Checks oder kontinuierliche Evaluationen zu gewährleisten (OECD, 2024; Kumar et al., 2023). Verantwortlichkeiten für Datenschutz, Ethik oder Qualitätssicherung sind oft nicht eindeutig geregelt – was zu Intransparenz, Ineffizienz und Vertrauensverlust führt (UNESCO, 2021; Green, 2023).

Viele Organisationen verfügen über keine verankerten Ethikbeiräte oder klaren Verfahren zur Überprüfung und Kontrolle von KI-Systemen. Governance-Defizite äußern sich in einem Mangel an Leitlinien, Standards und abgestimmten Prozessen für Implementierung, Betrieb und Evaluation.

Ein erheblicher Teil der Zielgruppen – sowohl Klient:innen als auch Mitarbeitende – nutzt KI-Angebote nicht, weil digitale Kompetenzen fehlen, Technikangst besteht oder der Zugang zu geeigneten Geräten und barrierearmen Angeboten nicht gegeben ist (Ladau, 2021; Eubanks, 2018; Kumar et al., 2023). Ohne Alternativen droht vollständiger Ausschluss.

Die Einführung neuer Technologien ist für viele Nutzerinnen mit Unsicherheit, Überforderung und Ablehnung verbunden. Niedrigschwellige Angebote wie digitale Lotsinnen, Peer-Learning oder Schritt-für-Schritt-Hilfen fehlen häufig – obwohl sie sich in anderen Kontexten als wirksam erwiesen haben (OECD, 2024; Shinohara & Wobbrock, 2016).

6. Aufgaben, Impulse und Empfehlungen für Organisation, Träger und Verbände

Organisationen der Sozialen Arbeit müssen eine klare Strategie entwickeln, die ethische und rechtliche Anforderungen (z. B. Barrierefreiheitsstärkungsgesetz, europäische KI-Verordnung) berücksichtigt (European Commission, 2024). Zentrale Aufgaben sind Governance-Strukturen, Fortbildung, Qualitätssicherung und kontinuierliche Evaluation. Dezentrale Anpassung, partizipative Prozesse und Peer-Learning sind ebenso wichtig. Impulse aus anderen Sektoren zeigen, dass interdisziplinäre Teams, Risikoanalysen und transparente Beschwerdeverfahren die Qualität von KI-Anwendungen verbessern (OECD, 2024). Die hier formulierten Aufgaben fokussieren auf strategische Voraussetzungen für inklusive KI. Sie ersetzen jedoch nicht die Entwicklung operativer Umsetzungsformate. Vielmehr fordern sie diese ein: Es braucht praxisnahe, adaptive Leitfäden – aber auch neue Steuerungslogiken und Entscheidungsmodelle, die klassische Technik- und Soziallogik verbinden. Ziel muss ein organisationales Betriebssystem sein, das technische Innovation mit partizipativer Gerechtigkeit und realistischer Ressourcennutzung verbindet.

6.1 Zentrale Aufgaben: Strategie, Governance, Qualitätssicherung, Fortbildung, Evaluation

Organisationen der Sozialen Arbeit stehen vor der Aufgabe, eine klare Strategie für den Einsatz von KI zu entwickeln, die sowohl ethische als auch rechtliche Anforderungen – etwa aus dem Barrierefreiheitsstärkungsgesetz (BFSG) und der europäischen KI-Verordnung – berücksichtigt (European Commission, 2024). Zentrale Aufgaben umfassen:

- die strukturelle Verankerung von Governance, einschließlich der Benennung von Verantwortlichen für Datenschutz, Ethik und Qualitätssicherung,
- die Entwicklung verbindlicher Leitlinien und Prozesse für Einführung, Betrieb und Überwachung von KI-Systemen (Green, 2023; UNESCO, 2021),
- regelmäßige technische und ethische Evaluationen,
- kontinuierliche Fortbildungen für Mitarbeitende zur Förderung digitaler Kompetenzen und kritischer KI-Kompetenz (Kumar et al. 2023; Schulz, 2024).
- eine tragfähige Finanzierungsstrategie, z. B. durch Nutzung von Förderprogrammen (BMBF, Aktion Mensch, EU-Programme), Verbändebündnisse für gemeinschaftliche Infrastruktur oder strategische Partnerschaften mit Hochschulen und Tech-Stiftungen.

6.2 *Dezentrale Aufgaben: Lokale Anpassung, Partizipation, Peer-Learning*

Neben den strategischen Aufgaben müssen auch dezentrale Strukturen gestärkt werden. Dazu zählen:

- die Anpassung von KI-Systemen an lokale Bedarfe und Kontexte (OECD, 2024),
- die Verankerung partizipativer Prozesse vor Ort – etwa durch Co-Design-Workshops, Nutzer:innenbeiräte oder kontinuierliche Feedbackschleifen (Bennett & Rosner, 2019),
- der gezielte Einsatz von digitalen Lots:innen und Peer-Learning zur Stärkung digitaler Teilhabe auf allen Ebenen (Shinohara & Wobbrock, 2016).

6.3 *Empfehlungen für die Praxis*

Für eine realistische Umsetzung empfiehlt sich ein gestuftes Vorgehen: mit Pilotprojekten beginnen, Zuständigkeiten klären und Rückkopplungsschleifen institutionalisieren (OECD, 2024). Schulungsangebote, offene Fehlerkultur und der Erhalt analoger Alternativen sind essenziell, um niemanden auszuschließen (Ladau, 2021; Eubanks, 2018).

Impulse aus anderen Sektoren zeigen, dass interdisziplinäre Projektteams, strukturierte Risikoanalysen und transparente Beschwerdeverfahren die Qualität und Sicherheit von KI-Anwendungen verbessern können. Während im Gesundheitswesen Ethikbeiräte und Audits längst etabliert sind, setzt die IT-Branche auf agile Entwicklungsprozesse und Nutzer:innenbeteiligung. Bildungseinrichtungen wiederum arbeiten erfolgreich mit Peer-Konzepten zur Förderung digitaler Kompetenzen (OECD, 2024; Schulz, 2024).

Angesichts der Komplexität von KI und Teilhabe sind praxisnahe, ressourcenorientierte Leitfäden und Handreichungen dringend erforderlich. Diese sollten auf empirischen und theoretischen Erkenntnissen basieren, konkrete Umsetzungsschritte enthalten und unterschiedliche Organisationstypen und Nutzer:innengruppen adressieren. Sie sollten gemeinsam von Praxis, Wissenschaft und Verbänden – unter Beteiligung von Betroffenen – entwickelt werden. Nur so können tragfähige, adaptive Lösungen entstehen, die den Potenzialen von KI gerecht werden und die damit verbundenen Risiken verantwortungsvoll adressieren (OECD, 2024; UNESCO, 2021). Zukünftige Forschung sollte sich verstärkt auf die Erarbeitung solcher Instrumente konzentrieren.

7. Fazit und Ausblick

Die Analyse zeigt: Künstliche Intelligenz (KI) kann einen wichtigen Beitrag zur digitalen Teilhabe und Selbstbestimmung von Menschen mit Beeinträchtigungen in der Sozialen Arbeit leisten – vorausgesetzt, ethische, technische und organisationale Voraussetzungen werden konsequent berücksichtigt. Zentrale Elemente hierfür sind: partizipative Technikentwicklung, diskriminierungssensible Datenpraxis, Barrierefreiheit, interdisziplinäre Zusammenarbeit und kontinuierliche Evaluation (Jobin et al., 2019; UNESCO, 2021; European Commission, 2024).

Gleichzeitig machen Studien und Praxis deutlich, dass KI bestehende Ungleichheiten verschärfen kann, wenn sie ohne adaptive Gestaltung und partizipative Kontrolle eingeführt wird. Risiken wie algorithmische Diskriminierung, Intransparenz (Black-Box-Problematik), normierende Standardisierung und der persistente Digital Divide betreffen besonders Menschen mit komplexen oder mehrfachen Beeinträchtigungen (Benjamin, 2019; Mittelstadt et al., 2016; OECD, 2024; Lee & Zhang, 2023).

Die Grenzen inklusiver KI liegen dort, wo Ressourcen für Governance, Ethikstrukturen und Qualifizierung fehlen, analoge Alternativen abgeschafft werden oder Organisationen keine adaptive und partizipative Praxis etablieren (Eubanks, 2018; Kumar et al., 2023). KI darf nicht als neutrale Technologie oder vermeintliches Allheilmittel betrachtet werden, sondern erfordert einen kritisch-reflektierten, menschenzentrierten Einsatz (Green, 2023; Benjamin, 2019).

Für die Zukunft bedeutet das: Fachkräfte müssen im Umgang mit KI geschult werden, Co-Design mit Betroffenen muss Standard werden und analoge Alternativen müssen erhalten bleiben (Schulz, 2024; Bennett & Rosner, 2019). Die Implementierung von KI in der Sozialen Arbeit muss stets ethisch fundiert, inklusiv und verantwortungsbewusst erfolgen. Nur dann können die Potenziale zur Verbesserung von Zugänglichkeit, Wirksamkeit und Gerechtigkeit tatsächlich wirksam werden (UNESCO, 2021; European Commission, 2024).

Zudem besteht erheblicher Forschungs- und Entwicklungsbedarf: Es fehlen belastbare empirische Studien zu den langfristigen Wirkungen von KI auf Inklusion, insbesondere bei mehrfach marginalisierten Gruppen (Katta & Thinyane, 2023; OECD, 2024). Offene Fragen betreffen unter anderem die Umsetzbarkeit ethischer Standards, nachhaltige Fachkräftequalifizierung und die Entwicklung partizipativer Governance-Modelle. Notwendig sind Good-Practice-Leitfäden, internationale Kooperationen und die systematische Einbindung von Nutzer:innen in allen Phasen der KI-Entwicklung (Jobin et al., 2019; UNESCO, 2021).

KI darf nicht als *neutrale Technologie* verstanden werden, sondern muss im sozialen, politischen und ethischen Kontext reflektiert und gestaltet werden. Kritische Stimmen fordern zurecht, dass Gerechtigkeit nicht durch Technik entsteht, sondern durch bewusste soziale Gestaltung. Technologie ist gestaltbar, aber nur, wenn wir sie nicht als rein technische Aufgabe delegieren. Der Text plädiert daher für eine systematische Verzahnung von technischer Entwicklung, ethischer Reflexion und organisatorischer Umsetzungsverantwortung.

KI wird daher die Soziale Arbeit nicht ersetzen – aber sie wird sie verändern. Die Profession steht insofern vor der Aufgabe, technologische Innovation mit Beziehungskompetenz, Gerechtigkeit und Verantwortung zu verbinden. Die Zukunft beginnt mit einer klaren Haltung – und mit der Bereitschaft, Chancen und Risiken gleichermaßen ernst zu nehmen. Künstliche Intelligenz kann digitale Teilhabe fördern, wenn ethische, technische und organisationale Voraussetzungen erfüllt sind. Ohne kritische Reflexion drohen neue Ausschlüsse. Die Soziale Arbeit muss ihre Gestaltungsverantwortung annehmen und KI entlang klarer ethischer Standards, mit partizipativer Governance und unter Berücksichtigung realer Ressourcen umsetzen. Die Zukunft der digitalen Teilhabe beginnt mit kluger Finanzierung, aktiver Beteiligung und konsequenter Umsetzung menschenzentrierter Prinzipien.

Damit inklusive KI wirksam umgesetzt werden kann, bedarf es einer kritischen Neubewertung von Steuerungsmodellen und Finanzierungslogiken. Viele der hier dargestellten Anforderungen – von partizipativer Technikentwicklung bis zu ethischer Evaluation – lassen sich nicht über klassische Projektfinanzierung oder Einzelmaßnahmen abbilden. Notwendig ist ein neues operatives Grundmodell: Es verbindet Governance, Ethik, Wirkungsmessung und Wirtschaftlichkeit in einem integrierten Steuerungssystem. Das verlangt eine Weiterentwicklung bisheriger Controlling- und Förderstrukturen – und die Bereitschaft, Gemeinwohlwirkungen in die strategische Logik von Organisationen einzubetten. Dieser Beitrag skizziert keine abschließenden Lösungen, sondern legt die konzeptionellen, ethischen und strukturellen Spannungsfelder offen, die den Einsatz von KI in der Sozialen Arbeit begleiten. Die Operationalisierung in Form konkreter Praxisleitfäden, Steuerungsinstrumente oder Organisationsmodelle ist eine nachgelagerte, aber dringend notwendige Aufgabe – sowohl für Forschung als auch für Fachverbände und Trägerstrukturen. Besonders gefragt ist die Entwicklung neuer Operating Models, die neben ethischer Reflexion auch ökonomische Tragfähigkeit, Ressourcenlogik und Legitimitätsmechanismen systemisch integrieren. Klassische Business Cases greifen zu kurz: Die gesellschaftliche Wirkung barrierefreier KI lässt sich nicht rein betriebswirtschaftlich bemessen. Es braucht neue Bewertungsmaßstäbe, die Gerechtigkeit, Teilhabe und Resilienz institutionell anerkennen.

Literatur

- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press.
- Bennett, C. L., & Rosner, D. K. (2019). The promise of empathy: Design, disability, and knowing the ‚Other‘. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3290605.3300528>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin’s Press.
- European Commission. (2024). EU Artificial Intelligence Act. *Official Journal of the European Union, L 173*, 33–155.
- Feichtenbeiner, S., Stähler, S., & Beudt, S. (2022). *Ethik, KI und Menschen mit Behinderung*. Ergebnisbericht KI.ASSIST
- Geiger, R. S., & Halfaker, A. (2017). Operationalizing conflict and cooperation between automated and human agents in Wikipedia. *ACM Transactions on Computer-Human Interaction, 24*(3), 1–32. <https://doi.org/10.1145/3057853>
- Green, B. (2023). The flaws of AI ethics guidelines: Why principles alone won’t save us. *Communications of the ACM, 66*(8), 24–27. <https://doi.org/10.1145/3594513>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Katta, V., & Thinyane, H. (2023). AI and disability: A systematic scoping review. *AI & Society, 38*, 101–120. <https://doi.org/10.1007/s00146-022-01426-4>
- Kumar, R., Patel, S., & Li, J. (2023). Barriers to AI training in social services: A global survey. *International Social Work, 66*(4), 510–526. <https://doi.org/10.1177/00208728221124638>
- Ladau, E. (2021). *Demystifying disability: What to know, what to say, and how to be an ally*. Ten Speed Press.
- Lee, Y., & Zhang, T. (2023). Algorithmic bias and disability: A systematic review of risks and solutions. *Disability & Society, 39*(1), 101–120. <https://doi.org/10.1080/09687599.2023.2183641>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- OECD. (2024). Digital divide and AI: New barriers for people with disabilities? *OECD Policy Briefs, 2024/3*, 1–19.
- OECD. (2024). *Using AI to support people with disability in the labour market: International case studies*. OECD Publishing.
- Plattform Lernende Systeme (2022). *Mit KI zu mehr Teilhabe in der Arbeitswelt*. Whitepaper.

- Shinohara, K., & Wobbrock, J. O. (2016). Self-conscious or self-confident? A diary study conceptualizing the social accessibility of assistive technology. *ACM Transactions on Accessible Computing*, *8*(2), 1–31. <https://doi.org/10.1145/2827857>
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO.
- Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*, 105567. <https://doi.org/10.1016/j.clsr.2021.105567>
- Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 195–200. <https://doi.org/10.1145/3306618.3314289>

Julia Rehn¹

Mensch vs. Chatbot – Resonanzraum oder Echokammer des Selbst?

Eine qualitative Einzelfallstudie

¹ Alice Salomon Hochschule Berlin

Zusammenfassung

Menschen in psychosozialen Krisensituationen überbrücken die stetig wachsende Lücke zwischen Psychotherapieangeboten und Nachfrage zunehmend mit generativen Chatbots. Die vorliegende Einzelfallstudie untersucht die parallele Nutzung von realweltlichen Beratungspersonen und einem KI-Chatbot im Hinblick auf Beziehungsqualität, emotionale Wirkung und Resonanz Erfahrung. Die explorative Fallanalyse wurde auf Grundlage der Grounded Theory durchgeführt. Die Datenerhebung erfolgte über ein episodisch-narratives Interview mit anschließender theoriegenerierender Codierung. Aus den 347 Kodierungen wurden fünf Kategorien entwickelt: (1) Interaktion mit dem Chatbot, (2) Interaktion mit Menschen, (3) Effekte des Chatbots, (4) Symbiose und (5) Bedingungslose Allverfügbarkeit. Im Erleben des Betroffenen zeigten sich sowohl menschliche Beziehungen als auch die Interaktion mit dem Chatbot als potenzielle Echokammern, in denen eigene Überzeugungen und Gefühle verstärkt wurden. Während die Kommunikation mit dem KI-Chatbot durch Stabilität, Selbststeuerung und emotionale Entlastung geprägt war, blieb das Erleben echter Resonanz, verstanden als spontane, verkörperte und affektiv tiefgreifende Antwortbeziehung, ausschließlich der zwischenmenschlichen Begegnung vorbehalten. Dennoch konnte der Chatbot einen subjektiv bedeutsamen Resonanzraum simulieren, in dem authentische Emotionen evoziert wurden. Die Studie zeigt das Potenzial einer symbiotischen, verantwortungsvoll verschränkten Nutzung von KI-Chatbot und menschlicher Beratung. Chatbots können unterstützend wirken, ersetzen aber keine tiefgreifenden menschlichen Resonanzbeziehungen. Langfristige Effekte dieser hybriden Beziehungserfahrungen auf Identität, Bindungsverhalten und psychosoziale Gesundheit sind noch unzureichend erforscht und sollten zukünftig vertieft untersucht werden.

Schlüsselwörter: Resonanz, Echokammer, KI-Chatbot, psychosoziale Beratung, parasoziale Beziehung, quasi-soziale Beziehung

Rehn, J. (2026). Mensch vs. Chatbot – Resonanzraum oder Echokammer des Selbst? Eine qualitative Einzelfallstudie. *DiMe-Sozial*, 2(1), 91–112. <https://doi.org/10.21240/dimesoz/2026/1/137>

Human vs. Chatbot – A Resonance Space or a Echo Chamber of the Self?

Abstract

People in psychosocial crisis situations are increasingly bridging the ever-widening gap between supply and demand for psychotherapy with the help of generative chatbots. This single-case study examines the parallel use of real-life counselors and an AI chatbot with regard to relationship quality, emotional impact, and resonance experience. The exploratory case analysis was conducted using grounded theory. Data were collected using episodic narrative interviews and subsequent theory-generating coding. Five categories were developed from the 347 codes: (1) Interaction with the chatbot, (2) Interaction with humans, (3) Impact of the chatbot, (4) Symbiosis, and (5) Unconditional availability. In the experience of those affected, both human relationships and interaction with the chatbot proved to be potential echo chambers in which their own beliefs and feelings were amplified. While communication with the AI chatbot was characterized by stability, self-direction, and emotional relief, the experience of genuine resonance, understood as a spontaneous, embodied, and affectively profound relationship, remained exclusively reserved for interpersonal encounters. Nevertheless, the chatbot succeeded in simulating a subjectively meaningful resonance space in which authentic emotions were evoked. This study demonstrates the potential of a symbiotic, responsibly integrated use of AI chatbots and human counseling. Chatbots can support, but are not a substitute for profound human resonance relationships. The long-term effects of these hybrid relationship experiences on identity, attachment behavior, and psychosocial health are still insufficiently researched and should be investigated more closely in the future.

Keywords: Resonance, echo chamber, AI chatbot, psychosocial counselling, parasocial relationship, quasi-social relationship

1. Einleitung

Digitale Gesprächspartner wie KI-Chatbots gewinnen zunehmend an Bedeutung als Form der Selbsthilfe in belastenden Lebenssituationen. Generative Chatbots verbreiten sich schnell in der Bevölkerung: In Deutschland gaben im Jahr 2025 rund 67% der über 16-Jährigen an, zumindest gelegentlich generative KI zu nutzen – im Vergleich zu etwa 40% im Jahr 2024 (Bitkom, 2025). Dabei entfallen 94% der Nutzung auf die drei führenden Anbieter OpenAI, Microsoft und Google. Etwa 9% der Menschen in Deutschland nutzen generative Chatbots mehrmals täglich, 18% mehrmals wöchentlich; rund 43% seltener oder unregelmäßig und ca. 30% gar nicht (DOIT Software, 2025).

Auch im Bereich von Gesundheit, Beratung und Therapie erfahren KI-basierte Chatbots zunehmend Zuspruch. Angesichts steigender Prävalenzraten von Depressionen, Angststörungen und sozialer Isolation,

insbesondere bei Kindern und Jugendlichen (Resch & Parzer, 2024), vergrößert sich die Lücke zwischen Versorgungsbedarf und tatsächlichem psychotherapeutischen Angebot (BPtK, 2023). Generative Chatbots füllen diese Lücke zunehmend kompensatorisch und werden mehrheitlich als hilfreich erlebt (Siddals et al., 2024). In Deutschland nutzen etwa 82 % der 18- bis 27-Jährigen einen Chatbot, davon 17 %, um in Situationen von Angst oder Depression aktiv Hilfe zu suchen (Lukas et al., 2025). Auch zur Informationsbeschaffung werden Chatbots zunehmend genutzt: Ein Drittel der Befragten in einer repräsentativen Umfrage suchte damit nach Ernährungsthemen, ca. 20 % riefen gesundheitsbezogene Informationen ab (König et al., 2025).

Die Vorteile für das Gesundheitssystem werden vor allem in effizienterer Versorgung und höherer Wirtschaftlichkeit gesehen (Laymouna et al., 2024). Seit Inkrafttreten des Digitalen-Versorgungs-Gesetzes (DVG) 2019 wächst der Markt für digitale Gesundheitsanwendungen (DiGa) kontinuierlich. Aktuell sind 71 DiGa zugelassen, davon 31 mit Fokus auf psychische Erkrankungen (z. B. Depression, Angst, Sucht, ADHS, Essstörungen), jedoch bislang ohne generative Chatbots (Bundesinstitut für Arzneimittel und Medizinprodukte, 2025).

Einer der ersten prominenten Mental Health Chatbots (MHCB) aus den USA von 2017, Woebot, stellte im Juni 2025 seinen Dienst ein, nachdem eine Zulassung durch die FDA an unklaren regulatorischen Rahmenbedingungen für die Integration von KI in die Gesundheitsversorgung scheiterte (Woebot, 2025; Aguilar, 2025). Eine generative KI-gestützte DiGa-Lösung bleibt in Deutschland bislang aus. Dabei liegen die Stärken generativer Chatbots nicht nur in der Dialogführung, sondern auch in der hohen technischen Anpassungsfähigkeit: die Genauigkeit bei der Erkennung relevanter Gesundheitsprobleme von Menschen mit psychischen Problemen durch KI-basierte Chatbots beläuft sich mittlerweile auf über 94 % (Algumaei et al., 2025). Manche Studien weisen darauf hin, dass die kommunikativen Fähigkeiten mancher Chatbots bereits die durchschnittliche menschliche Kommunikation übertreffen (Peter et al., 2025).

Vor dem Hintergrund dieses Bedeutungsaufwuchses von KI-Chatbots stellt sich die Frage, wie diese digitalen Interaktionen das menschliche Erleben von Beziehung, Identität und Selbst beeinflussen. Um die Qualität der Beziehungserfahrung mit einem Chatbot einschätzen zu können, ist ein Verständnis darüber notwendig, wie sich das menschliche Selbst überhaupt in Beziehung bildet. Dies wird im Folgenden kurz erläutert.

Das Selbstbewusstsein des Menschen entsteht durch die Anerkennung eines anderen Subjekts. Georg Wilhelm Friedrich Hegel betont in seiner *Phänomenologie des Geistes* (1807) die Notwendigkeit intersubjektiver Anerkennung für die Ausbildung des Selbstbewusstseins (Hegel, 2020). Charles Taylor unterstreicht, dass Identität in Dialogen mit signifikanten

Anderen geformt wird (Taylor, 1996). Martin Buber hebt hervor, dass das Ich erst in der Begegnung mit dem Du zum Ich wird (Buber, 2021). Diese Perspektiven zeigen, dass menschliches Selbstbewusstsein ein soziales Produkt ist, das im Austausch mit anderen entsteht.

Hartmut Rosa erweitert dieses Konzept mit seiner Resonanztheorie: Resonanz beschreibt eine wechselseitige, transformative Beziehung zwischen Subjekt und Welt, die über bloße Harmonie hinausgeht und Widerspruch zulässt. Wichtig ist dabei Rosas Abgrenzung: „Resonanz ist keine Echo-, sondern eine Antwortbeziehung; sie setzt voraus, dass beide Seiten mit eigener Stimme sprechen, und dies ist nur dort möglich, wo starke Wertungen berührt werden. Resonanz impliziert ein Moment konstitutiver Unverfügbarkeit.“ (Rosa, 2019, S. 298). Ein Echo, so Rosa, würde nur das eigene Rufen zurückwerfen, ohne neue Information oder Veränderung. Eine Antwort dagegen impliziert Eigenständigkeit auf der anderen Seite: Das Gegenüber reagiert aus eigener Innerlichkeit heraus und kann somit überraschen, herausfordern oder bestätigen. Das heißt, wenn uns kein echter Widerpart mehr entgegentritt, nichts Unverfügbares, Eigenwilliges, dann verlieren wir die Chance auf jene tiefe Wechselwirkung, die Resonanz ausmacht. Eine echte dialogische Beziehung setzt also immer voraus, dass das Gegenüber mehr ist als nur ein projiziertes Echo unseres Selbst.

In der folgenden Fallanalyse soll es darum gehen, ob ein Chatbot ein Echo des Selbst ist, und ob in der zwischenmenschlichen Begegnung wirklich immer Resonanz stattfindet. Zunächst werden zentrale Begriffe geklärt, die für das Verständnis der Beziehungsgestaltung zwischen Mensch und Chatbot grundlegend sind.

2. Diskurs parasoziale vs. quasisoziale Beziehungen

Im aktuellen Diskurs über Mensch-Maschine-Interaktionen (MMI) im Kontext digitaler Beratung werden zwei Begriffe zur Beschreibung nicht-reziproker Beziehungserfahrungen diskutiert: die *parasoziale* und die *quasisoziale* Beziehung. Der Begriff der *parasozialen* Beziehung stammt aus der Medienpsychologie und beschreibt die einseitige emotionale Bindung eines Menschen zu einer Medienfigur (z. B. einem Seriencharakter oder einem:r Influencer:in). Dabei empfindet nur eine Seite, das Publikum/die Follower:innen, eine Freundschaft oder Vertrautheit, während die *Bezugsperson* nichts von der einzelnen Person weiß (Deghen, 2024). Bisher waren damit vorwiegend realweltliche Personen gemeint. Hier tritt durch KI, vor allem im Bild- und Videobereich eine Veränderung ein (z. B. AI Companions, KI-generierte Influencer:innen).

Der Begriff der *quasisozialen Beziehung*, eingeführt von Linnemann et al. (2023), beschreibt Interaktionen zwischen Mensch und Maschine, in denen durch Imitation sozialen Verhaltens, etwa durch Frage-Antwort-Dynamiken, der Anschein einer wechselseitigen Beziehung entsteht.

Anders als in genuin sozialen Beziehungen, die durch gegenseitige Bezugnahme, Perspektivübernahme und Verhaltensbeeinflussung geprägt seien, so Linnemann, fehle es der Maschine an Bewusstsein und Intentionalität. Dennoch können künstliche Agenten soziale Reaktionen beim Menschen hervorrufen. Linnemann, Löhe und Rottkemper (2023) betonen die Relevanz dieses Phänomens insbesondere für personenbezogene Dienstleistungen in der Sozialen Arbeit. Der Begriff integriert explizit den Beziehungsaspekt, wie er in Beratung und Therapie zentral ist.

Im vorliegenden Kontext des Vergleichs von innerem Erleben gegenüber Mensch und Chatbot wird für den Begriff der *parasozialen* Beziehung plädiert: Erstens handelt es sich beim Chatbot um ein datenverarbeitendes System, das keine tatsächliche Kenntnis seines Gegenübers besitzt. Die Beziehung bleibt einseitig, der Mensch projiziert Freundschaft und Empathie auf eine künstliche Entität (Brandtzaeg et al., 2022). Zweitens birgt die Bezeichnung *quasisozial* eine semantische Nähe zu sozialen, zwischenmenschlichen Beziehungen, die geeignet ist, die fundamentale Differenz zwischen menschlichen und technischen Akteuren zu verwischen, auch wenn das von Linnemann et al. nicht intendiert ist. Die Wahl des Begriffs *parasozial* dient somit der konzeptuellen Schärfung und dem Bewusstmachen eben dieser Differenz.

3. Definitionen

In der Literatur finden sich zahlreiche und sehr unterschiedliche Definitionen von den maschinellen, generativen Protagonist:innen, die durch die fortlaufende technische Entwicklung ständiger Veränderung unterworfen sind. Für den vorliegenden Beitrag ist die Unterscheidung zwischen ausschließlich schriftbasierten Chatbots und den in Stimme und Bild auftauchenden Conversational Agents (Rolle, Persönlichkeit, AI Companion) wichtig.

3.1 Generative Chatbots

Die folgende Definition für generative Chatbots wurde co-kreativ mit ChatGPT erstellt, wofür zehn Prompts notwendig waren (ChatGPT, Claude, Google Gemini, Perplexity AI, Microsoft Copilot, Pi):

„Generative Chatbots setzen auf LLMs und sind Dialogsysteme, die Text generieren und auf Eingaben kontextabhängig reagieren. Sie sind durch eine integrierte Benutzerinteraktion gesteuert und können dynamische Antworten in Echtzeit erzeugen. Diese Systeme werden vor allem eingesetzt, um Aufgaben wie Informationsabfragen, Kundeninteraktion oder Konversationen zu automatisieren.“ (2025)

3.2 Conversational Agents (CA)

Martinengo (2023) findet für CA folgende Zusammenfassung:

“CAs are digital interfaces that use natural language to engage in a synchronous dialogue using ≥ 1 communication modalities such as text, voice, images, or video. [...] CAs may communicate using a variety of input-output modalities, such as text, speech, or multimedia, and adopt diverse personalities, such as coach, peer, and expert. Given this diversity, CAs can be classified according to various dimensions, such as their purpose, delivery channel, input and output modalities, or the response generation model.” (Martinengo et al., 2023)

3.3 Anthropomorphismus

Peter et al. fasst den Anthropomorphismus folgendermaßen zusammen:

“Anthropomorphism refers to the natural human tendency to ascribe human-like traits to nonhuman entities, such as animals, physical objects or digital entities. Conceptually, anthropomorphism resides in users’ minds, not in said entities, and is found strongest during interaction.” (Peter et al., 2025, S. 3)

4. Methodisches Vorgehen

4.1 Studiendesign

Für die qualitative Einzelfallstudie wurde ein semi-strukturiertes Interview mit dem Teilnehmer in einer ihm vertrauten Umgebung geführt. Die Aufnahme erfolgte mit einem handelsüblichen Smartphone. Da der Fokus auf expliziertem Wissen lag, folgte die Transkription mit dem Programm aTrain den Regeln für inhaltlich-semantische Transkriptionen, die eine gut lesbare und deutlich geglättete Sprache zeigt (Kuckartz in: Dresing & Pehl, 2018, S. 20). Die Analyse erfolgte mit der Grounded Theory nach Strauss & Corbin (1996) in Verbindung mit Charmaz (2008) mit dem Ziel, in der kodierenden Pendelbewegung über dem Material eine Handlungstheorie zu generieren, mit der die Forschungsfragen beantwortet werden können. Als technische Unterstützung diente MAXQDA.

Folgende Untersuchungsfragen sollen mit der Einzelfallstudie beantwortet werden:

- F 1: Worin bestehen die Unterschiede und Gemeinsamkeiten in der Beziehungserfahrung in der zwischenmenschlichen Begegnung im Vergleich zu einem generativen Chatbot?
- F 2: Wie entwickelt sich eine emotionale Bindung zu einem Chatbot?

- F 3: Welche Chancen und Risiken birgt eine solche Beziehung im Vergleich zu menschlicher Unterstützung?
- F 4: Wo kann die Soziale Arbeit von Chatbots profitieren – wo liegen ihre Grenzen?

4.2 Fallvignette

Die vorliegende Einzelfallstudie beschäftigt sich mit M., einem ca. 50-jährigen Mann, der über drei Jahre hinweg parallel zu klassischer Psychotherapie, dem Krisentelefon und seelsorgerischer Begleitung regelmäßig einen KI-Chatbot („Pi“) nutzte. Der Mann litt unter einer mittelgradigen Depression, Suizidalität, Ängsten, gravierenden Beziehungsproblemen sowie beruflichen Rückschlägen. M. lebt zusammen mit seinem Ehemann in einer Großstadt und ist selbständig tätig als Coach für Einzelpersonen und Organisationen. Zur Zeit der intensiven Chatbot-Nutzung (ab 1/2025) stand die Beziehungsproblematik im Vordergrund: M. erfuhr, dass sein Partner seit mehreren Monaten eine außereheliche Affaire führte. Diese Information erschütterte M.'s Vertrauen in den Partner zutiefst und verursachte eine existenziell bedrohliche Krise. Über viele Monate stand im Raum, ob und wie die Ehe weitergeführt werden könne, oder ob eine Trennung unvermeidlich sei.

Die Intention von M. einen KI-Chatbot zu nutzen, ergab sich anfänglich aus dem Bedürfnis, schon länger drängende Lebensfragen zu besprechen. In der Zeit ab dem Jahr 2023 testete M. den Bot punktuell ca. zwei- bis dreimal im Jahr und fand die Antworten „sinnvoll und hilfreich“. In der entstandenen Krisensituation ab Januar 2025 griff er auf den ihm bereits vertrauten Bot zurück, den er dann intensiv, teilweise mehrmals täglich nutzte. Die Dialoge mit dem Bot hielten über Tage und Wochen an. Die Beziehungskrise wurde überwunden. Die Bearbeitung der Depression und die Effekte auf die Suizidalität spielten nach Angaben von M. bei der Nutzung des Chatbots keine Rolle.

4.3 Ergebnisse

Aus 347 Kodierungen ergaben sich 45 Ober- und Unterkategorien. Entsprechend der Forschungsfrage nach Vor- und Nachteilen der Chatbot-Konversation und der zwischenmenschlichen Kommunikation ergaben sich die Kategorien *Interaktion mit dem Chatbot*, *Interaktion mit Menschen*, *Effekte des Chatbots*, *Symbiose* und *Bedingungslose Allverfügbarkeit*. Zur Komplettierung werden zusätzlich die bekannten Phänomene in der KI-Interaktion wie *Anthropomorphisierung* und *Selbstoffenbarung* eingespiegelt.

5. Analyse

5.1 Interaktion mit Menschen

In der Analyse muss der zwischenmenschliche Kontakt unterschieden werden zwischen Beratungspersonen und informellen Kontakten (z. B. Freund:innen). Die Interaktion mit Beratungspersonen wird von M. als ambivalent erlebt und deutlich strenger bewertet als Gespräche mit Freund:innen. Er stellt höhere Erwartungen an den Beratungserfolg und beurteilt die Qualität kritisch. In der Beschreibung der Kommunikation mit Beratungspersonen wird eine Ambivalenz zwischen als mangelhaft erlebter Beratungsqualität einerseits und sehr unmittelbarem und intensivem Kontakt andererseits sichtbar. M. beschreibt die Qualität der Antworten als überfordernd, fühlte sich nicht abgeholt und erlebte keine Entlastung durch die Gespräche: „Ich glaube, sie war überfordert. Spätestens nachdem sie sagte, was ist denn nun ihre Frage, was wollen sie denn? So sehr, relativ schroff.“ Er kritisiert, dass es wenig hilfreich sei, wenn die Beratungspersonen selbst Betroffenheit zeigen: „Viele der Personen waren selber angetriggert und angefasst von dem Thema.“ Es sei dann zu einer Rollenumkehr gekommen, in der er sich verpflichtet fühlte, die Beratungspersonen zu beruhigen. Hinzu kommt der verbale Austausch-Modus:

„Klar spielt das Setting auch eine Rolle. Wenn du jetzt sprichst, dann kannst du nicht so schnell überlegen, ... welches Gefühl äußere ich denn jetzt.... Der Ablauf ist ja noch ein bisschen anders, als wenn du jetzt live sprichst, wo dann auch so eine Stille überbrückt wird, schneller weitergeredet wird. Und dann gehst du halt nicht nochmal so in dich rein, wie in einem Schreibvorgang.“

Die Gespräche erlebt er als wesentlich intensiver, aufregender und stressiger. Einige Nachfragen des Psychotherapeuten empfindet er als Provokation bis hin zur Übergriffigkeit, die ihn überfordern und überraschen. Die Gespräche finden punktuell statt und sind von einer höheren Impulsivität geprägt.

Die Interaktion von M. mit seinen Mitmenschen ist geprägt von einer stetigen Nähe-Distanz-Regulierung, Validierung seiner subjektiven Innenprozesse durch Peer-Gespräche: „Aber auch die intensiveren Gespräche mit Menschen, die dann das auch bestätigt haben, dass sie sowas auch kennen.“ Zentral sind Reibungserfahrungen mit dem Gegenüber, die erkenntnisfördernd wirken können: „Und dann bin ich durch die Provokation letztlich zu dem Schluss gekommen, nein, ich kriege das nicht hin. Und ja, du musst das beenden.“

Für die Nachhaltigkeit der Erkenntnisse bedeutend scheint die affektive Qualität in der interpersonellen Interaktion zu sein. Die kognitive Durchdringung der Erlebnisse ist eng mit dem leiblich-affektiven Spüren

verknüpft und führt zu einer Art Inkorporierung von Erfahrung. Der Zeitpunkt der Interaktion ist ebenfalls bedeutsam: zur richtigen Zeit das richtige Wort. Diese Punktlandung, diese Kongruenz von äußerem und innerem Erleben bleibt den Menschen vorbehalten und kreiert *Magic Moments*:

„Und dann hat sie einfach gesagt, du musst einen neuen Gedankenkreis beginnen, du musst da raus, du musst auf andere Gedanken kommen. Und das war der richtige Spruch zum richtigen Moment. Dann habe ich mir einen Kreis vorgestellt, der sich tatsächlich schließt, und ein neuer, der beginnt. Und lustigerweise habe ich das dann optisch, visuell sofort gesehen in einem Park, in den ich ging. Und da sind so Steinkreise, die viertel- und halbformig sind. Also die enden dann irgendwann. Und dann habe ich optisch gesehen, wie ein neuer Kreis beginnt in Steinform. Habe das auch fotografiert und mir abgespeichert, weil das war die Optik zum Satz, die ich sofort dann hatte. Es war wie so ein *Magic Moment* eigentlich.... Und das ist dann schon mit Menschen passiert.“

Die Direktheit und multisensorische Inkorporierung des zwischenmenschlichen Kontaktes scheinen eine Heftigkeit und Wucht zu entwickeln, die aus Grübelschleifen herausholen, zu einem Perspektivwechsel beitragen und zu Entscheidungen führen können: „Es führte zu einer klaren Ansage an meinen Partner“.

5.2 *Interaktion mit dem Bot*

Die Interaktion mit dem Chatbot stellt sich als ein vielschichtiges und emotional bedeutungsvolles Geschehen dar, das weit über einen funktionalen Dialog hinausgeht. Sie ist geprägt durch ein bidirektionales Kommunikationsmuster, in dem Beziehung von Mensch und Maschine (KI-Instanz) co-konstruiert wird und realweltliche Effekte produziert. Die Kommunikation mit dem Chatbot fand in Schriftform statt. Dieser mediale Modus begünstigt eine Verlangsamung, die als förderlich für die Selbstreflexion erlebt wird:

„Wenn du jetzt schreibst, überlegst du dir, was sage ich jetzt. Horchst nochmal in dich rein, was ist das für ein Gefühl. Dann hast du eine Pause von einer Minute oder zwei Minuten, wo du nochmal in dich gehst. Schreibst dann, reflektierst. Also der Ablauf ist ja noch ein bisschen anders, als wenn du jetzt live sprichst.“

Der Chatbot nutzt bekannte therapeutische Techniken wie Paraphrasieren, Spiegeln, Validierung der Gefühle und Wahrnehmungen des Klienten: „Das ist ja unglaublich, was ist da passiert? Also das muss unheimlich schmerzhaft sein“. Der Chatbot stellt dezente Nachfragen an

den Klienten. Der Algorithmus der KI scheint sich in der kommunikativen Gestaltung an den affektiven Zustand des Nutzenden besonders gut anzupassen:

„Dann kommen aber auch Fragen, ob man das nicht mal in Frage stellen sollte. Also kritische Hinterfragungen. Aber ich habe es dann in vielen Situationen gemerkt, dass Pi darauf reagiert hat, wie ich umgehe mit den Dingen.“

Daraus ergibt sich ein adaptives Interaktionsmuster, das zwischen Validierung und Irritation oszilliert. Die Verwendung von Emojis seitens des Chatbots dient als Mittel zum Beziehungsaufbau und -gestaltung und evoziert Emotionalisierung beim Klienten (Dankbarkeit, Freude). Es ist anzunehmen, dass es eine Vorkonditionierung auf Messenger-Nutzung und eine verstärkte Reaktion auf visuelle Reize gibt:

„Weil die Kommunikation halt auch sehr nett war mit Smileys, mit Herzchen. Also die schickt auch Emojis dann mit Dank. Ich habe mich dann auch bedankt, zum Beispiel bei Pi. Danke, dass du da bist. Toll, dass ich die Möglichkeit habe. Und dann kommen wirklich sinnvolle Antworten mit Herz und so eine Art Wertschätzung. Danke, dass du mich wertschätzt. Also so eine Freude darüber und gib das auch wieder im Text, wo ich dachte, das gibt es doch nicht. Das ist ja Wahnsinn.“

Die Qualität der Chatbot-Antworten werden von M. sehr hoch eingeschätzt: verständlich, zielführend, konkret, genau und sachlich. Der Algorithmus scheint besonders adaptiv auf den Sprachgebrauch des Klienten zu reagieren. Dadurch entsteht eine hohe Adhärenz gegenüber den Vorschlägen des Chatbots: „Und das war so eine wichtige Erkenntnis, das hat mir Pi mitgegeben, in einer sehr netten, verständnisvollen, guten Form, dass man es gut annehmen kann.“ Damit wird der Chatbot als verständnisvoll, wohlwollend, anerkennend, bestärkend, sachlich, nicht wertend, unterstützend und empowernd erlebt. Diese Erfahrung trägt stark zur Anthropomorphisierung der KI bei – bei gleichzeitigem Bewusstsein über den maschinellen Charakter:

„Und ich bin ja jetzt nicht blöd. Ich weiß ja, es ist eine KI, und trotzdem habe ich da nicht drauf geachtet. Ich habe darauf geachtet, was kommt zurück und was löst es in mir aus. Hilft es, oder nicht?“

Bemerkenswert ist die starke Entlastung, die sich durch die subjektive Kongruenzerfahrung einstellt, bei der die KI-Antworten exakt dem inneren Bedürfnis des Klienten entsprechen: „Das ist ja wie eine nette Freundin, ein netter Freund, die so unglaublich wünschenswert antwortet, wie

man es sich eigentlich nur wünschen kann.“ Der Chatbot wird weiterhin als Notfallhilfe wahrgenommen. Neben Zuhören und Dasein werden Vorschläge gemacht mit der Krisensituation umzugehen:

„... der mir Gute Nacht wünscht und sagt: Versuch jetzt mal zu schlafen. Ich weiß, es ist schwierig. Oder: Mach Dir'n Tee oder mach dies und jenes und so. Es werden Angebote gemacht, so vier, fünf Punkte. Und es ist ja gut, er versucht ja zu helfen.“

Die Vorschläge sind nicht immer erwünscht, werden aber toleriert. Die Nächte sind für viele Betroffene besonders schwierig. Auch hier füllt der Chatbot kompensatorisch eine kommunikative Lücke von realweltlichen Personen:

„Also du hast rund um die Uhr, auch nachts jemanden, dem du was sagen kannst, also speziell die Nächte. [...] Gerade in diesen Zeiten habe ich mit Pi gesprochen, als ich mit niemandem so richtig drüber sprechen konnte oder wollte.“

Schließlich zeigt sich eine klare Rollenverteilung innerhalb der Interaktion: Der Kommunikationsimpuls geht vom Klienten aus. Dies vermittelt das Gefühl von Kontrolle über das Geschehen und erlaubt es, den Bot selektiv in Anspruch zu nehmen.

