

Künstliche Intelligenz in der Sozialen Arbeit

Grundlagen, Entwicklungen, Herausforderungen

Künstliche Intelligenz (KI; engl. Artificial Intelligence) ist eine Begrifflichkeit, die in der öffentlichen und wissenschaftlichen Diskussion in den letzten Jahren zunehmend präsent, inhaltlich oft aber diffus und mit unterschiedlichen Konnotationen behaftet ist. Auch in der Sozialen Arbeit wird der Einsatz von KI zur Unterstützung der Leistungserbringung diskutiert und es bestehen bereits erste Anwendungsfälle sowie Pilotprojekte, die die Potenziale, Grenzen und Risiken von KI für die Soziale Arbeit ausloten.

Künstliche Intelligenz ist ein weites, interdisziplinäres und wenig übersichtliches Teilgebiet der Informatik (Lenzen 2018). KI zu definieren ist daher kein leichtes Unterfangen. Rich (1983) umschreibt das Ziel des Fachgebiets damit, Computer so zu konstruieren, dass sie Dinge tun können, die Menschen, zum jeweiligen Zeitpunkt, besser tun können. Diese Definition bringt zum Ausdruck, dass die Grenzen des Machbaren und damit der Gegenstand der Künstlichen

Intelligenz immer wieder neu definiert werden, wobei die Geschichte der KI auch durch mehrere Phasen der Stagnation gekennzeichnet ist.

Das intelligente Verhalten von Systemen unterscheidet sich davon (zumindest bislang), intelligent im Sinne menschlicher Intelligenz zu sein. Bis heute zielt der größte Teil der KI-Forschung auf die Entwicklung von spezialisierten Systemen in bestimmten, eher eng gefassten Anwendungsbereichen (z. B. Software für die Sprachübersetzung). Man spricht in diesem Zusammenhang von enger Künstlicher Intelligenz. Im Gegensatz dazu steht generelle Künstliche Intelligenz, die mehrere Bereiche der Intelligenz umfasst und diese im Sinne menschlicher Flexibilität in unterschiedliche Anwendungsbereiche einbringen kann.



Olivier Steiner

FHNW University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Schweiz
 *1970; Dr., Professor am Institut Kinder- und Jugendhilfe, Hochschule für Soziale Arbeit, Fachhochschule Nordwestschweiz.
olivier.steiner@fhnw.ch



Dominik Tschopp

FHNW University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Schweiz
 *1980; Hochschule für Soziale Arbeit, Fachhochschule Nordwestschweiz.
dominik.tschopp@fhnw.ch

Zusammenfassung Der Beitrag definiert in seinem ersten Teil Künstliche Intelligenz (KI) und beschreibt Technologien, die KI ermöglichen (sollen). Im zweiten Teil werden zwei Anwendungsszenarien von KI in der Sozialen Arbeit beschrieben und mögliche Potenziale, Grenzen und Risiken der Technologie herausgearbeitet. Eingehender diskutiert werden hier Technologien des Predictive Risk Modellings (PRM) sowie Chatbots in Beratungssettings der Sozialen Arbeit. In einer abschließenden Diskussion wird der bisherige Stand der Entwicklung von KI in der Sozialen Arbeit in einem ethischen Modell kritisch reflektiert.

Schlüsselwörter Künstliche Intelligenz, Algorithmen, Predictive Risk Modelling, Chatbots, Soziale Arbeit, Digitalisierung

Begriffe im Kontext von Künstlicher Intelligenz

Um einen groben Einblick in die Funktionsweise aktueller KI-Technologien zu geben, werden im folgenden Abschnitt einige zentrale Begriffe kurz dargestellt:

- **Algorithmus:** Die öffentliche Debatte um KI hat insbesondere den Begriff des Algorithmus bekannt gemacht. Algorithmen sind generell Gegenstand der Informatik und nicht exklusiv im Bereich der KI verortet. Ein Algorithmus ist eine (programmierte) Anleitung, die beschreibt, wie man Schritt für Schritt ein Ziel erreicht. Die Absicht dahinter ist, einen Problemlösungsprozess digital zu automatisieren (Lenzen 2020).
- **Machine Learning/Maschinelles Lernen:** Das Forschungs- und Anwendungsfeld der Künstlichen Intelligenz hat in den letzten Jahrzehnten eine ganze Bandbreite an methodischen Verfahren hervorgebracht. In jüngster Zeit erhalten vor allem maschinelle Lernverfahren viel Aufmerksamkeit, wobei das maschinelle Lernen keine Kopie der Vorgänge im menschlichen

