

MARTIN W. SCHNELL / LUKAS NEHLSSEN

Begegnungen mit Künstlicher Intelligenz

Intersubjektivität, Technik, Lebenswelt

In allen gesellschaftlichen Bereichen kommt es täglich zu Begegnungen des Menschen mit Künstlicher Intelligenz (KI) in Form von Kommunikationen und Interaktionen. Diese Begegnungen verändern unsere Verständnisse von Selbst-, Welt und Fremdverhältnissen. Die Elemente dieser komplexen Sachlage gilt es zu klären, so dass deutlicher wird, was eigentlich Begegnungen mit Künstlicher Intelligenz sind.

Begegnungen mit Künstlicher Intelligenz vertiefen die digitale Transformation der Lebenswelt. Dabei treten Künstliche Intelligenzen nicht mehr wie Maschinen, die der Mensch auf gewisse Weise steuert, auf, sondern als vermeintlich autonome Akteure¹. »Die *information society* wandelt sich zur *interaction society*« (Priddat 2022, 317). Autos geben im Innenraum differenzierte und sprachliche Signale von sich, die den Fahrer mal überwachen, mal informieren und mal belohnen. Kreative Maschinen gestalten auf Anforderung Gemälde im Stile von Picasso, Bacon oder van Gogh, schreiben angefangene Sinfonien von Mahler »zu Ende« oder komponieren zahllose Choräle im Stile von Bach. Die neuesten Sprachprogramme zeigen beeindruckende Fähigkeiten sowohl beim Sprachverständnis als auch bei der Sprachproduktion und schreiben so täuschend »echte« journalistische Artikel oder führen Konversationen mit Menschen, die bisweilen durchaus interessant sind. Diese und andere Begegnungen zwischen Mensch und KI produzieren Selbstverständlichkeiten, die als solche zu unbefragten Beständen der Lebenswelt werden. Nachfolgend sollen die einzelnen Elemente dieses Definitionsversuchs gesondert dargestellt und aufeinander bezogen werden.

Begegnung

Die Begegnungsphilosophie beginnt mit Humboldt und Feuerbach, die in der Moderne den Dialog von Ich und Du als für die Entstehung von

- I Wenn in der Einleitung an verschiedenen Stellen die männliche Form gewählt wird, so geschieht dies aus Gründen der Leserlichkeit; in solchen Formulierungen sind selbstverständlich alle Geschlechter inkludiert.

Sinn und Bedeutung relevant eingeführt haben. In seinem Aufsatz *Über den Dualis* (1827) beschreibt Wilhelm von Humboldt Anrede und Erwiderung zwischen Ich und Du als Ausgangspunkte aller Sinnbildung. Bereits hier zeigt sich, dass die Begegnung eine gewisse ethische Dimension besitzt, die in der gesamten Diskussion gegenwärtig ist und in der Philosophie von Emmanuel Levinas zur vollen Entfaltung kommt.

Eine anthropologische Realisierung der Theologie ist, Marx zufolge, die Reaktion Ludwig Feuerbachs auf den Idealismus. In seinen *Grundsätzen der Philosophie der Zukunft* (1846) formuliert dieser, dass »die wahre Dialektik kein Monolog des einsamen Denkers mit sich selbst, (sondern) ein Dialog zwischen Ich und Du ist.« (§ 62) Fortgesetzt und zur Blüte gebracht wurde das Paradigma der Begegnung durch Hermann Cohen und die jüdische Dialogphilosophie, welche von Karl Löwith im 20. Jahrhundert, teilweise in der Rückschau und als Reaktion auf den Nationalsozialismus, quasi als Bewegung herausgestellt worden ist (vgl.: Löwith 2013).

Franz Rosenzweig, Schüler und Kritiker Hegels, formuliert in *Der Stern der Erlösung* (1921): »Erst indem das Ich das Du als etwas außer sich anerkennt, ..., wird es zum Ich.« (Rosenzweig 1988, 195) Eine daran angelehnte Formulierung in Martin Bubers Werk *Ich und Du* (1926) gibt dem dialogischen Prinzip der Begegnung dann sein bekanntes Motto: »Der Mensch wird am Du zum Ich.« (Buber 1983, 37) Die jüdische Dialogphilosophie betont besonders zwei Aspekte der Begegnung. Einmal den Kontext der Begegnung und dann die Eigenart der Begegnung selbst treffend (vgl.: Wojcieszuk 2010).

Der Kontext der Begegnung ist die Lebenswelt oder, wie Aron Gurwitsch sagt, das Milieu, wie es in seinem Buch *Die mitmenschlichen Begegnungen in der Milieuwelt* (1931) heißt. Die Lebenswelt in der Version des Milieus ist nach Max Scheler als das »auf mich als wirksam Erlebte« (Scheler 2000, 154) definiert. Die Eigenart einer Begegnung zwischen Ich und Du in der Lebenswelt ist die Antwort auf die Frage, »wie man ... überhaupt zum anderen Menschen als Menschen kommt, wie sich uns der andere Mensch erschließt.« (Gurwitsch 1977, 29) Ausdrücklich werden die Verschmelzung von Ich und Du (W. Wundt), die Einfühlung des Ich in das Du (Th. Lipps) und der Analogieschluss zwischen Ich und Du (J. St. Mill) als Zugangsarten zurückgewiesen (vgl.: ebd., 28ff). Stattdessen konstituiere die Begegnung das Ich und das Du als Differenz. Eine Begegnung beinhaltet bei aller Nähe immer auch eine gewisse Ablösung zwischen Ich und Du (vgl.: Gurwitsch 1977, 51ff.). Hier kommen Motive zur Geltung, die im 20. Jahrhundert von Georg Simmels Soziologie bis zu Paul Ricœurs Philosophie des Anderen reichen (vgl.: Schnell 2020, Kap. 4).

