

Die Macht der Muster

Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch ‚künstliche Intelligenz‘¹

Björn Görder (Tübingen)

Zusammenfassung

Dieser Beitrag gibt einen Überblick über ethische Fragen, die sich beim Einsatz von KI-Algorithmen in der Sozialen Arbeit stellen. In Bezug auf konkrete Anwendungsszenarien akzentuiert er insbesondere Fragen der Transparenz, der Diskriminierung und der Selbstbestimmung. Darüber hinaus wird erläutert, dass KI-Technologie insgesamt eine gesellschaftsprägende Kraft entfaltet und sich daher für die Soziale Arbeit relevante ethische Fragen weit über den konkreten Technikeinsatz hinaus stellen. Diese betreffen beispielsweise die Gestaltung des Wirtschafts- und Sozialsystems, den Bereich des Politischen, die Grundstruktur sozialer Interaktion, das Selbstverständnis des Menschen und letztlich auch metaethische Grundlegungsfragen.

Schlüsselwörter selbstlernende KI-Algorithmen, Soziale Arbeit, ethische Kriterien, Diskriminierung, Transparenz, Prognose, Sozialethik

¹ Dieser Beitrag wurde wesentlich angeregt durch den Austausch in der Fachgruppe Ethik der DGSA und beim Arbeitskreis für theologische Wirtschafts- und Technikethik. Diese Impulse können nicht einzeln ausgewiesen werden – doch den Kolleginnen und Kollegen sei an dieser Stelle ausdrücklich gedankt.

Einführung

Man stelle sich einmal das folgende (fiktive, aber nicht völlig unrealistische) Szenario vor: Alexa (oder Siri) wird in einer Familie genutzt, um die Rollläden und das Licht zu steuern, eine Einkaufsliste zu führen, Musik zu spielen und im Internet zu recherchieren. Damit der Voicebot erkennt, wann er angesprochen wird, hört er die gesamte Kommunikation in einem Haushalt ab. Anhand von Merkmalen wie Tonfall, Lautstärke oder bestimmten Schlüsselbegriffen stellt er nun fest, dass eine laufende Konversation mit sehr hoher Wahrscheinlichkeit in häuslicher Gewalt enden wird. Wie sollte das Gerät nun reagieren: Sollte eine für die Beteiligten erkennbare Intervention erfolgen (z.B. ein Warnton, eine Ansprache o.ä.)? Sollte ein:e Sozialarbeiter:in verständigt werden – oder gar die Polizei?

Was wie eine Szene aus einem dystopischen (oder utopischen?) Roman klingt, soll hier zunächst die Funktionsweise sogenannter ‚künstlicher Intelligenz‘ illustrieren. Gemeint ist damit nicht eine ‚starke‘ KI, die menschliche Fähigkeiten und menschliches Bewusstsein vollständig imitieren würde. Vielmehr geht es um „Algorithmen für bestimmte, abgegrenzte Problemstellungen [...]“. Dabei ist Lernfähigkeit eine wesentliche Anforderung“ (Buxmann/Schmidt 2021b, 7). D.h. es geht um Algorithmen, die nicht starr durch die Programmierung vorgegeben sind, sondern die auf der Grundlage großer Datensätze z.B. mit den Methoden des maschinellen Lernens trainiert werden und so selbst Verfahren der Mustererkennung entwickeln (vgl. ebd., 9f.). Im skizzierten Beispiel hätte sich der häusliche Chatbot also nicht zu einer menschenähnlichen Existenz entwickelt, die eigenständig das Ziel definiert, potentielle Gewaltopfer zu schützen (‚starke KI‘). Es läge aber auch nicht ein ‚klassischer‘ Algorithmus vor, bei dem fachkundige Menschen Vorgaben machen würden, bei welcher Lautstärke, welchen Begriffen etc. eine Gefahrensituation vorliegt. Ein KI-Algorithmus würde derartige Muster vielmehr aus einer Vielzahl von Beispielfällen eigenständig identifizieren und auf dieser Grundlage Prognosen treffen. Solche Programme werden vielfach, etwa in der Sprach- und Bilderkennung, eingesetzt. KI-Algorithmen sind eng verbunden mit dem Phänomen Big Data. Einerseits sind die Verfügbarkeit und schnelle Verarbeitbarkeit von großen Datenmengen Voraussetzungen für die Leistungsfähigkeit von KI-Technologie. Andererseits ermöglicht diese erst die zielgerichtete Nutzung umfangreicher, unter Umständen unstrukturierter Datensätze.

Das Eingangsbeispiel legt bereits nahe, dass sich beim Einsatz dieser Technologie ethische Fragen stellen. Dabei handelt es sich um vergleichsweise spezielle Anwendungsfragen innerhalb des großen Feldes der Digitalisierung. Dieser Beitrag nimmt seinen Ausgangspunkt von solchen Fragestellungen und beschreibt daher zunächst Einsatzszenarien von KI-Algorithmen in der Sozialen Arbeit (1.), die ihnen zugrunde liegende Logik (2.) und Kriterien für ihren ethisch verantworteten Einsatz (3.). Das eigentliche Ziel der Darstellung ist jedoch ein anderes. Es soll ausgehend von diesen konkreten Anwendungsfragen verdeutlicht werden, dass durch den verbreiteten Einsatz von KI-

Technologie keineswegs nur Spezialfragen für die Ethik der Sozialen Arbeit aufgeworfen werden. Vielmehr stellen sich darüber hinaus weitreichende sozialetische Fragen (4.) sowie Fragen, die das Selbstverständnis der Ethik betreffen (5.). Keine dieser Fragen kann hier ausführlich erörtert werden – es soll vielmehr darum gehen, die Vielfalt der Themenstellungen und ihren inneren Zusammenhang darzustellen.

1. ‚Künstliche Intelligenz‘ in der Sozialen Arbeit

Im Zuge ihrer zunehmenden gesellschaftlichen Verbreitung ist davon auszugehen, dass KI-Algorithmen weitgehend unbemerkt und unreflektiert auch Einzug im Alltag der Sozialen Arbeit halten. Aller Wahrscheinlichkeit nach werden auch Sozialarbeiter:innen KI-basierte Navigationssysteme oder Spracherkennungs- bzw. Übersetzungsprogramme sowie Smartwatches nutzen. Es wäre wohl naiv anzunehmen, dass diese in dienstlichen Kontexten nicht mitgeführt oder sogar gezielt eingesetzt werden. Auch dass (wie im Eingangsbeispiel) Sprachassistenten bei ambulanten Betreuungen, aber wohl auch in Wohngruppen, sozialarbeiterisches Handeln ‚begleiten‘, sollte angenommen werden.

Davon zu unterscheiden ist der bewusste und fachlich reflektierte Einsatz von spezialisierten KI-Systemen in der Sozialen Arbeit. Dabei ist zum einen an den Einsatz KI-basierter Assistenzsysteme etwa in der Pflege (vgl. Bendel 2018), aber auch zur Verbesserung von Teilhabe (vgl. Stock 2020; Kähler et al. 2021) zu denken. Besonders zu beachten sind dabei Einsatzformen, die über Service- und unmittelbar pflegerische Funktionen hinaus auch soziale Interaktion bis hin zur Sexualassistenz erfüllen sollen (vgl. Döring 2018; Lenz et al. 2019).

Zum anderen können KI-basierte Programme in der Sozialen Arbeit dazu genutzt werden, komplexe und umfangreiche Datensätze auszuwerten und auf dieser Basis Korrelationen zu identifizieren, Prognosen zu erstellen oder Empfehlungen abzugeben (vgl. Schrödter et al. 2020). Am intensivsten wird in diesem Zusammenhang derzeit wohl der Einsatz algorithmenbasierter Programme zum Kinderschutz diskutiert (vgl. Schrödter/Bastian 2019). So wird beispielsweise das an der University of Auckland entwickelte Allegheny Family Screening Tool eingesetzt, um bei eingehenden Hinweisen auf eine mögliche Kindeswohlgefährdung das vorliegende Risiko einzuschätzen. Dabei greift ein Algorithmus zur Prognoseerstellung auf eine Vielzahl von Daten beispielsweise über die Familiensituation, den Gesundheitszustand oder mögliche Vorstrafen zu (vgl. Vaithianathan et al. 2017). Ähnliche Einsatzszenarien ergeben sich z.B. bei der Suizidprävention (vgl. Kamenetz 2021), bei der Unterstützung Arbeitsloser oder der Kontrolle eines Missbrauchs von Sozialleistungen (vgl. Gapski 2020, 162). Unabhängig von der Prognoseerstellung wäre der Einsatz von KI-Algorithmen zur Strukturierung großer Datenmengen auch denkbar, um etwa das Risiko einer Überschuldung oder Ansprüche auf den Bezug von Sozialleistungen zu prüfen. Wiederum andere Einsatzmöglichkeiten ‚künstlicher Intelligenz‘ in der Sozialen Arbeit stellen

Chatbots dar. Diese können zur bloßen Informationsvermittlung, aber auch in therapeutischen Kontexten eingesetzt werden (vgl. Fiske et al. 2019).

