

Anna Strasser, Wolfgang Soest, Ralf Stapelfeld, Katja Stepec (Hg.): Künstliche Intelligenz: Die große Verheißung Buchbesprechung

„Die Gefahr, dass der Computer so wird wie der Mensch,
ist nicht so groß wie die Gefahr, dass der Mensch so wird wie der Computer.“
(Konrad Zuse)

Tatsächlich befasst sich der Band kaum mit der Erklärung der Funktionsweisen von KI-Systemen, sondern im ersten Teil mit dem Begriff der menschlichen Intelligenz und dem Vergleich mit der sogenannten künstlichen Intelligenz und in einem zweiten Teil mit Grenzen und Folgen der „Künstlichen Intelligenz“ (KI). Für beide Teile zentral ist die Auseinandersetzung mit dem Turing-Test (dem Imitationsspiel, das auf der Ununterscheidbarkeit der Kommunikation von Menschen mit einem Automaten basiert, dabei jedoch so etwas wie „Verstehen“ ausblendet) mit der Turing-These (Intelligenz sei durch solche Art von Kommunikation definierbar), auch starke KI-These genannt, und der schwachen KI-These. Erstere vermeint, alle Aspekte menschlicher Intelligenz, also auch Bewusstsein, Selbstbewusstsein und Verstehen, durch einen qualitativen Sprung erreichen, wenn nicht gar übertreffen zu können, während letztere die Simulation von Elementen der analytischen Intelligenz für realistisch hält, nie jedoch alle Aspekte menschlicher Intelligenz.

Aus Sicht des feministischen Materialismus sind für die Eigenschaften menschlicher versus künstlicher Intelligenz das *Embodiment*, das in-der-Welt-Sein, die Intentionalität der handelnden Akteur:innen wesentlich für den Vergleich. Es fällt auf, dass zwar die hermeneutische Geschichtlichkeit des Menschen und die menschliche phylogenetische Entwicklung in „wetware“ als zentral für Eigenschaften der Intelligenz gesehen werden, aber die Körperlichkeit von (Software in) Hardware von Computern und Serverfarmen für KI-Systeme, mit all ihren Anforderungen an Stromzufuhr und -erzeugung, Metallabbau, -isolation und -bearbeitung nicht in Betracht gezogen wird. Ganz besonders fehlen solche Überlegungen bei jenen KI-Phantasten, die die Ablösung des Menschen durch eine höhere Intelligenz erwarten oder befürchten, als ob KI-Systeme selbständig in der Erde nach seltenen Erden buddeln, Chips herstellen oder Windräder zur Stromerzeugung aufstellen würden. Vielleicht sollen das alles dann Menschensklaven in einer KI-Diktatur erledigen. Aber mit welchen Druckmitteln könnte KI die Menschen dazu knechten? Aus feministischer Sicht ist bei diesen Diskussionen die allenthalben fehlende Materialität und Körperlichkeit, vor allem der Maschinen zu bemängeln, während die Historizität von Ideen eines künstlichen Menschen durchaus in den Blick genommen wird.

Da es keinen einheitlichen fächerübergreifenden Intelligenzbegriff gibt, und die Beurteilung der Fähigkeiten von KI abhängig ist von Definitionen von Mensch und Maschinen, werden hier im ersten Teil jeweils Aspekte von Intelligenz erörtert und in Abgrenzung zu beziehungsweise Überschneidung mit KI diskutiert. Die aus der Psychologie bekannten Bereiche analytische, soziale und emotionale Intelligenz sind der Mathematisierung beziehungsweise Maschinisierung unterschiedlich gut zugänglich. Während man sich auf eine prinzipiell erreichbare Simulation der analytischen Intelligenz leichter einigen kann, erfordert die soziale Intelligenz beispielsweise Welthaltigkeit und Empathie und die emotionale Letzteres sowie Selbstbewusstheit. Aus philosophischer und kulturwissenschaftlicher Sicht kommen andere Aspekte von Intelligenz zur Sprache, wie die für Berücksichtigung von Individualität notwendige subjektive Wahrnehmung und Gefühllichkeit, die bewusste oder zumindest kulturell geprägte Geschichtlichkeit, oder die onto- und phylogenetische