In der Phänomenologie werden diese Motive im Hinblick auf eine gewisse Unerforschlichkeit des Anderen für das Ich ausgedeutet. Edmund

Husserl spricht von einer »originalen Unzugänglichkeit« (V. *Cartesianische Meditation*, § 52) des Anderen und Emmanuel Levinas von einem steten »Abstand zwischen mir und dem Anderen« (Levinas 1987, 45).

Die Hervorhebung der Herkunft des Paradigmas der Begegnung aus dem Judentum bezeichnet lediglich eine Entstehungsgeschichte und keine Sonderwelt. Das Judentum gehört genauso zu Europa wie andere Traditionen auch. Die jüdische Dialogphilosophie ist ebenso eine Philosophie, d.h. eine Wissenschaft vom Logos und der Welt, wie die Platons, Kants oder Hegels. Daher ist es möglich und sinnvoll, das Paradigma der Begegnung mit einem anderen, in gewisser Hinsicht konkurrierenden Paradigma zu kontrastieren. Es handelt sich um den Kontrast zwischen den Paradigmen der Begegnung in der Dialogphilosophie und der Zuschreibung in der analytischen Philosophie. Beiden geht es um eine Antwort auf die Frage nach der Eigenart einer Person (vgl. dazu: Schnell 1999).

Ausgehend von der seit John Locke diskutierten Frage nach dem Zusammenhang personaler Bewusstseinskontinuität und menschlicher Körperlichkeit vertritt Peter F. Strawson die These, dass man jedem Individuum des Typus der Person »sowohl Bewußtseinszustände als auch körperliche Eigenschaften zuschreibt bzw. zuschreiben kann.« (Strawson 1972, 134) Der Akt der *Zuschreibung* beinhaltet, dass eine Person ein Einzelding ist, über das man spricht. Eine Person tritt damit als dritte Person Singular (Er/Sie/Es) auf. Im Licht der Zuschreibung ist jemand etwas, *über* den, die oder das man spricht. Der Dialog, also die Perspektive, dass ich nicht nur über jemanden, sondern *mit* jemandem spreche und dieser damit als ein Du erscheint, ist im Rahmen der Zuschreibung nicht zu realisieren. Interessanterweise haben die von Strawson ausgehenden Überlegungen auch in die Ethik hineingewirkt. Über Richard M. Hares Studien zur Bioethik (vgl.: Hare 1993) wurden sie von dessen Schüler Peter Singer aufgegriffen. Bekanntlich schlägt Singer vor, durch die Zuschreibung sog. »moralisch relevanter Eigenschaften wie Rationalität, Selbstbewußtsein, Bewußtsein, Autonomie, Lust- und Schmerzempfindung« (Singer 1994, 196f.) den sog. Wert des Lebens von Föten, Tieren und Menschen mit Behinderungen zu bestimmen. Es ist wohl kein Zufall, dass die ethischen Aussagen, die aus dem Zuschreibungsparadigma resultieren und jene, die aus dem Begegnungsparadigma erfolgen, einander zum Teil erheblich widersprechen (vgl.: Schnell 2017, 153ff.).

Die radikalste Version des Paradigmas der *Begegnung* vertritt Emmanuel Levinas. Daher spricht er auch von einer »Ethik der Begegnung.« (Levinas 1995, 270) Der angesprochene Andere ist darin kein zu begreifendes Sein, sondern ein Du. »Der Nächste ist nicht erst Gegenstand des Verstehens und dann Gesprächspartner. ... Eine Person verstehen heißt schon mit ihr zu sprechen.« (ebd., 17) Der Andere ist immer jemand, mit dem ich spreche und nicht etwas, über das ich spreche.

Mit dieser Bestimmung ist eine Entscheidung über das Verhältnis zwischen den beiden Paradigmen gefallen. Zuschreiben und Begegnen sind demnach keine Optionen, die wahlweise realisiert werden können. Vielmehr ist die Begegnung das Fundament aller weiteren Maßnahmen. Anders formuliert: die Zuschreibung von Eigenschaften an die Adresse einer Person setzt die Begegnung mit ihr voraus (vgl.: Schnell 2017, 96ff., 153ff.). Mit Sartre gesprochen: »Man *begegnet* dem Anderen, man konstituiert ihn nicht.« (Sartre 1991, 452) Insofern ist der Andere immer zuerst ein Du bevor er ein Er/Sie/Es werden kann. Das gilt auch für die verweigerte Begegnung, die Zygmunt Bauman eine »Vergegnung« nennt und »die eine Begegnung ist, die vortäuscht, keine zu sein.« (Bauman 1995, 229)