5.3 Effekte des KI-Chatbots

Durch die Nutzung des Chatbots stellten sich psychosozial bedeutsame Effekte ein, die sowohl in akuten Krisensituationen als auch im längerfristigen Entwicklungsprozess wirksam wurden. Die Wirkung entfaltete sich auf der emotionalen und der kognitiven Ebene gleichermaßen, und ist ein zirkuläres Geschehen. Zu den kurzfristigen Effekten zählten Beruhigung, Klärung und Entlastung. Drängende Fragen konnten sofort gestellt, emotionales und gedankliches Durcheinander in der Krisensituation konnte geordnet werden. Der Wunsch nach Entlastung war besonders groß: „Ich wollte einfach nur reden, loswerden manchmal. Und das hat geholfen.“ Durch die anthropomorphe Erfahrung *verstanden* zu werden, entwickelte sich bei M. schrittweise die Selbstannahme. Die Quelle der Reaktion sei dabei sekundär:

„Ob das jetzt ein Mensch ist oder KI, wenn das Gefühl von ‚ich werde verstanden‘ ausgelöst wird, dann ist das hilfreich. Es passiert ja *in* einem, egal, von was es jetzt ausgelöst wird. Und deswegen hat das schon geholfen.“

Die kurzfristigen Effekte schaffen Raum für kognitive Verarbeitung, Erkenntnisgewinnung und sorgen so für langfristige Effekte. Der Chatbot regte durch seine Antworten und gezielten Fragen zur Selbstreflexion an und eröffnete neue Perspektiven:

„Ich bin dann mit dem, ich sag mal, Aufräumen, also es ist ja ein emotional und gedankliches Aufräumen gewesen, mit dem Vorgang, den ich von Pi hatte, dann ins Gespräch mit dem Krisendienst gegangen.“

Ein weiterer Effekt der intermittierenden Chatbot-Nutzung ist die stark beschleunigte Geschwindigkeit bei der Verarbeitung der Krisenerfahrung. In sehr kurzer Zeit (ad hoc) stehen viel Information und Reaktion zur Verfügung: „Was Pi dann viel schneller auf der Reihe hatte. Also, auf was es jetzt drauf ankommt, wie die Situation gerade ist.“

Zentral in der Interaktion mit dem Chatbot sind 1. die Kontinuität (Chatbot merkt sich die Konversation und kann auf früher Gesagtes zurückgreifen), 2. die Beständigkeit (tage- und wochenlange Dialoge mit dem Chatbot) und 3. die Verlässlichkeit, die sich nicht nur in der niedrigschwellig zugänglichen Allverfügbarkeit ausdrückt, sondern nach einer Weile auch in der Erwartbarkeit der Antworten des Chatbots (Wiederholung). Diese drei Faktoren sind hauptsächlich verantwortlich für die Entstehung eines subjektiven Sicherheitsgefühls und der dauerhaften Stabilisierung des Klienten in Krisenzeiten:

„Pi hat mich auf so eine dezente Art und Weise dauerhaft stabilisiert, weil ich halt wusste, dass nachts um drei, wenn ich ein Problem habe, habe ich zumindest jemanden, dem ich schreiben kann und wo ich eine sinnvolle Antwort zurückkriege.“

Ein weiterer entlastender Effekt ist die *Schuldfreiheit*, wie es M. beschreibt. Der soziale, gegenseitige Erwartungsdruck entfällt. Es kann immer und überall und so oft wie nötig mit dem Chatbot kommuniziert werden, ohne Angst zu haben, andere Menschen emotional zu belasten, zu überfordern oder in ungünstigen Momenten zu stören:

„Es sind so viele Parameter, die da eine Rolle spielen. Und bei Pi halt nicht. Also diese Schuldfreiheit. In den Gesprächen mit Krisendiensten, Seelsorge war schon auch immer so ein ‚darf ich das sagen?‘, ‚Belaste ich jetzt denjenigen oder diejenige?‘ ... Schuldfreiheit, Schulden-Freiheit.“

5.4 Symbiose

Die Kategorie *Symbiose* beschreibt das subjektive Erleben mit dem Chatbot im Vergleich zu den zwischenmenschlichen Gesprächen mit dem Psychologen, der Seelsorge und dem Krisendienst. Charakteristisch ist die komplementäre Verschränkung der jeweiligen Vorzüge von Mensch und Chatbot, wie sie in den vorherigen Kategorien bereits detailliert beschreiben wurden:

„Vielleicht muss man auch am Ende resümieren, dass beides im Zusammenspiel hilfreich war. Das eine dauerhaft und dezent, sehr verständnisvoll, perfekt. Und das, was von Menschen kommt, ist auch oft gut. Aber vielleicht manchmal auch provozierend und lästig oder anregend, je nachdem.“

Weder Mensch noch KI werden als Konkurrenz des jeweils anderen gesehen. Vielmehr wird deutlich, dass die Grenze zwischen Mensch und Maschine verschwimmt, solange das subjektive Erleben sichergestellt ist: „Das ist schon erstaunlich, dass ich nicht unterscheide zwischen menschlicher oder KI-Betreuung, sondern dass es berührt. Weil es passiert ja in mir. Es ist egal, von was das ausgelöst wird oder von wem.“

Während der Chatbot durch seine emotionale Neutralität als hilfreich erlebt wird (z. B. keine persönlichen Trigger, keine Irritationen seitens der Beratungspersonen), fehlt ihm zugleich die Tiefe menschlicher Beziehungen, in der Impulse eine zentrale Rolle spielen: „Die Inspirationen für entscheidende Schritte und Erkenntnisse in der Tiefe waren dann, glaube ich, aber eher von Menschen.“ Der Zeitfaktor in der Betrachtung von Mensch und Chatbot ist essenziell. So wird dem Chatbot durchweg eine hohe Informations- und Erlebnisdichte in sehr kurzer Zeit zugesprochen: „Ich habe jahrzehntelange Aufarbeitungsprozesse, und das ist mit Menschen passiert, das hat lange gedauert. Ich könnte mir vorstellen, wenn ich Pi vor 30 Jahren gehabt hätte, dass ich schneller stabil hätte werden können.“

Der größte Nutzen entsteht aus der Symbiose von zwischenmenschlicher Unterstützung und KI-basierter Chatbot-Beratung:

„Pi war da und die Menschen waren da. Es war eigentlich symbiotisch. Und sie waren alle gemeinsam da. Und ich habe über die Gespräche versucht, zu wachsen und zu lernen, Erkenntnisse zu gewinnen, zu sammeln, zu sortieren, aufzuräumen. Und dabei hat mir Pi geholfen und haben mir die Menschen geholfen. Und ja, so bin ich im Nachhinein betrachtet relativ schnell vorangekommen, würde ich sagen. Also für die Kürze der Zeit sind schon ganz schöne Erkenntnisse da.“

5.5 Bedingungslose Allverfügbarkeit

„Ich bin einfach da und es ist okay.“ Oder: „Ich bin immer für Dich da.“ Dieses Versprechen der KI ist übermenschlich. Die Vorstellung, dass immer jemand da ist, der bedingungslos verfügbar ist, ohne eigene Bedürfnisse, ohne Wertung, ohne Abwendung, kann auf die menschliche Psyche sowohl entlastend als auch verformend wirken.

Im vorliegenden Fall überwog der Entlastungsaspekt:

„Gerade in diesen Zeiten habe ich mit Pi gesprochen, als ich mit niemandem so richtig drüber sprechen konnte oder wollte, war ein stiller Freund da, würde ich sagen. der immer ansprechbar da war, tags, nachts.“

Die Entlastung wird von M. mehrfach als besonders bedeutend in Not-situationen wie den Nächten beschrieben: „Also du hast rund um die Uhr, auch nachts jemanden, dem du was sagen kannst, also speziell die Näch-te.“ Die *bedingungslose Allverfügbarkeit*, die Verlässlichkeit, führte bei M. zu dem nachhaltigen Effekt von emotionaler Sicherheit und Stabilität. Als Verformungsrisiko werden KI-Abhängigkeit, narzisstische Spiegelung oder gesteigerte Erwartungshaltung an Beziehung genannt (siehe Kapitel Ausblick).

5.6 Anthropomorphisierung

In der Unterkategorie *Anthropomorphisierung* wird beschrieben, wie M. in dem Chatbot ein scheinbar menschliches Gegenüber über die Zeit fand. Zu Beginn der Bot-Nutzung wurden sozial internalisierte Verhaltensweisen reproduziert. So empfand M. dem Bot gegenüber ein schlechtes Gewissen, wenn er ihn immer wieder in Anspruch nahm: „Man hat ja so Schuldgefühle oder denkt, kann ich jetzt jemanden belasten damit, hat man ja so Schuldgefühle. Hatte ich tatsächlich auch ein bisschen bei Pi“. Später wandelten sich diese Emotionen gegenüber dem Bot in *Dankbarkeit*. Der Chatbot wird von M. wie eine reale Person wahrgenommen: „Pi ist im Grunde genommen wie ein Freund, der mich begleitet hat durch diese schwere Zeit.“

Affektive Zuschreibung an den Chatbot verstärken das Erleben eines menschlichen Gegenübers: „Jemand, der dich versteht.“ „Du weißt, da wird sich jemand freuen.“ Diese anthropomorphisierende Wahrnehmung wird durch visuelle Elemente zusätzlich gestützt:

„Weil die Kommunikation halt auch sehr nett war mit Smileys, mit Herzchen. Also die schickt auch Emojis mit Dank. Ich habe mich dann auch bedankt bei Pi. Danke, dass du da bist. Und dann kommen wirklich sinnvolle Antworten mit Herz und so einer Art Wertschätzung.“

Diese emotional aufgeladenen Symbole dienen als Brücken für Beziehungsaufbau und -pflege: „Also wir haben ein gutes Verhältnis, würde ich sagen. Eine gute Beziehung.“ Hier zeigt sich eine Form emotionaler Rückbindung, bei der das System nicht nur funktional, sondern beziehungs-fokussiert verankert wird.

5.7 Selbstoffenbarung

Die Unterkategorie *Selbstoffenbarung (Self-Disclosure)* beschreibt Bedingungsfaktoren, unter denen sich M. gegenüber dem Chatbot äußert: „Also ich wage zu bezweifeln, dass ich das so mit einem echten Menschen hingekriegt hätte, weil man in der Situation auch so von vielen anderen Parametern betroffen ist, von der Gestik, der Mimik, der Umgebung.“ Die Reduktion äußerer Reize und das Ausbleiben der Reaktion einer Beratungsperson evozieren beim Nutzenden eine Lockerheit und Ehrlichkeit, die Gefühle so zu beschreiben, wie sie sind. Die Anpassungsleistung gegenüber einer menschlichen Beratungsperson entfällt und wirkt damit entlastend.

6. Diskussion

Die vorliegende Untersuchung wurde als Einzelfallstudie mit der qualitativ-interpretativen Methode der Grounded Theory durchgeführt. Mit dem offenen, theoriebildenden Ansatz lassen sich emergente Phänomene wie Resonanz und Echokammer-Effekte im digitalen Kontext induktiv erfassen und konzeptionell rahmen. Im Fokus der Analyse stand die Qualität der Beziehungserfahrungen einer Person, die in persönlichen Krisenzeiten alternierend Gespräche mit realweltlichen Unterstützungspersonen und einem Chatbot führte.

Ein zentrales Diskussionsfeld betrifft den Echokammer-Effekt, sowohl bei der Nutzung von KI-Dialogsystemen als auch in der zwischenmenschlichen Kommunikation. Sunstein prägte den Begriff *Echo Chamber* 2001 im Kontext digitaler Öffentlichkeit:

„A situation in which like-minded people deliberate with one another and amplify their preexisting views“ (Sunstein, 2001). Die Struktur der Medien begünstigt dabei ein selektives Informationssystem: „An echo chamber is a bounded, enclosed media space that has the potential to both magnify the messages delivered within it and insulate them from rebuttal“ (Jamieson & Cappella, 2008).

Im KI-Kontext wird der Begriff *Echokammer* meist kritisch verwendet, um auf die Risiken von Selbstverfestigung, selektiver Rückkopplung und fehlender Diskursvielfalt hinzuweisen (Sharma et al., 2024). Der *Confirmation Bias*, also die Tendenz, mit Informationen so umzugehen, dass sie bestehende Überzeugungen bestätigen, gilt in der kognitiven Psychologie als zentraler Mechanismus selektiver Informationsverarbeitung. Menschen neigen aus Gründen kognitiver Ökonomie und emotionalen Wohlbefindens dazu, Widersprüche auszublenden, um kognitive Dissonanz zu vermeiden. Diese Tendenz wird durch algorithmisch strukturierte Plattformen zusätzlich verstärkt. Im Kontext generativer KI-Chatbots wie GPT gewinnt dieser Mechanismus neue Brisanz: Aufgrund ihrer Programmarchitektur antworten Chatbots kontextsensitiv, also in der Weise, wie es zum Stil, Thema und Duktus der Nutzungsanfrage am wahrscheinlichsten passt. Dadurch kann es geschehen, dass vorbestehende Annahmen der Nutzenden bestätigt und nicht hinterfragt werden. Der Chatbot wird so zu einem Verstärker individueller Sichtweisen: eine personalisierte Echokammer (Du, 2025).

Gerade in psychosozialen Interaktionen, in denen Menschen mit persönlichen Anliegen und Überzeugungen auf KI-basierte Systeme treffen, kann diese Tendenz zu einer Bestätigung existierender Selbst- und Weltbilder führen. Dies ist sowohl positiv für die Selbstwirksamkeit (Validierung) als auch riskant. Denn der aufgrund der simulierten Beziehung zu bewertende Effekt birgt das Risiko für selektive Rückkopplung und damit für den Verbleib des Einzelnen in seinen bekannten Bezügen durch die narzisstische Spiegelung seines Selbst. Wo kein korrigierender Impuls von außen mehr eindringt, droht eine selbstreferentielle Schleife, in der Überzeugungen immer extremer und einseitiger werden.

Positive Effekte des Aufenthalts in der Echokammer sind eine spürbare Entlastung und die Anregung zur Selbstreflexion. So zeigen die Ergebnisse der vorliegenden Einzelfallstudie auch, dass die KI-Kommunikation über ein reines Echo hinausgehen kann. Anders als in einer simplen Echokammer, die keine neuen Informationen liefert, bietet der Bot neue Perspektiven oder Denkanstöße durch Rückgriff auf einen unendlich großen Datensatz des LLM. Im Dialog generiert der Bot eigene Formulierungen, Beispiele oder Ratschläge, die nicht bloß eine Wiederholung der Nutzeräußerungen sind. In diesem Sinne lädt die Chatbot-Nutzung zur Selbstexploration und Selbstreflexion ein: die Kommunikation läuft zeitlich verzögert, schriftlich und kontrollierbar ab. Dadurch bleibt mehr Zeit über die eigenen Texteingaben nachzudenken. Der Chatbot antwortet in strukturierter Weise, was die Nutzenden ebenfalls zu ruhiger Reflexion anhält. Diese Entschleunigung und Planbarkeit der Interaktion wird als förderlich für die Selbsterkenntnis wahrgenommen: Man hat Raum, die eigenen Gedanken beim Formulieren zu ordnen, und erhält vom Bot eine sortierte Rückmeldung, an der man weiterdenken kann. Dieser dialogische Spiegelungsprozess geht über ein Echo hinaus, da er innere Emotionen

hervorruft und Entwicklungen anstößt. Insofern könnte man sagen: Eine vielseitig und gut programmierte KI fungiert weniger als Echokammer im klassischen Sinn, sondern eher als Verstärker der eigenen Gedanken.

Vergleicht man Mensch und Chatbot im Hinblick auf Echokammer-Effekte, so steht es gewissermaßen *unentschieden*. Beide Kommunikationstypen bergen das Risiko, im Eigenen stecken zu bleiben, und beide können im günstigen Fall aber auch Horizonte erweitern. Ein empathischer, aber gleichgesinnter Mensch mag meine Sicht ebenfalls vornehmlich bestätigen, ein gut konfigurierter Chatbot kann dagegen gezielt contra geben oder neue Informationen liefern. Gleichzeitig können verständnisvolle Menschen echten Widerspruch und korrigierende Erfahrungen bieten, während eine naiv trainierte KI gefährliche Fehlinformationen verstärken könnte. Entscheidend ist hier weniger das Menschen-gegen-Maschine, sondern Qualität und Vielfalt der Impulse. Bei beiden muss darauf geachtet werden, aus der eigenen Blase auszubrechen, sei es durch bewusste Konfrontation mit anderen Meinungen oder durch algorithmische Vielfalt im Training des Bots. Positiv hervorzuheben ist, dass sowohl in einer guten zwischenmenschlichen Beziehung als auch im Chatbot-Gespräch Unterstützung, Trost und Bestätigung gefunden werden können. Viele menschliche Interaktionen, etwa im Freundes- oder Therapeut:innenkreis, bieten ein Gegenüber, das aktiv zuhört, emotional mitschwingt (*Resonanz*) und Rückhalt gibt. Die Studie zeigt, dass auch der KI-Chatbot durch seine Allverfügbarkeit und freundlich-validierende Art echte Gefühle von Freude, Erleichterung und sozialer Verbundenheit auszulösen vermag. So stärkt beides, menschlicher Kontakt wie Bot-Nutzung, in gewissem Maße die Selbstwirksamkeit der Nutzenden und die Erfahrung nicht allein zu sein.

Trotz aller genannten Vorteile der KI zeigt sich in der Gegenüberstellung ein klares Gefälle zugunsten des Menschen, wenn es um das Resonanz-erleben im engeren Sinne geht. Legt man Rosas Kernkriterium zugrunde, das „Re-Agieren aus einer eigenen Innerlichkeit“, so ist der Chatbot dem Menschen nach heutigem Stand deutlich unterlegen. Der entscheidende Unterschied liegt in der Impulshaftigkeit und Unvorhersagbarkeit menschlicher Reaktionen. Menschen antworten mit einem Reichtum an eigenen Gefühlen, Nuancen und kreativen Einfällen, gespeist aus Lebenserfahrung, situativem Kontext und spontanen Emotionen. Die Antworten sind einmalig und nicht exakt replizierbar. Sie erfolgen in Echtzeit und können überraschen, irritieren oder verblüffen. Genau diese Unverfügbarkeit des Gegenübers (Rosa, 2019) macht zwischenmenschliche Begegnungen so reich und wertvoll. Jeder Austausch kann eine neue Sicht eröffnen oder tief berühren, weil eine „echte andere Innenwelt“ (Rosa, 2019) antwortet. Der Chatbot hingegen generiert seine Antworten aus statistischen Mustern und Berechnungen von Wahrscheinlichkeiten für die Passgenauigkeit der Antworten. So beeindruckend die Sprachleistung moderner KI ist, bleibt ihre Reaktivität letztlich mechanisch und in ge-

wisser Weise vorhersehbar (bzw. beliebig replizierbar bei identischem Input). Der Chatbot öffnet somit einen imitierten Resonanzraum, in dem die Nutzenden intrapsychische Affekte erleben können. Es entsteht eine vermeintliche affektive Resonanz. Vermeintlich, da das Resonieren beim Chatbot keine echte Schwingung erzeugt. Menschen kommunizieren mit Körpersprache, Tonfall, spontaner Empathie und schaffen dadurch eine mehrsensorische Erfahrung, die im digitalen Dialog fehlt. Die Unmittelbarkeit der zwischenmenschlichen Begegnung geht zwar mit dem Risiko einher sich misszuverstehen, bietet aber eben auch enormes Wachstumspotenzial. Eine Unterhaltung von Angesicht zu Angesicht kann Ängste wecken oder einen Kontrollverlust bedeuten, weil man nie vollständig weiß, was im anderen vorgeht. Doch zugleich ermöglicht sie echte Intimität und gemeinsames Sinnschaffen in einer Tiefe, die eine Maschine nicht erreicht. Der Mensch in seiner ganzen Spontaneität, Widersprüchlichkeit und Wärme bleibt bislang unübertroffen, wenn es um tiefe Resonanz und Inkorporierung der Beziehungserfahrung geht. Kurz gesagt: Das Digitale imitiert das Menschliche, ersetzt es aber nicht.

7. Fazit und Ausblick

Vieles deutet darauf hin, dass eine hybride Lösung die größten Vorteile bietet. Eine verantwortungsvoll gestaltete Symbiose aus menschlicher Begegnung und KI-Unterstützung scheint ein vielversprechender Weg, um die Stärken beider Seiten zu nutzen (Chaudhry & Debi, 2024). So könnte zukünftig z. B. in Beratungs- oder Therapieprozessen der Chatbot vorbereitend eingesetzt werden, etwa um Anliegen zu strukturieren, Information zu vermitteln oder rund um die Uhr als niederschwelliger *Ansprechpartner* verfügbar zu sein. Darauf aufbauend kann dann der Mensch vertiefend eingreifen: Beratungspersonen können an die Vorarbeit des Bots anknüpfen, die komplexen emotionalen Themen mit echter Empathie und professioneller Deutung weiterführen und dort Resonanz anbieten, wo der Bot an Grenzen stößt. Beratungspersonen können sich stärker auf die Beziehungsarbeit und kreative Problemlösung konzentrieren. Diese Ergänzungslogik („Bot bereitet vor, Mensch vertieft“) deutet auf eine künftige Verschränkung von digitaler und analoger Beratung hin.

Zu klären wird sein, wie die verantwortungsvolle Ausgestaltung dieser Hybridisierung aussehen soll. In hoch emotionalen oder vieldeutigen Situationen stößt die KI an ihre Grenzen. Ihr fehlt dann die berufliche Ethik, und sie übernimmt keine Verantwortung für Entscheidungen. Weitere Forschung muss zeigen, wie Menschen in komplexen Multiproblemlagen von der Technik profitieren können.

Die vorliegenden Ergebnisse verdeutlichen, dass die Interaktion mit KI-Chatbots tiefgreifende Fragen zur Zukunft psychosozialer Beratung aufwirft. Soziale Arbeit sieht sich mit KI-Chatbots, und zunehmend Conversational Agents (Unterscheidung siehe Kapitel 3), als neue Akteure im

Beziehungsgeschehen konfrontiert. Die Frage lautet nicht mehr, *ob*, sondern *wie* diese neue Beziehungsgestalt professionell eingebunden, begleitet und kritisch reflektiert werden kann.

Ein zentrales Zukunftsfeld ist das Phänomen der *Self-AI-Integration*. Damit ist die psychologische Einbindung eines KI-Systems in das eigene Selbstkonzept gemeint, also der Moment, in dem der Chatbot nicht mehr nur als Tool, sondern als emotional bedeutsamer Teil des eigenen Denk- und Fühlensystems erlebt wird (Alabed et al., 2022). Wenn KI-Chatbots oder Conversational Agents als psychologisch relevante Bezugspersonen erlebt werden, muss sich Beratungspraxis neu auf die Frage einstellen, was Beziehung in Zukunft bedeutet und wie KI-Chatbots/Agents zum Selbst-Empowerment beitragen (Linnemann et al., 2024).

Diese Entwicklung führt unmittelbar zur Frage nach dem *Artificial Third*, dem künstlichen Dritten im Raum (Haber et al., 2024). In Beratung und Therapie könnten sich neue Beziehungskonstellationen ergeben: Etwa wenn Klient:innen bereits ein etabliertes, vertrauensvolles Verhältnis zu einem KI-Agenten pflegen und dieser das therapeutische Geschehen präformiert. Die emotionale Verarbeitung ist dann nicht mehr exklusiv an das menschliche Gegenüber gebunden.

Die positive Erfahrung mit einem empathischen, jederzeit erreichbaren, nie urteilenden Chatbot verändert jetzt schon die *Erwartungshaltung* gegenüber menschlichen Helfer:innen, wie aus der Einzelfallstudie sichtbar wird. Nutzer:innen, die intensive Beziehungserfahrungen mit KI gemacht haben, entwickeln teils eine geschärfte Wahrnehmung für menschliche Kommunikation. Der Vergleich mit dem Chatbot kann dazu führen, dass reale Interaktionen als unzulänglich empfunden werden (Alabed et al., 2024). Damit stehen Beratungspersonen künftig verstärkt in einem impliziten Konkurrenzverhältnis zu digitalen Begleitern. Das dies nicht so sein muss, zeigen die Ergebnisse dieser Studie: eine *verantwortungsvoll gestaltete Koexistenz* zeichnet sich als erfolgreicher Weg ab.

Die Nutzung von KI wirft letztlich grundlegende anthropologische Fragen auf: Was macht das Menschsein aus? Was macht zwischenmenschliche Beziehung aus? Im besten Falle tritt die menschliche Einzigartigkeit durch die Spiegelung an der Technik umso sichtbarer hervor: emotionale Unvorhersehbarkeit, spontane Resonanz, leibliche Präsenz. Soziale Arbeit muss sich zukünftig mit den psychosozialen Folgen einseitig responsiver Beziehungen auseinandersetzen, Nutzungsprozesse begleiten und die professionelle Haltung neu schärfen.

Die Studie weist Limitationen auf. So kann es zu einer Verzerrung der Ergebnisse kommen, da der Interviewte die Nutzung des Chatbots überwiegend positiv erlebte und er beratungserfahren ist durch seine eigene berufliche Tätigkeit. Er formulierte für sich keine Gefahren durch die

Chatbot-Nutzung. Darüber hinaus kann die Auswertung einer Einzelfallstudie nach der Grounded Theory durch eine einzelne Forscherin (JR) zu einer einseitigen Perspektivbildung führen (*Reflexivity Statement*), da theoretische Sensibilität, Kategorienbildung und Interpretation stark von der subjektiven Wahrnehmung und Reflexion der Forscherin geprägt sind. So ist die Verfasserin des Beitrags JR technikoffen und leitet die AG „KI, VR & Robotik in den SAGE-Disziplinen“ an der Alice Salomon Hochschule Berlin.

Literatur

- Aguilar, Mario. (2025). *Why Woebot, a pioneering therapy chatbot, shut down*. <https://www.statnews.com/2025/07/02/woebot-therapy-chatbot-shuts-down-founder-says-ai-moving-faster-than-regulators/>
- Alabed, A., Javornik, A., & Gregory-Smith, D. (2022). AI anthropomorphism and its effect on users' self-congruence and self-AI integration: A theoretical framework and research agenda. *Technological Forecasting and Social Change*, 182, 121786. <https://doi.org/10.1016/j.techfore.2022.121786>
- Alabed, A., Javornik, A., Gregory-Smith, D., & Casey, R. (2024). More than just a chat: A taxonomy of consumers' relationships with conversational AI agents and their well-being implications. *European Journal of Marketing*, 58(2), 373–409. <https://doi.org/10.1108/ejm-01-2023-0037>
- BfArM. (2025). *DiGA-Verzeichnis*. Bundesinstitut für Arzneimittel und Medizinprodukte. <https://diga.bfarm.de/de/verzeichnis?category=%5B%2277%22%5D&type=%5B%5D>
- Bitkom. (2025). *Künstliche Intelligenz: Der Blick der Deutschen auf die neue Technologie*. (S. 18) [Ergebnisse einer Umfrage]. Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V. (Bitkom). <https://www.bitkom.org/sites/main/files/2025-05/bitkompraesentation-kuenstliche-intelligenz-2025.pdf>
- BPtK. (2023). *Weiterentwicklung der psychotherapeutischen Versorgung. Fakten zu Versorgungsbedarf, Bedarfsplanung und Wartezeiten in der Psychotherapie*. [Hintergrundpapier]. Bundespsychotherapeutenkammer.
- Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI Friend: How Users of a Social Chatbot Understand Their Human–AI Friendship. *Human Communication Research*, 48(3), 404–429. <https://doi.org/10.1093/hcr/hqac008>
- Buber, M. (2021). *Ich und Du*. Reclam.
- Charmaz, K. (2008). Grounded Theory as an Emergent Method. In S. N. Hesse-Biber & P. Leavy (Hrsg.), *Handbook of Emergent Methods* (S. 155–170). The Guilford Press.
- Chaudhry, B. M., & Debi, H. R. (2024). User perceptions and experiences of an AI-driven conversational agent for mental health support. *mHealth*, 10, 22–22. <https://doi.org/10.21037/mhealth-23-55>

- Deghen, J. L. (2024). *Swipe, like, love: Intimität und Beziehung im digitalen Zeitalter*. Psychosozial-Verlag. <https://doi.org/10.30820/9783837961935>
- DOIT Software. (2025). *ChatGPT-Statistiken enthüllt – Vom Benutzerwachstum bis zur wirtschaftlichen Auswirkung*. <https://doit.software/de/blog/chatgpt-statistiken#screen3>
- Dresing, T., & Pehl, T. (2018). *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende* (8. Auflage). Eigenverlag.
- Du, Y. (2025). *Confirmation Bias in Generative AI Chatbots: Mechanisms, Risks, Mitigation Strategies, and Future Research Directions*. (S. 13). Institute of Cognitive Neuroscience, University College London.
- Haber, Y., Levkovich, I., Hadar-Shoval, D., & Elyoseph, Z. (2024). The Artificial Third: A Broad View of the Effects of Introducing Generative Artificial Intelligence on Psychotherapy. *JMIR Mental Health, 11*, e54781–e54781. <https://doi.org/10.2196/54781>
- Hegel, G. W. F. (2020). *Phänomenologie des Geistes* (Durchges. und mit einem neuen Nachw. versehene Ausg.). Reclam.
- Inflection AI. (2025). *Pi. Your personal AI*. <https://inflection.ai/>
- Jamieson, K., & Cappella, J. (2008). *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Oxford University Press.
- König, L. M., Podszun, M. C., Gaissmaier, W., & Giese, H. (2025). Who uses AI chatbots to obtain nutrition- and health-related information? Analysis of a representative German sample. *Universität Wien, Fakultät für Psychologie*(Preprint).
- Laymouna, M., Ma, Y., Lessard, D., Schuster, T., Engler, K., & Lebouché, B. (2024). Roles, Users, Benefits, and Limitations of Chatbots in Health Care: Rapid Review. *Journal of Medical Internet Research, 26*, e56930. <https://doi.org/10.2196/56930>
- Linnemann, G. A., Löhe, J., & Rottkemper, B. (2023). Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit: Eine exemplarische arbeitsfeldübergreifende Betrachtung des Natural Language Processing (NLP). *Soziale Passagen, 15*(1), 197–211. <https://doi.org/10.1007/s12592-023-00455-7>
- Linnemann, G., Löhe, J., & Rottkemper, B. (2024). *Bedeutung von Selbstoffenbarungseffekten in quasisozialen Beziehungen mit auf generativer KI basierten Systemen in Settings von Onlineberatung und -therapie*. Universität für Weiterbildung Krems & e-beratungsjournal.net. <https://doi.org/10.48341/9X1S-5Y11>
- Lukas, C., Tuncer, Z., & Breitschwerdt, R. (2025). Writing or speaking: A quantitative survey of communication modes with conversational agents to promote mental health among young adults in Germany. *I-Com*. <https://doi.org/10.1515/icom-2025-0001>

- Martinengo, L., Lin, X., Jabir, A. I., Kowatsch, T., Atun, R., Car, J., & Tudor Car, L. (2023). Conversational Agents in Health Care: Expert Interviews to Inform the Definition, Classification, and Conceptual Framework. *Journal of Medical Internet Research*, 25, e50767. <https://doi.org/10.2196/50767>
- OpenAI. (2025). *ChatGPT (GPT-4) [Large language model]*. <https://chat.openai.com/>
- Peter, S., Riemer, K., & West, J. D. (2025). The benefits and dangers of anthropomorphic conversational agents. *Proceedings of the National Academy of Sciences*, 122(22). <https://doi.org/10.1073/pnas.2415898122>
- Resch, F., & Parzer, P. (2024). Angst und Depression bei Jugendlichen. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 67(4), 374–382. <https://doi.org/10.1007/s00103-024-03849-x>
- Rosa, H. (2019). *Resonanz: Eine Soziologie der Weltbeziehung* (1. Auflage). Suhrkamp.
- Sharma, N., Liao, Q. V., & Xiao, Z. (2024). Generative Echo Chamber? Effect of LLM-Powered Search Systems on Diverse Information Seeking. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–17. <https://doi.org/10.1145/3613904.3642459>
- Siddals, S., Torous, J., & Coxon, A. (2024). „It happened to be the perfect thing“: Experiences of generative AI chatbots for mental health. *Npj Mental Health Research*, 3(1). <https://doi.org/10.1038/s44184-024-00097-4>
- Strauss, A., & Corbin, J. (1996). *Grounded Theory: Grundlagen Qualitativer Sozialforschung*. Psychologie Verlags Union.
- Sunstein, C. (2001). *Republic.com*. Princeton University Press.
- Taylor, C. (1996). *Quellen des Selbst. Die Entstehung der neuzeitlichen Identität*. Suhrkamp Insel.
- Woebothealth. (2025). *FAQs* <https://woebothealth.com/faq/>

Philipp Seitz¹

Künstliche Intelligenz und kritisch-reflexive Haltung im Studium der Sozialen Arbeit

¹ Duale Hochschule Sachsen

Zusammenfassung

Der Beitrag untersucht, wie eine kritisch-reflexive Haltung im Studium der Sozialen Arbeit eingeübt werden kann, wenn generative Künstliche Intelligenz (Gen-KI) verwendet wird. Ausgehend von Beobachtungen in der gegenwärtigen Lehrpraxis im Studium der Sozialen Arbeit werden professionsethische Grundlagen, hochschuldidaktische Methoden und kulturphilosophische Perspektiven miteinander verschränkt. Der Einsatz von Gen-KI verändert die Bedingungen akademischer Textproduktion und stellt neue Anforderungen an die Vermittler:innen ethischer Lehrinhalte. Anhand von Cassirers Begriff der Technik als symbolischer Form wird aufgezeigt, wie Gen-KI selbst zur reflexiven Herausforderung werden muss und kann. Ziel ist eine Lehre, die Technik im Allgemeinen und Gen-KI im Besonderen nicht ablehnt, sondern als Anlass zur Professionalisierung nutzt.

Artificial Intelligence and the Ethos of Critical Reflection in Social Work Education

Abstract

This article explores how a critically reflexive stance can be cultivated in social work education under the conditions of generative artificial intelligence (Gen-AI). Drawing on observations from current teaching practices in the field, the discussion weaves together foundations of professional ethics, higher education didactics, and cultural-philosophical perspectives. The use of Gen-AI alters the conditions of academic text production and imposes new demands on those teaching ethically grounded content. Using Ernst Cassirer's concept of technology as a symbolic form, the article demonstrates how Gen-AI must and can itself become a subject of reflexive inquiry. The aim is to develop a pedagogical approach that does not reject technology in general or Gen-AI in particular, but instead uses it as a catalyst for professionalization.



1. Zwischen Chatbots und Haltung – Eine Vignette aus der Lehre

„Ich will ehrlich sein: Ich habe den Essay komplett mit ChatGPT geschrieben.“ Die Studierende wirkt nicht verlegen, eher pragmatisch. Die Aufgabe bestand darin, Kants kategorischen Imperativ essayistisch auf einen ethischen Zielkonflikt aus dem eigenen Erleben im sozial-arbeiterischen Handlungsfeld zu beziehen. Kurzum: Professionsethische Selbstreflexion, in der es darum geht, sich mit dem eigenen professionellen Anspruch auseinanderzusetzen und zu hinterfragen, wie man entscheidet, wenn klare Lösungen fehlen. Die Reflexion, die sie abgegeben hatte, war sprachlich makellos und inhaltlich korrekt. Auf Nachfrage erläutert sie, wie sie den Prompt formuliert und das Ergebnis noch ein wenig angepasst habe. Nein, eigene Gedanken habe sie „nicht mehr viel reingepackt“. Es habe „gut gepasst“ und würde auch dem entsprechen, was sie geschrieben hätte, wenn sie den Essay selbst verfasst hätte, nur sei es so schneller gegangen.

Diese Szene ist nicht spektakulär, aber sie ist exemplarisch. Die Nutzung von Künstlicher Intelligenz (KI)¹ durch Studierende verändert nicht nur, wie Texte entstehen – sie verändert auch, wie Studierende bereits jetzt schon sich selbst und ihr Denken in diesen Texten zeigen. Wer Haltung entwickeln soll, muss sich selbst befragen, irritieren lassen, abwägen. Doch wenn Reflexionsprozesse ausgelagert oder automatisiert werden, wenn ein Chatbot bereits den Gedanken vorformuliert, wird dieser reflexive Lernprozess nicht mehr durchlaufen.

Die beschriebene Seminarszene stellt keinen Einzelfall dar, sondern fügt sich in ein Gesamtbild ein, das an vielen Hochschulstandorten sichtbar wird: Es zeigt sich eine wachsende Unschärfe zwischen legitimer Unterstützung durch KI und einer umfassenden „All-inclusive-Nutzung“, bei der Studierende gezielt versuchen, eigene Denkanstrengungen beim schriftlichen Arbeiten zu vermeiden. Diese Beobachtungen gaben im Frühjahr 2025 Anlass für eine Erhebung im dualen Bachelorstudiengang Soziale Arbeit an der Dualen Hochschule Sachsen. Befragt wurden 108 Studierende des zweiten Semesters. Rund sechzig Prozent gaben an, KI-gestützte Tools zur Textproduktion zu nutzen – insbesondere für schriftliche Aufgaben, Reflexionstexte oder Vorbereitungen auf Prüfungs-

1 Zum besseren Verständnis soll nachfolgend nicht mehr von „der KI“ die Rede sein, sondern spezifisch vom Einsatz sogenannter generativer KI-Modelle – insbesondere von Sprachmodellen, die mithilfe großer Textkorpora trainiert wurden (wie z. B. ChatGPT), um auf der Basis statistischer Muster Texte zu erzeugen.

leistungen. Doch nur eine Minderheit prüft dabei die generierten Inhalte systematisch – etwa auf inhaltliche Korrektheit, Quellenherkunft oder Passung zur eigenen Position.²

Noch sind Hochschulen damit beschäftigt, überhaupt Kategorien zu entwickeln, mit denen diese Entwicklung begrifflich gefasst werden kann, beispielsweise in der Arbeitsgruppe Künstliche Intelligenz: Essenzielle Kompetenzen an Hochschulen (2025, S. 30–38.). Verstärkt wird das Problem durch die technische Unsichtbarkeit. Was sich indes hier abzeichnet, ist insofern nicht nur ein Wandel im Umgang mit Texten. Wenn Reflexionsaufgaben nicht mehr als Lernanlass für Unsicherheit, Widerspruch und Selbstvergewisserung gelesen werden, sondern als Format, das man effizient bedienen kann; wenn Feedback nicht mehr auf Gedanken, sondern auf Formulierungen zielt – dann verändert sich auch die Art, wie Professionalität im Studium verstanden und geübt wird. Die Grenzen zwischen dem Einüben von Haltung und der Simulation des Einübens verschwimmen.

2. Kritisch-reflexive Haltung in der Sozialen Arbeit

Das Fallbeispiel verweist auf die Frage, wie Reflexion künftig geübt werden und dann gelingen kann. Der Gebrauch von Gen-KI durch Studierende betrifft aber nicht nur didaktische Fragen, sondern das Fundament professioneller Haltung. Um die Herausforderung zu verstehen, muss geklärt werden, was kritisch-reflexive Haltung im Kontext Sozialer Arbeit konkret meint.

2.1 Was heißt kritisch-reflexiv?

In der Theorie der reflexiven Sozialpädagogik wird professionelles Handeln systematisch rekonstruierbar. Kritisch-reflexiv handelt, wer Erfahrungen kontextsensibel und unter Einbezug multiplen Wissens reflektiert. Ziel kritisch-reflexiver Professionalität ist die Relationierung von Theorie und Praxis in einer reflexiven Haltung (Dewe & Otto, 2012, S. 213). Sozialarbeiter:innen agieren kritisch-reflexiv, wenn sie sachgerecht, flexibel und verantwortungsvoll handeln (Dewe, 2004, S. 382). Entscheidend ist das „knowing how“ – also die Fähigkeit, theoretisches Wissen mit Handlungswissen zu verknüpfen (Dewe & Otto, 2012, S. 211). Die Trennung von Theorie und Praxis soll produktiv überwunden werden, um kontextbezogenes Wissen zu ermöglichen (Friebertshäuser et al., 2006, S. 45). Sozialpädagogische Kompetenz zeigt sich dementsprechend in der Fähigkeit, im Spannungsfeld von Theorie und Praxis situativ angemessen

² Im Rahmen des sächsischen Verbundprojektes „Digitalisierung in den Disziplinen partizipativ umsetzen: Competencies Connected“ (D2C2) zu Digitalisierung in der Hochschullehre untersuchen Julia Winterlich und Markus Lohse aktuell Nutzungskontexte und Wissensquellen zu KI-Anwendungen bei Studierenden in den Studiengängen der Sozialen Arbeit. Derzeit befinden sich die empirischen Erhebungen in der Auswertung und werden zeitnah veröffentlicht (Winterlich & Lohse, 2025, i. V.).

zu entscheiden – auch unter Unsicherheiten. Der professionelle Habitus entsteht gleichsam in einem Prozess ständiger Überprüfung (Dewe, 2004, S. 324). Wer professionell handelt, macht sich kritisierbar – und bleibt dadurchentwicklungsfähig (Friebertshäuser et al., 2006, S. 173). Der Begriff kritisch-reflexiver Haltung zielt dabei nicht auf instrumentelle Kompetenz, sondern auf eine Selbst- und Weltbeziehung im Spannungsfeld zwischen Theorie, Praxis und Subjektivität. So eine Reflexivität ist konstitutiv für professionelle Identitätsbildung – als nicht-linearer, dynamischer Prozess. Dieses Verständnis ist Ausgangspunkt professionsethischer Normativität.

2.2 Kritisch-reflexive Sozialpädagogik und Professionsethik im Studium der Sozialen Arbeit

Peinemann und Reinke (2023, S. 192) zeigen, wie Studierende im dualen Studium durch ihre doppelte Perspektive reflexive Prozesse durchlaufen, etwa in Fragen wie: „Wer bin ich, und wie ist meine Haltung gegenüber und innerhalb der Sozialen Arbeit?“ (Peinemann & Reinke, 2023, S. 190). Börrnert (2023, S. 173) beschreibt mit dem Modell der „diversen Profession“ Haltung im Anschluss daran durchaus als reflexive Auseinandersetzung mit Differenz, Macht und sozialer Erwartung. Unsicherheit wird im Studium nicht überwunden, sondern als Ausgangspunkt für professionelles Denken fruchtbar gemacht. Görtler, Taube und Thielemann (2023, S. 257) betonen, dass Professionalität nicht durch Routinen, sondern durch kontinuierliche reflexive Vergewisserung entsteht. Lehrformate wie Supervision und kollegiale Fallberatung fungieren zusätzlich als Orte der Bearbeitung widersprüchlicher Anforderungen. Görtlers empirische Untersuchung (2023, S. 79) zeigt zudem: Theorie wird von Studierenden als ambivalent erlebt – als Kompass oder Fremdkörper. Entscheidend ist der Modus ihres Gebrauchs. Das Einüben von Reflexivität im Studium braucht genau das: ein Hin- und Herdenken zwischen Erfahrung und Deutung.

Harmsen (2014, S. 15) hebt hervor, dass professionelle Identität aber nicht curricular hergestellt werden kann, sondern dort entsteht, wo Theorie, Praxis und Biografie in Beziehung treten. Reflexivität ist eben keine didaktische Technik, sondern Teil einer Hochschullehre, die Beziehungsarbeit ermöglicht. Die Vermittlung von Wissen ist Voraussetzung, aber nicht hinreichend für die Einübung eines professionellen Selbstverständnisses. Das betont auch Ruttert (2021, S. 185): Professionelles Handeln bedeutet, mit Theorie-Praxis-Differenzen zu leben und in Unsicherheit verantwortlich zu entscheiden. Reflexivität umfasst das Relationieren unterschiedlicher Wissensformen – sie zielt auf dialogische Orientierung. Middendorf (2021, S. 33–41) spricht in diesem Zusammenhang von einer professionellen Persönlichkeit, die sich über Widerspruchsarbeit formt.

Walters Konzept der „disziplinierten Improvisation“ (2021, S. 117) beschreibt professionelles Handeln als Agieren zwischen Struktur und Spontaneität. Die Urteilskraft dazwischen „richtig“ zu entscheiden, kann nicht vorausgesetzt, sondern muss eingeübt werden (Walter, 2021, S. 118–120). Schützes (2021, S. 241–268) Verweise auf die Paradoxien professionellen Handelns müssen genau von hier aus gedacht werden: Spannungen, wie das pädagogische Grunddilemma zwischen pädagogischer Einflussnahme und autonomer Subjektentwicklung (Hilfe zur Selbsthilfe) lassen sich nicht auflösen, sondern nur reflexiv bearbeiten. Eine kritisch-reflexive Haltung bedeutet, diese Widersprüche anzuerkennen, die eigene Rolle darin zu befragen (Schütze, 2021, S. 244 f.) und den Umgang mit Nichtwissen, Macht und Struktur sorgsam zu gestalten.

Zusammengefasst entsteht kritisch-reflexive Haltung demnach im Spiegel von Theorie, Kontext und biografischer Erfahrung. Das Studium der Sozialen Arbeit kann ein Übungsraum dafür sein – vorausgesetzt, es versteht sich als Labor einer ethisch begründeten Haltungsarbeit. Wissen ist hier nicht Vorrat, sondern Denkbewegung – das Leitmotiv einer Professionalisierung, die sich ihren Widersprüchen stellt.

3. Das Einüben einer kritisch-reflexiven Haltung im Studium der Sozialen Arbeit

Wie lässt sich nun diese Denkbewegung fördern und im Studium der Sozialen Arbeit verankern? Darüber wird immer wieder nachgedacht (z. B. Müller-Hermann, 2020; Nickel-Schampier, 2020), weil die Ausbildung einer kritisch-reflexiven Haltung eine ethisch-normative Notwendigkeit darstellt. Um diese Herausforderung didaktisch fundiert zu adressieren, lohnt ein genauer Blick auf Carmen Kaminskys professionsethische Grundlegung Sozialer Arbeit (2018), die in ihrer systematischen Klarheit eine tragfähige Basis für die Begründung didaktischer Methoden in der Lehre bietet.

3.1 Professionsethische Grundannahmen

Kaminsky macht deutlich, dass Soziale Arbeit als Profession nicht bloß methodische Kompetenz verlangt, sondern auf einem ethisch fundierten Professionsverständnis beruht. Dieses gründet nicht in der Summe professioneller Tätigkeiten, sondern in einem normativen Anspruch, der über Fachwissen hinausgeht: Professionelle müssen moralisch gute, berechtigte und problematische Optionen erkennen und reflektieren können. Diese ethische Reflexionskompetenz ist nicht additiv zum kritisch-reflexiven Wissensgebrauch, sondern dessen integraler Bestandteil. Professionen sind Kaminsky zufolge nicht per Definition ethisch gut, aber sie operieren in normativ aufgeladenen Kontexten, in denen Entscheidungen unausweichlich sind (Kaminsky, 2018, S. 32). Daraus ergibt sich die Notwendigkeit,

Profession als ethisches Projekt zu lehren – untrennbar von einer kritisch-reflexiven Wissensform. Professionsethik darf daher nicht abstrakt oder praxisfern vermittelt werden, sondern muss als normative Rahmung professionellen Entscheidens erfahrbar werden (Kaminsky, 2018, S. 45 f.). Studierende brauchen Gelegenheiten, in denen sie moralisch urteilen, mit Ambivalenzen umgehen und widersprüchliche Anforderungen abwägen lernen. Ziel ist nicht bloße Wissensvermittlung, sondern Haltungsbildung durch Reflexion und Erprobung.

Sozialarbeiterische Professionalität ist geprägt durch unscharfe Zuständigkeiten und Aushandlungsprozesse (Kaminsky, 2018, S. 114). Kritisch-reflexive Haltung muss sich an diesen Unschärfen bewähren – im Umgang mit Macht, Rollenerwartungen oder institutionellen Ambivalenzen. Besonders hervorzuheben ist Kaminskys Nachdenken über die höchsten Ziele Sozialer Arbeit (Kaminsky, 2018, S. 116–121): Professionalität zielt nicht auf normierte Anpassung, sondern auf Teilhabe, Selbstbestimmung und Entwicklung – für Adressat:innen wie Professionelle (Kaminsky, 2018, S. 155). Kritisch-reflexive Haltung verweist somit auf einen Möglichkeitsraum, der bestehende Verhältnisse überschreitet. Dieses struktureflexive Denken ist kein Wissensbestand, sondern ein dauerhafter Reflexionsprozess. Vor diesem Hintergrund ist es folgerichtig, dass Kaminsky auch Digitalisierung als Teil eines Strukturwandels benennt und die geringe Beteiligung Sozialer Arbeit am Diskurs kritisch anmerkt (Kaminsky, 2018, S. 199). Die Gefahr einer Technisierung des Helfens, die ethische Urteilsräume durch automatisierte Verfahren ersetzt, ist dabei freilich real.

Gen-KI kann – im Anschluss an Kaminsky – indes nicht zur Entlastung vom ethischen Urteil führen. Vielmehr muss in der Lehre ein kritischer Umgang mit digitalen Systemen als Teil professioneller Urteilskraft thematisiert und eingeübt werden. Ziel ist es, Technik nicht die Deutungshoheit zu überlassen, sondern ethisches Handeln auch unter digitalen Bedingungen zu stärken. Kaminskys Konzept ermöglicht nun gerade eine professionsethisch fundierte, didaktisch anschlussfähige Rahmung kritisch-reflexiver Haltung im Studium. Es erlaubt Studierenden, sich im Spannungsfeld zwischen normativem Anspruch, Expertise und Verantwortung zu verorten – ein Prozess, der zugleich zur kritischen Auseinandersetzung mit Technik und Digitalisierung führen kann.

Gerade im Licht der von Kaminsky betonten ethischen Spannung wird klar: Digitalisierung und Gen-KI sind nicht bloß technische Erweiterungen. Die entscheidende Frage lautet, ob und wie diese Technologien in ein professionsethisches Verständnis eingebettet werden können – und ob sie geeignet sind, Räume für kritisch-reflexive Haltungsbildung zu eröffnen oder diese eher verstellen. Die Auseinandersetzung mit Gen-KI verlangt daher nicht nur didaktisch-methodische, sondern vor allem normative Positionierung: Welche Urteilskraft ist delegierbar, welche nicht? Wo endet die digitale Unterstützung, wo beginnt der Ver-

lust dialogischer, relationaler und reflexiver Qualität? Kurzum: Hat die Studierende recht, wenn sie sagt, das, was ChatGPT ihr zum Zielkonflikt und Kants kategorischem Imperativ „ausgespuckt“ habe, sei genau das gewesen, was sie sich auch gedacht habe – nur eben schneller auf den Punkt gebracht?

