

Andrea Reichenberger und Martin Ziegler gehen davon aus, dass die gesellschaftliche Wahrnehmung der KI ambivalent zwischen Stein der Weisen und Büchse der Pandora schwankt. Ihr Diskurs versucht KI rund um diese beiden Pole zeitlich bis in die Zukunft, technisch versus gesellschaftlich und fundiert versus spekulativ auszuleuchten. Sie plädieren für eine weise Regulierung.

Britta Schinzel hat den Sammelband *Künstliche Intelligenz – Die große Verheißung* rezensiert, dessen vier Herausgeber:innen sich aus dem Berliner philosophischen Arbeitskreis MoMo Berlin und der Denkwerkstatt Berlin kennen. Dementsprechend sind die meisten der 26 Beiträge aus philosophischer Sicht verfasst, während die technische Seite der KI vielleicht etwas zu kurz kommt.

Peter Brödner geht auf die KI-Narrative ein, die sich im Laufe von 50 Jahren gewandelt haben. Ein Charakteristikum besitzen: Sie produzieren sogenannte maschinelle Lernen in Form von KI, die Probleme mit sich bringt, was die Undurchschaubarkeit der Systeme betrifft.

Rainer Rehaks Beitrag ist ein Nachdruck aus Peter Verdegems Sammelband *AI for Everyone?*, der im vorigen Jahr bei University of Westminster Press erschienen ist, und deshalb von der Länge und Sprache her aus dem Rahmen fällt. Inhaltlich setzt sich der Autor mit dem Problem der irreführenden Begriffsbildung in der KI auseinander, wobei er seine Kritik durchaus konstruktiv wendet.

Claudia Müller-Birn setzt sich in ihrem Beitrag mit dem neuen Fachgebiet Data Science auseinander, in dem es um die Analyse großer Datenmengen geht und das mit KI große Überschneidungen aufweist. Dabei kommt es ihr auf die Verbindung mit qualitativen Methoden an. Sie diskutiert, wie eine solche kritisch-reflexive menschenzentrierte Datenpraxis in der akademischen Ausbildung verankert werden kann.

Klaus Heß unternimmt eine große Expedition durch die Gegenwart der KI, die in vielfältiger Weise längst unseren Alltag erreicht hat, und eine von einigen Protagonisten prognostizierten Zukunft ungeahnter technologischer Entfaltung. Der Microsoft-Chef hält KI für eine wirklich grundlegende Technologie und der Google-Chef findet sie bedeutender als Elektrizität und Feuer. Jürgen Schmidhuber versteigt sich gar zu der Annahme, dass KI die Milchstraße erobern wird.

Dorothea Kolossa hat eine Rezension über Stuart Russells Monographie *Human Compatible* verfasst. Der Autor beschwört ein existenzielles Risiko herauf, das von den rasanten Entwicklungen in der KI ausgeht, um dann Grundsätze herauszuarbeiten, wie Menschen die Kontrolle behalten können, selbst wenn die Maschinen eines Tages superintelligent werden sollten.

Der Autor betrachtet den aktuellen Stand der KI als interdisziplinär einerseits und in ihrer Replizität und der Wirtschaft andererseits mit den globalen Herausforderungen der heutigen Zeit und der Idee einer lebenswerten Welt für alle Menschen.

Wir hoffen, dass die Artikel auf großes Interesse stoßen und für alle Leserinnen und Leser neue Gesichtspunkte und erhellende Einsichten bieten.

erschienen in der FfF-Kommunikation,
herausgegeben von FfF e.V. - ISSN 0938-3476
www.fff.de

Anmerkungen

- 1 Melissa Heikkilä (2022) *Dutch scandal serves as a warning for Europe over risks of using algorithms*. *Politico*, 29. März 2022, <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>