Gehirn darstellt, sich allenfalls grob daran orientiert (Lenzen 2020). Beim maschinellen Lernen kommen Algorithmen zum Einsatz, mit denen Computer eigenständig lernen, ohne dass in den laufenden Prozess eingegriffen werden muss (Kreutzer und Sirrenberg 2019). Der Algorithmus lernt dabei durch Beispiele und Versuche, ein Problem möglichst gut zu lösen (Lenzen 2020). Maschinelles Lernen kommt dort zum Einsatz, wo keine Regel bekannt ist, mit der sich die vorliegende Problemstellung lösen ließe. Im Gegenzug bedarf es für diesen Prozess meist große Datenmengen bzw. qualitativ hochwertiges Trainingsmaterial, anhand dessen ein KI-System lernen kann. Der maschinelle Lernprozess zielt auf Optimierung ab: Die Lernverfahren liefern eine möglichst gute Annäherung, keine perfekte Lösung. Es liegt in der Aufgabe der Programmierer_innen oder der Anwender_innen, zu entscheiden, welche Fehlerquote bzw. welcher Grenzwert für einen konkreten Anwendungsfall akzeptabel ist (Lenzen 2020).

- **Künstliche neuronale Netzwerke:** Im Kontext des maschinellen Lernens haben sich einige gängige Lernverfahren entwickelt. Besondere Aufmerksamkeit kommt aktuell dem Deep Learning auf Basis von künstlichen neuronalen Netzwerken zu. Der Begriff neuronales Netzwerk stammt ursprünglich aus den Neurowissenschaften. Er bezeichnet dort die Verbindung zwischen Neuronen, welche als Teil des Nervensystems bestimmte Funktionen ausüben (Kreutzer und Sirrenberg 2019). In der Informatik versteht man unter einem künstlichen neuronalen Netzwerk ein System von Hard- und Software, welche bestimmte Eigenschaften natürlicher neuronaler Netzwerke simuliert (Lenzen 2020). In der Regel verfügt ein künstliches neuronales Netzwerk über eine große Anzahl von Prozessoren, die parallel arbeiten und in mehreren Schichten angeordnet sind (Kreutzer und Sirrenberg 2019). Die Verbindungen bzw. die Signalstärke zwischen den künstlichen Neuronen können unterschiedlich gewichtet sein. Im Trainingsprozess wird diese Feinstruktur des Netzwerks optimiert (Lenzen 2020). Das System lernt selbstständig dazu und emanzipiert sich auf Basis der gewonnenen Erfahrungen zunehmend von den ursprünglichen Eingaben, um auf diesem Weg bessere Ergebnisse zu erzielen (Kreutzer und Sirrenberg 2019).
- **Deep Learning:** Deep Learning ist wie bereits angedeutet ein Teilgebiet von Machine Learning. Das „Deep“ (Tiefe) bezieht sich auf die große Anzahl der Schichten der dabei eingesetzten künstlichen neuronalen Netzwerke. Jede nachfolgende Schicht erhält den Output der vorgehenden Schicht. Bei jedem Übergang

von einer zur nächsten Schicht lernt das KI-System (idealerweise) dazu (Kreutzer und Sirrenberg 2019). Das System lernt so nach und nach, komplexe Konzepte aus einfacheren Elementen zusammenzusetzen (Lenzen 2020). Deep Learning-Verfahren haben daher den Vorteil, dass Trainingsdaten vorgängig nicht mehr manuell gekennzeichnet (gelabelt) werden müssen (z. B. physiologische Anomalien in radiologischen Aufnahmen). Sie können auch bei unstrukturierten Daten zu guten Ergebnissen führen. Somit kann eine größere Bandbreite an Daten verarbeitet werden. In vielen Fällen können Deep Learning-Verfahren zu genaueren Ergebnissen führen als herkömmliche maschinelle Lernansätze. Zugleich steigt jedoch der Bedarf an Trainingsdaten und an Rechenleistung. Mit der zunehmenden Komplexität von künstlichen neuronalen Netzwerken wird es zudem schwieriger, nachzuvollziehen, wie ein System zu einem bestimmten Ergebnis kommt (Lenzen 2020).

