

Künstliche Intelligenz und Weltverstehen

Veranstalter: Leopoldina-Zentrum für Wissenschaftsforschung; Interdisciplinary Network for Studies Investigating Science and Technology (INSIST)

Datum, Ort: 30.09.2020–02.10.2020, Halle (Saale)

Bericht von: Florian Arnold, Staatliche Akademie der Bildenden Künste Stuttgart

Fragen zur Künstlichen Intelligenz (KI), die über technische Belange hinausgehen, treten in den letzten Jahren zusehends in den Vordergrund. Dabei wird unter dem Begriff „Künstliche Intelligenz“ ein Phänomen umrissen, das sich weniger durch Eindeutigkeit und feststehende Definitionen auszeichnet als durch einen anhaltenden diskursiven Wandel, der seit Jahrzehnten einen blinden Fleck umkreist: Lässt sich eine nicht-menschliche bzw. nicht-anthropomorphe Intelligenz (für uns) denken, ja herstellen, oder haben wir es hierbei eher mit Projektionen und Rekonstruktionen unserer selbst in technischen Medien zu tun? Beide Sichtweisen resultieren aus gravierenden Veränderungen innerhalb unserer Lebenswelt, und beide Sichtweisen führen wiederum zu unterschiedlichen Einstellungen und Herangehensweisen in Fragen der Beforschung und Entwicklung Künstlicher Intelligenzen heute.

Wie sich das Phänomen KI auf unser gesellschaftliches und persönliches Selbstverständnis auswirkt, war Thema einer Herbsttagung, bei der mehrheitlich NachwuchsforscherInnen in der erklärten Absicht zusammenkamen, zu einer „gesamtgesellschaftlichen Reflexion über KI“ beizutragen. In den Worten des Mitorganisators Christian Vater geht es um das „Desiderat“ einer „digitalen Aufklärung“ bzw. „Aufklärung des Digitalen“ und die Zusammenführung der verschiedenen Wissenskulturen im Zeichen des zu reaktualisierenden Fragenkatalogs Immanuel Kants: Was kann ich wissen? Was soll ich tun? Was darf ich hoffen? Was ist der Mensch? Zu hören waren entsprechend Vorträge aus der Soziologie, Rechts-, Politik- und Geschichtswissenschaft, der Medien- und Genderforschung, der Philosophie, aber auch aus theologischer Perspektive. Darüber hinaus prä-

sentierten NATALIE SONTOPSKI (Leipzig) und AMELIE GOLDFUSS (Halle/Salle) ein Design-Research-Projekt, den Sprachassistenten *MiauMiau*, der coronabedingt leider nicht in einer Ausstellung erprobt werden konnte.

Den Auftakt machte eine Keynote von ANDREAS KAMINSKI (Stuttgart), der aus philosophisch-ethischer Perspektive den Rahmen absteckte für die weitere Auseinandersetzung. Was am Problem der Reliabilität von KI-Systemen in Entscheidungssituationen und deren Bewertung insbesondere zu Tage tritt (aber dort nicht ausschließlich), ist das Problem der Schuldzuschreibung oder allgemein der Verantwortbarkeit von künstlichen Intelligenzen, die nicht nach Gründen, sondern Algorithmen, Korrelationen oder auch Ursachen entscheiden können. Die Diskussion entzündete sich an der Haltbarkeit einer Unterscheidung zwischen normativ-diskursiven Gründen einerseits, die sich im inferentiellen, intersubjektiven „Spiel des Gebens und Forderns von Gründen“ (Robert Brandom) zu behaupten haben, und mechanischen oder stochastischen Ursachen andererseits, die sich aus der Programmierlogik des jeweiligen Systems ergeben.