Vergegenwärtigt man sich die vielfältigen Einsatzformen oder -szenarien ‚künstlicher Intelligenz‘ in der sozialen Arbeit, so verbietet sich eine pauschale (negative oder positive) ethische Beurteilung quasi von selbst. Anstelle einer je einzeln zu leistenden ethischen Diskussion von Anwendungsfragen sollen im Folgenden grundsätzliche Fragen angesprochen werden, die sich für die Ethik der Sozialen Arbeit angesichts des gesellschaftlichen und professionellen Einsatzes ‚künstlicher Intelligenz‘ stellen. Dazu erfolgt zunächst eine Orientierung in Bezug auf die Funktionsweise ‚lernender‘ Algorithmen.

2. Mustererkennung durch KI-Algorithmen und in der Sozialen Arbeit

2.1 Anknüpfungspunkte zwischen ‚Künstlicher Intelligenz‘ und Sozialer Arbeit

In Theorie und Praxis der Sozialen Arbeit ist häufig eine gewisse Skepsis gegenüber dem Einsatz von Technik anzutreffen – schließlich habe man es ja ‚mit Menschen‘ zu tun. Diese Abwehrhaltung greift jedoch gleich zweifach zu kurz: Zum einen prägt Technik – und gerade Digitaltechnik – das menschliche Leben in hohem Maße – und Menschen die Technik natürlich ebenso. ‚Mensch‘ und ‚Technik‘ gehören also nicht zwei verschiedenen Welten an. Zum anderen gibt es in den Arbeitsweisen von KI-Technologie und Sozialer Arbeit mehr Berührungspunkte, als es auf den ersten Blick scheinen mag.

Armin Nassehi weist darauf hin, dass die Anfänge der Digitaltechnologie weit vor der eigentlichen Erfindung dieser Technik liegen, nämlich in der sozialwissenschaftlichen Erforschung von (unbewussten) Mustern gesellschaftlicher Strukturen und menschlichen Verhaltens.

„Die empirische Sozialforschung ist selbst Teil der Digitalisierungsgeschichte der modernen Gesellschaft. Wenn es stimmt, dass ‚Gesellschaft‘ als Chiffre für Regelmäßigkeiten, vor allem für dem direkten Blick nicht unmittelbar ansichtige Regelmäßigkeiten und Muster steht, dann ist die Entstehung der empirischen Sozialforschung selbst eine Reaktion auf diese Erfahrung der musterhaften Strukturiertheit der Gesellschaft.“ (Nassehi 2019, 54)

KI-Algorithmen funktionieren also deswegen, weil sie in einer hochkomplexen Gesellschaft Regelmäßigkeiten aufdecken, die dem subjektiven Bewusstsein oft verborgen sind. Das ist durchaus anschlussfähig für Arbeitsweisen der Sozialen Arbeit als Theorie und Profession. Auch diese fragt schließlich nach Mustern, die menschliches Verhalten in kollektiv-sozialer und individuell-psychologischer Hinsicht verstehbar und zu einem gewissen Grad vorhersehbar und steuerbar machen. Sowohl die Diagnoseklassifikation als auch die Behandlungsklassifikation stellen im Wesentlichen solche wissensbasierte Typologisierungen dar (vgl. Bastian 2014, 151f.; Schneider 2021, 128). Auch

wenn der Kern professioneller Tätigkeit in nicht standardisierbaren Akten und der Interpretation von Einzelfällen liegt, so greifen Professionelle dabei stets (bewusst oder unbewusst) auf erfahrungsbasierte Vorstellungen von letztlich statistisch identifizierten Mustern zurück (vgl. Bastian 2014, 157f.).

Insofern besteht hier zumindest eine Ähnlichkeit zum eingangs beschriebenen Algorithmus: Es geht um Kenntnisse darüber, welche Faktoren allgemein und bei einer spezifischen Person z.B. das Risiko gewalttätigen Verhaltens erhöhen und was kurz-, mittel- und langfristig getan werden kann, um dieses Risiko zu reduzieren. Wenn es um das Erkennen solcher Muster geht, haben selbstlernende Algorithmen den großen Vorteil, dass sie Muster in einer unglaublichen Menge von Daten identifizieren können.

2.2 Die Leistungsfähigkeit von KI-Algorithmen und ihre Grenzen

Natürlich ist stets kritisch zu prüfen, wie leistungsfähig ein KI-Algorithmus tatsächlich ist. Letztlich fördert er statistische Zusammenhänge zutage – kann aber bei deren Bewertung und insbesondere im Blick auf Einzelsituationen auch entscheidende Faktoren übersehen. So sind zumindest gegenwärtig gerade bei komplexen Aufgaben häufig noch geradezu grotesk anmutende Fehlleistungen festzustellen (vgl. Lenzen 2020, 55ff.). Es ist auch kaum abzusehen, in welchen Bereichen tatsächlich mit weiteren Leistungssprüngen gerechnet werden kann, durch die neue Funktionsweisen und Einsatzszenarien erschlossen werden könnten. Insbesondere in unübersichtlichen Situationen und für komplexe Zielsetzungen ist bei allzu optimistischen Erwartungen Zurückhaltung geboten. Es bleibt letztlich das Problem, dass KI-Algorithmen zwar viele Daten verarbeiten können, ihnen aber das für die Einordnung relevante Weltwissen fehlt.

Wenn aber für klar definierte Aufgaben das Urteil von Fachkräften durch eine ergänzende Nutzung statistischer bzw. algorithmenunterstützter Entscheidungsverfahren verbessert werden kann (vgl. Bastian 2014, 149f.), stellt sich letztlich die Frage, ob es aus ethischer Sicht vertretbar wäre, diese grundsätzlich *nicht* einzusetzen. Angenommen, es könnte tatsächlich eine Verbesserung des Kinderschutzes oder eine Eindämmung häuslicher Gewalt erreicht werden, so bräuchte es sehr starke ethische Gründe, um dennoch gegen eine Nutzung der entsprechenden Technologie zu votieren. Anstelle einer im Prinzipiellen verharrenden Debatte wäre demnach eine pragmatische Herangehensweise an die Frage wünschenswert, in welchen Bereichen der Sozialen Arbeit die Qualität der Arbeit durch den Einsatz ‚intelligenter‘ Algorithmen verbessert werden kann. Dies setzt die Kompetenz voraus, dass die Beteiligten deren Wirkweise in Grundzügen nachvollziehen und ihre Aussagekraft realistisch einschätzen können (vgl. Schneider 2021, 132f.). Zur Sicherung fachlicher Standards und orientiert am Wohlergehen der betroffenen Personengruppen scheint dabei zentral, dass sozialarbeiterisch-fachliche Expertise bereits in der Entwicklung professionsbezogener Programme umfangreich zur Geltung kommt.