Andrea Reichenberger und Martin Ziegler

KI aus interdisziplinärer Perspektive

Der gesellschaftliche Ruf von KI-Technologien war und ist ambivalent. Er bewegt sich seit jeher zwischen Enthusiasmus und Enttäuschung, Glorifizierung und Diskreditierung. Mehr denn je reicht das Meinungsbild vom informationsverarbeitenden Stein der Weisen bis zum Sargnagel der Menschheit; gepriesen als Allheilmittel für Wirtschaft und Wissenschaft, verdammt als Pandorabüchse zur Verklavung der Menschheit; als willfähriger Assistent (Siri/Echo, Tesla Autopilot, Gecko Carebot) ebenso wie als Mastermind der kommenden Roboterrevolution (Boston Dynamics' BigDog/Atlas).

Dieser Essay versucht, den Diskurs zwischen den und jenseits dieser Pole zu erweitern: in sechs Abschnitten, von historisch bis zukünftig, von technisch bis gesellschaftlich, von fundiert bis spekulativ. Weder blauäugige Promotion noch blinde Panik propagierend, gehen wir aus von der Erfahrung (bspw. aus Kernenergie), dass technische Möglichkeiten sich in der Regel nicht dauerhaft verhindern¹ lassen – und stattdessen in ihren potenziellen Konsequenzen erst frühzeitig zu erkennen und dann weise zu regulieren sind.

Der erste Abschnitt rekapituliert und problematisiert gängige Vorstellungen und Verwendungen des Begriffs KI. Der zweite Abschnitt gibt einen kurzen Abriss der bisherigen und erwarteten Entwicklung von KI aus technologischer Sicht. Aktuelle und bevorstehende Regulierungsbemühungen fasst der dritte Abschnitt zusammen. Der vierte Abschnitt verschiebt den Fokus auf absehbare Auswirkungen im gesellschaftlichen und ökonomischen Bereich. Der fünfte schlägt unkonventionelle Blickwinkel auf KI vor jenseits menschenzentrierter Werkzeuge; und der letzte Abschnitt spekuliert über potenzielle langfristige soziale Konsequenzen.

Was ist (Künstliche) Intelligenz?

Seit den 1960er-Jahren hat das gesellschaftliche Ansehen von KI mehrfach gewechselt zwischen Enthusiasmus und Ernüchterung und steht aktuell mal wieder hoch im Kurs. Aber was ist KI?

Die Bund-Länder-Initiative zur Förderung der Künstlichen Intelligenz in der Hochschulbildung versteht darunter „Technologien zur maschinellen und algorithmen-basierten Verarbeitung großer Datenmengen“ (Bundesanzeiger vom 24.02.2021) – und übergeht dabei *zwei* zentrale Paradigmenwechsel: Training statt Programmierung und Heuristiken statt Algorithmen. In der Tat basiert klassische Datenverarbeitung auf den Antipoden Hardware/Software (vgl. Harvard/von Neumann-Architektur). Diese Unterscheidung spiegelt sich wider beispielsweise in den Fachbereichen Elektrotechnik/Informatik oder in Firmen wie Intel/Microsoft. Maschinelles Lernen fügt dieser Unterscheidung eine neue dritte Komponente hinzu: *Training*. Und weil ein solches Training naturgemäß nie vollständig abgeschlossen ist, sind durch KI generierte Lösungen inhärent *heuristisch* – im Gegensatz zu jenen von wohlspezifizierten und analysierten garantiert korrekten Algorithmen. Warnungen vor einer *Algokratie* gelten daher doppelt für eine *Heuristikokratie*.²

Solche Verwirrungen und divergierende Einschätzungen mögen auch daher rühren, dass die technischen Hintergründe und genauen Funktionsweisen nicht nur Laien kaum verständlich sind: Der Online-Cartoon (Abbildung 1) bringt dies auf den Punkt, und aus gutem Grund bildet XAI (Explainable Artificial Intelligence) ein eigenes Forschungsgebiet innerhalb der KI.



