

ETHIK – KIRCHE – DIAKONIE

Digitalisierung als Herausforderung für Anthropologie und Ethik

1. Einleitung

Digitalisierung besagt zunächst die Umwandlung eines kontinuierlichen, nichtgestuften Signals in ein gestuftes Signal. Da die gestuften Signale maschinenlesbar, leicht kopier- und übertragbar sind, können Informationen heute fast gleichzeitig sehr vielen Nutzer*innen zur Verfügung gestellt werden. Weil sich dadurch Kommunikations-, Arbeits- und Lebensweisen gravierend ändern, steht Digitalisierung für einen radikalen gesellschaftlichen Wandel mit rechtlichen und ethischen Herausforderungen. Vielleicht ändert sich auch unser Selbstbild als denkende und moralisch handelnde Lebewesen: Als denkende Wesen scheinen wir Konkurrenz durch künstliche Intelligenz bekommen zu haben, als körperliche Wesen treten uns Roboter zur Seite.¹

Einige dieser Herausforderungen scheinen durchaus bereits in den „alten Medien“ auf: Radio, Fernsehen und Presse haben bei ihrer Einführung Neufigurierungen der Kommunikation, der Grenzziehung zwischen Privatem und Öffentlichem impliziert und Regeln erforderlich gemacht. Die

1 Die Herausforderungen von Digitalisierung, künstlicher Intelligenz und Robotik haben mich in den vergangenen Jahren vielfach beschäftigt, im Rahmen einer Ethik-Vorlesung für Studierende der Informatik, in einem philosophischen Seminar zur Robotik, in einer Sommerakademie zur Digitalisierung (Kremsmünster 2018), in einem Diskursprojekt mit der Academia Superior sowie in einigen Arbeitsgruppen des Europarates zu ethischen Herausforderungen emergenter Technologien. Im nun vorliegenden Beitrag knüpfe ich an Zwischenresultate an, die im Rahmen des Tagungsbandes von Kremsmünster *„Selbstlernende Systeme – ethische Fragen“*, in: Lederhilger, Severin J. (Hg.), *Gott und die digitale Revolution*, (=Schriften der Katholischen Privatuniversität Linz 6), Regensburg 2019, 114–131) sowie zweier kurzer Kolumnen (*Regeln für selbstfahrende Autos*, in: *Interesse. Soziale Information*, Nr. 1, hrsg. vom Sozialreferat der Diözese Linz, 2018/1, 1–2; *Digitalisierung und selbstlernende Systeme – ethische Fragen*, in: *Mitteilungen. Zeitschrift der Religionslehrerinnen und -lehrer der Erzdiözese Salzburg*, 1–2019, 5–7) bereits in gedruckter Form vorliegen. Den Studierenden der KU Linz und der JKU Linz sowie den Mitwirkenden der Sommerakademie in Kremsmünster 2018 sowie den Mitgliedern der Arbeitsgruppen in Linz, Straßburg und Paris danke ich herzlich für Rückmeldungen und Anregungen zum Weiterdenken.

Möglichkeiten, die neue Technologien den Menschen eröffnen, sind erneut ambivalent und teilweise mit Gefahren verbunden: Autonomes Fahren soll kommen, aber die Haftungsfragen sind ungeklärt. Rasenmäher, die als Roboter arbeiten, verletzen Igel und Bildschleichen. Algorithmen, die Kreditvergabeentscheidungen steuern, tun dies in einer Weise, dass der Verdacht auf Diskriminierungen bleibt.

Bei so vielen und teilweise so grundlegenden Problemen und Fragen ist es wichtig, sich den Hintergrund zu verdeutlichen und nach Möglichkeit eine Ordnung vorzunehmen. Sicher wird man Lukas Ohly zustimmen können, wenn er zu bedenken gibt, dass die „sozialen und ethischen Probleme [...] nicht einheitlicher Art“ sind, „nur weil sie die digitale Technik voraussetzen“ (Ohly 2019, 25). Er betont zu Recht, dass Themen wie „virtuelle Welten, Datenschutz und Eigentumsrechte angesichts einer hohen Datenmenge (Big Data), Verantwortungsfähigkeit Künstlicher Intelligenz oder die Delegation sozialer Verantwortung an Roboter“ die digitale Speicherung von Daten und ihre Vernetzung voraussetzen. Gleichwohl ist, und auch hier wird man Ohly folgen, das digitale Verarbeitungsverfahren nicht als Grundlage aller Probleme anzusehen. Vor allem macht er geltend, dass man durchaus zwischen Veränderungen, die sich in einer virtuellen Welt des Cyberspace abspielen, und solchen, die auf die analoge Welt abzielen, unterscheiden müsse. Alles andere wäre ein falscher Reduktionismus, man würde, so meint Ohly sogar, die Ethik auf Technik reduzieren (Ohly 2019, 27).

Wenn in diesem Beitrag die Herausforderungen der Digitalisierung für Anthropologie und Ethik thematisiert werden, dann heißt das weder, unsere Welt als ganze würde digital, noch wäre es angemessen, den Menschen insgesamt als einen *homo digitalis* zu beschreiben.² Dennoch hat sich die Welt verändert durch die Digitalisierung, und auch wir haben uns verändert.

Die Darstellung des Verhältnisses der Digitalisierung zum menschlichen Selbstverständnis und den ethischen Problemen lässt sich allerdings auch nicht einfach in separate Teilgebiete aufspalten. Es gibt Zusammenhänge zwischen der Entwicklung in der sog. Künstlichen Intelligenz und der Robotik. Indes ist es gerade für eine vorsichtige Orientierung ratsam, in Schritten vorzugehen. Hier wird vorgeschlagen, als einen ersten Schritt

2 Rafael Capurro, der diesen Terminus als Titel eines Sammelbands von einschlägigen eigenen Aufsätzen zur Digitalisierung, zur Medienethik und zu Fragen der künstlichen Intelligenz wählt, scheint eine solche starke These auch nicht zu teilen (vgl. Capurro 2017).

auf die Robotik zu schauen und dabei zunächst auf die technologischen Entwicklungen und dann auf ihre Reflexion.

2. Zur Entwicklung und Reflexion der Robotik

Die neue Generation der Staubsauger und Rasenmäher, die nach dem Einschalten ohne weitere menschliche Steuerung ihre Arbeit des Saugens oder Mähens verrichten, stellt nur den Anfang einer Entwicklung zu mehr Automatisierung des Alltags dar. Wohin genau diese führt und wie schnell sie dorthin voranschreitet, kann freilich niemand seriös vorhersagen. Wie die Staubsauger und die Rasenmäher sind die meisten der im Einsatz befindlichen Roboter keine Androiden. Sie sind nicht in der Absicht hergestellt, dem Menschen in besonderer Weise zu gleichen. Schaut man auf die Roboternutzung und auf Zukunftsszenarien dazu, dann fällt auf, dass es Bereiche gibt, in denen menschliche Tätigkeit aufgrund besonderer Gefahr vermieden und daher durch Robotereinsatz ersetzt werden soll. Mienenentschärfungen, Einsätze bei Brandkatastrophen, Einsätze bei drohenden Atomkatastrophen, militärische Einsätze, Expeditionen für die Klimaforschung in der Tiefsee, der Bau von Eisenbahntunneln durch das Gebirge oder unter Meerengen und Flüssen: all dies sind Einsatzbereiche, in denen Lebens- und Gesundheitsgefährdung menschlicher Einsatzkräfte erheblich wären und wo deshalb Roboter und Drohnen als Ersatz angedacht oder erprobt werden. Daran wird deutlich, dass Roboter keinen Schutzstatus haben, der dem Lebensschutz der Einsatzkräfte vergleichbar wäre. Ein weiterer Unterschied, der sich zeigt, ist, dass Roboter eben nicht ähnlich verletzlich, schmerzempfindlich, toxisch bedroht, sauerstoffabhängig etc. sind. Die Differenz bezieht sich also auf den unterschiedlichen Status von Robotern und Menschen und auf unterschiedliche Eigenschaften. Nicht die Ähnlichkeit, sondern die Unähnlichkeit macht den Robotereinsatz hier attraktiv.