Problematischer als mögliche Schwächen von KI-Technologie scheint paradoxerweise ihre enorme Leistungsfähigkeit: *Erstens* kann sie zu einem objektivistischen Missverständnis führen, wonach die von Algorithmen erkannten Muster die Wirklichkeit abbilden. Demgegenüber ist deutlich zu machen, dass Algorithmen immer nur auf Grundlage eines bestehenden Datensatzes operieren können. Was nicht erfasst ist bzw. ihnen nicht zur Verfügung steht, kann also schon gar keine Berücksichtigung finden. Die vorliegenden Daten werden dann priorisiert, klassifiziert, zugeordnet und gefiltert (vgl. Busch 2018, 19ff.). Durch die Zuschreibung von Bedeutung finden also interpretativ-konstruktive Akte statt, die zugleich mit normativen Wertungen einhergehen. Bei menschlichen Akteuren ist dies deutlicher im Bewusstsein. Man sollte sich aber bewusst sein, dass auch KI-Algorithmen die Wirklichkeit nicht einfach nur widerspiegeln. Sie präsentieren keinesfalls eine – weil technisch rekonstruierte – ‚objektive‘ Sicht der Wirklichkeit, die ‚subjektiven‘ menschlichen Wahrnehmungen kategorial überlegen und insofern der Kritik entzogen wäre. Die Leistungsfähigkeit von KI-Systemen bei der Interpretation komplexer, mehrdeutiger Systeme muss sich daher erst erweisen und fachlich kritisch untersucht werden.²

Dass Algorithmen die Wirklichkeit nicht nur abbilden, lässt sich *zweitens* daran festmachen, dass ihr Einsatz nicht nur zum Zweck bloßer Erkenntnis erfolgt (vgl. Nassehi 2019, 148). Hat man Handlungsmuster identifiziert, so kann man dies dazu nutzen, Verhalten gezielt zu steuern. Dadurch können bestehende Muster verstärkt, aber auch durchbrochen werden. Die Leistungsfähigkeit von KI-Algorithmen führt also dazu, dass sie individuelles und kollektives Verhalten gezielt beeinflussen können. Damit stellen sich zahlreiche ethische Fragen danach, ob überhaupt, zu welchem Zweck und unter welchen Umständen eine solche Einflussnahme legitim ist und ob ihre Ziele als wünschenswert erscheinen. Dem soll nun nachgegangen werden.

3. Ethische Kriterien zur Nutzung von KI-Algorithmen in der Sozialen Arbeit

In den letzten Jahren wurden in Politik, Wirtschaft und Wissenschaft zahlreiche Kriterienkataloge für ethisch vertretbare KI entwickelt (vgl. Hagendorff 2019). Auch wenn es im Detail und in der Konkretion teilweise erhebliche Unterschiede gibt, zeichnen sich dabei einige Kriterien ab, die allgemein anerkannt werden. Dazu gehören die sieben Leitlinien, die eine von der EU-Kommission eingesetzte Expert:innengruppe erarbeitet hat. Dies sind: (1) menschliche Entscheidungsfreiheit und Kontrolle, (2) technische Robustheit und Sicherheit, (3) Datenschutz, (4) Transparenz, (5) Diversität, Nichtdiskriminierung sowie Fairness, (6) gesellschaftliches und ökologisches Wohlergehen und (7) Zurechenbarkeit (vgl. Independent High-Level Expert Group on Artificial Intelligence 2020, 3). Auch in Bezug auf die Soziale Arbeit werden Kriterien für den

² Vgl. für ein modellhaftes interdisziplinäres Projekt zum Einsatz von KI-Technologie für die Einschätzung von Kindeswohlgefährdung <https://www.kaimo.bayern/>.

ethisch vertretbaren Einsatz von KI diskutiert (vgl. Fritz et al. 2019, 25ff.; Görder 2020). Hier sollen daher nur einige zentrale Aspekte dieser Fragestellung herausgehoben werden.

3.1 Das Problem ungerechtfertigter Diskriminierung

Unter dem Gesichtspunkt der Gerechtigkeit ist hier zumeist das Problem diskriminierender Algorithmen im Blick. Diese identifizieren in vorgegebenen Datensätzen ‚Muster‘ – welche Personengruppen etwa für bestimmte Stellen erfahrungsgemäß ausgewählt, häufiger von der Polizei kontrolliert oder bei Prognosen über eine zukünftige Straftat schlechter eingeschätzt werden. Werden diese Muster als Grundlage für weitere Entscheidungen verwendet (also z.B. die Auswahl von Bewerber:innen oder zielgerichtete Polizeikontrollen), ist dies nicht in den Eigenschaften und Verhaltensweisen der Menschen selbst begründet, sondern in gesellschaftlichen Vorurteilen. Diese werden durch KI-Algorithmen reproduziert und verstärkt, sodass es zu ungerechtfertigter Diskriminierung (z.B. von Frauen oder Menschen aus bestimmten ethnischen Gruppen) kommt (vgl. z.B. O'Neil 2016, 96ff., 112ff.; Busch 2018, 20ff.; Martini 2019, 47ff.; Beranek 2021a, 7f.). Dies ist in der Tat ethisch hoch problematisch und der Einsatz von Algorithmen ist nur vertretbar, wenn eine solche ungerechtfertigte Diskriminierung weitgehend ausgeschlossen werden kann.

3.2 Die Frage gerechtfertigter Diskriminierungen

Weniger im Blick scheint demgegenüber die Frage, in welchen Fällen Ungleichbehandlungen ethisch legitim sind. Der Einsatz ‚künstlicher Intelligenz‘ zielt letztlich nahezu immer auf Diskriminierung. Personen sollen je nach ihren bisherigen Verhaltensweisen und Eigenschaften unterschiedlich behandelt werden. Folgt man dem klassischen aristotelischen Gerechtigkeitsgrundsatz, wonach „Gleiches gleich und Ungleiches ungleich“ zu behandeln ist, kann dies aus Gründen der Gerechtigkeit sogar gefordert sein. Zum Potenzial eines KI-Einsatzes in der Sozialen Arbeit gehört ja gerade die Möglichkeit, dass Kombinationen von Risikofaktoren identifiziert werden können, die (womöglich gerade aufgrund von *menschlichen* Vorurteilen) bislang übersehen wurden. Im positiven Falle könnte eine effizientere Suizidprävention gelingen oder Menschen könnten in einer für sie verständlichen Sprache mit für sie wesentlichen Informationen adressiert werden. Auch dass Menschen mit erhöhtem Assistenzbedarf einen leichteren Zugang zu Förderprogrammen haben, dürfte im Rahmen von Bedürfnis- und Chancengerechtigkeit weitgehend unbestritten sein.

Unter Gerechtigkeitsgesichtspunkten ist also zu klären, welche Unterschiede in Bezug auf Eigenschafts- oder Verhaltensmerkmale welche Formen der Ungleichbehandlung legitimieren oder sogar verlangen. Macht es einen Unterschied, ob Menschen aufgrund von Alter, Geschlecht, Herkunft, religiöser Praxis, Konsumgewohnheiten, Be-

rufszugehörigkeit, Musikvorlieben usw. unterschiedlich behandelt werden? Ist es vertretbar, bei der Auswahl von Werbeanzeigen, bei der Berechnung von Versicherungstarifen, bei Preisen für Flugreisen oder bei der Auswahl für Bewerbungsgespräche gemäß der genannten Merkmale zu differenzieren?

So zeigt sich im Umgang mit KI-Technologie die Vielschichtigkeit von Gerechtigkeitsdiskursen und davon berührter Fragen: Wann ist Ungleichbehandlung wünschenswert, weil sie passgenau individualisierte Angebote und einen effizienten Einsatz von Ressourcen ermöglicht? Wo steht die Gleichheit aller Menschen im Vordergrund, so dass eine Gleichbehandlung (ungeachtet aller erkennbaren Muster oder Formen von Singularität) gefordert werden sollte? Und in welchen Bereichen erscheint es wünschenswert, Unterschiede zu nivellieren, sodass ausgleichende Maßnahmen zu fordern wären. Sollte also beispielsweise Werbung für einen naturwissenschaftlichen Studiengang primär Männern, allen Menschen gleichermaßen oder primär Frauen angezeigt werden? Solche Fragen können nicht allein auf Basis statistischer Zusammenhänge beantwortet werden, sondern berühren immer auch Fragen der Anthropologie und des guten (Zusammen-)Lebens.

Entsprechende Fragen stellen sich auch für einen möglichen Einsatz von KI-Technologie in der Sozialen Arbeit: Wie sollte man damit umgehen, wenn Muster identifiziert werden können, wonach bestimmte Kombinationen aus unterschiedlichen Eigenschafts- und Verhaltensmerkmalen ein statistisch erhöhtes Risiko für Sucht, Überschuldung oder Kindeswohlgefährdung induzieren? Welche typischen Merkmale dürfte ein Voicebot bei der Einstufung einer Situation als potenziell bedrohlich berücksichtigen? Sollten also Ungleichbehandlungen aufgrund von Merkmalen wie Alter, Geschlecht, Herkunft etc. vorgenommen werden? Oder sollte im Sinne des Gleichbehandlungsgrundsatzes die Berücksichtigung mancher Faktoren ausgeschlossen werden? Welche Faktoren dürften für eine Ungleichbehandlung legitimer Weise zugrunde gelegt werden und welche nicht? Während die Ablehnung einer Diskriminierung aufgrund statistischer Verzerrung auf der Hand liegt, scheint ein Umgang mit statistisch induzierten Ungleichheiten die ethisch herausfordernderen Fragen aufzuwerfen.

