

KI und Verantwortung

Fragestellung

Beim Einsatz von Künstlicher Intelligenz entstehen erhebliche Verantwortungs- und Haftungslücken. Wenn KI vernetzt arbeitet, wenn eine enge Mensch-Maschine-Kooperation stattfindet und wenn Entscheidungen an autonome Softwaresysteme delegiert werden, wird nach geltendem Recht nicht gehaftet. Diese Expertise greift das Thema der Verantwortungslücke auf und entwickelt Vorschläge, mit welchen Verantwortungskonzepten diese Lücke geschlossen werden kann.

Expert:innen



Prof. Dr. jur. Dr.h.c.mult. Gunther Teubner
Goethe-Universität Frankfurt am Main



Dr. Anna Beckers
Maastricht University

Ergebnisse

Der Verantwortungs- und Haftungslücke sollte mit drei Verantwortungsregimes begegnet werden: Die Delegation von Entscheidungen an autonome Softwaresysteme sollte mit einer Prinzipal-Agenten-Haftung verknüpft werden. Für die enge Mensch-Maschine-Kooperation sollte die Netzwerkhaftung angewandt werden. Bei der autonomen algorithmischen Vernetzung müsste die Haftung vergesellschaftet werden und Fonds eingerichtet werden.

Folgerungen

1. Bei der Forschung über KI sollte ein Ansatz gewählt werden, der die Technik-, Sozial- und Rechtswissenschaften verbindet. Methodisch wird ein interdisziplinärer Ansatz verfolgt, der Erkenntnisse aus der Informationstechnologie, den Sozialwissenschaften und dem Recht miteinander verknüpft. Informationswissenschaftliche Typisierungen des Maschinenverhaltens werden auf sozialwissenschaftliche Analysen sozialer Praktiken des Algorithmeinsatzes und deren Risiken bezogen. Passgenau dazu werden rechtliche Haftungsregimes entworfen.
2. KI sollte nicht einheitlich als Person oder Werkzeug verstanden werden. Die je spezifischen Risiken der neuen sozio-digitalen Institutionen – digitale Autonomie, Hybridität, Vernetzung – erfordern differenzierte Zuschreibungen des Rechtsstatus der Algorithmen und ebenso differenzierte Haftungsfolgen.
3. Wenn in dem nur scheinbar technischen Haftungsrecht über die Bekämpfung von Verantwortungsdefiziten, über Chancen und Gefahren der Personifizierung algorithmischer Prozesse, über die Sorgfaltsmaßstäbe für algorithmische Entscheidungen, über den Schutz von durch digitale Fehlentscheidungen gefährdeten individuellen Rechtspositionen, über das Ausmaß von Entschädigungsleistungen und über die notwendige Re-programmierung von risikobehafteten Algorithmen gestritten und entschieden wird, leistet das Haftungsrecht einen nicht unbeträchtlichen Teilbeitrag zu einer sich gerade entwickelnden Digitalverfassung.

Menschen- und Technikbilder in der KI-Diskussion

Fragestellung

Die Konstruktion von Künstlicher Intelligenz orientiert sich in unterschiedlicher Weise an der (manchmal auch nur vermuteten) menschlichen Denkfähigkeit und Denkungsart. Die »künstliche« Intelligenz wird analog zu der »natürlichen« verstanden. Zumeist gehen dabei bestimmte Bilder vom Menschen und seiner Intelligenz, aber auch der Technik ein. Vieles davon bleibt jedoch implizit – und gleichwohl nicht minder wirksam. Ziel der Expertise ist es:

- solche impliziten Bilder von Mensch und Technik zu explizieren und damit eine wissenschaftliche wie allgemeine Aushandlung darüber zu ermöglichen, was KI (nicht) ist, wozu wir sie effektiv brauchen und wie sie hierfür gestaltet sein soll,
- in der Ausdifferenzierung von Menschen-, Technik- und Weltbildern einerseits und wissenschaftlichen Erklärungen der Technik andererseits eine informierte und nüchterne Einschätzung der Leistungsfähigkeit von KI zu ermöglichen.

Expert:innen



Dr. Bruno Gransche
Karlsruher Institut für Technologie (KIT)



Prof. Dr. Arne Manzeschke
Evangelische Hochschule Nürnberg

Ergebnisse

Die Rede von Künstlicher Intelligenz basiert im Wesentlichen auf sprachlichen Metaphern und einer spezifischen Bildsprache, die das Unanschauliche des Phänomens veranschaulichen, die Funktionsweise über Analogien und Sprachbilder vermitteln soll. Dabei wird erkennbar, dass KI »sehr bindungsfreudig« in die unterschiedlichsten Richtungen unserer Kultur reagiert. D. h. ihre Darstellung geht mit (Sprach-)Bildern aus der Tier-, Menschen-, Pflanzen- und Götterwelt einher. Damit steigt insgesamt der Aufmerksamkeitswert des Phänomens KI, was bewusst strategisch genutzt wird; keineswegs wird dadurch jedoch eine höhere Erklärungskraft gewonnen. So klaffen Sach- und Debattenlage bei KI besonders und in problematischer Weise auseinander. In mancher Hinsicht gleicht das Phänomen KI auch einem Fetisch, einem selbst geschaffenen Artefakt, dem dann Fähigkeiten jenseits der menschlichen Erfindungskraft zugewiesen werden, was Anlass zu Idolatrie einerseits oder Ikonoklasmus andererseits führt.

Nicht zuletzt hat die unspezifische Übertragung von menschlichen Fähigkeiten oder Tätigkeiten auf eine KI eine Anthropomorphisierung zur Folge, die ein klares Urteil über ihre Leistungen eher trübt und auch Verwirrung in den Aussagen über menschliche Fähigkeiten stiftet.

Folgerungen

- Ein bewusster und sorgfältiger Umgang mit Sprache im allgemeinen und Metaphorik in der Rede über KI im Besonderen. Hierbei geht es nicht um sprachpolizeiliche Regelungen, sondern um eine sprachliche Präzision, die in jeder anderen wissenschaftlichen Domäne *conditio sine qua non* ist – und auch hier gelten sollte.
- Von dort aus könnten wichtige Impulse für den Alltagssprachlichen Bereich ausgehen, was für einen breiten informierten Diskurs über die Gestaltung von KI unerlässlich ist und einer allzu großen Spreizung in apokalyptische und utopische Positionen entgegenwirkt.
- Die Zielbestimmung, wofür wir Menschen KI einsetzen wollten und welchen breiteren soziotechnischen Wandel wir damit vorantreiben wollen, müsste kontinuierlich neben der Weiterentwicklung der Technik verhandelt und konsequent umgesetzt werden.

Dynamiken der KI Entwicklung

Fragestellung

KI ist als **Promising Technology** zu verstehen. Insbesondere der gegenwärtige Hype um die KI wird von einem Technologieversprechen getrieben, wonach KI das Potenzial für die Realisation kommerzieller und industriell-ökonomischer Gewinne und vor allem weitreichende gesellschaftliche Visionen bietet. Insofern gilt KI als „Schlüsseltechnologie“ für den wirtschaftlichen Erfolg der deutschen Industrien auf dem Weltmarkt und einen wünschenswerten gesellschaftlichen und auch den notwendigen ökologischen Wandel. Basis hierfür ist ein technologischer Entwicklungsschub, besonders die verschiedensten Methoden des Machine Learnings und Deep Learnings, die zweifellos neue Anwendungspotenziale für die KI eröffnen.

