

MARTIN W. SCHNELL

Anerkennung einer Künstlichen Intelligenz als meinesgleichen?

Beitrag zur Maschinenethik

»Anerkennen ist etwas anderes als naive Gewißheit.«

Edmund Husserl (1966, 39)

Von körperlosen Geistern zu künstlichen Menschen

Hubert Dreyfus (1985) antwortete Mitte der 80er Jahre als Leibphänomenologe auf die Visionen der KI-Forschung. Er legte dar, *what Computers can't do*. Demnach ist es nicht möglich, menschliches, intelligentes Verhalten und Handeln künstlich zu simulieren. KI-Forscher wie Hans Moravec (1990) gingen irrtümlich davon aus, dass Intelligenz eine rein kognitive Angelegenheit sei, die der Vulnerabilität des Leiblichen nichts Wesentliches verdankt. Man denke auch an die Fiktion eines Gehirns im Tank und an die Spekulation über einen sich erneuernden Cartesianismus (vgl. als Beispiel: Putnam 1993, 209ff.). Dreyfus, Schüler von Merleau-Ponty, hebt demgegenüber hervor, dass Intelligenz leiblich verkörpert und verankert ist, dass sie situativ zur Geltung kommt, Unschärfen produziert und auf Vagheiten spontan eingehen kann. Der Geist hat nicht nur eine Syntax, sondern auch eine Semantik, wie Searle betont (vgl.: Searle 1986, 29ff.). Der Geist ist zudem verkörpert (embodied Mind), wie Francisco Varela (1995) sagt, und somit an die Endlichkeit des Lebens gebunden. Die gesamte Natur des Menschen lässt sich nicht restlos formalisieren oder digitalisieren (vgl. zu dieser Debatte: Schnell 2002).

Die KI-Forschung ist inzwischen von einer solchen Vision körperloser Geister und Intelligenzen abgerückt. Es macht keinen Sinn, eine Postbiologie im wörtlichen Sinne anzustreben. Heute geht es vielmehr um die Schaffung künstlicher Menschen, die den Diskurs der Vulnerabilität nicht umgehen, sondern ihn durchlaufen. Damit verändert sich Isaac Asimovs erstes Gesetz der Robotik. Die Verletzlichkeit bezieht sich nicht allein auf den Menschen, wie wir ihn bisher kannten, sondern auch auf den kommenden künstlichen Menschen.

Maschinenethik

Der vorliegende Text knüpft an das Projekt »Digitalisierung der Lebenswelt« (Schnell/Dunger 2019) an, welches nach den Dimensionen der Transformation der Lebenswelt durch die Digitalisierung fragte. Er fügt ihm die Frage hinzu, ob ein Mensch eine Künstliche Intelligenz in Gestalt eines Roboters als seinesgleichen anerkennen kann oder gar muss, und zwar im Zuge von Begegnungen zwischen Mensch und KI in der Lebenswelt. Begriff und Sache der Anerkennung werden damit in das Feld der Maschinenethik eingeführt (vgl. dazu Bendel 2019).

Die Maschinenethik ist ein komplexes Feld, auf dem gewiss nicht nur die Sozialphilosophie agiert. Sie stößt auf Technik und Ingenieurwissenschaft, die nicht nur an der Konstruktion von Robotern arbeiten, sondern auch an der Beantwortung philosophischer Fragen. Im Ausgang von der auf Emil du Bois-Reymond zurückgehenden Unterscheidung zwischen *Ignoramus* (wir wissen es noch nicht) und *Ignorabimus* (wir werden es nicht wissen) neigen Ingenieure manchmal dazu, Antworten auf unbeantwortete Fragen dem erwarteten Fortschritt zu überlassen. Aktuell vermag ein Roboter das Menschsein noch nicht perfekt zu simulieren, aber eines Tages wird es so weit sein. Damit finden unbeantwortete Fragen eine rasche Antwort. Der vorliegende Beitrag möchte zeigen, dass auf diese Weise jedoch das Fragebegehren nicht erschöpft ist und daher für ein gewisses Offenhalten von Antworten plädieren.

Postbiologisches vs. Humanoides Paradigma

<i>Postbiologisches Paradigma</i>	<i>Humanoides Paradigma</i>
Die KI vermag etwas, was ein Mensch nicht leisten kann. Z.B.: sehr komplexe Rechenaufgaben zu lösen. Das IBM-Programm Watson gewinnt das Jeopardy-Quiz, indem es seinen Gegenspieler an Wissen, Reaktionen und Assoziationen erheblich übertrifft.	Die KI vermag etwas, was ein Mensch beherrscht. Z.B.: Alltagsassistent, wie die Steuerung von Küchengeräten oder das Herabsteigen einer Treppe.

Das postbiologische Paradigma verstand die Optimierung menschlicher Intelligenz als einen Prozess, der vom leiblichen Menschen abzulösen ist. Heute, wie es in Japan heißt, geht es darum, künstliche Menschen zu schaffen. Die Maschine passt sich dem Menschen an. Auch optisch werden Roboter quasi menschlich. Sie haben Körper mit entsprechender Anatomie, ein Gesicht mit Augen, Nase, Mund, Ohren, Wimpern und einen Haarschopf. Daher die Rede von humanoiden Robotern. Gabriele

Gramelsberger bezeichnet diesen Schritt als Übergang von der »künstlichen Intelligenz« zum »künstlichen Leben.« (Gramelsberger 2019, 125) Am Forschungszentrum für Informatik in Karlsruhe wird versucht, Robotern Selbstbewusstsein beizubringen. Die leibliche Innenwahrnehmung als Basis für Handlungsentscheidungen wird simuliert. Das Schwierige daran ist, dass diese sehr heterogen ist. Eine Treppe hinaufsteigen ist etwas anderes, als ein Fahrzeug im Straßenverkehr zu steuern.

Menschliche Gefühle vs. Kalte Technik?

Wie jede technische Innovation stößt auch die Digitalisierung auf eine gesellschaftliche Alltagsmentalität in der Lebenswelt, die nicht einfach übersprungen werden kann. Deutschland ist ein Land, das in der Tradition der Romantik steht. Natur, Authentizität und Begegnung gelten als Parameter für Wahrheit und Echtheit. Entsprechend wurde bei der Einführung des Fernsehens im letzten Jahrhundert nichts weniger als der Verlust der Realität befürchtet. Bis vor kurzer Zeit galt es als soziale Störung, wenn Menschen, zumeist junge, auf ihren Handybildschirm starrten, statt sogenannte echte Menschen anzuschauen. Anders verhält es sich in Japan. Dort existiert traditionell ein Verständnis von Natur und Technik, das eine hohe Alltagsakzeptanz für Technologien begünstigt (vgl.: Lévi-Strauss 2012). Japan ist daher ein Vorreiter von Digital Nursing im Bereich der Robotik.