Am Beispiel eines KI-Lügendetektors, wie er im Vortrag Kaminskis u.a. Thema war, lässt sich veranschaulichen, dass eine gewisse „epistemische Opazität“ besteht, etwa mit Blick auf die Parameter des Trainingsprogramms der KI, aber auch deren Mathematisierung, die es letztlich nicht erlaubt, tatsächlich von guten, richtigen Gründen zu sprechen, egal wie hoch die Reliabilität ist. Wie zu ergänzen wäre, hat das seinen Grund darin, dass, auch wenn ein KI-Lügendetektor gelernt hat, bestimmte Mikrogesten als Anzeichen für die (Un-)Wahrhaftigkeit einer Aussage zu lesen, dies bereits das Resultat gesetzter Parameter und Korrelationen darstellt, also selbst Resultat von Wertungen ist, die im unbedarften Anwendungsfall nicht nur epistemisch, sondern auch politisch und moralisch zumeist undurchsichtig bleiben.

Ferner klafft im konkreten Fall stets die empfindliche Lücke, dass etwa die Zuschreibung einer Lüge weiterhin der ergänzenden Beurteilung und Entscheidung durch menschliche Akteure bedarf. Denn normative Urteile, als deren Inbegriff Schuldzuschrei-

bungen gelten können, lassen sich nicht auf Tatsachenkorrelationen reduzieren. Schon die Semantik des Begriffs „Lüge“ erschöpft sich nicht in einer Tatsachenfeststellung der Art: Wangenzuckung ja oder nein. Vielmehr erfordert es eine kontextuelle Semantik etwa des Dialoggeschehens im Rückblick auf die nicht weniger komplexe Semantik des Tatgeschehens, in dem sich die unüberschaubare Zahl unterschiedlichster Parameter – und wer wüsste zu bestimmen, welche überhaupt die einschlägigsten sind? – zu einem normativen Urteil verdichten. Wie das aber geschieht, ist für KI-Systeme mehr als opak, es ist der konstitutive Graubereich menschlicher Urteilskraft, insbesondere der juristischen Intuition.

Neben JASCHA BAREIS (Berlin/Karlsruhe), der die letztlich menschenverachtende Vagheit national divergierender Rechtsdefinitionen von autonomen Waffensystemen herausarbeitete, problematisierte LUNA RÖSINGER (Köln) den rechtlichen Status von elektronischen Personen in Fragen der Verantwortbarkeit. Was als Minimum der Rechtsfähigkeit gilt, die Normbefolungsfähigkeit, wurzelt ihr zufolge in einem starken Autonomiebegriff, wie er aus der Tradition des deutschen Idealismus (Kant, Fichte) in unser heutiges Rechtsverständnis eingegangen ist. KIs dagegen können sich nicht nach eigenen Gründen entscheiden, d.h. überhaupt nach Gründen, die intersubjektiv verhandelt und kodifiziert, nicht nur erkannt, sondern anerkannt werden müssen. In einem Rechtsraum dagegen, der KIs Rechtsfähigkeit zubilligt, wie etwa beim automatischen Fahren, ist mit erheblichen Verantwortungslücken zu rechnen, die eine eindeutige Schuldzuschreibung verunmöglichen. So der Status quo.

Aber wäre hier, wie auch aus dem Publikum eingewendet wurde, der klassische Autonomiebegriff der Selbstgesetzgebung in seiner Exklusivität womöglich neu zu überdenken – etwa im erweiterten Sinne einer Selbstprogrammierung? Dass wir heute schon mit deutlichen Lücken zu kämpfen haben, die durch KIs zumindest mitverursacht werden, macht das chronische Problem des „informed consent“ in der digitalen Medizin deutlich. Ist es überhaupt noch möglich, vollständig infor-

miert Einschränkungen des eigenen persönlichen Datenschutzes zuzustimmen, um an dem versprochenen Segen einer digitalisierten Individualmedizin teilzuhaben? MAIKE JANSSEN (Weimar) und ECKHARD GEITZ (Freiburg) antworteten mit Nein, da die Datendistribution in den unterschiedlichen Informationsnetzwerken und Clouds längst unüberschaubar geworden sei. Eine technische Verunmöglichung des Nachvollzugs konterkariere den Rechtsanspruch auf Informativität; dieser Anspruch ist gegenwärtig also eher als eine Rechtsfiktion aufzufassen, die weniger der Gesundheit des Patienten als der Auswertung seines Datenkörpers durch Dritte dient.