3.3 Muster, Prognosen und Selbstbestimmung

Angesichts der hohen Wirksamkeit algorithmischer Systeme stellt sich das Problem der Manipulierbarkeit des Menschen. Wo durch das lenkende Gestalten von Entscheidungssituationen und Informationsaufbereitung („Nudging“) Verhalten gezielt beeinflusst oder durch algorithmisierte Entscheidungen Lebenschancen verteilt werden, kommt es zu Einschränkungen von Selbstbestimmungsmöglichkeiten. Es ist also jeweils zu prüfen, wo eine solche Beeinflussung in so einem Maße gegen Rechte und Interessen der Betroffenen verstößt, dass ein Einsatz entsprechender KI-Technologie moralisch nicht zu vertreten ist bzw. auch rechtlich reguliert werden sollte (vgl. Busch 2018, 56ff.; Müller 2021).

Auch in diesem Zusammenhang sollte man sich bewusst machen, dass eine solche Beeinflussung in der Funktionslogik von KI-Algorithmen angelegt ist. Diese basiert ja gerade darauf, dass Prognosen, Empfehlungen und Entscheidungen auf Grundlage von Daten über bisherige Verhaltensmuster getroffen werden. Insofern wird die Vergangenheit für die Zukunft gemäß der größten Wahrscheinlichkeit fortgeschrieben. Demgegenüber entspricht es den Zielen und Arbeitsweisen der Sozialen Arbeit, dass Menschen Möglichkeitsräume entdecken können und gerade nicht auf das Bisherige (seien es Routinen, Scheitern, Beziehungsdynamiken...) festgelegt sind. Es ist also insbesondere darauf zu achten, dass durch das Design und den Umgang mit KI-Algorithmen keine Festlegung auf Widerfahrnisse und Verhaltensmuster der Vergangenheit erfolgt und die Zukunftsoffenheit des menschlichen Lebens nicht unterminiert wird.

3.4 Das Gebot der Transparenz und das Black-Box-Problem

Angesichts der Funktionsweise ‚künstlicher Intelligenz‘ verdient außerdem die Frage der Transparenz besondere Aufmerksamkeit (vgl. Busch 2018, 56f.). Ethische Probleme ergeben sich angesichts des Black-Box-Phänomens. In der Regel kann gerade nicht identifiziert werden, aufgrund welcher Faktoren ein KI-Algorithmus zu einer Empfehlung oder Entscheidung kommt (vgl. Martini 2019., 28ff.). Gerade in sensiblen Lebensbereichen (z.B. der Familienhilfe) ist es für die Beteiligten jedoch von hoher Relevanz, dass sie zumindest die Gründe verstehen, aus denen andere u.U. in ihr Leben eingreifen oder gewünschte Hilfen verweigern. Auch aus diesem Grund erscheint der Einsatz eines Black-Box-Algorithmus jedenfalls in sensiblen Lebensbereichen ethisch kaum vertretbar. Hier stellt sich also die Frage, inwiefern die Explizierbarkeit technisch und rechtlich sichergestellt werden kann (vgl. Busch 2018, 61; Martini 2019, 41ff.; Buxmann/Schmidt 2021a, 219). Ist dies der Fall, kann ein gut programmierter KI-Algorithmus den Betroffenen und beteiligten Fachkräften sogar einen besseren Einblick in relevante Faktoren ermöglichen und so die Entscheidungsqualität auch in ethischer Hinsicht verbessern.

3.5 Offensichtliche und versteckte ethische Herausforderungen

Zuletzt ist auf die Gefahr hinzuweisen, einerseits sehr strenge Kriterien an explizit für soziale Handlungsfelder entwickelte KI-Programme anzulegen, andererseits den Einsatz von alltäglich gewordener Digitaltechnologie zu bagatellisieren. So stellen sich dringliche Fragen der Privatsphäre und des Datenschutzes nicht erst, wenn sich Alexa oder Siri (wie im Eingangsbeispiel) in sozialarbeiterische Tätigkeitsfelder ‚einmischen‘ würden, sondern bereits dann, wenn ihre Präsenz in den entsprechenden Handlungskontexten akzeptiert wird. Dasselbe gilt, wenn kommerzielle Apps (Übersetzungsprogramme, Messenger, Navigationssysteme etc.) im professionellen Kontext genutzt

werden. Bei den meisten dieser alltäglich genutzten Programme werden Daten erhoben und zur Beeinflussung genutzt – in der Regel orientiert am ökonomischen Interesse der Anbieter. So ist es denkbar, dass eher unproblematische oder sogar förderliche Potenziale im Bereich des professionellen Einsatzes nicht genutzt werden, während gleichzeitig ethisch bedenkliche Angebote weitgehend unreflektiert Verwendung finden.

Damit aber zeichnet sich zugleich ab, dass es eine Überforderung für alle Beteiligten darstellen würde, die Frage nach dem ethisch vertretbaren Einsatz ‚künstlicher Intelligenz‘ allein auf Ebene der Individual- und Organisationsethik zu verhandeln. Der weitverbreitete Einsatz von KI-Technologie transformiert alle Bereiche des gesellschaftlichen Zusammenlebens und sollte daher auch sozialetisch reflektiert werden.

4. Der Einsatz von KI-Algorithmen aus sozialetischer Perspektive

Die durch den Einsatz von KI ermöglichten oder beschleunigten gesellschaftlichen Transformationsprozesse sind für die Soziale Arbeit in unterschiedlicher Weise relevant. Zum einen führen sie zu Umstrukturierungen, die immer auch mit der Umverteilung von Macht zusammenhängen. Dies soll an den Bereichen der Wirtschaft und der Politik gezeigt werden. Die Soziale Arbeit ist hier insbesondere mit Blick auf ihr politisches Mandat gefordert. Zum anderen verändern sich allgemein soziale Interaktionsweisen, sodass sich grundlegende Fragen nach dem sozialen Zusammenhalt und den Arbeitsbedingungen sozialer Berufe ergeben.

4.1 Mustererkennung und ökonomische Macht

Der Einsatz von KI-Algorithmen ist im Bereich der Ökonomie am weitesten fortgeschritten. Am prominentesten ist hier sicher der Bereich des Marketings und der dynamischen Preisgestaltung im Onlinehandel. Doch auch bei der Berechnung von Versicherungstarifen, Kreditvergaben oder auch der Personalauswahl werden entsprechende Technologien eingesetzt (vgl. Busch 2018, 12ff.). Entscheidungen von großer Tragweite werden so anhand von (in der Regel auf Gewinnmaximierung hin programmierten) Algorithmen getroffen. Die Orientierung an bestehenden Mustern verfestigt dabei bestehende Ungleichheiten und wirft Fragen der Gerechtigkeit auf, die (wie oben dargestellt) über die der verzerrenden Diskriminierung hinausgehen. Auch auf dem Arbeitsmarkt ist zu erwarten, dass Arbeit, Beschäftigung, Einkommen und Status zunehmend ungleich verteilt sind und davon die bereits jetzt Begünstigten profitieren (vgl. Albrecht/Kellermann 2020). Durch ein Zusammenspiel vieler Einzelfaktoren kommt es so zu einer verstärkten Exklusion, Kriminalisierung und zur Reduzierung von Wahlfreiheiten für besonders vulnerable Gruppen (vgl. Kutscher 2020, 353ff., bes. 355). Gleichzeitig können sie besonders anfällig sein für die ökonomische Ausnutzung von Suchtpotenzial, wie es etwa bei der Gestaltung von Online-Games geschehen kann (vgl. Beranek 2021b, 160ff.).

Gesamtwirtschaftlich ist zu berücksichtigen, dass KI-Algorithmen besonders von denjenigen effizient genutzt werden können, die über große Datenmengen und Ressourcen für ihre technische Verwertung verfügen. Netzwerkeffekte und geringe Grenzkosten für zusätzliche Produkte führen in der Plattformökonomie und allgemein in der Digitalwirtschaft zu starken Monopolisierungstendenzen (vgl. Bauberger 2020, 37ff.). Sofern keine Regulierung erfolgt, wäre also eine Zunahme der Machtasymmetrie zugunsten großer Internetkonzerne sowohl gegenüber potenziellen Wettbewerbern als auch gegenüber Mitarbeitenden und Kund:innen zu erwarten.

