

EINLEITUNG

KI-Systeme

Aktuelle Trends und Entwicklungen aus Perspektive der Technikfolgenabschätzung

Bernhard G. Humm, Hochschule Darmstadt – University of Applied Sciences, Fachbereich Informatik,
Haardtring 100, 64295 Darmstadt, DE (bernhard.humm@h-da.de)  0000-0001-7805-1981

Stephan Lingner, Institut für qualifizierende Innovationsforschung und -beratung GmbH (IQIB), Bad Neuenahr-Ahrweiler, DE (stephan.lingner@iqib.de)

Jan C. Schmidt, Hochschule Darmstadt – University of Applied Sciences, Hochschule Darmstadt,
Fachbereich Gesellschaftswissenschaften, Darmstadt, DE (jan.schmidt@h-da.de)

Karsten Wendland, Karlsruher Institut für Technologie (KIT), Institut für Technikfolgenabschätzung und Systemanalyse (ITAS),
Karlsruhe, DE (karsten.wendland@kit.edu)  0000-0003-4812-0928

Zusammenfassung • Künstliche Intelligenz (KI) ist – ebenso wie damit verknüpfte Techniken wie maschinelles Lernen und Big Data – in aller Munde. Die große Dynamik und Tragweite dieser Entwicklungen zeigen sich bereits in zahlreichen Anwendungsgebieten von Wirtschaft, Gesellschaft und Wissenschaft. Technikfolgenabschätzung (TA) von KI hat in diesem Zusammenhang zunächst die Aufgabe, etwaige überzogene öffentliche Erwartungen und Befürchtungen in sachliche, realistische Perspektiven zu transformieren. In einem zweiten Schritt kann TA entlang begründbarer Entwicklungsziele von KI und legitimer gesellschaftlicher Wertvorstellungen Impulse für die weitere, wünschbare Gestaltung von KI geben. Wenn TA diese Orientierungsaufgabe nah am technologischen Kern wahrnimmt, findet sie dabei große gestalterische Freiräume in frühen Phasen der Technikentwicklung vor. Die damit zusammenhängenden Gedanken werden im vorliegenden Einleitungskapitel konkretisiert und auf die Beiträge zu diesem Themenschwerpunkt angewendet.

Designing and experiencing AI systems. Recent trends and developments from a technology assessment perspective

Abstract • Artificial intelligence (AI) is on everyone's lips – as well as the associated technologies of machine learning and big data. The enormous dynamics and consequences of these developments become already evident in numerous areas of application in business, society and science. In this context, technology assessment (TA) of AI initially has the task of transforming any excessive public expectations and fears to the factual level. In a second step, TA can provide impulses for

the further, desirable design of AI based on reasonable development goals of AI and legitimate societal values. If TA conducts this orientation task close to the technological core, it can consider wide scopes of options for action in the early phases of technology development. Related thoughts are put into concrete terms in this article and will be related to the authors' contributions to this topical focus.

Keywords • Artificial Intelligence, AI, Machine Learning, Big Data, Technology Assessment

Künstliche Intelligenz (KI) und die mit ihr verbundenen Ermöglichungstechniken wie maschinelles Lernen und Big Data sind derzeit von einer großen öffentlichen Aufmerksamkeit gekennzeichnet. Die hohe Entwicklungsdynamik ist zudem mit einem erheblichen Veränderungspotenzial für Wirtschaft, Gesellschaft und Wissenschaft verknüpft. In der öffentlichen Debatte kursieren zudem teilweise unreflektierte Erwartungen an KI und auch überzogene Befürchtungen vor ihrem breiten Einsatz. Angesichts ihrer möglichen Konsequenzen für alle gesellschaftlichen Lebensbereiche erscheint es daher geboten, Erforschung, Entwicklung und Einsatz von KI entlang begründbarer Zielvorstellungen und legitimer Wertvorstellungen zu rahmen und zu gestalten. Dies ist eine Aufgabe der Technikfolgenabschätzung (TA), die hier zur Klärung und Orientierung auf unsicherem und ambivalentem Terrain in ähnlicher Weise gefragt ist, wie bei anderen tiefgreifenden Entwicklungen in Bereichen der Nanotechnologie oder der Biotechnologie. Technikfolgenabschätzung von KI betrifft dabei nicht nur die Implementierung und Anwendungsfelder von bereits entwickelten KI-Technologien, sondern insbesondere auch die vorgelagerte, im öffentlichen Diskurs oft vernachlässigte Ebene der Forschung. In diesen Frühphasen der