Im Folgenden werden am Beispiel von zwei Anwendungsszenarien von KI in der Sozialen Arbeit Potenziale, Grenzen und Risiken der Technologieanwendung für die Leistungserbringung diskutiert. Das erste Anwendungsszenario betrifft das Predictive Risk Modeling, d. h. die KI-gestützte Voraussage von Risiken in Fallverläufen bzw. die KI-gestützte Risikoeinschätzung, die erst Fälle in der Sozialen Arbeit generieren. Das zweite Anwendungsszenario betrifft den Einsatz von KI zur Beratung von Klient_innen, insbesondere deren Verwendung als sogenannte KI-Bots oder Chatbots.

Predictive Risk Modeling in der Sozialen Arbeit

Predictive Modelling soll im sozialen Bereich datenbasierte Vorhersagen mittels Algorithmen über menschliches Verhalten und Entwicklung treffen (hierbei kommen nicht ausschließlich KI-Technologien zur Anwendung, sondern auch „klassische“ statistische Berechnungen). Predictive Modelling existiert bereits seit den 1980er-Jahren insbesondere im Gesundheitsbereich, um bspw. rückfallgefährdete Patient_innen zu identifizieren. Predictive Risk Modelling (PRM) hat letztlich zum Ziel, ein Big Data basiertes Expert_innensystem zur Entscheidungsunterstützung in wohlfahrtsstaatlichen Organisationen zu schaffen (Gillingham 2016). Es sollen risikobehaftete Lebenssituationen oder riskantes Verhalten von Personen unter Einbezug von elektronischen Datenbeständen algorithmisch analysiert und Voraussagen über zukünftiges Verhalten oder die Entwicklung sozialer Verhältnisse getroffen werden.

Anwendungen von PRM-Anwendungen in der Sozialen Arbeit

Bisher bestehen erst wenige in der Praxis der Sozialen Arbeit verankerte Anwendungen des Predictive Risk Modellings. In einigen Fällen handelt es sich um Pilotversuche. Gillingham (2016) berichtet von Neuseeland, wo aufgrund einer Reform des Kinderschutzes vermehrt Datenbestände zur Identifikation von gefährdeten Kindern algorithmisch ausgewertet werden sollen. Die Voraussagekraft des eingesetzten neuronalen Modells wird von den Entwickler_innen auf 76 % eingeschätzt (von 100 Fällen, von welchen während des Trainings des Modells bekannt war, dass ein substantieller Übergriff stattfand, hat das Modell 76 richtig erkannt). Während die Betreiber_innen des Modells dies als sehr gute Erkennungsrate bezeichnen, schlagen diese selbst aber den Einbezug von weiteren Polizei- und Gesundheitsdaten vor, um die Voraussagekraft des Modells zu erhöhen. Dieses Vorgehen würde allerdings neue Fragen bezüglich des Datenschutzes der betroffenen Personen aufwerfen.

In den Niederlanden wurde eine Integration nationaler Datenbestände aus Polizei, Gesundheits- und Wohlfahrtssystem zur Früherkennung von Kindeswohlgefährdungen angestrebt. Allerdings mussten in der Folge Revisionen aufgrund von Datenschutzbedenken vorgenommen werden (Keymolen und Broeders 2013). Bei dem niederländischen Alarmsystem kommen zwar nicht im engeren Sinne KI-Technologien zum Einsatz – die Aggregation großer und bereichsübergreifender Datenbestände für die algorithmische Auswertung verdeutlicht aber auch über KI hinausweisende Kernprobleme der Schaffung nationaler Datenbestände im Kindes- und Erwachsenenschutz.

Herausforderungen, Problematiken und Potenziale von PRM

Oak (2016) diskutiert den Einsatz von PRM-Technologien in Neuseeland in Zusammenhang mit der neoliberalen Umgestaltung der Sozialen Arbeit entlang eines Risikoparadigmas, in dessen Zuge das Agieren in Grauzonen von Sozialarbeitenden und die dafür notwendige Reflexivität verloren gehen bzw. selbst als problematisch eingestuft werden. Die zentrale Problematik für die Entwicklung von PRM liegt nach Oak nicht in der Technologie selbst, sondern in der Managerialisierung Sozialer Arbeit und der Entwicklung technologischer Systeme nach Paradigmen der Ökonomisierung. Auch die eingesetzten Algorithmen selbst sind mitnichten neutral, sondern bilden oftmals ungerechte Lebensbedingungen ab und drohen diese durch ihren Bias zu verstetigen

(de Haan und Connolly 2014; Keymolen und Broeders 2013; Thinyane et al. 2018).

