

OLIVER MÜLLER

Maschinelle Alterität

Philosophische Perspektiven auf Begegnungen mit künstlicher Intelligenz

1. Einleitung: Sozialität und Alterität in der Mensch-Maschine-Begegnung

Die Entwicklungen auf dem Gebiet der Künstlichen Intelligenz durchdringen viele unserer Lebensbereiche, von den inzwischen allgegenwärtigen Produktempfehlungen neunmalkluger Algorithmen über die Implementierung von Gesichtserkennungsmethoden in der Öffentlichkeit und den Auswüchsen eines digitalen Überwachungskapitalismus bis hin zu robotischer Unterstützung in Industrie und Militär. Und es steht außer Frage, dass wir nicht nur unseren rechtlichen Rahmen nachjustieren müssen – etwa im Bereich des Datenschutzes oder des Völkerrechts, was ›autonome‹ Waffensysteme betrifft –, sondern dass wir auch ethischer Selbstverständigungsdiskurse bedürfen, etwa in Bezug auf den ›bias‹, also die gesellschaftlichen Vorurteile, die in den ›selbstlernenden‹ maschinellen Verfahren reproduziert werden, oder ganz allgemein in Bezug auf Fragen des guten Lebens, der (globalen) Gerechtigkeit und der Werte des Zusammenlebens in der digitalisierten Welt.¹

- 1 Siehe exemplarisch für diese Debatten: Silja Vöneky, Philipp Kellmeyer, Wolfram Burgard, Oliver Müller (eds.): *The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives*. Cambridge 2022 (i. Ersch.); Shoshana Zuboff: *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs 2019; Kate Crawford: *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven 2021; Patrick Lin, Keith Abney, and Ryan Jenkins (eds.): *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. New York: Oxford University Press 2017; Marco Nørskov (ed.): *Social Robots*, London: Routledge 2017; Luciano Floridi et al.: »AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations«. *Minds and Machines*, 28(4) (2018): 689–707; David J. Gunkel and Joanna J. Bryson (eds.): »Machine Morality: The Machine as Moral Agent and Patient«. Special issue of *Philosophy & Technology*, 27(1) (2014): 1–142; Michael Anderson and Susan Leigh Anderson (Hg.): *Machine Ethics*. Cambridge: Cambridge University

Die philosophische Reflexion künstlicher Intelligenzen hat sich früh auf die Frage kapriziert, ob und inwiefern künstlich-intelligente Systeme den Status von rationalen und/oder moralischen Akteur:innen haben können.² So naheliegend eine solche Frageperspektive und die entsprechende Klassifizierungspraxis in Bezug auf den Akteur:innenstatus auch sein mögen, derartige Vorhaben haben auch ihre Tücken. Denn sie tragen dazu bei, dass wir den künstlich-intelligenten Systemen Entscheidungs-, Handlungs- und Verhaltenskompetenzen zuschreiben, die sich an den Fähigkeiten von Menschen orientieren. Mit den Begriffsbildungen um den Akteur:innenstatus behandeln wir die komplexen Algorithmen nicht mehr als nützliche Werkzeuge, sondern statten sie mit menschen- oder personenähnlichen Eigenschaften und Kompetenzen aus. Dazu gehört auch der Begriff der Autonomie, der die künstlichen Intelligenzen mit einer erhabenen Aura versieht und den Blick dafür verstellt, dass wir hier letztlich Anthropomorphisierungen aufsitzen. Denn wie beeindruckend eigenständig die Maschinen auch operieren können, mit der Autonomie, die wir Menschen zuschreiben, hat das wenig zu tun. Auch wenn wir den Akteur:innenbegriff in größtmöglicher analytischer Abstraktion behandeln, bleibt am Ende die Suggestion, die künstlichen Intelligenzen sind eigentlich so wie wir – nur eben etwas anders. Dieser Suggestion verfallen wir auch daher gerne, weil dem Begriff ›künstliche Intelligenz‹ der Assoziationsraum des ›Geistes in der Maschine‹ und ähnlicher (post-) cartesianischer Reflexionstraditionen anhaftet. Das heißt: im Zentrum steht die schauernde Faszination der Frage, ob die Maschinen einen uns Menschen vergleichbaren Geist haben (können). Dass auf diese Weise allein schon verdeckt wird, dass hinter den scheinbar körperlos geistähnlichen künstlichen Intelligenzen ein auf Ausbeutung (von Rohstoffen und Personen) ausgerichtetes Industriesystem steht, ist dabei eines der gewichtigen ethischen Probleme in diesem Feld (wie Kate Crawford in ihrem *Atlas of AI* eindrucksvoll gezeigt hat).³ Dies sind (in aller Kürze) einige der Gründe, warum wir dazu neigen, künstliche Intelligenzen anders als andere Technologien zu behandeln – weil wir ihnen gewisse agentielle Fähigkeiten zuschreiben.