Expert:innen



Prof. Dr. i.R. Hartmut Hirsch-Kreinsen
Technische Universität Dortmund

Ergebnisse

Mit dem besonderen Fokus auf die KI-Dynamik in Deutschland lassen sich zentrale Erkenntnisse der Expertise in folgenden Punkten bündeln:

1. Vorangetrieben wird das Technologieversprechen von einer einflussreichen KI-Community, es weckt nicht zuletzt in der öffentlichen Debatte weitreichende Zukunftserwartungen, begründet im wirtschaftlichen Bereich hohen Investitionen und führt zu vielfältigen innovationspolitischen Fördermaßnahmen.
2. Insgesamt ist die Entwicklung der KI bis heute von ungelösten anwendungspraktischen aber auch grundlegenden Grenzen gekennzeichnet. Als eine Zukunftsperspektive kann daher eine Entwicklung angesehen werden, die als „Veralltäglichung“ der KI zu fassen ist. Zugleich aber weist der alte Traum einer Artificial General Intelligence, d.h. die Weiterentwicklung der sog. starken KI in Richtung einer intelligenten Maschine für Viele eine überaus hohe Attraktivität auf.
3. Der Impact des Technologieversprechens auf Innovations- und Gesellschaftspolitik setzt eine rhetorische Architektur voraus, die überzeugend ist und eine breit gestreute Anschlussfähigkeit ermöglicht. Ihr Kern ist die Mehrdeutigkeit des Begriffs der KI. Damit kann es mit unterschiedlichsten Orientierungen und Interessen verknüpft werden und es eröffnet einen weiten Interpretationsspielraum. Letztlich aber erhält das Technologieversprechen seine Überzeugungskraft aus dem in der Menschheitsgeschichte tief verankerten Mythos von einer intelligenten Maschine, die der menschlichen Intelligenz ebenbürtig, wenn nicht gar überlegen ist.

Folgerungen

Voraussetzung und Folge der KI-Dynamik ist ein KI-spezifischer Innovationsmodus, der auf den Wandel des Innovationssystems drängt und der die Innovationspolitik vor große Herausforderungen stellt. Zu nennen sind hier besonders:

1. Transdisziplinarität und Wissenstransfer zwischen bislang voneinander disziplinär und sektoral abgeschotteten Bereichen und die Etablierung eines effektiven Wissenstransfers zwischen diesen.
2. Voraussetzung hierfür sind agile Innovationsprozesse, dynamischer Kooperationsprozesse und eine Offenheit gegenüber verschiedensten Akteure, d.h. die Schaffung sog. offener Innovation Ecosystems.
3. Dies impliziert ein erweitertes und nicht-technisches Verständnis von Innovationen und den Einbezug sozialer, rechtlicher und ethischer Kriterien in die Entwicklungsdynamik der KI.

Machine learning, political participation and the transformations of democratic practice

Fragestellung

The expertise addresses the epistemic interface between artificial intelligence technologies and democracy, particularly with regard to political participation.

Expert:innen



Prof. Dr. Jeanette Hofmann
Wissenschaftszentrum Berlin für
Sozialforschung (WZB)



Dr. Clara Iglesias-Keller
Wissenschaftszentrum Berlin für
Sozialforschung (WZB)

Ergebnisse

Democracy research often takes AI as a threat – be it because its technical nature is assumed to restrict rights, or because automated decisions could replace democratic decision making. While raising important concerns, these approaches overlook the malleability of both, AI and democracy and their relationship.

Democratic practice implicates constant (re)negotiations of rights and institutions, entailing the re-shaping of the scope of guarantees and deliberation processes. Technology is one of many factors that contribute to this process and transformations of democratic practice cannot be addressed solely to it.

Also, machine learning technologies themselves are far from reaching their maturity. Ongoing academic and policy discourses, as well as regulatory initiatives, have great potential to influence future forms of these technologies, which raises the stakes for current AI research.

The openness of the interface of AI and democracy is well illustrated by three critical perspectives that hold special importance for participation in public discourse:

1. algorithmic biases are considered as threats to rights, which ultimately affect citizens conditions for political participation. However, for systematic reasons we should also account for the possibility of AI mitigating bias- not only in automated processes, but also in humans led decision making processes within the democratic design;
2. the recent politicization of AI illustrates the potential of policy discourses and regulation to influence technical features of automated decision-making and to innovate with new forms of creating transparency and accountability;
3. regarding AI's epistemic dimension, political science ponders whether AI could compete with democratic decision-making – the risk being that political decisions no longer appear as the result of considering different possibilities of action, but as the execution of optimal solutions according to algorithmic models. Alternatively, algorithmic systems can be treated not as rivals, but as cooperation partners, a scenario where they can actually expand political discourse instead of restricting it, or making it superfluous.

Folgerungen

AI and democracy research could profit from a perspective that reflects the malleability of both our democratic institutions and of AI systems themselves.

Academic and policy discourses will be one of the sources shaping the final form of AI technologies. By reflecting the changing condition of democratic practice, current research can be productive in the sense of co-defining a shared understanding of new technologies and aim to set standards for their legitimate use.

Unlocking new possibilities for the realization of the democratic project through AI can only happen by approaching the epistemic interface of AI and democracy as a space open for meaning-making.

Kartographierung von Akteuren im aktuellen KI-Diskurs

Fragestellung

Der gesellschaftlich wirkmächtige Diskurs zu KI wird von vielfältigen Akteuren bestimmt, die nicht leicht zu unterscheiden sind hinsichtlich ihrer Motive und Interessen, der Reichweite ihrer Wissensproduktion und medialen Rezeption. Das erfordert eine Kartierung in der Fläche, erlaubt einen Überblick über das Feld der Akteure, wie eine Topografie der Machtverhältnisse und Hierarchien im Zusammenspiel der verschiedenen Akteursgruppen.

Nur so wird deutlich, wer das Verständnis von KI, die leitenden Bilder, Narrative und Zukunfts-Erwartungen an KI in welchem Maß prägt, aber auch, wer sich auf wen beziehen muss und wer die Deutungsmacht innehat. Ebenso wo Exklusionsdynamiken am Werk sind, wer nicht gehört wird im Diskurs oder gesellschaftlich exkludiert wird vom Zugang zu und der Weiterentwicklung von KI.

Expert:innen



Dr. Karin Hutflötz
Katholische Universität Eichstätt

Ergebnisse

KI erweist sich als ein von einseitigen Machtverhältnissen geprägtes Tätigkeits- und Diskursfeld. Populärwissenschaftliche Akteure ebenso wie große Tech-Unternehmen in Resonanz mit massenmedial wirkmächtigen KI-Visionen in Film und SF-Literatur bestimmen die Sprache, Leitbilder und Erwartungen an KI.

Diese prägen das gesellschaftliche Verständnis von KI maßgeblich, weniger die Akteure, die KI technisch entwickeln oder in wissenschaftlichen Beiträgen fundiert reflektieren. Diejenigen mit den stärksten, wenn auch oft fiktiven Versprechen und emotionalisierter Rhetorik werden maßgeblich gehört und rezipiert, haben die gesellschaftliche Deutungsmacht, nicht die, die den wissenschaftlichen Standards im Feld entsprechen. Die Kartierung der KI-Akteure zeigt auch, inwiefern blinde Flecken und strukturelle Ausschlüsse das Feld prägen – als Coded bias und Gender Gap.