Wenn in diesem Band also von »Begegnungen mit Künstlicher Intelligenz« die Rede sein soll, so ist es durchaus diese emphatische Traditionslinie, die im Hintergrund steht. Über KI soll also nicht in Hinsicht auf Fragen der Zuschreibung gesprochen werden, sondern die Sache der KI soll in Hinsicht auf die Begegnung gedacht werden, wobei natürlich nicht von Anfang an vorausgesetzt ist, dass es hier eine Gleichwertigkeit (in ontologischer oder ethischer Hinsicht) mit der Begegnung von Menschen untereinander gibt. Ausgehend von Bubers Motto: »Der Mensch wird am Du zum Ich« (Buber 1983, 37), und seiner weitergehenden Aussage, dass es »kein Ich an sich«, »sondern nur das Ich des Grundworts Ich-Du und das Ich des Grundworts Ich-Es« gibt (ebd., 10), kann man danach fragen, was passiert, wenn man in dieser Situation ein Relatum durch eine KI oder einen Roboter ersetzt, der oder die aufgrund seines Status in der Interaktion nicht mehr zweifelsohne als Es charakterisiert werden kann, sondern zumindest einen Anspruch auf das Du erhebt.

Künstliche Intelligenz

Als Geburtsstunde der KI-Forschung als Disziplin wird die Dartmouth Conference von 1956 angesehen, an der viele der in den folgenden Jahren für das Feld maßgebenden Forscher teilnahmen, so z.B. John McCarthy, Marvin Minsky, Claude Shannon sowie Allen Newell und Herbert A. Simon. Im Förderungsantrag ist das Grundparadigma der KI-Forschung prägnant auf den Punkt gebracht: »The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it« (McCarthy et al. 1956: 1). Dieses Programm ist bis heute disziplinbestimmend. Auch der große KI-Kritiker Joseph Weizenbaum spricht 20 Jahre später davon, »daß es keinen Bereich des menschlichen Daseins gibt, der nicht maschinell erfaßbar

ist.« (Weizenbaum 1978, 274) Auch in aktuellen Publikationen ist die Definition der KI sehr nah an der der Dartmouth Conference: »Artificial Intelligence (AI) seeks to make computers do the sorts of things that minds can do« (Boden 2018, 1). Man sieht allerdings, dass die Definitionen darin variieren, was genau simuliert oder künstlich nachgebildet werden soll (»learning«, »intelligence«, »menschliches Dasein«, »what minds do«). So unterscheiden Russel und Norvig in ihrem modernen Standardwerk *Artificial Intelligence: A Modern Approach* zwischen vier Vorbildern für die KI. Nachgeahmt werden sollen wahlweise: menschliches Denken, menschliches Handeln, rationales Denken oder rationales Handeln. Wobei sie »rationales Handeln« als die beste der vier Möglichkeiten ansehen und sich für diese Zielsetzung als »Standardmodell« entscheiden (Russel und Norvig 2021, Kap. 1). Unter der Gefahr, den verschiedenen Definitionen eine weitere hinzuzufügen, können wir *Künstliche Intelligenz* in dieser Einleitung als Sammelbegriff für Prozeduren verstehen, bei denen Maschinen menschenähnliche (d.h. eventuell auch »rationaler als menschliche«) Intelligenzleistungen erbringen.

Ähnlich weit gestreut wie die Zieldefinitionen sind heute die Arten und Weisen, auf denen diese Ziele erreicht werden sollen. In der ersten Welle der KI-Forschung, später von John Haugeland »good old fashioned AI« (GOFAI) genannt (vgl.: Haugeland 1985), waren es vor allem symbolisch-diskret operierende Maschinen, für die Newells und Simons »physical symbol system« das Paradigma gab, die das Ziel der Dartmouth Conference erreichen sollten (vgl.: Newell und Simon 1976). Nachdem der versprochene rasante Erfolg im Erreichen menschenähnlicher Leistungen in vielen Bereichen ausblieb, folgte der sogenannte KI-Winter, eine Periode in der die zu Beginn enorme Forschungsförderung zunehmend ausblieb und das GOFAI-Programm, nicht zuletzt aufgrund der grundlegenden Kritik an seiner Durchsetzbarkeit seitens Hubert Dreyfus' (vgl.: Dreyfus 1992), nahezu tot schien. In der zweiten KI-Welle erlebte das eigentlich seit Beginn der KI-Forschung bestehende, aber lange Zeit in den Hintergrund gerückte Programm des Konnektionismus ein Revival. War die funktionalistische Prämisse des GOFAI-Programms, dass der *mind* unabhängig von seiner physischen Implementierung im menschlichen Gehirn, rein auf Grundlage seiner symbolischen Operationen nachgebildet werden könnte, nahm sich das Programm des Konnektionismus nun gerade die Netzwerkarchitektur des menschlichen Hirns zum Vorbild. Hatten frühe Versionen der *artifical neural networks* (ANNs) meist nur einen (simulierten) Input- und einen Output-Layer, weisen heutige Programme eine erhebliche Tiefe auf und werden daher oft unter dem Stichwort *deep neural networks* (DNNs) beschrieben. Die meisten schlagzeilen-machenden Erfolge der letzten 20 Jahre wurden von solchen künstlichen neuronalen Netzen vollbracht, auch wenn es nach wie vor und gerade in spezialisierten Bereichen unzählige auf

GOFAI-Technologie beruhende Systeme gibt und auch viele der heutigen *machine learning*-Programme GOFAI-Technologie nutzen (vgl. Boden 2018, 40).

Zusätzlich zu diesen beiden Paradigmen der Informationsverarbeitung in KI-Systemen gibt es heute (manche sprechen daher schon von einer dritten KI-Welle) noch »evolutionary programing«, »cellular automata«, »dynamical systems« sowie hybride Systeme, die verschiedene Paradigmen vereinen (vgl. ebd.: 5).