Vor über zehn Jahren reagierte der japanische Staat auf die sogenannte Überalterung der Gesellschaft mit einer verstärkten Förderung der Entwicklung von Haushaltsrobotern. Weil es an echten und jüngeren Menschen mangelt, sollen »künstliche Menschen« (Wagner 2013, 38) anstehende Dienstleistungen, Hilfestellungen und Pflegeverrichtungen im Haushalt durchführen. Damit ist der Weg für die Lebensform posthumaner Familien geebnet. Das ist möglich, weil die Alltagsmentalität in Japan so ausgerichtet ist, dass etwa alte Menschen pflegerische Versorgungsverrichtungen, die Roboter mit und an ihnen vornehmen, als Unterstützungsleistungen erfahren und akzeptieren. Dem Digital Nursing kommt zugute, dass in Japan mechanische Servierpuppen, die Tee einschenken, und Mangamythologien verbreitet sind. In den 90er Jahren erreichte die Popsängerin Kyoko Date eine sehr große Popularität. Da sie die Hitparade anführte, interessierte sich die Öffentlichkeit auch für ihr Privatleben. Kyoko ist ein digitales Kunstprodukt, kaum anders als ein Tamagotchi. Das war dem Publikum völlig bewusst. Dennoch wollte es wissen, wie Kyoko lebt, was sie zum Frühstück isst. Somit erfand die Softwarefirma eine Lebensgeschichte für ihren Star.

Der Gebrauch von Robotern zur Assistenz im Alltag wird in Deutschland durchaus akzeptiert. Ein Roboter wird aber nur als reines Instrument

geduldet. Ein Roboter ist kein Mensch, er ist eine unpersönliche Maschine. In diesem Sinne soll er nur dort eingesetzt werden, wo es um den Einsatz von Kraft geht (z. B. Menschen aus dem Bett heben), wo das Grobe erledigt werden muss (z. B. staubsaugen) oder bei krankengymnastischen Übungen. Im Unterschied zum Roboter ist der Mensch ein Partner, der allein zu Ansprache, Berührung und Gefühl in der Lage ist. Darin liegen die eigentlich humanen Qualitäten begründet.

In dieser Ansicht wird die von der Romantik geprägte Auffassung Europas reproduziert. Sie zeigt sich im Dualismus Mensch vs. Maschine. Diese Thematik findet man um das Jahr 1800 sehr oft aufgenommen, etwa in E.T.A. Hoffmanns Geschichte vom redenden Türken in *Die Automate* (1814), in Heinrich von Kleists *Über das Marionettentheater* (1810) oder in Edgar Allen Poes *Maelzels Schachspieler* (1827). Ein Echo dieser Thematik klingt noch in Stefan Zweigs *Schachnovelle* (1938) an. Interessanterweise existieren vor 1800, also zeitlich vor der Romantik, durchaus andere Ansätze. Man denke nur an La Mettries programmatische Schrift *L'homme machine* von 1748.

Ein möglicher Transhumanismus durch eine Verbindung von Mensch und Maschine wird von deutschen Usern im Gesundheitswesen heute aus dem Bereich der häuslichen Pflege verbannt, und zwar im Lichte einer Abgrenzung von den Sitten in Japan. Für einen Europäer scheint es schwerer als für einen Japaner vorstellbar zu sein, soziale Beziehungen zu Computerspielfiguren zu unterhalten und Bordelle aufzusuchen, in denen ausschließlich Sexpuppen aus Silikon vorzufinden sind (vgl.: Wagner 2013).

Auszug aus der Forschung von Alexander Hochmuth, Interview 94 CZ (in: Schnell/Dunger 2019):

Forscher: »Sie hatten vorher im Vorgespräch anklingen lassen, dass sie sich schon mit dem auseinandergesetzt haben. Was verstehen sie unter dem Begriff Roboter?«
#00:11:14-5#

Proband: »Tja. Ich habs eigentlich nur gelesen, dass es von Japan rübergekommen ist. Und ich habe auch wie gesagt letzts eine große Reportage über Japan gelesen. Japan ist ja nun ein sehr automatisiertes Land. Japan ist nicht auf Menschen, auf zwischenmenschliche Beziehungen ausgerichtet. Ich war entsetzt was da eigentlich haben. Die heiraten sich selber. Die kaufen sich hübsche Kleider und lassen sich fotografieren. Aber sind nicht verheiratet. In Japan gibt es Männer eben halt die so in Spielsalons rumhängen, das ist der ihr Leben. Und hier gehen die Männer in den Puff, sag ich mal wenn sie Single sind. Dort gehen sie auch in eine Institution um mit einer hübschen Frau zu reden vielleicht, sie müssen dafür bezahlen und eventuell dürfen sie die Hand der Frau mal berühren. Da müssen sie auch für bezahlen und das höchste der Gefühle was sie machen dürfen ist den Kopf in den Schoß einer Frau legen. Da habe ich gedacht »was ist das für ein Land?«. Und deswegen hat es mich auch nicht gewundert als ich gelesen habe, dass die Roboter aus Japan kamen. Die werden das auch und dann werden die ja auch sehr alt dort. Und die werden eine große Anzahl von Menschen haben die sehr alt werden. Die werden diese Roboter anwenden und für die wird das komplett normal sein.«

Aufgrund der kulturellen Differenz ist die Akzeptanz der Mitwirkung von künstlichen Menschen bei der Erledigung alltäglicher Verrichtungen in Europa noch eher gering. Diese mangelnde Akzeptanz könnte in den nächsten 20 Jahren deutlich schwinden. Wir werden diesem Gedanken weiter unten, am Beispiel einer Fiktion von Ian McEwan, nachgehen.

Gefühlsarbeit

Der entscheidende Schritt auf den Weg zum künstlichen Menschen in westlichen Gesellschaften besteht in der Einsicht, dass Gefühle keine natürlichen und authentischen Äußerungsweisen des Menschen sind. Arlie Russell Hochschild plädiert deshalb dafür, besser von Gefühlsarbeit zu sprechen. Sie versteht darunter im Ausgang von Norbert Elias »soziale Konstrukte, die gesellschaftlich erlernt« (Hochschild 2006, 23) werden und somit veränderbar sind. Sie verdeutlicht Erlernen und Einsatz von Gefühlsarbeit am Beispiel der professionellen Zugewandtheit von Stewardessen gegenüber Fluggästen und am Beispiel des Selbstverständnisses von Krankenschwestern im Krankenhaus.

Beispiel: Gefühlsarbeit im Krankenhaus

Eine klassische, empirische Untersuchung von Anselm Strauss und seinen Mitarbeitern zur Gefühlsarbeit im Krankenhaus stellt heraus, dass diese (oft von Krankenschwestern ausgeübte) Gefühlsarbeit die technische Arbeit der (häufig männlich dominierten) Medizin (Untersuchung, Ermittlung von Laborwerten etc.) ergänzt und bewirkt, dass Patienten sich als Mensch und Person anerkannt fühlen (vgl.: Strauss et. Al. 1980). In der Folge dessen wirken sie bei Behandlung und Untersuchung produktiv mit.