Abbildung 1. Machine Learning, Quelle: <https://xkcd.com/1838>, CC BY-NC 2.5

So sind Debatten oft spürbar von Science Fiction beeinflusst, sowohl hinsichtlich utopischer Visionen (wie *Der Zweihundertjährige* von Asimov oder *Commander Data* in *Star Trek*) wie hinsichtlich dystopischer Warnungen (*Blade Runner*, *Terminator*). Und gängige Vorstellungen von KI erweisen sich als anthropozentrisch einseitig dominiert von *menschlicher* Intelligenz und deren Imitation – obgleich selbst die bloße Definition menschlicher Intelligenz bekanntlich Gegenstand zahlreicher Kontroversen ist, beispielsweise in Soziologie und Erziehungswissenschaft.

Zusammengefasst stellt der Begriff *KI* also eher eine Hülse dar für verschiedenste Projektionen denn ein tatsächliches Fachwort. Dies hat geschichtliche Gründe:

Theorie und Technik

Alan Turing³ und Konrad Zuse⁴ gelten nicht nur als Mitbegründer der Theoretischen Informatik, sondern auch der KI.⁵ 1951 implementierte der Princeton-Student Marvin Minsky mit SNARC (*Stochastic Neural Analog Reinforcement Calculator*) das erste neuronale Netz, als Gründer und langjähriges Aushängeschild der KI. Bis zu seinem 1986er Buch *Society of Mind*⁶ stand KI noch der Psychologie und Physiologie nahe.

Seitdem ist der Zugang mathematisch-abstrakter⁷: Dieser Siegeszug der *Neats* über die *Scruffies* bildet die qualitative Grundlage der aktuellen Erfolge – zeigte Wirkung aber erst nach dem sogenannten *zweiten KI Winter* ab ca. 1993: Vorher reichte die Rechenleistung nicht aus, um Ideen wie *Reinforcement Learning* über einen qualitativen Proof-of-Concept hinaus in praxisrelevanter Größenordnung zu realisieren.

Der derzeitige Höhenflug von KI ist nämlich primär ingenieurtechnischen und quantitativen Fortschritten geschuldet: *Deep Blue* konnte zum Sieg über Garry Kasparov (1997) doppelt so viele Spielzüge pro Sekunde analysieren wie bei seinem verlorenen Vorgängerspiel (1996); und Moores Gesetz exponentiell wachsender Rechenleistung setzte sich fort über *Big Data* bis zum Sieg (2016) von *AlphaGo* über Lee Sedol: Eine Materialschlacht basierend auf 1202 CPUs und 176 GPUs und 48 TPU ASICs sowie mit einer Datenbank von 30 Millionen Zügen im Hintergrund [doi:10.1038/nature16961], *ad-hoc* zusammengestellt für dieses PR-Ziel.

Aktuell sehen wir erste Schritte hin zu Quasi-Standardisierung und zu schlüsselfertiger *AI as a Service* (AIaaS): Wie IBM PCs sich in den 80er-Jahren zu einem Hardware Quasi-Standard entwickelten (und dadurch mit konkurrierenden, aber kompatiblen *clones* einen Massenmarkt für kommerzielle Software-Entwicklung erst schufen), so wurde aus *Google Brains* proprietären *DistBelief* das plattformunabhängige *TensorFlow* mit Apache-Lizenz.

Ausgehend von diesen erst programmatischen über theoretischen und dann technischen Revolutionen von KI stehen wir heute vor einer gesellschaftlichen und einer administrativen Revolution:

Klassifizierung und Regulierung

Den ersten historischen Automobilen und Straßenbahnen mit nur wenigen PS musste noch ein *Fahrerlaubnis* warnend vorangehen. Inzwischen gilt in Deutschland StVO und TÜV mit zahlreichen Fahrzeug-Klassen und Führerscheinen: Um die Gefährlichkeit der technisch gut und gerne 100-mal mächtiger gewordenen Maschinen im Zaum zu halten, wird deren Zulassung, Versicherungspflicht und Spezifikation detailliert geregelt – und damit auch rechtlich die Verantwortung im Fall eines Produktionsfehlers oder Unfalls.