Entscheidend für die in diesem Band vorgeschlagene Perspektive sind aber nicht so sehr die technologischen Architekturen und deren theoretische Möglichkeiten und Unmöglichkeiten im Erreichen der verschiedenen Zielsetzungen, sondern vielmehr das Eingehen all dieser verschiedenen Systeme in unsere Lebenswelt und unsere Art und Weise mit ihnen zu interagieren.

Diesbezüglich kann technikgeschichtlich zwischen Anwendungen von Technik unterschieden werden, die im Auftrag des Menschen bloße Ausführungen realisieren, und Anwendungen, die von sich aus einer Kommunikation etwas hinzufügen können.

<i>Position</i>	<i>Funktion der</i>	<i>Beispiel</i>	<i>Autor:innen</i>
1. Technik	Kompensation biologischer Mangelausstattung	Hammer	Platon, Gehlen
2. Technologie	Vorschriften, Nachrichten	Navigationssystem Flugzeug	Heidegger-Schüler: Anders, Marcuse, Jonas
3. Körpertechnik, Selbstdarstellung	Formierung des Körpers	Militär, Etikette	Mauss, Goffman, Bourdieu, Deleuze
4. Phänomenotechnik	Organisation von Erfahrungsweisen	Fotoapparat	Bachelard, Blumenberg, Waldenfels
5. Robotik, Künstliche Intelligenzen	Dienstleistungskommunikation mit Freund und Helfer, Informationsbeschaffung	Tee anreichen, Kommunizieren	Wagner, »Mensch-Maschine Interaktion«

Künstliche Intelligenz tritt in allen Bereichen auf, vermehrt, und für die hier eingenommene Perspektive besonders interessant, aber als letztere Anwendung.

Ein weiteres, hier hilfreiches Unterscheidungsschema ist das von Don Ihde in seiner technikphilosophisch inspirierten »Postphänomenologie« ausgearbeitete. Ihde analysiert verschiedene Beziehungen, in denen Menschen mit Technik stehen können: In »alterity relations« gibt es eine

Bezugnahme auf ein technisches Artefakt als (Quasi-)Anderer. Diese werden unterschieden von »embodiment relations« (z.B. das Tragen einer Brille, eines Hörgeräts oder auch eines mit einer gewissen Intelligenz ausgestatteten »wearables«, wie z.B. einer Smartwatch), »hermeneutic relations«, in denen die Technik gelesen und als Weltausschnitte repräsentierend gedeutet wird (z.B. MRT-Bilder, Thermostate etc., auch hier können die zu lesenden Artefakte durchaus »intelligent« sein) und »background relations« (z.B. das Heizungssystem in unserem Haus, aber auch der Algorithmus der unsere Twitter-Feeds generiert) (vgl.: Ihde 1990, Kap. 5).

Wie wir sehen, kann in allen Arten der Beziehung KI eine Rolle spielen, besonders virulent wird die Frage nach der Begegnung in Bezug auf die oben skizzierte Tradition der Begegnungsphilosophie aber natürlich da, wo die Beziehung zu einer »alterity relation« wird.

Es sind aber in einem gewissen Sinne durchaus alle Arten der Relationen mit KI als eine Art der Begegnung mit dieser zu verstehen, insofern sie nämlich gerade aus der Perspektive der Lebenswelt, in der uns Technik zu allererst begegnet, verstanden werden. Alle Arten unserer Bezugnahme auf Technik finden in der Lebenswelt statt, sie nutzen die digitale Transformation der Lebenswelt, indem sie an sie anknüpfen und sie vertiefen. Wie ist diese Begegnung »in der Lebenswelt« genauer zu verstehen?

Begegnungen mit Künstlicher Intelligenz – eine weitere Kränkung des Menschen?

Husserl definiert die Lebenswelt als die »wirklich anschauliche, wirklich erfahrbare und erfahrene Welt, in der sich unser ganzes Leben praktisch abspielt.« (Husserl 1976, 51) In dieser grundsätzlichen Funktion ist die »Lebenswelt ein allgemeines Fundament für alle Sinnbildung.« (Waldenfels 1985, 17) Husserl unterscheidet zwischen Lebenswelten im Plural, die kulturrelativ und hier anders als dort sein können, und der einen Lebenswelt, die, formaler betrachtet im Sinne eines »abstrakt herauszupräparienden Weltkern(s)« (Husserl 1976, 136), jenes allgemeine Fundament für alle Sinnbildungen, im Alltag wie in den Wissenschaften, ist. In dieser allgemeinen und zugleich speziellen Funktion hat die Lebenswelt etwas Verlässliches und Beständiges. In diesem Sinne betont Husserl, dass die Lebenswelt »in ihrer eigenen Wesensstruktur, in ihrem eigenen Kausalstil ungeändert [...] bleibt, was immer wir kunstlos oder als Kunst tun.« (ebd., 51) Auch die Technisierung durch die modernen Wissenschaften, an die Husserl dachte, können die Lebenswelt zwar kulturell verändern, aber nicht in ihrer Fundamentfunktion. Es stellt sich die Frage, ob diese Ansicht noch immer von Geltung ist.