Bei diesen Entwicklungen geht es nicht um einzelne Bereiche des Wirtschaftslebens, sondern um eine Transformation des Wirtschaftens insgesamt, die letztlich alle Lebensbereiche umfasst. Nicht nur wirtschaftliche Aktivitäten, sondern auch private Kommunikation und Verhaltensformen werden im Rahmen omnipräsenter KI-Technologie analysiert und durch diese strukturiert. Maßgeblich sind dabei die Interessen der diesen Markt dominierenden großen IT-Konzerne. Shoshana Zuboff beschreibt eindrücklich den Grundmechanismus dieses „Überwachungskapitalismus“: „Er erklärt Erfahrungen von Privatmenschen zum kostenlosen Rohstoff für Produktion und Verkauf“ (Zuboff 2019, 5).

Somit stellen sich angesichts technischer Innovationen auch sozialetische Fragen z.B. im Bereich der Wettbewerbspolitik (z.B. Monopolisierung, Regulierung des Einsatzes von KI-Technologie) und der Sozialpolitik (z.B. Finanzierung der Sozialsysteme durch Steuern oder Lohnnebenkosten, bedingungsloses Grundeinkommen). Zugleich wird deutlich, dass auch Fragen des Datenschutzes eine soziale Frage erheblicher Tragweite darstellen. Als Menschenrechtsprofession (vgl. Staub-Bernasconi 2007, 24ff.) sollte die Soziale Arbeit auch im Blick haben, dass durch den Einsatz von KI-Algorithmen Fragen nach der Geltung von Grundrechten wie Selbstbestimmungs- und Persönlichkeitsrechten oder der Meinungsfreiheit aufgeworfen werden, die sowohl gegenüber dem Staat als auch gegenüber Wirtschaftsakteuren zu schützen sind (vgl. Di Fabio 2016).

4.2 KI-Algorithmen als Mittel politischer Macht

Als wohl drastischstes Beispiel der Nutzung von KI-Algorithmen im politischen Raum kann das chinesische ‚Social Scoring‘-System dienen, mit dem die Bevölkerung zu einem im Sinne des Staates gewünschten Verhalten gelenkt werden soll (vgl. Buxmann/Schmidt 2021a, 221ff.). Doch spätestens mit der Wahl Donald Trumps und den Enthüllungen über gekaufte Datenbanken bei Cambridge Analytica wurde deutlich, dass die Mechanismen mustererkennender Algorithmen auch in demokratischen Auseinandersetzungen eingesetzt werden. Durch sogenanntes Microtargeting werden gezielt Personengruppen mit Informationen angesprochen, die ihr Verhalten in eine gewünschte Richtung lenken sollen (vgl. Christl 2019). Auch im Vorfeld der Bundes-

tagswahl 2021 konnte gezeigt werden, dass alle Parteien mithilfe des Facebook-Algorithmus gezielt geworben und zum Teil unterschiedliche Personengruppen mit unterschiedlichen Botschaften adressiert haben (vgl. ZDF Magazin Royal 2021).

Bedenklich ist solches Microtargeting zum einen schon deswegen, weil es politischen Akteuren, die bereit und in der Lage sind, entsprechende Ressourcen einzusetzen, erhebliche Möglichkeiten der Einflussnahme eröffnet. Dabei werden auch die großen Internetunternehmen zu zentralen politischen Akteuren. Fehlende demokratische Legitimation und Transparenz sind in der Sphäre des Politischen besonders kritisch zu beleuchten.

Plurale demokratische Gesellschaften basieren auf dem Ideal eines öffentlichen Diskurses, in dem unterschiedliche Interessen und Perspektiven im Modus des rationalen Argumentierens nach Verständigung und Ausgleich suchen (vgl. Habermas 2005, 124ff.). Demgegenüber „könnte datenbasiertes Microtargeting gesellschaftliche Debatten fragmentieren sowie das Vertrauen in Politik und demokratische Institutionen weiter unterminieren“ (Christl 2019, 48; vgl. O'Neil 2016, 179ff.). Gerade aus Sicht einer Ethik der Sozialen Arbeit ist bei einem solchen neuerlichen Strukturwandel der Öffentlichkeit kritisch darauf zu achten und entgegenzuwirken, wenn in der politischen Kommunikation Partikularinteressen und Vorurteile befördert und schwer zu mobilisierende Menschen tendenziell aus dem politischen Diskurs exkludiert werden.

4.3 Singularisierung und Standardisierung – KI-Algorithmen als Katalysator kultureller Veränderungen

Die zunehmende Fragmentarisierung politischer Diskurse ist eingebettet in einen umfassenden gesellschaftlichen Wandel. Andreas Reckwitz beschreibt, dass in der Spätmoderne die soziale Logik des Besonderen die Logik des Allgemeinen ablöst. Ökonomische, kulturelle und technologische Aspekte tragen zu einer zunehmenden Singularisierung bei, welche die Einzigartigkeit von Menschen, aber auch Objekten und Organisationen, in den Mittelpunkt rückt. KI-Algorithmen übernehmen in diesem Prozess eine wichtige Funktion: Anders als die Schlüsseltechnologien der Moderne erreichen sie nicht Effizienz und Rationalität durch Massenproduktion und Standardisierung, sondern erlauben gerade die Orientierung an je individuellen Profilen (vgl. Reckwitz 2018, 11ff., 228f., 243f.). KI-Algorithmen führen zu personalisierter Nachrichtenerstattung (vgl. Schicha 2021, 117) und strukturieren über Social Media und Dating-Plattformen soziale Kontakte. Auch hier führt die Logik der Mustererkennung dazu, dass wir verstärkt mit solchen Botschaften und Kontakten konfrontiert werden, die bisherigen Gewohnheiten entsprechen. Getrieben durch ökonomische Interessen der Plattformanbieter werden dabei Botschaften bevorzugt, die stark emotionalisieren und Aufmerksamkeit erregen. So kommt es zu Radikalisierungstendenzen und einer Zunahme von ‚Hate Speech‘, da provokantere Inhalte tendenziell häufiger abgerufen werden (vgl. Beranek 2021b, 164ff.).

In gewisser Weise gegenläufig zu dieser Tendenz wirken Algorithmen, sofern sie allgemeine Muster zur Norm machen und so zu einer Standardisierung beitragen. Dies geschieht etwa, wenn durch Filter ästhetische Normvorstellungen automatisch zur ‚Optimierung‘ von Fotos herangezogen werden (vgl. Beranek 2021a, 9f.) oder Menschen sich in Bezug auf eine gesunde Lebensführung an Fitness-Trackern orientieren. Hier werden versteckte Normen oder auch schlicht statistische Mittelwerte verdeckt zu normativen Orientierungsgrößen und prägen unser Selbstbild.

Die durch KI-Technologie begünstigte Singularisierung kann einen Zugewinn an Lebensqualität mit sich bringen und birgt auch emanzipatorisches Potenzial. Zugleich kann die Kombination aus Personalisierung, Fragmentarisierung und Standardisierung dazu führen, dass Menschen zunehmend weniger mit Personen und Inhalten konfrontiert sind, die ihnen fremd sind oder ihren Ansichten und Vorlieben nicht entsprechen. Gerade aus Perspektive einer Ethik der Sozialen Arbeit wird deutlich, dass Werte wie Toleranz, Solidarität und Konfliktfähigkeit für ein gelingendes Zusammenleben eine hohe Bedeutung haben. Zu fragen ist also, wie solche Werte in einer Gesellschaft verankert werden können, die durch „partikulare[] Kommunikationsgemeinschaften“ bzw. eine „Vielzahl von Parallelöffentlichkeiten“ (Reckwitz 2018, 261, 269) geprägt ist.

5. ‚Künstliche Intelligenz‘ und das Selbstverständnis der Ethik der Sozialen Arbeit

Angesichts ihrer umfassenden gesellschaftlichen Wirkungen, stellt sich die Frage, wie KI-Algorithmen ethisch verantwortlich gestaltet und genutzt werden können. Durch die Diskussion über ‚gute‘ KI-Algorithmen werden zugleich Fragen aufgeworfen, die Anlass geben, Grundprinzipien der Ethik allgemein und der Ethik der Sozialen Arbeit im Speziellen erneut zu bedenken oder ins Bewusstsein zu rufen.