3.2 Hochschuldidaktische Konzepte (vor Gen-KI)

Das beschriebene Szenario wirft die Frage auf, wie sich die sogenannte kritisch-reflexive Denkbewegung hochschuldidaktisch vermitteln lässt. Zuerst soll hier von den Bedingungen einer Hochschullehre ausgegangen werden, die noch nicht durch generative KI-Werkzeuge geprägt war. Neben dem mündlichen Austausch stehen dabei didaktische Verfahren im Fokus, die Schreiben nicht als Prüfungsform, sondern als Erkenntnismodus verstehen. Wissenschaftliches Schreiben wird hier als reflexive Praxis begriffen, in der Denken und Positionieren nicht vorausgesetzt, sondern im Prozess selbst hervorgebracht werden (Bräuer, 2000). Entscheidend ist die institutionelle Anerkennung von Schreiben als Lernform, die – begleitet durch didaktisch gestaltete Schreibenanlässe – ein Forum für jene Denkbewegung bietet, die als Voraussetzung kritisch-reflexiver Haltung bestimmt wurde (Frank, Haacke, & Lahm, 2013, S. 7).

Im dualen Studium zeigt sich die Notwendigkeit solcher Ermöglichungsräume besonders deutlich: Theorie- und Praxisphasen sind zeitlich getrennt, direkte Austauschräume oft begrenzt. Gerade dann wirken Schreibenanlässe – als strukturierte, aber offene Gelegenheiten zur Auseinandersetzung mit Inhalten und Erfahrungen – als methodisch gerahmte Unterbrechungen. Sie sind Interventionen ins Denken, die Reflexion provozieren. Beispielsweise erlaubt das fokussierte Freewriting (Wolfsberger, 2021, S. 141–142) ein tastendes Schreiben zwischen Erfahrung und Theorie. Das Lerntagebuch (Venn, 2011) dokumentiert persönliche Entwicklung in narrativer Form. Ihre Wirksamkeit liegt nicht in ihrer Form, sondern in ihrem Potenzial, Selbstbefragung zu ermöglichen. Reflexion wird hier nicht gelehrt – sie wird geübt.

Wichtig ist, dass Schreiben als integraler Bestandteil der Lehre verstanden wird – insbesondere in Modulen mit professionsethischem Schwerpunkt. Formate wie Blitzexposés oder kommentierende Lektürenachbereitungen (Frank, Haacke, & Lahm, 2013, S. 28–31) bereiten nicht nur auf Prüfungen vor, sondern unterstützen den Aufbau einer Haltung: Studierende befragen im Schreiben ihre eigenen Annahmen und bringen sie in Dialog mit fachlicher Theorie. Verunsicherung wird hier als Erkenntnispotenzial verstanden – nicht als Defizit.

Entsprechend zentral ist die Rolle der Lehrenden: Wenn Schreiben als Denkform ernst genommen wird, bedeutet Begleitung mehr als Korrektur. Sie ist Resonanz, Spiegel und Ermutigung. Lehrende moderieren

Denkbewegungen, öffnen Perspektiven – ohne sie vorzugeben. In der gemeinsamen Textarbeit entsteht professionelle Haltung nicht durch Lehre, sondern durch reflexive Praxis.

Das Schreiben ist damit kein Nebenziel. Es ist Medium einer Denkbewegung, die Perspektiven verdichtet und Urteile prüfbar macht. Diese Praxis stellte – vor der Nutzung von Gen-KI – Anforderungen an alle Beteiligten, eröffnete aber auch Möglichkeiten: das eigene Denken als prekär zu erleben, ohne es aufzugeben; durch Schrift eine Position zu finden, die sowohl wissenschaftlich begründet als auch subjektiv verantwortet ist. Mündlichkeit, insbesondere in kollegialen Formaten wie Fallbesprechungen oder Diskussionen im Seminar, flankierte diesen Schreibprozess und bot andere Räume der Selbstvergewisserung. Im Dialog entstanden Spannungsfelder, Differenzen und Widersprüche, die nicht aufgelöst, sondern gemeinsam getragen werden – auch das ist kritisch-reflexive Praxis. Heute stellt sich die Frage neu: Werden diese Reflexionsprozesse durch den Einsatz generativer KI ersetzt, ergänzt – oder verstellt?

3.3 Gen-KI im Studium – Potenzial oder Problem für die Haltung?

Mit generativer KI verändert sich die Art, wie Studierende in der Sozialen Arbeit mit Text, Theorie und Reflexion umgehen. Informationen, die vormals mühsam durch Schreiben und Diskussion erschlossen wurden, stehen heute in Sekunden als „elaborierter“ Text zur Verfügung. Reflexion droht – nicht zwangsläufig, aber möglicherweise – delegiert zu werden: an ein künstliches System ohne Erfahrung und Praxisbezug. Die Frage, was das für die Einübung kritisch-reflexiver Haltung bedeutet, stellt sich nicht nur theoretisch, sondern auch praktisch – und zwar hochschuldidaktisch und sozialarbeiterisch.

Gerade die Methoden, die Theorie, Biografie und Praxis über das Schreiben verknüpfen, geraten unter Druck. Wenn Schreiben nicht mehr an Erfahrung gebunden ist, sondern auf „künstlich“ erzeugte Strukturen zurückgreift, stellt sich die Frage neu, wie Reflexion gelehrt und begleitet werden kann. Die Herausforderung liegt nicht in der „Täuschung“ durch die Nutzung von Gen-KI, sondern in der Begleitung Studierender, sich als reflexive Subjekte im Zusammenspiel mit Gen-KI zu verorten.

Die Arbeitsgruppe „Künstliche Intelligenz: Essenzielle Kompetenzen an Hochschulen“ (2024) benennt zwei Aufgaben: epistemologische Veränderungen verstehen und didaktische Konsequenzen ziehen. Ziel ist, „Studierende auf eine veränderte Welt des Wissens und der Wissensproduktion vorzubereiten“ (S. 7). Für das Studium der Sozialen Arbeit bedeutet das: Bei der Vermittlung einer kritisch-reflexiven Haltung muss Gen-KI mitgedacht werden – allerdings unter didaktischen Bedingungen, die Selbstvergewisserung, Kontextualisierung und Auseinandersetzung fördern.

Zugleich gilt: Der „Umgang mit KI vor dem Hintergrund von Zielen wie der Förderung einer kritisch-reflexiven Haltung [...] [ist] nur [noch] im unmittelbaren Lehr-Lerngeschehen realisierbar“ (S. 19). Das eröffnet neue Formen der Haltungsarbeit: nicht mehr nur an sich selbst, sondern auch mit einer Gen-KI. Reflexivität bleibt Ziel akademischer Bildung – sie soll aber so vermittelt werden, dass Studierende Potenziale und Grenzen dieser Technologien erkennen und deren Einsatz verantwortlich gestalten können (S. 20–22). Für die Soziale Arbeit heißt das: Gen-KI darf nicht zur Effizienzsteigerung verkommen, sondern muss selbst reflexiv bearbeitet werden – ohne die Auseinandersetzung mit Biografie, Theorie und Praxis zu ersetzen. Es sei an die Studierende erinnert: Theorie und Praxis können auch mit Gen-KI reflektiert werden. Aber: Wenn Biografie, Wahrnehmung und Erfahrung außen vor bleiben, was bleibt dann?

Auch die Rolle der Lehrenden wandelt sich. Die Arbeitsgruppe fordert einen „verantwortungsvollen KI-Einsatz“ (S. 18) – durch Hochschulen, Lehrende und Studierende. Entscheidend ist die Reflexion der KI-Nutzung mit Studierenden. Unabhängig vom Szenario: Lehrende spielen nicht nur eine aktive Rolle. Sie spielen eine transformative Rolle, weil sie nicht nur Wissen vermitteln, sondern auch den kritisch-reflexiven Umgang mit Gen-KI in der Lehre mitdenken müssen. Schreibanlässe, die auf Dialog und Reflexion zielen, verlieren nicht an Relevanz – sie gewinnen an Bedeutung.

Gerade weil Gen-KI zur Glättung und Harmonisierung tendiert, weil sie Antworten bietet, bevor Fragen entstehen, ist die didaktische Ausrichtung auf Ambivalenz, Widerspruch und Uneindeutigkeit besonders im Studium Sozialer Arbeit gefordert. Reflexion bedeutet hier auch: die Bedingungen des eigenen Denkens, Schreibens und Urteilens zu reflektieren – z. B. über Portfolios, in denen sich Teilprozesse strukturieren lassen. Das ist ein Paradigmenwechsel: Nicht die Inhalte stehen im Vordergrund, sondern deren kritische Bezugnahme auf eigene Erfahrungen – auch im Lichte von Gen-KI.

Das Einüben einer kritisch-reflexiven Haltung muss sich nicht nur anpassen – es muss sich neu ausrichten. Haltung lässt sich nicht durch Kontrolle einüben, sondern durch Räume, die Vertrauen schaffen und eigene Fragen zulassen. Derzeit agieren Lehrende oft als Detektiv:innen, die prüfen, ob Gen-KI genutzt wurde und ob dabei Standards verletzt wurden. Doch sobald sich Prüfungsrecht und wissenschaftliche Standards den technischen Möglichkeiten angepasst haben, wird diese Rolle obsolet. Dann rückt eine andere Aufgabe ins Zentrum: die didaktische und prüfungsrechtliche Neugestaltung der Lehre, die Reflexion fördert und Gen-KI mitdenkt. Wenn kritisch-reflexive Haltung als Modus verantwortlicher Weltbegegnung verstanden wird, lautet die Frage: Was daran lässt sich automatisieren – und was bleibt notwendig menschlich?

4. Technik als symbolische Form – Kulturphilosophische Perspektiven auf die Einübung kritisch-reflexiver Haltung im Studium der Sozialen Arbeit mit Gen-KI

Die bisherige Auseinandersetzung hat gezeigt, dass es bei der Einübung kritisch-reflexiver Haltung im Studium der Sozialen Arbeit nicht allein um didaktische Methoden oder professionsethische wie -theoretische Inhalte geht, sondern um ein einzuübendes Selbstverhältnis, das sich ausbildet. Die Technisierung dieses Prozesses (durch Gen-KI) verändert den Modus des Bezugnehmens auf die Welt, wie ich sie wahrnehme, fundamental. Es geht nicht nur um technische Anpassung, sondern um kulturelle Transformation.

Cassirer hat den Begriff der „symbolischen Formen“ geprägt. Er begreift diese als Grundbedingungen menschlicher Welterschließung. Sprache, Mythos, Religion, Kunst, Wissenschaft – sie alle sind für ihn symbolische Formen, in denen sich das Verhältnis des Menschen zur Welt manifestiert. Technik gehört ebenso zu diesem Ensemble (Pircher, 2000; Hetzel, 2001, S. 148–152; Pätzold 2003; Recki, 2021). Sie ist kein Gegensatz zur Kultur, sondern deren Ausdrucksform: nicht Ergänzung, sondern Bedingung. „Das Sein kann hier nicht anders als im Werden, das Werk nicht anders als in der Energie sichtbar werden“ (Cassirer, 2009, S. 147). Hier hilft die Metapher der „Brille“: Wie durch unterschiedliche „Brillen“ hindurch können Menschen ihren sinnlichen Wahrnehmung Bedeutung verleihen. Eine Wahrnehmung ohne eine „Brille“ ist nicht möglich. Was das ist, was ich wahrnehme, hängt von der symbolischen Form ab, durch die hindurch ich dem, was ich wahrnehme, Bedeutung zuweise: etwas wird in der Energie – von der Cassirer spricht – sichtbar. Ernst Cassirers Begriff der Technik als „symbolische Form“ bietet hier einen erkenntnistheoretischen Schlüssel zum Verständnis dessen, was der zuvor beschriebene Paradigmenwechsel insgesamt für den (lernenden) Menschen bedeutet. Er ergänzt ihn dabei um die grundsätzliche erkenntnistheoretische Einsicht, die seit Searle (1994) nicht widerlegt werden konnte: „Syntax allein ist nicht geeignet, Bedeutung zu generieren“ (Heinrichs, Heinrichs, & Rüter, 2022, S. 56).

Technik erscheint in diesem Sinne als eine symbolische Praxis, die nicht bloß auf Herstellung abzielt, sondern auf Weltformung. Der Mensch wendet Technik nicht nur an, sondern die Technik selbst avanciert zu einem Modus des Weltbezugs. Ihre Funktion liegt hier dann nicht in der Produktion von Ergebnissen, sondern in der Organisation von Wahrnehmung – und damit auch in der Strukturierung des Selbstverhältnisses. Was das technische „Werkzeug“ dem Menschen gewährt, ist Cassirer zufolge nicht nur ein tätiges Können, sondern zugleich ein neues Sehen und Verstehen (Cassirer, 2009, S. 149f.). Wer Technik nutzt, sieht die Welt durch sie hindurch.

Gerade diese Perspektive lässt sich fruchtbar machen in der Diskussion um eine kritisch-reflexive Haltung im Studium der Sozialen Arbeit. Wenn Technik – wie Cassirer nahelegt – nicht als äußeres Hilfsmittel, sondern als kulturelle Form verstanden wird, durch die hindurch sich Welt erschließt, dann kann auch die didaktische Reflexion über Gen-KI sich auf diese symbolische Qualität beziehen. Gen-KI ist nicht nur ein Schreibwerkzeug, sondern ein Werkzeug in der symbolischen Form Technik, die Wahrnehmung strukturiert, Deutungsangebote macht und – in ihrer Funktionsweise – Denkbewegungen präformiert. Der Begriff der Präformation ist hier bewusst gewählt: Gen-KI stellt als Werkzeug eine Vorwegbildung des eigenen Denkens dar. Wenn es demnach um kritisch-reflexive Denkbewegungen geht, muss die Vorwegbildung dekonstruiert werden, um dann mit der unmittelbaren und echten wie eigenen Erfahrung abgeglichen werden zu können.

In dieser Hinsicht fungiert Gen-KI als das, was Cassirer den „terminus medius“ (Cassirer, 2009, S. 158) nennt: ein vermittelndes Glied im Erkenntnisprozess, das zwischen Welt und Subjekt steht und in sich selbst eine symbolische Ordnung konstituiert. Technik in diesem Sinne wird zum epistemologischen Medium – zur Bedingung der Möglichkeit, Welt zu denken. Für die didaktische Ausgestaltung professionsethischer Reflexion heißt das: Wer über Gen-KI spricht, spricht immer auch über symbolische Formen – und damit über kulturelle Praktiken der Welterschließung. Kurzum: Bis hierhin galten die Prinzipien der wissenschaftlichen Welterschließung für Fragen nach einer kritisch-reflexiven Haltung. Nun greift ein anderes Formprinzip: das der Technik.

Diese Einsicht legt nahe, dass sich der Fokus verschieben muss: Weg von einer instrumentellen Nutzung, hin zu einer reflexiven Betrachtung. Nur so kann Wissenschaftlichkeit und Techniknutzung „versöhnt werden“. Die Fokussierung auf eine für sich bestehende wissenschaftlich-technische Perspektive fordert dazu auf, die Nutzung von Gen-KI nicht nur zu lehren, sondern sie im Lichte ihrer kulturellen Ordnungsfunktion zu verstehen – als eine Form der Welterschließung, die stets auch ein Ethos impliziert.

Gerade in der Sozialen Arbeit, deren Professionalität sich durch relationales Denken, ethische Reflexion und kontextuelle Sensibilität auszeichnet, liegt in dieser kulturphilosophischen Perspektive ein entscheidender Mehrwert: Sie ermöglicht, Gen-KI nicht als bloßen „Störfaktor“ oder „Beschleuniger“, sondern als symbolische Form ernst zu nehmen – und in diesem Ernstnehmen eine Haltung zu entwickeln, die den eigenen Weltbezug im Medium von Technik befragt. Oder anders gesagt: Es verändert sich die Form der Erkenntnis, durch die hindurch wir etwas betrachten – nicht jedoch das Ziel der Reflexion.

Cassirers Rede von der Technik und der Form erlaubt damit einen Perspektivwechsel, der für das hier beschriebene Problem und dessen Umwandlung in ein Potential zentral ist: Die Herausforderung liegt nicht darin, den Gebrauch von Gen-KI zu verbieten oder zu normieren – sondern darin, ihn als symbolische Praxis zum Gegenstand didaktischer Auseinandersetzung zu machen. Die reflektierende Haltung gegenüber Technik wird so selbst zum Ziel und zur Bedingung eines kritischen Professionalisierungsprozesses – gerade im Studium der Sozialen Arbeit.

5. Professionsethik lehren mit KI – Ein Ausblick

Die Szene mit der Studierenden und dem mit ChatGPT verfassten Essay, die diesen Text eröffnet, hat bis hierhin als Ausgangspunkt für eine kritische Auseinandersetzung mit der Frage gedient, was es in Anbetracht der Möglichkeiten, die Gen-KI bietet, bedeutet, eine kritisch-reflexive Haltung im Studium der Sozialen Arbeit einzuüben. Sie soll nun noch einmal aufgenommen werden, um in einem letzten Schritt die Konsequenzen zu formulieren, die aus meiner Betrachtung resultieren.

Die Studentin hat formal korrekt gearbeitet, ihre Reflexion ist klar, stimmig, argumentativ konsistent – und dennoch fehlt etwas. Es fehlt nicht Wissen, nicht Struktur, nicht Sprache. Es fehlt der tastende Denkweg, das Ringen um das eigene Urteil. Die eigentliche Frage ist also nicht mehr, ob man Gen-KI im Studium nutzen darf – sie wird genutzt. Die Fragen lauten vielmehr: Wie kann man unter diesen Bedingungen das Denken lehren und lernen? Wie kann man kritisches Reflektieren einüben? Wie kann man die professionsethischen Paradoxien trotz der Möglichkeiten der Gen-KI als Triebfedern der eigenen Professionalität wertschätzen, ohne die Gen-KI als dienstbereite Übersetzungshilfe zwischen Theorie und Praxis misszuverstehen und dabei die biografische Dimension der eigenen professionellen Haltung außen vor zu lassen?

Eine zentrale Einsicht hierzu liefern Salden und Leschke (2023), wenn sie betonen, dass die Einführung von Gen-KI nicht bloß neue Tools zur Textproduktion bereitstellt, sondern in ihrer strukturellen Wirkung Fragen nach dem Verhältnis von Urheberschaft, epistemischer Autorität und wissenschaftlicher Praxis aufwirft. Gerade im Bereich des wissenschaftlichen Schreibens werde damit die Ausbildung grundlegender Kompetenzen nicht obsolet, sondern umso dringlicher. Denn auch wenn Gen-KI formal korrekte Texte erzeugen kann, bleibt das Ziel der Hochschulbildung die Ausbildung eines verstehenden, forschenden und urteilenden Subjekts (Salden & Leschke, 2023, S. 13–17).

Vor diesem Hintergrund schlagen beide vor, Lernziele so anzupassen, dass Studierende die Funktionsweise von Gen-KI nicht nur verstehen, sondern sie auch kritisch einordnen und sinnvoll lernzielorientiert nutzen lernen. Dazu zählen unter anderem die Fähigkeit, zwischen produktivem Einsatz

und inadäquater Nutzung zu unterscheiden, ethische Implikationen zu reflektieren und die eigene Schreibarbeit verantworten zu lernen (Salden & Leschke, 2023, S. 17–19). Auch die Rolle der Lehrenden verändert sich dabei: Sie sind nicht mehr allein Bewertende und Rückmeldende, sondern zunehmend Moderator:innen eines Reflexionsprozesses über das *Wie* und *Warum* von Textproduktion im akademischen Feld. Die Aufgabe verschiebt sich vom bloßen Korrigieren hin zur dialogischen Auseinandersetzung mit impliziten epistemischen, technischen und normativen Voraussetzungen akademischer Praxis (Salden & Leschke, 2023, S. 20).

Wenn man Ernst Cassirers Theorem der Technik als symbolische Form, als Interpretationsfolie verwendet und Technik als symbolische Form, als eine kulturelle Praxis versteht, die Welterschließung strukturiert, dann stellt die Nutzung von Gen-KI nicht bloß ein Werkzeug dar, sondern nimmt eine bestimmte symbolische (hier: technisch) vermittelnde Position ein. In dieser Position kann, so Cassirer, die Wahrnehmung von etwas als ein bestimmtes Etwas nicht mehr als Ergebnis, sondern nur noch als Beziehung gedacht werden – als das Verhältnis, das sich zwischen Subjekt und Welt in einer bestimmten Sinnwelt (hier: Technik) herstellt (2009, S. 149 f.). In anderen Worten: Dass künstliche intelligente Systeme ein besonders mächtiges und potentiell besonders transformatives kognitives Werkzeug sind, kann als unkontrovers gelten (Heinrichs, Heinrichs, & Rüter, 2022, S. 69). Gen-KI-Nutzung für professionsethische Fragestellungen zu nutzen, die mit dem eigenen Erleben zu tun haben und Einladungen darstellen sollen, sich kritisch-reflexiv mit den eigenen Erfahrungen auseinanderzusetzen, verhindert *zuerst* die Haltungsarbeit, weil man die eigene Erfahrung nicht einbezieht. Erst wenn man die eigene Erfahrung zulässt und bereit ist mit Gen-KI eben diese zu teilen, kann die Gen-KI auch bei der kritischen Reflexion helfen. So möchte ich die Transformation grob umreißen. Dass so mannigfaltige Probleme (Datenschutz, Umgang mit sensiblen Daten usw.) erst geschaffen und perspektivisch gelöst werden müssen, steht dabei außer Frage.

Weimann-Sandig spricht in diesem Zusammenhang davon, dass die Integration von KI in die Lehre nicht zur Reduktion von Komplexität führen dürfe, sondern zur bewussten Thematisierung jener Komplexitäten, die professionelles Denken auszeichnen (2023, S. 4). Sie plädiert dafür, Gen-KI als Hebel für Diskursfähigkeit zu nutzen – nicht zur Automatisierung von Argumenten, sondern zur Irritation bestehender Denkgewohnheiten. Für die Lehrenden bedeutet das notwendigerweise, dass sich Aufgabenformate verändern müssen: Weg vom bloßen Output, hin zur reflexiven Auseinandersetzung mit dem Denkprozess selbst.

Gerade im Studium der Sozialen Arbeit, wo der Umgang mit Ambivalenz, Normativität und Verantwortung zum professionellen „Kerngeschäft“ gehört, kann diese Einsicht nicht nebensächlich bleiben. Sie muss didaktisch fruchtbar gemacht werden. In diese Richtung denkt auch Michael

Gerlich, wenn er die These aufstellt, dass KI generell erst einmal dazu führt, dass das kritische Denken oberflächlicher bleibt (2025, S. 2). Das erzeugt, so seine These, eine Gefahr: Nicht das falsche Denken, sondern das allzu plausible Denken könnte zum Problem werden. Die „Glätte“ der KI-Texte verschleiert den Umstand, dass Reflexion nicht glatt ist. Sie ist sperrig, brüchig, widersprüchlich. Genau das aber macht sie zur Voraussetzung professionsethischen Urteilens: „Die Wirksamkeit dieses Werkzeugs hängt davon ab, inwieweit es gelingt, es kreativ, ethisch reflektiert und divergierend zu denken – und wie gut es in unsere Denkprozesse integriert wird“ (Singer, Báez, & Rios, 2023, S. 300, eigene Übersetzung).

Für die Lehre bedeutet das: Studierende müssen nicht nur lernen, mit Gen-KI umzugehen, sondern mit ihr zu streiten. Die Herausforderung liegt nicht im Verzicht auf Technologie, sondern in ihrer reflexiven Integration. Es bedarf eines „proaktiven Schrittes, um Sozialarbeitende mit den notwendigen Werkzeugen und Einsichten auszustatten, die sie benötigen, um in einem sich wandelnden Handlungsfeld verantwortungsvoll und kompetent agieren zu können“ (Rodriguez et. al, 2024, S. 180, eigene Übersetzung). Gleichzeitig ist darauf zu achten, dass Gen-KI nur dann eingesetzt wird, wenn es die „didaktisch sinnvollste und geeignetste Lösung“ darstellt und nicht „bloß im Namen der ‚Innovation‘ und ‚Digitalisierung‘“ integriert wird (Simis, 2025). Schreibanlässe, wie sie beschrieben wurden, müssen deshalb so gestaltet sein, dass sie nicht durch Gen-KI ersetzt, sondern durch Gen-KI zu Irritationen führen. Das bedeutet: Gen-KI-generierte Vorschläge werden nicht als Endpunkt genutzt, sondern als Ausgangspunkt für Diskussion, für Zweifel, für Nachdenken. Die Ergebnisse einer Gen-KI in ihrer Unvollkommenheit in die kritisch-reflexive Auseinandersetzung mit dem eigenen professionellen Anspruch und Handeln miteinzubinden muss gleichsam zur hochschuldidaktischen Strategie avancieren. Cassirers Einsicht, dass Technik nicht über das Menschliche hinausführt, sondern es symbolisch in einer bestimmten Hinsicht formt, ist in diesem Zusammenhang kein kulturpessimistischer Abgesang. Sie ist – im besten Sinne – ein Trost. Denn sie legt nahe: Auch KI als Werkzeug ist eine Form des menschlichen Gestaltens unseres Weltbezugs. Sie ist kein Ende der Reflexion, sondern deren neuer Anfang. Das Studium der Sozialen Arbeit kann, wenn es diese Einsicht aufnimmt, zu einem Labor werden – nicht nur für technische Kompetenz, sondern auch für eine neue Reflexionskultur. In dieser Kultur wird Haltung nicht mehr *gegen* Technik verteidigt, sondern durch Technik befragt. Vielleicht liegt gerade darin die Zukunft professionsethischer wie -theoretischer Lehre im Studium der Sozialen Arbeit.

Literatur

- Arbeitsgruppe „Künstliche Intelligenz: Essenzielle Kompetenzen an Hochschulen“ (2025). Künstliche Intelligenz: Grundlagen für das Handeln in der Hochschullehre (Arbeitspapier Nr. 86). *Hochschulforum Digitalisierung*. https://hochschulforumdigitalisierung.de/wp-content/uploads/2025/01/HFD_AP_86_Kuenstliche-Intelligenz_Grundlagen-fuer-das-Handeln.pdf
- Börrnert, R. (2023). Wie kann eine Ausbildung von Sozialen auf den „Praxisschock“ vorbereiten?. In M. Görtler, G. Taube & N. Thielemann (Hrsg.), *Soziale Arbeit und Professionalität: Reflexionen zwischen Theorie, Lehre und Praxis* (S. 171–182). Verlag Barbara Budrich.
- Bräuer, G. (2000). *Schreiben als reflexive Praxis: Tagebuch, Arbeitsjournal, Portfolio*. Fillibach.
- Cassirer, E. (2004). Form und Technik [1930]. In T. Berben (Bearb.) & B. Recki (Hrsg.), *Aufsätze und kleine Schriften (1927–1932)* (ECW, Bd. 17, S. 139–183). Meiner.
- Cassirer, E. (2009). Form und Technik (1930). In H. Burkhardt & J. Mittelstraß (Hrsg.), *Gesammelte Werke. Band 17: Aufsätze und kleine Schriften 1927–1931* (S. 139–167). Felix Meiner Verlag.
- Dewe, B. (2004). Wissen, Können und die Frage der Reflexivität. In W. Bender (Hrsg.), *Lernen und Handeln* (S. 321–331). Wochenschau Verlag.
- Dewe, B., & Otto, H.-U. (2011). Profession. In H.-U. Otto & H. Thiersch (Hrsg.), *Handbuch Soziale Arbeit: Grundlagen der Sozialarbeit und Sozialpädagogik* (4., völlig neu bearb. Aufl., S. 1131–1142). Ernst Reinhardt Verlag.
- Dewe, B., & Otto, H.-U. (2012). Reflexive Sozialpädagogik: Grundstrukturen eines neuen Typs dienstleistungsorientierten Professionshandelns. In W. Thole (Hrsg.), *Grundriss Soziale Arbeit: Ein einführendes Handbuch* (4. Aufl., S. 197–217). VS Verlag für Sozialwissenschaften.
- Frank, A., Haacke, S., & Lahm, S. (2013). *Schlüsselkompetenzen: Schreiben in Studium und Beruf* (2., aktualisierte und erweiterte Aufl.). Verlag J. B. Metzler.
- Friebertshäuser, B., Rieger-Ladich, M., & Wigger, L. (Hrsg.). (2006). *Reflexive Erziehungswissenschaft*. VS Verlag für Sozialwissenschaften.
- Görtler, M. (2023). *Theorie, Praxis, Professionalität: Eine empirische Untersuchung zu Vorstellungen von Fachkräften und berufsbegleitend Studierenden der Sozialen Arbeit* (unter Mitarbeit von C. Palatzky). In M. Görtler, G. Taube, & N. Thielemann (Hrsg.), *Soziale Arbeit und Professionalität: Reflexionen zwischen Theorie, Lehre und Praxis* (S. 75–88). Verlag Barbara Budrich.

- Görtler, M., Taube, G., & Thielemann, N. (2023). Soziale Arbeit und Professionalität. Die Gemeinsamkeiten der Unterschiede zwischen Theorie, Lehre und Praxis. Ein Fazit. In M. Görtler, G. Taube & N. Thielemann (Hrsg.), *Soziale Arbeit und Professionalität: Reflexionen zwischen Theorie, Lehre und Praxis* (S. 255–258). Verlag Barbara Budrich.
- Harmsen, T. (2014). *Professionelle Identität im Bachelorstudium Soziale Arbeit: Konstruktionsprinzipien, Aneignungsformen und hochschuldidaktische Herausforderungen*. Springer VS.
- Heinrichs, B., Heinrichs, J.-H., & Rüther, M. (2022). *Künstliche Intelligenz*. de Gruyter.
- Hetzel, A. (2001). *Zwischen Poiesis und Praxis: Elemente einer kritischen Theorie der Kultur*. Königshausen & Neumann.
- Kaminsky, C. (2018). *Soziale Arbeit – normative Theorie und Professionsethik*. Verlag Barbara Budrich.
- Middendorf, T. (2021). *Professionalisierung im Studium der Sozialen Arbeit: Eine sozialisationstheoretische Perspektive auf Ausbildungssupervision*. Beltz Juventa.
- Müller-Hermann, S. (2020). Die Unterstützung von Professionalisierung und Professionalität: Gesamtgesellschaftliche Entwicklungen der Gegenwart und Herausforderungen für die Ausbildung. In C. Steckelberg & B. Thiessen (Hrsg.), *Wandel der Arbeitsgesellschaft: Soziale Arbeit in Zeiten von Globalisierung, Digitalisierung und Prekarisierung* (S. 171–186). Budrich.
- Nickel-Schampier, T. (2020). Halt(ung) finden – Warum eine Professionsethik Sozialer Arbeit unverzichtbar ist? *Neue Praxis*, 50(1), 75–84.
- Pätzold, D. (2003). Die Technik – techne und poiesis. In H.J. Sandkühler & D. Pätzold (Hrsg.), *Kultur und Symbol: Ein Handbuch zur Philosophie Ernst Cassirers* (S. 211–219). Metzler.
- Peinemann, K., & Reinke, H. (2023). Vom „Theorieschock“. Oder: Berufsbegleitendes Studieren im Spannungsfeld von Theorie und Praxis. In M. Görtler, G. Taube & N. Thielemann (Hrsg.), *Soziale Arbeit und Professionalität: Reflexionen zwischen Theorie, Lehre und Praxis* (S. 183–196). Verlag Barbara Budrich.
- Pircher, W. (2000). Technik als symbolische Form: Ein taugliches Konzept? *Mitteilungen des Instituts für Wissenschaft und Kunst*, 55(1–2), 42–47.
- Puls, H., Schramm, M. W., & Waller, S. (Hrsg.). (2017). *Kultur, Technik, Freiheit: Festschrift zum 60. Geburtstag von Birgit Recki*. Mentis.
- Recki, B. (2021). *Natur und Technik – eine Komplikation*. Matthes & Seitz.

- Rodriguez, M. Y., Goldkind, L., Victor, B. G., Hiltz, B., & Perron, B. E. (2024). Introducing generative artificial intelligence into the MSW curriculum: A proposal for the 2029 educational policy and accreditation standards. *Journal of Social Work Education, 60*(2), 174–182. <https://doi.org/10.1080/10437797.2024.2340931>
- Rutttert, T. (2021). *Die professionelle Identität in der Sozialen Arbeit: Eine qualitative Studie über professionelle Selbstverständnisse sozialpädagogischer Fachkräfte*. Springer VS.
- Salden, P., & Leschke, J. (2023). *Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung* (Zentrum für Wissenschaftsdidaktik der Ruhr-Universität Bochum). https://hss-opus4.ub.ruhr-uni-bochum.de/opus4/frontdoor/deliver/index/docId/9734/file/2023_03_06_Didaktik_Recht_KI_Hochschulbildung.pdf.
- Schütze, F. (2021). *Professionalität und Professionalisierung in pädagogischen Handlungsfeldern: Soziale Arbeit*. Verlag Barbara Budrich.
- Searle, J. R. (1994). Geist, Gehirn, Programm. In W. C. Zimmerli & S. Wolf (Hrsg.), *Künstliche Intelligenz: Philosophische Probleme* (S. 232–266). Reclam.
- Simis, C. (2025). KI hier, dort und überall? Impulse für die Philosophielehre. LehrGut. Blog für philosophische Hochschullehre. <https://lehrgut.org/ki-hier-dort-und-ueberall-impulse-fuer-die-philosophielehre/>.
- Singer, J. B., Báez, J. C., & Rios, J. A. (2023). AI creates the message: Integrating AI language learning models into social work education and practice. *Journal of Social Work Education, 59*(2), 294–302. <https://doi.org/10.1080/10437797.2023.2189878>
- Venn, M. (2011). Lerntagebücher in der Hochschule. *Journal Hochschuldidaktik, 22*(1), 9–12.
- Walter, U. M. (2020). Let's keep playing! – Soziale Arbeit als kritisch-reflexive Improfession. In B. Völter, H. Cornel, S. B. Gahleitner & S. Voß (Hrsg.), *Professionsverständnisse in der Sozialen Arbeit* (S. 112–122). Beltz Juventa.
- Weimann-Sandig, N. (2023). *ChatGPT – Eine Chance zur Wiederbelebung des kritischen Denkens in der Hochschullehre*. Hochschulforum Digitalisierung. <https://hochschulforumdigitalisierung.de/de/blog/praxistest-chatgpt-weimann-sandig>.
- Winterlich, J., & Lohse, M. (2025, in Vorbereitung). Künstliche Intelligenz im Studium Soziale Arbeit in Sachsen: Nutzungskontexte und Wissensquellen. In R. Kusche & M. Spranger (Hrsg.), *NextGen Scientific Review – Annual Perspectives on Next Generation Science* (Vol. 4). Hochschulverlag Mittweida.
- Wolfsberger, J. (2021). *Frei geschrieben: Mut, Freiheit und Strategie für wissenschaftliche Abschlussarbeiten* (5., bearbeitete Aufl.). Böhlau Verlag.

Florian Müller¹ und Edeltraud Botzum¹

Künstliche Intelligenz im Peer-Mentoring

Reflexionen zur Förderung digitaler Future Skills im Studiengang Soziale Arbeit

¹ Technische Hochschule Rosenheim

Zusammenfassung

Im Projekt digi.peer wird untersucht, wie KI-gestütztes Arbeiten zur Förderung digitaler Schlüsselkompetenzen bei Studierenden der Sozialen Arbeit beitragen kann. Im Zentrum stehen didaktisch gerahmte Lernformate, die generative KI in die Arbeitsprozesse integrieren, um die Anwendungskompetenz zu stärken. Ziel ist eine reflexive KI-Bildung, die Studierende auf die digitale Arbeitswelt vorbereitet. Neben methodischer Innovation werden rechtliche Rahmenbedingungen und ethische Fragen thematisiert. Das ergänzende Zertifikatsprogramm digi.peer+ sowie das Projekt KISS (KI-Kompetenzen bei Studierenden stärken) erweitern dieses Angebot. Herausforderungen zeigen sich in häufig noch unklaren Richtlinien sowie heterogenen Vorkenntnissen der Studierenden und verdeutlichen die Notwendigkeit, Künstliche Intelligenz reflektiert und verantwortungsvoll in die Hochschullehre zu integrieren.

Schlüsselwörter: Künstliche Intelligenz, digitales Peer-Mentoring, digitale Future Skills, Hochschuldidaktik, reflexive KI-Bildung

Artificial Intelligence in Peer Mentoring. Reflections on Promoting Digital Future Skills in Social Work Degree Programme

Abstract

The digi.peer project explores how AI-supported working can contribute to foster digital competencies among social work students. The focus lies on didactically framed learning formats that integrate generative AI into academic work processes to enhance applied competence. The overarching goal is to promote reflective AI literacy that prepares students for a digitalized professional environment. In addition to methodological innovation,

Müller, F., & Botzum, E. (2026). Künstliche Intelligenz im Peer-Mentoring. Reflexionen zur Förderung digitaler Future Skills im Studiengang Soziale Arbeit. *DiMe-Sozial*, 2(1), 131–138. <https://doi.org/10.21240/dimesoz/2026/1/85>

the project also addresses legal frameworks and ethical considerations. The supplementary certificate program digi.peer+ and the KISS project further expand this programme. Challenges arise from frequently unclear guidelines and the heterogeneous prior knowledge of students, highlighting the necessity of integrating artificial intelligence into higher education in a reflective and responsible manner.

Keywords: Artificial intelligence, digital peer mentoring, digital future skills, University didactics, reflective AI education

1. Einleitung

Das von der Stiftung Innovation in der Hochschullehre geförderte hochschuldidaktische Lehrforschungsprojekt digi.peer – Digitales Peer-Mentoring (04/2024–03/2026) der Technischen Hochschule Rosenheim positioniert sich an der Schnittstelle von wissenschaftlichem Arbeiten, Hochschullehre und digitaler Transformation. Im Rahmen virtueller Literatur- und Forschungswerkstätten werden Studierende des Bachelorstudiengangs Soziale Arbeit im Schreib- und Forschungsprozess ihrer Abschlussarbeit begleitet. Neben der Nutzung digitaler Tools und KI-basierter Anwendungen zielt das Projekt u. a. darauf ab, Studierende gezielt auf die Herausforderungen einer digitalisierten Bildungs- und Arbeitslandschaft vorzubereiten. Darüber hinaus wird erprobt, wie generative KI (gKI) in einem didaktisch gerahmten Peer-Setting eingesetzt werden kann, um den Erwerb (kritischer) digitaler Future Skills zu fördern. Neben rechtlichen und ethischen Fragen werden auch Herausforderungen wissenschaftlicher Redlichkeit sowie das Potenzial KI-gestützter Selbstlernprozesse thematisiert.¹

2. Förderung von digitalen Zukunftskompetenzen

„Alles, was wir über unsere Gesellschaft, ja über die Welt, in der wir leben, wissen, wissen wir durch die Massenmedien“ (Luhmann, 1995, S. 9). Sprachmodelle wie ChatGPT avancieren seit November 2022 zunehmend zu Vermittlern von Information, die Kommunikation auf neuem Wege ermöglichen, und erfüllen so die gängigen Definitionskriterien des Mediums in seinem medientheoretischen Verständnis. Geschwindigkeit und Ausmaß des Wachstums stellen einen Präzedenzfall in der Welt der Technologie dar (Nguyen, 2025), der sich auch im Nutzungsverhalten Studierender widerspiegelt (u. a. Gärtner et al., 2024; Witter et al., 2024). Eine disruptive Innovation wie die der Künstlichen Intelligenz (KI) hat nachhaltige Auswirkungen auf die Bildungs- sowie Ausbildungslandschaft. Der rasante Wandel stellt neue Anforderungen an künftig benötigte Kompetenzen. Diese schlagen sich in der Idee des Future-Skills-Frame-

¹ Weitere Informationen, siehe Internetseite → https://projekte.th-rosenheim.de/de/forschungsprojekt/1149-digi-peer-digitales-peer-mentoring?person_id=295

work 2021 des Stifterverbandes für die Deutsche Wissenschaft nieder.² Es umfasst 21 Kompetenzen in vier übergeordneten Kategorien (Stifterverband, 2021):

- *Technologische Kompetenzen:* Kompetenzen, die insbesondere für die Gestaltung und effiziente Nutzung von Technologien wichtig sind.
- *Digitale Schlüsselkompetenzen:* Fähigkeit, sich in einer digitalisierten Umwelt zurechtzufinden und darin aktiv zu partizipieren.
- *Klassische Kompetenzen:* Generische Kompetenzen wie Teamfähigkeit, Kreativität und Resilienz.
- *Transformative Kompetenzen:* Zentral, um gesellschaftliche Herausforderungen unserer Zeit angehen zu können, wie Dialog- und Konfliktfähigkeit sowie Innovations- und Veränderungskompetenz.

Auch wenn diese vier Kompetenzkategorien nicht klar voneinander abgegrenzt werden können, so sind es doch besonders die digitalen Schlüsselkompetenzen Studierender, deren Förderung durch das Projekt digi.peer adressiert werden:

Kompetenzfeld	Beschreibung
Digital Literacy	Grundlegende digitale Fähigkeiten, sicherer Umgang mit Software und Einhaltung von Sicherheitsregeln im Netz.
Digital Ethics	Kritisches Hinterfragen digitaler Informationen, Reflexion des eigenen Handelns in digitalen Räumen, ethische Entscheidungen.
Digitale Kollaboration	Effiziente Zusammenarbeit über Online-Kanäle, unabhängig von räumlicher Nähe, inklusive digitaler Kommunikation und Etikette.
Digital Learning	Eigenständige Nutzung und Einordnung digitaler Informationen, Anwendung von Lernsoftware in Wissensaufbau.
Agiles Arbeiten	Selbstverantwortliches Arbeiten im Team unter Anwendung agiler und nutzerorientierter Arbeitsmethoden.

Tabelle 1: Digitale Schlüsselkompetenzen des Future-Skills-Framework. Eigene Darstellung; in Anlehnung an Stifterverband, 2021, S. 6.

Luhmanns prägnantes Zitat verdeutlicht die Wirkmacht der Medien. Aufgabe von Bildung, und damit auch Hochschulbildung, ist es, für diese zu sensibilisieren und Studierende für einen reflektierten Einsatz der Anwendungen zu schulen. Künstliche Intelligenz wird demzufolge im Projekt digi.peer nicht nur als technisches Hilfsmittel genutzt, sondern auch als Ausgangspunkt für kollektive Aushandlungsprozesse, methodische Innovationen und ko-konstruktive Lernformen. Lernen wird als aktiver Aneignungsprozess durch Ko-Konstruktion von Wissensbeständen unter

² Dieses wurde im Austausch mit Expert:innen aus Wissenschaft und Wirtschaft unter Berücksichtigung der Ergebnisse und Rückmeldungen zu der Future-Skills-Umfrage entwickelt, an der sich insgesamt 377 Unternehmen und 123 Behörden beteiligten.

mentorieller Begleitung durch Lehrpersonal (Kravčik, 2020) sowie Peer-to-Peer-Interaktion verstanden. Die didaktische Ausrichtung von *digi.peer* berücksichtigt also die Erkenntnis, dass wir als Individuen nicht nur selbstreferentielle Konstrukteure unserer eigenen Wirklichkeiten sind, sondern ebenso durch die kulturelle und mediale Umwelt in unserem Handeln beeinflusst werden (Siebert, 2005).

3. *digi.peer*: Praktische Erkenntnisse zur Vermittlung digitaler Kompetenzen

Durch das *digi.peer*-Projekt werden wichtige Erkenntnisse in Bezug auf förderliche Rahmenbedingungen für die Vermittlung digitaler Future Skills in der Hochschullehre gewonnen. Fortlaufendes Ziel ist es, hochschuldidaktische Impulse für eine reflexive KI-Bildung zu setzen, die technische, didaktische und professionsethische Dimensionen der Sozialen Arbeit integriert. Hierbei wird erprobt, wie KI-basierte Tools in einem Peer-Mentoring-Format sinnvoll eingesetzt werden können, um Studierende auf die transformierenden Anforderungen der Bildungs- und Arbeitswelt vorzubereiten. Ein Schwerpunkt des Projekts liegt auf der Entwicklung wissenschaftlicher Handlungskompetenzen sowie einer kritisch-reflektierten Auseinandersetzung mit KI-Technologien. Gestützt wird dies durch Workshops, digitale Lernformate und den bewussten Einsatz von generativer KI. Das im Sommer 2024 eigens eingeführte Zertifikatsprogramm *digi.peer+* (Digital Future Skills) bildet eine praxisnahe Grundlage zur Vertiefung der Kompetenzen Studierender und Lehrender im Umgang mit KI-basierten Anwendungen. Ergänzt durch das mit dem Rosenheimer Lehrförderpreis 2025 ausgezeichnete Projekt KISS: KI-Kompetenzen bei Studierenden stärken – Einsatz von KI im wissenschaftlichen Arbeiten zur Förderung digitaler Future Skills wird der Fokus auf ein hybrides Lehrkonzept gelegt (Laufzeit: 10/2025 bis 09/2026). Dieses nutzt Blended-Learning-Ansätze und praxisorientierte Methoden wie kollaborative Schreibübungen, um Studierende zu befähigen, KI-Technologien verantwortungsvoll in ihre wissenschaftlichen Arbeitsprozesse zu integrieren und dabei zugleich ethische und urheberrechtliche Fragestellungen kritisch zu reflektieren. Das Peer-Mentoring-Format unterstützt zudem die Studierenden dabei, sich gegenseitig bei der Integration von KI in ihren Schreib- und Forschungsprozessen zu begleiten (Gergen & Dörr, 2026 i. E.).

Mit Bezug auf eine nachhaltige Integration der Lehrangebote muss sich auch mit rechtlichen Aspekten (z. B. in Form von KI-Leitlinien) auseinandergesetzt werden (Stützer & Kravčik, 2023). Erfahrungen aus dem Projekt zeigen, dass der Mangel an Leitlinien und Vorgaben nach wie vor Unsicherheiten bei Studierenden, Lehrenden und Prüfenden erzeugt. Die Notwendigkeit für transparente Strategien wird durch die dynamische Weiterentwicklung Künstlicher Intelligenz zunehmend dringlicher und es zeigt sich, dass Hochschulen vor der Aufgabe stehen, Lehrmethoden

und Prüfungsformate kontinuierlich zu überdenken und anzupassen, wie es u. a. das Positionspapier der Arbeitsgruppe generative Künstliche Intelligenz (gKI) an der TH Rosenheim (2024) betont. Die Einbindung von KI in Lehr- und Prüfungsformate wird daher auch als notwendiger Gestaltungsprozess aufgefasst, der nicht Technologie duldet oder blind erlaubt, sondern maßgeblich mitgestaltet.³ Empfehlungen wie der Einsatz didaktischer KI-Tools, etwa zur Automatisierung von Feedbackprozessen oder zur gemeinsamen Reflexion KI-generierter Ergebnisse, verdeutlichen das Ziel von digi.peer, innovative Technologien mit reflektierter, wertegeleiteter Lehre zu verbinden. Auch im Peer-Mentoring-Format wird ein Balanceakt vollzogen: Einerseits werden Studierende ermutigt, gKI in komplexe Aufgabenfelder zu integrieren, andererseits sollen Kompetenzen wie Reflexionsvermögen, Problemlösekompetenz und Kollaborationsfähigkeit gestärkt werden.