Anna Strasser, Wolfgang Soest,
Ralf Stapelfeld, Katja Stepec Hg.
(2021) Künstliche Intelligenz:
Die große Verheißung, MoMo
Berlin Philosophische KonTexte
Bd. 8, xenomoi Verlag
510 Seiten
Preis € 29,80 (gebunden)
ISBN-13: 978-3942106795

Leiblichkeit, die gefühlsmäßig erfahrene Kommunikation mit der Umwelt und damit die räumliche und zeitliche Zukunfts- und Welt-Offenheit.

Dass mathematische Funktionen nicht nur auf Zahlen und Buchstaben operieren können, sondern u. a. auch auf Worten und Sätzen (allesamt in Zahlen kodierbar), begünstigt m. E. Missverständnisse hinsichtlich semantischer oder gar sozialer und emotionaler Kompetenz der KI.

Nadine Schumann und Yaoli Du schlagen eine solche Betrachtung in einem ganzheitlichen Simulationsmodell vor, das auch menschliche Subjektivität erfasst. *Hans Zillmann* analysiert dafür auch die neurowissenschaftliche Subjektbeschreibung Antonio Damasio mit Bezug auf die Subjektivität, die er nicht nur sozialer und kultureller Zuschreibung zuordnet, sondern auch im Kontext emotionaler Erfassung der Umwelt sieht. Vorsichtig vermutet er, dass Gefühle Maschinen wohl unzugänglich seien, wobei er – unzutreffend – Maschinen ein eher deduktives Vorgehen zuschreibt. Aber zwischen induktivem Schließen, Subsumption etc. und Gefühlen klafft, von ihm unbetrachtet, immer noch eine große Lücke.

Daniel Dennett stellt KI in den Kontext von Evolutionstheorie und Darwins Thesen und fordert zum Verständnis des Ver-

hältnisses von Maschine und Mensch einen Perspektivwechsel, gewissermaßen beider *bottom-up*-Evolution, womit er dann eine Art „als ob“-Perspektive einnimmt, ähnlich einem digitalen Zwilling. *Michael Meyer-Albert* ruft eine Reihe von Philosophen auf, um drei Eigenschaften von Intelligenz herauszustellen, die der Ko-Konstruktion von Denken und Wahrnehmen Schwierigkeiten machen könnten: einmal die körperliche Kommunikation mit der Außenwelt, welche Geist und Welt mit Gefühlen verbindet, zum zweiten die hermeneutische Dimension Intelligenz als ein geschichtliches Verstehen von Sinn, in Absetzung von archivarisch hinterlegter Information. Intelligenz erfordert demgemäß zeitliche und geschichtliche Reflexivität, um eine Interaktion der Gegenwart mit der Tradition zu ermöglichen. Diese sich in Zusammenhängen modifizierende „Sinnbewegung des Verstehens“ (Gadamer) erschließt eine als leiblich erfahrene und geschichtlich reflektierte Welt-Offenheit des Menschen. Harte KI-ler:innen könnten gegen die Behauptung: „Das Objekt des Verstehens, das Verständnis des Verstehens und das Subjekt des Verstehens verändern sich im maschinellen Verstehen nicht“ einwenden, dass sich beim *deep learning* das *Subjekt* und das *Verständnis* des Verstehens, also das Modell, und das *Objekt*, also das zu Erkennende sehr wohl verändern. Nur ist dagegen zu fragen, was bedeutet hier *Verstehen*?. Schließlich führt Meyer-Albert ethische Bedenken gegen die Erschaffung künstlicher Intelligenzen auf, indem er das *Existieren* mitsamt seinen Bedingungen von Sterblichkeit und Angst ins Treffen führt. *Gutmann, Haag und Wadephul* führen mit einem tätigkeitsbezogenen Wissensbegriff, der auf den Verwendungszweck abhebt, abgesetzt von den Begriffen Information und Daten, zu einer Problematisierung von Maschinenzwecken gegenüber reflexivem Wissen mit selbstständiger Zwecksetzung, Rationalität und Verantwortung. Solche zweckorientierte Verwendung führt *Christian Freska* zu einer weiteren Voraussetzung für soziale Intelligenz, nämlich für ökologisches Problemlösen.

