

THEOLOGISCHE REVUE

121. Jahrgang
– Oktober 2025 –

Künstliche Intelligenz und ethische Verantwortung, hg. v. Michael REDER / Christopher KOSKA. – Bielefeld: transcript 2024. 264 S., kt. € 37,00 ISBN: 978-3-8376-6905-3

Der Sammelbd. *Künstliche Intelligenz und ethische Verantwortung* enthält Aufsätze, die sich mit der Frage nach KI und Verantwortung in sozialen Praktiken beschäftigen. Werden diese Praktiken – in denen bislang menschliches Handeln und Entscheiden der Regelfall war – in algorithmische Strukturen überführt, kommen Fragen auf, die zu einer Kritik dieser Technologie führen.

Michael Reder, Nicholas Müller und Robert Lehmann skizzieren in ihrem einleitenden Beitrag das Problemfeld als eines, auf dem menschliches und maschinelles Entscheiden sowohl in Konkurrenz als auch als sich ergänzend betrachtet werden kann. Die Umsetzung von KI-basierten Prognosen sowie deren Chancen und Risiken in der Sozialen Arbeit werden anhand von Literatur herausgearbeitet und diskutiert (13–16).

In einem ideengeschichtlichen Beitrag stellt *Rudolf Seising* die Idee von *Artificial Intelligence* vor dem Hintergrund verschiedener Wissenschaftskulturen dar, und zeigt die grundlegende Interdisziplinarität dieses Konzepts (29–38), indem er es auf den begriffsgeschichtlichen Seziertisch legt. Dieser auch technikgeschichtliche Abriss (38–42) hilft dabei, KI auch jenseits ihrer technischen Leistungen und deren soziale Folgen kritisch einzuordnen.

Tanja Henking diskutiert in ihrem Beitrag die Möglichkeit, via KI moralische Entscheidungen zu treffen. Am Beispiel von Entscheidungsunterstützungssystemen im medizinischen Kontext fragt sie, inwieweit dort erkannte Dilemmastrukturen von einer KI bearbeitet werden können und mit welchen Folgen dies geschehen würde – sowohl auf der moralischen Ebene als auch auf der vorgelagerten Sachebene (55f). Henking kommt zum Fazit, dass eine ethische Entscheidung immer eine personengebundene bleibt (59).

Maximilian Kraus, Jennifer Burghardt und Christopher Koska argumentieren dafür, dass die KI-unterstützte Fallbearbeitung in der Kinderwohlfahrtspflege bei allem Nutzen den Menschen nicht ersetzen könne. KI-Unterstützung bedeute in diesem Zusammenhang, dass die Urteilskraft gestärkt werde und Fälle genauer und tiefgründiger bearbeitet werden könnten (63f). Unabhängig vom KI-Einsatz gebe es einen statistischen sowie einen hermeneutisch orientierten Zugang zur Risikobewertung (66). Mit KI werde ersterer bedient. Setze man KI auf dem Gebiet des Kinderschutzes ein, sei es wichtig, über die Risiken informiert zu sein. Anhand des KAImo-Projekts stellen die V.f.:innen ein unterstützendes KI-System vor und diskutieren dessen verantwortbaren und die damit einhergehenden Herausforderungen konstruktiv (73–79).

Kerstin Schlögl-Flierl und *Paula Ziethmann* schlagen in ihrem Beitrag die Verbindung von KI-Ethik und einer Ethik der Vulnerabilität vor, um dem Interesse des Kindeswohls bestmöglich zu begegnen. Ausgehend von einer Analyse der Situationen in Neuseeland und Deutschland zeigen die Autorinnen auf, dass die Frage, ab wann eine Behörde beim Verdacht auf Kindeswohlgefährdung intervenieren soll, nicht einfach algorithmisch entschieden werden könne (88–90). Der Status *vulnerabel* treffe auf alle am Prozess beteiligten zu (91–94). Ein reflektierter KI-Einsatz in diesem Feld könne den Ansprüchen aller Stakeholder begegnen, sollten solche Tools als Ergänzung eingesetzt werden, um die Effizienz und Präzision sozialer Arbeit zu steigern (97).