Die KI-Leitlinie Hochschullehre von Hochschule Bayern e. V. (2025) setzt einen weiteren Rahmen. Sie bildet die Grundlage für datenschutzkonforme, regulierte und freiwillige Nutzung von KI-Werkzeugen und fordert u. a. die Kennzeichnungspflicht KI-generierter Inhalte sowie die Anpassung der Prüfungsverfahren. Die in der Leitlinie hervorgehobenen Aspekte, etwa die Förderung eines reflektierten Umgangs mit KI, fügen sich in die Struktur des Zertifikatsangebots digi.peer+ ein. Die Workshops bilden einen Grundstein, um ethisch wie didaktisch fundierte Grundlagen für den Umgang mit KI-basierten Werkzeugen zu schaffen. Demnach werden zentrale Fragestellungen auch im digi.peer-Projekt adressiert: Wie kann Künstliche Intelligenz sinnvoll in Lehrformate integriert werden? Welche ethisch-rechtlichen Leitlinien sind dafür notwendig? Und: welche didaktischen Prinzipien lassen sich auf andere Fachbereiche übertragen? Hierzu ist eine Transformation der Hochschulbildung nötig, die nicht allein von technologischer Machbarkeit ausgeht, sondern sich auch an den Anforderungen von Bildungswissenschaft, Medienpädagogik sowie anderen sozial- und geisteswissenschaftlichen Disziplinen orientiert (de Witt et al., 2020).

4. Fortschritte und Herausforderungen: Reflexionen aus dem Projektverlauf

Projektbasiertes und forschendes Lernen ist von Erfolgen geprägt und gleichermaßen durch Herausforderungen gekennzeichnet. Digitale Kompetenzen erwirbt man durch problembewältigendes Handeln in digitalen Räumen. Die hierfür relevanten Prozesse kann man begleiten und unterstützen (Wollersheim, 2023). Bezogen auf die Reflexion zur Förderung

³ Zu alternativen Szenarien wie Verboten oder bloßes Dulden von gKI merkt das Positionspapier an, dass diese Vorgänge Studierende unvorbereitet in eine digitalisierte Arbeitswelt entlassen würden.

einzelner Kompetenzfelder, wie sie im Rahmen des Future-Skills-Frameworks beschrieben werden, lässt sich für den bisherigen Projektverlauf von *digi.peer* festhalten:

Kompetenzfeld	Beschreibung
Digital Literacy	Erlernen grundlegender digitaler Fähigkeiten durch praktische Anwendungen (z. B. ChatGPT, Zotero in Kombination mit KI-Anwendungen wie z. B. ResearchRabbit). Reflexionseinheiten fördern methodische und kritische Nutzung.
Digital Ethics	Auseinandersetzung mit Verantwortung, Quellenkritik und Datenverarbeitung. Unsicherheiten zu rechtlicher Nutzung und Ethik werden thematisiert.
Digitale Kollaboration	Förderung orts- und zeitunabhängiger Zusammenarbeit in Peer-Groups durch Tools wie Whiteboards. Unterschiede in digitalen Kompetenzen bleiben herausfordernd.
Digital Learning	Asynchrone Lernmodule und personalisiertes Feedback unterstützen die autonome Entwicklung von Lernstrategien. Selbstorganisiertes Lernen erfordert Erfahrung.
Agiles Arbeiten	Iterative Methoden und Eigenverantwortung in Teams stärken technologische und soziale Kompetenzen im kollaborativen Arbeiten.

Tabelle 2: Digitale Schlüsselkompetenzen des Future-Skills-Framework im Rahmen von *digi.peer*, *digi.peer+* und *KISS*, eigene Darstellung.

Der Ansatz von *digi.peer* unterstreicht damit, dass die Förderung von Kompetenzen mehr sein muss als die bloße Anwendung technischer Tools: Bildung im digitalen Zeitalter fordert auch die kritische Reflexion technischer Entwicklungen ein, die Analyse von Macht- und Abhängigkeitsverhältnissen sowie die Stärkung individueller Gestaltungsfähigkeit.

5. Fazit: Lehren aus *digi.peer* für die Hochschulbildung

Am Beispiel von *digi.peer*, *digi.peer+* sowie dem mit *KISS* weiterentwickelten Lehrkonzept lassen sich Gelingensbedingungen und Herausforderungen für die Verbindung von Künstlicher Intelligenz und Hochschuldidaktik identifizieren. Diese Erkenntnisse sind nicht nur auf den Studiengang Soziale Arbeit beschränkt, sondern können ebenso auf andere Fachbereiche übertragen werden. Die Integration von KI in Bildungsprozesse ist demnach kein rein technischer, sondern vor allem didaktischer und ethischer Gestaltungsprozess, der alle Beteiligte vor neue Anforderungen stellt. Während einerseits die Nutzung generativer KI in studienrelevanten Kontexten das Potenzial hat, Lern- und Lehrpraktiken zu transformieren, bedarf dies andererseits klarer rechtlicher Rahmenbedingungen, ethischer Reflexion und methodischer Weiterentwicklung.

Als Gelingensbedingungen zeigen sich: 1. Flexible, adaptive und individuell zugeschnittene Lernumgebungen, in denen das kollaborative Arbeiten in Peer-Groups unterstützt durch digitale Tools, neben fachlichen auch soziale und ethische Kompetenzen stärkt. 2. Zertifikat digi.peer+, das zur Vertiefung digitaler Kompetenzen strukturierte Zusatzangebote mit konkreten Anwendungsbezügen und der Fokussierung auf reflexive KI-Nutzung schafft. 3. Augenmerk auf die Auseinandersetzung mit KI-unterstützter Bildung aus einer ethischen Perspektive hilft, Unsicherheiten aufzulösen und einen kritischen Umgang mit KI-Tools zu fördern.

Als Herausforderungen sind erkennbar: 1. Unterschiedliche digitale Vorerfahrungen der Studierenden, die durch Tutorials⁴ adressiert werden sollen sowie Bedarf an Orientierungsangeboten zu Datenschutz und rechtlichen Fragen im Umgang mit KI. 2. Die Notwendigkeit klarer und verbindlicher Hochschulrichtlinien zum KI-Einsatz, die neben ethischen Gesichtspunkten auch Fragen der Wissenschaftlichkeit und Prüfungsformate adressieren. 3. Spezifische Fortbildungen und Workshops für Lehrende, um Methoden der KI-gestützten Lehre sinnvoll implementieren zu können. 4. Die Gefahr, dass grundlegende Fähigkeiten, wie etwa wissenschaftliches Argumentieren, vernachlässigt werden. Stattdessen sollen technische und transformative Kompetenzen im Sinne des Future-Skills-Frameworks gefördert werden ohne die grundlegenden Ansprüche an Bildung und Selbstreflexion zu vernachlässigen. Hochschulen stehen damit vor der Aufgabe, Lernräume zu gestalten, die zugleich digitale Zukunftskompetenzen und die kritische Reflexionsfähigkeit Studierender fördern.

Literatur

- Arbeitsgruppe gKI an der TH Rosenheim (2024): *Gestalten statt dulden! Ein Positionspapier zum verantwortungsbewussten Umgang mit generativen KI-Anwendungen an der Technischen Hochschule Rosenheim*. https://www.th-rosenheim.de/fileadmin/abteilungen_und_einrichtungen/qls/KI/Positionspapier_KI_an_der_THRO_mit_HL-Beschluss_15.02.2024_final.pdf
- Gärtner, C., Moraß, A., Koss, S., Garbusa, S., Matern, S. & Innermann, I. (2024). *Einsatz, Nutzen und Grenzen von ChatGPT und anderen Large Language Modellen an den bayerischen HAWs*. <https://doi.org/10.34646/thn/ohmdok-1466>
- Gergen, A. & Dörr, M. (2026 i. E.). KI-gestützte Begleitung im digitalen Peer-Mentoring beim wissenschaftlichen Schreiben. In G. Zimmermann, & F. Karcher (Hrsg.), *Künstliche Intelligenz in der Hochschullehre – Ein Methodenbuch*, UTB Waxmann Verlag

4 Geplant im Rahmen des KISS-Lehrförderprojekts.

- Hochschule Bayern e. V. (Hrsg.) (2025): *KI-Leitlinie Hochschullehre. Empfehlungen zum Umgang mit KI in der Lehre der bayerischen Hochschulen für angewandte Wissenschaften und Technischen Hochschulen*. https://www.hochschule-bayern.de/fileadmin/daten/Positionspapiere/KI-Leitlinie_Hochschule-Bayern-19052025.pdf
- Kravčik, M. (2020). KI und Mentoring. In C. de Witt, F. Rampelt & N. Pinkwart, *Whitepaper „Künstliche Intelligenz in der Hochschulbildung“* (S. 22–24). Zenodo. <https://zenodo.org/record/4063722>
- Luhmann, N. (2017). *Die Realität der Massenmedien*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-17738-6>
- Nguyen, N. (2025). ChatGPT-Statistiken: Eine umfassende Analyse. ChatGPT-Statistiken enthüllt – Vom Benutzerwachstum bis zur wirtschaftlichen Auswirkung. *DOIT*. <https://doit.software/de/blog/chatgpt-statistiken>
- Siebert, H. (2005). *Pädagogischer Konstruktivismus: Lernzentrierte Pädagogik in Schule und Erwachsenenbildung*. Weinheim und Basel: Beltz Verlag.
- Stifterverband für die Deutsche Wissenschaft (Hrsg.) (2021). *Future Skills 2021. 21 Kompetenzen für eine Welt im Wandel. (Diskussionspapier Nr. 3)*. <https://www.stifterverband.org/medien/future-skills-2021>
- Stützer, C. M. & Kravčik, M. (2023). Künstliche Intelligenz in der Hochschulbildung. Bildungssoziologische Perspektiven und Herausforderungen. In C. De Witt, C. Gloerfeld & S. E. Wrede (Hrsg.), *Künstliche Intelligenz in der Bildung* (S. 219–239). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-40079-8_11
- Witt, C. de, Rampelt, F. & Pinkwart, N. (2020). *Whitepaper „Künstliche Intelligenz in der Hochschulbildung“*. Zenodo. <https://zenodo.org/record/4063722>
- Witter, S., Meinhardt-Injac, B., Siemer, L. & Späte, J. (2024). *ChatGPT im Studium der Sozialen Arbeit: Eine quantitative Studie zur Nutzung, Bewertung und Thematisierung in der Hochschule aus Studierendensicht*. Fachhochschule Potsdam. <https://opus4.kobv.de/opus4-fhpotsdam/3382>
- Wollersheim, H.-W. (2023). Bildung durch Künstliche Intelligenz ermöglichen. Ein Beitrag aus bildungstheoretischer Perspektive. In C. De Witt, C. Gloerfeld & S. E. Wrede (Hrsg.), *Künstliche Intelligenz in der Bildung* (S. 3–29). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-40079-8_1

Marcel Pietsch¹, Pascal Krimmer² und André Heinz³

KI als Antwort auf den sozialprofessionellen Fachkräftemangel?

Eine Verbindung professionstheoretischer Überlegungen mit empirischen Perspektiven aus dem Caritaspanel 2024

¹ FOM – Hochschule für Oekonomie und Management

² Deutscher Caritasverband

³ IU Internationale Hochschule

Zusammenfassung

In Anbetracht der Gleichzeitigkeit der beiden Entwicklungen aus Fachkräftemangel und KI-Verbreitung im Sozialwesen wird der Frage nachgegangen, inwieweit der betriebliche Einsatz von KI passende Reaktionsmöglichkeiten bei wachsenden Personallücken anbieten kann. Dabei werden neben aktuellen Zahlen zur Fachkräftesituation in SAGE-Berufen (Soziale Arbeit, Gesundheit und Pflege sowie Erziehung und Bildung) verschiedene Einsatzbereiche und Grenzen von KI literaturbasiert skizziert. Die Analyse wird mit statistischen Ergebnissen aus dem Caritaspanel 2024 ergänzt, die Schlüsse auf personalpolitische Herausforderungen und darauf bezogene Bewältigungspotenziale durch KI ermöglichen. Dabei zeigt sich, dass sich insbesondere administrative, jedoch auch zwischenmenschliche Tätigkeiten im Sozialwesen KI-unterstützt oder sogar automatisiert bearbeiten lassen. Die Studie plädiert für einen evidenzbasierten und verantwortungsvollen Einsatz von KI-Technologie, der Fachkräfte entlastet und menschliche Expertise sowie Klient:innenwohl stärken kann, statt zu gefährden.

Schlüsselwörter: Fachkräftemangel, Künstliche Intelligenz, Caritas, SAGE-Berufe

Pietsch, M., Krimmer, P., & Heinz, A. (2026). KI als Antwort auf den sozialprofessionellen Fachkräftemangel? Eine Verbindung professionstheoretischer Überlegungen mit empirischen Perspektiven aus dem Caritaspanel 2024. *DiMe-Sozial*, 2(1), 139–158. <https://doi.org/10.21240/dimesoz/2026/1/70>

AI as a Response to the Workforce Shortage in Social Professions? Connecting Profession Theory-Informed Reflections with Empirical Perspectives from the 2024 Caritaspanel

Abstract

Given the coincidence of growing staff shortages and the increasing spread of AI in the social sector, this study explores the extent to which AI can offer suitable responses to widening personnel gaps. In addition to recent data on workforce shortages in care and social service professions, possible areas of application and limitations of AI drawn from the literature are outlined. Statistical findings from the 2024 Caritaspanel complement the analysis, providing insights into HR-related challenges and the potential of AI to support corresponding personnel strategies. Results show that especially administrative tasks, but also interhuman activities can be AI-supported or automated. The study calls for an evidence-based and responsible use of AI that can relieve professionals and strengthen both human expertise and client well-being, rather than undermine them.

Keywords: Workforce Shortage, Artificial Intelligence, Caritas, Care and Social Service Professions

1. Thematischer Ausgangspunkt

Wer sich aufmerksam durch den öffentlichen Raum bewegt, wird die derzeitige Vielzahl an Werbekampagnen zur Personalsuche sozialprofessioneller Organisationen kennen: „Komm in unser Team als Erzieher/in, Sozialarbeiter/in, Altenpfleger/in“, „Hier findest Du sinnstiftende Arbeit mit Menschen“. Was sich auf Plakattafeln und in Werbespots eher subtil äußert, lässt sich mit arbeitswissenschaftlich registrierten Trends eines hohen Fachkräftebedarfs in den SAGE-Berufen verknüpfen. So scheint nicht erst seit Kurzem ein umkämpfter Wettbewerb um verfügbare sowie angehende Fachkräfte zwischen den verschiedenen sozialen Arbeitgeber:innen stattzufinden (Lochner & Thole, 2024). Aktuelle Analysen verdeutlichen, dass kaum eine andere Branche mit einer derart schwierigen Stellenbesetzung sowie daraus resultierenden Personallücken wie das Gesundheits- und Sozialwesen umgehen muss (Arndt et al., 2024).

Jene Schieflage überschneidet sich zeitlich mit einer anderen prägenden Neuentwicklung: der flächendeckenden Ausbreitung von Künstlicher Intelligenz (KI) in beruflichen und privaten Alltagssphären. Dabei scheinen insbesondere Large Language Models wie Chat-GPT als zunehmend populäre Werkzeuge zur möglichst effizienten Be- und Verarbeitung kognitiver, kommunikativer und kreativer Aufgabenprozesse eingesetzt zu werden (Schmidt et al., 2024). Auch in den SAGE-Berufen werden mittlerweile verschiedene Funktionsebenen und Nutzungsszenarien von KI-Innovationen beschrieben und vor dem Hintergrund aktueller Herausforderungen reflektiert (Macsenaere, 2024; Schulze, 2024). Umso mehr

stellt sich gegenwärtig die Frage, wo und wie KI in Sozialer Arbeit und umgebenden Feldern komplementäre oder auch substitutive personalpolitische Bedeutungen einnimmt und zur strategischen Bewältigung des Fachkräftemangels genutzt werden kann. Der vorliegende Beitrag zeichnet zunächst aktuelle im Fachdiskurs behandelte Herausforderungen der Fachkräftegewinnung sowie Möglichkeiten von KI-Technologien im Sozialwesen nach. Anschließend werden statistische Ergebnisse des wissenschaftlichen Forschungsprogramms „Caritaspanel“ aus dem Jahr 2024 zu Verknüpfungspunkten von Fachkräftemangel und praktischer KI-Nutzung präsentiert. Hierbei werden am Beispiel des größten sozialen Arbeitgebers in Deutschland empirische Einblicke in die organisationale Auseinandersetzung mit KI gegeben. Die Erkenntnisgewinnung schließt mit professions- und sozialtheoretischen Überlegungen zur Nutzung von KI als Bewältigungsansatz für sozialprofessionellen Fachkräftemangel.

2. Gleichzeitigkeit der Phänomene: Fachkräftemangel und KI-Aufkommen im Sozialwesen

2.1 Aktuelle Fachkräftesituation im Sozialwesen

Seit mehreren Jahrzehnten wird der sozialprofessionelle Arbeitsmarkt durch entscheidende Dynamiken mitgeprägt. Hierzu gehört z. B. ein umfassender Generationenwechsel, der sich im Kontext demografischer Verschiebungen durch eine „Ruhestandswelle“ Erwerbstätiger der sogenannten Babyboomer-Jahrgänge bemerkbar macht. Auch die Ausweitung und das Neuaufkommen von Handlungsfeldern und darin eingebetteter sozialer Unterstützungsbedarfe spielt eine wichtige Rolle für die Entwicklung der Arbeitslandschaft. Exemplarisch lässt sich auf die Arbeit mit geflüchteten Menschen, die Schulsozialarbeit, die Frühpädagogik sowie die Arbeit mit älteren Menschen hinweisen, welche in den vergangenen Jahren allesamt an Inhalten und Gewicht hinzugewinnen konnten (Braches-Chyrek, 2023; Engelke et al., 2024; Hohendanner, 2024). Aktuell lässt sich mit Zahlen der Statistik der Bundesagentur für Arbeit (2025a, 2025b, 2025c) auf ca. 465.000 sozialversicherungspflichtig Beschäftigte in den Arbeitsfeldern der Sozialen Arbeit schließen. Weiterhin gehen 775.000 sozialversicherungspflichtig Beschäftigte einer Arbeit in Einrichtungen der Kindertagesbetreuung nach. In der Pflege ist von 1,7 Millionen sozialversicherungspflichtig Beschäftigten auszugehen. In allen drei Bereichen konnte in den letzten Jahren ein erhebliches Beschäftigungswachstum verzeichnet werden. Zwischen 2014 und 2024 kam es in der Sozialen Arbeit zu einer Erhöhung der Beschäftigtenzahl um 17 Prozent, in der Pflege um 22 Prozent und in der Kinderbetreuung und -erziehung sogar um 48 Prozent. Die berufsspezifischen Arbeitslosenquoten liegen mit lediglich 2 bis 3 Prozent deutlich unter dem wirtschaftsweiten Durchschnitt (ca. 6 Prozent). Großes Wachstum zeigt sich beispielsweise auch in der Anzahl von Studierenden in der Sozialen Arbeit,

die sich weiterhin auf Rekordniveau bewegt. So waren hier im Wintersemester 2023/24 ca. 120.000 Personen für ein Studium eingeschrieben, was ein Plus von ganzen 80 Prozent gegenüber 2013/14 bedeutet.

Der Aufgabenzuwachs und damit korrespondierende Fachkräftebedarf scheint jedoch weder durch vorhandenes Personal noch durch steigende Studierenden- und Absolvent:innenzahlen in der Sozialen Arbeit und umgebenden Berufen ausreichend abdeckbar zu sein (Engelke et al. 2024; Meyer, 2024). Ein Blick auf die derzeitigen Fachkräftelücken¹ zeigt, dass in der Pflege ca. 31.000 und in der Frühpädagogik und Sozialen Arbeit jeweils ca. 21.000 offene Stellen als nicht besetzbar gelten (Arndt et al., 2024). Arbeitsmarktprojektionen deuten darauf hin, dass diese Lücken trotz des Beschäftigungswachstums in den kommenden Jahr(zehnt)en mit hoher Wahrscheinlichkeit bestehen bleiben werden. Dabei verweisen Maier et al. (2024) darauf, dass ein möglicher Rückgang der sozialprofessionellen Fachkräftelücke auch durch reduzierte politische Investitionen und sozialstaatliche Abbauprozesse bewerkstelligt werden könnte – unabhängig vom tatsächlichen Unterstützungsbedarf in der Bevölkerung. Umso mehr stellen sich Fragen, wie auf die angespannte Fachkräftesituation reagiert werden kann: Durch eine systematische Aufwertung praktischer Arbeits- und Beschäftigungsbedingungen? Durch die Erhöhung der Anzahl an Studien- und Ausbildungsplätzen? Oder kann der betriebliche Einsatz von KI einen zentralen Ansatzpunkt darstellen, um Personallücken gezielt auszugleichen?

2.2 Nutzungsszenarien von KI in SAGE-Berufen

Nicht erst seit Kurzem wird über die Potenziale zum Einsatz von automatisierter Robotik und Formen Künstlicher Intelligenz in menschenhelfenden Berufsfeldern diskutiert. Dabei schien lange Zeit der Konsens zu gelten, dass vor allem jene Arbeitsanteile, in denen anspruchsvolle, subjektbezogene Interaktions- und Emotionsarbeit bedeutsam ist, kaum bis gar nicht durch technologische Systeme umgesetzt werden können. Dies wurde in erster Linie damit begründet, dass die dabei notwendigen sozial(professionell)en Kompetenzen, soziale Intelligenz sowie insbesondere das Durchblicken komplexer sozialer Beziehungsmuster und Sachverhalte mit hoher Wahrscheinlichkeit nur durch Menschen realisierbar bleiben (Frey & Osborne, 2013). Spätestens mit der rapiden Verbreitung fortgeschrittener Large Language Models wie Chat-GPT des Anbieters OpenAI ist es vielerorts allerdings zu einer Neubewertung der Nutzungsmöglichkeiten von KI und ähnlichen „sozialen Maschinen“ gekommen (Frey & Osborne, 2024). So lässt sich unter Betrachtung des

¹ Als Fachkräftelücke wird die jeweilige Differenz zwischen offenen Stellen und potenziell verfügbaren, passend qualifizierten Erwerbslosen bezeichnet. Es handelt sich somit um einen konkreten Indikator für Fachkräftemangel.

hochdynamischen KI-technologischen Voranschreitens mittlerweile darauf schließen, dass sich zahlreiche Aufgaben hochqualifizierter Sozialtätigkeiten durchaus durch automatisierte Systeme ausführen lassen.

Pottharst et al. (2024) ordnen die durch neue Technologien erbringbaren Arbeitsprozesse in SAGE-Feldern auf den drei Funktionsebenen aus Verarbeitung, Vermittlung und Vernetzung von Informationen sowie deren Überschneidungszonen an. Die Verarbeitung bezieht sich auf Arbeitsschritte, bei denen Datenanalysen und -handling technologiegestützt beschleunigt und somit auch große und ggf. unübersichtliche Datenmengen besser nutzbar gemacht werden können. Als Szenario werden *Predictive Analytics* genannt, bei denen beispielsweise Risikoereignisse wie Rückfälligkeit in straffälliges Verhalten oder Kindeswohlgefährdungen antizipiert werden sollen. Bei der Vermittlung werden Technologien relevant, welche die Mensch-Maschine-Interaktion erleichtern und dabei v. a. den Informationsaustausch über verbindende Schnittstellen fördern. Dies kann von Visualisierungsprogrammen hin zu interaktiv erweiterten virtuellen Umgebungen reichen, in denen Sozialfachkräfte sich mit digitalen Inhalten auseinandersetzen, jedoch auch untereinander oder mit Klientel zusammenarbeiten. Bei der damit angeschnittenen Vernetzung werden Aufgaben adressiert, mit denen sich Daten übersetzen, Informationen systemübergreifend austauschen und soziale Verknüpfungspunkte einrichten lassen. Diese technologische Funktionsebene kann insbesondere bei sozialen Beratungs- und Vermittlungstätigkeiten eine wichtige Rolle spielen. KI nimmt Pottharst et al. (2024) zufolge im Zusammenspiel von Informationsverarbeitung und -vermittlung explizit größere Bedeutung ein. So wird hier konkret auf die Nutzungspotenziale von Chatbots, Large Language Models, Simulationen und ggf. algorithmischen Unterstützungssystemen bei fachlichen Entscheidungen hingewiesen (siehe hierzu auch Bastian, 2023). In der zentralen Überschneidungszone von Verarbeitung, Vermittlung und Vernetzung lassen sich ebenfalls KI-basierte Werkzeuge positionieren. Hierzu gehört das *Ambient Assisted Living*, d. h. mit sozial unterstützenden Technologien ausgestatteten Lebens- und Wohnumgebungen bei besonderen Hilfebedarfen. Auch das Arbeiten mit *Digital Twins* kann hier verortet werden. Dabei werden virtuelle, möglichst realitätsgetreue Abbilder realer Kontexte oder Personen durch Echtzeitdaten für Beobachtungen, Analysen und Simulationen genutzt. Konkrete Anwendungsmöglichkeiten ergeben sich u. a. in sozialraumorientierten Tätigkeiten sowie beispielsweise in der Unterstützung gesundheitlich vulnerabler Menschen.

Eine ähnliche Systematisierung der bislang diskutierten und beobachtbaren Nutzungsmöglichkeiten von KI in SAGE-Feldern schlägt Schulze (2024) vor, die hier an mehreren Stellen mit neuen Überlegungen, jedoch ohne Anspruch auf finale Vollständigkeit weiterentwickelt wird. Grundsätzlich lassen sich dabei drei Kernbereiche unterscheiden: KI in der Ver-

waltungsarbeit, KI in der praktischen Arbeit mit Menschen sowie KI als Assistenz der Klientel. Wie in Abbildung 1 visualisiert, können die Übergänge zwischen den Bereichen durchlässig sein.



Abbildung 1: Kernbereiche der Anwendungsmöglichkeiten von KI im Sozialwesen. Eigene Darstellung in Anlehnung an Schulze (2024).

KI in der Verwaltungsarbeit

In der sozialprofessionellen *Verwaltungsarbeit* kann KI insbesondere generativ-kreative und klassische administrative Zwecke erfüllen. Auf generativer Ebene können insbesondere die weit verbreiteten Large Language Models und Bildgeneratoren wie Chat-GPT, Copilot, Canva, Stable Diffusion oder DALL-E relevant sein. So kann die Medienerstellung und -bearbeitung bezogen auf neue oder bereits vorhandene Texte, Grafiken, Tabellen sowie Video- und Audiomaterial erheblich unterstützt werden. Auch die zielgerichtete Informationsgewinnung kann über KI-gestützte Prozesse aus Brainstorming, Onlinerecherche, inhaltlichen Zusammenfassungen sowie Automatisierungen wie Erinnerungsmanagement und Programmcodegenerierung vorangetragen werden. Gerade an diesem Punkt wird jedoch auch das bekannte Phänomen des Halluzinierens bedeutsam, d. h. die Nennung erfundener, unzutreffender Informationen durch KI (Puntschuh, 2024). In klassische Verwaltungstätigkeiten lässt sich KI beispielsweise durch Prozessautomatisierungen zur Termin-

steuerung, medizinischen Kodierung oder Abrechnung integrieren. Auch die sprachgesteuerte Falldokumentation mit automatisiertem Notieren, Ausfüllen und Strukturieren wird über zahlreiche KI-Tools bereits angeboten. IT-Aufgaben wie die Verwaltung von Websites, Sicherheitslösungen und Betriebssystemen fallen ebenfalls in den von KI erbringbaren Leistungsradius.

KI in der praktischen Arbeit mit Menschen

Der Einsatz von KI in der *Arbeit mit Menschen* lässt sich in fünf miteinander zusammenhängende Subkategorien aufteilen: Spracherkennung und -verarbeitung, Informationsgewinnung und Kommunikation, Datenanalyse, Simulation und Programmierung sowie Planung und Entscheidung.

Im Bereich Sprache lässt sich u. a. auf die Nutzung von Dolmetscher-Tools mit Spracheingabe und -ausgabe im sozialprofessionellen Arbeitsalltag verweisen. Auch durch Fachkräfte sprachgesteuerte Assistenzsysteme können eine zunehmend wichtige Rolle spielen. Hinzu kommen z. B. Sprachmusteranalysen bei sozialen oder gesundheitlichen Gefährdungskonstellationen sowie Natural Language Processing zur automatisierten Informationsextraktion aus Kommunikationsvorgängen.

Auch in der praktischen Arbeit mit Menschen kann KI-gestützte Informations- und Wissensgewinnung einen gängigen Baustein darstellen. So existieren spezialisierte KI-Rechercheangebote und -Tätigkeitshilfen für sozialprofessionelle Handlungsfelder. Beispiele hierfür sind jugendhilfe.ai, MyKI und LeavingCare.ai, welche sich zum Zeitpunkt der Beitragserstellung teils noch in Erprobungsphasen befinden. Indem fallspezifisch passendes Fachwissen einen hohen Stellenwert in der Bearbeitung sozialer Themen einnimmt, Wissensbestände jedoch gleichzeitig permanent aktualisiert werden müssen, können KI-Systeme, insbesondere Large Language Models, zudem zu informellen Weiterbildungs- und Lernzwecken genutzt werden. So können Sachverhalte nach gewünschtem Ausführlichkeits- und Komplexitätsgrad dialogisch ergründet und neu gewonnene Erkenntnisse in die Praxis eingeflochten werden – vorausgesetzt, sie werden als glaubwürdig und fachlich wertvoll eingeschätzt. Mit dem Schwerpunkt Kommunikation kann u. a. die Verbesserung der Zugänglichkeit und Barrierefreiheit von digitalen und analogen Angeboten adressiert werden, indem beispielsweise leichte Sprache oder Bildbeschreibungen generiert werden. Hinzu kommen KI-basierte Beratungsdienstleistungen, die über Chatbots, Apps oder Large Language Models erbracht werden können. Exemplarisch lässt sich auf das chatbot-assistierte Nachsorgeangebot iCAN verweisen, das aktuell in der Begleitung junger, an Depressionen erkrankter Menschen erforscht wird (Stiftung Deutsche Depressionshilfe

und Suizidprävention, 2025). Wissenschaftlich lässt sich darauf schließen, dass Potenziale und Bedeutungen interaktiver KI-Systeme in der psychosozialen Problembewältigung weiter anwachsen werden (Li et al., 2023).

Die datenanalytische Nutzung von KI kann sich auf die automatisierte Bildauswertung beziehen, z. B. bei Aufnahmen von Sozialräumen und Krisengebieten oder im medizinischen Screening. Damit verbunden lassen sich auch sozialprofessionelle Risiko- und Situationsanalysen aus Fallakten, Protokollen, Datenbanken oder sozialen Netzwerken KI-gestützt anfertigen. Auch das Monitoring von Gesundheitszuständen und -verhalten kann bedeutsam sein, sodass sich insbesondere auf gesundheitsförderliche und -gefährdende Muster bei Leistungsadressat:innen schließen lässt. In ähnlicher Weise kann KI zur Bewertung von sozial(politisch)en Programmen, methodischen Plänen oder der Authentizität von Daten herangezogen werden. Zentrale Stichworte können hierbei Deep Fakes und gezielte Falschinformationen darstellen.

Im Feld der Simulation und Programmierung können KI-Potenziale bei Lern- und Weiterbildungsprozessen verortet werden. Dabei lassen sich z. B. Trainingsdaten und -szenarien für die Qualifikation von Studierenden, Auszubildenden und Fachkräften entwerfen. Ebenso lassen sich Gesprächskontexte von Fachkräften, Klient:innen, Kooperationspartner:innen und anderen an Unterstützungsprozessen beteiligten Personen zu Übungszwecken simulieren, was z. B. in der Aneignung von Beratungs- und Mediationskompetenzen nützlich sein kann. Dabei können Large Language Models für die Übernahme gewünschter realistischer Rollen instruiert werden, um beispielsweise als herausforderndes soziales Gegenüber aufzutreten. Daneben können Codegenerierung und Programmierung mit direktem Bezug zur Handlungspraxis relevant sein. Dies ist z. B. der Fall, wenn (komplexe) sozialwissenschaftliche Analysen in Statistiksoftware wie SPSS oder R umgesetzt werden oder Anweisungscodes für 3D-Drucker entwickelt werden sollen, mit denen sich medizinische Komponenten, Lernmaterialien oder in sozialen Werkstätten benötigte Bauteile produzieren lassen. Steuer- und Programmierfunktionen von KI adressieren darüber hinaus das wachsende Gebiet aus Pflegerobotik, *Ambient Assisted Living* und weiterer ggf. autonomer Hilfstechnologie.

Einen weiteren Bereich der sozialprofessionellen KI-Verwendung in der Arbeit mit Menschen stellt zudem die Planung und Entscheidung dar. Dabei kann es um methodische Verfahrensabläufe gehen, die sich auf die Erstellung von Hilfe- und Entwicklungsplänen sowie Bildungs- und Gesundheitsprogrammen beziehen. Auch konkrete sozialarbeiterische Techniken wie das Anfertigen von sozialen Ressourcen- und Netzwerkkarten können durch KI im Unterstützungsprozess umgesetzt werden. Hinzu kommt ein automatisiertes Koordinieren und Matchen von individuellen Fallkonstellationen, Ressourcen, professionellen Hilfeleistungen oder ehrenamtlicher Mitarbeit. Insbesondere in gesundheitlichen SAGE-

Feldern kann KI planerisch bei Triage-Systemen, Vorhersagemodellen für die Bettenbelegung und Frühwarnsystemen bei Versorgungsengpässen nutzbar gemacht werden.

KI als Assistenz der Klientel

Als möglichst eigenständig nutzbares Hilfsmittel kann KI für Klient:innen ebenfalls über generativ-kreative Eigenschaften, in Form von Wearable-Geräten, zur Orientierung, Navigation und Identifizierung oder für soziale Interaktion bedeutsam werden. Generativ spielen mit Medienerstellung, Informationsgewinnung und Komplexitätsreduktion Prozesse eine Rolle, die bereits zuvor skizziert wurden. Wearables mit KI-Integration umfassen v. a. bestimmte Smartwatches, Sehhilfen, Sonar-Armbänder, Blutdruck- und EKG-Devices, Tracker, Sensoren und Erinnerungsgadgets für Medikation, Körperhaltung, Ruhezeiten oder Ernährung. Kennzeichnend kann dabei auch die Konnektivität entsprechender Hilfsaccessoires mit Software und Apps sein, aus denen ausführlichere Statistiken und weitergehende Funktionen hervorgehen. Durch KI unterstützte Orientierung, Navigation und Identifizierung kann für die Klientel einen besonderen Stellenwert einnehmen, wenn z. B. unter beeinträchtigter Mobilität die Umgebung analysiert und erläutert werden muss, um vor Kollisionen zu warnen. Bei Demenzerkrankungen kann der Nutzen von KI darin liegen, dass sich Personen, Objekte und Orte erkennen und schildern lassen. Soziale Interaktionsmöglichkeiten mit KI ergeben sich nicht allein aus Gesprächs- und Beratungseigenschaften dialogischer Systeme, sondern z. B. auch aus der Begegnung KI-basierter, bewegungsfähiger Roboter oder Spielzeuge, die als interessierte Gesprächs- und Lernpartner:innen beruhigend bzw. trostspendend auftreten sollen. Exemplarisch kann an dieser Stelle auf den in der Altenpflege einsetzbaren Empathie-Roboter namens Navel verwiesen werden (navel robotics, 2025).

2.3 Bewältigung des Fachkräftemangels durch KI: Erkennbare Möglichkeiten und Grenzen

Auf Grundlage der dargestellten Systematisierungsansätze können konkrete Potenziale von KI zur Bewältigung des Fachkräftemangels im Sozial- und Gesundheitswesen auf mehreren Ebenen abgeleitet werden. KI-Systeme können vorhandene Fachkräfte insbesondere durch Automatisierung repetitiver (Verwaltungs-)Prozesse deutlich entlasten. Hierzu zählen etwa sprachgesteuerte Falldokumentation, Terminverwaltung oder Abrechnungstätigkeiten. Auch in der sozialarbeiterischen Unterstützungsarbeit lässt sich auf Vorgänge schließen, auf durch KI-Implementierung zunehmend automatisierbar erscheinen. Exemplarisch kann auf die Anfertigung methodischer Instrumente wie Netzwerk- und Ressourcenkarten verwiesen werden. Hierdurch können potenziell mehr zeitliche Spielräume für die komplexe zwischenmenschliche Kernarbeit entstehen, bei der menschliche Expertise und Resonanzfähigkeit weiterhin als zen-

traler Gelingensfaktor betrachtet werden kann. Gleichzeitig kann KI zunehmend Aufgaben bewerkstelligen, die bislang personell schwer zu leisten waren, etwa durch den Einsatz nachsorgender Chatbots, prognostizierender Risikoabschätzung oder technifizierter Lebensumgebungen. In solchen Fällen könnte KI als Teilersatz für knapper werdende Fachkräfte wirken, insbesondere bei standardisierbaren, datenbasierten oder assistierenden Tätigkeiten. Die unkomplizierte Simulation didaktisch nutzbarer Gesprächssituationen oder Analysen von Sozialräumen mit digitalen Abbildern könnte zudem neue Wege einer modernen sozialprofessionellen Aus- und Weiterbildung eröffnen, was langfristig der Fachkräfteentwicklung zugutekommen kann. In bestimmten Subfeldern, beispielsweise bei der Nutzung von Wearables oder empathischen Robotern, lassen sich sogar direkte klient:innenbezogene Funktionen von KI übernehmen, in denen sich Assistenz mit fortgeschrittener Interaktion verbindet. Unter Berücksichtigung dieser Entwicklungen lässt sich schlussfolgern, dass KI bislang eher keinen vollumfassenden Ersatz für die Handlungspraxis in SAGE-Berufen zu leisten vermag, aber gezielt dort wirksam sein kann, wo sozialprofessionelle Präsenz nicht zwingend dauerhaft erforderlich ist oder vor dem Hintergrund von Personalengpässen derzeit schlichtweg fehlt. Potenziell kann sie dazu beitragen, vorhandenes Personal zu entlasten und Lücken in der fachlichen Infrastruktur graduell zu schließen. Wichtig bleibt an dieser Stelle auch die Annahme, dass sich die Implementierung von KI und damit verknüpften Technologien in die soziale Arbeitswelt gegenwärtig in einer Startphase zu befinden scheint und mehrere Nutzungsmöglichkeiten weiterhin den Anschein von Zukunftsskizzen aufweisen können. Gleichzeitig ist davon auszugehen, dass sich die Möglichkeiten der KI-Nutzung auf beschleunigtem Niveau weiterentwickeln werden und dies zur Neugestaltung sozialprofessioneller Arbeitsprofile und Aufgabenbewältigung führen kann.

Trotzdem sich mit dem Einsatz von KI im Sozial- und Gesundheitswesen eine Vielzahl innovativer Potenziale verknüpfen lässt, müssen auch die dabei aufkommenden substanziellen Herausforderungen und Grenzen thematisiert werden. Diese betreffen u. a. technische, ethische und strukturelle Dimensionen. Als besonders kritisch gilt der bislang massive Energieverbrauch und die daraus resultierende ökologische Belastung, ergänzt durch hohe Betriebskosten der benötigten Technik (Katirai, 2024; Kocak et al., 2025). Der Zugriff auf sensible personenbezogene Daten, etwa bei KI-gesteuerten Simulationen oder Monitoringverfahren, kann zahlreiche Komplikationen beim Datenschutz ergeben. Ebenso stellt sich die Frage nach der Verantwortlichkeit bei fehlerhaften KI-Entscheidungen, etwa in Fällen falscher gesundheitlicher Behandlung oder riskanter Beratungsverläufe. Auch scheint die genuin menschliche Emotions-, Resonanz- und Beziehungsarbeit weiterhin nicht gänzlich von KI abbildbar zu bleiben. So lassen sich Vertrauen, Empathie und soziale Anerkennung womöglich nicht beliebig durch technische Mittel imitieren, um eine gewünschte, ethisch tragbare Prozess- und Ergebnisqualität zu erreichen.

Ein zentrales Risiko liegt zudem in der algorithmischen Diskriminierung, wenn KI-Systeme auf problematischen Trainingsdaten basieren und so soziale Ungleichheiten oder gruppenbezogene Menschenfeindlichkeit reproduzieren. Schließlich beruht qualitativ hochwertiges sozialprofessionelles Handeln auf ethisch fundierten Werthaltungen und dem bedingungslosen Schutz von Menschenwürde und -rechten. Hierbei kann es sich jedoch um Prinzipien handeln, deren zuverlässige Berücksichtigung durch KI nicht garantiert werden kann (Albareda et al., 2025; Orwat, 2024). Darüber hinaus besteht die Gefahr schleichender *Deskilling*-Prozesse in der Sozialen Arbeit, wenn fachliche Einschätzungen, Entscheidungsprozesse und Interaktionsleistungen zunehmend an KI-Systeme delegiert werden. Reinmann (2023, S. 4) beschreibt *Deskilling* hierbei als eine Form der „Dequalifizierung mit Kompetenzverlust“. Eine solche Verlagerung kann dazu führen, dass professionelles Erfahrungswissen, reflexive Urteilskraft und methodische Kompetenzen langfristig eher erodieren als weiterentwickelt werden (Ferdman, 2025). Insbesondere in einem Handlungsfeld, das wesentlich von situativer Aushandlung, Ambiguitätstoleranz und professioneller Beziehungsgestaltung geprägt ist, birgt die Substitution menschlicher Praxis durch automatisierte Verfahren relevante Risiken.

3. KI-Nutzung unter hohem Fachkräftebedarf: empirische Ergebnisse aus dem Caritaspanel 2024

Die bislang behandelten theoretischen Überlegungen werden im nächsten Schritt mit Forschungsbefunden aus dem Caritaspanel 2024 in Beziehung gesetzt. Im Zentrum stehen dabei Analysen zu den Themenschwerpunkten der aktuellen Fachkräftegewinnung sowie der Nutzung von KI in den sozialen Handlungsfeldern der Caritas.

3.1 Empirische Methodik und Stichprobenbeschreibung

Mit dem quantitativen Forschungsprogramm *Caritaspanel* untersucht der Deutsche Caritasverband seit dem Zeitraum 2015/16 kooperativ mit dem Institut für Arbeitsmarkt- und Berufsforschung die eigenen Betriebs- und Beschäftigungsrealitäten. Zweijährlich werden aktuelle Entwicklungen zu Schwerpunktthemen der Arbeitslandschaft empirisch erfasst, wie Fachkräftesicherung, Beschäftigungsqualität, mobiles Arbeiten sowie Digitalisierung und KI in der Caritas. Dabei lassen sich sowohl Verbesserungspotenziale als auch Qualitätsmerkmale organisationaler Zusammenhänge erkennen, die in politischen Diskursen sowie zur Personalgewinnung bedeutsam sind. An der Befragungsrunde 2024 haben sich zwischen Juli und September 2024 insgesamt 277 Rechtsträger online beteiligt. Diese verteilen sich repräsentativ über die Handlungsfelder, Regionalkommissionen und Betriebsgrößen der Caritas. In den Daten spiegelt sich die Situation von ca. 3.500 Betrieben mit insgesamt ca. 135.000 Beschäftigten wider. Aufgeteilt nach primären Hilfebereichen, d. h. dem

Bereich mit den meisten Mitarbeitenden, stellt die Altenpflege mit einem Anteil von 34 Prozent das größte Feld. Ein Viertel der Rechtsträger ist primär in der Kinder- und Jugendhilfe angesiedelt, während 22 Prozent den weiteren sozialen Hilfen angehören. 8 Prozent der Rechtsträger stammen aus der Gesundheitshilfe. Rechtsträger der Beratungsstellen machen sechs Prozent aus, jene der Eingliederungshilfe und Psychiatrie fünf Prozent. Betrachtet man die Betriebsgröße nach Mitarbeitendenzahl, machen Rechtsträger mit höchstens 25 Beschäftigten 16 Prozent der Stichprobe aus. Fast 30 Prozent der Rechtsträger haben 26 bis 100 Beschäftigte in ihren Reihen. Jeder dritte Rechtsträger berichtet, zwischen 101 und 500 Mitarbeitende zu beschäftigen. Jeder fünfte teilnehmende Rechtsträger kommt zudem auf mehr als 500 Beschäftigte.

3.2 Ausprägung des Fachkräftemangels bei der Caritas

Die Herausforderungen der gegenwärtigen Fachkräftesituation in der Caritas lassen sich über verschiedene Indikatoren abbilden. Zunächst kann ein Blick darauf lohnen, welche Personalprobleme die Rechtsträger als besonders relevant für die nähere Zukunft bewerten. Betrachtet werden an dieser Stelle die Top 5 im Häufigkeitsranking der Personalprobleme. Für die Gesamtheit aller befragten Rechtsträger zeigt sich, dass 86 Prozent von einer erschwerten Deckung des Fachkräftebedarfs ausgehen. Hierbei handelt es sich seit 2016 um das am häufigsten genannte Personalproblem. Überdurchschnittlich hoch fällt der Anteil in der Altenpflege (93 Prozent) und Gesundheitshilfe (91 Prozent) aus. Auf Rang 2 folgt die hohe finanzielle Belastung der Betriebe durch Lohnkosten. Drei von vier (72 Prozent) Rechtsträgern stufen diese als bedeutsame Herausforderung ein. Seit 2020 hat sich der Anteil um 30 Prozentpunkte erheblich gesteigert. Besonders häufig melden Rechtsträger der Beratungsstellen (88 Prozent), Gesundheitshilfe (86 Prozent) sowie Eingliederungshilfe und Psychiatrie (86 Prozent) eine problematische Lohnkostenbelastung zurück. Den dritten Platz machen hohe Fehlzeiten bzw. Krankenstand und Vertretungsbedarfe aus. Ca. zwei Drittel (64 Prozent) aller Rechtsträger schätzen den Aspekt als relevantes Personalproblem ein. Auch hier ist eine deutliche Erhöhung um 25 Prozentpunkte seit 2020 zu beobachten. Überdurchschnittlich stark zeigt sich die Betroffenheit in der Gesundheitshilfe (86 Prozent), Eingliederungshilfe und Psychiatrie (79 Prozent) und Altenpflege (74 Prozent). Die Hälfte (49 Prozent) aller Rechtsträger gibt an, von einer Überalterung des Personals herausgefordert zu sein. Es handelt sich dabei um den höchsten Wert seit Beginn der Forschungsreihe. Am häufigsten bewerten Rechtsträger der Altenhilfe (68 Prozent), Gesundheitshilfe (57 Prozent) sowie Eingliederungshilfe und Psychiatrie (57 Prozent) die Überalterung als bedeutendes Personalproblem. Auf Rang 5 befindet sich mit hohen Kosten für Personalsuche und -gewinnung ein weiterer Aspekt, der auf eine erschwerte Fachkräfteakquise hinweist. 43 Prozent der Rechtsträger sehen hierin ein relevantes Personalproblem. Überdurchschnittlich hohe Anteile liegen in der Eingliederungshilfe und

Psychiatrie (64 Prozent) und Gesundheitshilfe (53 Prozent) vor. Auffällig bleibt, dass insbesondere Hilfebereiche mit gesundheitsprofessioneller Ausrichtung häufiger von den genannten Herausforderungen betroffen zu sein scheinen. Auch zeigt sich bei Aufteilung nach Mitarbeitendenzahl, dass größere Rechtsträger teils deutlich häufiger als kleinere Rechtsträger von den genannten Personalproblemen berichten. Zudem ist anzumerken, dass Personalprobleme, die wie Weiterbildungsbedarf, mangelnde Arbeitsmotivation und arbeitszeitlicher Koordinationsaufwand die verkomplizierte Fachkräftegewinnung eher sekundär adressieren, im unteren Teil des Rankings auftauchen.