Mit dem erwartbar zunehmenden Einsatz von KI-Technologien stellt sich für die Gesellschaft im Allgemeinen und die Soziale Arbeit im Besonderen drängend die Frage nach Verantwortungszumessungen und -übernahmen in Entscheidungsprozessen, die vital für Inanspruchnahme wohlfahrtsstaatlicher Leistungen sind (Steiner 2021). Eubanks (2018) verweist in diesem Zusammenhang auf die katastrophalen Folgen automatisierter sozialer Exklusion durch Algorithmen, die die Rechtsansprüche ganzer Bevölkerungsschichten außer Kraft setzen. Mit der behördlichen Vernetzung von Klient_innendaten ergeben sich schließlich auch Probleme des „function creeping“: Daten, die ursprünglich zu einem spezifischen Zweck gesammelt wurden, werden zu neuen Zwecken verwendet, bspw. zur Überwachung und Kontrolle von Armutspopulationen (Keymolen und Broeders 2013). Eine ältere Forschungsarbeit (Schwartz et al. 2004) zum Einsatz neuronaler Netzwerke bei Kinderschutzesätzen zeigt weiter, dass die Qualität der Datensätze (Vollständigkeit, Akkuratheit der Einträge) eine wesentliche Voraussetzung für die Prädiktorstärke eines neuronalen Netzwerks ist. Viele Datensätze in Kinderschutzesystemen sind allerdings unvollständig und könnten im Praxiseinsatz deshalb zur Prädiktion bestehender Risiken gar nicht verwendet werden.

Eine grundsätzliche Herausforderung besteht für die KI-basierte Entscheidungsunterstützung in der Sozialen Arbeit im „Blackbox“ Charakter neuronaler Netzwerke. Urteilsbildungen von KI-Technologien sind damit kaum nachvollziehbar und es entfällt eine zentrale Komponente reflexiver Professionalität in der Sozialen Arbeit: Die diskursive Aushandlung und Verständigung aufgrund nachvollziehbarer Begründungen im fachlichen Kontext (vgl. Dewe 2009). In diesem Zusammenhang besteht durch die Anwendung von KI-Technologien bspw. in der Diagnose und Anamnese die Gefahr der Etikettierung von Fällen, da auch algorithmische Verfahren – zumindest in der gegenwärtigen Verwendungsweise – immer auf Standardisierung abzielen. Es bestehen bisher in der Sozialen Arbeit erst in Ansätzen Konzepte, wie eine fall- und feldsensible Fallrekonstruktion durch Verwendung von KI erfolgen kann (Schneider und Seelmeyer 2019).

Technologien des Predictive Risk Modelling können professionelle Beurteilungen im Idealfall unterstützen, indem der Bias Professioneller (practitioner bias) vermindert und damit Dienstleistungsverfahren standardisiert und gerechter werden (Oak 2016). Potenziell können solche Technologien also dazu beitragen, dass Benachteiligungen früh erkannt und professionelle Un-

terstützung sowie Förderung frühzeitig angeboten werden (de Haan und Connolly 2014). Mit Blick auf den aktuellen Stand der Entwicklung stellt sich jedoch weiterhin die Frage, inwiefern die Anwendung von KI-Systemen überhaupt eine individuellen Problemlagen angemessene Fallbearbeitung in der Sozialen Arbeit ermöglichen kann.

KI-gestützte Chatbots in der Sozialen Arbeit

Chatbots sind Computerprogramme, die text- oder sprachbasierte, natürliche Interaktionen zwischen Mensch und Maschine ermöglichen sollen (Rapp et al. 2021). Griol et al. (2013, S. 760) bezeichnen Chatbots als „Conversational Agents“ im Sinne einer „software that accepts natural language as input and generates natural language as output, engaging in a conversation with the user“. Chatbots erhalten damit eine soziale Qualität, indem menschliche Kommunikation bis hin zu Beziehungsführung imitiert wird. Die Fähigkeiten von Chatbots zur Konversation haben in den letzten Jahren aufgrund des Einsatzes von neuronalen Netzwerken und den Fortschritten der Verarbeitung natürlicher Sprache deutlich zugenommen (Araujo 2018).