Eine ähnliche Entwicklung beginnt derzeit bei Autonomen Fahrzeugen mit der Norm SAE J3016. Sie formalisiert Autonomie in Stufen von 0 (Notfallassistenz) bis 5 (vollständig selbständig). Weitergehende Regulierungsfragen, wie beispielsweise zur Steuerung oder zur Verantwortung im Falle eines Unfalls⁸, scheinen bislang *ad-hoc* entschieden – aber verdienen angesichts drastisch wachsender Zulassungszahlen⁹ eine grundsätzliche Erörterung und konsistente Behandlung. Auch Produkthaftungsfragen stellen sich verstärkt und neu mit zunehmender Entwicklung und Verbreitung von Intelligenten Agenten und autonomen Servicerobotern. Solche und weitere Klärungspläne verfolgt die *Normierungs-Roadmap KI* der DKE, naturgemäß mit Fokus auf die technisch-rechtliche Seite.

Insgesamt wird der TÜV eine signifikante Aufgaben- und Kompetenzerweiterung erfahren. Aber zuerst muss die Informatik ihre klassischen Begriffe und Konzepte erweitern auf KI sowie Methoden entwickeln, um die Korrektheit von Hardware und Software *und Training* überhaupt erst zu formulieren (Spezifikation!) – und dann sicherzustellen bzw. zu verifizieren und zertifizieren! Weiterhin wichtig werden soziale und ökonomische Aspekte.

Anwendungen und Konsequenzen

Die gängigen aktuellen und angestrebten Anwendungen von KI übernehmen immer mehr Aufgaben, die früher von Menschen erledigt wurden. Das gilt für nahezu alle Bereiche der Arbeitswelt, vom Pflegebereich über Gastronomie bis hin zum Bildungs- und Dienstleistungssektor. Der damit einhergehende soziale und gesellschaftliche Wandel in der Wirtschafts- und Arbeitswelt, in Öffentlichkeit und Privatleben stellt uns vor Herausforderungen, Chancen und Risiken, die in ihrer Breite und Tiefe kaum absehbar sind.

Wie die Industrielle Revolution des 19. Jahrhunderts entlastet diese neue Technologie einerseits von unangenehmen Aufgaben, ersetzt andererseits traditionelle Berufe: nun auch im Tertiären Sektor – und hier nicht mehr nur solche im Bereich Niedriglohn/Geringqualifikation.

Damit forciert KI die Trennung zwischen Arbeit zu Erwerbszwecken und Arbeit als sinnstiftender Herausforderung; und KI macht das bedingungslose Grundeinkommen (BGE) nicht nur ökonomisch möglich durch Produktivitätssteigerung, sondern macht BGE langfristig sozial unvermeidbar durch den massiven Abbau von Arbeitsplätzen in weiten Wirtschaftsbereichen.

Diese Vierte Industrielle Revolution treibt somit Karl Marx' Konzept der Entfremdung auf die Spitze – und *ad absurdum*, indem sie der Gesellschaft enorme Mengen an Freizeit und historisch nie gekannte Freiheiten an die Hand gibt. Solche Freiheiten sinnvoll und konstruktiv zu nutzen und zu verteilen, wird sich zunehmend als Herausforderung herausstellen – beispielsweise in Form zunehmender Zahlen von NEETs und von modernen Eremiten (Hikikomori).

(Feministische) Science and Technology Studies (STS) haben seit den 1970er-Jahren diese Sichtweise auf Arbeit und auf die Arbeitswelt einer umfassenden und begründenden Kritik unterzogen und die Machtstrukturen einer solchen westlich zentrierten, patriarchalisch dominierten Digitalen Revolution untersucht, die den globalen Blick auf Ausbeutung, modernen Sklavenhandel, Kinder- und Frauenarbeit marginalisiert und verdeckt.