Im Juni 2018 stellte eine Gruppe von Wissenschaftler*innen einen sogenannten micro-bot vor, der Stammzellen durch ein Magnetfeld in ein Zielgewebe transportieren kann (Li, Li, et al. 2018). Die vorgestellten Experimente beziehen sich auf in-vitro-Kulturen und auf das in vivo-System der Maus. Es geht um Eingriffe, wie sie bislang nicht möglich waren. Der Vorzug der micro-bots für eine künftige therapeutische Nutzung besteht gerade in ihrer Winzigkeit, durch die sie sich vom Menschen unterscheiden. Für nano-bots gilt dasselbe in noch verschärfter Weise. Micro-bots und nano-bots werden also durchaus in der Robotik geführt und gleiches gilt für

sogenannte co-bots, die als Kooperationspartner menschlicher Akteure eingesetzt werden. Der Markt als wichtiger Motor der zeitgenössischen Technologieentwicklung, so kann man mutmaßen, wird gerade die Einführung solcher Roboter begünstigen, die durch ihren Unterschied zum Menschen charakterisiert sind.³

Allerdings gibt es durchaus Bereiche des Einsatzes von Robotern, bei denen Ähnlichkeit zum Menschen angestrebt sein könnte und von den Entwickler*innen faktisch auch angestrebt wird. Dann nämlich, wenn wir nicht über Industrieroboter sprechen, sondern über Serviceroboter, und dieser Service eine menschliche Note erfahren soll, oder besser eine quasimenschliche Note, setzen Designer*innen auf Ähnlichkeit (Wynsberghe 2013). Für den Bereich der Pflege wird diskutiert, inwiefern hier der Einsatz von Servicerobotern den durch demographischen Wandel und Mangel an Pfleger*innen drohenden Notstand kompensieren könnte (Alves-Oliveira, et al. 2015).

So wird man festhalten müssen, dass es diverse Entwicklungstendenzen gibt, die gerade in der Verhältnisbestimmung von Mensch und Roboter nicht sofort auf einen Nenner zu bringen sind. Daher scheint es angemessen, nicht nur nach dem menschlichen Selbstbild und Selbstverständnis zu fragen, sondern auch näher auf das Bild vom Roboter einzugehen. Ein unter Ingenieur*innen immer noch gebräuchliches Verständnis ist jenes, das der Verein Deutscher Ingenieure in der Richtlinie 2860 ausgedrückt hat: „Ein Roboter ist ein frei und wieder programmierbarer, multifunktionaler Manipulator mit mindestens drei unabhängigen Achsen, um Materialien, Teile, Werkzeuge oder spezielle Geräte auf programmierten, variablen Bahnen zu bewegen zur Erfüllung der verschiedensten Aufgaben“ (VDI 1990). Eine andere Weise der Definition kann man dem Glossar entnehmen, welches Catrin Misselhorn ihren „Grundfragen der Maschinenethik“ beigefügt hat. Unter dem Stichwort Roboter findet sich folgende Erklärung: „Elektromechanische Maschine, die typischerweise programmierbar ist, über Sensoren verfügt, um die Umwelt wahrzunehmen, und Aktoren, um auf sie einzuwirken. Der Robotik kommt im Kontext des Verkörperungsansatzes (Embodiment) in der Künstlichen Intelligenz (KI) eine wichtige Rolle zu, die intelligentes Verhalten als abhängig vom Körper und seiner Interaktion mit der Umwelt betrachtet“ (Misselhorn 2018, 271).

3 „Market ‚pull‘ is coming from a very wide range of applications where autonomy promises to do things that could not be done before, or to do existing operations more efficiently or safely.” (Innovation in autonomous systems, Summary of an event held on Monday 22 June 2015 at the Royal Academy of Engineering, 16.)

Auch wenn Misselhorn hier die Sensorik anspricht und mit dem Hinweis zur KI weit über klassische Maschinenvorstellungen hinausweist, so ist das Bild, welches sich aus den eigentlichen definitiven Elementen ergibt, demjenigen der deutschen Ingenieure nicht unähnlich. Indes kann man zu bedenken geben, dass der Begriff des Roboters im Zusammenhang der Idee kreierte wurde, menschliche Arbeit durch maschinelle zu ersetzen. Der Begriff „Roboter“ wurde um 1920 von *Karel Čapek* in einem Theaterstück geprägt, dessen Titel übersetzt „Rossums universelle Roboter“ lautet, und ist vom tschechischen „*robota*“ für „Arbeit“ oder für „Fronddienst“ abgeleitet. In einigen Landschaften und so auch in Teilen Österreichs wird als Roboter-Arbeit eine unbezahlte für die Gemeinde geleistete Arbeit bezeichnet. In Čapeks Stück stellt das Unternehmen R. U. R. künstliche Menschen her. Diese „Robots“ werden als preisgünstige und rechtlose Arbeiter verwendet. Čapek schildert, wie ihr massiver Einsatz in der Industrie das gesamte Wirtschaftssystem verändert. Im Verlauf des Stücks lehnen sich die Roboter jedoch auf und bedrohen die Menschheit mit Unfreiheit und Tod.

Das Stück wird nicht mehr aufgeführt, und auch den gängigen literaturwissenschaftlichen Nachschlagewerken und Schauspielführern ist die Handlung nicht oder nur sehr verkürzt zu entnehmen. Deshalb sei sie hier kurz geschildert: Da Rossum ein Wortspiel mit dem tschechischen Wort für Verstand ist, lautet der Titel der einzigen greifbaren Übersetzung „W. U. R. Werstands universal Robots“.

Das Stück spielt in der Fabrik „Werstands Universal Robots“. Helene Glory ist aus Europa zu Harry Domin, dem Zentraldirektor, gekommen, um sich die Erzeugung künstlicher Menschen zeigen zu lassen. Domin erläutert zunächst die Vorgeschichte der Roboterproduktion. Im Jahre 1920 habe der alte Werstand versucht, durch chemische Synthese die lebendige Materie Protoplasma nachzubilden. 1932 habe er einen Stoff entdeckt, der sich wie lebendige Materie verhielt, allerdings bei anderer chemischer Zusammensetzung. Von diesem Ausgangspunkt habe Werstand dann versucht, *das* analog nachzubilden, was in der Entwicklung der Arten geschehen ist. Genannt werden Hund und Mensch. Andere Optionen seien durchaus offen gewesen.