This is an article distributed under the terms of the Creative Commons Attribution License CCBY 4.0 (<https://creativecommons.org/licenses/by/4.0/>) <https://doi.org/10.14512/tatup.30.3.11>
Received: Oct. 27, 2021; revised version accepted: Nov. 02, 2021; published online: Dec. 20, 2021 (editorial peer review)

Technikentwicklung ist TA allgemein mit großen gestalterischen Freiräumen nah am technikwissenschaftlichen Kern von KI konfrontiert, aber eben auch mit besonderen methodischen Herausforderungen der Vorausschau ihrer Folgen (Collinridge 1982).

So ist hier zunächst die Suche nach geeigneten Ansätzen *für eine prospektive Beurteilung von KI* zu nennen. Dabei stellt sich unter anderem die Frage, wie retrospektive Analysen von bereits abgeschlossenen – und somit evidenten – Digitalisierungsdebatten für heutige Ideen für KI-Zukünfte nutzbar gemacht werden können.

Für die gestaltungsorientierte Beurteilung von KI ist sicher von zentraler Bedeutung, wie sich die *Autonomie und Kreativität sowie die Komplexität und Opazität* künstlicher ‚intelligenter‘ Systeme im gesellschaftlichen Alltag und in der Forschungspraxis aus TA-Sicht darstellt. Hier sind zum Beispiel spezifische Probleme künstlicher Kommunikation durch technische Agenten in sozialen Medien zu nennen. Ein weiterer Punkt ist die Frage, wie vertrauenswürdige KI entwickelt und in die Praxis umgesetzt werden kann und welche Hemmnisse dabei zu überwinden sind.

Daran schließt sich unmittelbar auch die Notwendigkeit der Gestaltung von geeigneten Rahmenbedingungen für den Einsatz von KI-Systemen in der gesellschaftlichen Praxis an. Hier sind zunächst *aussichtsreiche Regulierungsansätze*, ihre ‚offenen Flanken‘ sowie mögliche Lösungsvorschläge zu erarbeiten, die wünschbare KI-Innovationen fördern können. Diese Vorschläge sollten dabei auch auf die unterschiedlichen immanenten Geschwindigkeiten gesetzgeberischer und technischer Entwicklungen eingehen.

Ein anderes praktisches Problem ist die Frage, wie *notwendige KI-Kompetenzen* für den aufgeklärten Einsatz entsprechender Systeme in der Bevölkerung vermittelt und verankert wer-

ansätze zu ertüchtigen; mögliche Anwendungsfälle aus vielen Beispielen sind die Abschätzung der Chancen und Risiken von KI in der biomedizinischen Diagnose und Therapie von Krankheiten.

Die folgenden Abschnitte erläutern den Kern von KI-Systemen und ihrer Nutzbarkeit vor dem Hintergrund ihrer öffentlichen Wahrnehmung. Es folgt sodann eine kursorische Übersicht der Themenbeiträge dieses Schwerpunkts mit Vorschlägen für die Gestaltung von KI-Systemen und ihres Anwendungsrahmens.

KI – Hype und Realität¹

Die hohe Öffentlichkeitswirksamkeit von KI ist mit überzogenen Erwartungen einerseits sowie mit übertriebenen Befürchtungen andererseits verknüpft. Hinzu kommt, dass KI ein beliebtes Thema für Kinofilme ist und so die Gefahr besteht, dass Fiktion und Wirklichkeit im Bewusstsein vieler verschwimmen.