Daniel Schlör und *Andreas Hotho* diskutieren in ihrem Beitrag „Verantwortungsvolle Empfehlungssysteme für die medizinische Diagnostik“ die Bedingungen für den Einsatz solcher Systeme anhand von Anwendungsszenarien und stellen Kriterien dafür vor. Die Vf. beziehen sich auf ein Empfehlungssystem für den klinischen Kontext und identifizieren mögliche Effizienzgewinne wie auch Risiken durch dessen Einsatz (107–109). Ihre Diskussion der Aspekte Verantwortung, Fairness, Rechenschaftspflicht, Transparenz, Compliance und Sicherheit führt sie zum Fazit (109–114), dass unter der Annahme, dass die Letztentscheidung bei den Ärzt:innen liegt, ein Einsatz vertretbar sei, wenn Compliance und Sicherheit gewährleistet würden – auch wenn es hinsichtlich des Aspekts Fairness offene Fragen gebe (144f).

„Zu viel Gewissheit?“ fragen *Max Tretter*, *Ulrich Freiherr von Ulmenstein*, *Christina Lauppert von Peharnik* und *David Ehrlich* mit Blick auf die Ergebnisse ihres Gedankenexperiments zu den Folgen des Einsatzes KI-gestützter personenbezogener Gesundheitsvorhersagen für die öffentliche Gesundheitsversorgung. Aus einer pragmatischen Perspektive argumentieren die Vf.:innen, dass durch adverse Selektion die Finanzierungsprobleme im Gesundheitssystem im Bereich der GKV zunehmen würden (122). Diese komme dadurch zustande, dass die nach Gewissheit hinsichtlich ihrer Gesundheit Strebenden, sich insbesondere dann für die GKV entschieden, wenn ihre Prognose negativ ausfalle, während solche, deren Risiko durch eine KI als gering beurteilt wird, in die PKV wechseln würden (131f).

Dem Thema „Algorithmische Differenzierung und Diskriminierung aus Sicht der Menschenwürde“ widmet sich *Carsten Orwat*. Basierend auf einer Darlegung der Bedeutung der Menschenwürde im verfassungsrechtlichen Kontext betrachtet er u. a. Restrisiken vor einem antidiskriminierungsrechtlichen Hintergrund (145–148) und diskutiert Möglichkeiten der Menschenwürdeverletzung durch Algorithmen. Sie erörtert die Problemlagen „schwerwiegende und strukturelle Diskriminierung“, die ohnehin marginalisierte Gruppen betreffe (148f) sowie die „statistische Diskriminierung“, die den Einzelfall nicht betrachte (150). Automatisierte Entscheidungen, die Machtasymmetrie zwischen Betroffenen und Staat (156f) sowie die Frage nach einer validen Zustimmungspraxis, die Zweifelhafte legalisiere (158f), können die Menschenwürde potentiell ebenso gefährden. Orwat plädiert dafür, die mögliche Verletzung sowie den Schutz der Menschenwürde unmittelbar sowie mittelbar Betroffener weiter zu untersuchen (160f).

Cecilia Colloseus stellt in ihrem Beitrag zur Normung von KI aus soziotechnischer Perspektive Operationalisierungsprobleme in den Mittelpunkt und geht auf den entsprechenden Teil der Normungsroadmap des DIN ein (169f). Sie unterstreicht die Wichtigkeit partizipativer Technikentwicklung für die Vermeidung von *Biases* (172f). Der Darstellung von normativen Anforderungen an sozio-technische Systeme folgt das Fazit, dass Betroffene „in jeden Schritt der Entwicklung der individuellen KI-Lösung einbezogen werden [müssen]“ (182).