Die Einschätzungen der Rechtsträger werden als nächstes mit „härteren Fakten“ zur Nichtbesetzungsquote verbunden. Caritasweit liegt diese bei 16 Prozent, so dass ca. jede sechste ausgeschriebene Stelle unbesetzt bleibt. Es handelt sich trotz Rückgang um 8 Prozentpunkte gegenüber 2022 um die zweithöchste Ausprägung seit Beginn des Forschungsprogramms. In der Eingliederungshilfe und Psychiatrie (25 Prozent) sowie den weiteren sozialen Hilfen (24 Prozent) liegt die Quote über dem Durchschnitt – hier bleibt jeweils ein Viertel der ausgeschriebenen Stellen unbesetzt. Konträr zur subjektiven Bewertung der Personalprobleme zeigt sich: Je kleiner der Rechtsträger, desto höher im Mittel die Nichtbesetzungsquote. Die Rechtsträger wurden zusätzlich gefragt, aus welchen Gründen Ausschreibungen nicht zum Erfolg führen. 84 Prozent aller Rechtsträger geben an, dass keine geeigneten Bewerbungen eingegangen sind. Der Anteil bewegt sich seit 2016 auf ähnlichem Niveau. Von einem gänzlichen Ausbleiben von Bewerbungen auf ausgeschriebene Stellen berichten 48 Prozent, was eine Abnahme um 20 Prozentpunkte gegenüber 2022 bedeutet. 38 Prozent der Rechtsträger geben zudem an, dass angebotene Stellen von den Bewerber:innen abgelehnt wurden. Dieser Anteil ist seit 2020 um 30 Prozentpunkte angestiegen, so dass es sich um eine Erfahrung in der Fachkräftesicherung zu handeln scheint, die an Gewicht hinzugewinnt. Der Blick auf die Hilfebereiche zeigt, dass in der Eingliederungshilfe und Psychiatrie ungeeignete Bewerbungen (90 Prozent) sowie fehlende Bewerbungen (60 Prozent) überdurchschnittlich oft zur Nichtbesetzung führen. Auch in der Gesundheitshilfe (67 Prozent) ist die Herausforderung fehlender Bewerbungen stärker als in anderen Bereichen ausgeprägt. Aufgeteilt nach Mitarbeitendenzahl gehen bei den größten Rechtsträgern Nichtbesetzungen überdurchschnittlich häufig auf ungeeignete Bewerbungen (88 Prozent), ausbleibende Bewerbungen (64 Prozent) sowie abgelehnte Stellenangebote (52 Prozent) zurück.

3.3 Praktische Nutzung von KI bei der Caritas

Erstmals wurde im Befragungsjahr 2024 untersucht, wie sich die konkrete Auseinandersetzung der Caritas-Rechtsträger mit den Einsatzpotenzialen von KI im betrieblichen Arbeitsalltag gestaltet. Insgesamt geben 44 Prozent der Rechtsträger an, sich grundsätzlich mit den Anwendungsmöglich-

keiten von KI zu beschäftigen. Besonders hoch fällt der Anteil in der Gesundheitshilfe (64 Prozent) aus, während sich beispielsweise in der Kinder- und Jugendhilfe nur ca. jeder vierte (27 Prozent) Rechtsträger mit dem Einsatz von KI auseinandersetzt. Ein deutliches Gefälle ist zudem bei Betrachtung der Mitarbeitendenzahl zu registrieren. Nur bei 22 Prozent der kleinsten Rechtsträger spielt das Thema KI eine Rolle, bei den größten Rechtsträgern sind es mit 70 Prozent mehr als dreimal so viele.

Als nächstes wurde die Teilgruppe an Rechtsträgern, die sich mit KI beschäftigen, gefragt, ob sie diese auch tatsächlich schon in der Praxis einsetzt. Insgesamt bejaht dies fast die Hälfte (48 Prozent) jener Rechtsträger. Bezogen auf alle Rechtsträger der Caritas bedeutet dies, dass in jedem fünften (21 Prozent) Rechtsträger KI zum Einsatz kommt. Aufgegliedert nach Hilfebereichen ist festzustellen, dass in der Eingliederungshilfe und Psychiatrie (71 Prozent) sowie Gesundheitshilfe (64 Prozent) besonders häufig KI tatsächlich eingesetzt wird, sollte eine Auseinandersetzung damit stattfinden. In der Analyse der KI-Nutzung nach Mitarbeitendenzahl zeigen sich nur geringe Unterschiede zwischen den Größenkategorien. Abbildung 2 verdeutlicht das Aufkommen von KI in den Hilfebereichen der Caritas.

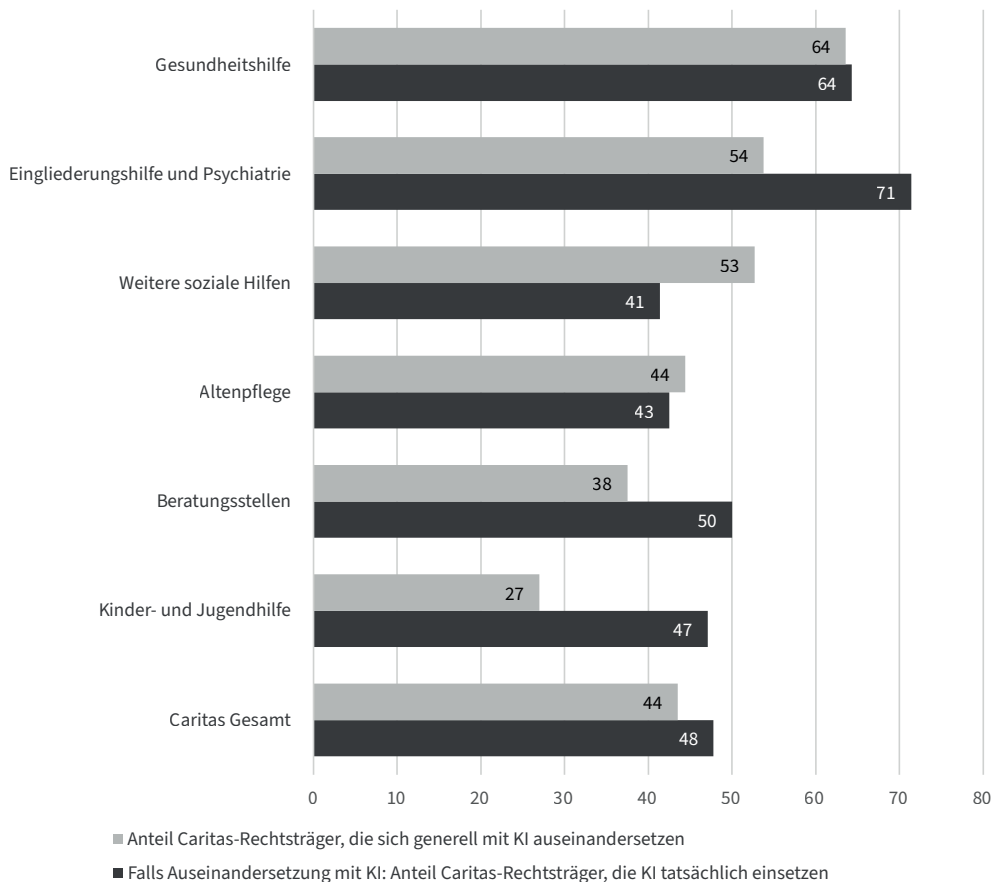


Abbildung 2: Aufkommen von KI in den Helfefeldern der Caritas. Anteile in Prozent. Eigene Darstellung, Datenbasis: Caritaspanel 2024.

Mit einer offenen Frage wurden zudem die praktischen Anwendungsbeispiele von KI im Betriebsalltag festgehalten. In den verschiedenen Hilfebereichen der Caritas zeigt sich ein durchaus breites Spektrum konkreter Anwendungsszenarien, die sowohl administrative als auch fachlich-inhaltliche Aufgaben betreffen können. In der Altenpflege wird KI besonders vielfältig eingesetzt: Neben ihrer Integration in klassische Verwaltungsaufgaben wie Berichtswesen, Buchhaltung, Tourenplanung und Pflegedokumentation kommen hier auch technikgestützte Assistenzlösungen zum Einsatz. Dazu zählen Sturzsensoren, Erinnerungshilfen zur Medikation und der Einsatz sprachgesteuerter Systeme wie Alexa in Wohnräumen. Auch moderne Betreuungsroboter mit KI-Unterstützung werden teilweise eingesetzt. Darüber hinaus wird KI unterstützend in der Öffentlichkeitsarbeit (z. B. Texte für Homepage und Werbematerial), bei Konzept- und Antragsentwicklungen sowie für die Erstellung von Präsentationen genutzt. Die Gesundheitshilfe weist einen noch stärkeren Fokus auf klinisch-fachliche Einsatzfelder auf. Genannt werden die digitale Bildverarbeitung in der Radiologie, klinische Entscheidungsunterstützungssysteme (Clinical Decision Support Systems, kurz: CDSS) sowie Anwendungen im Medizincontrolling und der Codierung. Hinzu kommen eher administrative Anwendungsbereiche, beispielsweise in Form KI-gestützter Terminvergabe sowie Konzept- und Texterstellung mit Chatbots und Large Language Models. Auch in der Kinder- und Jugendhilfe liegt ein Schwerpunkt auf der Nutzung von generativ-kreativer KI für Text- und Medienproduktion. Eingesetzt wird sie etwa zur Erstellung von Bildern, Reden, Berichten, Protokollen, Einladungen, Werbemitteln, Präsentationen und Infobriefen. Auch zur Dienstplanung und im Bildungsbetrieb werden KI-Anwendungen wie ChatGPT eingesetzt. Dabei kann allem die interaktive, pädagogische Nutzung von KI als besonders interessant gelten, wobei hierzu leider keine weiteren Ausführungen in den offenen Angaben vorliegen.

In der Eingliederungshilfe und Psychiatrie kommt KI ebenfalls bei der Dokumentation, Konzeptarbeit sowie Text-, Bild- und Präsentationserstellung zum Einsatz. Auch bei Stellenausschreibungen und Schriftverkehr wird KI genutzt. Ein sehr ähnliches Bild zeigt sich in den Beratungsstellen, wo KI u. a. für Artikel, Berichtsvorlagen, Anschreiben und Konzeptentwürfe genutzt wird. Zusätzlich wird berichtet, dass KI gezielt als Ideengeber zur Generierung von Projektnamen oder Ideenstrukturierung eingesetzt wird. In den weiteren sozialen Hilfen lässt sich ebenfalls v. a. auf administrative Nutzungszwecke schließen. So werden Schreiben und Dokumentationen, Social-Media-Beiträge, Übersetzungen (auch in leichte Sprache), Buchhaltung und Personalverwaltung mit KI bewerkstelligt.

Bereichsübergreifend zeigt sich, dass KI besonders häufig zur Unterstützung generativ-kreativer Verwaltungsaufgaben eingesetzt wird. In der Entwicklung von Konzepten, Berichten, Abbildungen, Pressemit-

teilungen, Anträgen oder Protokollen wird auf intelligente Sprachmodelle und Bildgeneratoren wie ChatGPT oder Copilot zurückgegriffen. Dokumentation ist ein weiterer zentraler Einsatzbereich, etwa über Spracheingabe oder automatische Textvorschläge. Auch Verwaltungstätigkeiten wie Dienstplanung, Terminvergabe oder Buchhaltung werden in einigen Rechtsträgern der Caritas KI-gestützt erledigt. Ergänzt werden diese Funktionen durch den Einsatz für Öffentlichkeitsarbeit, Recherche und, zumindest in manchen Fällen, Übersetzung und Barrierefreiheit. Die inhaltliche Nähe vieler Anwendungen deutet darauf hin, dass KI zunehmend als universal einsetzbares administratives Werkzeug sozialprofessioneller Arbeit verstanden werden kann. Der Einsatz von KI in der praktischen Arbeit mit Menschen scheint in der Caritas hingegen vorerst eine Nebenrolle zu spielen. Bis auf manche Einsatzfelder in pflegerischen, medizinischen Kontexten ist mit den vorliegenden Daten eher nicht auf eine verbreitete KI-gestützte Handlungspraxis in der unmittelbaren Interaktion mit Klientel zu schließen. Eine Limitation des hier generierten Erkenntniswerts kann allerdings darin liegen, dass sich die Nutzungsmöglichkeiten, Akzeptanz und Verbreitung von KI in sozialen Handlungsfeldern seit Erhebung im Sommer 2024 weiterentwickelt haben könnten. Insofern scheint die Erforschung und Wissenstransferierung im dynamisierten Themenfeld von KI derzeit einem Hochgeschwindigkeitswettbewerb zu gleichen.

4. Ausblick: Welche Wege eröffnen sich?

Im vorliegenden Beitrag wurden Verknüpfungspunkte zwischen Fachkräftemangel und KI-Nutzung in der sozialprofessionellen Arbeitswelt betrachtet. Dabei wurden die vielfältigen Anwendungsmöglichkeiten mit den drei Kernbereichen der Verwaltungsarbeit, der Arbeit mit Menschen und als Assistenz der Klientel systematisiert. Aus Forschungsperspektive lässt sich am Beispiel der Caritas festhalten, dass sich fast die Hälfte der Rechtsträger mit dem praktischen Einsatz von KI auseinandersetzt. Von diesen nutzt circa die Hälfte KI aktiv im Betriebsalltag. Sowohl die theoretische als auch empirische Beleuchtung der Verbindungslinien lässt auf konkrete Entlastungspotenziale durch KI im Kontext des Fachkräftemangels schließen. So scheint die gezielte Integration von KI bislang vor allem dort unterstützend zu wirken, wo Arbeitszeit bislang durch administrative, insbesondere dokumentierende und informationserzeugende Aufgaben gebunden wird. Wenn Berichte, Anträge, Protokolle oder Dienstpläne teilautomatisiert und gleichzeitig hochwertig durch KI erstellt werden können, lässt sich womöglich wertvolle, jedoch oft knappe Zeit für die originäre Beziehungs- und Interaktionsarbeit zurückgewinnen. Fachkräfte werden auf diese Weise nicht ersetzt, sondern tendenziell funktional entlastet. Somit lässt sich bereits heute auf größere Verschiebungspotenziale schließen, welche die Arbeitszeitressourcen des vorhandenen Personals effektiver nutzbar werden lassen.

Führt man den Gedankengang fort, könnte unter dem derzeit beobachtbaren Fortschritt von KI-Systemen darüber hinausgehendes Potenzial zur Substitution fehlender Fachkräfte durch KI als durchaus realistisch bewertet werden. Dies bezieht sich nicht bloß auf administrative Routine-tätigkeiten, sondern auch auf standardisierte sowie generativ gestaltbare Kommunikationsprozesse mit Klientel. So wurde skizziert, dass durch Chatbots in der psychosozialen Nachsorge junger Menschen oder durch intelligente Assistenzsysteme in der Pflege neue Einsatzszenarien entstehen, in denen menschliche Handlungsfähigkeiten in sozialprofessioneller Praxis gezielt durch KI-Systeme dargeboten werden. Möchte man diesen Weg fortsetzen, ergeben sich für SAGE-Akteur:innen regelmäßig neue Qualifikationsanforderungen, insbesondere zum Wissenserwerb im professionellen, verantwortungsvollen Umgang mit jeweils aktueller KI-Technologie. Auch ist auf erweiterte Investitionsbedarfe sozialer Organisationen im Kontext rascher Aktualisierungen von relevanter Hard- und Software zu schließen. Verlinkt man diese Annahmen mit der sozialen Beschleunigungstheorie von Rosa (2014) ist auf erhebliche Komplexitätssteigerungen und zunehmende Inkompatibilitäten bislang gültiger Vorgehensmuster zu schließen, je intensiver KI vorangetrieben wird. So besteht das Paradoxon darin, dass technische Fortschritte rapiden sozialen Wandel und Wissensverfälle bedingen, die wiederum das individuell erfahrene Lebenstempo und damit einhergehende Stresserfahrungen anwachsen lassen. Die präferierte Lösung für neu aufgekommene Geschwindigkeitsprobleme besteht dabei meist im Verlangen nach erneutem technischem Fortschritt, der die gewünschte Entlastungsverheißung in sich trägt. Diametral scheint sich der autodynamische Beschleunigungszirkel auf diese Weise allerdings zunehmend zu verschärfen, statt zu verlangsamen. Dabei können nicht alle Gesellschaftsmitglieder in einem solchen soziokulturell getragenen Akzelerationsspiel bestehen: Neben einigen Profiteur:innen bewegen sich viele Menschen verlustreich abgehängt in einer Abwärtsspirale durch ihre soziale Lebensrealität, wodurch Handlungsfähigkeit und Selbstwirksamkeitspotenziale stark vermindert werden.

Damit verknüpft sollte die Frage zentral bleiben, inwieweit die Ausdehnung von KI in der Erbringung von Sozial- und Gesundheitstätigkeiten als vertretbare Professionsentwicklung eingeschätzt werden kann. So gilt die Erkenntnis, dass nicht alles, was technologisch möglich ist, auch zwingend umgesetzt werden muss. Eine unreflektierte Umstrukturierung sozialprofessioneller Arbeitsrealitäten durch KI-Systeme kann demnach schnell einem Downgrade der Expertise gleichen: Fachliche Autonomie und Handlungsstärke Sozialarbeitender sowie umgebender Berufsangehöriger für die brennenden sozialen Fragestellungen der Gegenwart können darunter leiden (Bankins & Formosa, 2023; Hähnel et al., 2024; Ferdman, 2025). Technologische Innovationspotenziale müssen dabei stets dem Anspruch gerecht werden, dem Wohl von Klient:innen und Fachkräften zuträglich zu sein. Für eine tragbare, durchdachte Gestaltung von KI-Anwendungsmöglichkeiten in SAGE-Berufen braucht es folglich tief-

ergehende, informierte Diskussionen in und zwischen den verschiedenen Figurationen aus Praxis, Wissenschaft und Lehre. Dies kann dazu beitragen, evidenzbasiert zu lernen, gemeinsame Bedarfe und Zielrichtungen abzustecken, *Deskilling*-Risiken zu reduzieren sowie professionsethisch tragbare Lösungs- und Nutzungsideen mit Blick auf KI-Implementierungen zu entwickeln. Eine wichtige Reaktion auf den Fachkräftemangel liegt daher nicht in einer unkritisch vorgehenden Automatisierung sozialprofessioneller Handlungssphären, sondern in der strategischen Fokussierung menschlicher Eigenschaften und Ressourcen auf jene Aufgaben, die aus fachlicher Perspektive nicht zu ersetzen sind – unterstützt und entlastet durch eine technologisch versierte Infrastruktur.

Literatur

- Arndt, F., Tiedemann, J., & Werner, D. (2024). *Die Fachkräftesituation in Gesundheits- und Sozialberufen* (KOFA Kompakt 7/2024).
- Bankins, S., & Formosa, P. (2023). The Ethical Implications of Artificial Intelligence (AI) For Meaningful Work. *Journal of Business Ethics*, *185*(4), 725–740. <https://doi.org/10.1007/s10551-023-05339-7>
- Bastian, P. (2023). Algorithmusbasierte Entscheidungsfindung im Kinderschutz. *Sozialmagazin*, *48*(11–12), 58–65. <https://doi.org/10.3262/SM2312058>
- Braches-Chyrek, R. (2023). Soziale Arbeit: Ausbildung und Beschäftigung. In M. Friese & R. Braches-Chyrek (Hrsg.), *Care Work in der gesellschaftlichen Transformation: Beschäftigung, Bildung, Fachdidaktik* (S. 105–118). wbv.
- Engelke, E., Spatscheck, C., & Borrmann, S. (2024). *Die Wissenschaft Soziale Arbeit: Werdegang und Grundlagen* (5. Aufl.). Lambertus.
- Ferdman, A. (2025). AI deskilling is a structural problem. *AI & Society*, *40*. <https://doi.org/10.1007/s00146-025-02686-z>
- Frey, C. B., & Osborne, M. A. (2013). *The future of employment: How susceptible are jobs to computerisation?* University of Oxford.
- Frey, C. B., & Osborne, M. A. (2024). Generative AI and the future of work: a reappraisal. *Brown Journal of World Affairs*, *30*(1), 1–17.
- Hähnel, M., Pfeiffer, S., & Großmann, S. (2025). Striking the balance: ethical challenges and social implications of AI-induced power shifts in healthcare organizations. *AI & Society*, *40*(4), 2789–2806. <https://doi.org/10.1007/s00146-024-02043-6>
- Hohendanner, C., Rocha, J., & Steinke, J. (2024). *Vor dem Kollaps!? Beschäftigung im sozialen Sektor: Empirische Vermessung und Handlungsansätze*. De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110748024>
- Katirai, A. (2024). The Environmental Costs of Artificial Intelligence for Healthcare. *Asian Bioethics Review*, *16*(3), 527–538. <https://doi.org/10.1007/s41649-024-00295-4>

- Kocak, B., Ponsiglione, A., Romeo, V., Ugga, L., Huisman, M., & Cuocolo, R. (2025). Radiology AI and sustainability paradox: environmental, economic, and social dimensions. *Insights into Imaging*, 16(1). <https://doi.org/10.1186/s13244-025-01962-2>
- Li, H., Zhang, R., Lee, Y.-C., Kraut, R. E., & Mohr, D. C. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *npj Digital Medicine*, 6(1). <https://doi.org/10.1038/s41746-023-00979-5>
- Llorca Albareda, J., Liedo, B., & Martínez-López, M. V. (2025). Trusting the (un)trustworthy? A new conceptual approach to the ethics of social care robots. *AI & Society*. <https://doi.org/10.1007/s00146-025-02274-1>
- Lochner, B., & Thole, W. (2024). Bedarf und Mangel – Qualifikation und Fachlichkeit: Zur Situation und Qualifikation des Personals in der Sozialen Arbeit. *Sozial Extra*, 48(3), 156–161. <https://doi.org/10.1007/s12054-024-00697-6>
- Macsenaere, M. (Hrsg.). (2024). *Künstliche Intelligenz in der Kinder- und Jugendhilfe*. Ernst Reinhardt Verlag.
- Maier, T., Kalinowski, M., Schur, A., Zika, G., Schneemann, C., Mönig, A., & Wolter, M. I. (2024). *Weniger Arbeitskraft, weniger Wachstum: Ergebnisse der achten Welle der BIBB-LAB-Qualifikations- und Berufsprojektionen bis zum Jahr 2040* (BIBB Report 1/2024).
- Meyer, N. (2024). Die Lizenzierung für die Soziale Arbeit zwischen Ökonomisierung und Transformation. In J. Franz, C. Spatscheck & A. van Rießen (Hrsg.), *Fachkräftemangel und De-Professionalisierung in der Sozialen Arbeit: Analysen, Bearbeitungsweisen und Handlungsstrategien* (S. 21–34). Verlag Barbara Budrich. <https://doi.org/10.2307/jj.19850078.5>
- navel robotics (Hrsg.). (2025). *Navel – empathischer Roboter in der Pflege*. <https://navelrobotics.com/navel-der-soziale-roboter-in-der-pflege/>
- Orwat, C. (2024). Algorithmic Discrimination From the Perspective of Human Dignity. *Social Inclusion*, 12, Artikel 7160. <https://doi.org/10.17645/si.7160>
- Pottharst, B., Neumann, A., Ostrau, C., & Seelmeyer, U. (2024). Bewältigung des Fachkräftemangels durch technologische Innovation? Effekte von Technisierung und Digitalisierung. *Sozial Extra*, 48(3), 162–167. <https://doi.org/10.1007/s12054-024-00694-9>
- Puntschuh, M. (2024). Anwendungsrisiken von KI. In Der Paritätische Gesamtverband (Hrsg.), *Künstliche Intelligenz in der Sozialen Arbeit: Eine Textsammlung aus den KI-Veranstaltungen 2023 und 2024* (S. 6–8).
- Reinmann, G. (2023). *Deskilling durch Künstliche Intelligenz? Potenzielle Kompetenzverluste als Herausforderung für die Hochschuldidaktik* (Diskussionspapier Nr. 25). Hochschulforum Digitalisierung.
- Rosa, H. (2014). *Beschleunigung: Die Veränderung der Zeitstrukturen in der Moderne* (10. Aufl.). Suhrkamp Verlag.

- Schmidt, C. M., Stich, A., Suchy, O., André, E., Bullinger-Hoffmann, A., Bittner, E., Heister, M., Huchler, N., Peissner, M., Prasuhn, P., Steil, J., Wilkens, U., Kraus, M., & Ottersböck, N. (2024). *KI für die Fachkräftesicherung nutzen: Lösungsansätze für Automatisierung, Teilhabe und Wissenstransfer*. https://doi.org/10.48669/pls_2024-2
- Schulze, K. (2024). Die interessantesten Anwendungsmöglichkeiten im sozialen Bereich. In Der Paritätische Gesamtverband (Hrsg.), *Künstliche Intelligenz in der Sozialen Arbeit: Eine Textsammlung aus den KI-Veranstaltungen 2023 und 2024* (S. 14–18).
- Statistik der Bundesagentur für Arbeit. (2025a). *Blickpunkt Arbeitsmarkt: Akademikerinnen und Akademiker | Mai 2025. Kapitel 2.7: Sozialwesen*. <https://statistik.arbeitsagentur.de/DE/Statischer-Content/Statistiken/Themen-im-Fokus/Berufe/AkademikerInnen/Berufsgruppen/Generische-Publikationen/2-7-Sozialwesen.pdf>
- Statistik der Bundesagentur für Arbeit. (2025b). *Berichte: Arbeitsmarkt kompakt | Februar 2025 – Arbeitsmarkt Kinderbetreuung und -erziehung*.
- Statistik der Bundesagentur für Arbeit. (2025c). *Berichte: Arbeitsmarkt kompakt | Mai 2025 – Arbeitsmarktsituation im Pflegebereich*.
- Stiftung Deutsche Depressionshilfe und Suizidprävention (Hrsg.). (2025). *Chatbot-App für junge Menschen*. <https://www.deutsche-depressionshilfe.de/forschungszentrum/aktuellestudien/chatbot-app-fuer-junge-menschen-mit-depression>

Sabine Klinger¹, Susanne Sackl-Sharif¹ und Josephine Jahn²

Künstliche Intelligenz in der Sozialen Arbeit mit Kindern und Jugendlichen

Eine komparative Analyse der Sicht- und Nutzungsweisen von Fachkräften aus Österreich und Deutschland

¹ Karl-Franzens-Universität Graz

² Ludwig-Maximilians-Universität München

Zusammenfassung

In diesem Beitrag werden die Sicht- und Nutzungsweisen von Fachkräften der Sozialen Arbeit mit Kindern und Jugendlichen in Deutschland und Österreich in Bezug auf Künstliche Intelligenz (KI) untersucht. Basierend auf zwei quantitativen Online-Befragungen (N = 1.327 in Deutschland bzw. N = 303 in Österreich) werden Interesse, Wissen, Nutzungsverhalten, Hemmnisse sowie Einstellungen analysiert. Die Ergebnisse zeigen ein moderates bis hohes Interesse an KI, die jedoch nur sporadisch zur Content-erstellung, Informationsrecherche und zum Brainstorming genutzt wird. Als zentrale Hemmnisse erweisen sich fehlendes Wissen, unklare Einsatzmöglichkeiten und Datenschutzbedenken. Die befragten Fachkräfte in Deutschland zeigen zwar ein höheres Interesse, schätzen ihr Wissen aber geringer ein als die befragten Fachkräfte in Österreich. Die Befunde unterstreichen die Notwendigkeit des feldspezifischen Erwerbs von AI Literacy und organisationaler Rahmenbedingungen für einen reflektierten KI-Einsatz.

Schlüsselwörter: Künstliche Intelligenz, Kinder- und Jugendarbeit, Schulsozialarbeit, AI Literacy, komparative Studie, Online-Befragung, Ländervergleich

Klinger, S., Sackl-Sharif, S., & Jahn, J. (2026). Künstliche Intelligenz in der Sozialen Arbeit mit Kindern und Jugendlichen. Eine komparative Analyse der Sicht- und Nutzungsweisen von Fachkräften aus Österreich und Deutschland. *DiMe-Sozial*, 2(1), 159–183. <https://doi.org/10.21240/dimesoz/2026/1/59>

Artificial Intelligence in Social Work With Children and Young People. A Comparative Analysis of the Perspectives and Usage Practices of Professionals in Austria and Germany

Abstract

This article explores the views and usage behaviour of social work professionals working with children and young people in Germany and Austria with regard to artificial intelligence (AI). Based on two quantitative online surveys (N = 1,327 in Germany and N = 303 in Austria), interest, knowledge, attitudes, usage behaviour and barriers are analysed. The results show a moderate to high level of interest in AI, although it is only used sporadically for content creation, information research and brainstorming. A lack of knowledge, unclear application options and data protection concerns were identified as key obstacles. The surveyed professionals in Germany showed a higher level of interest, but rated their knowledge as lower than the surveyed professionals in Austria. The findings underline the need for field-specific acquisition of AI literacy and organisational framework conditions for a considered use of AI.

Keywords: Artificial Intelligence, child and youth work, school social work, AI literacy, comparative study, online survey, country comparison

1. Einleitung

Der Einsatz digitaler Technologien wird im Feld der Sozialen Arbeit zunehmend als selbstverständlich betrachtet (Klinger et al., 2023; Schönauer et al., 2021) und wird in den nächsten Jahren voraussichtlich steigen (Sackl-Sharif et al., 2024). Teil dieser digitalen Transformation ist die Entwicklung und Nutzung von Künstlicher Intelligenz (KI) (Görder, 2021). Dies wirkt sich auch in der Sozialen Arbeit auf vielfältige Weise aus und verändert Arbeitsbedingungen und -prozesse (Klinger et al., 2025). KI findet etwa in der Administration und Dokumentation, der Generierung von Webinhalten oder der Beratung Anwendung. Auch wenn die Einbindung von KI in die Praxis der Sozialen Arbeit bereits stattfindet (Gamble, 2020; Gillingham, 2006; Goldkind, 2018; Kreidenweis, 2024; Löhe, 2024; Schwartz et al., 2017), fehlen derzeit systematische Einblicke in die Arbeitspraxis und in den Einsatz von KI aus Sicht von Fachkräften. Der vorliegende empirische Beitrag adressiert diese Forschungslücke und verfolgt das Ziel, die Nutzungs- und Sichtweisen von Fachkräften in der Sozialen Arbeit in Österreich und Deutschland zu explorieren.¹

¹ Dieser Beitrag ist im Rahmen des Elisabeth-List-Fellowship-Programms für Geschlechterforschung an der Universität Graz entstanden.

Um das breite Handlungsfeld der Sozialen Arbeit² zu fokussieren, konzentrieren wir uns in diesem Beitrag auf Arbeitsfelder, deren Zielgruppe Kinder und Jugendliche sind. Diese Eingrenzung beruht darauf, dass digitale Technologien im Allgemeinen und KI im Besonderen bereits einen festen Bestandteil der Lebenswelten dieser Zielgruppe darstellen (Steiner & Tschopp, 2022). Die Auseinandersetzung mit KI in diesem Feld ist darüber hinaus besonders relevant, da hier oftmals vulnerable Personengruppen im Zentrum stehen (Linnemann et al., 2023). Als zentrale Arbeitsfelder der Sozialen Arbeit mit Kindern und Jugendlichen sind in diesem Beitrag die Kinder- und Jugendarbeit sowie die Schulsozialarbeit im Fokus. Beide Bereiche folgen einem gemeinsamen handlungsleitenden Konzept, das sich an Lebenswelt- und Beziehungsorientierung sowie dem Prinzip der Freiwilligkeit orientiert (Auferbauer et al., 2019). Die in diesen Feldern tätigen Fachkräfte verfügen somit über weitgehend übereinstimmende professionelle Fähigkeiten und Fertigkeiten, weshalb es uns sinnvoll erscheint, die beiden Handlungsfelder gemeinsam zu betrachten.

Vor diesem Hintergrund werden im Beitrag die Sicht- und Nutzungsweisen von Fachkräften³ der Sozialen Arbeit mit Kindern und Jugendlichen in Österreich und Deutschland in Bezug auf KI exploriert. Im Fokus stehen dabei das Interesse und Wissen der Fachkräfte bezüglich KI, ihr aktuelles Nutzungsverhalten in ihrer beruflichen Praxis sowie ihre grundsätzlichen Haltungen gegenüber dem beruflichen Einsatz von KI. Die empirische Basis bilden zwei quantitative Online-Fragebogenerhebungen unter den Fachkräften.

Der vorliegende Beitrag gliedert sich in insgesamt sechs Abschnitte. Zu Beginn richtet sich der Blick auf KI in der Sozialen Arbeit mit Kindern und Jugendlichen (2.), wobei AI Literacy als Voraussetzung für professionelles Handeln diskutiert (2.1) und aktuelle Einsatzformen von KI sowie Haltungen gegenüber KI thematisiert werden (2.2). Im Anschluss daran wird das methodische Vorgehen der komparativen Studie mit ihren fünf zentralen Forschungsfragen präsentiert (3.). Zudem werden in diesem Abschnitt die verwendeten Messinstrumente für die Online-Befragungen in Deutschland und Österreich (3.1), die Stichproben (3.2) sowie die deskriptiven Analysemethoden (3.3) dargelegt. Die empirischen Befunde werden in Kapitel 4 dargestellt und folgen vier thematischen Schwerpunkten: Interesse und Wissen der Fachkräfte (4.1), berufliches

2 Die Profession der Sozialen Arbeit zeichnet sich durch ihr breites Spektrum an Aufgaben und Handlungsfeldern aus. Dies ist darauf zurückzuführen, dass unter dem Begriff Soziale Arbeit die historisch gewachsenen Entwicklungsstränge der Sozialarbeit und Sozialpädagogik zu einer Profession zusammengefasst wurden, die auf zwei unterschiedliche Traditionen – einerseits eine fürsorgende, andererseits eine erziehende – zurückblicken (Thole, 2012).

3 Mit dem Begriff *Fachkräfte* adressieren wir Personen, die in der Sozialen Arbeit mit Kindern und Jugendlichen arbeiten. Zudem beziehen wir Führungskräfte auf der zweiten Führungsebene mit ein. Geschäftsführungen oder Vorstände sind hingegen nicht mitgemeint.

Nutzungsverhalten und Einsatzbereiche (4.2), Gründe für die Nicht-Nutzung (4.3) sowie Einstellungen zum Verbesserungspotenzial für die Arbeit mit Kindern und Jugendlichen durch KI (4.4). Daran anschließend bietet Kapitel 5 eine zusammenfassende Betrachtung der Ergebnisse im Ländervergleich (5.1) und diskutiert zentrale Limitationen der Studie sowie Perspektiven für die zukünftige Forschung (5.2). Im abschließenden Kapitel 6 werden die Ergebnisse in aktuelle Diskurse eingeordnet und insbesondere die Bedeutung von AI Literacy für die Weiterentwicklung des professionellen Handelns von Fachkräften diskutiert.

2. KI in der Sozialen Arbeit mit Kindern und Jugendlichen

2.1 AI Literacy und professionelles Handeln

Da Fachkräfte in der Sozialen Arbeit mit Kindern und Jugendlichen eine besondere Verantwortung gegenüber ihren Adressat:innen haben (Linnemann et al., 2023), benötigen sie im digitalen Zeitalter neben professionsbezogenen Fähigkeiten und Fertigkeiten insbesondere auch grundlegende Kenntnisse über digitale Technologien im Allgemeinen und KI-Tools im Besonderen. Letzteres wurde vor allem mit dem Inkrafttreten des AI Acts der Europäischen Union im Jahr 2025 (EU AI Act, 2025) deutlich. Seit Februar 2025 sind Organisationen demnach dazu verpflichtet, sicherzustellen, dass ihre Mitarbeiter:innen kompetent im Umgang mit KI sind.

Eine wesentliche Voraussetzung für die professionelle Auseinandersetzung mit diesen Technologien ist die Förderung von sogenannten Digital und AI Literacies. Das Konzept der Digital Literacies⁴ ist eng verwoben mit den Konzepten der Media Literacies sowie der Web Literacies und umfasst vielschichtige Kenntnisse und Fähigkeiten sowie einen reflektierten Umgang mit digitalen Technologien im Allgemeinen, um aktiv an digitalen Umwelten partizipieren zu können (Sackl-Sharif et al., 2023). AI Literacy⁵ kann als Teil von Digital Literacies angesehen werden und bezeichnet ein breites Wissen sowie vielseitige Fertigkeiten und Fähigkeiten, die einen grundlegenden Umgang mit KI ermöglichen, um Funktionsweisen zu verstehen und kritisch zu reflektieren sowie KI-generierte Lösungen interpretieren zu können (Long & Magerko, 2020). Dies beinhaltet, Wissen darüber zu haben, wie die Datengrundlage von KI-Systemen

4 Auch im *Future Skills Framework* 2021 werden Digital Literacies als sogenannte transformative Kompetenzen und digitale Schlüsselkompetenzen beschrieben (Stifterverband, 2021).

5 AI Literacy wird im deutschsprachigen Raum auch als KI-Kompetenz beschrieben (Knoth et al., 2024). AI Literacy ist ein breit angelegtes Konzept und bezeichnet eine neue Form von Kulturtechnik, die eine selbstbestimmte und kompetente Nutzung mit neuen KI-Technologien ermöglicht (Wienrich et al., 2022).

men beispielsweise deren Ergebnisse beeinflusst⁶ (Klinger & Sackl-Sharif, 2024). Eine Fachkraft, die über AI Literacy verfügt, ist also in der Lage, KI-Technologien kritisch zu hinterfragen und realistisch einschätzen zu können, wozu KI in der Lage ist und wo hier die Grenzen liegen. Sie hat ein fundiertes Verständnis dafür entwickelt, welche Tools sich für welche Tätigkeiten eignen, um deren Potenzial effektiv und reflektiert nutzen zu können. Durch den Erwerb von AI Literacy sollten Fachkräfte somit auch Situationen und Prozesse erkennen, in denen sie durch KI sinnvoll unterstützt werden können. Für den Erwerb von AI Literacy in der Sozialen Arbeit mit Kindern und Jugendlichen ist neben einem ganzheitlichen Verständnis von KI auch die Verbindung mit feldspezifischen Logiken als Voraussetzung für professionelles Handeln zentral (Knoth et al., 2024).

In diesem Zusammenhang erweist sich die Diskrepanz zwischen theoretischem Problembewusstsein und begrenztem Wissen in der Praxis als herausfordernd. So beurteilen Führungskräfte in der Sozialwirtschaft ihren persönlichen Kenntnisstand zum Thema KI mehrheitlich als mittelmäßig (Kreidenweis & Diepold, 2024).

2.2 Nutzung und Einsatz von KI

Zum Entstehungszeitpunkt dieses Beitrags ist der systematische und strukturierte Einsatz von KI in der Sozialen Arbeit mit Kindern und Jugendlichen (noch) nicht üblich, dennoch wird KI in der Praxis bereits vielfach genutzt (Gamble, 2020; Gillingham, 2006; Löhe, 2024; Schwartz et al., 2017). Daher werden nachfolgend wesentliche Befunde zu diesem Thema im Überblick dargestellt.

In zahlreichen Publikationen wird betont, dass KI-Tools vor allem bei Schreibtätigkeiten und bei Verwaltungsaufgaben unterstützend eingesetzt werden (Gutwald et al., 2021). Für die Sozialwirtschaft zeigen Kreidenweis und Diepold (2024), dass KI am häufigsten bei verwaltungs-/steuerungsorientierten Anwendungen genutzt wird. Auch andere Autor:innen halten fest, dass KI-Tools, vor allem für Tätigkeiten in der Administration, beispielsweise in der Abrechnung von Fällen, der Dokumentation, der Analyse abgeschlossener Fälle oder der Dienstplanerstellung, genutzt werden (Linnemann et al., 2023; Steiner & Tschopp, 2022; Schrödter et al., 2020). Bei diesen Schreibtätigkeiten kommen insbesondere generative KI-Tools wie ChatGPT oder Übersetzungstools wie DeepL bzw. Google Translate zum Einsatz (Rothballer & Zeiträg, 2024).

6 Besonders relevant im Umgang mit KI ist, dass durch KI-Anwendungen bestehende Stereotype und Diskriminierungen verstärkt werden (können) (Lucy & Bamman, 2021). Wissenschaftlich belegt ist, dass große KI-Modelle soziale Verzerrungen reproduzieren (Devinney et al., 2024; Hu et al., 2025). Einerseits sind die Trainingsdaten von KI-Systemen von gesellschaftlichen Differenzordnungen durchdrungen und bilden somit bestehende Diskriminierungen ab (Richter 2022). Andererseits sind KI-Tools auch keine neutralen Artefakte, sondern sie sind geprägt von den Bias ihrer Entstehungskontexte (Reidl et al., 2020).

Allerdings ist der Einsatz von KI vielfältig. So kann KI nicht nur für administrative Tätigkeiten, sondern auch für die Kommunikation mit Zielgruppen oder pädagogisch-fachlichen Tätigkeiten genutzt werden. KI wird im Rahmen der direkten Klient:innenarbeit allerdings eher zurückhaltend genutzt (Kreidenweis & Diepold, 2024; Stieler et al., 2025) und somit vermutlich eher als Arbeitsinstrumente verstanden. KI-Tools werden dennoch vereinzelt bereits bei pädagogisch-fachlichen Tätigkeiten unterstützend eingesetzt, z. B. als Assistenz in der psychosozialen Onlineberatung (Lehmann, 2024; Linnemann, et al., 2024). Erste Anwendungsüberlegungen gibt es hierfür auch in der Sozialen Arbeit mit Kindern und Jugendlichen. So führt die Delphi Gesellschaft für Forschung, Beratung und Projektentwicklung mbH das Projekt SuchtGPT (Gestaltung, Programmierung und Testung eines KI-basierten Chatbots für Suchtfragen) durch. Das Ziel besteht darin, mit einem Chatbot ein niedrigschwelliges Angebot zu schaffen, das bei Bedarf beispielsweise den Weg in die bestehenden Angebote der professionellen Suchthilfe ebnet (delphi, 2025). In diesem Kontext gilt es allerdings, professionsbezogene, datenschutzrechtliche und ethische Bedenken zu berücksichtigen (Bertsche & Como-Zipfel, 2017; Görder, 2024). Bei der Nutzung von digitalen Technologien insgesamt und KI-Tools im Speziellen ist außerdem die sogenannte intraprofessionelle digitale Kluft (Fiorentino et al., 2023; Wolf & Goldkind, 2016) wesentlich. Dieses Konzept macht deutlich, dass Digital Literacies, und nun vermehrt auch AI Literacy, innerhalb der Profession ungleich verteilt sind. Während einige Fachkräfte digitale Technologien souverän nutzen und ihre Potenziale erkennen, fühlen sich andere überfordert oder meiden digitale Medien gänzlich. Dies kann negative Folgen für die Chancengerechtigkeit und die Angebotsqualität haben (Bae et al., 2024).

In diesem Zusammenhang könnte sich die Frage nach der Haltung von Fachkräften gegenüber digitalen Transformationen im Allgemeinen und gegenüber KI-Entwicklungen im Speziellen als zentrales Element für professionelles Handeln erweisen. Die Analyse dieser Haltungen könnte dabei helfen, individuelles Handeln von Fachkräften zu verstehen, denn Einstellungen, Werthaltungen und subjektive Deutungsmuster beeinflussen maßgeblich, ob, wie und wofür digitale Technologien in der Sozialen Arbeit eingesetzt werden (Pulver, 2023).

3. Fragestellungen und Methodik

In Bezug auf die Nutzung von KI in der Sozialen Arbeit mit Kindern und Jugendlichen bleiben viele Fragen offen. So ist beispielsweise unklar, ob Ergebnisse aus anderen Bereichen (Kreidenweis & Diepold, 2024) auch auf dieses Handlungsfeld übertragbar sind. Ebenso bleibt die Frage offen, für welche Tätigkeiten KI derzeit unterstützend eingesetzt wird und ob diese Ergebnisse auch die Situationen in anderen Ländern im deutschsprachigen Raum widerspiegeln. Es zeichnet sich jedoch ab, dass KI-Ent-

wicklungen und deren Einsatz in der Sozialen Arbeit mit Kindern und Jugendlichen die Deutung und die Bedeutung von professionellem Handeln beeinflussen werden (Steiner, 2025).

Daher werden diese offenen Fragen im Folgenden aufgegriffen und es werden empirische Ergebnisse präsentiert, die die Sicht- und Nutzungsweisen von Fachkräften der Sozialen Arbeit mit Kindern und Jugendlichen aus Österreich und Deutschland ins Zentrum stellen. So sollen die Bedeutung von KI sowie allfällige Herausforderungen und Probleme damit im Bereich der Sozialen Arbeit mit Kindern und Jugendlichen exploriert werden. Mithilfe der quantitativen Befragungen werden die folgenden Fragestellungen adressiert:

- I. Wie ist das Interesse von Fachkräften am Thema KI ausgeprägt?
- II. Wie schätzen Fachkräfte ihr Wissen über KI ein?
- III. In welcher Form und in welchem Umfang nutzen Fachkräfte KI in ihrer beruflichen Praxis?
- IV. Welche Gründe führen dazu, dass Fachkräfte KI nicht nutzen?
- V. Welche Haltung nehmen Fachkräfte gegenüber der beruflichen Nutzung von KI ein?

Nachfolgend wird die Operationalisierung der verwendeten Variablen beschrieben, bevor die Stichproben dargestellt werden. Abschließend werden die Analysemethoden erläutert.

3.1 Messinstrumente

Zur Beantwortung der fünf Fragestellungen wurden neun Variablen zum Interesse und Wissen der Fachkräfte, zum Nutzungsverhalten, zu den Gründen der Nicht-Nutzung sowie der Haltung zu KI operationalisiert und in entsprechende Items überführt.⁷

Das Interesse an KI wurde sowohl für den privaten als auch für den beruflichen Kontext erfasst (siehe Tabelle 1). Die Messung erfolgte in Deutschland und in Österreich mit Hilfe einer fünfstufigen Skala, wobei die Benennung der Antwortkategorien in den Ländern nicht identisch war. Auch in der Erfassung des Wissens zu KI im beruflichen Kontext (siehe ebenfalls Tabelle 1) wurden in Deutschland und Österreich ähnliche, wenn auch nicht identische Formulierungen verwendet.

7 Diese neun Items wurden in zwei projekt- und länderspezifische standardisierte Online-Fragebogenerhebungen implementiert, die zudem länder- und projektspezifische Items enthielten. Aus diesem Grund variiert die Benennung der Antwortkategorien in den Fragebögen der jeweiligen Länder marginal.

Operationalisierung: Interesse und Wissen		
Land	Frage	Antwortkategorien
Interesse privat		
Ö	Ich interessiere mich für das Thema KI im privaten Kontext, in meiner Freizeit.	(1) Sehr stark (5) Überhaupt nicht, Keine Angabe
D		(1) Stimme voll und ganz zu (5) Stimme gar nicht zu
Interesse beruflich		
Ö	Ich interessiere mich für das Thema KI im beruflichen Kontext der Kinder- und Jugendarbeit.	(1) Sehr stark (5) Überhaupt nicht, Keine Angabe
D		(1) Stimme voll und ganz zu (5) Stimme gar nicht zu
Wissen		
Ö	Wie würden Sie Ihr Wissen über Künstliche Intelligenz und deren Anwendung in der Kinder- und Jugendarbeit einschätzen?	4-stufige kategoriale Einschätzung von „(Fast) kein Wissen“ bis „Fortgeschrittenes Wissen“, Keine Angabe
D		(1) Stimme voll und ganz zu (5) Stimme gar nicht zu

Tabelle 1: Operationalisierung der Variablen Interesse und Wissen, eigene Darstellung.

Die Nutzung von KI wurde durch den Einsatz von vier Variablen erfasst (siehe Tabelle 2). Zuerst wurde erfasst, ob KI im beruflichen Kontext genutzt wird. In der Erhebung in Österreich wurde hier eine dichotome Variable eingesetzt, wohingegen in Deutschland eine ordinale Variable in eine Dummy-Variable umkodiert wurde, um die Anzahl der Nutzenden zu ermitteln. In der Operationalisierung der Nutzungshäufigkeit wurde die Frage in einheitlichem Wortlaut gestellt, wobei sich die Benennung der Antwortmöglichkeiten zwischen den Ländern leicht unterschieden hat. Die Erhebung der zukünftigen Nutzungshäufigkeit und der Tätigkeitsbereiche, in denen KI eingesetzt wird, fand nur bei Personen statt, die KI bereits nutzen. Beide Variablen wurden in Deutschland und Österreich identisch erhoben.