Daran scheint jedoch nicht jeder gleichermaßen Anstoß zu nehmen, wie etwa NELE WULF (Furtwangen) in ihrem Beitrag zur Frage des Gedächtnisses in Zeiten seiner digitalen Bewirtschaftung darlegte: KI-gestützte Datenbibliotheken stellen nicht nur eine Eins-zu-Eins-Abbildung von Lebensläufen in Aussicht, sondern speisen sie in ein kollektives digitales Gedächtnis ein. Nicht wenigen Nutzern geht es dabei um einen „Upload“ der eigenen Persönlichkeit in eine Art Menschenbibliothek, die wiederum die Verheißung wahr machen soll, den Wunsch nach Unsterblichkeit auf lange Sicht zu digitalisieren – ungeachtet des Problems etwa, dass menschliche Erinnerungen durch Interpretationen und Lücken gebildet werden, die allererst den narrativen Spielraum des je eigenen Selbstentwurfs ausmachen. Wulf sprach in diesem Kontext treffend von einer „Dialektik zwischen Entlastung und Entmündigung“.

Damit brachte sie begrifflich auf den Punkt, was HENNING MAYER (München) seinerseits als persönliche Erfahrung in einem Startup schilderte, das sich mit der Entwicklung eines Chatbots zum Coaching von Managern beschäftigt. Unter der Leitfrage, wie die Welt maschinenlesbar werden kann, kam er im Zuge seiner Feldstudie zu der bezeichnenden Antwort, dass die Programmierung des Chatbots nicht zuletzt den Entwicklern selbst ein gewisses Einfühlungsvermögen für dessen „Denkweise“ abverlangt, um eine möglichst reibungslose Implementierung zu ermöglichen. „Was würde der Bot jetzt denken?“ scheint so gesehen zu-

gleich schon die halbe Antwort auf die Frage: „Wie macht man einen User glücklich?“ – vorausgesetzt, man vernachlässigt bereitwillig den Unterschied zwischen menschlichem User und Bot. In einer Variation von Thomas Nagels berühmtem Aufsatz zum Verständnisproblem nicht menschlicher Bewusstseinsformen könnte man also fragen: „What is it like to be a bot?“ Wäre mit dieser Fragestellung jedoch schon der Grund erreicht, von dem aus menschliche Autonomie und Selbstverantwortung im Spiel des Gebens und Forderns von Gründen triftig problematisiert, gar revidiert werden müsste?

Bevor man sich darauf ernstlich einlassen mag, bräuchte es allererst eine Auseinandersetzung mit dem Begriff „Künstliche Intelligenz“ in seiner Geschichte, wie zwei weitere Vorträge zeigten. Ausgehend von René Descartes' methodologischem Rationalismus, Gottfried Wilhelm Leibniz [U+02BC] totaler Symbolsprache und Michel Foucaults „Episteme“ als Präfiguration der E-Humanities diskutierte LORIANA BUHR (Göttingen) das neuzeitliche Konzept einer „mathesis universalis“. SANDRO HERR (Wuppertal) bemühte sich um eine genealogische Aufarbeitung der Frage, was eigentlich unter dem „Künstlichen“ Künstlicher Intelligenzen zu verstehen sei. Im Rückgang etwa auf das Phänomen der Schrift als Kulturtechnik, die für sich schon ein Prozeptionsmedium darstellt, das nicht mehr notwendigerweise von menschlicher Vorstellungskraft begleitet sein muss (etwa im Fall von Codes), lässt sich begreifen, dass intelligentes Verhalten seit jeher auf die Künstlichkeit – oder besser: Kunstfertigkeit – von Medien und Methoden angewiesen ist. Laut Herr ist es letztlich eine Frage der Organizität, ob sich Intelligenz auf leiblicher Basis „menschlich“ oder durch andersartige Medien und Materien „künstlich“ ausbildet.