Entsprechende Äußerungsformen der Gefühlsarbeit sind:

- Motivation (»Halten Sie durch!«).
- Wertschätzung (»Das haben Sie gut gemacht!«),
- Fürsorge (»Sagen Sie mir, wenn es Ihnen weh tut.«),
- Tröstung (»Das tut mir leid.«).

Motivieren, Wertschätzen, Fürsorgen, Trösten sind ethisch relevant, weil sie auf die Beziehungen zwischen Personen bezogen sind. Die ethische Relevanz ist aber nicht identisch mit der Gefühlsarbeit. Diese ist auf die Produktion bestimmter Inhalte ausgerichtet (Ausdauer, Güte einer Handlung, Schmerz, Unglück), die von Programmen der Emotionserkennung identifiziert und durch simuliertes Verhalten (Worte des Trostes bei

Schmerzbekundung) beantwortet werden können. Das Gefühl ist damit allein kein Kriterium zur Abgrenzung des Menschen von der Maschine.

Das Ethische geht über das Emotionale hinaus, da es der Person des Anderen gilt und nicht nur seinen Lebensäußerungen im Sinne Diltheys. Künstliche Intelligenzen können nur so kommunizieren, dass sie auf die Sprechakte des Anderen eingehen, aber nicht auch auf die Person des Anderen (vgl.: Misselhorn 2021). Maschinen empfinden keine Emotionen, sie vermögen es aber, am Menschen Gefühle im Sinne von Lebensäußerungen zu identifizieren und ihrerseits Gefühle bei Menschen auszulösen. Eine KI erkennt, *was* geäußert wird. Furcht zeigt sich im Zusammenziehen der Augenbrauen, im Heben des oberen Augenlids, dem Anspannen des unteren Augenlids und im Aufeinanderpressen der Lippen. Diese Lebensäußerungen sind digital erfassbar. Wer sich allerdings auf diese Weise fürchtet, Peter oder Paul, vermag die KI nicht zu berücksichtigen, weil sie nicht weiß, was es heißt, *jemand* zu sein. Eine sozial agierende KI in Gestalt eines Roboters »bleibt solipsistisch und führt zu einer verdinglichten Einstellung gegenüber anderen und sich selbst.« (ebd., 162) Ein »Fremdbewusstsein« vom Anderen als Anderen ist ihr »nicht gegeben.« (103)

Vom Turing-Test zur Ethik

Im Jahre 1950 entwickelte der Mathematiker Alan Turing einen nach ihm benannten Test. Ein Mensch kommuniziert demnach per Computer mit einem anderen Menschen und einem anderen Computer, also mit einer Maschine. Der Mensch hat die Aufgabe zu entscheiden, welcher seiner beiden Kommunikationspartner der andere Mensch und welcher die Maschine ist. Der andere Mensch und der Computer versuchen indes beide, den Menschen davon zu überzeugen, dass sie selbst Menschen sind. Wenn der Mensch diese Aufgabe genügend oft nicht lösen kann, Mensch und Maschine also nicht unterscheiden kann, dann könne davon ausgegangen werden, dass Maschinen gleich Menschen denken können. Heutzutage ist diese Versuchsanordnung weniger interessant aufgrund von Umstellungen in der angebahnten Interaktion.

Während der Turing-Test die Interaktion quasi verblindet, da der Mensch seine Partner nicht sieht, sondern nur deren Kommunikationsinhalte auf seinem Computer lesen kann, geht die Versuchsanordnung in dem Film *Ex_Machina* aus dem Jahre 2014 einen Schritt weiter. Der Mensch kommuniziert nun ganz bewusst mit einem humanoiden Roboter. Er kommuniziert mit ihm wie mit einem anderen Menschen, obwohl der Mensch sieht und weiß, dass dieser eine Maschine ist. Damit wäre gezeigt, dass es zwischen Mensch und Maschine zu gelingenden

intersubjektiven Beziehungen kommen kann, die der Mensch von sich aus als gelungen und erstrebenswert erlebt und bewertet. Der Roboter ist kein Instrument, sondern Partner einer Lebensform. Er hat Persönlichkeit und ist darin Popikonen wie R2 D2 weit überlegen.

Dieses Ergebnis ist durch den technischen Fortschritt bedingt (Maschinen können inzwischen Emotionen und Differenzierungen in Sinngehalten angemessener kommunizieren als dieses 1950 möglich gewesen ist), aber auch durch die Umstellung der Perspektive. Es geht nicht mehr darum, dass der Mensch die Maschine erkennt, sondern um ein Gelingen der Kommunikation mit ihr. Kommunikation ist ethisch relevant.

Hier erheben sich zahlreiche Fragen. Kann ein künstlicher Mensch, also eine KI, die in der Gestalt eines Roboters auftritt, nicht als voller Mensch anerkannt werden? Wer und/oder was ist das Ich, das am künstlichen Menschen zum Du wird? Und: wie gestaltet sich Intersubjektivität in dieser Hinsicht? Die Sozialphilosophie hat sich diese Fragen bislang noch nicht intensiv genug gestellt, weil sie in der Sache der Intersubjektivität Martin Buber oder Edmund Husserl gefolgt ist.

Ich und Du

»Der Mensch wird am Du zum Ich.« (Buber 1983, 37) Buber setzt voraus, dass sich als Ich und als Du zwei Wesen begegnen, die Menschen sind, und das heißt damit auch, vulnerable, leibliche und endliche Wesen. Diese sind konstitutiv für das Personsein des Anderen. »Wer man ist, wenn man mit jemanden spricht, bestimmt sich nicht unabhängig davon, zu wem man spricht, auf wen man hört und wem man sich schweigend verschließt.« (Waldenfels 2019, 123) Es muss somit einen Unterschied bedeuten, ob der Andere überhaupt ein Mensch ist oder nicht.

Wir geraten mitten in die Problematik der Digitalisierung, wenn wir an diese Formulierung die Frage anschließen: »Was bedeutet es, wenn man in einer Ich/Du-Beziehung eines der beiden Elemente durch einen Roboter ersetzt.« (Schnell/Dunger 2019, 196) Kann das Ich selbst zur Maschine werden?

Herausforderung

Die wesentlich neue Herausforderung besteht in einer komplett neuen Frageperspektive, da all die fundamentalen Aspekte und Qualitäten, die die Sozialphilosophie seit dem späten 18. Jahrhundert aufgegeben hat, um die Beziehung zwischen Ich, Du und dritter Person zu beschreiben, grundsätzlich neu überdacht werden müssen. Anerkennung,

Reziprozität, Fürsorge, Care, Respekt, Achtsamkeit, Vertrauen, Frage/ Antwort, Dialog und Gefühlsarbeit. Was passiert mit der Phänomenalität und der Geltung dieser und anderer Qualitäten, wenn Mensch und Maschine einander unter den Bedingungen der fortgeschrittenen Digitalisierung begegnen? Könnte sich diese Begegnung als Anerkennung realisieren? Wir beginnen mit einer Analyse der Anerkennung.