Domin führt dann aber aus, dass der Weg einer anatomisch möglichst exakten Replikation des Menschen nicht gewählt wurde. Der Neffe des alten Physiologen Werstand setzt sich über die Strategie hinweg und geht ingenieurmäßiger vor. Was er will, sind lebende und intelligente Arbeitsmaschinen. Dazu braucht es zunächst eine Vereinfachung der Anatomie. „Der junge Werstand erfand einen Arbeiter mit der kleinsten Menge Bedürfnisse. [...] Roboter sind keine Menschen. Sie sind mechanisch voll-

kommener als wir, haben eine erstaunliche Vernunftintelligenz, aber sie haben keine Seele“ (Čapek 2017, 12).

Es stellt sich dann aber heraus, dass doch vieles nachgebildet wird – Adern, Nerven, Haut –, und zwar in einer Weise, dass Helene nicht erkannt hat, dass die Sekretärin in Wirklichkeit kein Mensch, sondern ein Roboter ist. Und Helene gibt bekannt, dass der eigentliche Zweck ihrer Mission die Befreiung der Roboter ist; sie reist nämlich im Namen der Humanitätsliga. Zwischen dem Leitungsteam der Firma und Helene entspannt sich eine Debatte, ob die Roboter eine Seele haben und ob sie schmerzempfindlich gemacht werden sollten zum Zwecke des industriellen Nutzens. Und die Leiter der Firma streiten selbst, ob die Vision, dass die Roboter die menschliche Arbeit unnötig machen, nicht auch negative Komponenten hat, wie vor allem der Baumeister Alquist, Chef der Bauten von W. U. R., vertritt.

Zehn Jahre später, Helene ist Domins Frau geworden, haben sich die Dinge zum Schlechten gewandelt. Die Arbeiter haben sich gegen die Roboter empört und diese zerschlagen; die Fabrikbetreiber gaben den Robotern daher zur Selbstverteidigung Waffen, und die bewaffneten Roboter erschlugen viele Menschen. Die Regierungen setzten Roboter als Soldaten ein und vermehrten die Kriege. Einer der Roboter, Radius, hat als Sonderkonstrukt ein besonders großes Gehirn. Bei ihm hat sich ein Herrschaftsdrang entwickelt. Er ist reizbarer, zeigt offenbar Affekte.

Helene verbrennt die Pläne des alten Werstand, und kurz darauf teilen ihr die Leitungspersonen der Firma mit, dass dieses ein Pfand gegen die revoltierenden Roboter hätte sein können. Alle Pläne scheitern: die Flucht, die Konstruktion von Robotern, die einander gegenseitig bekämpfen, ein Stopp des Baus neuer Roboter. Die Fabrikzentrale ist umstellt von Robotern, von außen gibt es Botschaften, dass sie Menschen in großer Zahl töten. Gall, Leiter der Experimentalabteilung, bekennt, dass er bei einer Reihe von Robotern die „Irritabilität“ eingeführt habe, also offenbar eine gewisse Affektivität oder Emotionalität, und es stellt sich heraus, dass er damit dem Drängen von Helene nachgab, den Robotern eine Seele zu geben. Die Roboter dringen ein, töten alle Menschen außer Alquist. Der wird verschont, weil er arbeiten könne.

Im letzten Aufzug ist Alquist als letzter Mensch allein unter Robotern. Die Roboter können sich nicht vermehren, auch Alquist scheint keine Formel für neues Leben zu finden. Die Roboter nötigen ihn, einen von ihnen selbst zu sezieren, um einen Weg zu finden; auch dies offenbar ohne Erfolg. Doch es treten zwei Roboter auf, „Helene“ und „Primus“, die starke Gefühle füreinander zeigen, zudem Empfänglichkeit für die Schönheit der Natur und die Schönheit des anderen und die sogar anbieten, sich für den

anderen zu opfern. Alquist hofft, dass das Leben nicht untergeht. Soweit das Theaterstück.

Čapeks Idee ist, dass ein technisches Erzeugnis des Menschen dem Menschen Entlastung bringen soll, im Ergebnis aber den Menschen bedroht und ersetzt. Der Ansatzpunkt der Dramenhandlung liegt dabei im Bereich der Reflexion auf die Rolle der Arbeit im Kontext der Industrialisierung. Darüber hinaus erinnert *Čapeks* Szenario an Motive aus Mythos und Literatur, wo der Mensch zum Schöpfer menschenähnlichen organischen Lebens wird, man denke an Faust oder an den 1818 erschienenen Roman „Frankenstein oder der neue Prometheus“ von Mary Shelley.⁴ Auch bei *Čapek* geht es um Selbstüberschätzung und ihre Gefahren. Zudem klingt das Thema der Optimierung oder Selbstoptimierung an.

Dies alles gilt bei *Čapek*, aber es gilt nicht durchgängig für androide Roboter. Nicht immer erscheint die Ähnlichkeit zwischen Mensch und Roboter bedrohlich. Wir können durchaus auch spielerisch und ästhetisch mit der Ähnlichkeit von Roboter und Mensch umgehen. Schon seit vielen Jahren gibt es Wettkämpfe zwischen Fußball spielenden Robotern. Hier ist die Vision, sie könnten menschlichen Fußballern bald überlegen sein, offenbar sehr fern. Die Roboter wirken unbeholfen. Zuschauende beim Roboterfußball sind keine Fußballfans, sondern Technikfans.

Beim Ars Electronica Festival im September 2009 stellte Hiroshi Ishiguro Geminoid HI-1 vor, einen Roboter, der seiner eigenen Gestalt nachempfunden ist. Ishiguro spielt mit der Verwechselbarkeit. Es wird berichtet, er reise weiterhin um die Welt und habe sich selbst liften lassen, um dieses Spiel erfolgreich fortsetzen zu können (Ramge 2018, 78). Die Ähnlichkeit zwischen zwei Relata hängt von der Beschaffenheit beider Relata ab, Angleichung kann in verschiedene Richtungen erfolgen. Ähnlichkeit aber heißt auch nicht qualitative Identität oder besser qualitative Identität in jeder Hinsicht. *Čapeks* Stück macht dies deutlich durch die unterschiedlichen Visionen des alten Werstand und seines Neffen. Der Physiologe zielt auf die Gleichheit oder Ähnlichkeit der materialen Grundlage, sein Neffe, der Ingenieur, auf funktionale Entsprechungen. Diese funktionalen Entsprechungen werden durch gestalthafte Ähnlichkeiten möglich; die Roboter haben Arme und Beine etc. Aber sie haben kein Schmerzempfinden, keine Bedürfnisse, keine Seele. Darin steckt auch eine Anthropologie: Was uns Menschen von Robotern unterscheidet, ist nicht der Verstand, der Vernunftverstand, wie es in der zitierten deutschen Übersetzung heißt, sondern die Seele,

4 Shelley 2017. Der Beitrag zum Romantext durch Percy Shelley, Marys späteren Ehemann, wird in der literaturwissenschaftlichen Forschung unterschiedlich beurteilt.

die mit Bedürfnissen sowie mit positiven und negativen Affekten in Verbindung gebracht wird. Wir könnten also sagen: KI kann Gefühle lesen. Roboter können sich so verhalten, als hätten sie Gefühle. Aber KI und Roboter *haben* keine Gefühle. Diese Position wird nicht nur in der Literatur vertreten, sondern durchaus auch in der philosophischen Reflexion (vgl. Bendel o. D.).