Folgerungen

1. Es bedarf der Freilegung hegemonialer Strukturen und Deutungsabhängigkeiten hinsichtlich diverser Akteursgruppen, ihrer Motive und Interessen, damit diejenigen Akteure maßgeblich Gehör und Stimme finden in der KI-Entwicklung und-Debatten, die den wissenschaftlichen Standards im Feld hinreichend entsprechen.
2. Akteure der KI brauchen ein wissenschaftskommunikativ verbindliches Ethos, damit sie sich eingedenk ihrer fachlichen Expertise und gesellschaftlichen Verantwortung viel mehr dem Wirklichkeitsabgleich und der Kommunikation realer Entwicklungen verpflichten können, statt sich willfährig (Drittmittel-diktierten) Marketing-Logiken zu unterwerfen oder sich als Meta-Reflexion zu bloßen KI-Fantasmen zu verstehen.
3. KI benötigt die Sichtbarmachung von Exklusionsdynamiken, um sie einer gesellschaftlichen Auseinandersetzung zugänglich zu machen. Aufgrund dessen können Maßnahmen und Regelungen entwickelt werden, die dem gender gap und coded bias im Feld der KI gezielt entgegen wirken und die Voraussetzung schaffen für eine gerechtere, Demokratiekompatible und zukunftsfähige Gestaltung von KI.

KI und gesellschaftliche Machtverhältnisse

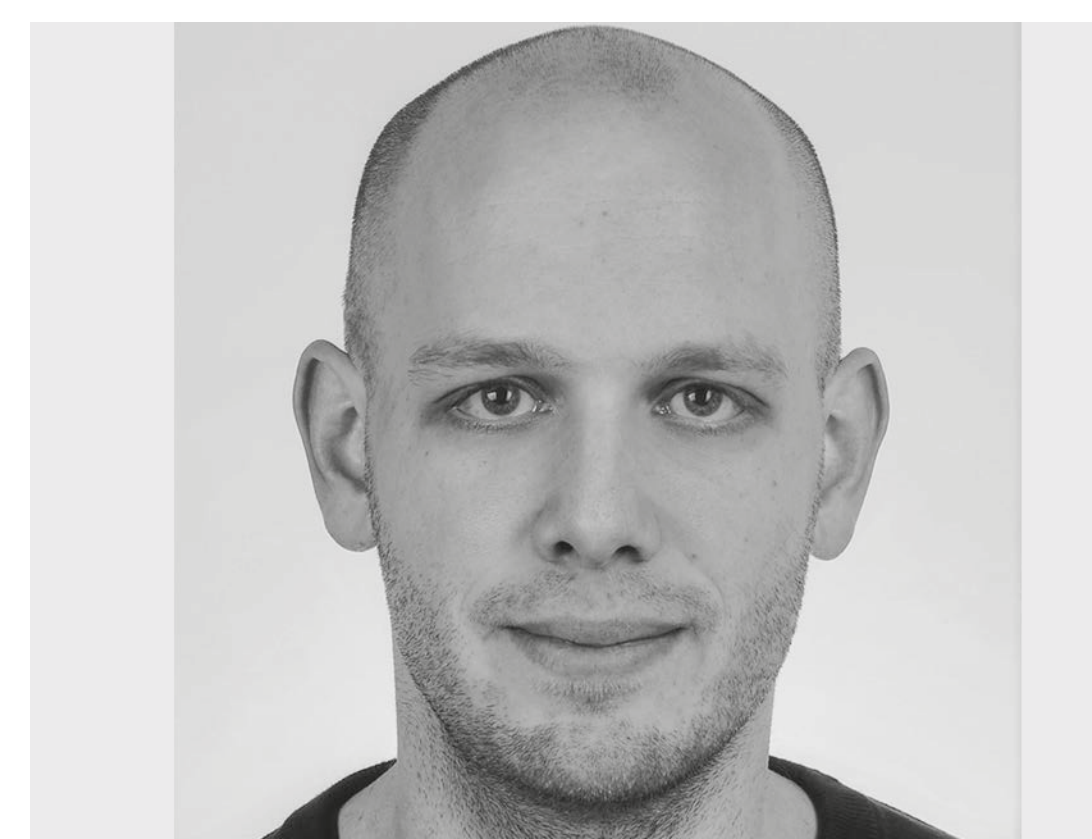
Fragestellung

- Die grundsätzliche Frage, ob und inwieweit Technik im Allgemeinen und KI im Besonderen ein gesellschaftsveränderndes Potenzial zukommt.
- Die konkreten Fragen, wie KI in gesellschaftliche Machtverhältnisse eingebettet ist, inwieweit KI ein Macht- und Herrschaftsinstrument darstellt und wie Künstliche Intelligenz Machtstrukturen und Herrschaftsverhältnisse in Politik, Wirtschaft und Gesellschaft transformiert.
- Die machtheoretisch interessanten Fragen, ob KI selbst Macht besitzt und (etwa als Akteur) Macht ausübt, ob Macht der KI etwa als „Sachzwang“ oder als „Nebenfolge“ eingeschrieben ist oder ob Menschen technisch vermittelt handeln und über den Besitz, die Beherrschung oder Anwendung von KI Macht ausüben.

Expert:innen



Prof. Dr. Peter Imbusch
Bergische Universität Wuppertal



Dr. Joris Steg
Bergische Universität Wuppertal



Lea-Sophie Natter, MA
Bergische Universität Wuppertal

Ergebnisse

- KI darf nicht allein als technologische Innovation, sondern muss als gesellschaftstheoretische Kategorie mit weitreichenden technologischen, ökonomisch-sozialen, politisch-kulturellen und habituellen Folgewirkungen verstanden und analysiert werden.
- Der Einsatz von KI wirkt in verschiedenen gesellschaftlichen Bereichen höchst spezifisch.
- KI ist (noch) kein eigenständiger und autonom agierender Akteur mit einem Willen zur Macht, aber KI ist ein Machtfaktor und eine gesellschaftliche Macht mit dem Potenzial, die Macht- und Herrschaftsverhältnisse einer Gesellschaft substantiell umzugestalten und zu revolutionieren.
- Es ist aber davon auszugehen, dass durch den Einsatz von KI gesellschaftliche Macht- und Herrschaftsverhältnisse stabilisiert, zentralisiert, monopolisiert und potenziert werden.
- Mit dem Einsatz von KI ist ein qualitativer Sprung in Bezug auf Machtpraxen und Herrschaftstechniken verbunden, der sich als problematisch für moderne demokratische Gesellschaften erweisen können.
- Es bedarf einer neuen Machtheuristik und Herrschaftstypologie, die dem gesellschaftlichen Stellenwert und der Bedeutung von Künstlicher Intelligenz Rechnung trägt.
- Mit dem Einsatz von KI sind Chancen und Risiken verbunden, die wahrgenommen und adressiert werden müssen, damit sich Technik nicht verselbstständigt.

Folgerungen

- Es bedarf einer theoretisch fundierten und empirisch gesättigten interdisziplinären Forschung über die Zusammenhänge, Wirkungsmechanismen und Konsequenzen von KI in einzelnen gesellschaftlichen Feldern und Bereichen.
- Es bedarf einer effektiven demokratischen Regulierung mit hohen ethischen Standards in Bezug auf die Entwicklung und Anwendung von KI.
- Es ist zu erwarten, dass die KI-induzierten Veränderungen von Macht- und Herrschaftsverhältnissen soziale Konflikte verursachen bzw. auf neuer Grundlage reproduzieren. Diese müssen politisch eingehegt und wissenschaftlich erforscht werden.

Gesellschaftliche Risiken und Gefahren von KI

Fragestellung

Unter dem Etikett Künstliche Intelligenz werden eine Reihe von unterschiedlichen neuen Technologien in verschiedenen gesellschaftlichen Handlungsfeldern implementiert. Dadurch verändern sich die Rahmenbedingungen sozialer Kommunikation und individueller Entscheidungen.

Ein durch unübersichtlichen und undurchsichtigen Technikeinsatz modularer Handlungsraum erzeugt problematische Nebeneffekte, die als solche von den Akteuren im Alltag nicht immer wahrgenommen werden. Voraussetzung für den großflächigen Einsatz von KI ist die Maschinenlesbarkeit des Menschen, seine Transformation in ein techno-soziales Hybridwesen durch zunehmende Digitalisierung und Datafizierung, die wiederum zur Voraussetzung für die Teilhabe am gesellschaftlichen Verkehr wird.