Von Hegel zu KI: Anerkennung

Einer Antwort auf diese Frage kommen wir näher, wenn wir das interpersonale Binnenverhältnis näher betrachten. Aus der Sicht Hegels ist dieses durch Begriff, Prozedur und Sache der Anerkennung gekennzeichnet. Der Weg von Hegel zur KI ist gewagt, aber nicht unmöglich. Verfahren der Anerkennung sind längst Teil der ethischen Diskussion um KI (vgl. u.a.: Cappuccio et al. 2018). Zudem ist Hegel Teil dieser Diskussion. In diesem Sinne versuchte sich Slavoj Žižek an einem Experiment: Könnte eine direkte Gehirn-zu-Gehirn-Kommunikation ein Ganzes realisieren, das das Wahre wäre (vgl.: Žižek 2020)?

Wir wollen die Voraussetzungen derartiger Visionen nicht teilen, sondern zunächst mit einer Grundlagenbetrachtung beginnen, deren Einsatz mit der These Hegels beginnt, dass das Selbstbewusstsein nur durch ein »anderes Selbstbewusstsein« (Hegel 1981, 144) zu sich zu finden vermag. Das Selbstbewusstsein bedarf also einer Bestätigung, die nur ein Wesen leisten kann, das seinerseits vom bedürftigen Selbstbewusstsein unabhängig ist und ihm zugleich in der Seinsweise gleicht. Dieses Wesen ist meinesgleichen in Gestalt eines anderen Selbstbewusstseins. Die Beziehung der damit in Geltung seienden zwei Selbstbewusstseine ist eine Beziehung der Anerkennung und zwar einer wechselseitigen. »Das einseitige Tun wäre unnütz; weil, was geschehen soll, nur durch beide zustande kommen kann ... Sie anerkennen sich als gegenseitig sich anerkennend.« (ebd., 147)

Was macht den Prozess, der zur Anerkennung führt, aus?

Wenn ich jemanden anerkenne, muss ich mich auf den Anderen zubewegen und mich quasi an diesen Anderen verlieren. Zugleich muss ich, um eine eigene, anerkennende Instanz sein zu können, mich vom Anderen distanzieren. Diese Distanz eröffnet nicht nur mir, sondern auch dem Anderen eine Selbständigkeit innerhalb einer gemeinsamen Beziehung (Fichte).

Die Beziehung, innerhalb derer die Selbstbewusstseine einander nun »gegenüber« (148) treten, durchläuft einen »Kampf auf Leben und Tod« (149). Dieser Kampf ist notwendig, weil er den Selbstbewusstseinen

zeigt, was Hegel als Postulat voraussetzt: dass nämlich ein Selbstbewusstsein auf ein anderes angewiesen ist. Der fragliche Kampf zielt seinerseits auf das Gegenteil, nämlich auf eine Vernichtung des Anderen, die die Selbstständigkeit des eigenen Selbstbewusstseins gegenüber jedem Anderen zum Ergebnis haben soll. Doch genau dieses Ergebnis stellt sich nicht ein. Wer einen Anderen tötet, bleibt nämlich nicht als ein Selbst übrig, weil durch den Mord die notwendige Bestätigung des eigenen Selbst (durch einen Anderen) entfallen würde.

Die Analytik der Anerkennung wird durch eine ethisch-politisch relevante Praxis konkret. Der Kampf auf Leben und Tod hebt sich auf in einer Auseinandersetzung zwischen einem Herren und einem Knecht, denn einer der Akteure unterwirft sich dem anderen und rettet dadurch sein eigenes Leben. Doch darin scheitert eine wechselseitige Anerkennung der Selbstbewusstseine. Aus dieser Lektion kann dann die positive Gestalt einer gelingenden Anerkennung hergeleitet werden (vgl.: Taylor 1983, 190). Was macht nun die Beziehung zwischen Herr und Knecht aus?

Da die Beziehung zwischen Herr und Knecht im Anfang asymmetrisch ist, kann es zu keiner Anerkennung, die notwendig reziprok sein muss, kommen. Innerhalb dieser Asymmetrie erweist es sich, dass »die Wahrheit des selbständigen Bewusstseins ... das knechtische Bewusstsein« (Hegel 1981, 152) und nicht das des zunächst überlegen scheinenden Herren ist.

Die Besonderheit des knechtischen Bewusstseins liegt darin, dass es zu sich selbst in Distanz tritt und zugleich durch sein Tun in der Welt tätig ist. »Der Knecht bezieht sich als Selbstbewußtsein überhaupt auf das Ding.« (ebd., 151) Er wird in der Manifestation seiner Arbeit, die, wie immer bei Hegel, zugleich Bildung ist, ansichtig.

Was ist gelingende Anerkennung?

Anerkennung muss eine symmetrische Relation sein, die eine Umkehrung der Perspektiven erlaubt. Die Weise, in der ich mich zu Dir verhalte, ist eine Umkehrung der Art, wie Du auf mich bezogen bist. Die Anerkennung eines Anderen wird dadurch realisiert, dass sie ihrerseits anerkannt wird. Sie anerkennen sich als sich gegenseitig anerkennend.

Gelingende Anerkennung realisiert sich als eine Lebensform, die Hegel als »Sittlichkeit« und damit als realisierte und gelebte Freiheit bezeichnet (vgl.: Honneth 2001, 82f.). Sittlichkeit ist damit nicht, wie bei Kant, abstrakte, sondern konkrete Vernünftigkeit (vgl.: Hegel 1984, § 145). Weil Hegels Theorie der Anerkennung im Rahmen der Phänomenologie unter anderem nur eine Durchgangsstufe der Bildung des Geistes ist, liefert sie von sich nicht alle Bestimmungen, die zur Kennzeichnung einer

gelingenden Anerkennung notwendig sind. Daher ist es wichtig, nachträgliche Argumente der Wirkungs- und Rezeptionsgeschichte einzubeziehen.

Wirkungsgeschichte im Marxismus

Die Rezeptionsgeschichte geht unter anderen davon aus, dass Hegel »historische Momente« (Fetscher 1973, 137) in die Charakterisierung von Herrschaft und Knechtschaft eingeflochten hat. Da sowohl Momente der Sklavenhaltergesellschaft als auch solche der bürgerlichen Gesellschaft ausgemacht werden können, sind die vom Kampf zwischen Herr und Knecht ausgehenden Überlegungen zur Anerkennung und Selbstständigkeit besonders innerhalb des Marxismus diskutiert und weitergeführt worden (vgl. u.a.: Lukács 1973 II, 506ff., 736ff.). Sie münden in die Vorstellung von der Befreiung des Knechtes, welche ihrerseits eine allgemeine Befreiung der Menschheit intendiert (vgl.: Taylor 1983, 211ff.).