3. Mensch und Maschine

Doch sollten wir uns damit zufriedengeben? Gibt es andere Unterscheidungen, auf die wir zugreifen können, um das Verhältnis zwischen Mensch und Roboter zu verdeutlichen? Hier mag ein geistesgeschichtlicher Rückgriff hilfreich sein. Der Vergleich zwischen Mensch und Roboter kann nämlich auf einen anderen Vergleich Bezug nehmen, den Vergleich zwischen Mensch und Maschine. Während es im 19. und 20. Jahrhundert vor allem der Vergleich zum Tier war, der den Menschen verunsicherte, so beschäftigt sich die Philosophie des 17. und 18. Jahrhunderts intensiv mit dem Maschinenvergleich. Das 19. Jahrhundert wird durch Theorie und Empirie dazu gezwungen, die alte Überzeugung von der Konstanz der Arten aufzugeben. Charles Darwin und viele andere verschaffen der Abstammungstheorie den wissenschaftlichen Durchbruch. Mit Wolfgang Köhlers Studien zum Werkzeuggebrauch von Schimpansen auf Teneriffa Anfang des 20. Jahrhunderts wird fraglich, ob der Mensch als einziges Wesen *animal rationale* ist, also ein vernunftbegabtes Sinnenwesen bzw. ein vernunftbegabtes Tier. Läßt sich der Werkzeuggebrauch bei Schimpansen und die Herstellung einfacher Instrumente anders deuten als durch Intelligenz (vgl. Köhler 1917)? So die Frage im 19. und 20. Jahrhundert. Das 17. und 18. Jahrhundert dagegen war – ganz ähnlich wie wir – durch Automaten fasziniert.⁵ Die radikalste Antwort auf die anthropologische Frage, die sich aus dieser Faszination und aus dem daraus gewonnenen Vergleich ergibt, ist jene von Julien Offray de La Mettrie: „Ziehen wir also die kühne Schlußfolgerung“, so formuliert er, „daß der Mensch eine Maschine ist, und daß es im ganzen Universum nur eine einzige Substanz – in unterschiedlicher Gestalt – gibt“ (La Mettrie 2009, 137). De la Mettrie hat sich damit zum *enfant terrible* des 18. Jahrhunderts gemacht, wenngleich eine genauere Lektüre seines Textes Anlass gibt, die Einfachheit seiner These zu bezweifeln. Sicher will

5 Einen begriffsgeschichtlich-philosophischen Überblick gewährt der Artikel „Maschine“ von W. Schmidt-Biggemann im Historischen Wörterbuch der Philosophie (vgl. Schmidt-Biggemann 1980).

er zu einer einfachen Ontologie kommen. Es gibt für ihn keine Geisterwelt und damit auch keine Seele außerhalb der körperlichen Welt. Wenn er den Menschen die „aufrecht kriechende Maschine“ nennt, dann ist das durchaus abfällig. Dennoch spricht er von der „erleuchteten“ Maschine und das heißt einer Maschine, in der ein Licht wirkt, ein Licht der Vernunft (vgl. Tetens 1999). Und schon der menschliche Körper ist nicht ganz wie die Maschinen seiner Zeit, denn Tiere und Menschen sind Maschinen, „die selbst ihre Triebfedern“ aufziehen. De La Mettrie zwingt uns dennoch, auch den Denkvorgang als Vorgang in der Körperwelt zu denken und damit das zu tun, was auch heutige Neurowissenschaftler*innen tun.

Anders verhält sich dies bei René Descartes. Er treibt den Vergleich zwischen Tier und Maschine sehr weit. Die Lebensvorgänge, Wahrnehmung und Vorstellung, Wachheit und Schlaf, dies alles, so meint er, lasse sich mit einer komplexen Mechanik und Materialkenntnis nachbilden. Er schreibt:

„Alle diese Dinge hatte ich ziemlich detailliert in der Abhandlung erklärt, die zu veröffentlichen ich ehemals den Plan gehabt hatte. Ich hatte anschließend gezeigt, welches die Machart der Nerven und Muskeln des menschlichen Körpers sein mußte, um zu bewerkstelligen, daß die Lebensgeister (*les esprits animaux*) in ihm die Kraft haben, die Glieder zu bewegen – wie man ja sieht, daß Köpfe, kurz nachdem sie abgeschlagen sind, sich noch regen und ins Gras beißen, obwohl sie nicht mehr beseelt sind –; welche Veränderungen im Gehirn vonstatten gehen müssen, um den Wachzustand zu verursachen, den Schlaf und die Träume; wie das Licht, die Töne, die Gerüche, die Geschmäcke, die Wärme und alle anderen Qualitäten der äußeren Objekte durch Vermittlung der Sinne in ihm verschiedene Ideen einprägen können; wie der Hunger, der Durst und die anderen inneren Leidenschaften ebenfalls ihre Ideen in es senden können; was in ihm als diese Ideen empfangender Gemeinsinn aufzufassen ist; als das sie aufbewahrende Gedächtnis; und als Phantasie, die sie verschiedentlich verändern, aus ihnen neue zusammensetzen kann, dadurch die Lebensgeister auf die Muskeln verteilt und so die Glieder des Körpers sich auf dementsprechend viele verschiedene Weisen bewegen läßt; und wie sich unsere Körperteile anlässlich der sich den Sinnen des Körpers präsentierenden Objekte und der inneren Leidenschaften bewegen können, ohne daß der Wille sie leitet.“ (Descartes 2011, 95–97)

Die Lebensgeister, von denen er spricht, sind daher als Antriebe gedacht, die durchaus technologisch nachbildbar sind und allein auf materialer und mechanischer Grundlage agieren. Deshalb schreibt er weiter:

„All dies wird denjenigen überhaupt nicht seltsam erscheinen, die wissen, wie viele verschiedene *Automaten* oder selbstbewegliche Maschinen der Einfallsreichtum der Menschen bewerkstelligen kann, der dafür im Vergleich mit der großen Menge an Knochen, Muskeln, Nerven, Arterien, Venen und all der anderen Bestandteile im Körper jedes Tieres nur sehr wenige Teile verwendet; und die deshalb diesen Körper als eine von den Händen Gottes hergestellte Maschine betrachten, die

unvergleichlich viel wohlgeordneter ist und bewundernswere Bewegungen aufweist als irgendeine von denen, die von den Menschen erfunden werden können. Hier hatte ich mich besonders aufgehalten, um zu zeigen, daß, wenn es solche Maschinen mit den Organen und der Gestalt eines Affen oder irgendeines anderen Tieres ohne Vernunft gäbe, wir keinerlei Mittel besäßen, zu erkennen, daß sie nicht in allem dieselbe Natur hätten wie diese Tiere.“ (Descartes 2011, 97)

Descartes sieht also nicht nur eine Analogie zwischen Tier und Maschine, sondern es gibt auch keine Unterscheidungsmöglichkeit, sofern die Komplexität der Maschine sich jener der Tiere angleicht. Anders als bei den Tieren ist dies hingegen, wie Descartes deutlich machen will, beim Menschen. Menschen unterscheiden sich von ihnen nachgebildeten Automaten:

„Gäbe es hingegen solche, die Ähnlichkeit mit unseren Körpern besäßen und unsere Handlungen soweit nachahmten, wie es praktisch (moralement) möglich wäre, hätten wir immer zwei sehr sichere Mittel, um zu erkennen, daß sie deswegen keineswegs schon wahre Menschen sind. Das erste ist: Sie könnten niemals Worte oder andere Zeichen gebrauchen, indem sie sie zusammensetzen, wie wir es tun, um anderen unsere Gedanken kundzutun. Denn man kann sehr gut verstehen, daß eine Maschine so gebaut sein soll, Worte zu äußern, und man kann sogar verstehen, wenn sie einige Worte anlässlich körperlicher Vorgänge äußert, die irgendeine Veränderung in ihren Organen verursachen: etwa daß sie, wenn man sie an irgendeiner Stelle berührt, fragt, was man ihr sagen wolle, oder daß sie, berührt man sie an einer anderen Stelle, schreit, man tue ihr weh und dergleichen. Aber man kann nicht verstehen, daß sie Worte verschieden zusammenstellt, um auf den Sinn alles dessen zu antworten, was in ihrer Gegenwart gesagt werden wird, wie es selbst die stumpfsinnigsten Menschen tun können.“ (Descartes 2011, 97)

Descartes nennt ein zweites Argument: Auch wenn solche Maschinen viele Dinge ebenso gut oder vielleicht sogar besser als irgendeiner von uns verrichten würden, so meint er, dass diese dennoch bei einigen anderen versagten. So ließe sich entdecken, dass sie nicht aus Kenntnis tätig sind, sondern nur aufgrund der Disposition ihrer Organe.

„Denn anders als die Vernunft, die ein Universalinstrument ist, das bei allen Arten von Begebenheiten benutzt werden kann, benötigen diese Organe eine ganz bestimmte Anordnung für jede besondere Tätigkeit, und deshalb ist es praktisch (moralement) unmöglich, daß es genügend verschiedene Organe in einer Maschine gibt, um sie in allen Vorfällen des Lebens in derselben Weise wie unsere Vernunft tätig sein zu lassen.“ (Descartes 2011, 99)

Syntax und Semantik und das Verstehen von Syntax und Semantik sollen also Alleinstellungsmerkmale sein sowie auch der Vernunftgebrauch in seiner universellen Anlage. Gottfried Wilhelm Leibniz ist noch weit vorsichtiger als Descartes. Schon auf der Ebene der Wahrnehmung sieht er

Probleme. Perzeption, so heißt das bei ihm, könne nicht durch mechanische Gründe erklärt werden:

„Man muß übrigens zugestehen, daß die Perzeption und was von ihr abhängt *durch mechanische Gründe*, d. h. durch Figuren und durch Bewegungen, *unerklärbar* ist. Angenommen, es gäbe eine Maschine, deren Struktur zu denken, zu fühlen und Perzeptionen zu haben erlaubte, so könnte man sich diese derart proportional vergrößert vorstellen, daß man in sie eintreten könnte wie in eine Mühle. Dies vorausgesetzt, würde man, indem man sie von innen besichtigt, nur Teile finden, die sich gegenseitig stoßen, und niemals etwas, das eine Perzeption erklären könntet. Also muß man danach in der einfachen Substanz und nicht im Zusammengesetzten oder in der Maschine suchen. Es gibt zudem überhaupt nur Perzeptionen und deren Veränderungen in einer einfachen Substanz. Und daraus allein können alle *inneren Tätigkeiten* der einfachen Substanzen bestehen.“ (Leibniz 1998, § 17, 19)

Können uns diese Antworten zufriedenstellen? Gehen nicht beide, Descartes und Leibniz, von einer Eigenständigkeit des Geistigen aus, die wir als Rezipienten natur- und neurowissenschaftlicher Erkenntnisse nicht akzeptieren können? Sind sie nicht durch technische Entwicklungen widerlegt? Verfügen nicht Roboter über Sensoren und über Spracherkennungsprogramme? Meines Erachtens kann man sich diesen Fragen eher nähern, wenn wir den Blick über die Robotik hinaus erweitern. Einen wichtigen Schritt hierzu hat Alan Turing getan. Er greift die Fragen von Descartes und Leibniz auf, gibt dem Vergleich von Mensch und Maschine aber eine neue Wendung. Wie beide war auch er nicht nur an philosophischen Fragen interessiert, sondern er war Mathematiker. Wie Leibniz ging es auch ihm um formalisierte Operationen. Leibniz und Turing stehen in der Geschichte der theoretischen Entwicklung des Computers. Nicht zuletzt ist Turing berühmt, weil er die Entschlüsselung des deutschen Codierungssystems „Enigma“ im 2. Weltkrieg bewerkstelligte. Seine Idee der Turingmaschine, eine Maschine also, die so gedacht wird, dass sie in der Lage ist, die Tätigkeit jeder anderen Maschine zu simulieren (vgl. Turing 1937), bringt die Maschine theoretisch in die Nähe dessen, was Descartes nur dem universalen Instrument der Vernunft zutraute. Wichtiger noch für unsere Fragestellung ist, dass er das Argument der Ununterscheidbarkeit, das bei Descartes für die Nähe von Maschine und Tier benutzt wurde, auf die Nähe von Maschine und Mensch bezieht. Turings Frage lautet: Können Maschinen denken? Er weicht dieser Frage zunächst aus, weil er Schwierigkeiten in der Definition des Denkens anführt. Stattdessen schlägt er ein Imitationsspiel vor, das später unter dem Namen Turing-Test zitiert wurde. Bei dem Spiel ist der Blick auf den Körper, das körperliche Substrat verstellt. An sprachlicher Äußerung, der Performanz also, sollen wir ent-

scheiden, ob die Sprecher*in ein Mensch oder eine Maschine ist. Man wird vielleicht an Lautsprecher in der Straßenbahn oder am Bahnhof denken, bei denen sich diese Frage heutzutage oftmals stellt, doch das ist nicht gemeint. Auch die Stimmqualität darf keine Hilfestellung bei der Antwort bieten. Turing empfiehlt deshalb einen Fernschreiber als Vermittler. Der Autor ist sich der Vereinfachung bewusst: Der Konstrukteur einer denkenden Maschine braucht sich nicht mit künstlichem Fleisch aufzuhalten. Er kann sich sozusagen auf das maschinelle Denken beschränken. Turing geht davon aus, dass Rechenmaschinen mit potentiell unbegrenzter Kapazität das Imitationsspiel erfolgreich spielen können, d.h., dass sie Antworten auf menschliche Fragesteller formulieren können, die für den Fragesteller nicht als bloß maschinelle Antwort identifiziert werden können. Der Mensch meint, es handle sich um einen Menschen. Turing formuliert eine Reihe von Einwänden, die er teilweise abwendet, teilweise relativiert. Jedenfalls wirbt er dafür, sich die Vorstellung einer lernenden Maschine plausibel zu machen und sieht im Schachspiel ein erstes Feld, in dem eine solche Maschine erprobt werden könnte.

Damit ist klar: Nicht nur der Roboter als „Verkörperung der Digitalisierung“ (Steil 2019, 15), sondern auch der Computer, bei dem es auf die Verkörperungsweise nicht ankommt, wirft die Frage nach einer Abgrenzung auf. Diese Abgrenzungsfrage ist eine anthropologische Frage, eine Frage danach, was der Mensch ist, und welches Bild wir uns von uns machen.