5.1 Entscheidungsalgorithmen und Fragen nach der Verantwortlichkeit und Subjektivität

Vereinzelt werden KI-Algorithmen angesichts ihrer Leistungsfähigkeit und durch sie getroffener Entscheidungen als moralische Subjekte oder sogar Träger von Rechten anerkannt. In der ethischen Diskussion wird jedoch mehrheitlich die Position vertreten, dass Verantwortung nur personalen Subjekten zugeschrieben werden kann, also keinen Algorithmen (vgl. Lenzen 2020, 62ff.; Müller 2021). Offen bleibt dann aber, inwiefern und wem dann eine (Teil-)Verantwortung für die Wirkungen von KI-Algorithmen zugeschrieben werden kann. Wer ist ‚schuld‘, wenn Alexa eine eskalierende Situation nicht rechtzeitig erkannt hat? Oder, wenn sie ‚falschen Alarm‘ gegeben hat und es nun Ärger mit der Bewährungshelferin gibt? Infrage kommen z.B. Personen oder auch Organisationen, die die KI-Programme programmiert, verkauft, installiert oder genutzt haben. Die Frage nach der Zuschreibung stellt sich auch in juristischer

Perspektive in Bezug auf die Möglichkeit einer ‚Willenserklärung‘ durch KI-Algorithmen und ihren zivil- und strafrechtlichen Status (vgl. Krimphove 2021).

Wichtig ist daher die Unterscheidung zwischen Systemen automatisierter Entscheidungsfindung und algorithmischen Systemen der Entscheidungsunterstützung, bei denen letztlich die Entscheidung aber von Menschen getroffen werden (vgl. Schneider 2021, 123). Zumindest bei Entscheidungen von großer Sensibilität oder Tragweite sollten letztere den Vorzug erhalten, um einer Verantwortungsdiffusion entgegenzuwirken.

5.2 ‚Künstliche Intelligenz‘ und die kategoriale Differenz von Mensch und Maschine

Auch wenn KI-Algorithmen menschliche Denkprozesse nachahmen und je nach Programmierung auch menschliches Verhalten zunehmend gut imitieren (vgl. Chatbots), bleiben sie von Menschen doch fundamental verschieden. Die Umwelterfassung durch Sensoren ist etwas anderes als Leiblichkeit, Wahrscheinlichkeitsaussagen sind etwas anderes als Zweifel und eine Selbstkorrektur ist etwas anderes als die Erfahrung von Scheitern und Schuld.

Den meisten Menschen dürfte dies durchaus bewusst sein. Dennoch fällt eine zum Teil weitreichende Anthropomorphisierung von Technik auf. Dies gilt für verbreitete Begriffe wie künstliche ‚Intelligenz‘, ‚Autonomie‘, ‚entscheiden‘ oder maschinelles ‚Lernen‘. Sie sind daher in diesem Text in Anführungszeichen gesetzt, weil es nahezu unmöglich scheint, auf sie ganz zu verzichten. Doch auch über solche einprägsamen Begriffe hinaus lässt sich eine Vermenschlichung beobachten – etwa wenn Chatbots Namen bekommen, die Sprachausgabe möglichst ‚menschlich‘ erscheinen soll oder auch ihre graphische Oberfläche als Mensch dargestellt wird (bis hin zu humanoiden Robotern). Solche Darstellungen mögen als witzig wahrgenommen werden und – vielleicht gerade für Adressat:innen sozialen Hilfehandelns – den Zugang zur Technologie erleichtern. Es scheint aber eine wichtige ethische Frage, inwiefern es hier zu einer problematischen und zum Teil bewusst eingesetzten Verwässerung der Differenz zwischen Mensch und Maschine kommt. Dies gilt auch umgekehrt, wenn Menschen sprachlich und gedanklich in die Nähe von Maschinen gerückt werden (wenn sie etwa ‚funktionieren‘, ihren ‚Akku aufladen‘ oder ‚ein Update bekommen‘).

5.3 ‚Moralische Maschinen‘ und Kriterien des ethischen Urteilens

Häufig wird gefordert, die Ethik müsse „in die KI integriert werden“. Diskutiert wird diese Frage etwa in Bezug auf Dilemmasituationen, die beim ‚autonomen‘ Fahren in Unfallsituationen entstehen können. Hier wird zum Teil der Eindruck erweckt, man könne ethische Probleme dadurch lösen, dass KI-Algorithmen die moralisch richtigen Normen einprogrammiert werden. Dies ist insofern richtig, als Algorithmen, wie in den vorangegangenen Abschnitten gezeigt, nicht neutral sind, sondern einen erheb-

lichen Einfluss auf gesellschaftliche Prozesse haben und sie (insbesondere bei Optimierungen) vorgegebene Wertungen verfolgen. Diese normative Dimension der Programmierung bzw. allgemein des Technikdesigns ist zu reflektieren und bei der Entwicklung von Algorithmen zu berücksichtigen. In diesem Sinne kann ein KI-Algorithmus tatsächlich in einem gewissen Sinne ‚moralisiert‘ werden.

Ein Missverständnis wäre jedoch die Vorstellung, ein Algorithmus könne so programmiert werden, dass er selbst eine ethisch verantwortete ‚Entscheidung‘ trifft. Das ist nicht nur aufgrund der fehlenden Subjekthaftigkeit eine irreführende Vorstellung. Der Gedanke setzt nämlich außerdem voraus, dass das ethisch Entscheidende überhaupt in algorithmische Strukturen überführt werden kann. Dies mag für streng deontische oder utilitaristische Ethiken denkbar sein. Gerade die Ethik der Sozialen Arbeit knüpft aber meist aus gutem Grund (zumindest auch) an andere Ethiktraditionen an. Dazu zählen tugend- und auch situationsethisch orientierte Ansätze, die Haltungen, Motive, Beziehungsqualität, Gefühle oder moralische Intuitionen berücksichtigen (vgl. Dallmann/Volz 2017, 23ff.). Werte wie Empathie, Loyalität, Zugewandtheit oder Nächstenliebe lassen sich nicht in Algorithmen überführen und setzen gerade die zwischenmenschliche Begegnung voraus.

Die zunächst vielleicht plausibel erscheinende Absicht, eine „moralische Maschine“³ zu programmieren, könnte also dazu führen, dass gerade diejenigen Faktoren verloren gehen, die für eine ethisch qualifizierte Entscheidung besonders bedeutsam sind.

5.4 Ethik-KI und die Kunst des ethischen Urteilens

Einen anderen Ansatz, KI-Algorithmen zu einem moralischen Urteil zu befähigen, verfolgt ein Projekt des Allen Institute for Artificial Intelligence an der University of Washington. Hier sollen den Algorithmen keine Werte einprogrammiert werden, sondern sie sollen selbst mit Mitteln der KI-Technologie auf der Grundlage mehrerer Datenbanken ‚lernen‘ ethische Urteile zu fällen (vgl. Jiang et al. 2021). Mit ‚Ask Delphi‘ steht seit Oktober 2021 ein Prototyp online, der die Funktionsweise des Algorithmus demonstriert und Urteile zu moralischen Fragen gibt („it’s good“, „it’s normal“, „it’s wrong“ etc.).⁴

Jedenfalls im gegenwärtigen Entwicklungsstand fällt es recht leicht, den Algorithmus dahingehend ‚auszutricksen‘, dass er offensichtlich fragwürdige oder widersprüchliche ethische Urteile abgibt (vgl. Vincent 2021). Bereits in den ersten Wochen konnten die in dieser Hinsicht jedoch gravierendsten Fehler behoben und die Qualität der Urteile merklich verbessert werden (vgl. Gault 2021). Wichtiger scheint demgegenüber, dass das Modell grundlegende Fragen aufwirft, die sich gerade bei einer qualitativen

³ Vgl. dazu das MIT-Projekt „Moral Machine“: <https://www.moralmachine.net/hl/de>.

⁴ Vgl. <https://delphi.allenai.org/>.

Verbesserung stellen werden. Erstens verfolgt die Forschungsgruppe ausdrücklich einen deskriptiv-ethischen Ansatz. Das Programm soll nicht selbst ein ethisches Urteil abgeben, sondern – der Funktionsweise der Mustererkennung folgend – statistisch auswerten, wie die jeweilige Frage (im englischsprachigen, US-amerikanischen Kontext) mehrheitlich beantwortet würde (vgl. Jiang et al. 2021, 4). Die Darstellung der Antwort verzichtet jedoch auf diese Differenzierung, sodass sehr wohl der Eindruck eines normativen Urteils entstehen kann. Ein gewisses Unbehagen daran lassen die Autor:innen des Papers selbst erkennen, wenn sie es als problematisch erachten, dass Delphi gesellschaftliche Stereotype und Vorurteile reproduziert (vgl. Jiang et al. 2021, 24). Zweitens kommt es zu einer geradezu grotesken Verzerrung dessen, was ethische Urteilsbildung ausmacht. Im Mittelpunkt steht ein Ergebnis, also eine Aussage darüber, was moralisch richtig, zulässig oder falsch ist. Dabei werden weder Ambivalenzen ausgewiesen noch Kontextualisierungen vorgenommen. Auch die Gründe für das jeweilige moralische Urteil werden nicht expliziert. Das Problem der Black Box wird hier in Bezug auf ein ethisches Urteil auf die Spitze getrieben. Es handelt sich (dem gewählten Namen durchaus entsprechend) letztlich eher um ein Ethik-Orakel als um eine ernst zu nehmende Form der Urteilsbildung.