Expert:innen



PD Dr. Reinhard Kreissl
Vienna Centre for Societal Security
(VICESSE)



PhD Roger von Laufenberg
Vienna Centre for Societal Security
(VICESSE)

Ergebnisse

Je enger und präziser das praktische Einsatzgebiet einer auf KI basierenden Technologie ist, als desto erfolgreicher erweist sie sich. Diese Erfolge verleiten dazu, KI wirkmächtiger wahrzunehmen als sie eigentlich ist.

Denn außerhalb von den engen und präzisen Einsatzgebieten offenbaren sich schnell die Grenzen der Handlungsfähigkeit der KI. Vergleicht man nämlich die Wirkungsweise menschlicher Intelligenz mit den Fähigkeiten sogenannter „maschineller Intelligenz“ zeigt sich, dass die menschliche Fähigkeit, gezielt und sozial koordiniert praktische Handlungsprobleme in einer realen, dreidimensionalen und kulturell kodierten Umwelt zu bearbeiten jeder Form maschinenbasierter Technologie überlegen ist.

Folgerungen

1. KI benötigt ein besseres Verständnis situierter menschlicher Problemlösungs- und Informationsverarbeitungsprozesse unter besonderer Berücksichtigung der Dimension sozialer Kooperation und kognitiver Arbeitsteilung.
2. KI-Entwicklung erfordert in allen Phasen eine Berücksichtigung der Nutzer dieser Technologie. Dadurch lässt sich nicht nur der Gebrauchswert KI-basierter Lösungen erhöhen, die Einbeziehung der Nutzer erhöht auch die Transparenz der zu entwickelnden Lösungen und Systeme.
3. Ein Zertifizierungsverfahren durch kompetente öffentliche Stellen sollte für KI-Lösungen als Voraussetzung für ihren Einsatz verbindlich sein. Ebenso sollte, analog zu den Vorgaben der Datenschutzgrundverordnung, die Nutzer von KI-basierten Lösungen über deren Wirkungsweise umfassend informiert sein und die Möglichkeit eines nicht diskriminierenden opt-out haben.

Methoden der gesellschafts- und sozialwissenschaftlichen KI-Forschung

Fragestellung

Die Expertise fragt danach, welche Beiträge gesellschaftswissenschaftliche Methoden und Methodologien zur Erforschung künstlicher Intelligenz leisten können. Künstliche Intelligenz stellt die Forschung vor einige Herausforderungen. Viele KI-Systeme sind zu intransparent, proprietär verschlossen, umfangreich, kurz- oder schnelllebig, um traditionellen Forschungsansätzen zugänglich zu sein.

Dabei operieren solche Systeme zentral oder verteilt an gesellschaftlichen Schnittstellen von mehr und weniger unmittelbarer Bedeutung: als automatischer Filter unserer Handykamera, Medienkuratorin und Herausgeberin, administrative Mitarbeiterin, die über Zugang zu Leistungen und Anerkennung von Ansprüchen (mit)entscheidet, und vieles mehr. Noch vor der Anwendung stecken unzählige gesellschaftliche Voraussetzungen und -annahmen in KI-Technologien, die sich insbesondere aus einer sozialwissenschaftlichen Perspektive erschließen lassen.

Expert:innen



Prof. Dr. Jan-Hendrik Passoth
Europa-Universität Viadrina



Peter Kahlert, MA
Europa-Universität Viadrina

Ergebnisse

Wir verstehen Künstliche Intelligenz selbst als eine sozio-technische Schnittstelle. Die ihr zurechenbaren Technologien und Anwendungen sind vielfältig und jeweils fallspezifisch. Es macht einen eklatanten Unterschied, ob eine Bild-Erkennungssoftware, ein automatisierter Entscheidungsbaum oder ein digitaler Zwilling den Forschungsgegenstand darstellt. Allein der Anwendungsfall selbst hat einen entscheidenden Einfluss auf die Eignung und Ausgestaltung des Forschungsdesigns: ob man etwa Empfehlungsmodelle und -algorithmen einer Mediathek oder einer Social Media Plattform betrachtet.

Das Potential sozialwissenschaftlicher Methodologie für die Erforschung künstlicher Intelligenz erstreckt sich von Design, über Implementierung bis hin zu ihrer (Alltags-)Nutzung. Darüber hinaus haben Gesellschaftswissenschaften ein erprobtes Repertoire an Konzepten, Methoden und Theorien, die speziell für den Umgang mit komplexen Prozessen in Echt-Zeit, Intransparenz (z.B. der subjektiven Erfahrung) und der Gleichzeitigkeit von verschiedenen Sinnwelten und ihren Widersprüchen gebaut sind. Diese Herausforderungen finden sich beispielsweise in folgenden Gegenständen: Der Computer als intransparente Blackbox- Trends und Netzwerke, die sich fortlaufend und in Echt-Zeit verändern – das SEO-optimierte Thumbnail, in dem Sinnwelten von Nutzer*innen, Content-Anbieter*innen, und statistischer Berechnung zusammenlaufen.

Folgerungen

1. KI benennen: Genau differenzieren und benennen, um welche Art von Technologie und Anwendungsfeld es sich handelt. Forschung soll sich spezifisch daran ausrichten, anstatt Gemeinplätze zu verhandeln.
2. Digitale Themen auch digital beforschen: Digitale Methoden müssen fester Bestandteil der Methodenausbildung und -programmatisierung in den Humanwissenschaften sein. Eine solche Methodenausbildung muss den wissenschaftlichen Nachwuchs auch mit den Funktionsweisen ihrer Forschungswerkzeuge sowie -gegenstände vertraut machen.
3. Die Entwicklung von und Forschung zu Künstlicher Intelligenz braucht gesellschaftswissenschaftliche Perspektiven. KI-Systeme sind immer sozio-technische Komplexe aus gesellschaftlich produzierten Daten, Konzepten, Normen, Diskursen und Machtverhältnissen innerhalb einer belebten, dynamischen Umwelt (z.B. Recommender-Systeme auf Onlineplattformen), die sich nur im Wechselspiel mit sozialen Sinndimensionen hinreichend verstehen lassen.

KI als Kollegin? (KIK) – Repräsentative Beschäftigtenbefragung zu Künstlicher Intelligenz am Arbeitsplatz

Fragestellung

Künstlicher Intelligenz (KI) wird auch und vor allem in der Arbeitswelt eingesetzt werden. Wer (noch) Arbeit hat, wie Arbeit gestaltet ist und wie sehr sie „gute“ Arbeit ist – das hat immense Folgen für das gesellschaftliche Leben.