4. Computer und selbstlernende Systeme

Die Geschichte des Computers zeigt in beeindruckender Weise, wie ein technisches Gerät in kurzer Zeit leistungsfähig werden kann und dadurch immer mehr Bereiche des menschlichen Lebens mitbestimmt. Diese Leistung kann als Rechenleistung oder als Informationsverarbeitung beschrieben werden. Der Computer kann Aufgaben übernehmen, die zuvor nur der Mensch leisten konnte und welche dieser seinen kognitiven oder rationalen Vermögen zugeordnet hat. Deshalb ist die Rede von künstlicher Intelligenz gut nachvollziehbar. Auch ist nicht überraschend, dass Formen künstlicher Intelligenz in einen Vergleich zur natürlichen Intelligenz als der Intelligenz des Menschen gesetzt wurden.

Dass wir technische Systeme personifizieren, ist den meisten Zeitgenoss*innen wahrscheinlich nicht fremd. Schon lange ist unser Personal Computer nicht nur ein Computer *für* die Person, sondern ein technisches Konstrukt, in dem wir Intentionalität vermuten. Wir fluchen über diesen

nicht wie über schlechtes Wetter oder einen kaputten Rasierapparat, sondern wie über einen Handelnden. Freilich ist uns zumeist klar, dass solche Personalisierungen inadäquat sind. Doch dies könnte mit künstlicher Intelligenz (KI) anders sein. Im Umgang mit der KI-Forschung muss jedenfalls auch gefragt werden, welche Leistungen es sind, die als intelligent gelten. Starke KI scheint umfassende Leistungen vor Augen zu haben, die kognitive, emotionale und moralische Einstellungen und Operationen umfassen. Starke KI ist bislang nur Vision und in ihrer Realisierbarkeit umstritten. Schwache KI dagegen existiert und kann vorweisen, im Bereich des Bilderkennens und Analysierens erstaunliche Fähigkeiten zu konstruieren: Es gibt Möglichkeiten der Gesichtserkennung, des Abgleichs von Fingerabdrücken, der Iris oder von Handschriften. Muster können in bestimmten Zusammenhängen nicht nur erkannt, sondern auch vorhergesagt werden. Menschliche Gesichtsausdrücke können typisierten Gefühlszuständen zugeordnet werden. Im Bereich der natürlichen Sprache können Lautfolgen in Schrift umgesetzt werden und umgekehrt sowie Texte verschiedener natürlicher Sprachen übersetzt werden. Es werden also nicht nur Laute und Buchstaben assoziiert, sondern auch Semantik und Syntax analysiert. Die Systeme Amazon-Echo und Siri nutzen diese Möglichkeiten der Bedeutungsanalyse.

Als besonders beeindruckend gilt Alpha Go. Das Go-Spiel hat KI Expert*innen beschäftigt, nachdem 1997 der IBM Computer Deep Blue den amtierenden Schach-Weltmeister Garri Kasparov besiegte. Allerdings muss man sich Folgendes klarmachen: Wahrscheinlich würde Deep Blue das Imitationsspiel nicht einmal für den Bereich des Schachspiels gewinnen. Zumindest eine erfahrene Spieler*in würde die Strategien von Computer und Mensch zuordnen können. Aber Deep Blue sollte Kasparov nicht imitieren, sondern besiegen. Und auch eine Überlegenheit des Computers ist nicht ohne anthropologische Relevanz. Go hat weit mehr Zugmöglichkeiten als das Schachspiel. Man sagt, dass gute Go-Spielende nicht nur kalkulieren können, sondern Logik und Intuition verbinden. Die Wissenschaftler*innen bei Google verbanden zunächst Programmierung und Selbstlerneffekte mit dem Ergebnis, dass das Programm den Weltmeister besiegte (2016). Dann aber verzichtete man auch auf alle Vorgaben, die über die Spielregeln hinausgingen. Zu diesen Regeln gehört auch die Vorgabe, dass eine Spieler*in versucht, das Spiel zu gewinnen. Ende 2017 wurde in der Zeitschrift *Nature* publiziert, das allein auf Selbstlernen gegründete System spiele nun erfolgreich. Es habe gegen das datengestützte Programm 100 zu 0 gespielt. Das Programm lernt, indem es gegen sich selbst spielt. Die Autor*innen beschreiben das Programm als Lehrer*in

seiner selbst. Und dieses Programm besiegt das Programm, welches den Weltmeister besiegt hat (Silver et al. 2017).

Was heißt das nun für uns? Sollten wir als Menschen bescheiden auf jede Sonderstellung verzichten? Oder sollten wir, wie Karel Čapek es durch sein Stück suggeriert, alle Besonderheit in unserer Lebendigkeit und in unseren authentischen Gefühlen erblicken? Ich denke nicht, dass dies die einzigen möglichen Strategien sind.

5. Zur Begriffsverwendung von „Autonomie“, „Intelligenz“ und „Lernen“

Als sich die *European Group on Ethics in Science and New Technologies* bei der Europäischen Kommission 2018 mit dem Fragenkomplex dieses Beitrags befasst hat, kam dabei ein „Statement on Artificial Intelligence, Robotics and ‚Autonomous‘ Systems“ heraus. Dabei wird „autonomous“ in Anführungszeichen gesetzt. Aber eigentlich verdienen alle Begriffe, die im Titel aufscheinen, eine nähere Analyse. In der Tat gibt es nicht erst durch die autonomen Systeme Schwierigkeiten mit dem Autonomiebegriff.

„In der Systemtechnik ist der Begriff der Autonomie mathematisch eindeutig definiert. Autonom ist ein gegebenes technisches System dann, wenn es alleine auf Grund seiner inneren Zustände zur Zeit t_0 auch in Zukunft ($t > t_0$) vollständig beschreibbar ist. Ein Beispiel ist ein störungsfrei aufgehängtes und schwingendes Pendel: wenn der Pendelausschlag zu Beginn der Betrachtung bekannt ist, lässt sich das weitere Verhalten vorhersagen, da es nach einer inneren Gesetzmäßigkeit, ohne Einfluss von außen, abläuft.“ (Christaller et al. 2001, 36)

In der neueren speziellen Ethik gilt Vielen Autonomie als Synonym für Selbstbestimmung. Manchmal wird es auch als Synonym für die Fähigkeit zur Selbstbestimmung verwendet. Und schließlich, das gilt Philosoph*innen vom Europäischen Kontinent zumeist als Krone der Bedeutungserklärungen, besagt „Autonomie“ Selbstgesetzgebung. Griechische Stadtstaaten waren dann autonom, wenn sie ihre Gesetze selbst erließen. Eine Person ist bei Kant dadurch ausgezeichnet, dass sie sich nicht einfach frei von Fremdbestimmung verhält, sondern dass sie sich selbst ein Gesetz gibt, kraft ihrer Vernunft. Sie stellt sich unter das Gesetz (Kant 1984). Auch wenn autonome Systeme gerade dem systemtechnischen Verständnis von Autonomie entsprechen und künstliche Intelligenz in die Lage versetzt werden kann, Regeln anzupassen und fortzuschreiben, so scheint dies nicht eine moralische Autonomie zu sein, wie sie für Kant mit dem Begriff der Person gegeben ist.

Auch der Begriff der Intelligenz verdient eine längere Erläuterung. Bei den Intelligenzprüfungen von Menschenaffen ging es vor allem auch darum, ob sie Problemlösungen nicht durch Versuch und Irrtum, sondern spontan zustande bringen. Sie erfassen die Eignung von Stöcken oder Kisten bei geeigneter Lage im Raum, als Stapel oder als Anstückung geeigneter Werkzeuge zu sein. Wie ist das mit einem Computer, der verschiedene Möglichkeiten durchrechnet?