Auch wenn Delphi ausdrücklich nicht als Ratgeber in moralischen Fragen fungieren soll, zeigt dieses Projekt doch deutlich die Grenzen einer ‚moralisierten‘ KI auf. Eine dialogisch-diskursive ethische Reflexion im eigentlichen Sinne kann sie weder durchführen noch ersetzen. Gerade diese aber ist für eine Professionsethik unabdingbar (vgl. Lob-Hüdepohl 2011, 19f.). Wird z.B. in der Werbung für eine Suchmaschine suggeriert, dass diese Antworten auf die großen Fragen des Lebens geben könne, wird eine scheinbar durch Technik zu gewinnende Eindeutigkeit simuliert. Die vermeintlich größere Objektivität und Verlässlichkeit von technischen Systemen kann durchaus dazu führen, dass Menschen gewichtige Entscheidungen an diese abgeben, um auf der ‚sicheren Seite‘ zu stehen. Eine Ethik der Sozialen Arbeit sollte demgegenüber Menschen befähigen und ermutigen, Ambivalenzen auszuhalten und eigenverantwortliche Entscheidungen zu treffen. Es ist wichtig, dass KI-Algorithmen moralische Normen berücksichtigen. Doch ethisch verantwortete Entscheidungen können – weder auf dem Wege einer ‚inprogrammierten‘ Moral noch durch einen eigenen Ethik-Algorithmus – an diese delegiert werden.

Fazit

‚Künstliche Intelligenz‘ identifiziert Muster, also Korrelationen und Regelmäßigkeiten in unserer Gesellschaft in einem bisher nicht gekannten Ausmaß. Sie ermöglicht es daher, menschliches Verhalten zunehmend präzise zu prognostizieren und zu steuern. Dies verändert unser Selbstverständnis, zwischenmenschliche Beziehungen und gesellschaftliche Strukturen. Dieser durch KI-Algorithmen wesentlich begünstigte, umfassende und vielschichtige Transformationsprozess kann kaum als ganzer ethisch beurteilt werden – und ganz sicher nicht als ‚gut‘ oder ‚schlecht‘.

Für die Ethik der Sozialen Arbeit stellen sich einerseits sehr konkrete, anwendungsorientierte Fragen nach fachlich und ethisch wünschenswerten Einsatzszenarien von KI-Technologie. Dabei sind technische Fragen wie die Auswahl von Trainingsdaten, die Aufbereitung von Daten, die Explizierbarkeit von Algorithmen, aber auch rechtliche Aspekte (beispielsweise zum Datenschutz, zur Transparenz, Diskriminierungsverbote oder den Grenzen des Einsatzes von KI-Algorithmen) zu bedenken. Auch die Verantwortung Einzelner für eine reflektierte Nutzung und damit verbundene Bildungsaufgaben sollten in den Blick genommen werden.

Doch angesichts ihrer breiten gesellschaftlichen Wirkung wäre es verkürzt, ‚künstliche Intelligenz‘ als ein Spezial- oder Sonderthema innerhalb der Ethik der Sozialen Arbeit zu begreifen. Sie prägt individuelles Leben und soziale Prozesse und wirft damit Fragen u.a. nach dem Menschenbild, der Sozial- und Wirtschaftsethik sowie nach ethischen Grundprinzipien auf. In all diesen Bereichen sollte die Ethik der Sozialen Arbeit also fragen, inwiefern KI-Algorithmen und die von ihnen ausgelösten Effekte einem gelingenden Leben für alle Menschen und besonders für die Zielgruppen Sozialer Arbeit dienlich sind und wo auf individueller, organisationaler oder gesamtgesellschaftlicher bzw. rechtlicher Ebene steuernd eingegriffen werden sollte.

Literatur

Albrecht, Thorben/ Kellermann, Christian (2020), Künstliche Intelligenz und die Zukunft der digitalen Arbeitsgesellschaft. Konturen einer ganzheitlichen Technikfolgenabschätzung, Hans Böckler Stiftung (Forschungsförderung Working Paper, 200), https://www.boeckler.de/de/faust-detail.htm?sync_id=HBS-007902 (abgerufen 14.11.2021).

Bastian, Pascal (2014), Statistisch Urteilen – professionell Handeln. Überlegungen zu einem (scheinbaren) Widerspruch, in: Zeitschrift für Sozialpädagogik 12 (2).

Bauberger, Stefan (2020), Welche KI? Künstliche Intelligenz demokratisch gestalten. München: Hanser (Hanser eLibrary).

Bendel, Oliver (Hg.) (2018), Pflegeroboter, Wiesbaden: Springer Gabler Open.

Beranek, Angelika (2021a), Auch eine Frage der Menschenrechte. Was ist Künstliche Intelligenz?, in: Politikum (1), 4–11, http://medienpaedagogik-bayern.com/wp-content/uploads/2021/02/Beranek_Politikum_1_2021_k.pdf (abgerufen 14.11.2021).

Beranek, Angelika (2021b), Zwischen Algorithmen und Wertediskurs. Auswirkungen der Digitalisierung auf die Profession der Sozialen Arbeit, in: Hammerschmidt, Peter/

Sagebiel, Juliane/ Hill, Burkhard/ Beranek, Angelika (Hg.), Big Data, Facebook, Twitter & Co. und Soziale Arbeit, 2. Aufl., Weinheim: Beltz (Aktuelle Themen und Grundsatzfragen der Sozialen Arbeit), 153–175.

Busch, Christoph (2018), Algorithmic Accountability, <https://www.abida.de/sites/default/files/ABIDA%20Gutachten%20Algorithmic%20Accountability.pdf> (abgerufen 14.11.2021).

Buxmann, Peter/ Schmidt, Holger (2021a), Ethische Aspekte der Künstlichen Intelligenz, in: Buxmann, Peter/ Schmidt, Holger (Hg.), Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg, 2. Aufl., Berlin, Heidelberg: Springer Berlin Heidelberg, 215–229.

Buxmann, Peter/ Schmidt, Holger (2021b), Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens, in: Buxmann, Peter/ Schmidt, Holger (Hg.), Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg, 2. Aufl., Berlin, Heidelberg: Springer Berlin Heidelberg, 3–24.

Christl, Wolfie (2019), Microtargeting. Persönliche Daten als politische Währung, in: APuZ 69 (24–26), 42–48.

Dallmann, Hans-Ulrich/ Volz, Fritz-Rüdiger (2017), Ethik in der Sozialen Arbeit, Berlin: Wochenschau Verlag (Grundlagen Sozialer Arbeit).

Di Fabio, Udo (2016), Grundrechtsgeltung in digitalen Systemen. Selbstbestimmung und Wettbewerb im Netz, München: C.H. Beck.

Döring, Nicola (2018), Sollten Pflegeroboter auch sexuelle Assistenzfunktionen bieten?, in: Bendel, Oliver (Hg.), Pflegeroboter, Wiesbaden: Springer Gabler Open, 249–267.

Fiske, Amelia/ Henningsen, Peter/ Buyx, Alena (2019), Your Robot Therapist Will See You Now. Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology and Psychotherapy, in: Journal of Medical Internet research 21 (5), <https://www.jmir.org/2019/5/e13216/> (abgerufen 14.11.2021).

Fritz, Olaf/ Weber, Carsten/ König, Angelika/ Wolf, Jan (2019), Ethische Aspekte der Künstlichen Intelligenz, Essen: MA (KCT Schriftenreihe, 1).

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

Gault, Matthew (2021), Ethical AI Trained on Reddit Posts Said Genocide Is OK If It Makes People Happy (Vice Motherboard), <https://www.vice.com/en/article/v7dg8m>

/ethical-ai-trained-on-reddit-posts-said-genocide-is-okay-if-it-makes-people-happy (abgerufen 14.11.2021).