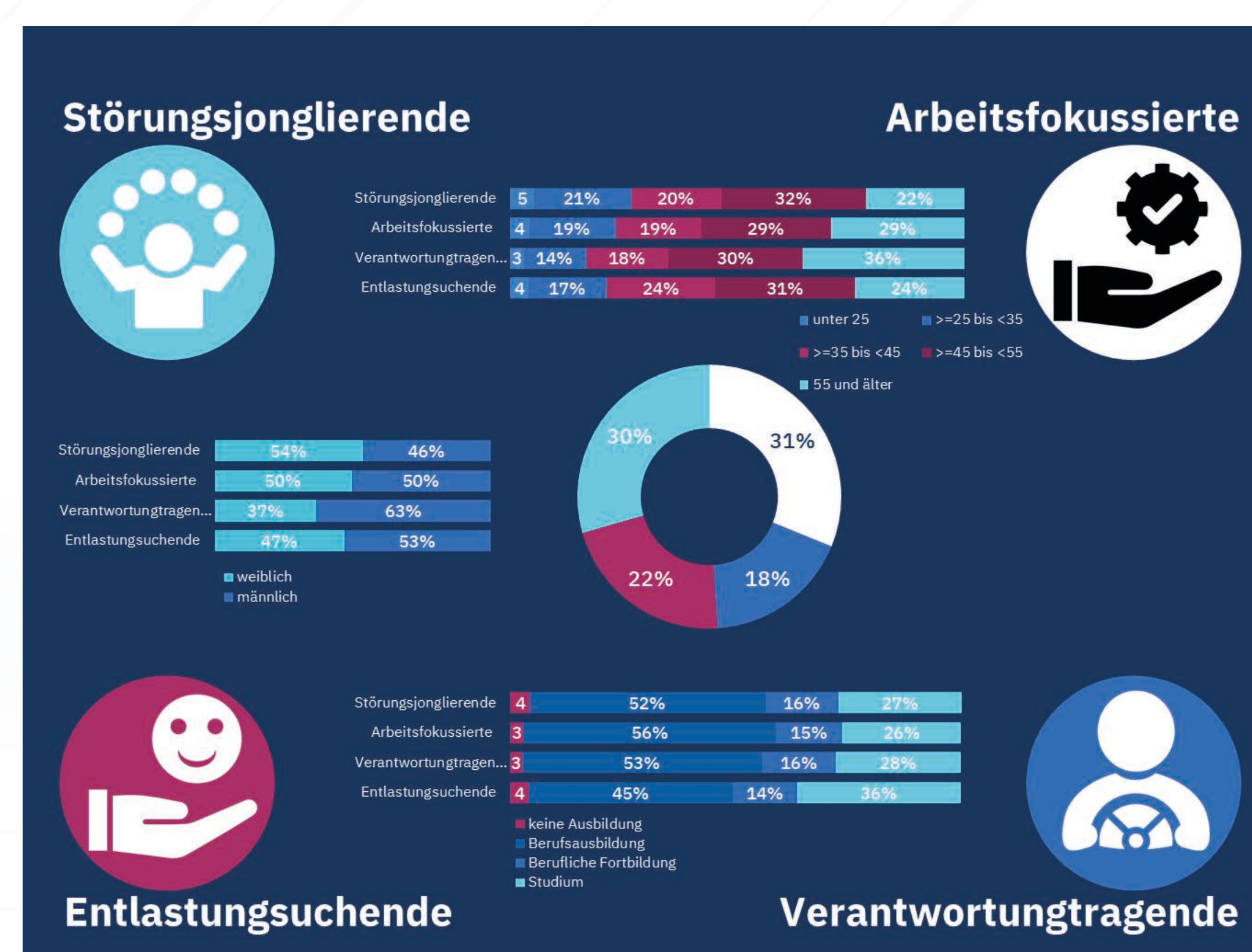
Über die Folgen von Künstlicher Intelligenz (KI) für Arbeit gibt es zwar einen breiten medialen und interdisziplinären wissenschaftlichen Diskurs. Im Vordergrund stehen aber ökonomische Prognosen zu Arbeitsmarkteffekten oder die Debatte um ethische Aspekte.

Was aber Beschäftigte über KI in ihrem direkten Arbeitsumfeld denken, ist in der Forschung bislang nur rudimentär erforscht. Diese Lücke füllt unsere 2022 durchgeführte Erhebung „KI als Kollegin? (KIK) – Repräsentative Beschäftigtenbefragung zu Künstlicher Intelligenz am Arbeitsplatz“, eine quantitative webbasierte Primärstudie mit N=2.018 Erwerbstätigen.

Expert:innen



Prof. Dr. Sabine Pfeiffer
FAU Erlangen-Nürnberg



Ergebnisse

Bei 11% der Beschäftigten ist KI schon am Arbeitsplatz angekommen, bei weiteren 12% ist dies in der konkreten Planung.

69% sehen positive Auswirkungen durch KI für das eigene Unternehmen, aber nur 24% sehen positive Auswirkungen für die Beschäftigten.

65% der Beschäftigten haben die Sorge, dass KI die Arbeit vieler Menschen ersetzen wird, aber nur 32% sorgen sich dabei um den eigenen Arbeitsplatz.

Nur 7% wünschen sich KI als Kollegin, 2% sehen KI gerne als Chefin – am ehesten besteht der Wunsch zu KI als Assistentin (22%).

Auf Basis von multivariaten Analysen konnten vier Einstellungstypen identifiziert werden: Die Störungsjonglierenden, die Arbeitsfokussierten, die Verantwortungstragenden und die Entlastungssuchenden. Für diese vier Typen zeigt sich trotz aller Unterschiede: die Ansichten zu KI als Kollegin sind stärker vom Arbeitskontext geprägt als von allgemeinen Vorstellungen zu Technik und KI. Wie Menschen KI am Arbeitsplatz beurteilen hängt also weitgehend ab von ihrer Arbeit, ihren Arbeitsaufgaben und den Folgen für andere bei Fehlentscheidungen der KI. Allgemeine Aussagen über KI generell sind daher nicht einfach auf die Anwendung von KI innerhalb der Arbeitswelt zu übertragen.

Beschäftigte wünschen sich sehr eindeutig eine starke Partizipation, wenn KI am Arbeitsplatz eingesetzt wird: 86% wollen vor der Einsatzentscheidung gefragt und 81% aktiv in den Gestaltungsprozess einbezogen werden.

73% wünschen sich bei dem Thema mehr Mitbestimmung. Beschäftigte vertrauen bei ökonomischen, sozialen und ethischen Fragen rund um den Einsatz von KI in der Arbeitswelt am meisten den Betriebsräten und am wenigsten der Politik.

Folgerungen

Bei der Gestaltung von KI am Arbeitsplatz sind drei Prämissen zentral:

- Welche Rolle KI am Arbeitsplatz spielen soll ist vor allem geprägt vom jeweiligen Arbeitskontext – daher muss KI konkret an die jeweilige Arbeitsaufgabe angepasst werden. Ein „one size/way fits all“-Ansatz ist der falsche Weg.
- Angst ist nicht – wie oft unterstellt – ein großes Problem. Gesunder Skepsis aber sollte mit guter Gestaltung der KI begegnet werden. Auch deshalb sind Beschäftigte aus dem späteren Anwendungskontext (nicht nur) in der Implementierungsphase von Anfang partizipativ einzubeziehen.
- Die Erhebung ist eine Momentaufnahme. Die Einstellungen zu KI am Arbeitsplatz werden sich im Laufe der Zeit verändern, wenn mehr Menschen mit KI erste Erfahrungen machen. Das macht weitergehende Erhebungen notwendig.

Folgen von KI für soziale Ungleichheit, Vielfalt und Inklusion

mit Fokus auf Alter, Geschlecht, Behinderung und Herkunft

Fragestellung

In der Expertise werden KI-gestützte Entwicklungen im Kontext gesellschaftlicher Differenzverhältnisse diskutiert. Im Rahmen dessen wird KI mit Blick auf marginalisierte Personengruppen untersucht. Im Mittelpunkt stehen die Differenzkategorien Geschlecht, Behinderung, Alter, „race“ und Herkunft. Hintergrund bildet die Dynamik der Entwicklungen im Bereich der informatischen Technologien und gesellschaftlichen Praktiken seit dem späten 20. Jahrhundert.