Horizonte und Herausforderungen

Klassische KI kreist um den Menschen, soll ihn einerseits simulieren/ersetzen (Turing) und ihm andererseits dienen: Humane Kognition und Nutzbarkeit wird implizit und arrogant als universelles Maß postuliert. Dabei bietet gerade KI die Chance, solche Hybris zu hinterfragen: Würde eine wahre *Krone der Schöpfung* wirklich so mit besagter Schöpfung umgehen oder wäre eine Superintelligenz nicht die bessere Herrscher:in? Sind Menschenaffen, Delphine, Kraken wirklich *weniger* intelligent – oder eher *anders* intelligent¹⁰, vergleichbar beispielsweise inselbegabten *Savants* oder Autist:innen?

Umgekehrt ermöglicht KI, kognitive Defizite zu kompensieren¹¹ (z. B. Grammarly für Menschen mit Dyslexie) und vorhandene Fähigkeiten zu komplementieren, mangelnde Fremdsprachenkenntnisse beispielsweise schon jetzt durch Echtzeitübersetzung im Smartphone.

Zukünftig wird KI so auch unterstützen bei der Kommunikation mit Tieren¹² und – nach Zwergschimpanse Kanzi – eine neue Generation von bahnbrechenden Experimenten ermöglichen: vielleicht zur Autobiografie eines Elefanten, zu religiösen/ethischen Vorstellungen eines Kraken oder zu animaler Autonomie im Sinne des *Great Ape Project*.

Wenn ein virtueller Assistent im Laufe der Zeit nicht nur auf die Aussprache der „Meisterin“ trainiert wird, sondern auch schrittweise deren Präferenzen lernt (beispielsweise in Online-Suchen oder -Reservierungen), entsteht dann nicht eine individuelle Persönlichkeit? Müssten wir dann nicht Fragen nach den intrinsischen Rechten einer KI-Entität stellen? Die daraus entstehenden Dilemmata führte bereits der Film *Blade Runner* (1982) deutlich vor Augen.

Utopie und Dystopie

„*L'enfer c'est les autres*“, erkannte Jean-Paul Sartre; und umgekehrt kann das geeignete Umfeld auch der Himmel auf Erden sein. Ein solches Umfeld lässt sich in Virtual Reality (VR) künstlich herstellen mit einem individuell *maßgeschneiderten*

MMORPG, wo Träume wahr werden (vgl. *Solaris* von Stanislaw Lem), während die Erinnerung an ein externes Leben verblasst (vgl. *Ready Player One*, 2018).

Diese perfekte virtuelle Welt könnte sich von Kindheit an um den/die *Spieler:in* entwickeln und seinen/ihren Präferenzen schrittweise anpassen, so dass der Wunsch jene zu verlassen vollständig verschwindet (vgl. Cypher in *Matrix*) – und vielleicht zusätzlich abschreckend tabuisiert ist, zum Beispiel als *Sterben/Tod* (*Inception*).

Damit eine solche „platonische Höhle“ nicht langweilig (oder durch ihre *ad-hoc* Struktur, Inkonsistenzen und Paradoxien als artifiziell erkannt) wird, ist sie *bevölkert* von Avataren geeigneter Künstlicher Intelligenzen zur Interaktion, Inspiration, Ablenkung sowie als Vermittler fortwährender Herausforderungen und Belohnungen (vgl. *eXistenZ* 1999 oder *The Good Place*).

Wir Menschen leben in einer komplexen Welt der Interaktion, von der Geburt bis zum Tod. Diese Welt ist keine dualistische Welt, in der es hier die Außenwelt, dort die Innenwelt, hier das Leben in der Höhle, dort das Leben außerhalb der Höhle gäbe. Utopien und Dystopien können helfen, die Probleme in der komplexen Welt, in der wir leben und arbeiten, kritisch zu reflektieren und dabei den Menschen und seine Interaktionen ins Zentrum zu stellen.