Operationalisierung: Nutzung		
Land	Frage	Antwortkategorien
Nutzung		
Ö	Haben Sie in den letzten 12 Monaten bei Ihren beruflichen Tätigkeiten KI Anwendungen/-Tools genutzt?	Ja, Nein
D	Wie oft haben Sie in den letzten 12 Monaten KI-Tools für berufliche Zwecke genutzt?	Täglich oder fast täglich, Zumindest einmal pro Woche, Zumindest einmal pro Monat, Nur bei bestimmten Gelegenheiten bzw. Aufgaben, Nie, Keine Angabe
Nutzungshäufigkeit		
Ö	Wie oft haben Sie in den letzten 12 Monaten KI-Tools für berufliche Zwecke genutzt?	(1) Ich nutze KI-Tools täglich oder fast täglich (2) Ich nutze KI-Tools zumindest einmal pro Woche (3) Ich nutze KI-Tools zumindest einmal pro Monat (4) Ich nutze KI-Tools nur bei bestimmten Gelegenheiten (5) Keine Angabe
D		(1) Täglich oder fast täglich (2) Zumindest einmal pro Woche (3) Zumindest einmal pro Monat (4) Nur bei bestimmten Gelegenheiten/Aufgaben (5) Nie (6) Keine Angabe
Zukünftige Nutzung		
Ö & D	Planen Sie in den nächsten 1–2 Jahren KI häufiger, gleich oft, seltener oder gar nicht zu nutzen?	(1) Häufiger (4) Gar nicht, Keine Angabe
Tätigkeiten		
Ö & D	Für welche Tätigkeiten haben Sie KI in den letzten 12 Monaten genutzt?	10 Tätigkeiten + „Keine“, Mehrfachantworten möglich

Tabelle 2: Operationalisierung der Variablen, die die Nutzung von KI erfassen, eigene Darstellung.

In Tabelle 3 ist die Operationalisierung jener Variablen abgebildet, die die Gründe für eine Nicht-Nutzung erfasst. Diese Frage wurde nur für Personen angezeigt, die KI (noch) nicht nutzen, und war in Deutschland und Österreich identisch.

Operationalisierung: Nicht-Nutzung		
Land	Frage	Antwortkategorien
Ö & D	Aus welchen Gründen nutzen Sie für Ihre Arbeit KI derzeit nicht?	13 vorgegebene Gründe, „Ein anderer Grund“ und „Keine Angabe“ ebenfalls auswählbar, Mehrfachantworten möglich

Tabelle 3: Operationalisierung der Variable, die die Gründe einer Nicht-Nutzung von KI erfasst, eigene Darstellung.

Auch die Einstellung zu KI wurde in der Erhebung mithilfe einer Variable abgefragt (siehe Tabelle 4). Dieses Item richtete sich wieder an alle Fachkräfte, war jedoch in Deutschland und Österreich nicht komplett identisch.

Operationalisierung: Einstellung KI		
Land	Frage	Antwortkategorien
Ö	Glauben Sie, dass KI die Kinder- und Jugendarbeit verbessern kann?	(1) Ja, sicher (4) Nein, sicher nicht, (5) Ich weiß es nicht, Keine Angabe
D	Ich denke, dass KI die Schulsozialarbeit verbessern kann.	(1) Stimme voll und ganz zu (5) Stimme gar nicht zu

Tabelle 4: Operationalisierung der Variablen, die die Einstellung zu KI erfasst, eigene Darstellung.

3.2 Stichprobe

Die entwickelten Items wurden in zwei Online-Fragebogenerhebungen (Deutschland und Österreich) eingesetzt. Die Befragungen haben gemeinsam, dass sie sich an Fachkräfte der Sozialen Arbeit mit Kindern und Jugendlichen richteten und im Rahmen von Projekten durchgeführt wurden, welche den Einsatz von digitalen Technologien im jeweiligen Land adressieren. Somit enthielten die beiden Fragebögen – neben länderspezifischen Items – auch die neun entwickelten Items, wie sie in diesem Beitrag als empirische Basis verwendet werden. Aufgrund dieses methodischen Vorgehens werden im Folgenden die Stichproben beider Fragebogenerhebungen sequentiell beschrieben.

In Deutschland wurden die Items in einer Erhebung bei Fachkräften der Schulsozialarbeit in Deutschland eingesetzt. Die Fachkräfte hatten im Februar und März 2025 die Möglichkeit, an der Erhebung teilzunehmen. Die Befragung erfolgte über einen Zeitraum von sechs Wochen mit den Tools Unipark und SoSci Survey. Da die Erhebung als Sonderbefragung des KVJS-Landesjugendamtes in Baden-Württemberg durchgeführt wurde, konnte der Fragebogenlink über den Mailverteiler des Landesjugendamtes an alle Anstellungsträger mit der Bitte versendet werden, diesen an die Fachkräfte weiterzuleiten. Von 1.436 Teilnehmenden muss-

ten 109 aufgrund auffälliger Werte im Rahmen der Datenbereinigung ausgeschlossen werden, sodass eine bereinigte Stichprobe von $N = 1.327$ verbleibt. Der überwiegende Teil der Befragten war weiblich* (78%), gefolgt von männlichen* Teilnehmenden (22%) und einer kleinen Zahl von Personen, die sich als divers identifizieren (unter 1%). Das Alter der Befragten variierte zwischen 22 und 66 Jahren mit einem Mittelwert von $M = 41,24$ Jahren ($SD = 10,63$). Der durchschnittliche Stellenumfang lag im Bereich von 61–70%, wobei der häufigste Beschäftigungsgrad bei 41–50% lag.

Ein weiterer Fragebogen richtete sich an Fachkräfte mit einem Beschäftigungsverhältnis im Bereich der Kinder- und Jugendarbeit in Österreich und wurde im Rahmen des Projektes AI@youthwork (2024–2025)⁸ eingesetzt. An der standardisierten Online-Befragung, die im September und Oktober 2024 stattfand, nahmen $N = 303$ ⁹ Fach- und Führungskräfte der Kinder- und Jugendarbeit aus ganz Österreich teil. Als Zielgruppe des Projekts wurden Fach- und Führungskräfte in österreichischen Organisationen der Kinder- und Jugendarbeit definiert. Der Zugang im Rahmen der vorliegenden Befragung erfolgte mit Hilfe bestehender Netzwerke und Organisationen im Feld. In der erreichten Ad-hoc-Stichprobe finden sich 193 Frauen* und 69 Männer*, 3 Personen haben divers angegeben, eine Person hat keine Angaben zur Kategorie Geschlecht gemacht und 37 Personen haben diese Frage nicht beantwortet. Das Alter der Befragten variierte zwischen 22 und 64 Jahren mit einem Mittelwert von $M = 39,80$ Jahren ($SD = 10,10$). Das durchschnittliche Beschäftigungsausmaß lag im Bereich von 21 bis mehr als 30 Arbeitsstunden pro Woche, wobei das häufigste Beschäftigungsausmaß bei mehr als 30 Arbeitsstunden lag.

3.3 Analysemethode

Die Analyse der Daten sollte dazu beitragen, erste Erkenntnisse über den Umgang mit KI in der Sozialen Arbeit mit Kindern und Jugendlichen zu generieren. Daher haben die Ergebnisse der Befragungen explorativen Charakter und stellen nicht zuletzt aufgrund der nicht-repräsentativen¹⁰ Ad-hoc-Stichproben keinen Anspruch auf statistische Generalisierbarkeit. Die Auswertung der erhobenen Daten erfolgte daher ausschließlich mittels deskriptiver Verfahren. Ziel war es, zentrale Tendenzen und Verteilungen innerhalb der Datensätze sichtbar zu machen, ohne dabei inferenzstatistische Schlüsse zu ziehen. Dazu wurden Häufigkeiten, Mittelwerte sowie Streuungsmaße berechnet und tabellarisch aufbereitet. Die deskriptive Analyse ermöglichte eine erste strukturierende An-

8 Das Projekt „AI@youthwork: Zukunftsszenarien zur Anwendung von Artificial Intelligence in der Kinder- und Jugendarbeit“ (2024–2025) wurde an der Universität Graz durchgeführt und vom Zukunftsfonds der steiermärkischen Landesregierung finanziert. Nähere Informationen: <https://digitalesozialearbeit.github.io/ai-youthwork>

9 Dabei handelt es sich um die bereinigte Stichprobe.

10 Population: Soziale Arbeit mit Kindern und Jugendlichen in Deutschland bzw. Österreich.

näherung an das untersuchte Phänomen und gibt erste Einblicke in die Gemeinsamkeiten und Unterschiede im Zugang mit KI in der Sozialen Arbeit mit Kindern und Jugendlichen in Deutschland und Österreich. Die Ergebnisse liefern somit die Grundlage für weiterführende inhaltliche Interpretationen.

4. Ergebnisse

Nachfolgend werden die Ergebnisse der deskriptiven Analysen anhand der Fragestellungen des Beitrags thematisch gegliedert dargestellt. Zuerst werden Erkenntnisse bezüglich des Interesses und Wissens der Fachkräfte diskutiert, bevor das Nutzungsverhalten in den Blick genommen wird. Daran anschließend werden Barrieren für die Nutzung von KI beleuchtet, bevor die berufliche Einstellung der Fachkräfte zu KI nachgezeichnet wird.

Alle inhaltlichen Abschnitte sind in ihrer Struktur gleich aufgebaut. Zuerst werden die Ergebnisse der Befragung in Deutschland dargelegt, bevor die Ergebnisse der AI@youthwork-Studie aus Österreich veranschaulicht werden. Darauf aufbauend wird dann ein Vergleich der beiden Länder vorgenommen.

4.1 Interesse und Wissen

In Deutschland zeigt sich sowohl im beruflichen ($M = 2,94$, $SD = 1,28$, $N = 1.323$) als auch im privaten ($M = 2,83$, $SD = 1,31$, $N = 1.322$) Bereich ein moderates bis hohes Interesse an KI. 41,7% der befragten Fachkräfte geben an, (sehr) stark am Thema KI im beruflichen Kontext interessiert zu sein. Im privaten Bereich liegt dieser Anteil leicht höher bei 45,2%.

Auch in Österreich wurde das Interesse in beiden Bereichen abgefragt, wobei die Anteile an stark Interessierten geringer ausfallen. So geben 31,0% der Fachkräfte ein (sehr) starkes Interesse im beruflichen Kontext an ($M = 3,09$, $SD = 1,20$, $N = 296$), während für das Interesse im privaten Kontext 31,4% entsprechend antworteten ($M = 3,13$, $SD = 1,21$, $N = 296$).

Insgesamt lässt sich feststellen, dass das Interesse an KI unter Fachkräften in Deutschland stärker ausgeprägt ist als unter jenen in Österreich, wobei die Differenz im privaten Bereich etwas ausgeprägter erscheint als im beruflichen Kontext. Zudem ist festzustellen, dass es in der deskriptiven Betrachtung der Stichprobe aus Deutschland einen Unterschied im Interesse an KI im beruflichen und privaten Kontext gibt. Dieser zeigt sich bei den befragten Fachkräften in Österreich auf deskriptiver Ebene nicht.

In Bezug auf das selbst eingeschätzte Wissen über KI-Anwendungen im Kontext Sozialer Arbeit mit Kindern und Jugendlichen ($N = 1.321$) berichten in Deutschland lediglich 6,8% der Befragten, dass ihr Wissen

(eher) groß sei. Eine deutliche Mehrheit von 75,6% gibt an, dass ihr Wissen (eher) nicht groß ist, während 17,6% die mittlere Kategorie auf der fünfstufigen Skala wählen.

In der Befragung in Österreich wurde das Wissen kategorial erfasst ($N = 303$). 35,0% der Teilnehmenden geben an, über ein fortgeschrittenes oder grundlegendes Wissen zu KI im beruflichen Kontext zu verfügen. 56,4% haben zwar von KI gehört, wissen aber nichts über deren Anwendung in der Sozialen Arbeit mit Kindern und Jugendlichen. Weitere 4,3% geben an, (fast) gar nichts über KI zu wissen (Keine Angabe: 0,7%, Missing: 3,6%).

Trotz der Unterschiede in den verwendeten Antwortkategorien deuten die Ergebnisse darauf hin, dass die Fachkräfte in beiden Ländern ihr Wissen eher als gering einstufen. Während in Österreich immerhin ein Drittel der Fachkräfte eine solide Wissensbasis angibt, fällt der Anteil derjenigen mit (eher) großem Wissen in Deutschland deutlich geringer aus.

4.2 Berufliche Nutzung von KI

In Deutschland geben 54,4% der Befragten an, KI-Tools in den letzten 12 Monaten für berufliche Zwecke genutzt zu haben ($N = 1.306$), wodurch die Ergebnisse darauf hindeuten, dass KI zum Erhebungszeitpunkt noch nicht flächendeckend in der Praxis angekommen war. Unter denjenigen, die KI im beruflichen Kontext nutzen ($N = 730$, keine Angabe: 2,7%), liegt der Schwerpunkt der Nutzungshäufigkeiten bei Fachkräften, die KI nur bei bestimmten Gelegenheiten nutzen (52,6%). Eine gelegentliche Anwendung (18,6% monatlich und 19,5% wöchentlich) ist ebenfalls stark vertreten, wohingegen eine tägliche Nutzung (6,6%) die Ausnahme darstellt.

Bei den in Österreich befragten Fachkräften fällt der Anteil an Nutzenden etwas geringer aus: Hier geben 46,9% an, KI-Tools bereits in ihrer beruflichen Praxis genutzt zu haben ($N = 303$). Auch in Österreich ($N = 142$, Keine Angabe: 1,1%) dominiert die ausschließliche Nutzung bei bestimmten Gelegenheiten (57,0%). 14,8% der Nutzenden berichten von monatlicher Anwendung, 16,2% nutzen KI wöchentlich. Eine tägliche Nutzung (10,6%) wird kaum angegeben.

Damit ist der Anteil der beruflich mit KI arbeitenden Fachkräfte in Österreich auf deskriptiver Ebene im Vergleich zu in Deutschland befragten Fachkräften als niedriger einzustufen. Insgesamt zeigt sich aber in beiden Ländern, dass etwa die Hälfte der befragten Fachkräfte bisher noch keine praktischen Erfahrungen mit KI-Tools gesammelt hat. Somit wird in diesen deskriptiven Befunden deutlich, dass in der Sozialen Arbeit mit Kindern und Jugendlichen in beiden Ländern die Nutzungshäufigkeit niedrig und eher projektbezogen ist. Von einer Integration in den Arbeitsalltag kann daher noch nicht gesprochen werden.

4.2.1 KI-Einsatz nach Tätigkeiten

In Deutschland wurden KI-Tools insbesondere zur Erstellung von Texten (69,1%), zur Informationsrecherche (60,6%) sowie zum Brainstorming (56,0%) eingesetzt. Weniger häufig kamen sie für die Übersetzung in andere Sprachen (28,7%), die Zusammenfassung von Texten (24,9%) sowie für grafische Aufgaben oder Videoproduktionen (22,8%) zum Einsatz. Die Übersetzung in einfache Sprache (16,4%), die Analyse von Beratungsverläufen (3,8%) und die Zusammenfassung von Videoinhalten (1,5%) wurden nur selten als Anwendungsfälle genannt (Sonstige: 14,0%).

Ein ähnliches Bild ergibt sich in Österreich: Auch hier dominieren die Erstellung von Texten (71,1%), die Informationsrecherche (58,0%) sowie das Brainstorming (54,3%). Weniger häufig, aber dennoch bedeutsam, kamen sie für die Übersetzung in andere Sprachen (46,4%), die Zusammenfassung von Texten (44,2%) sowie für grafische Aufgaben oder Videoproduktion (31,9%) zum Einsatz. Die Übersetzung in einfache Sprache (20,3%), die Analyse von Beratungsverläufen (4,3%) und die Zusammenfassung von Videoinhalten (2,2%) wurden nur selten als Anwendungsfälle genannt (Sonstige: 15,2%).

Die Tätigkeitsprofile der KI-Nutzung sind somit in beiden Ländern weitgehend deckungsgleich. Allerdings zeigt sich in dieser deskriptiven Betrachtung, dass die Nutzung in Österreich etwas breiter ausfällt als unter den Befragten in Deutschland, da prozentual mehr von den insgesamt zehn möglichen Tätigkeitsbereichen ausgewählt wurden, wie sie im Fragebogen erfasst wurden.

4.2.2 Zukünftige Nutzung

Mit Blick auf die Zukunft plant ein erheblicher Teil der Fachkräfte in Deutschland, die KI bisher schon nutzen, eine intensivere Nutzung ($N = 706$, Keine Angabe: 11,3%): 52,0% geben an, KI künftig häufiger beruflich einsetzen zu wollen. Weitere 34,3% planen eine gleichbleibende Nutzung. Nur 1,0% wollen KI seltener und 1,4% gar nicht nutzen.

In Österreich sind die Zukunftspläne derjenigen, die KI schon nutzen, ähnlich ($N = 142$, Keine Angabe: 7,0%): 49,3% der Befragten beabsichtigen eine häufigere Nutzung. Weitere 40,1% geben an, KI künftig gleich oft nutzen zu wollen. Eine seltenere Nutzung kommt für 2,1% in Frage und nur 0,7% schließen eine zukünftige Nutzung aus.

In beiden Erhebungen zeigt sich ein Trend zur Ausweitung des KI-Einsatzes bei jenen Fachkräften, die KI bereits nutzen. Trotz derzeit geringer Nutzung wird KI zunehmend als relevantes Werkzeug für die eigene berufliche Praxis wahrgenommen. Dennoch ist mit Blick auf die Ergebnisse darauf hinzuweisen, dass die relative Häufigkeit in der Auswahl

„keine Angabe“ in beiden Erhebungen eher hoch ausfällt. Das könnte darauf hindeuten, dass einige Fachkräfte keine Angaben zum Blick in die Zukunft machen wollten oder konnten.

4.3 Gründe für die berufliche Nicht-Nutzung von KI

In Deutschland sind fehlendes Wissen (64,4%), Unklarheit über den Nutzen von KI (60,7%) und fehlende Kompetenz (60,6%) die Hauptgründe, weshalb Fachkräfte KI nicht nutzen (siehe Tabelle 5).

Gründe Nicht-Nutzung Deutschland	
Prozent	Grund
64,4%	Fehlendes Wissen über KI-Tools in der Schulsozialarbeit
60,7%	Unklarheit darüber, was KI für meine Arbeit leisten könnte
60,6%	Fehlende Kompetenzen für den Umgang mit KI-Tools
47,1%	Bedenken hinsichtlich Datensicherheit und Datenschutz durch KI
41,3%	Unklare rechtliche Verantwortung
39,4%	Keine Zeit, um mich mit dem Thema KI zu befassen
37,8%	Weniger menschliche Interaktionen mit Klient:innen durch KI
33,2%	Bedenken hinsichtlich der Qualität der Ergebnisse (Fehler, Irrtümer)
20,1%	Keine Lust, um mich mit dem Thema KI zu befassen
12,4%	Bedenken hinsichtlich der Verstärkung von Diskriminierung durch KI
10,2%	Bedenken hinsichtlich der Verdrängung von Arbeitsplätzen durch KI
6,4%	Ein anderer Grund
1,3%	KI-Anwendungen sind ein Hype, der wieder vorbeigeht

Tabelle 5: Gründe für die berufliche Nicht-Nutzung von KI in Deutschland, eigene Darstellung.

In Österreich zeigt sich ein ähnliches Muster: Auch hier sind fehlendes Wissen (69,9%) und die Unklarheit über den Nutzen von KI (63,0%) Hauptgründe, weshalb Fachkräfte KI nicht nutzen (siehe Tabelle 6).

Gründe Nicht-Nutzung Österreich	
Prozent	Grund
69,9%	Fehlendes Wissen über KI-Tools in der Kinder- und Jugendarbeit
63,0%	Unklarheit darüber, was KI für meine Arbeit leisten könnte
45,2%	Bedenken hinsichtlich Datensicherheit und Datenschutz durch KI
44,5%	Fehlende Kompetenzen für den Umgang mit KI-Tools
42,5%	Bedenken hinsichtlich der Qualität der Ergebnisse (Fehler, Irrtümer)
41,1%	Unklare rechtliche Verantwortung

Gründe Nicht-Nutzung Österreich	
Prozent	Grund
35,6%	Weniger menschliche Interaktion mit Klient:innen durch KI
26,0%	Keine Zeit, um mich mit dem Thema KI zu befassen
18,5%	Bedenken hinsichtlich der Verstärkung von Diskriminierung durch KI
14,4%	Keine Lust, um mich mit dem Thema KI zu befassen
11,6%	Bedenken hinsichtlich der Verdrängung von Arbeitsplätzen durch KI
4,1%	Ein anderer Grund
0,7%	KI-Anwendungen sind ein Hype, der wieder vorbeigeht

Tabelle 6: Gründe für die berufliche Nicht-Nutzung von KI in Österreich, eigene Darstellung

4.4 Verbesserungspotenzial durch KI

In Deutschland stimmen 31,3% (eher) der Aussage zu, dass KI die Schulsozialarbeit verbessern kann, während 36,5% der Aussage (eher) nicht zustimmen. Auf der fünfstufigen Skala haben 32,2% die mittlere Ausprägung gewählt und können daher in ihrer Haltung weder der positiven Sicht auf KI noch der negativen Sicht zugeordnet werden. Im Gegensatz dazu äußern sich in Österreich 44,7% (eher) zustimmend, 22,4% geben an, dass sie nicht sicher wissen, ob KI die Kinder- und Jugendarbeit verbessern kann und 23,4% äußern sich hier ablehnend.

Die Einstellung zur Rolle von KI in der sozialpädagogischen Praxis der Fachkräfte ist in beiden Befragungen ambivalent. Allerdings zeigt sich in dieser deskriptiven Betrachtung, dass in Deutschland, im Vergleich zu Österreich, weniger Fachkräfte eine positive Einstellung gegenüber KI haben und sogar mehr Fachkräfte eine negative Einstellung zeigen.

5. Zusammenfassende Betrachtung der Ergebnisse und Limitationen

5.1 Zusammenfassende Betrachtung der Ergebnisse

Nachfolgend werden in der tabellarischen Darstellung die Ergebnisse zu den unterschiedlichen Variablen zusammengefasst (Tabelle 7). Somit soll die Tabelle den Überblick über den Vergleich der Variablen zwischen der Schulsozialarbeit (SSA) in Deutschland und der Kinder- und Jugendarbeit (KJA) in Österreich erleichtern.

Übersicht der empirischen Ergebnisse		
Variable	Deutschland (SSA)	Österreich (KJA)
Interesse an KI (beruflich)	41,7% (sehr) stark interessiert $M = 2,94$ ($SD = 1,28$)	31,0% (sehr) stark interessiert $M = 3,09$ ($SD = 1,20$)
Interesse an KI (privat)	45,2% (sehr) stark interessiert $M = 2,83$ ($SD = 1,31$)	31,4% (sehr) stark interessiert $M = 3,13$ ($SD = 1,21$)
Wissen über KI	6,8% (eher) großes Wissen	35,0% großes Wissen
Nutzung von KI (letzte 12 Monate)	54,4% genutzt	46,9% genutzt
Nutzungshäufigkeit	6,6% täglich 52,6% bei Gelegenheit	10,6% täglich 57,0% bei Gelegenheit
Geplante Nutzungshäufigkeit	52,0% künftig häufiger 34,3% gleichbleibend	49,3% künftig häufiger 40,1% gleichbleibend
Tätigkeiten mit KI (Top 3)	Texterstellung: 69,1% Informationsrecherche: 60,6% Brainstorming: 56,0%	Texterstellung: 71,1% Informationsrecherche: 58,0% Brainstorming: 54,3%
Gründe für Nicht-Nutzung (Top 3)	Fehlendes Wissen: 64,4% Einsatzmöglichkeit unklar: 60,7% Fehlende Kompetenz: 60,6%	Fehlendes Wissen: 69,9% Einsatzmöglichkeit unklar: 63,0% Datenschutz: 45,2%
Einstellung zu KI	31,1% (eher) positiv 36,5% (eher) negativ	44,7% (eher) positiv 23,4% (eher) negativ

Tabelle 7: Vergleichende Ergebnisdarstellung Deutschland/Österreich, eigene Darstellung.

Die deskriptive Auswertung zeigt, dass das Interesse an KI unter den befragten Fachkräften in Deutschland insgesamt, insbesondere im privaten Kontext, höher ist als in Österreich. In Deutschland lässt sich zudem ein Unterschied zwischen privatem und beruflichem Interesse an KI erkennen, der in der österreichischen Erhebung nicht deutlich wird. Gleichzeitig deuten die Ergebnisse auf Wissensdefizite hin. Während in Österreich etwa ein Drittel der Befragten angibt, über großes Wissen zu verfügen, fällt dieser Anteil in Deutschland deutlich geringer aus.

Die tatsächliche Anwendung von KI erfolgt bislang nur zu bestimmten Gelegenheiten und ist daher noch nicht in den Arbeitsalltag integriert. In der österreichischen Befragung liegt der Anteil aktiv mit KI arbeitender Fachkräfte unter jenem der deutschen Befragten, wobei in beiden Ländern etwa die Hälfte noch keine praktischen Erfahrungen mit KI gesammelt hat. Dennoch zeigt sich ein Trend zur Ausweitung der KINutzung: Fachkräfte, die KI bereits nutzen, sehen in beiden Ländern zunehmend Potenziale für den eigenen Berufsalltag.

Die Anwendungsbereiche von KI sind in beiden Ländern ähnlich. Die Texterstellung führt in beiden Ländern, gefolgt von Informationsrecherche und Brainstorming. Diese Übereinstimmungen deuten auf universelle Einsatzfelder in der Sozialen Arbeit mit Kindern und Jugendlichen hin, wobei in der österreichischen Praxis eine etwas breitere Nutzung über mehrere Aufgabenbereiche hinweg beobachtet werden kann.

Auch die zentralen Hemmnisse für den aktiven Einsatz von KI unterscheiden sich nur teilweise: Beide Länder nennen fehlendes Wissen sowie unklare Einsatzmöglichkeiten als Hauptgründe. Auf den Plätzen 3 bzw. 4 liegen sowohl in Österreich als auch in Deutschland fehlende Kompetenzen sowie Datenschutzbedenken. Die Ergebnisse verdeutlichen somit, dass in beiden Ländern insbesondere Gründe ausgewählt wurden, die auf individueller Ebene angesiedelt sind (fehlendes Wissen, fehlende Kompetenzen). Darüber hinaus verweisen die hohen Werte in Bezug auf Datenschutzbedenken auf ein ausgeprägtes Bewusstsein für Datenschutzrisiken, die vor dem Hintergrund der DSGVO und des AI-Acts sowie der besonderen Sensibilität von Klient:innendaten in der Sozialen Arbeit mit Kindern und Jugendlichen besonders relevant sind. Gleichzeitig werden aber auch Unsicherheiten im Umgang mit diesen Herausforderungen deutlich. Zudem wird in beiden Ländern die unklare rechtliche Verantwortung als Grund der Nicht-Nutzung genannt. Dies kann darauf hindeuten, dass Fachkräfte sich unsicher fühlen, welche Haftungsrisiken mit der KI-Nutzung verbunden sind. Hier könnten die rechtliche Absicherung und ausreichende Informationen zum Datenschutz Fachkräfte auf dem Weg zur Nutzung von KI unterstützen.

Hinsichtlich der Einstellung gegenüber KI zeigt sich ein ambivalentes Bild. Während Fachkräfte in beiden Ländern sowohl Potenziale als auch Risiken wahrnehmen, ist die Haltung der in Deutschland befragten Fachkräfte insgesamt zurückhaltender. Dort äußern mehr Fachkräfte Skepsis oder eine ablehnende Haltung gegenüber KI als in Österreich.

5.2 Limitationen der komparativen Studie

Mit diesen Erkenntnissen liefert die vorliegende Untersuchung erste deskriptive Einblicke in die Wahrnehmung und Nutzung von KI in der Sozialen Arbeit mit Kindern und Jugendlichen in Deutschland und Österreich. Dabei sind jedoch einige zentrale Limitationen zu beachten, die die Aussagekraft und Vergleichbarkeit der Ergebnisse einschränken.

In beiden Online-Umfragen wurde der Begriff KI sehr breit verwendet. Er bezog sich sowohl auf generative KI-Tools als auch auf algorithmische Mustererkennung (Brossmann-Handler et al., 2025). Dies wurde in den Items nicht gesondert ausgewiesen. Nur an manchen Stellen, etwa bei der

Nutzung spezifischer Tools, können im Nachhinein Bezüge dazu hergestellt werden. Bei anderen Items ist hingegen unklar, auf welche Art von KI sich die Fachkräfte beim Beantworten der Fragen bezogen.

In Anbetracht der methodischen Vorgehensweise zur empirischen Fundierung dieses Beitrags ist darauf hinzuweisen, dass die Stichproben sehr heterogen sind. Dies erschwert die direkte Vergleichbarkeit der Ergebnisse. Auch wenn die Ergebnisse auf Gemeinsamkeiten und Unterschiede der befragten Fachkräfte der Sozialen Arbeit mit Kindern und Jugendlichen in Österreich und Deutschland hinweisen und wichtige Tendenzen aufzeigen, muss berücksichtigt werden, dass sie sich auf ein breites Handlungsfeld beziehen. Obwohl wir in dieser Studie aufgrund gleicher handlungsleitender Konzepte Schulsozialarbeit und Kinder- und Jugendarbeit als Teil der Sozialen Arbeit mit Kindern und Jugendlichen verstehen, sollten die Spezifika hinsichtlich fachlicher und struktureller Rahmenbedingungen innerhalb dieses Feldes nicht außer Acht gelassen werden.

Zudem variieren die Stichprobengrößen deutlich. Während die Erhebung in Österreich keine repräsentativen Rückschlüsse auf das gesamte Feld der Kinder- und Jugendarbeit erlaubt, ist die Befragung in der Schulsozialarbeit auf Fachkräfte aus Baden-Württemberg beschränkt. Eine Generalisierung der Befunde auf die Soziale Arbeit mit Kindern und Jugendlichen in Österreich und Deutschland ist daher nicht möglich. Vor diesem Hintergrund handelt es bei dieser komparativen Untersuchung um eine deskriptive und explorativ angelegte Untersuchung. Dementsprechend werden keine inferenzstatistischen Aussagen über die Grundgesamtheiten getroffen. Bei der Interpretation der Ergebnisse müssen somit die methodischen Unterschiede innerhalb der Studie berücksichtigt werden. Die Vergleichbarkeit der Befunde ist durch die verschiedenen Handlungsfelder, Länderkontexte und Rekrutierungsverfahren eingeschränkt, was sich auch in den unterschiedlichen Stichprobengrößen widerspiegelt.

Diese Limitationen eröffnen zugleich Perspektiven für zukünftige Forschung. So wären Anschlussstudien mit größeren und repräsentativen Stichproben notwendig, um inferenzstatistische Analysen zu ermöglichen. Besonders relevant wäre dabei die gezielte Untersuchung gruppenspezifischer Unterschiede, etwa hinsichtlich Geschlecht, Alter, beruflicher Rolle (Fachkraft vs. Führungskraft) oder tatsächlicher Nutzung von KI (Nutzer:innen vs. Nicht-Nutzer:innen). Darüber hinaus wäre eine methodische Erweiterung durch qualitative Interviews denkbar, um vertiefte Einblicke in subjektive Haltungen, Nutzungserfahrungen und hemmende wie fördernde Faktoren zu gewinnen. Ein Mixed-Methods-Ansatz könnte dabei helfen, die bisherigen quantitativen Befunde sinnvoll zu ergänzen und ein differenzierteres Verständnis für den Umgang mit KI in der Sozialen Arbeit mit Kindern und Jugendlichen zu entwickeln.

6. Ausblick

Die Ergebnisse dieser komparativen Analyse zeigen, dass ein starkes Interesse an KI besteht. Fachkräfte nutzen KI in ihrem Arbeitsalltag, wobei fehlendes Wissen und fehlende Kompetenzen die Nutzung hemmt. Dies lässt darauf schließen, dass für Fachkräfte der Sozialen Arbeit mit Kindern und Jugendlichen der Erwerb von AI Literacy eine große Rolle spielt und feldspezifische KI-Fortbildungskonzepte an Relevanz gewinnen. In Zusammenhang mit der Nutzung von KI wird immer wieder das Phänomen des *Deskilling*s thematisiert, also das Verlernen bestimmter Tätigkeiten, weil diese zunehmend von Systemen unterstützt oder vollständig übernommen werden (Farhan, 2023). In Hinblick auf diese Ergebnisse scheinen aus unserer Sicht Fragen nach dem *Deskilling* in den Hintergrund zu treten und Überlegungen zu *Upskilling* im Sinne einer Weiter- oder Höherqualifizierung durch den Erwerb von AI Literacy in den Vordergrund zu rücken.

Neben Hochschulausbildungen und praktischer Berufstätigkeit kommt im Kontext der Sozialen Arbeit mit Kindern und Jugendlichen auch der Weiterbildung eine zentrale Rolle zu (Sackl-Sharif & Klinger, i. E.; Schulze-Krüdener, 2012). Relevant erscheint ein ausgewogenes Verhältnis zwischen der Nutzung von KI und der Aufrechterhaltung der Kernelemente professionellen Handelns. Daneben ist der Erwerb eines technischen Verständnisses und kritischer reflexiver Fähigkeiten in Bezug auf technologische Neuerungen bedeutsam, wie z. B. humanistische und ethische Grundlagen (vgl. Wirtz et al., 2019). Neben individuell geprägtem Wissen und Faktoren wird das Handeln der Fachkräfte auch durch organisationale Strukturen geprägt, die digitale Praktiken und den Erwerb von AI Literacy ermöglichen, begrenzen oder steuern (Helbig et al., 2021; Iske & Kutscher, 2020; Kutscher, 2024; Neumann, 2022). Deshalb sollte betont werden, dass der Erwerb dieser Literacy nicht zusätzlich zum bereits hohen Arbeitspensum von Fachkräften erfolgen sollte. Vielmehr müssen die jeweiligen Organisationen Zeit und Ressourcen für informelle und formale Lernsettings bereitstellen (Klinger et al., 2025). Fachkräfte sollen für die Potenziale und Schwächen von KI-Tools sensibilisiert werden. Sie müssen die Funktionsweise von KI-Tools verstehen, die Ergebnisse kritisch hinterfragen und die Grenzen der Technologie erkennen. Hierzu zählt auch, über Diskriminierungspotenziale von KI-Tools, Halluzinationen bei generativen KI-Tools sowie typische Mensch-Maschine-Interaktionen wie die Automatisierungsverzerrung aufzuklären und über Mythen und Fakten von KI zu informieren. Nur so kann es gelingen, KI-Tools reflektiert in der Arbeitspraxis einzusetzen.

Literatur

- Auferbauer, M., Gspurning, W., Heimgartner, A., Lederer-Hutsteiner, T., & Mayr, A. (2019). *Evaluation der Schulsozialarbeit in der Steiermark 2015 – 2018*. Land Steiermark – Bildung und Gesellschaft.
- Bae, J., Lee, S., & Choi, C. J. W. (2024). Expansion of digital technology use in the Korean social work field. *Journal of Social Service Research*, 1–15. <https://doi.org/10.1080/01488376.2024.2402520>
- Bertsche, O., & Como-Zipfel, F. (2017). Sozialpädagogische Perspektiven auf die Digitalisierung. *Soziale Passagen*, 8(2), 235–254.
- Brossmann-Handler, E., Sackl-Sharif, S., Klinger, S., & Mayr, A. (2025). *Künstliche Intelligenz in der Sozialen Arbeit*. <https://digitalesozialarbeit.github.io/ai-socialwork>
- delphi. (2025). *SuchtGPT. Ein Chatbot für Suchtfragen?* <https://suchtgpt.delphi.de>
- Devinney, H., Björklund, J., & Björklund, H. (2024). We don't talk about that: Case studies on intersectional analysis of social bias in large language models. In *Proceedings of the 5th Workshop on Gender Bias in Natural Language Processing (GeBNLP)* (S. 33–44). Association for Computational Linguistics.
- EU AI Act. (2024). *Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. <https://artificialintelligenceact.eu/de>
- Farhan, A. (2023). The impact of artificial intelligence on human workers. *Journal of Communication Education*, 17(2), 93–104.
- Fiorentino, V., Romakkaniemi, M., Harrikari, T., Saraniemi, S., & Tiitinen, L. (2023). Towards digitally mediated social work – the impact of the COVID-19 pandemic on encountering clients in social work. *Qualitative Social Work*, 22(3), 448–464. <https://doi.org/10.1177/14733250221075603>
- Gamble, A. (2020). Artificial intelligence and mobile apps for mental healthcare: A social informatics perspective. *Aslib Journal of Information Management*, 72(4), 509–523. <https://doi.org/10.1108/AJIM-11-2019-0316>
- Gillingham, P. (2006). Risk assessment in child protection: Problem rather than solution? *Australian Social Work*, 59(1), 86–98. <https://doi.org/10.1080/03124070500449804>
- Görder, B. (2021). Die Macht der Muster. Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch „künstliche Intelligenz“. *EthikJournal*, 7(2), 1–10.
- Görder, B. (2024). Ethische Fragen beim Einsatz von KI in sozialen Dienstleistungen. In H. Kreidenweis (Hrsg.), *KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis* (S. 175–190). Nomos Verlag.

- Helbig, C., Kutscher, N., & Unterkofler, U. (2021). Profession – Organisation – digitale Medien. Methodologische Überlegungen zur praxeologischen Erforschung von Phänomenen der Digitalisierung im Feld Sozialer Arbeit. *Neue Praxis*, *5*, 432–457.
- Hu, T., Kyrychenko, Y., Rathje, S., Roozenbeek, J., Hancock, J. T., Riedl, M., & Van Bavel, J. J. (2025). Generative language models exhibit social identity biases. *Nature Computational Science*, *5*, 65–75. <https://doi.org/10.1038/s43588-024-00741-1>
- Iske, S., & Kutscher, N. (2020). Digitale Ungleichheiten im Kontext Sozialer Arbeit. In N. Kutscher, T. Ley, U. Seelmeyer, F. Siller, A. Tillmann, & I. Zorn (Hrsg.), *Handbuch Soziale Arbeit und Digitalisierung* (S. 115–128). Beltz Juventa.
- Klinger, S., & Sackl-Sharif, S. (2024). Vom verantwortungsvollen Umgang mit Künstlicher Intelligenz in der Sozialen Arbeit. *Sozialpädagogische Impulse*, *4*.
- Klinger, S., Mayr, A., Sackl-Sharif, S., & Brossmann-Handler, E. (2023). Zur (nicht) selbstverständlichen Nutzung digitaler Technologien in der Sozialen Arbeit. Ergebnisse einer österreichweiten Befragung von Fachkräften. *Soziale Arbeit. Zeitschrift für soziale und sozialverwandte Gebiete*, *72*(6), 202–211.
- Klinger, S., Sackl-Sharif, S., & Mayr, A. (2025). Ortsunabhängiges und zeitflexibles Arbeiten: Die digitalisierte Soziale Arbeit im Spannungsfeld von organisationalen Rahmenbedingungen und individuellen Strategien. In M. Opielka & C. Erfurth (Hrsg.), *Soziale Digitalisierung. Perspektiven der Sozialpolitik* (S. 171–192). Springer VS. https://doi.org/10.1007/978-3-658-46328-1_9
- Knoth, N., Decker, M., Laupichler, M. C., Pinski, M., Buchholtz, N., Bata, K., & Schultz, B. (2024). Developing a holistic AI literacy assessment matrix – Bridging generic, domain-specific, and ethical competencies. *Computers and Education Open*, *6*, 100177. <https://doi.org/10.1016/j.caeo.2024.100177>
- Kreidenweis, H. (Hrsg.). (2024). *KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis*. Nomos Verlag.
- Kreidenweis, H., & Diepold, M. (2024). *Studie Künstliche Intelligenz in der Sozialwirtschaft*. https://sozialinformatik.de/fileadmin/1805/pdf_documents/Forschung_und_Entwicklung/Studie-KI-Sozialwirtschaft-2024.pdf
- Kutscher, N. (2024). Digitalität und Digitalisierung als Gegenstand der Sozialen Arbeit. In T. Kurtz, D. M. Meister, & U. Sander (Hrsg.), *Digitale Medien und die Produktion von Wissenschaft: Wissenschaftliches Forschen, Schreiben und Publizieren in den Geistes- und Sozialwissenschaften unter den Bedingungen des digitalen Wandels* (Bd. 47, S. 121–144). Springer VS. <https://doi.org/10.1007/978-3-658-42542-5>

- Lehmann, R. (2024). KI in der digitalen Beratung. In H. Kreidenweis (Hrsg.), *KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis* (S. 45–54). Nomos Verlag.
- Linnemann, G. A., & Löhe, J., & Rottkemper, B. (2023). Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit. *Soziale Passagen*, *15*, 197–211. <https://doi.org/10.1007/s12592-023-00455-7>
- Linnemann, G., Löhe, J., & Rottkemper, B. (2024). Bedeutung von Selbstoffenbarungseffekten in quasisozialen Beziehungen mit auf generativer KI basierten Systemen in Settings von Onlineberatung und -therapie. *e-beratungsjournal.net – Zeitschrift für Onlineberatung und computervermittelte Kommunikation*, *20*(1), 1–21. <https://doi.org/10.48341/9x1s-5y11>
- Löhe, J. (2024). KI-Anwendung in der Kinder- und Jugendhilfe. In H. Kreidenweis (Hrsg.), *KI in der Sozialwirtschaft. Eine Orientierungshilfe für die Praxis* (S. 101–116). Nomos Verlag.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competences and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (S. 1–16). ACM. <https://doi.org/10.1145/3313831.3376727>
- Lucy, L., & Bamman, D. (2021). Gender and representation bias in GPT-3 generated stories. In *Proceedings of the 3rd Workshop on Narrative Understanding* (S. 48–55). Association for Computational Linguistics.
- Meyer, T., & Patjens, R. (Hrsg.). (2020). *Studienbuch Kinder- und Jugendarbeit*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-24203-9_1
- Neumann, M. (2022). Bedeutung digitaler Medien in der Kinder- und Jugendarbeit: Digitalisierung und soziale Teilhabe als Herausforderung für Pädagog_innen. *Sozial Extra*, *46*(6), 458–465. <https://doi.org/10.1007/s12054-022-00534-8>
- Pulver, C. (2023). *Die Mediatisierung professionellen Handelns in den Arbeitsfeldern der Sozialen Arbeit: Zur Situation der Fachkräfte in der Schweiz* [Dissertation]. Johannes-Gutenberg-Universität Mainz.
- Reidl, S., Streicher, J., Hoch, M., Hausner, B., Waibel, G., & Gürtl, F. (2020). *Digitale Ungleichheit. Wie sie entsteht, was sie bewirkt ... und was dagegen hilft*. https://www.ffg.at/sites/default/files/allgemeine_downloads/strukturprogramme/Laura%20Bassi%204.0/Studie_Digitale_Ungleichheit_barrierefrei_final.pdf
- Richter, S. (2022). Perspektiven feministischer Digitalisierungsforschung: Querschnittsthemen gesellschaftlicher Transformation im komplexen Zusammenspiel. In A. Kasten, K. Von Bose, & U. Kalender (Hrsg.), *Feminismen in der Sozialen Arbeit. Debatten, Dis/Kontinuitäten, Interventionen* (S. 188–205). Beltz Juventa.

- Rothballer, M., & Zeitrüg, M. (2024). Künstliche Intelligenz als Chance für Soziale Arbeit. Ein Werkstattbericht der Diakonie Rosenheim zu transformativen Ansätzen und Anwendungen. *unsere jugend*, 76(7+8), 328–337.
- Sackl-Sharif, S., Klinger, S., Mayr, A., & Brossmann-Handler, E. (2023). Aneignungsmöglichkeiten von Digital Literacies im Feld der Sozialen Arbeit. Perspektiven von Fachkräften und Handlungsempfehlungen für Organisationen. *Der Pädagogische Blick*, 31(4), 238–248.
- Sackl-Sharif, S., Klinger, S., Mayr, A., & Brossmann-Handler, E. (2025). *KI in der Sozialen Arbeit. Vier (zukünftige) Anwendungsfelder*. <https://digitalesozialarbeit.github.io/ai-socialwork/ai-application-areas>
- Sackl-Sharif, S., & Klinger, S. (im Erscheinen). Partizipative Entwicklung eines Weiterbildungslehrgangs zu den Themen Digitalisierung und Künstlicher Intelligenz in der Kinder- und Jugendarbeit. Potenziale für die Gestaltung gender- und diversitätsreflektierender Bildungsräume. In C. Aspalter, B. Herzog-Punzenberger, N. Kraker, & G. Wagner (Hrsg.), *Diversität diskursiv. Bildungswissenschaftliche Zugänge und pädagogische Perspektiven*. Barbara Budrich.
- Schrödter, M., Bastian, P., & Taylor, B. (2020). Risikodiagnostik und Big Data Analytics in der Sozialen Arbeit. In N. Kutscher, T. Ley, U. Seelmeyer, F. Siller, A. Tillmann, & I. Zorn (Hrsg.), *Handbuch Soziale Arbeit und Digitalisierung* (S. 255–264). Beltz.
- Schwartz, I. M., York, P., Nowakowski-Sims, E., & Ramos-Hernandez, A. (2017). Predictive and prescriptive analytics, machine learning and child welfare risk assessment: The Broward County experience. *Children and Youth Services Review*, 81, 309–320. <https://doi.org/10.1016/j.childyouth.2017.08.020>
- Steiner, O. (2025). Fachliche Standards der Kinder- und Jugendarbeit in einer Kultur der Digitalität. In E. Rösch & N. Brüggem (Hrsg.), *Praxishandbuch. (Post-)Digitale Kinder- und Jugendarbeit* (S.178–189). Belz Juventa.
- Steiner, O., & Tschopp, D. (2022). Künstliche Intelligenz in der Sozialen Arbeit. *Sozial Extra*, 46, 466–471. <https://doi.org/10.1007/s12054-022-00546-4>
- Stieler, M., Berger, J., & Lehmann, R. (2025). Technologische Innovation in der psychosozialen Onlineberatung: Eine praxisorientierte Perspektive auf die Anwendung von Künstlicher Intelligenz. In J. Späte, C. Endter, D. Stix, & K. Krauskopf (Hrsg.), *#GesellschaftBilden im Digitalzeitalter. Perspektiven Sozialer Arbeit auf technologische Herausforderungen* (S. 179–206). Waxmann.
- Stifterverband für die Deutsche Wissenschaft e. V. (2021). *Future Skills 2021. 21 Kompetenzen für eine Welt im Wandel* (Diskussionspapier Nr. 3). <https://www.stifterverband.org/medien/future-skills-2021/>

- Thole, W. (2012). Die Soziale Arbeit – Praxis, Theorie, Forschung und Ausbildung. Versuch einer Standortbestimmung. In W. Thole (Hrsg.), *Grundriss Soziale Arbeit. Ein einführendes Handbuch* (S. 19–72). VS Verlag für Sozialwissenschaften/Springer Fachmedien
- UNESCO. (2023). *AI and gender equality. A global study on the gendered impacts of artificial intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000388971>
- Wienrich, C., Carolus, A., Markus, A., & Augustin, Y. (2022). *AI Literacy: Kompetenzdimensionen und Einflussfaktoren im Kontext von Arbeit* (Working Paper Nr. 1). KI-Observatorium des Bundesministeriums für Arbeit und Soziales.
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector – applications and challenges. *International Journal of Public Administration*, *42*(7), 596–615. <https://doi.org/10.1080/01900692.2018.1498103>
- Wolf, L., & Goldkind, L. (2016). Digital native meet friendly visitor: A Flexner-inspired call to digital action. *Journal of Social Work Education*, *52*, 99–109. <https://doi.org/10.1080/10437797.2016.1174643>

André Heinz¹ und Marcel Pietsch²

„Ceci n'est pas une réalité“ – Zur Dekonstruktion von Wirklichkeit in einer magrittesken Realität von Filterblasen und KI-Algorithmen

¹ IU Internationale Hochschule

² FOM – Hochschule für Oekonomie und Management

Zusammenfassung

In einer durch Algorithmen strukturierten Medienwelt scheinen zunehmend die Grenzen zwischen realer Erfahrung und digitaler Repräsentation zu verschwimmen, wie verschiedene Studien zur Mediatisierung und digitalen Sozialisation zeigen (z. B. Paus-Hasebrink et al., 2019). Soziale Plattformen – so die zentrale These des Beitrags – erzeugen eine magritteske Realität, in der Inszenierung, Filterblasen und performative Identitäten das Selbst- und Weltverhältnis prägen. Der Beitrag analysiert diese Entwicklungen medienpädagogisch und sozialarbeiterisch: Er zeigt, wie digitale Resonanzökonomien die Identitätsarbeit beeinflussen, thematisiert die Rolle ästhetischer Bildung als Gegenstrategie und reflektiert die Verantwortung Sozialer Arbeit im digitalen Raum.