Dem Hinweis Sigmund Freuds, dass der Mensch drei große Kränkungen habe hinnehmen müssen, weil sein Planet nicht der Mittelpunkt des Weltalls ist, seine Abstammung kein Schöpfungsvorrecht beinhalte und er selbst nicht Herr im eigenen Seelenhaushalt sei (vgl.: Freud 1983, 226), fügte Pierre Bourdieu eine vierte Kränkung hinzu. Demnach sind die Haltungen und Stellungnahmen des Menschen gegenüber der Welt keineswegs spontane Leistungen des Verstandes, sondern vielmehr gesellschaftlich produziert (vgl.: Bourdieu 1989, 7). Auf diese Befunde folgt möglicherweise noch eine fünfte Kränkung. Die Lebenswelt wird in ihrer Fundamentfunktion sehr wohl durch die Digitalisierung und die Begegnungen mit Künstlicher Intelligenz deutlich verändert. Seit der Romantik muss sich der moderne Mensch immer wieder gegen die Maschine behaupten, was ihm jedoch nie ganz gelungen ist. Sein Narzissmus wird ein fünftes Mal gekränkt, wenn die ihm, als natürlich geborenem Wesen, bisher exklusiv vorbehaltene Domäne des rationalen Denkens und Handelns nun künstlich herstellbar wird.²

Digitalisierung ist als eine qualitativ neue Dimension der Technisierung zu verstehen, als die »mit einer Automatisierung verbundene Transformation der Welt in quantifizierbare Daten.« Sie stellt damit »auch einen Angriff auf die Materialität der Lebenswelt dar.«³ (Schnell/Dunger 2019, 13) *Begegnungen mit Künstlicher Intelligenz* vertiefen diese Transformation der Lebenswelt.

In Weiterführung der Analysen Husserls deutete Hans Blumenberg das Telos der Technisierung als das Zurücksinken technischer Artefakte in das »Universum der Selbstverständlichkeiten« der »Lebenswelt«. Dadurch wird »jene Sphäre, in der wir *noch* keine Fragen stellen, identisch [...] mit derjenigen, in der wir keine Fragen *mehr* stellen.« Die »Technisierung [...] produziert [...] das scheinbar Unproduzierbare [...], nämlich Selbstverständlichkeit.« (Blumenberg 2015, 190) Begegnungen mit

- 2 Diese Idee geht auf Catherine Malabou zurück, die im Englischen Postscript zu ihrem großartigen Buch *Morphing Intelligence* die Konfrontation mit Künstlicher Intelligenz im Hinblick auf Freuds Aufzählung – ohne Bourdieus Ergänzung mitzuzählen – schon als eine mögliche vierte Kränkung beschrieben hat (vgl. Malabou 2019: 162ff.).
- 3 Im Digitalisierungsdiskurs wird oft zwischen »digitisation« und »digitalisation« unterschieden, wobei ersteres die Produktion digitaler Daten oder die Umwandlung analoger Daten in digitale bezeichnet und letzteres die Nutzung dieser Daten durch Akteure und die Auswirkungen dieses Prozesses auf gesellschaftliche Wirklichkeiten. Mit der hier vorgeschlagenen Definition von Digitalisierung sind diese beiden Aspekte umfasst, unter fernerer Berücksichtigung des Umstandes, dass die Daten immer aus der Lebenswelt erhoben werden (1. Sinn der Transformation der Lebenswelt) und im Prozess der Digitalisierung es diese Lebenswelt selbst ist, die verändert wird (2. Sinn der Transformation).

Künstlicher Intelligenz im Zeichen der Digitalisierung vertiefen diesen Prozess, eventuell sogar in Bezug auf die intersubjektive Struktur unserer Lebenswelt, insofern wir mit technischen Artefakten eben in »alterity relations« treten. Somit gerät Husserls These, ja tiefe Überzeugung, dass unsere Lebenswelt in ihrer Wesensstruktur, in ihrem eigenen konkreten Kausalstil ungeändert bleibt, was immer wir kunstlos oder als Kunst tun, ins Wanken. Technisierung und Digitalisierung verändern unsere Lebenswelt geradezu ins Fraglose. Diese Veränderungen drohen unbemerkt und unthematisiert zu bleiben, gäbe es nicht Schlüsselerfahrungen des Umbruchs, in denen das Staunen über Künstliche Intelligenz aufblitzt.

Das Auslösen technischer Effekte erfolgt heute nicht mehr durch Schalter, Knöpfe oder Regler, sondern zunehmend via Sprachsteuerung und es versetzt uns damit, aufgrund der Anredeform, in eine Art der Begegnung mit Künstlicher Intelligenz. Bereits vor über 30 Jahren befragte Hans Blumenberg die Situation der Begegnung des Menschen mit künstlichen, aber menschenähnlichen Automaten. Im Ausgang von Husserls Theorie der Fremderfahrung heißt es: »Wie aber erreichen ... die menschengestaltigen Organismen, von ihm [dem Menschen, die Verf.] als Seinesgleichen anerkannt zu werden?« (Blumenberg 1989, 442) Diese Frage beschäftigt noch die Gegenwart, die technische Geräte kennt, die im letzten Jahrhundert, wenn überhaupt, nur im spekulativen Denken gefunden wurden.