Im Gegensatz dazu argumentiert *Jan Tobias Fuhrmann*, dass Kommunikation nicht notwendigerweise Bewusstsein und Verstehen voraussetzt, weshalb er Programmen die Fähigkeit zugesteht, einen Turing-Test zu bestehen, ohne dass sie für sich selbst Sinn erzeugen müssten. Ja, ihm scheint die fehlende Reflexion gerade ein Vorzug für die Anschlussfähigkeit von Kommunikation zu sein. Aber genügt das für die Analogie beider Intelligenzen?

Christoph Merdes befasst sich mit der Frage, ob man Maschinen Moral beibringen kann, und modifiziert dafür einen Turing-Test für moralische Urteile, der sich im Grunde von einem sprachlichen Turing-Test nicht unterscheidet, aber für die Rechtfertigung des Urteils auf einen Gruppen-Turing-Test erweitert. Dabei stellt er fest, dass es zu Inkonsistenzen kommen kann, denn die Wertvorstellungen innerhalb der Gruppe können sich unterscheiden, aha! Wenn er auch die Möglichkeit einer programmierten oder freigelassenen KI für Emotionalität und Urteilsvermögen nicht ganz ausschließt, stellt er fest, dass die bisherigen Erfolge diesbezüglich nicht ausreichen.

Daniel Wenz diskutiert das mathematische Beweisen mittels KI und setzt es gegen das verstehende Nachvollziehen und ganzheitliche Verstehen der bewiesenen mathematischen Behauptung durch eine Anzahl von mathematischen Peers ab. So begründen Mathematiker:innen die Akzeptanz der Gültigkeit von

Beweisen in Abgrenzung von noch so langen Computer-Rechnungen, aber auch von immer möglichen Hardware-Fehlern.

Der philosophische Logiker *Walther Zimmerli* befasst sich schließlich mit der KI zwischen sinnstiftendem Mythos und der Verselbständigung in einem Hype. Dieser kulminiert schließlich in der Behauptung des immer wieder verschobenen Zeitpunkts der Kurzweilschen „Singularität“, also des Zeitpunkts, an dem die künstliche die menschliche Intelligenz übertrifft, und weiter sogar eines „postanalogen Menschseins“, in dem auch einem religiös verbrämten Trans- und Posthumanismus gehuldigt oder in dystopisch Bostromscher Manier Furcht vor einer Super-Intelligenz erzeugt wird. Er zeigt dazu die ideengeschichtlichen, literarischen und philosophischen Erzählungen, Voraussetzungen und Phantasmen, und auch die keineswegs unerheblichen Mittel, die für die Realisierung dieser Mythen eingesetzt werden. Wo und wie, um alles in der Welt, werden diese ungeheuren Materialisierungen in Servern für Big Data und (u. a. Lern-) Software enden? Wozu das alles, ist die Frage.

Der zweite Teil des Bandes beginnt mit einem Überblick von *Ralf Stapelfeld* über Diskussionen zur KI, ehe er einen Überblick über die folgenden Texte gibt. Dabei kommen insbesondere auf den Seiten 244-250 wichtige, vor allem logisch-philosophische Überlegungen zur Sprache, die danach nicht mehr behandelt werden.