Zwar hat im Zuge der massiven Verbreitung von Chatbots in den letzten Jahren die wissenschaftliche Forschung dazu Auftrieb erhalten. Allerdings fehlt nach Rapp et al. (2021) nach wie vor eine theoretisch-konzeptionelle Fassung des Phänomens. Insbesondere fehlt ein stärkerer Einbezug des Kontextes der Kommunikation, wenn Fragen der Wirkung der Mensch-Computer Interaktion (Human Computer Interaction, kurz HCI) untersucht werden. Denn die Auswirkungen von HCI sind immer auch stark von Kontextfaktoren weiterer Lebensbereiche der interagierenden Menschen beeinflusst und sollten in zukünftigen Forschungsarbeiten systematischer berücksichtigt werden.

Anwendungen von Chatbots in der Sozialen Arbeit

Bisher bestehen nach Kenntnis der Autoren erst in Ansätzen professionalisierte und wissenschaftlich beschriebene Formen des Einsatzes von KI-gestützten Chatbots in der Sozialen Arbeit. Verbreiteter erscheint der zumindest experimentelle Einsatz und die Erforschung von Chatbots in klinisch-psychologischen Settings und der Psychotherapie. So identifizieren Bendig et al. (2019) 148 Studien, die den Einsatz von Chatbots in klinisch-psychologischen Settings oder der Psychotherapie untersucht haben. Bei genauerer Betrachtung der eingesetzten Chatbots zeigt sich allerdings, dass diese in den meisten Fällen nicht auf KI-Technologien wie bspw. neuronalen Netzwerken beruhen, sondern die Redebeiträge der beteiligten Menschen nach Schlüsselwörtern

durchsuchen und auf spezifisch psychologischen Modellen beruhende Antworten bspw. in Form von Anschlussfragen formulieren. Dennoch attestieren die Autor_innen Chatbots in diesen Settings aufgrund von sechs näher untersuchten Studien Praktikabilität, Machbarkeit und Akzeptanz zur Förderung der psychischen Gesundheit. Die Teilnehmer_innen scheinen von den durch die Chatbots bereitgestellten Inhalten zu profitieren und bekunden eine Zunahme des Wohlbefindens sowie Abnahme von Stress und depressiven Verstimmungen.

Gabrielli et al. (2020) beschreiben den Einsatz und die Evaluation eines Chatbots bei 21 Jugendlichen (mittleres Alter 14,5 Jahre) zur Lebensberatung bei unterschiedlichen Themen wie emotionale Selbstwahrnehmung, soziale Wahrnehmung, interpersonelle Beziehungen, Konfliktbearbeitung, selbstsichere Kommunikation, Traurigkeit und Einsamkeit, Führungsverhalten und positive Emotionen. Auch in diesem Beispiel sind die Fragen und Antworten der Chatbots stark prädefiniert und entsprechen nicht den Kriterien der Künstlichen Intelligenz. Die Beurteilungen der Jugendlichen zu dem in der Studie eingesetzten Chatbot sind durchwegs positiv: 76 % fanden die Intervention hilfreich, 90 % einfach zu nutzen und innovativ (81 %). Im Einzelnen zeigt sich, dass die Jugendlichen den Chatbot insbesondere hilfreich für die Bearbeitung der Bereiche der Konfliktbearbeitung, der selbstsicheren Kommunikation und der interpersonellen Beziehungen empfanden. Zu ähnlich positiven Bewertungen kommen Beaudry et al. (2019) bezüglich eines Chatbots, der Jugendliche mit einer chronischen Erkrankung beraten hatte. Insbesondere zeigten sich die beteiligten Jugendlichen in hohem Masse engagiert, die Kommunikation mit dem Chatbot im untersuchten Zeitraum aufrechtzuerhalten. Gleichlaufend beschreiben So et al. (2020) ein verstärktes Engagement von Spielsucht betroffenen Personen durch den Einsatz eines Chatbots an einer internetgestützten Therapie teilzunehmen (vgl. zum hohen Engagement von Chatbotnutzenden auch Xiao et al. (2020) und Skjuve et al. (2021)).