Herbert Simon, einer der Gründerväter von KI meinte bereits 1965: „Machines will be capable [...] of doing any work that a man can do“ (Allen 2001). Und sein Kollege Marvin Minski prophezeite 1970: „Within 10 years computers won't even keep us as pets“ (Allen 2001). Keine dieser Prophezeiungen hat sich auch nur annähernd erfüllt. Doch das hält zeitgenössische Autor*innen und Visionär*innen auch heute nicht davon ab, überzogenen Erwartungen und Befürchtungen medienwirksam auszurufen. Prominente Vertreter gewagter Positionen sind Ray Kurzweil, Nick Bostrom, Yuval Noah Harai, Elon Musk, sowie der verstorbene Stephen Hawking. So behauptet Ray Kurzweil (2015, S. 9): „Mit der Singularität werden wir die Grenzen unserer biologischen Körper und Gehirne überschreiten. Wir werden

Die hohe Öffentlichkeitswirksamkeit von KI ist mit überzogenen Erwartungen einerseits sowie mit übertriebenen Befürchtungen andererseits verknüpft.

den können. Dieses Anliegen richtet sich darauf, in der Breite der Gesellschaft ein KI-Verständnis zu stärken, das technischen Laien erlaubt, Möglichkeiten und Risiken von KI-Systemen besser einschätzen können, um ihre Handlungs- und Entscheidungskompetenzen auch zukünftig einsetzen zu können. Auf Unternehmensebene stellt sich auch die Frage, wie der mitarbeiterfreundliche Einstieg in betriebliche KI-Anwendungen gelingen kann und welche flankierenden Forschungsansätze hierzu beitragen können. Diese sind insbesondere auf die Analyse und Abschätzung kritischer *Vertrauens- und Akzeptanzbedingungen* hin zu entwickeln.

Die vorstehenden Überlegungen sind aber auch auf konkrete Problemkontexte hin zu beziehen und für mögliche Lösungs-

die Gewalt über unser Schicksal erlangen. Unsere Sterblichkeit wird in unseren Händen liegen. Wir werden so langen leben können wie wir wollen [...]. Bis zum Ende des Jahrhunderts wird die nichtbiologische Komponente unserer Intelligenz Trillionen Mal mächtiger sein als bloße menschliche Intelligenz“.

Die meisten professionellen Einschätzungen von KI sind demgegenüber deutlich nüchterner. Uns Gastherausgebern ist Stand heute keinerlei Evidenz bekannt, dass sogenannte ‚starke KI‘, also KI-Anwendungen, die sich, wie oben prognostiziert,

¹ Dieser Abschnitt ist in weiten Teilen sinngemäß der Studie von Gethmann et al. (2021) entnommen, an der Bernhard Humm, Stephan Lingner und Jan Schmidt maßgeblich mitgewirkt haben.

eigenständig weiterentwickeln, wirklich möglich sind. In jedem Fall sind diese aus unserer Sicht nicht in greifbarer Nähe. Besser werden lediglich die Imitationen und Inszenierungen von Intelligenz. Auch der hohe Anspruch, dereinst KI mit Bewusstsein zu erschaffen, geht stark mit der Reduzierung des Bewusstseinsbegriffs auf Funktionalitäten einher, nicht aber mit etwaigen Erlebnisqualitäten. Alle KI-Anwendungen, die bis heute entwickelt wurden, werden der sogenannten ‚schwachen KI‘ zugeordnet. Sie sind auf spezielle Aufgaben zugeschnitten, deren Lösung bislang menschliche Fähigkeiten erforderten, und lösen diese selbstständig.

Klar ist, dass KI bereits heute eine hohe Bedeutung in Wirtschaft und Gesellschaft hat und diese Bedeutung voraussichtlich weiter zunehmen wird. Für viele Unternehmen hat KI eine strategische Bedeutung. So stammt beispielsweise folgende Aussage von Amazon: „Without ML [machine learning], Amazon.com couldn't grow its business, improve its customer experience and selection, and optimize its logistic speed and quality.“ (Marr 2018). Ähnlich äußern sich Google, Facebook, IBM und andere Technologie-Konzerne (Marr 2018). Auch die deutsche Bundesregierung hat 2018 eine KI-Strategie (BMBF 2021) verabschiedet mit dem Ziel, die Erforschung, Entwicklung und Anwendung von künstlicher Intelligenz in Deutschland auf ein weltweit führendes Niveau zu bringen und zu halten.