Hans-Jörg Kreowskis und Wolfgang Kriegers Beitrag: „Künstliche Intelligenz – ‚künstlich‘ ja, ‚Intelligenz‘ wohl kaum“: Nach einer Darstellung der 50-jährigen Geschichte der KI und Robotik und ihrer institutionellen und militärischen Förderungen, großartigen Prognosen und Rückschlägen, begann in den letzten Jahrzehnten der große Hype nach einzelnen Erfolgen in eng begrenzten Anwendungsbereichen, wie Schach und Go, aber auch Entscheidungssystemen und Szenariensimulationen, ermöglicht durch die enorm vergrößerte Speicherkapazität und Rechengeschwindigkeit. Große Beachtung finden insbesondere konnektionistische Systeme (deep learning) im Zusammenhang mit Sprach- und Bilderkennung im Verein mit dem so genannten *Dataismus*, also der Fütterung lernender Maschinen mit ungeheuren Mengen von Bild-, Text- und Sprach-Daten aus all unseren Bewegungen im Digitalen, von denen man sich Anwendungen und Wertschöpfung in fast allen Bereichen heutigen Lebens erhofft. Die Autoren stellen die staatlichen Hoffnungen und stattlichen Förderungen für politische Wunschträume in aller Breite genüsslich dar, bevor sie sich dem Vergleich zwischen künstlicher und natürlicher Intelligenz widmen. Hier wird auch eine Kritik am ersten Teil des Buches geäußert, dass nämlich dort Intelligenz kaum direkt diskutiert wird, sondern eher Vernunft, Geist und Seele, weshalb man an dieser Stelle eher von technischer Realisierbarkeit von künstlicher Vernunft, künstlichem Geist und künstlicher Seele sprechen sollte.

Bei der Einordnung der Fähigkeiten von KI beginnen die Autoren mit der Turing-These als der Definition von Berechenbarkeit via Turing-Maschinen, die dadurch gestützt wird, dass bisher alle universellen Computermodelle und Programmiersprachen beweisbar äquivalent zu Turing-Maschinen sind. Die Beschränkungen der Berechenbarkeit (so kann mit reellen Zahlen nur approximativ gerechnet werden, oder es lässt sich im Allgemeinen nicht nachweisen, ob ein Programm im Sinne der Absicht kor-

rekt läuft) und die praktischen Verengungen durch Komplexitätsüberschreitung setzen Allmachtsphantasien der Computerei harte Grenzen, insbesondere innerhalb der Mathematik.

Diskutierbar wäre allerdings die Churchsche These, welche alle denkbaren Berechenbarkeitsmodelle hinsichtlich ihrer Funktionalität für äquivalent hält, weil alle bisher definierten es waren. Akzeptiert man die Churchsche These, wie es für die Berechenbarkeitstheorie zwar nicht notwendig, aber sinnvoll erscheint, also jener (durch Komplexitätseinschränkungen enorm zu reduzierenden, daher funktional nochmals ungleich größeren) Basistheorie der Computabilität (Berechenbarkeit, Nicht-/Entscheidbarkeit), so kollabiert der Vergleich zwischen Turing-Test als (angeblichem) Nachweis menschlicher analytischer (soziale oder emotionale Intelligenz ohnedies ausgenommen) und künstlicher Intelligenz. Die starke KI-These geht ja davon aus, dass kognitive Prozesse nichts anderes sind als Informationsverarbeitungs- und damit Berechnungsprozesse auf symbolischen Repräsentationen. Diese These impliziert die prinzipielle Simulierbarkeit menschlicher Intelligenz auf Computern (sehen wir von Komplexitätsproblemen einmal ab), sonst stünde sie im Widerspruch zur Churchschen These.

Ob KI der natürlichen Intelligenz gleichen kann, erörtern Kreowski und Krieger anhand der Turing-These – und lassen eine Antwort zwar offen, halten aber eine nachmalige Bejahung doch für unwahrscheinlich, in jedem Fall beide Begriffszugänge für sehr verschieden. Am Ende werden die Gefahren der Anwendungen von KI in der Gesellschaft (Überwachung) in Waffen und Wirtschaft und Politik diskutiert.