Es wandeln sich soziale, politische, kulturelle und ökonomische Verhältnisse. Diese Prozesse bergen Risiken, beinhalten aber auch transgressive und emanzipatorische Potenziale, die freilich bislang kaum umgesetzt werden. Vielmehr sind Verschärfungen oder wenigstens Reproduktionen traditioneller binärer Konstruktionen und Stereotype zu beobachten, die einmal mehr Ungleichheiten hervorbringen. Insofern besteht die gesellschaftliche Herausforderung für diese Expertise darin ein differenziertes Verständnis von KI-Verfahren in ihrer komplexen Einbettung in gesellschaftliche und kulturelle Wirklichkeiten zu entfalten. Dies geschieht auf zweifache Weise: Geschlecht, Behinderung, Alter, „race“ und Herkunft werden als Differenzkategorien ins Visier genommen. Außerdem werden technologische Logiken als Querschnittsthema verhandelt.

Ausgangspunkt bildet die These, dass diskriminierende Muster und Machtverhältnisse sich lediglich modernisieren, aber nicht unbedingt auflösen. Es entsteht die Gefahr neuer Formen von Ungleichheit entlang von Kriterien der technologischen Verwertbarkeit von Vielfalt und Differenz. Gleichwohl wird KI-Technik als nicht determiniert verstanden.

Expert:innen



Dr. Heike Raab
Independent Scholar

Ergebnisse

Die Expertise ist von der Annahme getragen, dass KI das Soziale verändert. Denn, künstliche Intelligenz steht an der Schwelle zum Massenmarkt und zur Alltagstechnologie. Insofern sollten verstärkt sozial- und kulturwissenschaftliche Analysen zur allgemeinen gesellschaftlichen Bedeutung von KI-Technologien entstehen und sich auf Aspekte von Differenz, Diversität und Ungleichheit beziehen. Beispielsweise bleiben Gründe für die Entwicklung KI-gestützter Instrumentarien, in Anbetracht einer nach wie vor hohen „technologischen Unreife“ vieler Produkte, häufig vernachlässigt. Die gesellschaftlichen Voraussetzungen der Konstruktions- und Aushandlungsprozesse zu Künstlicher Intelligenz geraten aus dem Blick. KI verändert zudem nicht zwangsläufig gesellschaftliche Differenzverhältnisse positiv, sondern ruft ebenfalls Teilungsmechanismen hervor.

Insgesamt ist davon auszugehen, dass die besprochenen Personengruppen keine homogene Ganzheit sind, sondern in sich ein jeweils heterogenes Gebilde darstellen. Dabei verdeutlicht sich, dass Technik (KI) immer nur eine Möglichkeit darstellt, gesellschaftliche Differenzverhältnisse zu artikulieren. Aus diesem Grund wird in der Expertise dafür plädiert soziale Ungleichheit sublimierter zu konturieren. Neue Formen sozialer Ungleichheit entstehen entlang von Ressourcen, Zugang zur Teilhabe sowie Gestaltung von KI-Produkten oder entlang der Problematik KI-kompatibler Körper. Letzteres bezieht sich auf die Frage, welche Körper relevante Körper im Bereich der KI-Entwicklung sind. Aus diesem Grund ist es wichtig KI-Technologien nicht isoliert zu begutachten, vielmehr ist der jeweilige Kontext, in der KI-Produkte hergestellt werden, zu berücksichtigen. Die größte Herausforderung scheint zu sein, zu erkennen, welchen Vergesellschaftungsmodus KI generiert, bzw. welche sozialen Praktiken und Aneignungsprozesse hervorgerufen werden.

KI führt tendenziell zu neuen Formen der Standardisierung der Umwelt, des Wissens und von Praktiken. Dies begründet sich zum einen darin, dass KI-Technologie das gesamte Leben in eine mathematische Logik übersetzt, um Big Data zu konstruieren. Jene Daten und Phänomene werden wiederum entlang der üblichen, binären Differenzkategorien eingeteilt. Daran gekoppelt ist ein Glaube an Objektivität und eine Priorisierung technizistischer Lösungen.

Folgerungen

Unter inklusivem, diversitätssensiblen Produkt- und Anwendungsdesign von KI ist nicht Assimilation an binär-ableistische Normen zu verstehen. Bei Fragen zu den Gestaltungsmöglichkeiten sollten aus diesem Grund benachteiligte Personengruppen für die Umsetzung von Universal Design einbezogen werden - von der Produktentwicklung bis hin zur Nutzer*innenperspektive. Insofern wäre vertiefend zu behandeln, wie Software und Informationstechnologien inklusiv gestaltet werden können. Wie können Produkte und Inhalte der Informatik, z.B. Roboter, Textverarbeitung und Modellierung differenzsensibel gestaltet werden, ohne Binaritäten festzuschreiben? Insbesondere wären die Teilhabe und Partizipationsprozesse in diesem Geschehen zu erforschen. Gestaltungsmöglichkeiten von KI sind vorhanden, so vorherrschende Normen hinterfragt oder unterlaufen werden.

Es sollte nicht allein auf individuelle Chancen zur Teilhabe und Inklusion fokussiert werden, sondern eher auf materielle und strukturelle Ungleichheiten. Gleichzeitig sollte ein Verständnis von Inklusion und Barrierefreiheit in Frage gestellt, welches die Normen der bestehenden Umwelt affirmiert. Beispielsweise sollte nicht auf die technologische Überwindung von Alter oder Behinderung gezielt werden. Stattdessen sollte die Anerkennung von Alter und Behinderung im Vordergrund stehen. Insgesamt ist bei KI-Produkten nicht allein das isolierte technische Gerät zu erforschen, sondern vielmehr die Einbettung einer KI-Technologie in den jeweiligen soziokulturellen Kontext. Ungeachtet der allgemeinen Pluralisierungsthese in der Ungleichheits- und Sozialstrukturforschung, in der davon ausgegangen wird, dass im Zeitalter des Neoliberalismus gesellschaftliche Differenzverhältnisse nicht mehr allein durch Ausschließung und Verwerfung reguliert würden, ist daher dringlichste Aufgabe zu erforschen, wie und auf welche Weise KI Differenz reguliert. Dabei ist sowohl die Makroebene des Sozialen, als auch die Intersektionalität von Macht- und Herrschaftsverhältnissen, mit Blick auf die jeweilige marginalisierte Personengruppe zu erforschen, als auch der Vergleich zwischen marginalisierten Personengruppen nicht zu vernachlässigen.

Folgen von KI für Sozial- und Gesellschaftstheorien

Fragestellung

Die Expertise will dazu beitragen, die Folgen von KI für soziales Handeln zu verstehen.

Expert:innen



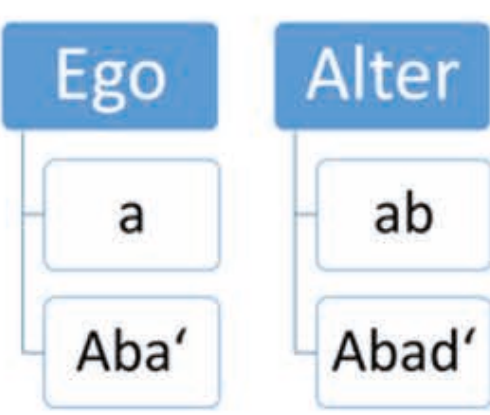
Prof. Dr. habil. Valentin Rauer
Türkisch-deutsche Universität Istanbul

Ergebnisse

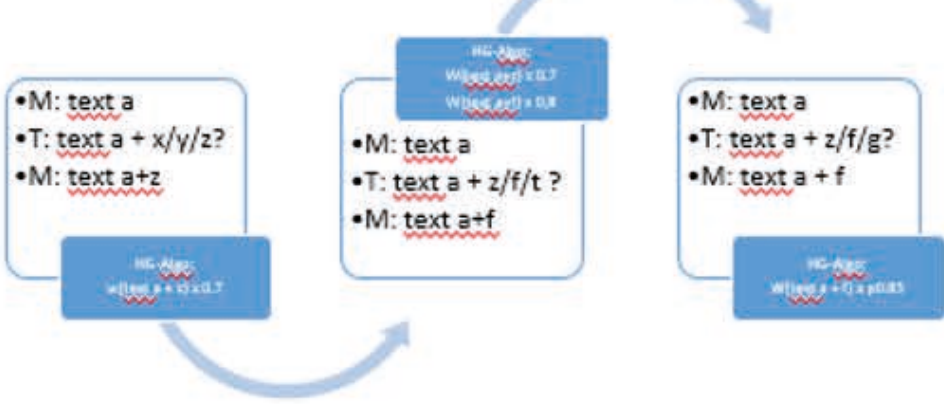
Der umfassende und alltägliche Einsatz von KI überführt klassische sozialer Interaktion in soziale „Transformaktion“

Weil: KI wird im Alltag in Interaktion mit realen Akteuren in ML-Systemen appliziert. Dieser Einsatz führt zu Optimierung des KI System und deren Transformation, was wiederum die sozialen Akteure in ihren Entscheidungsfindungen transformiert.