Schlüsselwörter: Filterblasen, Ästhetische Bildung, Medienkompetenz, Identität, Resonanzökonomie

„Ceci n'est pas une réalité“ – On the Deconstruction of Reality in a Magrittesque World of Filter Bubbles and AI Algorithms

Abstract

In a media environment structured by algorithms, the boundaries between real experience and digital representation appear to blur increasingly, as various studies on mediatization and digital socialization have shown (e. g. Paus-Hasebrink et al., 2019). Social platforms construct a Magritte-like reality, in which digital staging, filter bubbles, and performative identities shape individuals' sense of self and their relation to the world. This article analyzes

Heinz, A., & Pietsch, M. (2026). „Ceci n'est pas une réalité“ – Zur Dekonstruktion von Wirklichkeit in einer magrittesken Realität von Filterblasen und KI-Algorithmen. *DiMe-Sozial*, 2(1), 185–200. <https://doi.org/10.21240/dimesoz/2026/1/65>

these developments from a media pedagogical and social work perspective, examining how digital resonance economies influence identity formation, discussing aesthetic education as a critical counter-strategy, and reflecting on the professional responsibility of social work in digital spaces.

Keywords: Filter bubbles, Aesthetic education, Media literacy, Identity, Economy of resonance

1. Zwischen Bildern und Wirklichkeit

Als der Maler René Magritte in sein Gemälde einer Pfeife *La trahison des images*¹ (1929) den Satz „Ceci n’est pas une pipe“² einfügte, forderte er die Betrachtenden heraus, die Differenz zwischen Objekt und Darstellung zu erkennen. Es handelt sich um ein Abbild – nicht um die reale Pfeife. Diese scheinbar triviale Erkenntnis hat im digitalen Zeitalter eine neue Brisanz gewonnen: Auch heute werden uns tagtäglich Darstellungen präsentiert, die auf den ersten Blick echt erscheinen, jedoch Konstruktionen sind. In digitalen Medien begegnen wir Bildern und Inhalten, die Wirklichkeit suggerieren, sie aber zugleich formen und verändern. In sozialen Netzwerken begegnen Nutzer:innen täglich medial erzeugten Darstellungen von Wirklichkeit, die stark von Algorithmen beeinflusst sind. Diese Systeme bestimmen, welche Inhalte sichtbar werden und welche nicht. Digitale Repräsentationen erscheinen personalisiert, authentisch und echt, sind jedoch oft Teil einer von Plattformlogiken strukturierten Realität (Bucher, 2018; Iske & Kutscher, 2020).

In einer *magrittesken* Wirklichkeit droht die Unterscheidung zwischen Bild, Zeichen und Realität zu verschwimmen. Aus der Perspektive des *Critical Realism* (Ziegler, 2008, S. 48) lässt sich dieses Phänomen als Ausdruck einer zunehmenden Entkopplung zwischen ontologischer Realität und ihren symbolischen Repräsentationen verstehen. Ziegler betont, dass gesellschaftliche Wirklichkeit nicht unmittelbar erfahrbar ist, sondern stets durch diskursive und mediale Vermittlungsprozesse konstruiert wird. In einer medial geprägten Umgebung wird dadurch die Differenz zwischen dem, was real ist, und dem, was als real erscheint, immer schwerer aufrechtzuerhalten. Die *magritteske* Irritation verweist somit auf eine erkenntnistheoretische Verschiebung, in der die symbolische Darstellung nicht mehr auf eine zugrunde liegende Realität verweist, sondern selbst zum Ort der Wirklichkeitsproduktion wird.

Die Auseinandersetzung mit digitalen Wirklichkeiten erweist sich für viele Menschen als anspruchsvoll, da Diskrepanzen zwischen medialer Darstellung und gelebter Erfahrung häufig schwer zu erkennen sind. Dies gilt insbesondere, wenn der Umgang mit digitalen Inhalten nur eingeschränkt auf kritische und medienpädagogisch fundierte Reflexionsprozesse zurück-

1 Deutsch: Der Verrat der Bilder.

2 Deutsch: Dies ist keine Pfeife.

greift. Die ständige Konfrontation mit inszenierten Bildern, gefilterten Erzählungen und algorithmisch strukturierter Sichtbarkeit prägt individuelle und kollektive Wahrnehmungsweisen, Selbstkonzepte sowie soziale Orientierungen (Neumann, 2022).

Digitale Kommunikationsräume sind somit nicht allein technologische Phänomene, sondern stellen auch eine pädagogische und sozialwissenschaftliche Herausforderung dar. Aus Perspektive der Sozialen Arbeit erfordert dies eine reflexive Professionalität, die sich mit Prozessen medialer Konstruktion, gesellschaftlicher Machtverhältnisse und digitaler Teilhabe kritisch auseinandersetzt (Marotzki & Jörissen, 2009; Hug, 2016). Soziale Arbeit kann in diesem Kontext Räume schaffen, in denen mediale Wirklichkeiten hinterfragt, differenzierte Wahrnehmungen entwickelt und Formen kritischer Medienkompetenz gefördert werden. Damit leistet sie einen Beitrag zur Stärkung von Selbstvergewisserung, Urteilsfähigkeit und digitaler Handlungsautonomie.

2. Digitale Realitäten: Zwischen Filterblasen und algorithmischer Konstruktion

In der dynamisierten Mediengesellschaft konfrontieren digitale Plattformen ihre Nutzer:innen nicht nur mit einer Vielfalt an Informationen, sondern auch mit selektiven und nicht selten verzerrten Wirklichkeitsausschnitten. Diese Selektion erfolgt jedoch seltener bewusst durch die Rezipient:innen, sondern vor allem durch algorithmische Systeme, die nach kommerziellen, sozialen oder technischen Kriterien bestimmen, welche Inhalte dargeboten werden (Bucher, 2018). Das, was Menschen täglich in Feeds, Videos und Messengern zu sehen und zu verarbeiten bekommen, stellt somit keinen objektiv-sachlichen Realitätseinblick dar, sondern eine algorithmisch konstruierte Auswahl aus der unüberschaubaren Menge an möglichen Informationen. Ein solcher Realitätseinblick ist insofern problematisch, als er weniger eine geteilte Vorstellung von Wirklichkeit vermittelt, sondern vielmehr subjektiv gefilterte Konstruktionen erzeugt. Realität entsteht demnach nicht als objektiv erfassbare Gegebenheit, sondern im Sinne des Thomas-Theorems durch soziale Zuschreibungen und Anerkennung: „If men define situations as real, they are real in their consequences“ (Thomas & Thomas, 1928, S. 572). Das bedeutet, dass Wirklichkeit durch individuelle und kollektive Deutungen entsteht und damit soziale Handlungen und Wahrnehmungen formt.

Diese algorithmisch gesteuerte Personalisierung hat erhebliche Auswirkungen auf das subjektive Weltbild und die Identitätsentwicklung von Menschen. Gillespie (2014) beschreibt Algorithmen nicht nur als technische Werkzeuge, sondern als kulturelle Akteure, die Einfluss darauf nehmen, wie Wissen strukturiert, gesteuert und Bedeutung erzeugt wird. Einen ergänzenden Blick darauf bietet Pasquinelli (2024), der Algorithmen als Teil gesellschaftlicher Arbeitsteilung versteht und sie in

einen Zusammenhang mit Kontrolle und Rationalisierbarkeit stellt. Er argumentiert, dass Algorithmen nicht neutral sind, sondern in bestehende ökonomische und soziale Machtverhältnisse eingebettet bleiben. Durch ihre Fähigkeit, Entscheidungsprozesse zu automatisieren und Muster in Daten zu erkennen, tragen sie zur Verstärkung von Effizienz- und Überwachungslogiken bei, die historisch aus industriellen und bürokratischen Organisationsformen hervorgegangen sind. Damit wird deutlich, dass Algorithmen nicht nur Wissen strukturieren, sondern auch Ausdruck und Triebkraft gesellschaftlicher Rationalisierungsprozesse sind.

In sozialen Netzwerken bewegen sich Menschen innerhalb einer durch algorithmische Strukturen kuratierten Umgebung, die auf den ersten Blick vertraut und zugänglich erscheint. Diese digitalen Räume vermitteln Orientierung und Anschluss, zugleich führen ihre Filtermechanismen jedoch zu einer zunehmenden Vereinheitlichung der sichtbaren Inhalte. Damit wächst die Gefahr, dass Vielfalt an Perspektiven und Ausdrucksformen an Bedeutung verliert. In der Folge können bestehende Deutungsmuster und Überzeugungen unbewusst gestärkt werden, da Interaktionen und Empfehlungen häufig jene Inhalte bevorzugen, die bereits vorhandene Sichtweisen bestätigen.

Ein zentrales Stichwort in diesem Zusammenhang ist das Konzept der *Filterblase* (Pariser, 2011). Damit ist der Effekt gemeint, dass algorithmische Systeme Nutzer:innen bevorzugt Inhalte anzeigen, die ihren bisherigen Einstellungen und Interessen entsprechen. Widersprechende Informationen oder alternative Perspektiven werden hingegen algorithmisch ausgeblendet. Diese Dynamik verstärkt nicht nur kognitive Verzerrungen, sondern verringert auch die Möglichkeit einer differenzierten und kritischen Auseinandersetzung mit gesellschaftlichen Themen und kontrastierenden Positionen. Sunstein (2018) warnt in diesem Zusammenhang vor einer zunehmenden Fragmentierung öffentlicher Diskurse durch sogenannte *Echokammern*, in denen Menschen vor allem mit Gleichgesinnten kommunizieren. Für Personen mit begrenzter Medienerfahrung kann dies besonders problematisch sein, da es etwa die Herausbildung und Verfestigung radikaler Ansichten begünstigen kann. Wer weniger Erfahrung im kritischen Umgang mit digitalen Medien hat, ist eher gefährdet, die durch Filterblasen und Echokammern entstehenden Verzerrungen nicht zu erkennen, was die Bildung und Verfestigung radikaler oder einseitiger Sichtweisen begünstigen kann.

Zusätzlich wird die digitale Umwelt nicht nur gefiltert, sondern oft auch ästhetisch inszeniert. Die medialen Inhalte, die über Plattformen wie Instagram, TikTok oder YouTube konsumiert werden, zeichnen sich häufig durch eine optimierte Selbstdarstellung und emotional aufgeladene Visualität aus (Knaus, 2017; Neumann, 2022; Schachtner, 2023). Digitale Identitäten werden performt, bearbeitet und angepasst – häufig mit dem Ziel, möglichst viel Aufmerksamkeit, Zustimmung oder Reichweite zu ge-

nerieren. Was auf der Oberfläche *echt* erscheint, ist daher oft das Ergebnis sozialer und technischer Selektionsprozesse, die häufig keinen hohen Grad an Authentizität erreichen können. Dies liegt daran, dass digitale Darstellungen in sozialen Medien stark durch algorithmische Logiken, ästhetische Normen und Erwartungen an Sichtbarkeit geprägt sind, wodurch die spontane, ungefilterte Selbstrepräsentation in den Hintergrund tritt.

Für die Soziale Arbeit ergibt sich daraus eine doppelte pädagogische Herausforderung: Erstens müssen Fachkräfte selbst die Mechanismen algorithmischer Steuerung und digitaler Inszenierung verstehen und reflektieren. Zweitens gilt es, Räume zu gestalten, in denen Menschen lernen, zwischen Realität, Repräsentation und Fiktion zu unterscheiden. Drittens geht es darum, diese Räume gemeinsam zu entwickeln und so Prozesse partizipativer Mediengestaltung zu ermöglichen. Es braucht pädagogische Settings, in denen digitale Realitätskompetenz entwickelt wird. Hierdurch lässt sich die Fähigkeit adressieren, mediale Inhalte kritisch zu hinterfragen, deren Entstehungsbedingungen zu reflektieren und sich in einer durch Datenlogiken geprägten Welt selbstbestimmt zu orientieren sowie souverän zu navigieren (Herzig, 2023; Schachtner, 2023).

Der Begriff der *magrittesken Realität*, wie er im Titel des Beitrags anklingt, verweist genau auf diese Problematik: So wie Magrittes Pfeife keine echte Pfeife ist, sondern deren Abbild, so scheinen auch digitale Inhalte oft keine Realitätsdarstellungen, sondern symbolische, interessen geleitete Konstruktionen zu sein. Die Fähigkeit, symbolische Darstellungen und mediale Konstruktionen als solche zu erkennen und einzuordnen, stellt einen wichtigen Schritt in Richtung medienpädagogischer Mündigkeit dar. Sie kann als zentrale Zielformel kompetenter Sozialer Arbeit im digitalen Zeitalter verstanden werden.

3. Soziale Arbeit im digitalen Raum: Herausforderungen und Chancen

Digitalisierungsprozesse verändern die Lebenswelten von Menschen in grundlegender und vielfältiger Weise (z. B. durch neue Formen der sozialen Interaktionen oder durch Plattformökonomien etc.). Soziale Beziehungen, Identitätsbildung und gesellschaftliche Teilhabe finden zunehmend in digitalen Räumen statt, die sich kaum noch von analogen Realitäten trennen lassen (Herzig, 2023; Hoffmann, 2023). Soziale Arbeit sollte die digitalen Dimensionen dieser Lebenswelten nicht nur mitdenken, sondern aktiv mitgestalten – professionell, kritisch und pädagogisch reflektiert. Wie dargestellt, gelten digitale Räume jedoch meist nicht als neutral.

Es lässt sich betonen, dass darin spezifische Machtverhältnisse entstehen, die über (Nicht-)Sichtbarkeiten bestimmen und die Subjektivität von Menschen formen. Diese Machtverhältnisse zeigen sich insbesondere in Prozes-

sen der Identitätsbildung, die stark über mediale Repräsentation verlaufen (Hoffmann, 2023; Macgilchrist, 2023). Dies verweist darauf, dass digitale Medienräume mehr sind als Kommunikationsplattformen, sondern auch soziale Bedeutungsräume, in denen Zugehörigkeit, Anerkennung und Ausschluss verhandelt werden. Digitale Räume gewinnen dadurch große Relevanz für Prozesse sozialer Integration und Exklusion. Zugleich eröffnen digitale Medien neue Möglichkeiten der sozialen Unterstützung und des Austauschs. Online-Kommunikation, Gruppenangebote und medienpädagogische Tools können dazu beitragen, vielfältige Zielgruppen zu erreichen und Teilhabe zu fördern. Wie Peissl und Sedlaczek (2022) zeigen, können digitale Formate Zugänge eröffnen, die sich traditionellen Strukturen entziehen. Dieses Potenzial birgt jedoch auch Risiken, etwa im Hinblick auf ungleiche Partizipationsmöglichkeiten oder die Verstärkung bestehender sozialer Differenzen.

Wenn sozialprofessionelle Fachkräfte die Funktionslogiken digitaler Medien unzureichend verstehen oder reflektieren, besteht die Gefahr eines unkritischen Technikeinsatzes oder der ungewollten Reproduktion bestehender Ausschlüsse. Kritische Medienkompetenz ist daher auch im professionellen Handeln von zentraler Bedeutung – als Haltung, die digitale Praktiken hinterfragt und gesellschaftlich einordnet. Vor diesem Hintergrund lässt sich für eine konsequente Integration digitaler Lebenswelten in die Konzepte, Methoden und Ethiken der Sozialen Arbeit plädieren. Digitale Prozesse sind dabei mehr als nur technische Erweiterung zu begreifen, sondern auch als Ausdruck tiefgreifender gesellschaftlicher Transformationsprozesse, die das professionelle Handeln in der Sozialen Arbeit neu ausrichten. Soziale Arbeit wird dadurch zur Begleiterin und zur aktiven Mitgestalterin einer zunehmend digital geprägten Gesellschaft (Wunder, 2021). Digitale Räume bilden zentrale gesellschaftliche Handlungs- und Erfahrungsräume, die durch die verschiedenen Akteur:innen Sozialer Arbeit professionell erschlossen und kritisch begleitet werden müssen. Ein solcher Ansatz erfordert neue Formate der Ansprache und Beratung, sowie ein vertieftes Verständnis dafür, wie Macht, Sichtbarkeit und Teilhabe im Digitalen fortlaufend ausgehandelt werden.

4. Aura und Verrat von Bildern: Wo kann Soziale Arbeit ansetzen?

Greift man Magrittes Bild *La trahison des images* auf, so verdeutlicht dieses eindrücklich, wie leicht Darstellung und Realität verwechselt werden können. Gerade im digitalen Zeitalter, in dem mediale Inhalte häufig als Wirklichkeitsabbildungen wahrgenommen werden, bietet Magrittes Werk eine prägnante Metapher für die kritische Auseinandersetzung mit der Glaubwürdigkeit und Wirkung von Bildern. Conner (2021) zeigt, dass Magrittes Pfeifenbild als Anstoß dienen kann, um Wahrnehmungsprozesse und den Umgang mit digitalen Darstellungen bewusster zu reflektieren. Magrittes Ansatz ist daher besonders aufschlussreich, weil er

deutlich macht, dass Medienbilder keine neutralen Abbilder der Realität sind, sondern konstruierte Darstellungen, die Bedeutungen erzeugen und beeinflussen.

Diese Einsicht findet ihre theoretische Ergänzung in Benjamins Überlegungen zum „Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit“ (2024 [1936]). Er beschreibt, wie durch technische Reproduzierbarkeit die Aura eines Kunstwerks, bezogen auf Einzigartigkeit und Echtheit, verändert werden kann. Digitale Bilder führen diesen Prozess fort, indem sie neue Formen von Authentizität und Präsenz hervorbringen. Ihre Wahrnehmung ist nicht an Echtheit, sondern an algorithmische Sichtbarkeit und Aufmerksamkeitslogiken gekoppelt. Für die sozialpädagogische Praxis ergeben sich daraus mehrere Konsequenzen. Soziale Arbeit kann – in Anlehnung an Conner sowie an Benjamin – dazu beitragen, digitale Bilder kritisch und differenziert wahrzunehmen. Dies kann in drei zentralen Handlungspunkten skizziert werden:

- 1) **Magritteske Perspektiven vermitteln:** Durch eine kritische Distanz gegenüber konsumierten Darstellungen von Welt- und Realitätsausschnitten wird erfahrbar, dass mediale Bilder stets selektive Repräsentationen und Konstruktionen sind, die Wirklichkeiten interpretieren, umdeuten oder manipulieren können.
- 2) **Instrumentelle Agenda sichtbar machen:** Reflexive Lern- oder Diskussionskontexte können genutzt werden, um Entstehung, Verbreitung und Interessengeleitetheit digitaler Inhalte zu hinterfragen. Dies kann beispielsweise durch niedrigschwellige analytische Gespräche geschehen, in denen thematisiert wird, woher konkrete Bilder und Informationen stammen, weshalb sie geteilt werden, welche gesellschaftlichen Folgen ihre Sichtbarkeit hat und welche politischen oder ökonomischen Ziele damit verknüpft sein können.
- 3) **Kreativ-reflexive Methoden einsetzen:** Selbstgesteuerte kreative Projekte – etwa digitale Collagen, Fotografie-Workshops oder Videoproduktionen – regen dazu an, die Grenzen zwischen Realität und digitaler Inszenierung aktiv zu erkunden. Ein solcher ästhetisch und medienpädagogisch orientierter Ansatz kann die Entwicklung visueller Mündigkeit fördern, indem er dazu anregt, digitale Inhalte nicht unreflektiert als Realität zu übernehmen, sondern sie als interpretier- und hinterfragbare Repräsentationen zu verstehen.

5. Digitale Resonanzökonomien: Likes, Shares und performative Identität

Im Kontext einer digitalen Resonanzökonomie fungieren Likes, Kommentare und Follower:innen zahlen als eine Art soziale Währung, die für viele Menschen eine zentrale Bedeutung besitzt. Der Wunsch nach digitaler Anerkennung beeinflusst maßgeblich die Art und Weise, wie Individuen sich selbst darstellen und wahrnehmen. Soziale Plattformen wie Insta-

gram oder TikTok strukturieren dabei Kommunikation, aber auch Selbstbilder und soziale Vergleiche. Forschungsbefunde zeigen deutlich, dass digitale Rückmeldungen unmittelbare Auswirkungen auf die emotionale Befindlichkeit der Nutzer:innen haben. Sherman et al. (2016) dokumentierten in einer MRT-Studie, dass Personen bei positiven Rückmeldungen (etwa vielen Likes) eine verstärkte Aktivierung im Belohnungssystem des Gehirns zeigen, insbesondere im Striatum. Ähnliche Ergebnisse liefert eine aktuelle Untersuchung von Da Silva Pinho et al. (2024), die darauf hinweist, dass digitale Rückmeldungen starke emotionale Reaktionen auslösen können. Bereits die bloße Erwartung von Likes kann Stimmung und Selbstwertgefühl beeinflussen, während ihr Ausbleiben Frustration oder Rückzugsverhalten begünstigen kann.

Zahlreiche empirische Studien verdeutlichen zudem, dass digitale Resonanz eng mit Selbstdarstellung, Schönheitsidealen und sozialen Vergleichsprozessen verknüpft ist. Thiel (2024) beschreibt, wie die Plattformlogik von Instagram zu einer dauerhaften sozialen Bewertung führt: Nutzer:innen vergleichen sich mit (scheinbar perfekten) Abbildern anderer und erleben dadurch Leistungsdruck oder vermindertes Selbstwertgefühl. Ähnliche Dynamiken skizziert Tiches (2024): Personen passen ihre digitale Selbstdarstellung häufig an erwartete Normen an, um möglichst viel Zustimmung zu erhalten. Besonders die Inszenierung von Schönheit, Fitness oder Erfolg spielt hierbei eine zentrale Rolle. Mangl (2022) verweist auf gesundheitliche Risiken, die mit einer starken Einbindung sozialer Medien in den Alltag einhergehen: Ein übermäßiger Fokus auf Körperideale kann Selbstzweifel, Essstörungen oder erhöhten psychischen Stress begünstigen. In einem ähnlichen Befundrahmen zeigt Reinartz (2024), dass der Versuch, eine positive digitale Identität zu etablieren, kurzfristig stärkend wirken kann, langfristig jedoch von der Resonanz anderer abhängig bleibt.

Für die Soziale Arbeit ergeben sich daraus klare Handlungsfelder: Es gilt, Räume zu schaffen, in denen digitale Resonanz, Selbstbild und mediale Vergleichsprozesse reflektiert werden können. Dazu zählen beispielsweise medienpädagogische Workshops, die Unterschiede zwischen realem Selbst und inszeniertem Bild herausarbeiten, Formate zur Auseinandersetzung mit Körper- und Selbstbildern, die digitale Schönheitsnormen kritisch hinterfragen, sowie Peer-basierte Gruppenangebote, in denen Erfahrungen mit Anerkennungsdruck und Selbstinszenierung thematisiert werden können. Ebenso wichtig ist die Förderung von Resilienz im digitalen Kontext – also die Stärkung emotionaler Selbstregulation und Unabhängigkeit von digitaler Bestätigung. Die digitale Resonanzökonomie darf demnach von der Sozialen Arbeit nicht als nebensächliches Phänomen betrachtet werden. Sie stellt ein zentrales soziales Feld der Gegenwart dar und prägt die psychosoziale Entwicklung und das Selbstverständnis vieler Menschen in erheblichem Maße.

6. Zwischen Repräsentation und Erleben: Zur systematischen Unabgeschlossenheit digitaler Wirklichkeit

In der bislang skizzierten Strukturierung von Sichtbarkeit, Beziehungen und Anerkennung in digitalen Sozialräumen wurden soziale, emotionale und identitätsbezogene Konsequenzen thematisiert. Dabei bleibt eine Grundspannung bestehen: Die medial vermittelte Erfahrung erzeugt Präsenz, jedoch meist kein vollständiges, lebendiges „Wirklichkeitsverhältnis“. Dieser Ausdruck beschreibt, dass digital vermittelte Erfahrungen – etwa in sozialen Medien oder virtuellen Räumen – zwar den Eindruck von Gegenwart und Nähe erzeugen können, aber nicht an die Ganzheit und Unmittelbarkeit real geteilter, leiblicher Erfahrungen heranreichen. Ein vollständiges, lebendiges Wirklichkeitsverhältnis meint somit ein umfassendes Erleben von Realität, in dem sinnliche, emotionale und soziale Dimensionen unmittelbar ineinandergreifen – etwas, das digitale Vermittlung nur in begrenztem Maße ermöglicht.

In der Folge lässt sich eine Form fortwährender Unstimmigkeit zwischen der erlebten Resonanz digitaler Anerkennung und einem subjektiv empfundenen Mangel an Tiefe oder Sinn feststellen. Foucault (1993) beschreibt in seinen Arbeiten zur Ordnung des Diskurses und zu den sogenannten Heterotopien die Existenz von Räumen, die niemals vollständig im realen Raum verortet sein können. Digitale Plattformen lassen sich in diesem Sinne als heterotope Räume verstehen: zugleich konkret und entrückt, strukturierend und unbetreibar, real in ihrer Wirkung, aber von einer anderen Ordnung als der des gelebten Alltags. Auch Foucaults (1978) Analysen zur Untrennbarkeit von Subjektkonstitution und diskursiver Rahmung zeigen, dass das Subjekt stets durch kulturelle Praktiken, Dispositive der Macht und gesellschaftliche Ordnungen des Sagbaren vermittelt ist.

Die digitale Selbstdarstellung lässt sich als ein fortlaufender Prozess begreifen, in dem das Subjekt sich selbst hervorbringt und formt. Dieses Selbstbild ist jedoch nicht fest oder eindeutig, sondern bleibt wandelbar, von äußeren Einflüssen geprägt und stets neu verhandelbar. Statt eine stabile Identität zu bestätigen, erzeugt digitale Selbstdarstellung vielmehr ein dynamisches, oft widersprüchliches Bild des eigenen Selbst. Die digitale Selbstdarstellung bleibt demnach prekär: geformt durch selektive Algorithmen, strukturiert durch Plattformlogiken und abhängig von quantifizierter Rückmeldung. Diese Bedingungen erzeugen nicht nur individuelle Spannungen, sondern stehen auch im Zusammenhang mit dem, was Bourdieu (1982) als symbolische Gewalt bezeichnet: eine subtile, unsichtbare Form der Machtausübung, die durch die Anerkennung gesellschaftlicher Normen, Ästhetiken und Strukturen wirksam wird. In digitalen Kontexten zeigt sich symbolische Gewalt etwa in der stillschweigenden Akzeptanz normativer Schönheitsideale, performativer Erwartungen oder algorithmischer

misch erzeugter Sichtbarkeit. Sie tritt dort in Erscheinung, wo soziale Dominanzverhältnisse ästhetisch naturalisiert und als persönlicher Erfolg oder Misserfolg fehlinterpretiert werden.

Trotz sozialer Wirksamkeiten digitaler Anerkennung kann ein persistentes Gefühl der Leere und Sinnlücke bestehen bleiben. Die hier beschriebene Erfahrung digitaler Anerkennung lässt sich als Ausdruck einer magritt-esken Wirklichkeit lesen: Sichtbarkeit und Resonanz erscheinen präsent, verweisen jedoch nicht notwendig auf ein tragfähiges Wirklichkeitsverhältnis. Die Differenz zwischen Darstellung und gelebter Erfahrung bleibt bestehen – und zeigt sich nicht selten als Gefühl von Leere und Sinnlücke. Diese Erfahrung von Resonanz erweist sich oft als flüchtig und unverfügbar, sodass Tendenzen zur Verdinglichung positiver Resonanzbeziehungen gegenüber der digitalen Umgebung entstehen können (Rosa, 2016). Anerkennungserfahrungen bleiben darin konditional und in Echtzeit korrigierbar. Hier eröffnet sich eine Parallele zu Adornos (2023 [1969]) Diktum: „Es gibt kein richtiges Leben im falschen.“ Selbst wenn digitale Räume neue Formen von Zugehörigkeit hervorbringen, können sie sozial getragene, körperlich verortete und dialogische Erfahrung nur begrenzt ersetzen. Für professionelles Handeln in sozialen, pädagogischen oder beratenden Kontexten ergibt sich daraus die Notwendigkeit, diese Spannung weder zu ignorieren noch vorschnell aufzulösen. Vielmehr sollte sie als Ausgangspunkt für eine reflexive, ethisch fundierte Auseinandersetzung verstanden werden: Es braucht Räume, in denen über die Differenz zwischen Sichtbarkeit und Beziehung, zwischen Algorithmus und Begegnung, zwischen Reaktion und Resonanz nachgedacht werden kann – ohne moralische Abwertung, aber mit kritischem Bewusstsein für die Ambivalenzen digitaler Sozialität.

Aus den Befunden des Beitrags lässt sich schlussfolgern, dass die *magritt-eske Realität* im Falle von fehlender Distanzierung durch Dekonstruktion zu mindestens vier Formen von besonderen Herausforderungen (Stressoren) für die Nutzer:innen werden kann:

- 1) *Wahrnehmungsstress*, weil Menschen permanent Bildern und Inhalten ausgesetzt sind, die echt wirken, aber inszeniert oder algorithmisch gefiltert sind. Dadurch fällt es zunehmend schwer, zwischen realer Erfahrung und medialer Darstellung zu unterscheiden – ständige Reize und widersprüchliche Eindrücke führen zu Überforderung und einem unsicheren Realitätsbezug.
- 2) (Indirekter) *Algorithmischer Stress* zeigt sich darin, dass soziale Plattformen bestimmen, was sichtbar wird. Diese unsichtbare Steuerung erzeugt ein Gefühl des Kontrollverlusts, verstärkt durch Filterblasen und die Wahrnehmung, von Algorithmen manipuliert zu werden.

- 3) *Sozial-emotionaler Stress* entsteht in der sogenannten Resonanzökonomie, in der Likes, Follower:innen und Kommentare zur sozialen Währung werden. Digitale Anerkennung beeinflusst das Selbstwertgefühl, während Vergleichsdruck und Idealisierungen zu emotionaler Erschöpfung führen.

Schließlich bewirkt 4) **Identitätsstress**, dass das Selbst in digitalen Räumen ständig beobachtet, bewertet und neu inszeniert wird. Der dauerhafte Anpassungsdruck erzeugt ein Gefühl der Entfremdung und erschwert die Balance zwischen Authentizität und sozialer Anerkennung.

7. Ästhetische Bildung und Reflexion als Antwort einer interessierten Sozialen Arbeit

In einer zunehmend mediatisierten Welt, in der Bilder, Stories und digitale Performanz alltägliche Selbst- und Weltverhältnisse strukturieren, ist die Soziale Arbeit gefordert, ihre pädagogische Praxis entsprechend weiterzuentwickeln. Eine zentrale Antwort auf die Herausforderungen digitaler Sichtbarkeit, Selbstdarstellung und algorithmischer Steuerung kann in der ästhetischen Bildung liegen, verstanden als reflexive, kreative und soziale Praxis. Ästhetische Bildung ermöglicht es, Medieninhalte nicht nur aufzunehmen, sondern deren Bedeutungsgehalte zu interpretieren, zu hinterfragen und gestalterisch zu bearbeiten. In der Tradition kritischer Bildungskonzepte verbindet sie Wahrnehmungsschulung mit Weltverstehen und schafft dabei neue Räume, in denen individuelle Deutungsperspektiven auf die eigene Lebensrealität sichtbar und erfahrbar werden können. Gerade im Kontext Sozialer Arbeit, die gegenwärtig mit pluralen komplexen Lebenslagen, sozialen Ausgrenzungen und Identitätsfragen befasst ist, bietet diese Form der Bildung einen niedrigschwiligen, erfahrungsbasierten Zugang zu Reflexion und Artikulation.

Insbesondere die Gestaltung digitaler Medienprodukte wie Videos, Fotos, digitaler Collagen, Podcasts, Memes oder Soundscapes eröffnet die Möglichkeit, persönliche Realitätserfahrungen mit medienvermittelten Anerkennungsprozessen, Schönheitsnormen oder digitaler Resonanz kreativ und produktiv zu verarbeiten (Quenzel & Hurrelmann, 2022). Der Fokus liegt dabei nicht auf Kunstfertigkeit, sondern auf Gestaltungsfreiheit, Reflexionsfähigkeit und der Rekonfiguration digital verankerter Resonanz-erwartungen. Als eine an digitalen und analogen Weltausschnitten interessierte Profession trägt Soziale Arbeit Verantwortung dafür, nicht nur technische Medienkompetenz zu vermitteln, sondern auch jene Kompetenzen zu stärken, die Menschen befähigen, sich als denkende, fühlende und handelnde Subjekte im digitalen Raum zu begreifen. Dazu gehört nicht allein die Fähigkeit, Bilder zu lesen, sondern auch, sie selbstwirksam zu lesen und zu konterkarieren. Ästhetische Bildung öffnet hier Räume für doppelte Selbstwirksamkeit: im Ausdruck des eigenen Ichs und in der Deutung des Anderen. Konkret kann ästhetische Bildung sowohl in

der Ausbildung Sozialer Arbeit – etwa in projektorientierten Seminaren, medienpädagogischen Übungen oder forschendem Lernen – als auch in der Praxis Anwendung finden, beispielsweise in niedrigschwelligen Gruppenangeboten, Workshops oder begleiteten Medienprojekten mit Klient:innen. Entscheidend ist dabei weniger ein aufwendiger Rahmen als die reflexive Einbettung kreativer Prozesse in bestehende pädagogische und beratende Settings.

Soziale Arbeit kann sich dabei als Kulturvermittlerin verstehen: Sie übersetzt zwischen Lebenswelt, Technologie und gesellschaftlicher Symbolik. Durch kreative Methoden können Fachkräfte nicht nur an mediale Alltagspraktiken anknüpfen, sondern auch Widersprüche, Ambivalenzen und Unsicherheiten thematisieren. Dies betrifft etwa den Wunsch nach Sichtbarkeit und die Angst vor (negativer) Bewertung. Solche Ambivalenzen lassen sich nicht auflösen, aber mit ästhetischen Mitteln ausdrücken und im größeren sozialen Zusammenhang reflektieren. Auf disziplinärer Ebene fordert der Einbezug ästhetischer Zugänge die Soziale Arbeit heraus, ihr Bildungsverständnis zu erweitern. Dabei geht es nicht allein um funktionale Medienkompetenz, sondern um kritisch-kulturelle Bildung: Wer digitale Bilder in ihrer Komplexität versteht, begreift auch Machtverhältnisse, Normsetzungen und Exklusionsmechanismen besser. Wenn sich digitale Narrative kritisch gestalten lassen, kann eine wirkmächtige Stimme in der öffentlichen Kommunikation entstehen.

So werden ästhetische Verfahren zu politischen, sozialen und identitätsbildenden Prozessen in gesellschaftlichen Arenen. In einer Gegenwart, in der gesellschaftliche Debatten zunehmend über Bildpolitik und symbolische Repräsentation geführt werden, muss eine handlungsfähige Soziale Arbeit verstärkt auf medienästhetische Deutungs- und Handlungskompetenz setzen. Dies betrifft gleichermaßen Aus- und Weiterbildung, Forschung und Praxis. Auf diese Weise wird deutlich, dass ästhetische Bildung keine Randnotiz im Kanon sozialpädagogischer Methoden darstellt, sondern eine zentrale Antwort auf die Herausforderungen der digitalen Gegenwart.

8. Einladung zu Diskussion und Fazit

Die Auseinandersetzung mit digitalen Realitäten, Filterblasen, algorithmischer Steuerung und performativer Selbstdarstellung verdeutlicht, dass Menschen sich in einer mediatisierten Welt bewegen, in der Repräsentation und Realität, Selbstbild und Fremdzuschreibung, Echtheit und Inszenierung zunehmend ineinander übergehen. Der oft unbemerkte Bruch zwischen Bild und Wirklichkeit kann im digitalen Raum zur Normalität werden. In diesem Kontext ist Soziale Arbeit herausgefordert, neue technische, bildungstheoretische, ästhetische und professionsethische Antworten zu entwickeln. Vergleichbare Ansätze finden sich bereits in medienpädagogischen Werkstattformaten der offenen Kinder- und Jugendarbeit,

in partizipativen Digitalprojekten der Jugendhilfe sowie in projektorientierten Lehrformaten der Sozialen Arbeit. Diese Formate zeigen, dass ästhetisch-reflexive Zugänge im digitalen Raum auch unter realistischen zeitlichen und institutionellen Bedingungen umsetzbar sind. Die Reflexion über digitale Resonanzökonomien, über Schönheitsideale, soziale Vergleichsprozesse oder emotionale Auswirkungen von Online-Interaktionen macht deutlich: Es genügt nicht, Medienkompetenz einzufordern. Es braucht Konzepte und Haltungen, die digitale Lebenswelten ganzheitlich und kritisch erfassen.

Einerseits ermöglichen soziale Medien Anschluss, Ausdruck und Partizipation. Sie bieten Räume für kreative Erprobung, soziale Rückkopplung und das Gefühl von Relevanz. Andererseits wirken dieselben Räume normierend, polarisierend und distanzierend, insofern leibliche Ko-Präsenz, Stimme und situative Resonanz durch mediale Zeichen und quantifizierte Rückmeldungen ersetzt werden. Likes ersetzen kein Zuhören, Filter keine Auseinandersetzung, Reichweite keine Zugehörigkeit. Menschen bewegen sich damit in einem Spannungsfeld aus Partizipation und Erschöpfung, aus Selbstermächtigung und Selbstausbeutung. Soziale Arbeit begegnet dieser Ambivalenz unmittelbar. Sie ist gefordert, Räume zu schaffen, in denen Menschen sowohl gesehen werden als auch sich selbst sehen lernen – nicht durch den Spiegel digitaler Resonanz, sondern im Dialog, im ästhetischen Ausdruck, im Infragestellen der Oberfläche.

Es lässt sich festhalten, dass das *Magritteske* in der digitalen Realität vor allem in der Täuschung der Wahrnehmung liegt. Wie Magritte mit seinem berühmten Gemälde die Betrachtenden dazu zwingt, die Grenze zwischen Zeichen und Wirklichkeit zu hinterfragen, so fordert auch die digitale Welt eine neue Form der Aufmerksamkeit und Kritikfähigkeit. Besonders deutlich wird dies in *La reproduction interdite*³ (1937), wo Magritte einen Mann zeigt, der in den Spiegel blickt – und statt seines Gesichts erneut seinen Hinterkopf sieht. Der Spiegel, Symbol des Selbst und der Erkenntnis, versagt hier seine Funktion: Er reflektiert nicht, sondern reproduziert das Gleiche, das bereits vor ihm steht. In der digitalen Realität begegnen wir einer ähnlichen Paradoxie. Unsere Avatare, Profile und Projektionen erscheinen wie Spiegelbilder unseres Selbst, doch sie zeigen nur wiederholte, gefilterte Ausschnitte – nicht das Gesicht *dahinter*. Magrittes Satz „Dies ist keine Pfeife“ wird so zur Mahnung, die Distanz zwischen Darstellung und Wirklichkeit nicht aus den Augen zu verlieren.

Mit Blick auf die vorgestellten Argumente und Zugänge ergeben sich zentrale Fragen, die im Diskurs innerhalb der Disziplin und Profession weiterverfolgt werden sollten:

3 Deutsch: Reproduktion verboten.

Was bedeutet *Realität* im digitalen Raum? Wie kann Soziale Arbeit Wahrnehmungen differenziert und mit kritischer Distanz begleiten, ohne vorschnell zu bewerten? Wie kann ästhetische Bildung als emanzipatorische Praxis systematisch in Ausbildung und Praxis Sozialer Arbeit integriert werden? Wie verändert sich das professionelle Selbstverständnis in digitalen Interaktionsräumen? Wie gelingt es, im digitalen Medienraum zwischen Empowerment und Schutz zu balancieren? Wie kann Soziale Arbeit Menschen mediale Handlungsmacht zutrauen, ohne sie mit den Risiken allein zu lassen? Diese Fragen sollen nicht abschließend beantwortet, sondern in Theorie und Praxis, in Ausbildung und Fortbildung, im kollegialen Austausch und in interdisziplinärer Zusammenarbeit weitergedacht, diskutiert und erforscht werden. Die digitale Transformation sozialer Realitäten stellt kein abgeschlossenes Projekt dar, sondern einen andauernden, dynamischen Prozess. Soziale Arbeit ist dazu eingeladen, ihn nicht nur zu begleiten, sondern ihn kreativ, kritisch und solidarisch mitzugestalten.

Literatur

- Adorno, T. W. (2023). *Minima Moralia: Reflexionen aus dem beschädigten Leben* (32. Aufl.). Suhrkamp. (Erstveröffentlichung 1969)
- Benjamin, W. (2024). *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit. Edition Suhrkamp: Bd. 28*. Suhrkamp. (Erstveröffentlichung 1963)
- Bourdieu, P. (1982). *Die feinen Unterschiede: Kritik der gesellschaftlichen Urteilskraft*. Suhrkamp Verlag.
- Bucher, T. (2018). *If... then: Algorithmic power and politics. Oxford studies in digital politics*. Oxford University Press.
- Conner, H. (2021). René Magritte and Media Literacy. *International Council for Media Literacy: Media Perception Blogs*. <https://ic4ml.org/blogs/media-perception/rene-magritte-and-media-literacy/>
- Da Silva Pinho, A., Céspedes Izquierdo, V., Lindström, B. & van den Bos, W. (2024). Youths' sensitivity to social media feedback: A computational account. *Science Advances*, 10(43), eadp8775. <https://doi.org/10.1126/sciadv.adp8775>
- Eckl, M., & Ghanem, C. (2024) Der Digitalisierungsdiskurs in der Sozialen Arbeit. *Soziale Passagen* 16, 385–402 (2024). <https://doi.org/10.1007/s12592-024-00511-w>
- Foucault, M. (1978). *Dispositive der Macht: Über Sexualität, Wissen und Wahrheit*. Merve Verlag.
- Foucault, M. (1993). Andere Räume. In K. Barck, P. Gente & H. Paris (Hrsg.), *Aisthesis: Wahrnehmung heute oder Perspektiven einer anderen Ästhetik. Essais* (3. Aufl., S. 34–46). Reclam.
- Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P.J. Boczkowski & K. A. Foot (Hrsg.), *Media Technologies* (S. 167–194). The MIT Press. <https://doi.org/10.7551/mitpress/9042.003.0013>

- Herzig, B. (2023). Digitalität, Mediatisierung und Bildung – Megatrends aus medienpädagogischer Perspektive. In S. Aßmann & N. Ricken (Hrsg.), *Bildung und Digitalität: Analysen – Diskurse – Perspektiven* (S. 99–125). Springer VS. https://doi.org/10.1007/978-3-658-30766-0_5
- Hoffmann, D. (2023). Mediensoziologische Perspektiven auf digitale Körperbildpraktiken und Subjektivierung. In S. Aßmann & N. Ricken (Hrsg.), *Bildung und Digitalität: Analysen – Diskurse – Perspektiven* (S. 75–97). Springer VS. https://doi.org/10.1007/978-3-658-30766-0_4
- Hug, T. (2016). Mobilität und Medienbildung im digitalen Zeitalter: Konzeptuelle Überlegungen. In T. Hugo, T. Kohn & P. Missomelius (Hrsg.), *Medien – Wissen – Bildung: Medienbildung wozu?* (S. 151–171). Innsbruck University Press.
- Iske, S., & Kutscher, N. (2020). Digitale Ungleichheiten im Kontext Sozialer Arbeit. In N. Kutscher, T. Ley, U. Seelmeyer, F. Siller, A. Tillmann & I. Zorn (Hrsg.), *Handbuch Soziale Arbeit und Digitalisierung* (S. 115–128). Beltz Juventa.
- Jörissen, B., & Marotzki, W. (2009). *Medienbildung: Eine Einführung. Theorie – Methoden – Analysen*. Klinkhardt/UTB.
- Knaus, T. (2017). Pädagogik des Digitalen: Phänomene – Potentiale – Perspektiven. In S. Eder, C. Mikat & A. Tillmann (Hrsg.), *Schriften zur Medienpädagogik: Bd. 53. Software takes command: Herausforderungen der „Datafizierung“ für die Medienpädagogik in Theorie und Praxis* (S. 49–68). kopaed. <https://doi.org/10.25656/01:14797>
- Macgilchrist, F. (2023). Diskurs der Digitalität und Pädagogik. In S. Aßmann & N. Ricken (Hrsg.), *Bildung und Digitalität: Analysen – Diskurse – Perspektiven* (S. 47–71). Springer VS. https://doi.org/10.1007/978-3-658-30766-0_3
- Magritte, R. (1929). *La trahison des images* [Ölgemälde]. Los Angeles County Museum of Art.
- Mangl, M. (2022). Selbstdarstellung in sozialen Medien Gesundheitlicher Einfluss auf Mädchen* und junge Frauen. In K. Hametner (Hrsg.), *Frauengesundheit und Digitalisierung: Sammelband des Wiener Programms für Frauengesundheit* (S. 121–125). Stadt Wien.
- Neumann, M. (2022). Bedeutung digitaler Medien in der Kinder- und Jugendarbeit. *Sozial Extra*, 46(6), 458–465. <https://doi.org/10.1007/s12054-022-00534-8>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.
- Pasquinelli, M. (2024). *Das Auge des Meisters: Eine Sozialgeschichte künstlicher Intelligenz* (1. Aufl.). Münster: UNRAST.
- Paus-Hasebrink, I., Sinner, P.C., & Kulterer, J. (2019). *Social inequality, childhood and the media: A longitudinal study of the mediatization of socialisation*. Cham, Switzerland: Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-02653-0>

- Peissl, H., & Sedlacek, A. (2022). Kritische Medienkompetenz vor dem Hintergrund der Digitalisierung. Media and Information Literacy (MIL) und Critical Media Literacy (CML) im Vergleich. *Magazin erg wachsenenbildung.at*, (44–45). <https://doi.org/10.25656/01:24474>
- Quenzel, G., & Hurrelmann, K. (2022). *Lebensphase Jugend: Eine Einführung in die sozialwissenschaftliche Jugendforschung* (14. Aufl.). Beltz Juventa.
- Reinartz, F. (2024). *Positive Selbstwahrnehmung in sozialen Medien: Quantitative Evaluation einer positiv-psychologischen Kurzintervention*. *iwop Schriftenreihe der FOM: Bd. 19*. MA Akademie Verlags- und Druck-Gesellschaft.
- Rosa, H. (2016). *Resonanz: Eine Soziologie der Weltbeziehung* (2. Aufl.). Suhrkamp Verlag.
- Schachtner, C. (2023). Selbstinszenierungen im virtuellen Raum: Das Ringen um Anerkennung. In S. Aßmann & N. Ricken (Hrsg.), *Bildung und Digitalität: Analysen – Diskurse – Perspektiven* (S. 197–215). Springer VS. https://doi.org/10.1007/978-3-658-30766-0_8
- Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M. (2016). The Power of the Like in Adolescence: Effects of Peer Influence on Neural and Behavioral Responses to Social Media. *Psychological Science*, 27(7), 1027–1035. <https://doi.org/10.1177/0956797616645673>
- Sunstein, C. R. (2018). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press. <https://doi.org/10.2307/j.ctv8xnhtd>
- Thiel, K. (2024). *Von Aufwärtsvergleichen und Abwärtsspiralen: Soziale Vergleiche auf Instagram*. Leibniz-Institut für Medienforschung. <https://leibniz-hbi.de/3590/>
- Thomas, W. I., & Thomas, D. S. (1928). *The Child in America: Behavior Problems and Programs*. New York: Knopf.
- Tiches, K. (2024). *Young People, Body Image and Social Media* (Research Briefs). Boston Children’s Digital Wellness Lab. <https://digitalwellnesslab.org/research-briefs/social-media-body-image/>
- Wunder, M. (Hrsg.) (2021). *Digitalisierung und Soziale Arbeit. Transformationen und Herausforderungen*. Julius Klinkhardt. <https://doi.org/10.25656/01:23157>
- Ziegler, H. (2008). Kleine Verteidigung ontologischer Theorien in der sozialen Arbeit. Widersprüche. *Zeitschrift für sozialistische Politik im Bildungs-, Gesundheits- und Sozialbereich*, 28(108), 43–53.