Während Kreowski und Krieger auf den Seiten 272 und 273 (empfehlenswert!) sehr einfach und verständlich die Prinzipien und Ergebnisse der Berechenbarkeitstheorie erklären, greift Dieter Mersch, nach einer kurzen Revue der Versuche, das menschliche neuronale System nachzubilden, und der philosophischen Abwägungen zwischen intentionalem Handeln, Denken und Rechnen, deutlich formaler in die Kiste der logischen Theorie des Konstruktiven. Um für die Beurteilung von Intelligenz die Möglichkeiten formaler Sprachen einzuordnen, kommt er in Anlehnung an die von Lucas 1961 angestoßene Diskussion über die Gödelschen Sätze, die die Unentscheidbarkeit wahrer Sätze innerhalb genügend ausdrucksstarker Formaler Systeme nachweisen, zu dem Schluss, dass auch für maschinelles Argumentieren menschliches Denken teilweise unerreichbar sein müsse. Stimmt, wenn Maschine und Mensch getrennt bleiben, aber dem lässt sich entgegenhalten, dass KI großenteils auf dem Vergleich von in Big Data gespeicherten Texten oder Bildern beruht, wobei Letztere zwar kein *Verständnis* für sich selbst haben, sich in ihrer Bedeutung für die beobachtenden Menschen aber der Einordnung in formale Systeme entziehen, während der auto-

matische Text- oder Bild-Abgleich selbst sich innerhalb der formalen Systeme bewegt. Dass „die weithin geteilten Einsichten von Kurt Gödels Unvollständigkeitssatz“ zu dem Schluss führen, dass der menschliche Geist prinzipiell die Möglichkeiten eines formalen Systems übersteigt, demonstriert ein Missverständnis mancher Philosophen, mathematische Erkenntnisse für diskutierbar zu halten, die ja nur innerhalb explizit definierter formaler Systeme Gültigkeit beanspruchen. Dies gilt vor allem für die konstruktive Mathematik, innerhalb derer sich Gödels Theorie der Berechenbarkeit bewegt. (L. E. J. Brouwers Einwand, dass die Mathematik des *tertium non datur* oder eine, die sich des Auswahlaxioms bedienen muss, auf tönernen Füßen stehe, bezieht sich ja gerade auf jenen ungeheuer größeren Teil der Mathematik, der sich indirekter, also nicht konstruktiver Beweise bedient. Allerdings vertritt er mit seinem Intuitionismus der intuitionistischen Logik, die ein mathematisches Objekt als Produkt der konstruierenden Intuition des Mathematikers entstanden und von ihm dabei überhaupt erst hergestellt sieht – ein für die starke Turing-These der KI befremdlicher Gedankengang.) Am Ende erwägt Mersch die (Un-) Möglichkeit algorithmischer Kreativität in Kunst und Literatur vor dem Hintergrund eines für soziale Menschen notwendigen ästhetischen Ethos.

Elektra Wagenrad vergleicht Computer mit dem klugen Hans, einem sprechenden und rechnenden Pferd, dem sie ein Verständnis für seine geklopften Antworten abspricht. Die Dialoge des klugen Hans mit seinem Besitzer hat der Schriftsteller Clemens J. Setz in seiner ungewöhnlichen Büchner-Preis-Rede rekonstruiert¹ und als Metapher für die (Un-)Möglichkeit menschlicher Verständigung herangezogen. Oder kann diese Metapher – er ist auch Mathematiker – gerade auch zur Stützung von Turings These gelten? Bedeutet denn dasselbe Wort einem anderen Menschen dasselbe? Die Klopfzeichen des Hans werden nach Aufforderung für eine Antwort-Interpretation genauer, und ähneln so den durch wiederholtes Nicken angegebenen Buchstaben und -folgen von Locked-in-Patient:innen mit Brain-Computer-Interface. Weitergehend befasst Wagenrad sich – den Mangel an Eigenwillen von KI im Vergleich zum menschlichen Entscheidungsverhalten problematisierend – mit dem neurowissenschaftlichen Libet-Experiment, das die Intentionalität bewusster menschlicher Entscheidungen infrage stellt. Allerdings sind die Versuchsbedingungen für das elektrophysiologisch gemessene Bereitschaftspotenzial so eng und unnatürlich, dass der Umfang ihrer Bedeutung gegen bewusste freie Entscheidungen durchaus infrage steht. So muss man noch nicht Wittgensteins unterschiedliche Sprachspiele bemühen, um die Unvergleichbarkeit beider Systeme von Intelligenzbegrifflichkeit festzustellen. Schließlich geht sie auf das erstaunlich „einfühl-same“ Sprachsystem GPT-3 ein, das sie selbst in einem spielerischen Experiment nicht nur zu sich selbst widersprechendem