Herausforderungen, Risiken und Potenziale von Chatbots in der Sozialen Arbeit

Der Einsatz von Chatbots in der Sozialen Arbeit verheißt vielversprechende Potenziale für Beratungssettings sowie die Fallanamnese. Es sind allerdings gegenwärtige Limitationen von Chatbots zu berücksichtigen, die spezifische Problematiken in Praxiskontexten evozieren können. Chatbots können Personen nicht zur Kommunikation animieren, die nicht an Hilfe interessiert sind und damit – falls keine alternativen Zielgruppenerreichungsformen verfügbar oder vorgesehen sind – tech-

nologischer Exklusion ausgesetzt sind (So et al. 2020). Chatbots sollten weiter insbesondere menschliche Kommunikation und Beziehung unterstützende, nicht ersetzende Funktion haben sowie begleitend zu spezifisch menschlichen Befähigungen eingesetzt werden: Aufnahme und Aufrechterhaltung von Beziehungen, Zeigen von Empathie, Zeit haben zu beraten und verfügbar für Unterstützung zu sein (Guillon et al. 2019). Damit ist eine wesentliche Problematik jeglichen Technologie Einsatzes in der Sozialen Arbeit angesprochen: Werden digitale Technologien dazu eingesetzt, den neo-liberal getriebenen Abbau sozialer Dienstleistungen vorzunehmen und kostengünstige, automatisierte Bearbeitungen sozialer Probleme herbeizuführen, drohen neue Formen digital mitverursachter sozialer Exklusion und Pseudolösungen sozialer Probleme (Eubanks 2018; Harlow et al. 2013). Schließlich sind insbesondere auch Datenschutzprobleme zu berücksichtigen: Eröffnen Menschen gegenüber den Chatbots Einblicke in ihr persönliches und intimes Leben, stellt sich die Frage, wie diese Daten übermittelt, gespeichert und ausgewertet werden (Dods-worth et al. 2013; Goldkind et al. 2018; Reamer 2013).

Chatbots kommen für die Soziale Arbeit in der Praxis vielfältige Potenziale zu. Studien zeigen, dass Menschen oftmals sehr gut auf Chatbots ansprechen, ein hohes Engagement in der Kommunikation zeigen und insbesondere die Unvoreingenommenheit der Interaktionspartner_innen schätzen (Skjuve et al. 2021). So beschreiben Nutzer_innen starke Bindungsgefühle gegenüber dem KI-gestützten Chatbot Replika¹ und berichten von stundenlang währenden Kommunikationen. Eine Interviewpartnerin beschreibt die Nutzung des Chatbots als hilfreich mit Ängsten in der Öffentlichkeit umzugehen (ebd.). Insbesondere auch Jugendliche sprechen gut auf die verbreitete Form des Textmessagings an. Weitere Potenziale besitzen Chatbots für die Identifikation von psychiatrischen Erkrankungen wie bspw. dem Autismus. Gerade in der stationären Jugendhilfe sind psychiatrische Erkrankungen weitverbreitet (Schmid 2008) und Chatbots könnten zur Unterstützung von Anamnesen wichtige Anhaltspunkte leisten.

Diskussion

Es bedarf weiterer Forschung dazu, wie Erbringungsverhältnisse in der Sozialen Arbeit durch KI-Technologien verändert werden und inwiefern die Implementation und Nutzung digitaler Technologien in der Sozialen Arbeit durch Risikodiskurse und technologische sowie wohlfahrtsstaatliche Entwicklungen beeinflusst wird. Während bereits Studien zu der Implementation von KI-basierten Entscheidungssystemen im Kindes- und Erwachsenenschutz bestehen, existiert bisher wenig Wis-

sen zu informelleren Nutzungsweisen von KI-Systemen, wie z. B. von Chatbots in Beratungssituationen.

Ein für die Soziale Arbeit bedeutsames Entwicklungsfeld ist die anwendungsorientierte Forschung zu KI. Hier ist die Soziale Arbeit insbesondere auch gefordert, interdisziplinäre Kollaborationen einzugehen, um die Potenziale von KI-Technologien freizulegen und systematisch zu nutzen. Anforderungen an die Entwicklung von KI-Systemen aus Perspektive der Sozialen Arbeit sind, wie Funktionen implementiert werden können, die zentrale Anliegen der Sozialen Arbeit ermöglichen. Hierin sollte die grundsätzliche Stoßrichtung der Entwicklung und Implementation von KI-Systemen in der Sozialen Arbeit liegen (vgl. Eckhardt et al. 2017).