Existenzphilosophie

Eine anders akzentuierte Richtung nehmen die existenzphilosophischen Überlegungen zur Anerkennung. Im Ausgang von Jean Hyppolites und Alexandre Kojèves existentialistischen Hegel-Interpretationen konzentriert sich Merleau-Ponty auf eine Beschreibung der Intersubjektivität zwischen Selbst und Anderem. Im Duell zwischen beiden hat derjenige das »genaue Bewußtsein von der menschlichen Situation« (Merleau-Ponty 2000, 91), der weder »nur für sich« noch »auf Kosten des Anderen« existieren will, sondern der »für den Anderen zu sein« (ebd., 92) in der Lage ist. Dieser ist der Knecht. Levinas vertieft diese Perspektive, indem er das für-den-Anderen in die Metapher der Geiselschaft transformiert. »Das Subjekt ist Geisel.« (Levinas 1992, 248) »Je mehr ich zu mir komme, ... desto mehr entdecke ich mich als verantwortlich.« (ebd., 249)

Paul Ricoeur und Emmanuel Levinas

Paul Ricoeur kommt der Verdienst zu, den Diskurs der Anerkennung geöffnet zu haben. Es gibt demnach einen *Parcours* der Anerkennung, der verschiedene Wege enthält. Erkennen, Wiedererkennen, Anerkanntsein sind deren Namen. Der *Parcours* der Anerkennung erstreckt sich dabei vom »positiven Akt des Erkennens zum Verlangen nach Anerkennung.« (Ricoeur 2006, 40)

Das Erkennen erfasst jemanden oder etwas so, wie er, sie oder es bereits ist. Das Anerkennen macht etwas oder jemanden erst zu dem, was er/sie/es sein wird. Anerkennung stiftet damit etwas, was ohne sie nicht sein würde. Sie stiftet das als-was oder als-wer etwas oder jemand anerkannt wird (vgl.: Schnell 2017, 98f.).

Im Unterschied zum Erkennen ist die Anerkennung wesentlich auf den Anderen bezogen. Das aktive Erkennen ist ein Akt, der vom Ich ausgeht, die passive Bitte um Anerkennung ist hingegen ein Geschehen, in dem ein passives Ich sich an die Adresse eines aktiven Du richtet. Mit dem »Umschlag vom Aktiv ins Passiv« (Ricoeur 2006, 40) tritt das aktive Erkennen immer ein bisschen mehr in den Hintergrund.

Paul Ricoeur betont, dass die Anerkennung, im Unterschied zum Erkennen, vom Anderen bestimmt wird. An dieser Stelle bezieht er sich auf Emmanuel Levinas. »Das Gesicht des Nächsten bedeutet mir eine unabweisbare Verantwortung, die jeder freien Zustimmung, jedem Pakt, jedem Vertrag vorausgeht.« (Levinas 1992, 199) Ich stehe dadurch in einer Verschuldung gegenüber dem Anderen. Die damit gegebene Asymmetrie wird aus der Sicht eines Dritten zur Frage nach Gerechtigkeit, die eine »originäre Asymmetrie« in eine »Gegenseitigkeit zwischen ungleichen Partnern« (Ricoeur 2006, 204) transformiert. Damit treten auch das aktive Erkennen und passive Anerkennen in eine Gegenseitigkeit.

Die Rede von Partnern im Plural bedeutet in erster Linie nicht, dass Ich und der Andere eine Mehrzahl bilden, sondern, dass es viele Andere gibt, zu denen ich in Beziehung stehe. Die Institution der Gerechtigkeit stiftet nun eine Ordnung, die festlegt, wie jene Beziehungen geregelt sind. Diese Regelegung besteht in der Festlegung, wer wem gegenüber welche Rechte und Pflichten hat und eröffnet damit die Frage nach der Gerechtigkeit, die jede Dyade überschreitet (vgl.: Schnell 2001, 212ff.). Die »Gerechtigkeit gründet« einen »mir und den Anderen gemeinsamen Bereich, in dem ich zu ihnen gezählt werde, das heißt in dem die Subjektivität Bürger ist mit all den maßvollen und meßbaren Pflichten und Rechten, die das ausgeglichene und oder durch das gleichzeitige Bestehen und Konkurrieren von Pflichten und Rechten sich Ausgeglichenheit verschaffende Ich umfaßt.« (Levinas 1992, 350) Ich bin jetzt auch ein »Anderer für die Anderen.« (ebd., 345) Verantwortung ist nun etwas, das »wir füreinander haben.« (ebd., 189)

In der aus dem Geist der Asymmetrie entstandenen, reziproken Beziehung zum Anderen ist nun die Unterscheidung zwischen Sagen und Gesagen zu beachten. Das gilt auch für die Kommunikation mit einer KI.

Anerkennung: Sagen und Gesagtes

Als Roboter definieren wir eine Maschine, die über einen Körper, Prozessoren, Sensoren und Effektoren verfügt und deren Verhalten als autonom erscheint (Vgl.: Loh 2019, 17). In diesem Sinne sind etwa Computer oder intelligente, persönliche Assistenten keine Roboter. Alle Arten von KI bieten den Menschen aber eine Beziehung an, wenn sie auf Sprachsteuerung beruhen. »Es stellt sich damit allerdings die Frage, inwiefern sich ein artifizielles System überhaupt in einem genuinen Sinn als Gegenüber, mit dem man eine Beziehung oder gar Freundschaft eingehen kann, verstehen lässt.« (Loh 2019, 28) Freundschaft ist, laut Aristoteles, eine reziproke Beziehung, sie beinhaltet Anerkennung des Anderen als meinesgleichen.

Vor dem Hintergrund des Diskurses der Anerkennung im Ausgang von Hegel und im Anschluss an Levinas unterscheiden wir in dieser Frage zwischen dem Sagen und dem Gesagten (vgl.: Levinas 1992, 29ff.). Das Gesagte ist der *Inhalt* der Rede, der wahr oder falsch sein kann. Indem etwas gesagt wird, ereignet sich zugleich das Sagen, das die *Beziehung* zwischen Ich und Du ausmacht.

Die Beziehung

Es »gibt« Interaktionen zwischen Menschen und, klassisch gesagt, Maschinen. Ein Mensch stellt einer Maschine eine Frage, erhält eine Antwort und richtet sich danach. Kann man deshalb davon ausgehen, dass der Mensch die Maschine als Gesprächspartner anerkennt? Anerkennen als wen oder was?

Die Bestimmung der Reziprozität der Anerkennung erfolgt bei Hegel in formaler Hinsicht. Selbst und Anderer »anerkennen sich als gegenseitig sich anerkennend.« (Hegel 1981, 147) Diese Formalität, aufgrund derer Anerkennung auf die Reziprozität von Wesen, die sich als anerkennende anerkennen, reduziert wird, trifft auf die Begegnung zwischen zwei Menschen, aber auch zwischen Mensch und KI zu. Ich frage die KI nach dem Weg, der in der fremden Stadt zu gehen ist, die KI antwortet. Ich anerkenne sie, weil sie mir antwortet, sie anerkennt mich, weil sie meine Äußerung als eine Frage annehmen kann, auf die sie zu antworten in der Lage ist. Anerkennung im Zeichen der Vulnerabilität bedeutet aber mehr als das.