Bei aller Skepsis, die wir gegenüber dieser Zuschreibung geistiger und agentieller Fähigkeiten an die intelligenten Maschinen haben mögen,

Press 2007; Sarah Spiekermann: *Digitale Ethik. Ein Wertesystem für das 21. Jahrhundert*. München 2019.

- 2 Siehe Luciano Floridi und J.W. Saunders: »On the Morality of Artificial Agents«. *Minds and Machine*, 14 (2004): 349–379; Catrin Misselhorn: *Grundfragen der Maschinenethik*. Stuttgart 2018, S. 70–90; Johanna Wagner: *Künstliche Intelligenzen als moralisch verantwortliche Akteure*. Paderborn 2020.
- 3 Kate Crawford: *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven 2021.

bleibt die Lage aber verzwickter, wenn wir uns nicht nur die Technik für sich genommen anschauen, sondern die Interaktionen zwischen Menschen und Maschinen in den Blick nehmen. Denn wir gehen mit den Maschinen unweigerlich Beziehungen ein, die wir als ›sozial‹ bezeichnen können. Wir Menschen bilden als *zoa politica* nicht nur mit anderen Menschen und anderen Tieren, sondern immer auch mit Maschinen Gemeinschaften, weil es gewissermaßen zu unserer Natur gehört, zu interagieren, zu kommunizieren und soziale Strukturen aufzubauen und uns so, mit Hannah Arendt gesprochen, in ein »Gewebe« von sozialen Bezügen zu begeben.⁴ Und ein solches ›Sozialgewebe‹ können wir auch mit Maschinen weben, nicht nur, aber auch (und vermutlich in besonderer Weise), wenn diese in irgendeiner Form als ›intelligent‹ eingestuft werden.

Sollten wir die Maschinen partout nicht als ›intelligent‹ in einem gehaltvollen Sinne ansehen wollen, bleibt es aber doch dabei, dass sie mitunter so programmiert werden können, dass sie uns in einer mehr oder weniger ›menschenanalogen‹ Verhaltensweise begegnen, auf die wir als *soziale Wesen* reagieren und in komplexer Weise *antworten* können (›Antwort‹ als philosophischer Begriff wird unten näher erläutert). Gegen diese Beobachtung könnte man einwenden, dass wir es hier niemals mit ›echten‹ sozialen Beziehungen zu tun haben und bei derartigen Beschreibungen des Mensch-Maschine-Verhältnisses eher Übertragungen oder sogar (gezielte) Täuschungen vorliegen, die aus solchen ›menschenanalogen‹ Verhaltensweisen von Maschinen noch längst keine ›echte‹ soziale Situation machen.

In diesem Text will ich jedoch von der Hypothese ausgehen, dass einiges dafürspricht, dass wir Menschen unsere sozialen Beziehungen auch auf Unbelebtes ausdehnen können und daher die Grenzen von Mensch-Mensch-, Mensch-Tier-, Mensch-Ding- und Mensch-Maschine-Begegnungen graduell zu verstehen sind und all diese Sphären zusammengekommen ›die‹ Sozialität des menschlichen Lebens ausmachen. Um dies zu verdeutlichen, will ich ein Argument von Mark Coeckelbergh aufgreifen. In einem seiner Texte diskutiert der Technikphilosoph die Interaktionen zwischen Menschen und Robotern im Blick auf den Begriff des Vertrauens und vertritt die These, dass es keineswegs absurd ist, in diesem Kontext den Begriff der sozialen Beziehung zu verwenden.⁵ Coeckelbergh bezeichnet das Vertrauen in diesem Kontext als ›default mode‹⁶, also als eine Art ›Voreinstellung‹ oder ›Standardeinstellung‹, in der wir den Robotern begegnen und dabei immer schon eine soziale

4 Hannah Arendt: *Vita activa oder vom tätigen Leben*. Piper: München 2002, S. 222ff.

5 Mark Coeckelbergh: »Can we trust robots?«, *Ethics Inf Technol* 14 (2012): 53–60.

6 Ebd., S. 56.

Beziehung zu ihnen voraussetzen, wenn wir mit ihnen interagieren. Daher können wir mit gewissen Recht sagen, dass wir Menschen als *zoa politica* so disponiert sind, dass wir selbst mit Maschinen eine wie auch immer geartete soziale Beziehung eingehen – auch wenn wir es vielleicht nicht einmal wissen oder wahrhaben wollen. Diese Art von sozialer Beziehung mag sich von derjenigen mit anderen Menschen unterscheiden, kann aber aufgrund des oben genannten Sozialgewebes nichtsdestotrotz als eine soziale Beziehung verstanden werden.