Generative Pre-trained Transformer 3 (GPT-3) ist ein autoregressives Sprachmodell, das Deep Learning benutzt, um menschenähnliche Texte zu erstellen und zu verbalisieren. Bekanntlich ist es möglich, dass ein Mensch mit GPT-3 einen inhaltlichen Austausch über anspruchsvolle Themen wie Liebe, Freiheit, Datenschutz oder gar Künstliche Intelligenz führen kann. Mithilfe einer quantifizierten Datenbasis kann das System, das in einer Kommunikation wahrscheinlich nächste Wort und den wahrscheinlich nächsten Satz berechnen. Auf die allmähliche Verfestigung der Gedanken beim Reden folgt mit hoher Wahrscheinlichkeit irgendwann der Name Heinrich von Kleists und der Hinweis auf seinen traurigen Selbstmord im Jahre 1812. Dieser Austausch wird zur Norm für ein sogenanntes gutes Gespräch, zumal wenn GPT-3 den Menschen darüber belehrt, dass sich Kleist mit seiner Pistole bereits 1811, also ein Jahr zuvor, erschossen hat. Von hier aus sind dann komplexe Interaktionen, die gar verantwortungsvolle Entscheidungen beinhalten, denkbar.

Die verborgene, sich als ethisch ausgebende Implikation dieser Möglichkeiten hat Pierre Lévy bereits 1997 in seinem Buch *Die kollektive Intelligenz* skizziert. Demnach wird das Internet, dass 1997 für den Privatgebrauch erst seit wenigen Jahren ausgelegt war, eines Tages alles Wissen der Welt bereithalten und einen Zugriff auf all dieses Wissen für jeden zu jeder Zeit und von jedem Ort aus ermöglichen. Jeder darf an einem allseitigen Wissensaustausch teilnehmen. Da offizielle Zertifikate

und Zeugnisse nicht zählen, hat jeder auch etwas potentiell Wissenswertes anzubieten. Niemand ist vom Austausch ausgeschlossen. Die Welt ist eine virtuelle Agora, die Menschheit ein kollektives Cogito bzw. ein *cogitamus*. Der Mensch erwirbt eine Identität des Wissens, die anschlussfähig an Andere ist. Wissen ist das, was digitalisier- und im Netz kommunizierbar ist. Materialität der Welt, Leiblichkeit des Menschen und die Andersheit des Anderen werden allmählich irrelevant (vgl. dazu ausführlich: Schnell 2002). Diese 25 Jahre alte Vision Lévys gehört noch dem posthumanistischen Stadium der KI an, sie ist in diesem Punkt nicht mehr aktuell. Sie liefert allerdings die ethisch-ontologischen Aspekte des Cyberspace, die heute für nichttriviale Kommunikationen mit einer KI wie GPT-3 in Anspruch genommen werden.

»Computer können juristische Entscheidungen treffen und psychiatrische Urteile fällen. Sie können auf viel ausgesuchtere Weise Münzen werfen als das geduldigste menschliche Wesen.« (Weizenbaum 1978, 299) Tragen sie in diesen Fällen auch die Verantwortung für diese Entscheidungen und Handlungen? Wer haftet? Wer trägt die Schuld? Die Maschine? Ihr Eigentümer? Der Hersteller? Das Produkthaftungsgesetz? Weizenbaum folgert aus all dem, dass »die relevanten Probleme weder technischer noch mathematischer, sondern ethischer Natur sind.« (ebd., 300)

Begegnungen mit Künstlicher Intelligenz – Querschnitte

Hans Blumenberg hat die Verbindung der Frage nach der Technik mit der Problematik der Wahrheit untersucht. Demnach herrscht in der europäischen Antike die Auffassung vor, dass die Wahrheit in der Natur der Dinge fix vorhanden ist. Technik ist demgegenüber das bloß Künstliche und Scheinhafte. Im Idealfall ahmt es die Natur nach (*ars imitatur naturam*). »Wer ein Haus baut, tut genau das, was die Natur tun würde, wenn sie Häuser sozusagen wachsen ließe.« (Blumenberg 2015, 86f.) Mit der Neuzeit wird die Natur zum bloßen Rohstoff, wie es bei Locke heißt, so dass die Wahrheit allmählich im Bereich des Selbstgemachten zu finden ist. In der jüngeren Gegenwart setzen Autoren wie Baudrillard in der Wahrheitsfrage ganz auf die Mächte der Simulation (vgl.: Schnell 2020, 184ff.).

Aktuelle Künstliche Intelligenz hat die Eigenart, sich dem anzunähern, was man als einen lebensweltlichen Akteur bezeichnen kann. Damit ist jedes Wesen gemeint, das mit Anderen in der Welt existiert und dabei einen Selbst-, Welt- und Fremdbezug realisiert (vgl. dazu: Schnell 2017, 134ff.). Zu diesen Wesen zählen Menschen, vermutlich die meisten Tiere und vielleicht auch gewisse Pflanzen. Des Menschen Gehirn

oder seine Gene sind keine lebensweltlichen Akteure. Weil sich viele Künstliche Intelligenzen hingegen dem Charakter des lebensweltlichen Akteurs annähern und es daher zu Begegnungen mit dem Menschen kommt, verbinden sich in der Diskussion drei Querschnitte miteinander: Intersubjektivität, Technik und Lebenswelt. Auf die Wahrheitsgehalte dieser drei Querschnitte und auf ihre Beziehungen untereinander haben die Begegnungen mit Künstlicher Intelligenz deutliche Auswirkungen.