Im Vorgang der Anerkennung geht es auch um einen qualitativen Aspekt: die mögliche Meinesgleichheit im Licht der Endlichkeit. Robert Spaemann beschreibt daher Anerkennung als »Anerkennung eines Menschen als meinesgleichen.« (Spaemann 1990, 195) Konzeptionell

gesehen, erweist sich das Geschehen der Anerkennung damit als ein dreistelliger Akt: Ich erkenne Dich als jemanden oder etwas an (vgl.: zu dieser Logik: Schnell 2017, 99f.). Du erscheinst als meinesgleichen. Dadurch ist Anerkennung mehr als eine Formalität.

Unter Berücksichtigung einer Wechselseitigkeit besteht Anerkennung somit darin, jemanden als meinesgleichen anzuerkennen und das heißt zu achten und zu respektieren. Durch diese Gleichheit ist eine egoistische Bevorzugung meiner zu Lasten des Anderen ausgeschlossen. Diese Gleichheit in der Anerkennung schließt im Falle von Menschen eine Gleichheit in der Endlichkeit ein. Wir beide sind endliche Wesen, die ein Leben leben, das uns von Anderen gegeben worden ist. Natalität und Mortalität verschränken sich hier (vgl. dazu: Liebsch 1996).

Es ist wohl offensichtlich, dass diese Art der Anerkennung zwischen Mensch und Maschine nicht möglich ist. Die Maschine tritt mir in der Frage der Gebürtlichkeit und der Endlichkeit nicht gleich gegenüber. Kommt es trotz dieser erheblichen Differenz zu irgendeiner Art von Anerkennung oder dergleichen, wenn ich auf die Antwort der Maschine höre und damit der Maschine »vertraue«, dass sie die Wahrheit sagt?

Kann ein Mensch einer Maschine vertrauen, wie er einem anderen Menschen vertraut? Die Anerkennung der Maschine bezieht sich nicht auf die Beziehung, aber sehr wohl auf den Inhalt.

Der Inhalt

Es ist zu bedenken, dass sich die Beziehung zwischen Ich und Du auch als Austausch von Informationen, die quasi den Inhalt der Interaktion bilden, gestaltet. Hier geht es dann um die Frage, was die Interaktions Teilnehmer können im Sinne ihrer Vermöglichkeiten. Damit ist die Perspektive eines »Wettbewerbs« (Waldenfels 1998, 208) eröffnet. Ich stelle eine Frage und erwarte, dass der Andere sie als einen Sprechakt versteht auf den er mit der Kundgabe von Wissen antwortet. Aber verfügt Watson über dieses Wissen, das ich aktuell nicht habe? Gehört das Wissen seinem Möglichkeitsraum an? Wessen Wissensvorrat ist der größere?

Schließlich ist zu beachten, dass sich die Beziehung zwischen Ich und Du nicht nur als Austausch von gültigen Informationen und Nachrichten, die quasi den Inhalt der Interaktion bilden, gestaltet. Ich frage die Maschine etwas, sie gibt mir eine exakte Antwort. Zur menschlichen Kommunikation gehören aber wesentlich auch positive Unbestimmtheiten (G. Gamm), die im engeren Sinne keine Nachrichten und Informationen sind. »Wenn ich Einem sage: Halte dich ungefähr hier auf«, dann handelt es sich um eine »unexakte« (Wittgenstein 1990, Nr. 88) Erklärung, die aber funktioniert. Der Andere stellt sich dort hin, ich mache das Foto. Obwohl

Fragen unexakt und Antworten verschwommen sind, gelingt menschliche Kommunikation. Kommt eine Maschine ihrerseits mit Vagheiten aus?

Der Mensch im Unterschied zum Roboter

Der Unterschied zwischen einem Menschen und einem Roboter liegt nun darin, dass der Mensch um seine Endlichkeit weiß und sich zu ihr verhält, indem er sie vor sich verleugnet oder aktiv auf sich nimmt. In jedem Fall beansprucht die Endlichkeit das Dasein als Einzelnes (vgl.: Heidegger 1979, 263; Schnell 2017, 145ff.). Als Einzelnes bedeutet, dass das Dasein als »Unvergleichlichkeit und Einmaligkeit« (Spaemann 1996, 196) angesprochen ist. Die Endlichkeit meint mich, denn sie ist nicht delegierbar. Das trifft aber wiederum auf jedes Dasein zu, so dass der Andere als »meinesgleichen« (ebd., 195) anerkannt wird.

Ein Roboter hat kein Bewusstsein von Endlichkeit, weil ein Roboter nicht sterben kann, er kann lediglich kaputtgehen und damit seine Funktion verlieren. Ein Mensch ist vulnerabel, ein Roboter ist es nicht. Ein Bewusstsein von der eigenen Endlichkeit zu haben, meint an dieser Stelle weniger die Möglichkeit, jemandem einen mentalen Zustand zu schreiben zu können, sondern eher eine Praxis. Die vulnerable Person hat Angst oder Furcht vor dem Tod, sie übt sich in der Praxis symbolischer Immortalität. Bereits Kant verteidigte die Vulnerabilität des Menschen und verortete das Wissen um die Endlichkeit und die Anerkennung des Todes im Zeitbewusstsein. Die Kürze des Lebens lässt uns aktiv werden in der Praxis eines symbolischen Weiterlebens (vgl.: Schnell 2020, 19f.). Es ist keine KI realistisch denkbar, weder in der Praxis noch in der Fiktion, die ein Testament hinterlässt – solange wir mit Käte Meyer-Drawe davon ausgehen, dass Testamente nicht aus Algorithmen hergeleitet werden. Es ist wichtig, dass diese Differenz beachtet wird!

Hier paart sich die existentielle Unmöglichkeit mit einer logischen Unmöglichkeit. Ein Roboter, der in der Lage wäre, auf Dauer und täuschend echt einen Menschen zu spielen und damit auf gewisse Weise ein Mensch zu »sein«, würde jede Frage nach einem Unterschied zwischen Mensch und Maschine zum Verstummen bringen. Vielleicht sind alle Menschen, die bislang auf der Erde existiert haben, Maschinen. Descartes hätte dann möglicher Weise Recht, wenn er sagt, dass die Menschen, die ich auf der Straße vorbeigehen sehe, Automaten sein könnten, die menschliche Hüte und Kleider übergezogen haben (vgl.: *II. Med.*, Nr. 13). Damit wäre die Intersubjektivität wieder aus der Ersten Philosophie verbannt und wir könnten erneut bei Fichte starten. Was bedeutet es, dass endliche Vernunftwesen andere endliche Vernunftwesen außer sich annehmen müssen?