Nicht zuletzt verdient der Begriff des Lernens besondere Beachtung. Während ich immer versucht habe, Ausdrücke wie Autonomie oder Intelligenz nicht oder nur mit Vorbehalten auf künstliche Systeme anzuwenden, gilt das beim Lernen nicht. Ein Computer lernt, wie er im Schach unter gegebenen Bedingungen gewinnen kann. Will man für das Lernen eine allgemeine Bestimmung finden, unter die auch das Lernen künstlicher Systeme fällt, so kann man Jochen Steil vom Institut für Robotik und Prozessinformatik der Technischen Universität Braunschweig zitieren, der Spezialist für Roboterlernen ist: „Im Sinne einer Arbeitsdefinition bezeichne ich mit Lernen [...] einen Prozess, der Erfahrungen in der realen Welt für zukünftige Handlungen so verarbeitet und konsolidiert, dass sie generalisierbar, d. h. auf neue Situationen anwendbar sind“ (Steil 2019, 18).

Aber ein Computerprogramm und auch ein Roboter lernen anders, als Menschen dies tun. Ich nehme die bekannte Bekehrungsgeschichte des Buddha: „Der Prinz sieht *einen* Armen, *einen* Kranken, *einen* Toten, nachdem er im Palaste des Vaters jahrelang allen negativen Eindrücken ferngehalten“ worden war; „er erfaßt aber jene drei zufälligen ‚jetzt-hier-so-seienden‘ Tatsachen sofort als bloße Beispiele für eine an ihnen erfaßbare essentielle Weltbeschaffenheit“ (Scheler 2002, 46). Ob die Geschichte stimmt oder nicht, wir können sie jedenfalls verstehen und deuten sie auch nicht als ein Wunder. Vielmehr scheint es mir, dass wir selbst in ähnlicher Weise Einsichten gewinnen können.

Für moralische Autonomie, für Intelligenz in einem bestimmten Sinne und auch für menschliches Lernen sehe ich daher zumindest derzeit keine vergleichbaren Phänomene in der von uns geschaffenen Welt. Und vielleicht ist noch ein weiterer Punkt zu nennen, an dem ein Unterschied sichtbar wird. Fragt man Expert*innen aus dem Bereich der KI-Forschung nach Alpha Go und Alpha Go Zero, dann beschreiben sie das Lernen wie folgt: Das Programm lernt in kurzer Zeit mit dem Erfolg, dass es nach wenigen Stunden dem besten menschlichen Go-Spieler überlegen ist. Es holt also das Erfahrungswissen und Know how, was in Jahrtausenden erspielt und gesammelt wurde, in ganz kurzer Zeit auf. Es spielt gegen sich

selbst. Was heißt das? Es spielt gegen andere Sektionen seiner selbst, ohne zu erkennen, dass es Sektionen seines Selbst sind.

Freilich wäre es wohl überzogen zu erklären, solche Erkenntnis könne durch künstliche Systeme niemals erlangt werden. Derzeit jedoch gibt es sie nicht. Es gibt keine Selbstbezüglichkeit, keinen Ich-Begriff, kein praktisches Selbstverhältnis. All dies wäre nötig, um von Selbstbewusstsein oder auch von Gewissen zu sprechen, wie wir es beim Menschen tun.

6. Menschliche Selbstachtung und Selbstschätzung

In einer neuen Phase der Technik stärkt und verlängert der Mensch nicht nur seine Organe, sondern er delegiert ganze Prozesse, auch Analysen und Entscheidungsprozesse an künstliche Systeme. Im Idealfall bietet diese Stufe der Technik auch eine höhere Stufe der Entlastung. Solche Übertragung menschlicher Funktionen auf künstliche, technische Systeme setzt aber voraus, dass die natürlichen Funktionen technikartig modelliert werden (vgl. Hubig 2017, 87). Nur dann sind sie anschlussfähig und übertragbar. Auf das Konzept der selbstlernenden Systeme und der Roboter übertragen könnte man sagen, dass wir uns als Maschinen modellieren, damit wir technischen Systemen beibringen können, uns zu imitieren und mit uns zu kooperieren. Roboter und selbstlernende Systeme haben also quasi menschliche Gestalt. Die Modellierung beinhaltet vor allem eine Reduktion. Wir vernachlässigen jene Momente an uns, die sich nicht nach dem Modell technischer Systeme verstehen lassen. Dazu scheinen bestimmte Formen authentischer Emotionalität, Selbstbewusstsein oder Willensfreiheit zu gehören. Zudem könnte man auch das Geborensein oder das Bedürfnis zu teilen und mitzuteilen nennen. Empirische Wissenschaft lebt von derartigen Modellbildungen und Reduktionismen. Technisch-wissenschaftliches Handeln wäre niemals so effizient geworden, hätte man stets nur Komplexität, Unübersichtlichkeit, Ganzheit beschworen. Die ethischen Probleme entstehen, wenn man diese Reduktion nicht mehr als methodische Reduktion wahrnimmt, sondern sie als adäquate Selbstbeschreibung empfindet. Wir haben die Chance, uns als „aufrecht kriechende Maschinen“, „die selbst ihre Triebfedern“ aufziehen, zu verstehen oder eben auch nicht.

7. Der Status von Maschinen und Robotern

Damit sind wir auch bei einer weiteren Frageebene, die bislang nicht genannt wurde, die aber für einige Philosoph*innen derzeit die erste ist, wenn über Roboterethik oder Maschinenethik gesprochen wird. Robotereithik behandelt eben nicht nur, wie wir Roboter einsetzen, um Menschen oder Tieren nicht zu schaden und möglichst zu nutzen. Vielmehr wird diskutiert, ob wir Roboter als moralische Agenten, also als moralisch Tätige ansehen müssen, die moralische Ansprüche und Rechte uns gegenüber haben. Oft geschieht dies im Gestus der Erweiterung der Gruppe der moralisch zu beachtenden Wesen. Aus den Hinweisen zu den Unterschieden zwischen moralischen Akteur*innen und Maschinen ergibt sich indes, dass es verfehlt ist, wenn man Roboter in unsere moralische Gemeinschaft aufnehmen will. KI lernt Normativität aus den Daten, Erfahrungen, Stereotypen. Die Normativität im Sinne moralischer Geltung muss allen künstlichen Systemen vorgegeben werden, alles Lernen ist hier Anpassung und Anwendung. Eine Änderung dieses Umstandes ist auch bei großzügiger Betrachtung der Möglichkeiten nicht in Sicht. Die Anpassung an gelebte Einstellungen, die in Datensammlungen vorliegen, ist aber kein moralisches Lernen. Insofern gehe ich davon aus, dass Roboter und künstliche Systeme keine Moralität im Sinne einer autonomen moralischen Gesetzgebung haben und nach meinem Verständnis auch nicht haben können. Sie sind Instrumente unseres moralischen oder unmoralischen Verhaltens, weder Objekte noch Subjekte von Moral.