Perspektiven von KI

KI steht in der Tradition von Automatisierungs-Technologien, bei der menschliche Arbeit von Maschinen übernommen wird. Besonders seit der industriellen Revolution verändert diese Automatisierung in rasanter Weise alle Lebens- und Arbeitsbereiche. Viele Berufsgruppen sind komplett verschwunden, neue sind entstanden, und Menschen sowie Gesellschaften mussten sich stets an veränderte Gegebenheiten anpassen, die durch technische Entwicklungen verursacht wurden. Während im Zuge der industriellen Revolution vorwiegend physische Tätigkeiten von Maschinen übernommen wurden, werden im Zuge der Digitalisierung, insbesondere mit KI-Techniken, zunehmend auch kognitive Tätigkeiten automatisiert. Damit verbunden sind Chancen und Risiken, die oft eng beieinander liegen, sich sogar gegenseitig bedingen können. Wir möchten dies anhand von Anwendungsbeispielen erläutern.

KI wird zunehmend in der *Medizin* eingesetzt; so werden beispielsweise bildgebende Verfahren mittels Röntgenstrahlung, Ultraschall und MRT in Kombination mit KI-Verfahren zur (semi-)automatischen Analyse beziehungsweise Diagnose eingesetzt. Diese übertreffen in speziellen Gebieten teilweise sogar menschliche Ärzt*innen. Die Chancen hier liegen in der verbesserten Diagnostik, einer höheren Reichweite (z. B. bis hin zu Hausarztpraxen) und reduzierten Kosten im Gesundheitswesen. Demgegenüber steht das Risiko von Fehldiagnosen z. B. aufgrund von Verzerrungen (bias) in den Trainingsdaten für machine learning-Anwendungen. Ironischerweise kann dies eine

indirekte Folge der Chance verbesserter Diagnostik sein, wenn über einen längeren Zeitraum der erfolgreichen Verwendung einer KI-Anwendung sich schleichend ein blindes Vertrauen einstellt, insbesondere bei einer neuen Generation von Ärzt*innen. Dann kann schrittweise die Expertise zur menschlichen Diagnose, welche der maschinellen Diagnose kritisch gegenübergestellt wird, sinken.

Als nächstes Beispiel betrachten wir das sogenannte *autonome Fahren*, derzeit noch in der Experimentierphase, aber mit Ausblick auf breiten Einsatz in vielen Ländern. Die Chancen bestehen in einer höheren Effizienz und damit verbunden geringeren Kosten, z. B. durch Wegfall von Taxi-Fahrer*innen. Man kann auch hoffen, dass mit der Automatisierung car sharing-Konzepte umgesetzt werden können, sodass sich perspektivisch die Anzahl der Fahrzeuge, somit auch die damit verbundene Umweltbelastung sowie die Verkehrsverdichtung der Städte reduziert. Mit fortschreitender Technik könnten autonome Fahrzeuge im Durchschnitt sicherer fahren als menschliche Fahrer*innen. Allerdings bleibt auch bei immer besserer Technik grundsätzlich ein Risiko von Unfällen, verursacht durch ‚autonome‘ Fahrzeuge. Hier bestehen noch große ungeklärte juristische Fragen, z. B. nach der Verantwortung und Haftung. Die Sicherheit kann aber auch drastisch sinken, da mit durchgängigem autonomem Fahren das Risiko großflächiger, verheerender Cyberangriffe im Verkehr steigt. Auch die Hoffnung auf eine Reduzierung der Umweltbelastung könnte sich in das Gegenteil umkehren, wenn die Möglichkeit von einfach verfügbarem Individualtransport mit ‚autonomen‘ Fahrzeugen den Bedarf nach Mobilität ständig wachsen lässt – ein sogenannter Rebound-Effekt.

So genannte *Business Intelligence Systeme* unterstützen das Management von Unternehmen, Geschäftsentscheidungen zu treffen. Häufig werden dabei KI-Techniken eingesetzt, z. B. die Sentiment-Analyse von social media-Kanälen. Chancen sind das frühzeitige Erkennen, bessere Verständnis und schnelle Reagieren auf Kundenmeinungen – sowohl Wünsche als auch Kritik. Aber darin liegen auch Risiken. Werden Kunden wirklich oder nur vermeintlich besser verstanden? Es besteht die Gefahr einer Schein-Objektivität beziehungsweise Zahlengläubigkeit, die echter Einsicht entgegensteht. Insbesondere sind solche Analysen nur bedingt für strategische Zukunftsplanung geeignet, da sie nur die Vergangenheit abbilden können. Außerdem wächst mit der Bedeutung von solchen automatisierten Analysen die Gefahr der absichtlichen Manipulation, z. B. durch fake news, welche von Computerprogrammen (bots) generiert werden.