Görder, Björn (2020), Wann ist KI ok? Der Einsatz künstlicher Intelligenz in der Sozialen Arbeit aus ethischer Perspektive, in: *Soziale Arbeit* (1), 16–22.

Habermas, Jürgen (2005), Religion in der Öffentlichkeit. Kognitive Voraussetzungen für den „öffentlichen Vernunftgebrauch“ religiöser und säkularer Bürger, in: Habermas, Jürgen (Hg.), *Zwischen Naturalismus und Religion. Philosophische Aufsätze*, Frankfurt am Main: Suhrkamp, 119–154.

Hagendorff, Thilo (2019), The Ethics of AI Ethics. An Evaluation of Guidelines, <https://arxiv.org/abs/1903.03425v2> (abgerufen 14.11.2021).

Independent High-Level Expert Group on Artificial Intelligence (2020), The Assessment List For Trustworthy Artificial Intelligence. Unter Mitarbeit von European Commission, Brüssel, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (abgerufen 14.11.2021).

Jiang, Liwei/ Hwang, Jena D./ Bhagavatula, Chandra (2021), Delphi. Towards Machine Ethics and Norms, Paul Allen Institute for Artificial Intelligence, University of Washington, <https://arxiv.org/pdf/2110.07574.pdf> (abgerufen 14.11.2021).

Kähler, Marco/ Feichtenbeiner, Rolf/ Beudt, Susan (2021), Facilitating the Implementation of AI-Based Assistive Technologies for Persons with Disabilities in Vocational Rehabilitation. A Practical Design Thinking Approach, in: Roll, Ido/ McNamara, Danielle/ Sosnovsky, Sergey/ Luckin, Rose/ Dimitrova, Vania (Hg.), *Artificial intelligence in education. 22nd international conference, AIED 2021, Utrecht, Cham: Springer (Lecture notes in computer science, 12749)*, 224–228.

Kamenetz, Anya (2021), Software Flags ‚Suicidal‘ Students, Presenting Privacy Dilemma. National Public Radio, <https://text.npr.org/470840270>, zuletzt aktualisiert am 14.11.2021.

Krimphove, Dieter (2021), Künstliche Intelligenz im Recht – eine Übersicht, in: *Juristische Ausbildung* (7), 764–769.

Kutscher, Nadia (2020), Ethische Fragen Sozialer Arbeit im Kontext von Digitalisierung, in: Kutscher, Nadia/ Ley, Thomas/ Seelmeyer, Udo/ Siller, Friederike/ Tillmann, Angela/ Zorn, Isabel (Hg.), *Handbuch Soziale Arbeit und Digitalisierung*, Weinheim: Beltz, 347–361.

Lenz, Gaby/ Lüssem, Jens/ Eilers, Hannes/ Wachter, Hannah (2019), Soziale Robotik in der Altenpflege. Zwischen Unbehagen und Neugier, in: *Soziale Arbeit* 68 (11), 402–409.

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

Lob-Hüdepohl, Andreas (2011), Nur moralisch – oder auch noch ethisch? Ethische Reflexionskompetenz als unabdingbares professionelles Werkzeug der Sozialen Arbeit, in: *SozialAktuell* (3).

Martini, Mario (2019), *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz*, Berlin, Heidelberg: Springer.

Müller, Vincent C. (2021), *Ethics of Artificial Intelligence and Robotics* (The Stanford Encyclopedia of Philosophy, Summer 2021 Edition), <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>, zuletzt aktualisiert am 26.05.2021 (abgerufen 14.11.2021).

Nassehi, Armin (2019), *Muster. Theorie der digitalen Gesellschaft*, München: C.H. Beck.

O’Neil, Cathy (2016), *Weapons of math destruction. How big data increases inequality and threatens democracy*, London: Penguin Books.

Reckwitz, Andreas (2018), *Die Gesellschaft der Singularitäten. Zum Strukturwandel der Moderne*, 6. Aufl., Berlin: Suhrkamp.

Schicha, Christian (2021), Informationsethische Herausforderungen durch Algorithmen, in: *Hammerschmidt, Peter/ Sagebiel, Juliane/ Hill, Burkhard/ Beranek, Angelika* (Hg.), *Big Data, Facebook, Twitter & Co. und Soziale Arbeit*. 2. Aufl., Weinheim: Beltz (Aktuelle Themen und Grundsatzfragen der Sozialen Arbeit), 95–117.

Schneider, Diana (2021), Ein Schritt in Richtung De-Professionalisierung? Plädoyer für eine intensive Diskussion über algorithmische Systeme in der professionellen Praxis, in: *Wunder, Maik* (Hg.), *Digitalisierung und Soziale Arbeit. Transformationen und Herausforderungen*, Bad Heilbrunn: Julius Klinkhardt, 122–139.

Schrödter, Mark/ Bastian, Pascal (2019), Risikodiagnostik durch „Big Data Analytics“ im Kinderschutz, in: *ARCHIV für Wissenschaft und Praxis der sozialen Arbeit* (2), 40–49.

Schrödter, Mark/ Bastian, Pascal/ Taylor, Brian (2020), Risikodiagnostik und Big Data Analytics in der Sozialen Arbeit, in: *Kutscher, Nadia/ Ley, Thomas/ Seelmeyer, Udo/ Siller, Friederike/ Tillmann, Angela/ Zorn, Isabel* (Hg.), *Handbuch Soziale Arbeit und Digitalisierung*, Weinheim: Beltz, 255–264.

Staub-Bernasconi, Silvia (2007), Soziale Arbeit: Dienstleistung oder Menschenrechtsprofession? Zum Selbstverständnis Sozialer Arbeit in Deutschland mit einem Seitenblick auf die internationale Diskussionslandschaft, in: *Lob-Hüdepohl, Andreas/*

Lesch, Walter (Hg.), Ethik Sozialer Arbeit. Ein Handbuch, Paderborn, München, Wien: Schöningh.

Stock, Jessica (2020), Mit Künstlicher Intelligenz berufliche Teilhabechancen verbessern. Lern- und Experimentierräume als partizipativer Gestaltungsansatz, in: Trimpop, Rüdiger/ Fischbach, Andrea/ Seliger, Iris (Hg.), Gewalt in der Arbeit verhüten und die Zukunft gesundheitsförderlich gestalten! 21. Workshop Psychologie der Arbeitssicherheit und Gesundheit, Kröning: Asanger Verlag, 545–548.

Vaithianathan, Rhema/ Putnam-Hornstein, Emily/ Jiang, Nan (2017), Developing Predictive Models to Support Child Maltreatment Hotline Screening Decisions. Allegheny County Methodology and Implementation, <https://www.alleghenycountyanalytics.us/wp-content/uploads/2017/04/Developing-Predictive-Risk-Models-package-with-cover-1-to-post-1.pdf> (abgerufen 14.11.2021).

Vincent, James (2021), The AI oracle of Delphi uses the problems of Reddit to offer dubious moral advice. A fascinating project that's best understood as a cautionary tale. The Verge, <https://www.theverge.com/2021/10/20/22734215/ai-ask-delphi-moral-ethical-judgement-demo> (abgerufen 14.11.2021).

ZDF Magazin Royal (2021), TargetLeaks, Sendung vom 24. September 2021, <https://targetleaks.de/> (abgerufen 14.11.2021).

Zuboff, Shoshana (2019), Surveillanca Capitalism – Überwachungskapitalismus, in: APuZ 69 (24–26), 4–9.

Über den Autor

Prof. Dr. Björn Görder ist Professor für Theologie und Ethik sozialer Handlungsfelder an der Evangelischen Hochschule Ludwigsburg. Der Theologe und Wirtschaftswissenschaftler ist Vorsitzender des Arbeitskreises für theologische Wirtschafts- und Technikethik (ATWT).

Über www.ethikjournal.de

EthikJournal ist eine Onlinezeitschrift für Ethik im Sozial- und Gesundheitswesen. Ausgehend von aktuellen Themen werden grundlegende theoretische und handlungsorientierte Fragen zur Diskussion gestellt. Die Zeitschrift erscheint zweimal jährlich online. Herausgeber der Zeitschrift ist das Berliner Institut für christliche Ethik und Politik (ICEP).

ISSN 2196-2480

Zitationsvorschlag

Görder, Björn (2021), Die Macht der Muster. Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch ‚künstliche Intelligenz‘, in: EthikJournal Ausgabe 2/2021, Heft 1, Download unter: [Link zum pdf Onlinedokument](#) (Zugriff am).