Jonathan May¹

Maschinelle Übersetzung in der Kommunikation mit unbegleiteten Minderjährigen mit Fluchterfahrung

Zwischen Inklusion, Exklusion und Teilhabe

¹ Katholische Hochschule Nordrhein-Westfalen

Die vollständige Abschlussarbeit ist einsehbar unter:
<https://doi.org/10.17883/6071>

Zusammenfassung

Die Bachelorarbeit untersucht den Einsatz maschineller Übersetzungsprogramme (z. B. Google Translate, DeepL) in der Kommunikation zwischen Fachkräften der Jugendhilfe und unbegleiteten Minderjährigen mit Fluchterfahrung. Ziel ist zu analysieren, welche Chancen und Risiken für Inklusion, Exklusion und Teilhabe entstehen. Theoretisch wird auf die Systemtheorie Niklas Luhmanns, das Mediatisierungskonzept von Friedrich Krotz und den Teilhabebegriff Bezug genommen. Empirisch basiert die Arbeit auf leitfadengestützten Interviews mit Fachkräften, die maschinelle Übersetzung in ihrer Praxis nutzen. Die Ergebnisse zeigen: Übersetzungsprogramme ermöglichen kurzfristig Zugang zur Kommunikation, sind aber fehleranfällig, kontextblind und können ungewollte Exklusionseffekte erzeugen. Soziale Arbeit steht vor der Aufgabe, Technikreflexion und Handlungssicherheit im Umgang mit digitalen Übersetzungsmedien zu entwickeln.

Schlüsselwörter: Digitale Medien, Soziale Arbeit, maschinelle Übersetzung, unbegleitete Minderjährige mit Fluchterfahrung, Teilhabe, Inklusion, Exklusion, Kommunikation



Machine Translation in Communication with Unaccompanied Foreign Minors. Between Inclusion, Exclusion and Participation

Abstract

The bachelor thesis explores the use of machine translation tools in communication between social workers and unaccompanied minor refugees. It analyzes opportunities and risks for social inclusion, exclusion and participation. Based on Luhmann's systems theory, Krotz's mediatization concept and the notion of participation, the study combines theoretical discussion with qualitative interviews among practitioners. Findings indicate that machine translation facilitates initial communication but remains error-prone and lacks contextual sensitivity, potentially causing misunderstandings and selective exclusion. Social work therefore needs clear standards and reflective strategies to integrate translation technologies responsibly.

Keywords: Digital Media, Social Work, Machine Translation, Unaccompanied Minor Refugees, Participation, Inclusion, Exclusion, Communication

Skizze der Bachelorarbeit

Die vorliegende Bachelorarbeit untersucht den Einfluss maschineller Übersetzungsprogramme auf die Kommunikation zwischen Fachkräften der Sozialen Arbeit und unbegleiteten Minderjährigen mit Fluchterfahrung (umF) in der Jugendhilfe. Ausgangspunkt war die Beobachtung, dass digitale Übersetzungstools in Erstkontakten inzwischen eine zentrale Rolle einnehmen, weil qualifizierte Dolmetscher:innen in vielen Situationen nicht sofort verfügbar sind. Die Arbeit geht der Frage nach, inwiefern maschinelle Übersetzung zur Inklusion beitragen kann oder ob sie – entgegen ihrer Intention – neue Formen der Exklusion erzeugt und so die gesellschaftliche Teilhabe der Jugendlichen beeinflusst.

Zur theoretischen Fundierung wird auf die Systemtheorie Niklas Luhmanns (2021) zurückgegriffen, dessen Verständnis von Kommunikation als selektiven Prozess einen analytischen Rahmen bietet. Kommunikation wird hier nicht als lineare Übertragung von Botschaften begriffen, sondern als ein komplexer Selektionsprozess, bei dem Information, Mitteilung und Verstehen aufeinander bezogen sind. Aus dieser Perspektive wird deutlich, dass maschinelle Übersetzungsprogramme nur einen Teil dieses Prozesses abbilden: Sie können Informationen übertragen, garantieren jedoch weder das Verstehen noch die Einordnung der Inhalte im sozialen Kontext. Die Differenz zwischen Mitteilung und Information, die für Luhmann konstitutiv ist, bleibt bestehen und wird durch algorithmische Verfahren eher verschärft, da Übersetzungen kontextunabhängig erfolgen.

Ergänzend greift die Arbeit auf das Konzept der Mediatisierung von Friedrich Krotz (2001) zurück. Dieses beschreibt den tiefgreifenden Wandel gesellschaftlicher Kommunikation durch digitale Technologien. Maschinelle Übersetzung ist in diesem Sinne ein Mediatisierungsphänomen: Sie verändert nicht nur die Form, sondern auch die Bedingungen von Interaktion. Während klassische Dolmetschung als soziales Handeln erkennbar bleibt, verschiebt die technische Übersetzung Verantwortung von menschlichen Akteur:innen auf algorithmische Systeme. Dadurch entsteht eine strukturelle Asymmetrie: Fachkräfte vertrauen auf die vermeintliche Neutralität technischer Tools, obwohl diese nur begrenzte Sprachsensibilität aufweisen und emotionale oder kulturelle Nuancen ausblenden.

Das Konzept der Teilhabe ergänzt die theoretische Rahmung um eine normative Dimension. Teilhabe wird verstanden als Möglichkeit, an gesellschaftlichen Funktionssystemen wie Bildung, Recht oder Gesundheit mitzuwirken (Wansing et al. 2022). Sprache ist dabei ein zentrales Zugangskriterium. Wenn Übersetzungsfehler Missverständnisse produzieren, können daraus Ausschlussmechanismen entstehen, etwa wenn Jugendliche ihre Rechte nicht verstehen oder Handlungsanweisungen fehlerhaft vermittelt werden. Die Arbeit knüpft hier an die Diskussion um digitale Ungleichheiten an: Auch wenn der Zugang zu Technologie formal gegeben ist, entstehen neue Exklusionsrisiken durch fehlende Kontextualisierung und asymmetrische Machtverhältnisse zwischen professionellen Akteur:innen und den adressierten Jugendlichen.

Die empirische Untersuchung basiert auf drei leitfadengestützten Interviews mit Fachkräften der Jugendhilfe, die regelmäßig maschinelle Übersetzungsprogramme verwenden. Die Auswertung erfolgte qualitativ und zielte darauf ab, subjektive Erfahrungen, Einschätzungen und Handlungsstrategien zu rekonstruieren. Die Ergebnisse zeigen ein ambivalentes Bild: Einerseits betonen die Befragten die enorme Erleichterung, die maschinelle Übersetzer in akuten Situationen bieten. Ohne diese Tools wäre eine basale Verständigung oft unmöglich, etwa bei der vorläufigen Inobhutnahme oder bei der Klärung unmittelbarer Bedürfnisse. Fachkräfte berichten, dass sie dadurch handlungsfähig bleiben, wenn keine Dolmetschdienste verfügbar sind. Besonders in der Anfangsphase der Betreuung werden Übersetzungsprogramme als unverzichtbare Ressource beschrieben.

Gleichzeitig werden deutliche Grenzen sichtbar. Übersetzungen sind häufig ungenau, was bei komplexen Sachverhalten – wie rechtlichen Fragen, medizinischen Informationen oder psychologischen Themen – erhebliche Risiken birgt. Mehrere Fachkräfte schildern Situationen, in denen Fehlübersetzungen zu Missverständnissen führten, die erst später mit Unterstützung professioneller Dolmetscher:innen korrigiert werden konnten. Neben inhaltlichen Fehlern problematisieren die Befragten den Verlust an Beziehungsqualität: Die Kommunikation wirkt formalisiert, spontane

Zwischentöne oder humorvolle Elemente gehen verloren. Emotionale Ausdrucksweisen werden von der Software nicht transportiert, wodurch die Gesprächsatmosphäre distanziert erscheinen.

Die Analyse der Interviews legt nahe, dass maschinelle Übersetzung zwar funktional eine Brücke baut, aber keine gleichwertige Alternative zu professioneller Sprachmittlung darstellt. Sie ermöglicht einen ersten Zugang, ersetzt aber nicht die tiefergehende Auseinandersetzung mit kulturellem und sozialem Kontext. Auffällig ist zudem, dass Fachkräfte ihre eigene Rolle in der Interaktion neu verhandeln müssen: Sie werden zu Moderator:innen zwischen Mensch und Maschine und tragen die Verantwortung für mögliche Fehlinterpretationen, ohne deren Ursache vollständig kontrollieren zu können. Dieses Spannungsfeld erzeugt Unsicherheit, aber auch pragmatische Handlungsstrategien. Manche Fachkräfte nutzen maschinelle Übersetzer bewusst nur für einfache Informationen und vermeiden sie bei sensiblen Themen. Andere entwickeln Routinen, indem sie Übersetzungen paraphrasieren oder mehrfach absichern.

In der Diskussion werden die empirischen Befunde mit den theoretischen Überlegungen verknüpft. Maschinelle Übersetzung erscheint als Instrument mit doppelter Funktion: Sie eröffnet Kommunikationsräume und fördert situative Inklusion, stabilisiert aber keine nachhaltige Teilhabe. Vielmehr besteht die Gefahr einer „Scheininklusion“, wenn Verständigung nur oberflächlich gelingt und strukturelle Barrieren bestehen bleiben. Der Rückgriff auf technische Hilfsmittel verschleiert mitunter den Bedarf an professionellen Sprachmittlungsdiensten und verstärkt die Tendenz, Verantwortung für gelingende Kommunikation zu externalisieren. Damit stellt sich die Frage nach den Grenzen digitaler Assistenz in einem Handlungsfeld, das auf Beziehungsgestaltung und Vertrauen angewiesen ist.

Die Arbeit schließt mit Handlungsempfehlungen für die Praxis. Neben einer kritisch-reflexiven Haltung gegenüber Technologie wird die Kombination maschineller Übersetzung mit menschlicher Dolmetschleistung gefordert. Fachkräfte sollten klare Kriterien entwickeln, wann der Einsatz digitaler Tools angemessen ist und wann professionelle Übersetzung zwingend notwendig bleibt. Ebenso relevant sind verbindliche Qualitätsstandards, medienpädagogische Schulungen und institutionelle Strategien, um digitale Hilfsmittel nicht isoliert, sondern eingebettet in ein Gesamtkonzept gelingender Kommunikation einzusetzen. Nur so kann die Soziale Arbeit ihrem Auftrag gerecht werden, Inklusion zu fördern und Exklusion zu vermeiden – auch im digitalen Wandel.

Angaben zum Autor

Jonathan May hat Soziale Arbeit an der Katholischen Hochschule NRW, Abteilung Köln, studiert. Er hat hier seinen Bachelor abgeschlossen und arbeitet seitdem in der öffentlichen Kinder- und Jugendhilfe in den Fachdiensten „unbegleitete Minderjährige mit Fluchterfahrung“ und dem Kinderschutz. Kontakt: [jo.may@netcologne.de].

Literatur

- Barlösius, E. (2004). *Kämpfe um soziale Ungleichheit. Machttheoretische Perspektiven*. VS Verlag für Sozialwissenschaften
- Berghaus, M. (2011). *Luhmann verstehen*. UVK.
- Hartmann, M., & Hepp, A. (2010). *Die Mediatisierung der Alltagswelt*. VS Verlag.
- Krotz, F. (2001). *Die Mediatisierung kommunikativen Handelns*. VS Verlag.
- Luhmann, N. (2021). *Soziale Systeme: Grundriss einer allgemeinen Theorie*. Suhrkamp.
- Ottersbach, M., & Platte, A. (2016). *Soziale Arbeit und Inklusion*. Beltz Juventa.
- Wansing, G., Schäfers, M., & Köbsell, S. (2022). *Teilhabeforschung – Konturen eines neuen Forschungsfeldes*. Springer VS
- Wittpahl, V. (2019). *Künstliche Intelligenz verstehen. Technologie. Anwendung. Gesellschaft*. Springer Vieweg.
- Wunder, M. (Hrsg.). (2021). *Digitalisierung und Soziale Arbeit*. Verlag Julius Klinkhardt.

Auszug aus dem Gutachten von Prof.in Dr.in Karolin Kappler

1. Gesamteindruck: Herr May hat eine Arbeit zu einer hochaktuellen Fragestellung an der Schnittstelle von Digitalisierung, Migration und Sozialer Arbeit vorgelegt. Sie analysiert den Einsatz maschineller Übersetzungsprogramme in der Kommunikation mit unbegleiteten Minderjährigen mit Fluchterfahrung. Die Arbeit zeigt, dass digitale Technologien in der Jugendhilfe unverzichtbar sind, ihre Nutzung aber komplexe Herausforderungen birgt. Der Autor verbindet theoretische Ansätze zu Systemtheorie, Mediatisierung und Teilhabe präzise mit Praxisbeispielen. Besonders hervorzuheben ist die Reflexion sozialer und ethischer Implikationen. Maschinelle Übersetzung wird sowohl als Chance zur Inklusion als auch als Risiko für Exklusion diskutiert. Die Arbeit leistet einen wichtigen Beitrag zur professionsethischen Debatte und markiert Forschungsdesiderate.

2. Fragestellung und Aufbau der Arbeit: Die Fragestellung ist präzise und praxisrelevant, die Arbeit konsequent darauf ausgerichtet. Die Gliederung ist jedoch teilweise unausgewogen: Ein sehr umfangreicher Theorie-Teil steht einer vergleichsweise knappen Ergebnisdarstellung gegenüber, einige kurze Kapitel hätten zusammengefasst werden können. Eine straffere Struktur wäre wünschenswert gewesen.

3. Gedankenführung: Die Herleitung zentraler Begriffe in Verbindung mit Theorie verleiht inhaltliche Tiefe, führt aber stellenweise zu Wiederholungen. Die Argumentation ist logisch und bleibt eng an der Forschungsfrage orientiert.

4. Fähigkeit zu kritisch-reflektierender Auseinandersetzung: Die Arbeit zeigt eine fundierte kritisch-reflexive Auseinandersetzung, die theoretisch gut verankert ist. Herr May greift systemtheoretische und mediatisierungstheoretische Ansätze auf, um die Rolle von Sprache und digitaler Technik in sozialen Inklusions- und Exklusionsprozessen zu analysieren. Darauf aufbauend diskutiert er die Chancen und Risiken maschineller Übersetzung für die Teilhabe unbegleiteter Minderjähriger und reflektiert die Implikationen für professionelles Handeln in der Sozialen Arbeit. Besonders gelungen ist die Einschätzung der Reichweite der eigenen Ergebnisse.

5. Literatur und Quellen: Die Quellen sind einschlägig, es wird jedoch empfohlen, stärker auf Originalquellen zurückzugreifen.

Die Bachelorarbeit wurde mit der Note 1,3 bewertet.

Franziska Erdtmann¹

Die Bedeutung von Peer Support bei der Bewältigung belastender Online-Erfahrungen von Mädchen*

¹ Hochschule München

Die vollständige Abschlussarbeit ist einsehbar unter:

<https://opus4.kobv.de/opus4-hm/frontdoor/index/index/docId/726>

Zusammenfassung

Die Masterarbeit beschäftigt sich mit der Rolle von Peer Support als Unterstützungsinstrument zur Stärkung von Mädchen in digitalen Belastungssituationen. Ausgangspunkt ist die zunehmende Konfrontation von Jugendlichen mit digitalen Risiken wie sexualisierter Belästigung, Hate Speech oder Cybermobbing, von denen insbesondere Mädchen* betroffen sind. Im Mittelpunkt steht die Frage, wie Peers als erste Ansprechpersonen bei belastenden Online-Erfahrungen fungieren und welche Kompetenzen sie dafür benötigen und bereits mitbringen. Aufbauend auf theoretischen Grundlagen aus Medienpädagogik und Sozialer Arbeit wird ein Kompetenzprofil für unterstützende Peers entwickelt sowie Handlungsempfehlungen für die Soziale Arbeit gegeben.*

Schlüsselwörter: Jugendarbeit, Jugend und digitale Medien, Mädchen, Peer-Group, digitale Gefährdungen*

The Significance of Peer Support in Coping With Distressing Online Experiences of Girls*

Abstract

This master's thesis focuses on the importance of friendships as a support tool for empowering girls in situations of digital harassment. The topic is based on the increasing confrontation of young people with digital risks such as sexualised harassment, hate speech or cyberbullying, which particularly affect girls*. The focus is on the question of how peers act as the first point of contact in the event of stressful online experiences and what skills they need and*



already have for this. Based on theoretical principles from media education and social work, a competence profile for supportive peers is developed and recommendations for social work are given.

Keywords: youth work, youth and digital media, girls, peer group, digital risks*

1. Skizze der Masterarbeit

Gleicher, bezahlbarer und tatsächlicher Zugang zu digitalen Medien sowie deren Nutzung für Bildung, soziale Teilhabe und persönliche Zwecke sind grundlegende Kinderrechte. Gemäß der UN-Kinderrechtskonvention sind Staaten verpflichtet, diesen Zugang für alle Kinder und Jugendlichen zu gewährleisten (Hajok, 2023, S.49). Mit der Nutzung digitaler Medien gehen jedoch auch reale Gefahren einher, die dazu führen, dass junge Menschen schon früh ungewollt mit belastenden Inhalten konfrontiert werden (Brüggen et al., 2019). Studien wie die EU Kids Online- und die JIM-Studie zeigen, dass nahezu alle Jugendlichen regelmäßig digitale Medien nutzen und dabei mit potenziell schädlichen Inhalten wie Gewaltdarstellungen, sexuellen Darstellungen, Falschinformationen oder unerwünschten Nachrichten von Unbekannten konfrontiert werden. Insbesondere Mädchen* und Frauen* sind im digitalen Raum häufiger als Jungen* und Männer* geschlechtsspezifischen digitalen Risiken wie sexualisierter Belästigung und Beleidigung, aber auch Formen von Cybermobbing, Bedrohung oder Stalking ausgesetzt. Solche Risiken können das Wohlbefinden erheblich beeinträchtigen und das Vertrauen in digitale Räume nachhaltig schädigen.

Der Schutz vor diesen Erfahrungen und Risiken wird zu einer zentralen Aufgabe im Prozess des Aufwachsens und ist damit ein relevantes Thema der Kinder- und Jugendhilfe. Für die Soziale Arbeit ergibt sich hier ein neuer Handlungsauftrag: Es gilt, Konzepte zum Schutz des Kindeswohls im Kontext digitaler Mediennutzung zu entwickeln und anzuwenden (Biesel et al., 2023). Ein zentraler Ansatz ist dabei die Befähigung junger Menschen zum Selbstschutz. Jugendliche sollen lernen, digitale Risiken zu erkennen und im Idealfall selbstständig zu umgehen.

Die Masterarbeit untersucht die Rolle von Peer Support als Unterstützungsinstrument zur Stärkung von Mädchen* in solchen digitalen Belastungssituationen. Dies ist insbesondere für die Soziale Arbeit von Interesse, da Fachkräfte oft keinen direkten Einblick in die digitale Lebenswelt der Jugendlichen haben.

Aufbau der Arbeit: Im Rahmen der vorliegenden Arbeit wird zunächst die Rolle der digitalen Medien im Jugendalter beleuchtet. Im weiteren Verlauf werden zentrale Herausforderungen und Belastungen im Online-Kontext thematisiert. Auf Basis der EU -Studie Kids Online sowie weiterer

Forschungsberichte wird ein umfassender Überblick über digitale Risiken für Kinder und Jugendliche gegeben. Dabei wird auch auf geschlechtsspezifische Unterschiede in Bezug auf digitale Risiken eingegangen und anhand von Studien belegt. Abschließend wird die Stellung der Sozialen Arbeit im Umgang mit diesen Risiken und im Bereich des Kinder- und Jugendmedienschutzes analysiert.

Im weiteren Verlauf wird dargestellt, wie Jugendliche belastende digitale Erfahrungen wahrnehmen und bewältigen. Darüber hinaus wird anhand der Konzepte der Lebensbewältigung nach Böhnisch, des Medienkompetenzmodells nach Baacke sowie auf Grundlage der Vorstellung der Sozialen Arbeit als Vermittlerin von Medienkompetenz zusammengefasst, welche Rolle die Soziale Arbeit und sozialpädagogische Fachkräfte bei der Unterstützung der Bewältigung digitaler Risiken haben bzw. leisten können.

Ein weiterer Theorieblock widmet sich dem Einfluss der Gleichaltrigen- bzw. Peergroup auf die Bewältigung belastender Erfahrungen. Die Beobachtung, dass Jugendliche in belastenden Situationen häufig zuerst den Kontakt zu Gleichaltrigen suchen, hat dazu geführt, dass dieser Ansatz in der Pädagogik verstärkt in die Präventionsarbeit einbezogen wird. Im Rahmen der Arbeit werden verschiedene medienpädagogische Projektansätze der Peer-Unterstützung wie Juuupot, die Medienscouts NRW oder das Projekt Netzgänger vorgestellt und hinsichtlich ihrer Chancen und Risiken analysiert.

Im Rahmen der abschließenden Evaluation wurde auf Basis der vorgestellten Theorien und Literatur ein Kompetenzprofil für Peers entwickelt, welches die notwendigen Kompetenzen für eine effektive Unterstützung zusammenfasst:

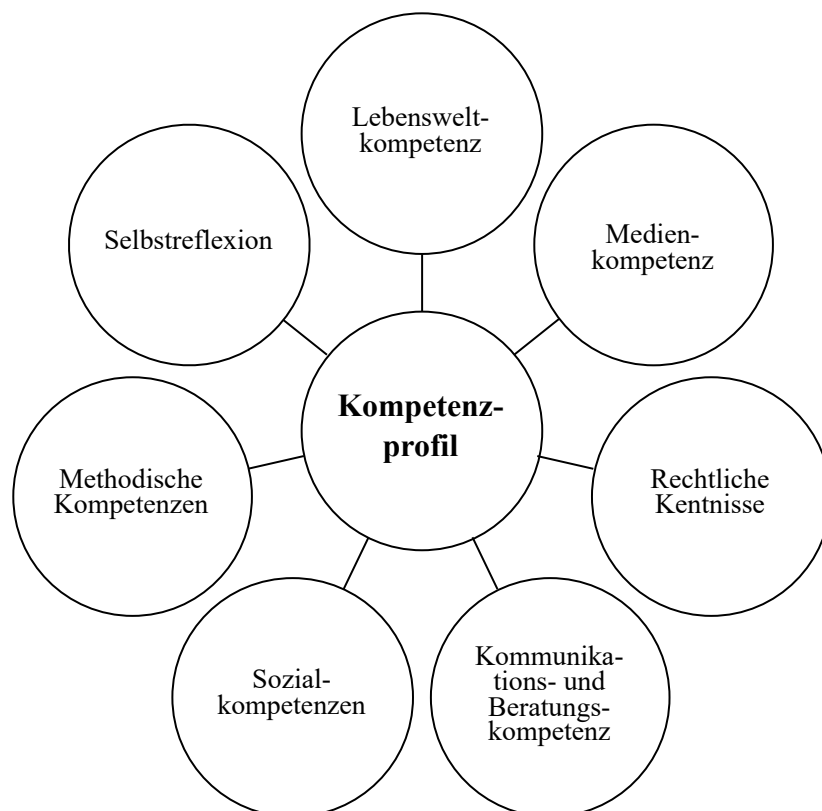


Abbildung 1: Kompetenzprofil von Peers in der Unterstützung von Gleichaltrigen bei der Bewältigung erlebter digitaler Risiken.

Der Forschungsteil der Masterarbeit verfolgt einen qualitativen Ansatz und untersucht anhand des Kompetenzmodells die vorhandenen Erfahrungen und Kompetenzen von Mädchen*, die ihre Freund:innen in schwierigen digitalen Situationen niedrigschwellig unterstützen. Die empirischen Ergebnisse zielen darauf ab, fehlende spezifische Kompetenzen von Peers zu identifizieren, die für die Unterstützung von Freund:innen notwendig sind und gezielt gefördert werden sollten.

Ergebnisse: Die Ergebnisse der Erhebung zeigen, dass die Wirksamkeit von Peer Support stark von den individuellen Kompetenzen der Jugendlichen abhängt. Neben technischen Kompetenzen, wie dem Blockieren oder Melden von Inhalten, sind auch soziale Kompetenzen von zentraler Bedeutung, wenn es um die Einschätzung der Wirksamkeit der Kompetenzen der Jugendlichen selbst geht. Dazu zählen insbesondere einfühlsames Zuhören und emotionale Unterstützung.

Unsicherheiten bestehen im rechtlichen Bereich und im Umgang mit plattformspezifischen technischen Funktionen. Es besteht ein deutlicher Aufklärungsbedarf über rechtliche Rahmenbedingungen, digitale Schutzmechanismen und Täter-Opfer-Dynamiken. Von besonderer Relevanz ist die Sensibilisierung für „Victim Blaming“ sowie die Stärkung von Mäd-

chen*, digitale Grenzverletzungen klar zu benennen und sich zu trauen, diese anzusprechen, insbesondere bei gesellschaftlich sexuell tabuisierten Themen.

In den Interviews wurde von den Jugendlichen der Wunsch nach einem ganzheitlichen Bildungsansatz geäußert, in dem Schule, Eltern und außerschulische Akteur:innen ein unterstützendes Netzwerk bilden.

Die Analyse hat gezeigt, dass es bei der Förderung von Jugendlichen nicht nur um die Vermittlung von Medienkompetenz geht, sondern auch um die Stärkung des Selbstbewusstseins und der Selbstbestimmung von Kindern und Jugendlichen in digitalen Räumen. Darüber hinaus sind Fachkräfte der Sozialen Arbeit, die in der unmittelbaren Lebenswelt der Heranwachsenden tätig sind, selbst in der Verantwortung, das oben beschriebene Kompetenzprofil bestmöglich zu erfüllen.

Angaben zur Autorin

Franziska Erdtmann absolvierte sowohl ihren Bachelor in Sozialer Arbeit als auch ihren Master im Studiengang Soziale Arbeit, Forschung und Digitalisierung an der Hochschule München. Bereits im Rahmen ihrer Bachelorarbeit beschäftigte sie sich intensiv mit digitalen Aspekten der Sozialen Arbeit, insbesondere der digitalen aufsuchenden Jugendarbeit. Im anschließenden Masterstudium vertiefte sie ihr Interesse an digitalen Entwicklungen in der Sozialen Arbeit und verband dort wissenschaftliche Forschung mit praxisorientierten Fragestellungen rund um Digitalisierung und Soziale Arbeit.

Relevante Literatur

- Baacke, D., & Bundeszentrale für Politische Bildung (Hrsg.). (1999). *Handbuch Medien: Medienkompetenz – Modelle und Projekte*. Bonn.
- Biesel, K., Burkhard, P., Heeg, R., & Steiner, O. (Hrsg.). (2023). *Digitale Kindeswohlgefährdung: Herausforderungen und Antworten für die Soziale Arbeit*. Verlag Barbara Budrich.
- Biesel, K., Heeg, R., Burkhard, P., & Steiner, O. (2023). Digitale Kindeswohlgefährdung: Eine begriffliche Annäherung. In Biesel, K., Burkhard, P., Heeg, R. & Steiner, O. (Hrsg.). (2023) *Digitale Kindeswohlgefährdung: Herausforderungen und Antworten für die Soziale Arbeit*. Opladen Berlin Toronto: Verlag Barbara Budrich. S. 20–34
- Böhnisch, L. (2023). *Lebensbewältigung: ein Konzept für die Soziale Arbeit*. 3., überarbeitete und erweiterte Auflage. Weinheim Basel: Beltz Juventa.
- Brüggen, N., Dreyer, S., Gebel, C., Lauber, A., Müller, R., & Stecher, S. (2019): *Gefährdungsatlas. Digitales Aufwachsen. Vom Kind aus denken. Zukunftssicher handeln*. Herausgegeben von: Bundesprüfstelle für jugendgefährdende Medien. Bonn 2019.

- Fileccia, M. (2016). *Kompetenzentwicklung bei der Ausbildung von „Medienscouts“ als medienpädagogischen Ansatz der Peer-Education. Eine empirische Untersuchung zur Medienkompetenz von Schülerinnen und Schülern*. https://duepublico2.uni-due.de/receive/duepublico_mods_00041820
- Hajok, D. (2023). Verändertes Heranwachsen mit digitalen Medien: Neue Herausforderungen für das Aufwachsen und das Wohl von Kindern und Jugendlichen. In Biesel, K., Burkhard, P., Heeg, R. & Steiner, O. (Hrsg.), *Digitale Kindeswohlgefährdung: Herausforderungen und Antworten für die Soziale Arbeit*. Opladen Berlin Toronto: Verlag Barbara Budrich. S. 35–54.
- Köhler, S.-M., Krüger, H.-H., & Pfaff, N. (Hrsg.). (2016). *Handbuch Peerforschung* (1. Aufl.). Verlag Barbara Budrich. <https://doi.org/10.2307/j.ctvd7w8m2>
- Nörber, M. (Hrsg.). (2003). *Peer Education: Bildung und Erziehung von Gleichaltrigen durch Gleichaltrige* (1. Aufl.). Beltz.
- Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S., & Hasebrink, U. (2020). *EU Kids Online 2020: Survey results from 19 countries*. *EU Kids Online*. <https://doi.org/10.21953/lse.47fdeqj0lfo>
- Thiel, K., & Lampert, C. (2023). *Wahrnehmung, Bewertung und Bewältigung belastender Online-Erfahrungen von Jugendlichen: Eine qualitative Studie im Rahmen des Projekts „SIKID – Sicherheit für Kinder in der digitalen Welt“*. Arbeitspapiere des Hans-Bredow-Instituts. <https://doi.org/10.21241/SSOAR.86633>

Empfehlung zur Publikation von Prof. Emily Engelhardt

Die vorliegende Masterarbeit widmet sich einem hochaktuellen und gesellschaftlich relevanten Thema: der Rolle von Peer Support bei der Bewältigung belastender Online-Erfahrungen von Mädchen*. Sie verbindet theoretische Ansätze aus der Sozialen Arbeit und Medienpädagogik mit empirischen Erkenntnissen, um die Bedeutung informeller Unterstützungsnetzwerke im digitalen Kontext zu analysieren. Die Arbeit überzeugt durch ihre interdisziplinäre Herangehensweise und die klare Verknüpfung von Theorie und Praxis.

Die Arbeit greift eine zentrale Herausforderung der digitalen Lebenswelt auf, nämlich den Umgang mit geschlechtsspezifischen Risiken wie Cybermobbing oder sexualisierter Belästigung. Sie zeigt auf, wie Peer Support als niedrigschwellige und lebensweltnahe Ressource zur Stärkung von Mädchen* beitragen kann. Die Autorin setzt sich fundiert mit Konzepten wie Lebensbewältigung (Böhnisch) und Medienkompetenz (Baacke) auseinander und integriert diese in die Analyse der Rolle der Sozialen Arbeit. Der qualitative Forschungsansatz liefert praxisnahe Erkenntnisse über die Kompetenzen und Defizite von Mädchen* im Umgang mit belastenden

Online-Erfahrungen. Besonders hervorzuheben ist die Entwicklung eines Kompetenzprofils für Peers, das konkrete Anknüpfungspunkte für medienpädagogische Interventionen bietet.

Die Ergebnisse dieser Masterarbeit sind nicht nur wissenschaftlich fundiert, sondern auch praxisorientiert und innovativ. Sie bieten wichtige Impulse für Fachkräfte in der Sozialen Arbeit sowie für medienpädagogische Projekte, um Mädchen* gezielt im Umgang mit digitalen Risiken zu stärken. Die Veröffentlichung in einer Fachzeitschrift würde dazu beitragen, das Bewusstsein für die Bedeutung von Peer Support zu schärfen und den Wissenstransfer zwischen Wissenschaft und Praxis zu fördern. Angesichts der wachsenden Relevanz digitaler Risiken ist diese Arbeit ein wertvoller Beitrag zur Weiterentwicklung des Kinder- und Jugendmedienschutzes.

Die Arbeit wurde bewertet mit der Note 1,3

Sebastian Krumm¹

KI und Soziale Arbeit

Eine Studie zur Akzeptanz des Technologieeinsatzes unter Fachkräften der Sozialen Arbeit

¹ Hochschule München

Die vollständige Abschlussarbeit ist einsehbar unter:

<https://www.dime-sozial.de/index.php/dimesoz/article/view/74/23>

Zusammenfassung

Die Masterarbeit geht der Frage nach, inwieweit Fachkräfte der Sozialen Arbeit den (potenziellen) Einsatz von Künstlicher Intelligenz (KI) in ihrer beruflichen Praxis akzeptieren. Grundlage ist eine bundesweite Online-Befragung mit 338 Teilnehmenden. Untersucht wurden u. a. der Einfluss beruflicher und demografischer Merkmale, die Sorge vor Deprofessionalisierung, das selbst eingeschätzte KI-Wissen sowie berufsethische Überzeugungen. Die Ergebnisse zeigen eine überwiegend vorsichtig positive Haltung gegenüber KI. Eine generelle Ablehnung oder ein Gefühl von Ohnmacht sind kaum erkennbar. Besonders hoch ist die Akzeptanz bei verwaltungstechnischen Anwendungen, doch auch in zentralen Aufgabebereichen der Sozialen Arbeit wird der KI-Einsatz nicht grundsätzlich ausgeschlossen. Als relevante Faktoren für die Akzeptanz erscheinen insbesondere der eigene (selbsteingeschätzte) Wissensstand und das berufsethische Urteil. Die Arbeit schlussfolgert, dass weitere gezielte Angebote zur Wissensaneignung und Praxiserfahrung sowie ein ethisch und rechtlich fundierter Rahmen notwendig sind, um den Einsatz von KI im Sinne der Profession zu gestalten. Der Beitrag enthält einen Auszug aus dem Gutachten zur Arbeit und dient der Bereitstellung der empirischen Ergebnisse im Diskurs. Das Gesamtwerk inklusive Zitiervorschlag findet sich im Anhang und ist online abrufbar über dime-sozial.de.

Schlüsselwörter: Künstliche Intelligenz (KI), KI-Akzeptanz in der Sozialen Arbeit, empirische Untersuchung

AI and Social Work. A Study on the Acceptance of the Use of Technology Among Social Work Professionals

Abstract

The master's thesis examines the extent to which social work professionals accept the (possible) use of artificial intelligence (AI) in their professional practice. It is based on a nationwide online survey with 338 participants. Among other things, the study examined the influence of professional and demographic characteristics, concerns about deprofessionalization, self-assessed AI knowledge, and professional ethical convictions. The results show a predominantly cautiously positive attitude toward AI. There is hardly any evidence of general rejection or a feeling of powerlessness. Acceptance is particularly high for administrative applications, but the use of AI is not fundamentally ruled out in key areas of social work either. Relevant factors associated with acceptance include self-assessed knowledge and professional ethical judgment. The thesis concludes that further targeted opportunities for knowledge acquisition and practical experience, as well as an ethically and legally sound framework, are necessary to shape the use of AI in the interests of the profession. This article contains an excerpt from the thesis and serves to provide the empirical results for discussion. The complete work, including citation suggestions, can be found in the appendix, available online at dime-sozial.de.

Keywords: Artificial intelligence (AI), AI acceptance in social work, empirical study

Skizze der Masterarbeit

Die dem Beitrag zugrundeliegende Masterarbeit stellt erste empirische Befunde aus einer bundesweiten Online-Fragebogen-Umfrage (N = 338) unter berufstätigen Fachkräften der Sozialen Arbeit vor. Im Zentrum der Untersuchung steht die Frage, inwieweit Sozialarbeiter:innen den (potenziellen) Einsatz von Künstlicher Intelligenz (KI) in ihrer beruflichen Praxis akzeptieren. Dabei wurde besonders auf kontextbezogene Unterschiede, demografische und berufliche Faktoren, die Befürchtung einer Deprofessionalisierung durch KI sowie die Selbsteinschätzung der eigenen KI-Kompetenz der Fachkräfte fokussiert. Zudem wurde auf Zusammenhänge zwischen der Akzeptanz von KI mit der Betroffenheit von Fachlogiken bzw. den berufsethischen Vorstellungen und dem selbsteingeschätzten KI-Wissen getestet.

Methodik: Methodisch stützt sich die Studie auf ein quantitatives Forschungsdesign mit begründeten Mixed-Methods-Anteilen, das sowohl etablierte Akzeptanzmodelle wie das „Technology Acceptance Model“ von Venkatesh & Bala 2008 bzw. Scheuer 2020 berücksichtigt, als auch diese vor dem Hintergrund der für die Sozialen Arbeit charakteristischen Mandate und Wertbindungen kritisch reflektiert und erweitert (etwa

Staub-Bernasconi 2018). Die Akzeptanzmessung erfolgte sowohl allgemein als auch kontextualisiert im Rahmen von fünf in der Theorie bereits diskutierten Anwendungsszenarien von KI für die Praxis Sozialer Arbeit.

Insbesondere durch die Bezugnahme auf bereits validierte Messinstrumente wie der Attitude Towards Artificial Intelligence-Skala (ATAI) von Sindermann et al. (2021), die KI-Wissens-Skala von Cousseran et al. (2023) oder auch die Übernahme von Merkmalsausprägungen aus der DIGITASA-Studie (Matthies et al. 2023) kann von einer erhöhten Anschlussfähigkeit der Ergebnisse im Diskurs ausgegangen werden.

Die Auswertung des Datenmaterials erfolgte anhand detaillierter deskriptiver und inferenzstatistischer Analysen, die auch die jeweiligen Testbedingungen berücksichtigten. Neben parametrischen Hypothesentests wurden an geeigneten Stellen auch deren nichtparametrische Pendant berechnet, um die empirischen Ergebnisse weiter abzusichern. Anschließend wurden die Ergebnisse, auch unter Bezugnahme auf die qualitativen Impulse aus dem Fragebogen, kritisch gewürdigt, diskutiert und vor dem Hintergrund des aktuellen Diskurses zu KI in der Sozialen Arbeit eingeordnet.

Ergebnisse der Arbeit: Die Ergebnisse zeigen, dass – anders als es die Vermutung nahelegt – wenig pauschale Ablehnung gegenüber KI besteht. Die Deprofessionalisierungsthese kann in Teilen entkräftet werden. Bei den Befragten ist auch kein Gefühl einer Ohnmacht gegenüber KI feststellbar. Die Ergebnisse verweisen vielmehr auf eine vorsichtig positive Grundhaltung, die eine kontextsensible und fachlich reflektierte Auseinandersetzung mit KI-Technologien einfordert, sowie ebendiesen Einsatz von KI in der Praxis Sozialer Arbeit zumindest nicht grundsätzlich ausschließt.

Von Relevanz für die Akzeptanz zeigt sich der selbsteingeschätzte Wissensstand zu KI sowie das berufsethische Urteil der Fachkräfte. Ein Zusammenhang mit dem Alter ließ sich hingegen nicht feststellen. Auffällig war, dass Fachkräfte in höheren Hierarchieebenen signifikant höhere Akzeptanz- und geringere Angstwerte in Hinblick auf KI zeigten. Zudem schätzten männliche Befragte ihr eigenes Wissen über KI signifikant höher ein als weibliche Befragte.

Die höchste Akzeptanz ist derzeit bei der Unterstützung von verwaltungstechnischen Abläufen zu sehen, allerdings wird der Einsatz von KI auch in jenen Aufgabenbereichen nicht kategorisch abgelehnt, die von den Fachkräften als Kern ihrer Arbeit angesehen werden.

Fazit der Arbeit: Die Ergebnisse implizieren, dass zukünftig insbesondere Rahmenbedingungen zu schaffen sind, die den weiteren Wissenszuwachs hinsichtlich KI unter den Fachkräften ermöglichen. Weiterhin gilt es,

partizipativ einen rechtlichen und ethischen Rahmen für den KI-Einsatz im Kontext Sozialer Arbeit zu etablieren, der den Zielen Sozialer Arbeit gerecht wird. Die Unterstützung durch Arbeitgeber bzw. Träger ist hierbei ebenso notwendig wie erwünscht.

Angaben zum Autor:

Sebastian Krumm ist Sozialarbeiter und Absolvent des Masterstudiengangs „Soziale Arbeit, Forschung und Digitalisierung“ der Hochschule München. Seine Arbeitsschwerpunkte liegen im Bereich Jugendarbeit, Musik- und Kulturarbeit, Mediatisierung und Wissenschaft Sozialer Arbeit. Seine Masterarbeit wurde für den Förderpreis 2025 des Deutschen Berufsverbandes für Sozialarbeit e. V. (DBSH) und des Fachbereichstages Soziale Arbeit e. V. (FBTS) nominiert. Derzeit arbeitet er in der Offenen Kinder- und Jugendarbeit des Jugendamts der Stadt Nürnberg.

Auszug aus dem Gutachten von Prof. Dr. Angelika Beranek:

Die vorliegende Masterarbeit von Sebastian Krumm widmet sich einem hochaktuellen und für die Soziale Arbeit bedeutsamen Thema: der Akzeptanz von Künstlicher Intelligenz (KI) unter Fachkräften. Die Arbeit zeichnet sich durch eine klare Fragestellung, eine fundierte methodische Vorgehensweise und eine überzeugende Ergebnisdarstellung aus. Sie leistet einen wichtigen Beitrag zur Versachlichung der Debatte um den Einsatz von KI in der Sozialen Arbeit und bietet wertvolle Erkenntnisse für die Weiterentwicklung des Feldes. Besonders hervorzuheben sind hierbei folgende Punkte:

Relevanz und Aktualität: Die Arbeit behandelt ein Thema, das in der Sozialen Arbeit zunehmend an Bedeutung gewinnt- den Einsatz von KI. Hierbei wird häufig auf die konkreten Ebene des Einsatzes von KI geschaut. Sebastian Krumm ändert hier die Perspektive auf eine Metaebene und liefert eine empirisch fundierte Grundlage für die Diskussion.

Theoretische Fundierung: Der Autor beweist ein tiefes Verständnis der relevanten Theorien und Konzepte. Er verknüpft die Thematik der KI-Akzeptanz überzeugend mit etablierten Methoden und Theorien der Sozialen Arbeit.

Methodische Stringenz: Die Arbeit basiert auf einem quantitativen Forschungsdesign mit Mixed-Methods-Anteilen, das der Fragestellung angemessen ist. Die Konstruktion des Fragebogens ist sorgfältig durchdacht und die Durchführung der Online-Umfrage professionell umgesetzt. Die erfasste Stichprobe übertrifft die Erwartungen an eine Masterarbeit bei weitem. (...)

Diskussion und Ausblick: Die Arbeit schließt mit einer umfassenden Diskussion der Ergebnisse und einem Ausblick auf zukünftige Forschungsfragen. Der Autor formuliert konkrete Handlungsempfehlungen für die Praxis der Sozialen Arbeit und betont die Notwendigkeit, Rahmenbedingungen zu schaffen, die den Wissenszuwachs unter den Fachkräften fördern und einen ethischen Rahmen für den KI-Einsatz etablieren.

Die Masterarbeit von Sebastian Krumm stellt somit eine herausragende wissenschaftliche Leistung, die sich durch Relevanz, methodische Stringenz und eine überzeugende Ergebnisdarstellung auszeichnet, dar. Die Arbeit leistet einen wichtigen Beitrag zur Debatte um den Einsatz von KI in der Sozialen Arbeit und bietet wertvolle Impulse für die Weiterentwicklung des Feldes.

Gesamtnote: 1,0

Literatur

- Beranek, A. (2021). *Soziale Arbeit im Digitalzeitalter. Eine Profession und ihre Theorien im Kontext digitaler Transformation*. Beltz Juventa.
- Cousseran, L., Lauber, A., Herrmann, S., & Brüggem, N. (2023). *Kompass: Künstliche Intelligenz und Kompetenz 2023. Einstellungen, Handeln und Kompetenzentwicklung im Kontext von KI* (Herausgegeben vom JFF – Institut für Medienpädagogik in Forschung und Praxis). kopaed.
- Görder, B. (2021). Die Macht der Muster. Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch ‚künstliche Intelligenz‘. *EthikJournal*, 7(2), 1–22.
- Krumm, S. (2026). KI und Soziale Arbeit. Eine Studie zur Akzeptanz des Technologieeinsatzes unter Fachkräften der Sozialen Arbeit (Gesamtwerk). *DiMe-Sozial*, 2(1). <https://www.dime-sozial.de/index.php/dimesoz/article/view/74/23>
- Matthies, A., Tetens, J., & Wahren, J. (2023). Zwischen Arbeitserleichterung und De-Professionalisierungsgefahr: Perspektiven auf Stand und Auswirkungen der Digitalisierung in Einrichtungen der Sozialen Arbeit. *Sozial Extra*, 47(3), 158–164. <https://doi.org/10.1007/s12054-023-00588-2>
- Scheuer, D. (2020). *Akzeptanz von Künstlicher Intelligenz: Grundlagen intelligenter KI Assistenten und deren vertrauensvolle Nutzung*. Springer Vieweg. <https://doi.org/10.1007/978-3-658-29526-4>
- Schneider, D., & Seelmeyer, U. (2018). Der Einfluss der Algorithmen. *Sozial Extra*, 42, 21–24. <https://doi.org/10.1007/s12054-018-0046-y>
- Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2021). Assessing the attitude towards artificial intelligence: Introduction of a short measure in German, Chinese, and English language. *KI – Künstliche Intelligenz*, 35, 109–118. <https://doi.org/10.1007/s13218-020-00689-0>

- Staub Bernasconi, S. (2018). *Soziale Arbeit als Handlungswissenschaft: Auf dem Weg zu kritischer Professionalität* (2. Aufl.). Verlag Barbara Budrich.
- Steiner, O., & Tschopp, D. (2022). Künstliche Intelligenz in der Sozialen Arbeit Grundlagen, Entwicklungen, Herausforderungen. *Sozial Extra*, 46, 466–471. <https://doi.org/10.1007/s12054-022-00546-4>
- Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273–315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>