Beim *Hochfrequenzhandel* werden umfangreiche Börsentransaktionen automatisch mittels KI-Anwendungen durchgeführt. Chancen bestehen für einzelne Unternehmen hier auf zusätzliche Gewinne durch Ausnutzen von kurzlebigen Kurschwankungen. Für den Markt und damit die gesamte Wirtschaft bestehen aber auch erhebliche Risiken durch nicht kalkulierbare Rückkopplungseffekte. So können einzelne Fehlentscheidungen von vielen anderen Systemen wiederholt werden; ein gegenseitiges Aufschaukeln kann nicht vorhersehbare Folgen mit sich bringen.

Robotik wird intensiv in der industriellen Fertigung eingesetzt. Die Chancen liegen sowohl in einer gesteigerten Effizienz als auch in einer höheren Präzision. Aber grundsätzlich bedingt eine gesteigerte Effizienz das Risiko einer geringeren Stabilität. So birgt hocheffiziente just in time-Produktion mit minimierter Lagerhaltung die Gefahr eines Zusammenbruchs der Lieferketten, wie es bei der Corona-Pandemie der Fall war. Auch hier gilt, dass komplette Vernetzung – die Grundlage für die hohe Automatisierung – gleichzeitig Einfallstore für umfassende Cyberangriffe bietet.

Überwachungssysteme setzen zunehmend KI-Methoden ein, z. B. Systeme zur Gesichtserkennung an Flughäfen. Chancen solcher Systeme sind erhöhte Sicherheit, auch bei der Detek-

gen zurück, die in frühen Dokumenten angelegt wurden. Heute wirkmächtig und analytisch aufschlussreich sind also technische Zukünfte der Vergangenheit. Heutige Darstellungsformen von KI prägen dementsprechend auch zukünftige Wirkungen und Entwicklungen.

Die zunehmende Verbreitung von Sprachassistenten, Chatbots und anderen sprachverarbeitenden Programmen nimmt Sascha Dickel zum Anlass, kommunizierende Technik kritisch zu beleuchten – vor allem im Hinblick auf menschliche Deutungen der ihr zugeschriebenen Autonomie, Kreativität und (In-)Transparenz. Mithin stellt sich der Bedarf nach Identifizierbarkeit des Menschen besonders in hybriden Kommunikationsensembles moderner mediatisierter Gesellschaften, in denen maßge-

Komplette Vernetzung bietet Einfallstore für Cyberangriffe.

tion von potenziellen Straftätern, die weltweit agieren. Damit verbunden ist aber auch das Risiko der ‚falschen Sicherheit‘, da solche Erkennungssysteme nicht perfekt sind und auf vielfältige Weise manipuliert werden können.

Schließlich kann KI vielfältig für *militärische Anwendungen* eingesetzt werden, z. B. in Form autonomer Gefechtsdrohnen. Die Chance besteht darin, dass Länder sich besser verteidigen und ihre eigenen Soldaten besser schützen können. Es gibt aber auch enorme Risiken. Zum einen wird das Töten von Menschen vereinfacht, was durch die Verbreitung jener autonomer Waffen auf alle Staaten zurückfallen kann. Aus ethischer Sicht besonders zu kritisieren ist, die Entscheidung über Leben und Tod Maschinen zu überlassen. Die Gefahr besteht, dass technische Möglichkeiten neue Realitäten hervorbringen; fangen die einen an, so müssen die anderen nachziehen.

Beiträge zum Schwerpunktthema

Zukunftsvorstellungen aus der Vergangenheit gehen Christian Vater und Eckhard Geitz auf den Grund und analysieren in ihrem Beitrag, wie in früheren Berichten der Enquete-Kommissionen des Deutschen Bundestags über künstliche Intelligenz und Digitalisierung gedacht wurde, welche Schwerpunkte gesetzt wurden und wie diese ausgestaltet wurden. Heutige Debatten zu Technik- und KI-Zukünften binden sie an Technikvergangenheiten an, stellen Querbezüge über mehrere Jahrzehnte her und arbeiten Bedeutungen früherer Formatierungen heutiger Konzeptverständnisse heraus. Wie KI heute angelegt und ausgerichtet wird, ist oft in solchen Kartierungen und Präsentationsformen der Enquete-Kommissionen begründet, die vor vielen Jahren bereits Vorschläge zum Umgang mit künstlicher Intelligenz im politischen Raum gemacht haben. Die Ausprägung heutiger Debatten und heutige Entscheidungen gehen somit mitunter auf Prägun-