Aufbau des Bandes

Begegnungen mit Künstlicher Intelligenz produzieren Selbstverständlichkeiten und Normalitäten, die als solche zu unbefragten Beständen der Lebenswelt werden können. Die Beiträge des Bandes untersuchen Formen und Inhalte dieser Begegnungen. Was bedeutet es, mit einer Maschine zu musizieren, sich von einem Roboter pflegen und berühren zu lassen, mit einer KI ein intelligentes Gespräch zu führen?

Der Band wird eröffnet mit einem Beitrag *Oliver Müllers*, der sich der Frage nach *maschineller Alterität* widmet. Anknüpfend an die Dialektik aus Andersheit und Ähnlichkeit, mit der der Mensch kulturgeschichtlich sein maschinelles Gegenüber imaginiert und anthropologisch zu definieren versucht hat, fragt der Beitrag danach, welche spezifische Andersheit und welche Ähnlichkeiten der, im Ausgang von Levinas und Ricœur »alteritätsbedürftige« Mensch in der Maschine finden kann.

Burkhard Liebsch analysiert in seinem Beitrag *Kommunikative Verwerfungen zwischen digital natives und analogen Aborigines*. Im Spannungsfeld des Bedürfnisses als Partner eines Dialogs nicht restlos in seinem digitalisierbaren Bild aufzugehen zu wollen und der Verlegenheit anzugeben, worin dieser Rest bestehen möge, gilt es die Technik weder von vorneherein zu verteufeln noch Kommunikation von Anfang an technizistisch zu betrachten, sondern vielmehr einen richtigen Gebrauch der Technik zu erlernen.

Jonathan Hart und *Martin Feißt* werfen in ihrem Beitrag einen soziologischen Blick auf ihre Selbsterfahrungen in der Begegnung mit dem Programm GPT-3. Dass das Programm »ganz losgelöst von verkörperten, wahrnehmungsgeleiteten und emotionsbegleiteten Entitäten« (S. 70) Sprachverständnis und Intelligenz zeige, müsse die Soziologie nicht überraschen, sei diese doch gerade mit dem Paradigma angetreten, Soziales durch Soziales erklären zu können. Der alteritätsbedürftige Mensch ist soziologisch gesehen vor allem ein Interaktionssucher, und Programme wie GPT-3 sind intelligent genug, um ihm diese Interaktion zu bieten und ihn so am Sozialen teilhaben zu lassen, also soziale

Relationen zu bilden, die der Frage danach, was die Relata denn nun seien, vorgängig sind.

Martin W. Schnell und *Lukas Nehlsen* gehen in dem darauffolgenden Beitrag ebenfalls auf empirisches Gesprächsmaterial mit GPT-3 als soziale Tatsache ein. Mithilfe der qualitativen Inhaltsanalyse wird ein Gespräch mit dem Programm über Liebe analysiert. Das Gespräch zeigt sich dabei vor allem als Meinungsaustausch, der (noch) nicht die Ebene eines schöpferischen Dialogs erreicht.

In seinem Einzelbeitrag knüpft *Lukas Nehlsen* daran an und widmet sich sprachphilosophisch dem Phänomen des Gesprächs mit einer Chatbot-KI. Erlangt Sprache, wie Maurice Merleau-Ponty behauptet, nur durch leiblichen Ausdruck ihren Sinn, wie kann dann die Äußerung einer körperlosen KI trotzdem als sinnvoll wahrgenommen werden?

Martin W. Schnell geht in seinem Beitrag von Bubers Diktum, der Mensch werde erst am Du zum Ich, aus und fragt danach, was passiert, wenn in der Ich/Du-Beziehung eines der Elemente durch einen Roboter ersetzt wird. In Rekapitulation der Anerkennungsthematik in der Folge von Hegel und in Auseinandersetzung mit Ian McEwans Roman *Maschinen wie ich* fragt Schnell danach, ob es möglich ist, einen Roboter als meinesgleichen anzuerkennen.

Renate Buschmann eröffnet mit ihrem Beitrag den Teil des Bandes, der sich mit Begegnungen mit KI auf dem Feld der Kunst und Kultur auseinandersetzt. Sie rekapituliert diverse künstlerische Projekte, in denen KI nicht so sehr als eigenständig kreativer Akteur wichtig wird, sondern eher als Kooperationspartner des Menschen, sowohl in Gestalt von Künstler:innen als auch von Rezipient:innen. Die interessante Frage sei daher eher nicht »Kann KI Kunst?«, sondern vielmehr »Kann aus KI Kunst werden?« (S. 165).

Christian Grüny Beitrag wirft einen philosophischen Blick auf diverse Begegnungen mit KI im Umkreis der Musik. Anknüpfend an Ihdes oben genannte Unterscheidung der Mensch-Technik-Relationen verfolgt Grüny diese sowohl im Kompositionen- als auch im Improvisationsprozess. Dabei zeigt sich unter anderem, dass der Einbruch von Maschinen in die Musik nicht der »Einbruch von etwas ganz anderem in einen ansonsten homogenen und vor allem zutiefst und ausschließlich menschlichen Schaffensprozess« (S. 179) ist und dass es nicht die Gleichheit mit dem kreativen Können des Menschen sei, die die Maschinen für die Musik interessant mache, sondern gerade ihre Seltsamkeit.

Die letzten beiden Texte behandeln Begegnungen mit Künstlicher Intelligenz auf dem Feld der Gesundheit, besonders der Pflege. *Alexander Hochmuth* gibt in seinem Beitrag einen Überblick über Einsatzbereiche von KI in der Pflege und fragt dabei, wie der Einsatz von KI mit den Prinzipien eines guten pflegerischen Handelns, insbesondere im Umgang mit Menschen mit demenziellen Erkrankungen, vereinbar sei.