Statt Interaktion



„Transformaktion“



Folgen von KI (ML-NLP) für Theorien sozialen Handelns:
Neuverteilung von ‚KI-‘ verteilten Handlungsträgerschaften

	Interaktion	Transformaktion
Situationsgrenze	Wechselseitig aushandelbar	Unbestimmbar
Akteuresgrenzen	Indexikal begrenzt	Indexikal entgrenzt
Zeitgrenzen	Symmetrische Chronoferenz	Asymmetrische Chronoferenz
Transparenz	hoch	niedrig

Folgerungen

1. Die Intentionen und sozialen Sinne werden interaktiv generiert, sie entstehen nicht sui generis. Also müssen ML-Systeme, die sich in diese Interaktionsdynamik transformativ einbetten in den Fokus der sozialwissenschaftlichen und interdisziplinären Aufmerksamkeit gerückt werden.
2. Zu betonen, dass nur der Mensch die Verantwortung trägt und die letzte Entscheidung fällt ist leider einem solchen fehlleitenden dichotomen handlungstheoretischen Modell geschuldet. Transformaktionsketten sollten nach Verantwortungs- und Handlungszonen suchen, die sich über menschliche und nicht-menschliche Sinngenerierungssystem verteilen.
3. Problem der Datenqualität: ML-Systeme werden an Trainingsdaten optimiert, die zu groß sind, dass sie noch analog von menschlichen Akteuren überprüft und kontrolliert werden können. Die Grenzbestimmungen und Qualitätsfrage von Trainingsdaten sollte besser sozialwissenschaftlich verstanden werden.

Maschinelles und menschliches Lernen

Fragestellung

Systeme mit „Künstliche Intelligenz“, seien es Sprachassistenten oder Roboter, durchdringen zunehmend alle Lebensbereiche. Vielfach werden sie als eigenständige Akteure angesehen, mit denen sich große Hoffnungen oder noch größere Befürchtungen verbinden. So oder so scheint die Übernahme vieler menschlicher Tätigkeiten durch die „lernenden Systeme“ unausweichlich, insbesondere in der Arbeitswelt.

Dieser Herausforderung stellt sich die Expertise mit den Fragen:

- Was ist „machine learning“?
- Welche Leistungen, Grenzen sowie (prinzipiellen) Schwächen haben Systeme, die auf „machine learning“-Algorithmen beruhen?
- Welche Bezüge bestehen zu menschlichem oder generell biologischem Lernen?

Expert:innen



Dr. Irmhild Rogalla
Institut für praktische Interdisziplinarität

Ergebnisse

- „Machine Learning“ ist ein Oberbegriff für eine bestimmte Art von Algorithmen, die auf statistischen Optimierungsverfahren beruhen.
- Auf leistungsfähigen, vernetzten Systemen implementiert und unter Nutzung extrem großer Datenmengen können diese Algorithmen für Aufgaben eingesetzt werden, die mit klassischen Algorithmen nicht lösbar waren. Dazu gehören beispielsweise autonome Steuerungen, Spielstrategien sowie Sprach- und Bilderkennung. Gleichzeitig gilt „Machine learning“-Algorithmen eignen sich, wie alle anderen Algorithmen auch, nur für spezifische Aufgaben. Ihre Leistungsfähigkeit hängt zudem von sehr vielen Faktoren ab, zu denen neben Verfügbarkeit und Qualität der Daten wesentlich auch die Kenntnisse und Fähigkeiten der jeweiligen Programmierer gehören, meist als „Data Analysts“ oder „Data Scientists“ bezeichnet.
- Die Benennung dieser algorithmischen Verfahren als „learning“ geschah von Anfang an aus Gründen der besseren Vermarktbarkeit. Es handelt sich dabei maximal um sprachliche Analogien. Keine KI-System erfüllt die für menschliches oder auch nur biologisches Lernen konstitutiven Bedingungen.

Folgerungen

- Für einen adäquaten Umgang mit KI-Systemen ist ihre realistische Einschätzung und Darstellung als leistungsfähige informationstechnische Systeme, die spezielle Funktionen ausführen können, erforderlich. Nur so ist eine angemessene Diskussion und Aufklärung über ihre Grenzen und Schwächen, aber auch über sinnvolle Einsatzgebiete möglich.
- Für einen solchen kritisch hinterfragenden Umgang mit KI-Systemen, auch hinsichtlich der hinter ihrer Entwicklung und Verbreitung stehenden wirtschaftlichen, gesellschaftlichen und politischen Interessen, ist eine deutlich bessere informationstechnische („digital literacy“) sowie mathematisch-statistische („statistical“ oder „data literacy“) Bildung erforderlich. Erst sie ermöglicht es, Gestaltungsoption zu sehen, wahrzunehmen oder eben auch begründet abzulehnen.
- Die informationstechnischen KI-Systeme erfordern erweiterte und neue Ansätze in der Soziotechnik damit letztendlich Mensch, Technik und Organisation erfolgreich zusammenwirken können. Dabei kommt es darauf an, die Informationstechnik als solche zu verstehen und zu behandeln, um Interaktion und Kooperation zwischen Menschen und technischen Systemen und letztlich die digitale Transformation zukunftsfest zu gestalten.

KI als gesellschaftliche Verheißung

Fragestellung

Die gesellschaftliche Herausforderung besteht darin, ein vorurteilsfreies und multiperspektivisches Verständnis heterogener Zukunftserzählungen über KI zu entwickeln.

Vor dem Hintergrund aktueller Krisendiagnosen, die individuelle, kollektive und planetarische Erschöpfungssyndrome als zentrale Herausforderung für „Zukunftsdesign“ adressieren, wurde nach der Herkunft und den Folgen von Bedeutungszuschreibungen an KI gefragt, die mit Heilsbotschaften aufgeladen sind. Verheißungserzählungen buchstabieren Erwartungshorizonte und Zukunftshoffnungen zwischen (alten) Technikversprechen und (neuer) Technikgläubigkeit immer wieder neu aus – womit permanent auch die gesellschaftliche Wahrnehmung von KI geprägt wird.

Verheißungsnarrative liegen in Form utopischer Fortschrittsgeschichten, Meta-Analysen, Hype-Zyklen, politisch-normativen Leitbildern, Visionen von KI-Schaffenden, Medienberichten oder auch zahlreichen fiktionalen Darstellungen vor und gehören zum gesellschaftlich vorrätigen Wissensraum über KI.