gschneiderte KI-Systeme bis dato genuin menschliche Rollen einnehmen. Der regelmäßige Umgang mit entsprechenden künstlichen Agenten lässt den Schluss zu, dass die Anwender*innen (und Gestalter*innen) diesen Artefakten – zumindest implizit – kontextuelle Sinnverständnisse zuschreiben. In die Zukunft gedacht, verschwimmt die bislang kulturell bedeutsame Grenze zwischen dem Menschlichen und dem Künstlichen in der allgemeinen Wahrnehmung oder sie wird durch eskalierende Turing-Tests immer wieder nachzuschärfen sein. Der Autor schlägt daher zur Klärung dessen, was wir von ‚sprechender Technik‘ erwarten dürfen, einen hermeneutischen Ansatz des ‚Verstehens des Verstehens von Technik‘ vor, der die Gestaltung, Wahrnehmung und Praxis künstlicher Kommunikation systematisch untersuchen soll.

Bernd Beckert behandelt das Thema vertrauenswürdiger KI. Es gibt bereits eine ganze Reihe praktischer Leitfäden für die Implementation vertrauenswürdiger KI-Anwendungen. Bei der Betrachtung von Umsetzungsprojekten zeigt sich jedoch durchweg ein noch geringer Konkretisierungsgrad. Was können die Gründe für das Umsetzungsdefizit sein? Bernd Beckert führt drei Gründe an: time to market-Überlegungen seitens der Unternehmen, Unklarheit darüber, welche Aspekte des Konzepts der vertrauenswürdigen KI bei welchen Anwendungen überhaupt relevant sind sowie die Tatsache, dass die Umsetzung von KI-Projekten komplexer ist als die Umsetzung ‚normaler‘ Software-Projekte und dass deshalb spezifische Vorkehrungen notwendig sind.

Stefan Strauß widmet sich der kritischen KI-Kompetenz für die konstruktive Nutzung von KI. Ziel seines Beitrags ist es, Problembewusstsein und kritische KI-Kompetenz zu stärken. Gesellschaftlicher Diskurs sowie Forschung und Entwicklung zu den Risiken künstlicher Intelligenz sind aus seiner Sicht oft einseitig: Entweder mit Fokus auf wichtigen, aber praxisfernen ethischen Aspekten, oder auf technokratischen Ansätzen, um gesellschaftliche Probleme nur mit Technologie zu lösen.

Es braucht jedoch auch praktikable, problemorientierte Perspektiven zur Analyse KI-bedingter Risiken. Im Kern sind Diskrepanzen zwischen Systemverhalten und Nutzungspraktiken in KI-automatisierten Anwendungskontexten zentrale Auslöser für Bias und gesellschaftliche Risiken. Stefan Strauß formuliert in dem Zusammenhang ein zentrales Meta-Risiko von KI-Systemen: Verzerrungen (automation bias). Dafür stellt er einen analytischen Rahmen vor mit Indikatoren zur Erkennung von Verzerrungen in KI-Anwendungen.

Eine besondere Aufgabe in der Gestaltung von IT-Systemen kommt jenen Führungskräften zu, die KI-Implementierungen und Einführungen verantworten und aktiv steuern. Ihr Einsatz in der Vermittlung von Zielklarheit und Erwartbarem inklusive der erwarteten Veränderungen sind nach Marie Jung und Jörg von Garrel wesentliche Einflussfaktoren für die Vertrauensbildung und die Unterstützung von Handlungsakzeptanz bei den Mitarbeitenden. Eine in diesem Sinne personalfreundliche Implementierung von KI-Systemen ist der Schwerpunkt ihres Beitrags, in dem sie anhand eines in ihrer Forschung entwickelten Modells aufzeigen, wie Vertrauensbildung und Akzeptanz bereits in frühen Entwicklungs- und Einführungsphasen angelegt und gefördert werden und somit zu einer erhöhten Handlungsakzeptanz beitragen können.