Britta Schinzel

Britta Schinzel promovierte in Mathematik, arbeitete in der Computerindustrie und habilitierte sich in der Informatik. Im Rahmen ihrer Professur für Theoretische Informatik an der RWTH Aachen arbeitete sie zunehmend interdisziplinär. Sie war von 1991 bis 2008 Professorin für Informatik und Gesellschaft und Gender Studies in Informatik und Naturwissenschaft an der Universität Freiburg.

Plappern, sondern auch teilweise zu philosophisch sinnvollen Antworten bringen konnte. Allerdings, so lässt sich einwenden, weiß GPT-3 nichts davon, ob es sinnvolle oder unsinnige Antworten gibt, und Menschen sollten die für sie sinnvollen nicht als systemimmanent sinnvoll missdeuten.

Vladova und Friesike konzentrieren sich auf das Lernen aus Irrtümern als wichtigen, für KI unerreichbaren Aspekt menschlicher Intelligenz. Auch dagegen lässt sich einwenden, dass die Methode beim Backtracking im Deep Learning gerade das Annähern an fehlerhafte Ergebnisse ist. Aber es ist wohl gemeint, dass Fehler neue oder unerwartete Lösungen in Gang setzen.

Sybille Krämer lenkt den Blick auf einfache Anwendungen von schwacher KI (wobei sie diese mit allgemeiner Software identifiziert) und sieht für diese eine weitere Verflachung der menschlichen Kultur in Text und Bild auf das Maschinenlesbare. Im Weiteren problematisiert sie innerhalb der KI das Maschinelle Lernen und stellt sehr in Frage, ob es je möglich wäre, dass Maschinen Sinn und Bedeutungszusammenhänge verstehen können sollten.

Rico Hauswald vergleicht Deep-Learning-Ergebnisse mit Orakeln und sieht darin Vorteile (Demokratisierung von Wissen) und Gefahren, da die Black Box von Urteilsfähigkeit entwöhnt. Allerdings ist beides nicht KI-typisch, dafür genügt schon Google und klassische Entscheidungsalgorithmen. Thomas Weiss hinwiederum diskutiert die Vollautomatisierung im Hinblick auf Arbeitsplatzverluste und phantasiert über sich emanzipierende Artificial General Intelligence von einem KI-Proletariat.

Reinhard Kahle problematisiert die rückwärtsgewandten Entscheidungen durch KI, wenn sie auf vergangenen Datensammlungen beruhen, beispielsweise im Hinblick auf ein Ausbleiben von Innovationen sowie den Autonomieverlust von Menschen bei der Übernahme von Aufgaben durch KI.

Engel und Schultheis befassen sich ebenfalls mit KI-basierten Entscheidungen anhand einer Delphi-Studie zu den Erwartungen der Befragten über die entsprechenden Veränderungen. Daran anschließend erörtern sie, wieviel Vertrauen in diese Technologie gesetzt werden kann oder sollte, was sie weiter zu moralischen Fragestellungen führt.