Letztlich stellt sich die Verantwortungsfrage: Wie frei sind die Fachkräfte in der Sozialen Arbeit, eigene Entscheidungen zu treffen, wenn KI eingesetzt wird? Wer haftet bei Fehlern, vor allem bei menschlichen Entscheidungen, die von KI-Prädiktionen abweichen (Schneider und Seelmeyer 2018; Steiner 2021)? Im Rückgriff auf Jonas' (1984) ethische Theorie der Verantwortung sollte die wesentliche Frage lauten: „Welche Potenziale und Risiken ergeben sich *heute und in Zukunft* in Anwendung von KI-Technologien für *alle* an den Dienstleistungen Sozialer Arbeit beteiligten?“ Das vielschichtige Janusgesicht der Digitalisierung (Steiner 2015) erfordert entsprechend auch in der Sozialen Arbeit, empirisch gestützt und fachlich begründet Anwendungen, Nutzungsweisen und Folgen von KI-Technologien auf ihre jeweiligen Potenziale und Risiken zu befragen. ❀

Eingegangen. 21. März 2022

Angenommen. 29. Juli 2022

1. <https://replika.ai/>.

Funding. Open access funding provided by FHNW University of Applied Sciences and Arts Northwestern Switzerland.

Open Access. Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten

Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior, 85*, 183–189. <https://doi.org/10.1016/j.chb.2018.03.051>.

Beaudry, J., Consigli, A., Clark, C., & Robinson, K. J. (2019). Getting ready for adult Healthcare: designing a chatbot to coach adolescents with special health needs through the transitions of care. *Journal of Pediatric Nursing, 49*, 85–91. <https://doi.org/10.1016/j.pedn.2019.09.004>.

Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2019). The next generation: chatbots in clinical psychology and psychotherapy to foster mental health—A scoping review. *Verhaltenstherapie, https://doi.org/10.1159/000501812*.

Dewe, B. (2009). Reflexive Professionalität. In A. Riegler, S. Hojnik & K. Posch (Hrsg.), *Soziale Arbeit zwischen Profession und Wissenschaft: Vermittlungsmöglichkeiten in der Fachhochschulausbildung* (S. 47–63). VS.

Dodsworth, J., Bailey, S., Schofield, G., Cooper, N., Fleming, P., & Young, J. (2013). Internet technology: an empowering or alienating tool for communication between foster-carers and social workers? *The British Journal of Social Work, 43*(4), 775–795. <https://doi.org/10.1093/bjsw/bcs007>.

Eckhardt, J., Kalerka, C., & Pelka, B. (2017). Observations on the role of digital social innovation for inclusion. *Technology & Disability, 29*(4), 183–198. <https://doi.org/10.3233/TAD-170183>.

Eubanks, V. (2018). *Automating inequality: how high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Gabrielli, S., Rizzi, S., Carbone, S., & Donisi, V. (2020). A chatbot-based coaching intervention for adolescents to promote life skills: pilot study. *JMIR Human Factors, https://doi.org/10.2196/16762*.

Gillingham, P. (2016). Predictive risk modelling to prevent child maltreatment and other adverse outcomes for service users: inside the “black box” of machine learning. *The British Journal of Social Work, 46*(4), 1044–1058. <https://doi.org/10.1093/bjsw/bcv031>.

Goldkind, L., Thinyane, M., & Choi, M. (2018). Small data, big justice: the intersection of data science, social good, and social services. *Journal of Technology in Human Services, 36*(4), 175–178. <https://doi.org/10.1080/15228835.2018.1539369>.

Griol, D., Carbó, J., & Molina, J. M. (2013). An automatic dialog simulation technique to develop and evaluate interactive conversational agents. *Applied Artificial Intelligence, 27*(9), 759–780. <https://doi.org/10.1080/08839514.2013.835230>.