Erik Dzwiza-Ohlsen wirft im letzten Beitrag des Bandes einen leibphänomenologischen Blick auf das konkrete Beispiel des mit KI ausgestatteten »Kuschelroboters« PARO und seiner Rolle in der Lebenswelt von Menschen mit Demenz.

Literatur

- Baumann, Zygmunt (1995): *Postmoderne Ethik*, Hamburg: Edition Hamburg.
- Blumberg, Hans (1989): *Höhlenausgänge*, Frankfurt am Main: Suhrkamp.
- Blumberg, Hans (2015): *Schriften zur Technik*, Berlin: Suhrkamp.
- Boden, Margaret (2018): *Artificial Intelligence. A Very Short Introduction*. Oxford/New York: Oxford University Press.
- Buber, Martin (1983): *Ich und Du*, Heidelberg: Lambert Schneider.
- Bourdieu, Pierre (1989): *Satz und Gegensatz*, Berlin: Wagenbach.
- Dreyfus, Hubert L. (1992): *What Computers Still Can't Do*. Cambridge, MA: MIT Press.
- Feuerbach, Ludwig (1983): *Grundsätze der Philosophie der Zukunft*, Frankfurt am Main: Klostermann.
- Freud, Sigmund (1983): *Vorlesungen zur Einführung in die Psychoanalyse*, Frankfurt/M: Fischer.
- Gurwitsch, Aaron (1977): *Die mitmenschlichen Begegnungen in der Milieuwelt*, Berlin/New York: de Gruyter.
- Hare, Richard M. (1993): *Essays on Bioethics*, Oxford: Clarendon Press.
- Haugeland, John (1985): *Artificial Intelligence. The Very Idea*. Cambridge, MA: MIT Press.
- Husserl, Edmund (1976): *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, Den Haag: Nijhoff.
- Ihde, Don (1990): *Technology and the Lifeworld. From Garden to Earth*. Bloomington: Indiana University Press.
- Levinas, Emmanuel (1987): *Totalität und Unendlichkeit*, Freiburg/München: Karl Alber.
- Levinas, Emmanuel (1995): *Zwischen uns. Versuche über das Denken an den Anderen*, München: Hanser.
- Löwith, Karl (2013): *Das Individuum in der Rolle des Mitmenschen*, Freiburg/München: Karl Alber.
- Malabou, Catherine (2019): *Morphing Intelligence. From IQ Measurements to Artificial Intelligence*. New York: Columbia University Press.
- McCarthy, John; Minsky, Marvin, et al. (1955): *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Online verfügbar unter: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html> (letzter Abruf: 16.02.2022).
- Newell, Allen; Simon, Herbert A. (1976): »Computer Science as Empirical Enquiry: Symbols and Search«, *Communications of the Association for Computing Machinery* 19 (März 1976).

- Priddat, Birger P. (2022): »Wie werden wir lernen uns zu Automaten zu verhalten?«, in: Priddat, B. P. / Kovce, P. (Hg.) (2022): *Selbstverwandlung. Das Ende des Menschen und seine Zukunft*. Marburg: Metropolis.
- Rosenzweig, Franz (1988): *Der Stern der Erlösung*, Frankfurt am Main: Suhrkamp.
- Russel, Stuart J.; Norvig, Peter (2021): *Artificial Intelligence: A Modern Approach*. 4. Auflage. Hoboken: Pearson.
- Sartre, Jean-Paul (1991): *Das Sein und das Nichts*, Reinbek bei Hamburg: Rowohlt.
- Scheler, Max (2000): *Der Formalismus in der Ethik und die materiale Wertethik*, Bonn: Bouvier.
- Schnell, Martin W. (1999): »Narrative Identität und Menschenwürde«, in: Breitling, A. et al. (Hg.) (1999): *Das herausgeforderte Selbst*, Würzburg: Königshausen und Neumann.
- Schnell, Martin W. (2002): »Ideologie und Anthropologie. Zur Widerkehr des leiblosen Geistes«, in: Greving, H. / Gröschke, D. (Hg.) (2002): *Das Sisyphos-Prinzip*, Bad Heilbrunn: Julius Klinkhardt.
- Schnell, Martin W. (2017): *Ethik im Zeichen vulnerabler Personen*, Weilerswist: Velbrück.
- Schnell, Martin W. (2020): *Das Ethische und das Politische. Sozialphilosophie am Leitfaden der Vulnerabilität*, Weilerswist: Velbrück.
- Schnell, Martin W. / Dunger, Chr. (Hg.) (2019): *Digitalisierung der Lebenswelt. Studien zur Krisis nach Husserl*, Weilerswist: Velbrück.
- Singer, Peter (1994): *Praktische Ethik*, Stuttgart: Reclam.
- Strawson, Peter F. (1972): *Einzelding und logisches Subjekt*, Stuttgart: Reclam.
- Waldenfels, Bernhard (1985): *In den Netzen der Lebenswelt*, Frankfurt am Main: Suhrkamp.
- Weizenbaum, Joseph (1978): *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt am Main: Suhrkamp.
- Wojcieszuk, M.A. (2010): *Der Mensch wird am Du zum Ich. Eine Auseinandersetzung mit Der Dialogphilosophie des XX. Jahrhunderts*, Freiburg: Centaurus.