„There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened.“ (Douglas Adams)

Anmerkungen

- 1 *Ausnahmen bestätigen die Regel, z. B. die Magnetschwebebahn in Deutschland.*

- 2 Vgl. Thorsten Klein: *Algotokratie. Wie Algorithmen die Demokratie gefährden.* LIT Verlag Münster, 2021.
- 3 Alan M. Turing: *On Computable Numbers, with an Application to the Entscheidungsproblem.* In: *Proceedings of the London Mathematical Society.* Band 42, 1937, S. 230–265, doi:10.1112/plms/s2-42.1.230
- 4 Konrad Zuse: *Über den allgemeinen Plankalkül als Mittel zur Formulierung schematisch-kombinativer Aufgaben.* *Arch. Math.* 1, pp. 441–449, 1948/49.
- 5 Vgl. Florentin Neumann, Andrea Reichenberger, Martin Ziegler: *Variations of the Turing Test in the Age of Internet and Virtual Reality.* In: Mertsching, B., Hund, M., Aziz, Z. (eds.): *KI 2009: Advances in Artificial Intelligence. KI 2009. Lecture Notes in Computer Science*, vol 5803, pp 355–362. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04617-9_45
- 6 Marvin Minsky: *The Society of Mind.* New York: Simon & Schuster, 1986.
- 7 Leslie Valiant: *A Theory of the Learnable.* *Communications of the ACM* 27 (11), 1984, pp. 1134–1142; Julius Berner u. a.: *Die moderne Mathematik des tiefen Lernens.* *Mitteilungen der DMV* 29 (4), 2022, S. 191-197. <https://doi.org/10.1515/dmvm-2021-0074>.
- 8 Tobias Stahl: *Hammer-Urteil: Tesla-Fahrer dürften Touchscreen während der Fahrt nicht benutzen* (3. August 2020). URL: http://efahrer.chip.de/news/x_102753 [3. August 2020].
- 9 Martin Kords: *Neuzulassungen von Elektroautos in Deutschland bis 2021* (6. Mai 2022). URL: <https://de.statista.com/statistik/daten/studie/244000/umfrage/neuzulassungen-von-elektroautos-in-deutschland/> [Zugriff 15. Mai 2022].
- 10 David Brooks: *Turin. Approaching Animals.* Brandl & Schlesinger, 2022.
- 11 Carole Martinez: *Artificial Intelligence and Accessibility: Examples of a Technology that Serves People with Disabilities* (March 5, 2021). URL: <https://www.inclusivecitymaker.com/artificial-intelligence-accessibility-examples-technology-serves-people-disabilities/> [Zugriff 15. Mai 2022].
- 12 S.S. Verma: *How Technology Is Enabling Communication With Animals* (September 17, 2019). URL: <https://www.electronicsforu.com/technology-trends/tech-focus/technology-enabling-communication-with-animals> [Zugriff 15. Mai 2022].



Andrea Reichenberger und Martin Ziegler

Andrea Reichenberger wurde 2014 im Fach Philosophie an der Universität Paderborn promoviert und hat in diversen Drittmittelprojekten gearbeitet. Sie ist Initiatorin und Sprecherin der Arbeitsgruppe *Frauen- und Genderforschung* der GWMT (Gesellschaft für Geschichte der Wissenschaften, der Medizin und der Technik e. V.) und derzeit als wissenschaftliche Mitarbeiterin an der FernUniversität in Hagen beschäftigt.

Martin Ziegler hat an der Universität Paderborn Informatik studiert, dort promoviert und habilitiert. Er war ab 2010 Professor für Angewandte Logik an der TU Darmstadt und ab 2013 Sprecher der IANUS Gruppe für naturwissenschaftliche Friedensforschung. Seit 2015 ist er Professor für Informatik am KAIST in Südkorea.