Neuere Ansätze des Machine Learning, so SIMON EGBERT (Berlin), scheinen genau hier anzusetzen, indem sie eine neue Epistemologie mit sich bringen: ein „data-behaviorismus“ (Antoinette Rouvroy), der von einer Tiefenhermeneutik menschlichen Sinnverstehens auf eine „logic of the surface“ umstellt. Mit N. Katherine Hayles unterschied Egbert dabei zwischen „thinking“ und „cognition“, um auf eine wesentliche Verände-

rung in unserer Denkökologie aufmerksam zu machen: Der Begriff Kognition markiert in diesem Kontext die Differenz zur anthropomorphen Reflexion als (selbst-)bewusstseinszentriert. Dagegen stehen „human-technical cognitive ensembles“ für eine Neuverteilung von Kognition über Menschen und Maschinen in einem durchaus neuen qualitativen Grad. So lässt sich der Intelligenzbegriff weiter fassen, ohne Gefahr zu laufen, konzeptuelle Bestimmungslücken mit Blick auf künstlichen Intelligenzen durch anthropomorphe Projektionen zu stopfen oder gar als metaphysische Schlupflöcher zu nutzen.

Wozu Letzteres führen kann, veranschaulichte FLORIAN HOFFMANN (Speyer) mit einem Strukturvergleich. Ausgehend von Niklas Luhmanns systemtheoretischem Ansatz, argumentierte er überzeugend dafür, das Phänomen KI als Lückenbüßer für traditionelle Theologumena verstehen zu lernen, genauer: als strukturellen Fluchtpunkt im Sinne des Luhmannschen Systems Religion. An die diskursive Stelle von „Transzendenz“ als disjunktivem Gegenpart zur „Immanenz“ innerhalb der Systeme/Umwelt-Unterscheidung, d.h. an die altehrwürdige Systemstelle Gottes, tritt heute bisweilen das Phänomen KI, indem es eine analoge Semantik entfaltet. Das diskursive Phänomen KI erzeugt auf diese Weise erneut geradezu „religiöse Problemlagen“, wie sie etwa an technoaffinen Glaubensgruppen, allen voran den kalifornischen Transhumanisten und anderen New-Age-Ausläufern, heute lediglich im Extrem deutlich werden.

Darauf, so will man meinen, kann nur sinnvoll mit dem alteuropäischen Projekt der Aufklärung reagiert werden, etwa indem der metaphysische Bias einschlägiger Diskussionen aufgezeigt oder überhaupt das historische Bewusstsein für die KI-Entwicklungen und ihre bereits seit Beginn der 1940er-Jahren anhaltende Kritik geschärft wird. In dieser Hinsicht boten HELEN PIEL und DINAH PFAU (beide München) zumindest cursorisch Aufschluss, indem sie die Geschichte der KI-Forschung der letzten acht Jahrzehnte skizzierten.

SUSANNE SCHREGEL (Köln) erweiterte diese historische Perspektive um Genderfragen, wodurch deutlich wurde, dass vor allem ein kritisches Geschichtsbewusstsein unabdingbar ist, um überkommene vage Verhei-

lungen nicht für neuerliche Offenbarungen einer Zukunft zu halten, in der sich grundlegende Probleme wie der „gender trouble“ – KI sei Dank – letztlich von selbst neutralisieren würden. Dazu zählt insbesondere die Korrektur von historischen Narrativen, deren naheliegendstes Beispiel wohl das einer deutschen Verspätung in der technischen Entwicklung von KIs ist. Hier gilt es erneut kritisch nachzufragen und auch ideengeschichtliche Faktoren zu würdigen, um aus einer vermeintlichen Verschlafenheit nicht stante pede in eine Hyperaktivität unbesonnener Aufholjagden überzugehen.

Die Podiumsdiskussion machte das zentrale Problem namhaft. RUDOLF SEISING (München), GABRIELE GRAMELSBERGER (Aachen) und CHRISTIAN VATER (Karlsruhe) konstatierten in überraschender Einhelligkeit, dass es KI in einem anspruchsvollen Sinne zurzeit nicht gebe. Stattdessen gebe es ein Forschungsgebiet, das diesen Namen erhalten habe – aus unterschiedlichen, nicht nur uneigennütigen Gründen. Wie Gramelsberger ergänzte, wäre es dagegen interessant zu erforschen, warum Menschen glauben, dass es KI gebe oder geben sollte. Letztlich seien in den Diskussionen, auch denen der Tagung, vor allem Metaphern im Spiel gewesen, die es zu entmystifizieren gelte. Erst in diesem Zuge ließen sich Fragen nach den Begriffen „Intelligenz“ und „Entscheidung“, aber auch „Wissenschaft“ innerhalb einer KI-gestützten Forschungspraxis sinnvoll stellen.