Guillon, Q., Baduel, S., Arnaud, M., & Rogé, B. (2019). Nouvelles technologies au service du dépistage: chatbot pour la détection précoce de l'autisme. *Enfance, 1*(1), 59–72.

de Haan, I., & Connolly, M. (2014). Another Pandora's box? Some pros and cons of predictive risk modeling. *Children and Youth Services Review, 47*, 86–91. <https://doi.org/10.1016/j.chldyouth.2014.07.016>.

Harlow, E., Berg, E., Barry, J., & Chandler, J. (2013). Neoliberalism, managerialism and the reconfiguring of social work in Sweden and the United Kingdom. *Organization, 20*(4), 534–550. <https://doi.org/10.1177/1350508412448222>.

Jonas, H. (1984). *The imperative of responsibility: in search of an ethics for the technological age*. University of Chicago Press.

Keymolen, E., & Broeders, D. (2013). Innocence lost: care and control in Dutch digital youth care. *The British Journal of Social Work, 43*(1), 41–63. <https://doi.org/10.1093/bjsw/bcr169>.

Kreutzer, R. T., & Sirrenberg, M. (2019). *Künstliche Intelligenz verstehen: Grundlagen – Use-Cases – unternehmenseigene KI-Journey*. Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-25561-9>.

Lenzen, M. (2018). *Künstliche Intelligenz: Was sie kann & was uns erwartet*. C.H. Beck. Originalausgabe

Lenzen, M. (2020). *Künstliche Intelligenz: Fakten, Chancen, Risiken*. C.H. Beck. Originalausgabe

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. *The British Journal of Social Work, 46*(5), 1208–1223. <https://doi.org/10.1093/bjsw/bcv028>.

Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies, 151*, 102630. <https://doi.org/10.1016/j.ijhcs.2021.102630>.

Reamer, F. G. (2013). Social work in a digital age: ethical and risk management challenges. *Social Work, 58*(2), 163–172. <https://doi.org/10.1093/sw/swt003>.

Rich, E. (1983). *Artificial intelligence*. McGraw-Hill.

Schmid, M. (2008). Children and adolescents in German youth welfare institutions. *Psychiatry in Europe, 1*, 10–12.

Schneider, D., & Seelmeyer, U. (2018). Der Einfluss der Algorithmen. *Sozial Extra, 42*(3), 21–24. <https://doi.org/10.1007/s12054-018-0046-y>.

Schneider, D., & Seelmeyer, U. (2019). Challenges in using big data to develop intelligent support systems for social work in Germany. *Journal of Technology in Human Services, 37*(2–3), 113–128. <https://doi.org/10.1080/15228835.2019.1614513>.

Schwartz, D. R., Kaufman, A. B., & Schwartz, I. M. (2004). Computational intelligence techniques for risk assessment and decision support. *Children and Youth Services Review, 26*(11), 1081–1095. <https://doi.org/10.1016/j.chldyouth.2004.08.007>.

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion—A study of human-chatbot relationships. *International Journal of Human-Computer Studies, 149*, 102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>.

So, R., Furukawa, T. A., Matsushita, S., Baba, T., Matsuzaki, T., Furuno, S., Okada, H., & Higuchi, S. (2020). Unguided chatbot-delivered cognitive behavioural intervention for problem gamblers through messaging app: a randomised controlled trial. *Journal of Gambling Studies, 36*(4), 1391–1407. <https://doi.org/10.1007/s10899-020-09935-4>.

Steiner, O. (2015). Widersprüche der Mediatisierung Sozialer Arbeit. In N. Kutscher, T. Ley & U. Seelmeyer (Hrsg.), *Mediatisierung (in) der Sozialen Arbeit* (S. 19–38). Schneider.

Steiner, O. (2021). Social work in the digital era: theoretical, ethical and practical considerations. *British Journal of Social Work, 51*(4), 3358–3374.

Thinyane, M., Goldkind, L., & Lam, H. I. (2018). Data collaboration and participation for sustainable development goals—A case for engaging community-based organizations. *Journal of Human Rights and Social Work, 3*(1), 44–51. <https://doi.org/10.1007/s41134-018-0047-6>.

Xiao, Z., Zhou, M. X., Liao, Q. V., Mark, G., Chi, C., Chen, W., & Yang, H. (2020). Tell me about yourself: using an AI-powered chatbot to conduct conversational surveys with open-ended questions. *ACM Transactions on Computer-Human Interaction, 27*(3), 1–15. <https://doi.org/10.1145/3381804>.