Angesichts des komplexen und unübersichtlichen Rechtsrahmens für Datenschutz, Dateneigentum und Urheberrecht fragt sich Thomas Wilmer, ob rechtssichere Rahmenbedingungen für neue KI-Entwicklungen weiterhin durch klassische, politisch veranlasste Regulierungsansätze formuliert werden sollen oder ob hier niederschwellige Verfahrensweisen womöglich besser geeignet sind. Dabei tangieren die spezifischen Datenbedarfe und -angebote vieler anwendungsorientierter KI-Entwicklungen direkt oder zumindest indirekt auch persönliche Interessen Einzelner und sind insofern von rechtlicher Relevanz. Dies betrifft insbesondere die Bereiche der Arbeitswelt und der medizinischen Forschung und Versorgung. In diesen sensiblen Bereichen sind Transparenzgebote und Grenzen für den angemessenen Einsatz von KI-Systemen durch geeignete Regulierungen zu formulieren. Entsprechende Maßnahmen sollen der Datensouveränität der Betroffenen und der Rechtssicherheit der Systementwickler und -betreiber dienen. Diese Aufgabe ist angesichts der unterschiedlichen Geschwindigkeiten heutiger technischer und legislativer Entwicklungen und der Heterogenität der Einsatzgebiete von KI eine große Herausforderung für die Verantwortlichen. Der Autor schlägt daher bestimmte vertragliche bzw. Control by Design-Lösungen zur Schließung von Regelungslücken im Bereich der KI vor: Eine pragmatische Lösung für zulässige Datenauswertungen und -übertragungen durch KI-Systeme wäre die Formulierung privatrechtlicher Klauseln in entsprechenden Vereinbarungen mit den Beteiligten (z. B. in Form von AGBs), die einen Rahmen für den Verkehr sowohl personenbezogener als auch anonymer Daten setzen. Diese Lösung könnte gegebenenfalls auch eine (finanzielle) Anreizregelung für die Betroffenen beinhalten. Ein ergänzender, vorgelagerter Ansatz greift die Idee des Privacy by Design der Europäischen Datenschutz-

Grundverordnung auf. Im Unterschied dazu soll der hier vorgeschlagene Control by Design-Ansatz die Souveränität auch über anonymisierte Daten bis hin zu entsprechenden Haftungsfragen regeln. Konkret könnte dies z. B. durch transparente und anerkannte Instrumente der informierten Einwilligung (Cookies etc.) unterstützt werden.

KI hat ein erhebliches Nutzungspotenzial in Forschung und Entwicklung – insbesondere auch in den Lebenswissenschaften und in der Biomedizin. Entsprechende Anwendungsbereiche zeigen sich in Medikamentenentwicklung, klinischer Praxis und vorgelagerter humangenetischer Forschung. Gleichwohl stehen diesen Potenzialen auch zahlreiche Herausforderungen entgegen, die besondere Handlungsnotwendigkeiten nach sich ziehen. Der Frage, wie die Entwicklung von KI mit der sich ebenfalls rasch entwickelnden Humangenomik konvergiert und welche praktische Bedeutung diese Konvergenz für die biomedizinische Praxis hat, haben sich Reinhard Heil und sechs weitere Ko-Autor*innen gewidmet. Auf Basis einer umfangreichen Literaturanalyse untersucht das Team, inwieweit der korrelative Zugang von KI zu neuem biomedizinischem Wissen dem Bedarf an begründbarem Wissen z. B. für die Diagnose und Therapie von Krankheiten gerecht wird und welche Fragen sich hieraus für Handlungsebene ergeben. Die Autor*innen heben dabei u. a. folgende ungelöste Probleme einer KI-unterstützten Forschung hervor: fragliche Qualität des Wissens und dessen Interpretation sowie dessen Erklärbarkeit, Evaluierung und Anerkennung. Auf anderer Ebene kommen noch ungeklärte Folgen von KI-assistierter Forschung für Wissenschaft, Wirtschaft und Gesellschaft dazu – auch hinsichtlich ethischer Herausforderungen sowie Fragen der adäquaten Gestaltung entsprechender Forschungs- und Dateninfrastrukturen und damit verknüpfter Regulierungsnotwendigkeiten (s. o.). Angesichts dieser Gemengelage empfehlen Heil et al. vertiefende Analysen, breite gesellschaftliche Dialoge zu kritischen Einzelproblemen sowie explorative Politikansätze, um einschlägige Forschungsprogramme und Regulierungen in allgemein wünschbarer Weise zu verbessern.