Die voreilige Gleichstellung von Robotern und Menschen zieht auch den Verlust eines Respekts gegenüber den Menschen mit sich, weil gerade die wichtigen Besonderheiten des Menschen dann nicht hinreichend gewürdigt werden. Auch wenn Roboter Roboter bauen, steht am Anfang der Kette ein Mensch, dessen Kreativität und Planungsfähigkeit nicht nur den Anfang gemacht hat, sondern alle vorhandenen Rahmen vorgegeben hat. Dies bedeutet im Blick auf die sogenannten autonomen Systeme, dass moralische Verantwortung da gebündelt werden muss, wo die Voraussetzungen für das System geschaffen werden. Die Erkenntnis, dass wir moralisch autonome Wesen sind, gebietet in einem moralischen Sinne die Selbstachtung.

8. Handlungsregeln für den Einsatz von Robotern und KI

Diese Überlegungen vermitteln ein Verständnis vom Menschen, das im Begriff der Würde prägnant ausgedrückt ist. Es geht davon aus, dass diese Würde mit dem im Menschsein gegebenen Potential zur Moralität gegeben ist. Es verlangt keinen Akt der Anerkennung durch andere. Der Vorschlag, am Konzept der Würde des Menschen als einer besonderen Würde festzuhalten, gibt dann für viele Fragen, die sich an den Einsatz von Robotern und künstlicher Intelligenz richten, eine gute Richtschnur. Zum Beispiel legt es nahe, Pflegeroboter nicht so einzusetzen, dass die vermeintlich Nutznießenden getäuscht werden über deren Natur als Roboter. Freilich sind das Konzept der Menschenwürde und das mit ihm einhergehende Verbot, Menschen bloß als Mittel und nicht auch immer als Zweck zu behandeln nicht ausreichend, um alle Fragen zu klären. Maßgeblich und dennoch schwierig in der Konkretion ist die Menschenwürde bei der Beurteilung von Cyborgs und auch jeder anderen Art von Mensch-Maschine-Schnittstellen. Wann gefährden wir durch radikale Prothetik und Computerimplantate, dass wir uns selbst noch als Menschen mit lebendigem Leib und eigenem Denk- und Urteilsvermögen wahrnehmen? Eine vorausschauende Ethik wird diese Frage im Blick behalten müssen und vielleicht schon bald eine provisorische Antwort geben müssen.

Literaturverzeichnis

- Alves-Oliveira, Patrícia. et al. 2015. "Social robots for older adults: Framework of activities for aging in place with robots." In *Social Robotics: 7th International Conference*, edited by Tapus, A. et al. ICSR 2015. Paris, France. October 26–30, Proceedings. Heidelberg, 11–20.
- Bendel, Oliver. O. D. „Soziale Roboter.“ <https://wirtschaftslexikon.gabler.de/definition/soziale-roboter-122268>.
- Čapek, Karel. 2017. *W.U.R. Werstands universal Robots* (Originaltitel: «R.U.R. Rossum's Universal Robots». Erstdruck 1920.) Neuausgabe hrsg. von Karl-Maria Guth. Berlin: Hofenberg.
- Capurro, Rafael. 2017. *Homo Digitalis. Beiträge zur Ontologie, Anthropologie und Ethik der digitalen Technik*. Wiesbaden: Springer VS.
- Christaller, Thomas et al. 2001, „Was erwarten wir von intelligenten Robotern?“ In *Robotik: Perspektiven für menschliches Handeln in der zukünftigen Gesellschaft*, 31–45. Berlin/Heidelberg: Springer.

- Descartes, René. 2011. *Discours de la Méthode*. Frz./dt., übers. u. hrsg. v. Christian Wohlers, Hamburg: Meiner.
- European Group on Ethics in Science and New Technologies. 2018. *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*. Brussels. https://ec.europa.eu/info/research-and-innovation_en.
- Hubig, Christoph. 2017. „Der ‚biofaktische‘ Mensch zwischen Autonomie und Technomorphie.“ In *Gute Wissenschaft, Theorie, Ethik und Politik*, hrsg. von Michael Spieker und Arne Manzeschke, 87-101. Baden-Baden: Nomos.
- Innovation in autonomous systems, Summary of an event held on Monday 22 June 2015 at the Royal Academy of Engineering.
- Junyang Li, Xiaojian Li et al. 2018. “Development of a magnetic microrobot for carrying and delivering targeted cells.” *Science Robotics*, 27 June (Vol. 3), Issue 19. <http://dx.doi.org/10.1126/scirobotics.aat8829>.
- Kant, Immanuel. 1984. *Grundlegung zur Metaphysik der Sitten*, hrsg. V. Theodor Valentiner, Stuttgart: Reclam.
- Köhler, Wolfgang. 1917. *Intelligenzprüfungen an Anthropoiden*. Berlin: Verlag der Königlichen Akademie der Wissenschaften.
- La Mettrie, Julien Offray de. 2009. *Die Maschine Mensch*, übersetzt u. hrsg. v. Claudia Becker, Hamburg: Meiner.
- Leibniz, Gottfried Wilhelm. 1998. *Monadologie*, Frz./Dt., übersetzt und hrsg. v. Hartmut Hecht, Stuttgart: Reclam.
- Misselhorn, Catrin. 2018. *Grundfragen der Maschinenethik*. Stuttgart: Reclam.
- Ohly, Lukas. 2019. *Ethik der Robotik und der Künstlichen Intelligenz*. Berlin: Lang.
- Ramge, Thomas. 2018. *Mensch und Maschine. Wie künstliche Intelligenz und Roboter unser Leben verändern*. Stuttgart: Reclam.
- Scheler, Max. 2002. *Die Stellung des Menschen im Kosmos*. 15. Aufl. Bonn: Bouvier.
- Schmidt-Biggemann. 1980. Wilhelm, „Maschine.“ In *Historisches Wörterbuch der Philosophie*, Bd. 5 : 790–802.
- Shelley, Mary. 2017. *Frankenstein: Annotated for Scientists, Engineers, and Creators of All Kinds*, ed. by David H. Guston et al., Cambridge MA: MIT press.
- Silver, David et al. 2017. “Mastering the game of Go without human knowledge.” *Nature* 550, Nr. 7676 (2017): 354–359.
- Steil, Jochen. 2019. „Roboterlernen ohne Grenze? Lernende Roboter und ethische Fragen.“ In *Roboter in der Gesellschaft: technische Möglichkeiten*

- und menschliche Verantwortung*, hrsg. von Christiane Woopen und Marc Jannes, 15–33. Berlin: Springer.
- Tetens, Holm. 1999. “L’Homme Machine: Die erleuchtete Maschine. Das neurokybernetische Modell des Menschen und die späte Ehrenrettung für den Philosophen Julien Offray de La Mettrie.” In *Die Zeit*, 10. Juni 1999.
- Turing, Alan. 1937. “On Computable Numbers, with an Application to the Entscheidungsproblem.” In *Proceedings of the London Mathematical Society*, Volume 42: 230–265.
- Turing, Alan. 1994. „Kann eine Maschine denken?“ In *Künstliche Intelligenz. Philosophische Probleme*, hrsg. von Walther Ch. Zimmerli und Stefan Wolf, 39–78. Stuttgart: Reclam.
- Verein Deutscher Ingenieure 1990. VDI-Richtlinie 2860. <https://www.vdi.de/richtlinien>.
- Wynsberghe, Aimee van, 2013, “Designing robots for care: care centered value-sensitive design”, *Science and engineering ethics*, Vol. 19 (2): 407–433.