Expert:innen



Prof. Dr. Stefan Selke
Hochschule Furtwangen

Ergebnisse

Verheißungsnarrative sind weder Bedienungsanleitung noch Technikfolgenabschätzung, sondern spekulative Bedeutungszuschreibungen. Ihre Wirkung beruht auf der Erzeugung trostspendener Zukunftseuphorie. Vor diesem Hintergrund sind Verheißungsnarrative Fortsetzungsgeschichten zwischen Gegenwart und Zukunft – einerseits noch (teils empirisch und theoretisch fundierte) Gegenwartsdiagnose, andererseits bereits prognostische und (hoch)spekulative Zukunftsszenarien. Sie dienen als Katalysator für entlastende Zukunftseuphorie, die in säkularisierten, sozial erschöpften und sich transformierenden Gesellschaften als Trostersatz wirkt. Verheißungen über KI werden von stark unterschiedlichen Zukunftsnarrativen gerahmt:

- **Anpassungs-Narrative** betonen inkrementelle Entwicklungsschritte von KI innerhalb der (politisch gewollten) Reproduktion von Standardwelten im Kontext globaler Wettbewerbsarenen.
- **Quest-Narrative** präsentieren diesseitig-rekontextualisierte Heilsverkündungen in Zukunfts- und Deutungsarenen mit ausgeprägtem spekulativem Potenzial und hohen ökonomischen Gewinnerwartungen.
- **Aufbruchs-Narrative** betten die KI-Entwicklung in Modelle gesellschaftlicher Transformation ein, die oftmals gemeinwohlorientiert sind.
- **Dagegen-Narrative** basieren auf (latenter oder manifester) Modernisierungsverweigerung und stellen als Kontrastfolie zu intentional überbelichteten Verheißungsnarrativen Unheilsverkündungen und antizipierte Zerstörungspotenziale in den Mittelpunkt.

Verheißungen lassen sich auf drei Ebenen rekonstruieren. KI nimmt hierbei die Rolle als Werkzeug, Assistenz, Partner oder Über-Ich ein, womit jeweils unterschiedliche sozial-ethische Folgen und gesellschaftliche Implikationen verbunden sind:

- **Funktional-operative Ebene**, z.B. Effizienz- und Optimierungsversprechen, Akzeptanztests für kulturelle Disruptionen, Verheißungen als strategische Inszenierungen in kompetitiven Märkten etc.
- **Kognitiv-epistemologische Ebene**, z.B. intentional überhöhte Visionen künstlicher Wesen, KI im Kontext von Subjektsimulationen, Versprechen im Sinne von Heiligkeitsäquivalenten oder neoreligiösen Verkündungen.
- **Zivilisatorisch-transformative Ebene**, z.B. KI als soziale Innovation, Erlösungshoffnungen im Kontext techno-utopischer Machbarkeitsfantasien und Ideen des Redesigns bzw. Resets von Gesellschaft, Überbelichtung von KI im Zusammenhang mit zivilisatorischen Menschheitsexperimenten und Stilisierung von KI als „kybernetische Regierungsform“ oder „Retter des Planeten“.

Folgerungen

- **Perspektivwechsel und Reframing des Diskurses:** Gefordert wird erstens die systematische Berücksichtigung von Effekten des schleichenden Wandels (“shifting baselines”) unterhalb der Wahrnehmungsschwelle in lebensweltnahen Anwendungsfeldern von KI (z.B. Bildung, Erinnerung) jenseits der hypnotischen redundanten Rhetorik über technologische „Disruptionen“.
- **Zukunftsethik im Kontext agiler Wissenschaft:** Gefordert wird zweitens das Bezugsproblem von KI vermehrt aus der Sicht von Bürger*innen sowie deren Affekt- und Begehrenskonstellationen (Wollen und Wünschen) zu definieren. Der dominanten Beschäftigung mit dem „Wie“ („know-how“) von KI, sollte verstärkt die Beschäftigung mit sinnhaften Zwecken hinzugefügt, wobei Fragen nach dem „Warum“ („know-why“) im Mittelpunkt stehen. Hinzu kommen Vorschläge für Veränderungen und Gestaltungsoptionen („know-what“). Verheißungsfreies Zukunftsdesign kann daher nur im Rahmen von Aufbruchs-Narrativen stattfinden und nicht im Kontext von Anpassungs-Narrativen.
- **Transformative Szenarien und Betonung der diskursiven Dimension von KI:** Gefordert wird drittens die produktive Verbindung eines empirisch fundierten Krisenbewusstseins und utopisch gerahmter Zukunftseuphorie jenseits technikdeterministischer Verheißungsnarrative. Dazu braucht es die Steigerung der gesellschaftlichen Leitfähigkeit für sozial-utopische Gestaltungsrahmen (z.B. im Bildungssystem) sowie zunächst die Rehabilitation des utopischen Denkens. Grundlegend hierfür wäre die wert- und vorurteilsfreie Berücksichtigung fiktionaler Perspektiven auf KI, z.B. im Kontext künstlerischer Forschung oder Scientific-Fiction (statt Science-Fiction). Idealerweise entsteht auf diese Weise ein nivellierender Umgang mit Argumenten über das Neue zwischen Problematisierung und Visionierung. Verheißungsnarrative gehören daher genuin zum gesellschaftlichen Verständigungsprozesses hinzu und dienen der notwendigen Diskurserneuerung.

Formen und Möglichkeiten gesellschaftlicher Normierung von KI

Fragestellung

Die Expertise behandelt zwei Fragen:

1. Wie kann die Entwicklung und vor allem die Nutzung von KI normiert bzw. reguliert werden? Dabei werden ethische Leitlinien bzw. Ethikodizes, Selbstverpflichtungen sowie staatliche und supranationale Regulierung untersucht.
2. Welche normierende Wirkung bspw. auf das menschliche Handeln hat KI? Die Existenz und Nutzung von KI kann selbst normierende Wirkung bspw. in Hinblick auf den Sprachgebrauch oder Arbeitsabläufe haben; dies sollte in einem Überblick beschrieben werden.

Expert:innen



Prof. Dr. phil. habil. Karsten Weber
OTH Regensburg



Nadine Kleine, MA
OTH Regensburg

Ergebnisse

Entsprechend der zwei behandelten Fragen kann man die Ergebnisse differenzieren. Zum einen gibt es sehr viele Ansätze der Normierung bzw. Regulierung von KI, wobei Leitlinien und Selbstverpflichtungen in Hinsicht auf deren Bindungskraft und Wirksamkeit skeptisch anzusehen sind. Einzelstaatliche Maßnahmen wiederum sind mit dem Problem konfrontiert, dass die Entwicklung und Nutzung von KI globalisiert sind; die Reichweite und Durchsetzmöglichkeit einzelstaatlicher Maßnahmen sind also begrenzt.

Supranationale Regulierungsansätze wie bspw. der Artificial Intelligence Act der EU stellen ohne Zweifel den besten Weg dar, bringen aber ebenfalls Probleme mit sich; Dauer des Regulierungsverfahrens, Umsetzung in nationales Recht oder Unschärfen bei wichtigen Definitionen stellen nur einige dieser Probleme dar. Gleichzeitig ist der Prozess der Normierung durch KI kaum aufzuhalten – hierin unterscheidet sich KI allerdings nicht grundsätzlich von anderen Technologien.

Folgerungen

Auch wenn diese Frage nicht im Zentrum der Expertise stand, lassen sich hierzu Aussagen treffen:

1. Trotz der genannten Probleme bleibt Normierung und Regulierung unverzichtbar.
2. Beim Einsatz von KI sollte das Vorsorgeprinzip (engl.: precautionary principle) beachtet werden; in Hinblick auf mögliche negative Folgen müssen plausible Worst-Case-Szenarien berücksichtigt werden.
3. Angesichts des großen Potenzials nützlicher Anwendungen von KI-Systemen sollte die Forschung an solchen Systemen gefördert werden; dabei sollte im Sinne des Vorsorgeprinzips ELSA-Begleitforschung als zentraler Bestandteil entsprechender Projekte einbezogen werden.