Dem kann man nur zustimmen. Was digitale Aufklärung betrifft, stehen wir erst am Anfang, und zu vermuten ist, dass dieser Prozess in den nächsten Jahrzehnten schwerlich seinen Abschluss finden dürfte. Im Gegenteil ist nicht abzusehen, welche Veränderungen uns bevorstehen und auf welchen Wegen sich ein kritischer Diskurs angesichts der rasanten Entwicklungen noch zu behaupten haben wird. Das ist nicht als Zukunftsverheißung gemeint, sondern als Ermahnung, weiterhin wirksame und förderliche Traditionen nicht ohne Not zu kappen. Sofern man sich nicht einem digitalen Goldrush und den damit einhergehenden Wildwestverhältnisse vorbehaltlos überantworten will, führt wohl kein Weg daran vorbei, das weite Territorium der europäischen Technik-, Politik und

auch Ideengeschichte erneut zu durchqueren. Gerade eine geschichtsvergessene, metaphysische Aufladung der KI-Debatte dient heute kaum jemanden; sie schadet vielmehr im Hinblick auf eine selbstbestimmte Auseinandersetzung mit den herausziehenden Herausforderungen der Zukunft. Wie die Tagung eindrücklich belegen konnte, sind zumindest erste Schritte zur diskursiven Aufklärung eines noch meist im Dunkeln tappenden Welt- und Selbstverstehens unter den Auspizien Künstlicher Intelligenzen bereits getan. Auf diesem Weg gilt es fortzuschreiten – ad fontes.

Konferenzübersicht:

Andreas Kaminski (Stuttgart): Gründe geben. Maschinelles Lernen und Entscheidungen

Helen Piel / Dinah Pfau (beide München): Geschichte(n) der KI

Lorina Buhr (Göttingen): Künstliche Intelligenz als *mathesis universalis*? Eine historisch-diskursive Untersuchung der KI

Sandor Herr (Wuppertal): „Künstliche“ Intelligenz? – Eine kleine begriffsgenealogische Naturalisierung von KI

Luna Rösinger (Köln): Der Autonomiebegriff im Kontext Künstlicher Intelligenz als Prüfstein für die Rechtsphilosophie

Maïke Janssen (Weimar) / Eckhard Geitz (Freiburg): Zur Unvereinbarkeit von KI und Informed Consent

Nele Wulf (Furtwangen): Künstliche Intelligenz und menschliche Imagination. Biographische und öffentliche Erinnerungskultur zwischen Assoziationsfähigkeit und Algorithmisierung

Klaus Angerer (Tübingen): Akzeptierbarkeit als Grundlage der Verantwortung von KI-Software: Eine Idee für ein Forschungsprojekt

Henning Mayer (München): Sozialität als Code. Zur Analytik der Verschaltung von Intelligenztypen im Labor

Susanne Schregel (Köln): Köpfe ohne Körper, Männer ohne Geschlecht, und die Schönheit der Frauen. Überlegungen der Geschlechtergeschichte der Künstlichen Intelligenz

Podiumsdiskussion

Rudolf Seising (München) / Gabriele Gramelsberger (Aachen) / Christian Vater (Karlsruhe)

Natalie Sontopski (Leipzig) / Amelie Goldfuß (Halle/Saale): The Privilege of Anger. Bias bei digitalen Sprachassistenten

Simon Egbert (Berlin): Epistemische Implikationen Künstlicher Intelligenz: Maschinelles Lernen und „human-technical cognitive ensembles“

Florian Hoffmann (Speyer): Exalting a new God? Künstliche Intelligenz als religiöses Problem

Jascha Bareis (Berlin / Karlsruhe): Die Abgabe der Souveränität über Leben und Tod? Letale autonome Waffensysteme als normative Problemstellung zwischen Fakt und Fiktion

Tagungsbericht *Künstliche Intelligenz und Weltverstehen*. 30.09.2020–02.10.2020, Halle (Saale), in: H-Soz-Kult 25.11.2020.