Mit Ethik und Moral im Zusammenhang mit KI befassen sich die weiteren Texte, wobei zu unterscheiden ist, ob sich die ethischen Fragen auf KI-Ergebnisse beziehen, oder ob versucht werden soll, ethische Anforderungen ins Design von KI-Entwicklungen einzubeziehen. *John Michael und Hendrik Kempt* diskutieren den Aufbau von Beziehungen und Vertrauen zwischen Menschen und Robotern in Grenzen durchaus positiv, während *Orphelia Deroy* gerade die Möglichkeit von Vertrauen in KI problematisiert und dazu Rechtfertigung verlangt. *Catrin Misselhorn* hält hingegen eine Maschinenethik für möglich, fordert dazu aber drei Grundsätze, von denen zwei meines Erachtens im Widerspruch zur Funktionsweise von KI stehen und so ihre Anwendung aushebeln würden, etwa wenn der Mensch immer die Verantwortung für KI-Entscheidungen tragen soll.

Ein wenig enttäuschend ist, dass das Buch wenig auf die Funktionsweise von KI zum Verständnis ihrer Möglichkeiten und Unmöglichkeiten eingeht. Ausnahmen bilden nur Hans-Jörg Krewowskis und Wolfgang Kriegers Darstellung der Entscheidbarkeit und Berechenbarkeit S. 272 ff. und Sybille Krämers Darstellung von Deep Learning ab S. 346, dies alles auch nur im Bereich des Formalen. Hingegen scheint es mir erstaunlich, dass die Körperlichkeit von (Soft- in) Hardware von Computern und Serverfarmen für KI-Systeme, mit all ihren ungeheuren Konsum von und Anforderungen an Stromzufuhr und seine Erzeugung, Metallabbau und seine Isolation und Bearbeitung nicht in Betracht gezogen wird. Ganz besonders fehlen solche Überlegungen bei jenen KI-Phantasten, die die Ablösung des Menschen durch eine höhere Intelligenz erwarten oder befürchten, als ob KI-Systeme selbständig in der Erde nach seltenen Erden buddeln, Chips herstellen oder Windräder zur Stromerzeugung aufstellen würden. Vielleicht sollen das alles dann Menschensklaven in einer KI-Diktatur erledigen. Aber mit welchen Druckmitteln könnte KI die Menschen dazu knechten?

Anmerkung

1 <https://www.sueddeutsche.de/kultur/clemens-j-setz-georg-buechner-preis-woyzeck-karl-krall-denkende-tiere-pferde-1.5457889>



Peter Brödner

Die Illusionsfabrik der »KI«-Narrative

Derzeit sind medial verbreitete »KI«-Narrative wieder en vogue. In den 1980er-Jahren versuchten Ansätze symbolischer KI, explizites Wissen über Praktiken kooperativer kognitiver Arbeit und daraus zu ziehende Schlüsse in Gestalt wissensbasierter oder Expertensysteme zu modellieren. Im Unterschied dazu richtete sich heutige Ansätze auf die Gewinnung von Berechnungsverfahren zur Bewältigung kognitiver Aufgaben ab. Diese Verfahren maschinellen Lernens (ML) zu umgehen – tatsächlich aber die hohen Hürden hinreichend zu überwinden – ist, wie die neuen Ansätze zeigen, erneut unüberwindlich erschienen. Dies zeigt, dass »KI«-Protagonisten nicht einmal wissen können, worin sich die künstlich intelligent genannten Systeme eigentlich genau von herkömmlichen Computersystemen unterscheiden – ein Umstand, aus dem viele Illusionen über Funktionsweisen und Leistungspotenziale dieser Systeme erwachsen. Neue Probleme ergeben sich einerseits aus der kaum einschätzbaren Relevanz und Validität der Daten, zudem aus der Intentionalität und Kontingenz sozialer Praktiken, andererseits aus einer höheren Art der Undurchschaubarkeit des Systemverhaltens. Das wirft zudem eine Reihe neuer, freilich noch ungelöster ethischer Fragen auf.

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