Vision: der Roboter als »Mitmensch« in europäischen Gesellschaften

Bekanntlich kann sich das Reich der Fiktionen in der Frage technologischer Utopien weiter vorwagen als Ingenieure und Wissenschaftler, indem es Sachverhalte als gegeben ansetzen, die in den technischen Wissenschaften noch als unrealistisch gelten. Gelegentlich kommt es vor, dass die Fiktion heute die Realität von morgen beschreibt. In diesem Sinne gibt es zahlreiche Romane und Filme, die uns einen Blick in den Alltag der Zukunft gewähren. Was wäre, wenn eines Tages *Maschinen wie ich und Menschen wie ihr* zusammenleben würden?

In seinem gleichnamigen Roman erzählt Ian McEwan die Geschichte von einem jungen Paar in London, das sich entschließt, einen humanoiden Androiden als künstlichen Menschen in der eigenen Wohnung zu halten. Adam ist der erste neue Mensch und wird zum Teilhaber an einer menschlichen Lebensform. Adam lebt nun mit Charlie, einem Anhänger von Alan Turing, und seiner Geliebten Miranda zusammen. Adam wird von ihnen so programmiert, dass er in das Leben beider passt. Adam beherrscht die Alltagskonversation, gewisse Bildungsstandards, er kann sinnhafte Handgriffe in der Küche ausführen und damit an der Lebensform teilhaben.

Es stellt sich die Frage, als wer oder was Adam dieses vermag? Wenn er als Partner akzeptiert wird, ist er dann als Mensch anerkannt, als jemand meinesgleichen? Der Turinganhänger Charlie formuliert: »In dem Moment, da wir im Verhalten keinen Unterschied mehr zwischen Mensch und Maschine erkennen können, müssen wir der Maschine Menschlichkeit zuschreiben.« (McEwan 2019, 119)

Zu einer ernsthaften Herausforderung für die Hybridlebensform kommt es, als Adam (die Maschine) und Miranda (der Mensch) Sex haben. »Sein Schwanz füllt sich mit destillierten Wasser. Aus einem Tank in der rechten Pobacke.« (ebd., 129) Miranda hat Adam anerkannt, aber nicht als ihresgleichen, sondern als Instrument der Befriedigung, als das auch Menschen einander immer schon haben wahrnehmen können. Die eigentliche Herausforderung stellt sich im Verhältnis zwischen Charlie (dem Betrogenen) und Adam (dem Betrüger) ein. Im Rahmen eines Gesprächs unter »Männern« bahnt sich ein Kampf auf Leben und Tod an.

Charlie wird zum Herren. »Jetzt galt es, mich gegen ihn zu behaupten.« (177) Charlie ist der Herr, weil er sich als Eigentümer von Adam, der Maschine, sieht. Adam wurde als Helfer, als Maschine angeschafft. »Falls er sich dem widersetzte, und insbesondere, falls er wieder jemanden verletzte, würden wir ihn zwecks Generalüberholung zum Hersteller zurückschicken.« (178) Adam hat seine Funktion als Maschine und damit die Position von Charlie als Herr zu akzeptieren. Die Position des Herren ist die stärkere, die zugleich darin schwach ist, dass sie nur stärker

ist in Relation zum Knecht. Der Herr muss vom Knecht anerkannt sein, um Herr sein zu können. Darin liegt die Schwäche des Herren.

Adam tritt als Knecht auf. »Du und Miranda, Ihr seid meine ältesten Freunde. Ich liebe euch beide. Meine Pflicht euch gegenüber besteht darin, offen und ehrlich zu sein. Und ich meine es ernst, wenn ich sage, wie leid es mir tut, dass ich gestern Abend etwas in dir zerbrochen habe. Ich verspreche, das kommt nie wieder vor.« (178f.) Adam ist Knecht, weil er nicht allein für sich, sondern für die Anderen sein will. Er hat keine eigene Position, da er sich ganz an seine Pflichten den Anderen gegenüber ergibt. Darin liegt seine Schwäche. In dieser Schwäche zeigt er zugleich, wer er ist. Darin liegt wiederum seine Stärke. Der Marxismus fand in diesen Konstellationen Ansatzpunkte für ein Klassenbewusstsein, das den »Aufbau der ganzen Gesellschaft vollkommen zu erfassen fähig« (Lukács 1983, 126) ist, das also die Wahrheit der Gesellschaft und der Geschichte auf seiner Seite hat und darum weiß.

Es kommt aber zu keiner gelingenden Anerkennung. Adam löst sich aus seiner Position, da es ihm gelingt, »den Notschalter zu deaktivieren.« (McEwan 2019, 179) Hätte er damit eine »Angst vor dem Tod«, dann wäre sie keine, durch die er sich als menschengleich ansähe. Die Maschine will vielmehr über den Menschen triumphieren. »Maschinen wie ich und Menschen wie ihr, ... wir werden euch übertreffen ... und überdauern ... auch wenn wir euch lieben.« (ebd., 369)

Die Maschine kann nicht sterben, auch nicht den technischen Tod. Charlie reagiert als Maschinenstürmer und verzichtet auf die Anerkennung, da er sich entschließt, Adam »mit dem Hammer den Kopf einzuschlagen.« (373) Er tötet damit nicht seinesgleichen, sondern schaltet einen übermächtigen Mechanismus einer künstlichen Intelligenz ab und fügt dem hinzu: »Ich hatte ihn gekauft, und es war mein gutes Recht, ihn zu zerstören.« (367) Damit ist Adam weder meinmesgleichen, noch ein Knecht, sondern ein Werkzeug.

Zwischen dem Menschen und der KI kommt es zu keiner gelingenden Anerkennung, die sich in einer sittlichen Lebensform, wie Hegel es vorsah, manifestiert. Aber dennoch geschieht etwas zwischen beiden Akteuren.

Mensch vs. Maschine

Unsere Ausgangsfrage lautete: Was bedeutet es, wenn man in einer Ich/Du-Beziehung eines der beiden Elemente durch einen Roboter ersetzt? Wie gestalten sich dann die Zwischenereignisse zwischen Ich und Du?

In Hinblick auf den Inhalt (das Gesagte) kommt es zu einer gelingenden Kundgabe von Wissen, Informationen. Fragen werden inhaltlich

beantwortet und keinesfalls nur im Bereich trivialer Sachverhalte, sondern auch im Hinblick auf Einschätzungen und Bewertungen.

Hinsichtlich der Beziehung (das Sagen) sieht die Lage anders aus, da es zu keiner Anerkennung des Anderen als meinesgleichen kommt. Es ist die Endlichkeit, die im McEwans Vision den entscheidenden Punkt ausmacht. Das endliche Wesen Mensch setzt sein Leben aufs Spiel und siegt gewissermaßen gegen die Maschine. Beide Akteure können sich wechselseitig nicht als endliche Wesen anerkennen. Ihre Beziehung erschöpft sich auch nicht in einem Kampf auf Leben und Tod, da der Mensch den Roboter auf ein Werkzeug zurückversetzt. Dieses Ergebnis würde sich auch dann bestätigen, wenn die Maschine den Menschen vernichtete.