KI-Anwendungen haben bereits heute große Auswirkungen auf Gesellschaft, Wirtschaft und Wissenschaft und dies wird noch zunehmen. KI-Anwendungen richtig einzuschätzen und verantwortlich zu gestalten ist eine wichtige Aufgabe. Dieses Themenheft soll dazu einen Beitrag leisten.

Literatur

- Allen, Frederick (2001): The myth of artificial intelligence. In: American Heritage 52 (1). Online verfügbar unter <https://www.americanheritage.com/myth-artificial-intelligence>, zuletzt geprüft am 04.11.2021.
- BMBF – Bundesministerium für Bildung und Forschung (2021): Nationale Strategie für Künstliche Intelligenz. Online verfügbar unter <https://www.ki-strategie-deutschland.de>, zuletzt geprüft am 04.11.2021.
- Collingridge, David (1982): The Social Control of Technology. London: Pinter.
- Gethmann, Carl Friedrich et al. (2021): Künstliche Intelligenz in der Forschung. Neue Möglichkeiten und Herausforderungen für die Wissenschaft. Cham: Springer Nature (in Druck).
- Kurzweil, Ray (2015): Menschheit 2.0. Die Singularität naht. Berlin: Lola Books.

Marr, Bernard (2018): The key definitions of artificial intelligence (AI) that explain its importance. In: Forbes, 04.02.2018. Online verfügbar unter <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#54d0f59a4f5d>, zuletzt geprüft am 04.11.2021.



PROF. DR. BERNHARD G. HUMM
ist Informatik-Professor an der Hochschule Darmstadt und geschäftsführender Direktor des Instituts für Angewandte Informatik Darmstadt (aiDa). Sein Forschungsschwerpunkt ist Angewandte KI. Er führt nationale und internationale KI-Forschungsprojekte im Industrie- und Hochschulumfeld durch und publiziert regelmäßig.



PROF. DR. JAN C. SCHMIDT
ist promovierter Physiker und habilitierter Philosoph. Seit 2008 ist er Professor für Wissenschafts- und Technikphilosophie an der Hochschule Darmstadt. Schmidt ist Mitglied verschiedener Beiräte und Kuratorien, etwa dem Transdisziplinaritätsbeirat der Schweizerischen Akademie der Wissenschaften oder dem Konvent der ev. Akademie Frankfurt.



DR. STEPHAN LINGNER
verantwortet seit vielen Jahren den Forschungsbereich Technology Assessment am Institut für qualifizierende Innovationsforschung und -beratung GmbH (IQIB) in Bad Neuenahr-Ahrweiler. Er ist zudem Mitglied im Koordinations-Team des Netzwerks TA (NTA) sowie im Wissenschaftlichen Beirat der Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis (TATuP).



PROF. DR. KARSTEN WENDLAND
ist Informatiker, Humanwissenschaftler und Technikfolgenabschätzer. An verschiedenen Universitäten und Hochschulen lehrt er in den Grenz- und Überlappgebieten von Technik- und Humanwissenschaften und forscht zu Sonderthemen im Feld Digitale Technologien und gesellschaftlicher Wandel am Institut für Technikfolgenabschätzung und Systemanalyse (ITAS).



Liebe Leserinnen und Leser,

danke, dass Sie auch im Jahr 2021 *TATuP – Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* mit Interesse begleitet haben. Im neuen Jahr blicken wir voraus auf Hefte mit den *TATuP*-Themen Technikfolgenabschätzung und Lehre, Energiesuffizienz sowie nukleare Endlager, dazu wie immer Forschungsbeiträge, Interviews, Buchrezensionen, Konferenzberichte und vieles mehr.

Mit diesem Ausblick wünschen Ihnen die *TATuP*-Redaktion und der oekom verlag ein gesundes Jahr 2022!

Dr. Ulrich Ufer,
Redaktionsleitung, ITAS/KIT

Dr. Ulrike Sehy,
Leitung Fachzeitschriften, oekom verlag

Foto: © s. mccutcheon /unsplash.com