Susanne Wolf, Jan Fiete Schütte, Marc Hauer und Koska legen ein Konzept zur Operationalisierung und Messung der Ressource Vertrauen vor. Sie klären zunächst die möglichen verschiedenen Kontexte des Vertrauensbegriffs (191–193), um dann in einem Szenario eine vertrauenswürdige Datennutzung auszubuchstabieren (193–200). Um Vertrauen zu messen, können Tests eingesetzt werden, u. a. um zunächst die Akzeptanz von etwas zu ermitteln und anschließend mit Audits eine Kriterienerfüllung zu überprüfen (201f). Dabei bleibe der Kontext wichtig. Auch diese Vf.:innen plädieren dafür, Stakeholder in alle Prozesse mit einzubeziehen, um Vertrauen auf der Mikro- und der Meso-Ebene – und damit auch das Systemvertrauen zu erhalten (211).

Nicole Kunkel fragt in ihrem Beitrag „Humaner als der Mensch?“ nach den sozialen Vorstellungen autonomer Waffen. Sie geht dabei von den Filmen *Terminator* und dessen zweitem Teil aus sowie davon, dass Technik nicht den Wert menschlichen Lebens lernen könne. Kunkel setzt dieser „absurde[n] Fantasie“ die *imago Dei* des Menschen als Korrektiv entgegen (226–230). Dafür nimmt sie Charles Taylors Theorie der sozialen Imagination in Anspruch, um zu zeigen, wie ein technikinduziertes Missverständnis des Menschen von sich selbst entstehen kann. Ihr Gegenvorschlag für eine Anthropologie besteht darin, nicht von der Maschine auszugehen, sondern von der Fähigkeit des Menschen mitzuleiden und mitzufühlen. Daraus ergebe sich die Aufforderung, Maschinen nicht zu vermenschlichen und ihnen keine menschlichen Fähigkeiten zu überlassen (230f).

Sophie Jörg greift in ihrem Aufsatz die Frage nach der Autonomie menschlichen Handelns vor dem Hintergrund algorithmeninduzierter Veränderungen in der Mensch-Computer-Interaktion auf. In diesem Zusammenhang will sie menschliche Autonomie als notwendige Voraussetzung demokratischer Praxis verstanden wissen (236).

Da Sammelbd.e nur in Ausnahmefällen von Anfang bis Ende gelesen werden, stört es nicht, dass hier eine Sammlung einigermaßen unverbunden nebeneinanderstehender Aufsätze vorliegt, die jedoch alle etwas aus ihrer je eigenen Perspektive zum Thema KI und Verantwortung beitragen.

Auf einem Gebiet, das von der Unschärfe der Ethik geprägt ist, geht es gar nicht anders, als viele Fragen offen zu lassen – der Gewinn liegt aber in der Vorbereitung ihrer (vorläufigen) Klärung durch die Angebote, die in den Beiträgen vorgelegt werden.

Für die theol. Ethik der Gegenwart sind die Beiträge auch dann relevant, wenn sie nicht als explizit theol. erkennbar sind. Obwohl mit Kunkels Beitrag nur ein dezidiert theol. enthalten ist – der auch gerade nicht das Rahmenthema der Sozialen Arbeit berührt –, bietet der Bd. für eine Rezeption in der Theol. doch vielfältige und wichtige Anknüpfungspunkte für ein weiteres Nachdenken. Dies betrifft insbesondere das Gebiet der Anthropologie, wie Kunkel einerseits, sowie Schlögl-Flierl und Ziethmann andererseits herausgearbeitet haben. Vor dem Hintergrund, immer mehr ursprünglich originär menschliche Tätigkeiten durch Maschinen zu ergänzen oder gar zu ersetzen, wird eine Frage umso akuter: Was ist der Mensch?

Über den Autor:

Hermann Diebel-Fischer, Dr., Wissenschaftlicher Mitarbeiter am Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) an der Technischen Universität Dresden (hermann.diebel-fischer@tu-dresden.de)