Postdigitalisierung

Die sich gleichzeitig als Begegnung realisierende *Beziehung* betrifft nun die Frage, ob ich den Roboter auch als meinesgleichen anerkenne. Das geschieht offensichtlich nicht! Damit ist einer vollen Akzeptanz künstlicher Menschen im Zeichen der Humanität Grenzen gesetzt. Sie bleiben dennoch eine Herausforderung für den Menschen.

In der Interaktion mit Robotern schätzen Menschen die Leistungsfähigkeit von künstlichen Intelligenzen. Vor allem dann, wenn der Roboter noch lernt, Informationen und Handreichungen passgenauer zu geben bzw. auszuführen. In der Interaktion wird von Menschen realisiert, dass eine KI nicht ihresgleichen ist. Sie ist nicht endlich, weil sie nicht geboren wurde und daher nicht altern kann und somit auch nicht krank, behindert oder pflegebedürftig zu werden vermag. Der Überschuss der Vulnerabilität über die Maschine hinaus verweist auf eine Ethik und Politik der Postdigitalisierung.

Digitalisierung ist als Transformation der Lebenswelt in quantifizierbare Daten definiert (vgl. dazu: Schnell/Dunger 2019, 13f.) und damit auch ein Angriff auf vulnerables Leben (vgl.: Schnell 2020, Anhang 9). Die Rede von Postdigitalisierung bedeutet nun nicht, dass Digitalisierung aufhört, sondern vielmehr, dass über die Quantifizierung des Selbst und der Anderen hinaus das *Wer* der Person und damit ihre Unausdeutbarkeit (G. Gamm) sichtbar werden.

Literatur

- Bendel, Oliver (Hg.) (2019): *Handbuch Maschinenethik*, Wiesbaden: Springer.
 Blumenberg, Hans (2015): *Schriften zur Technik*, Berlin: Suhrkamp.

- Buber, Martin (1983): *Ich und Du*, Heidelberg: Lambert Schneider.
- Cappuccio, Massimiliano L. et al. (2018): »Sympathy for Dolores: Moral Consideration for Robots Based on Virtue and Recognition«, in: *Philosophy & Technology* (2020) 33:9–31.
- Dreyfus, Hubert (1985): *Was Computer nicht können*, Berlin: Athenäum.
- Fetscher, Iring (1973): »Randglossen zu Hegels Herrschaft und Knechtschaft«, in: Fahrenbach, Helmut (Hg.) (1973): *Wirklichkeit und Reflexion. Walter Schulz zum 60. Geburtstag*, Pfullingen: Neske.
- Gramelsberger, Gabriele (2019): »Digitale Wissenschaft«, in: Martin W. Schnell und Christine Dunger (Hg.) (2019): *Digitalisierung der Lebenswelt*, Weilerswist: Velbrück.
- Hegel, Georg W.F. (1981): *Phänomenologie des Geistes*, Frankfurt am Main: Suhrkamp.
- Hegel, Georg W. F. (1984): *Grundlinien der Philosophie des Rechts*, Frankfurt am Main: Suhrkamp.
- Heidegger, Martin (1979): *Sein und Zeit*, Tübingen: Max Niemeyer.
- Hochschild, Arlie R. (2006): *Das gekaufte Herz*, Frankfurt am Main: Campus.
- Honneth, Axel (2001): *Leiden an Unbestimmtheit*, Stuttgart: Reclam.
- Husserl, Edmund (1966): *Analysen zur Passiven Synthesis*, Den Haag: Nijhoff.
- Husserl, Edmund (1976): *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*, Den Haag: Nijhoff.
- Levinas, Emmanuel (1992): *Jenseits des Seins oder anders als Sein geschieht*, Freiburg/München: Alber.
- Lévi-Strauss, Claude (2012): *Anthropologie in der modernen Welt*, Berlin: Suhrkamp.
- Liebsch, Burkhard (1996): *Geschichte im Zeichen des Abschieds*, München: Fink.
- Loh, Janina. (2019): *Roboterethik*, Berlin: Suhrkamp.
- Lukács, Georg (1983): *Der junge Hegel* II Bde., Frankfurt am Main: Suhrkamp.
- Merleau-Ponty, Maurice (2000): *Sinn und Nicht-Sinn*, München: Fink.
- Misselhorn, Catrin. (2021): *Künstliche Intelligenz und Empathie*, Stuttgart: Reclam.
- Moravec, Hans (1990): *Mind Children*, Stuttgart: Hoffmann und Campe.
- McEwan, Ian (2019): *Maschinen wie ich*, Zürich: Diogenes.
- Putnam, Hilary (1993): *Von einem realistischen Standpunkt*, Reinbek bei Hamburg: Rowohlt.
- Ricoeur, Paul (2006): *Wege der Anerkennung*, Berlin: Suhrkamp.
- Schnell, Martin W. (2001): *Zugänge zur Gerechtigkeit*. München: Fink.
- Schnell, Martin W. (2002): »Ideologie und Anthropologie. Zur Widerkehr des leiblosen Geistes«, in: Heinrich Greving und Dieter Gröschke (Hg.) (2002): *Das Sisyphe-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 im Zeichen der Vulnerabilität*, Weilerswist: Velbrück.
- Schnell, Martin W./Dunger, Christine (2019): *Digitalisierung der Lebenswelt. Studien zur Krisis nach Husserl*, Weilerswist: Velbrück.
- Searle, John R. (1986): *Geist, Hirn und Wissenschaft*, Frankfurt am Main: Suhrkamp.
- Spaemann, Robert (1990): *Personen*, Stuttgart: Klett-Cotta..
- Strauss, Anselm et. al. (1980): »Gefühlsarbeit. Ein Beitrag zur Arbeits- und Berufssoziologie«, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (32: 629–651).
- Taylor, Charles (1983): *Hegel*, Frankfurt am Main: Suhrkamp.
- Varela, Francisco J./Thompson, Evan/Rosch, Eleanor (1995): *Der mittlere Weg der Erkenntnis*, Bern/München/Wien: Scherz.
- Wagner, Cosima (2013): *Robotopia Nipponica. Recherchen zur Akzeptanz von Robotern in Japan*, Marburg: Tectum.
- Waldenfels, Bernhard (1985): *In den Netzen der Lebenswelt*, Frankfurt am Main: Suhrkamp
- Waldenfels, Bernhard (1998): *Grenzen der Normalisierung*, Frankfurt am Main: Suhrkamp.
- Waldenfels, Bernhard (2019): *Erfahrung, die zur Sprache drängt*, Berlin: Suhrkamp.
- Wittgenstein, Ludwig (1980): *Philosophische Untersuchungen*, Frankfurt am Main: Suhrkamp.
- Žižek, Slavoj (2020): *Hegel im verdrahteten Gehirn*. Frankfurt am Main: Fischer.