Doch auch wenn dem so sein sollte, kommen wir nicht umhin, hier philosophisch genauer hinzuschauen, um Beschreibungsformen und Verstehenshorizonte zu entwickeln, die das Eigentümliche dieser Art sozialer Beziehungen erfassen können. Ich will in diesem Text dazu beitragen, indem ich die spezielle Form von ›Alterität‹ untersuche, die die Mensch-Maschine-Interaktionen charakterisieren kann. Dass der Begriff ›des Anderen‹⁷ in der Einschätzung von künstlich-intelligenten Systemen hilfreich sein kann, wurde schon verschiedentlich bemerkt. David J. Gunkel spricht in seinem Buch *The Machine Question* etwa von ›*machinic others*‹ und beschreibt im Rückgriff auf Emmanuel Lévinas verschiedene Formen einer ›*machinic otherness*‹.⁸ Und bereits in früheren philosophischen Ansätzen wurde die Bedeutung der Alterität in der Technikreflexion hervorgehoben. Um nur ein Beispiel zu nennen: In seiner postphänomenologisch ausgerichteten relationalen Ontologie identifiziert der Philosoph Don Ihde neben ›*embodiment relations*‹, ›*hermeneutic relations*‹ und ›*background relations*‹ auch ›*alterity relations*‹.⁹ Mit dem (ebenfalls von Lévinas inspirierten) Alteritätsbegriff lotet er aus, ob und inwiefern wir von ›*technology-as-other*‹ oder auch von einer ›*quasi-otherness*‹ in der Begegnung mit Maschinen sprechen können.¹⁰

Ihdes tentativ-vorsichtiger Begriff der ›*quasi-otherness*‹ zeigt jedoch an, dass wir es hier möglicherweise mit einem tieferen philosophischen Problem zu tun haben. Denn ist es wirklich sinnvoll, von einer *Quasi*-Anderheit zu sprechen? Was soll das ›quasi‹ genau zum Ausdruck bringen? Ist damit eine Situation bezeichnet, in der wir mit einem Roboter oder

7 Die Rede von ›dem Anderen‹ hat sich in der Phänomenologie etabliert, auch wenn die im Singular stehende maskuline Form nicht unproblematisch ist. Im Folgenden werde ich daher diesen Begriff immer wieder in Anführungszeichen setzen oder alternativ durch einen Plural oder die weibliche Form ersetzen. Auch den Begriff ›das Andere‹, der in dieser Auseinandersetzung mit der maschinellen Alterität naturgemäß eine zentrale Rolle zukommt, wird immer wieder Verwendung finden.

8 Siehe David J. Gunkel: *The machine question. Critical perspectives on AI, robotics, and ethics*. Massachusetts 2017, S. 197ff.

9 Don Ihde: *Technology and the Lifeworld. From Garden to Earth*. Indiana 1990.

10 Ebd., S. 97ff.

einer anderen interaktiven Maschine nur so tun ›als ob‹? Sprechen wir also von einer Situation, in der wir nur so tun, als ob die Maschine nicht nur ein Werkzeug, sondern ein ›Anderes‹ wäre (in der wir in der Maschine ein über den einfachen Technologiegebrauch hinausgehendes ›Gegenüber‹ sehen), das sie aber ›eigentlich‹ gar nicht ist? Oder soll mit dem ›quasi‹ ein Moment der Täuschung erfasst werden und die Maschine würde etwas zu sein vorgeben, was sie gar nicht ist? In entgegengesetzter Richtung könnten wir aber auch fragen: wenn es so sein sollte, dass wir die Maschine bereits in ihrer ›Quasi-Andersheit‹ verstehen und uns entsprechend verhalten – könnten wir dann das ›quasi‹ nicht einfach problemlos streichen, weil es sich ohnehin bereits um eine echte Form von Andersheit handelt?

Wenn uns das Streichen des ›quasi‹ in der Quasi-Andersheit allerdings spontan Unbehagen bereiten sollte – vielleicht weil es lebensweltlich und im alltäglichen Verständnis offenkundig zu sein scheint, dass wir Menschen und Maschinen *clare et distincte* zu unterscheiden wissen –, könnten wir in einer ontologisch präzisen Klassifizierung übersehen, dass wir Menschen offenbar dazu neigen, mit Maschinen und anderen Dingen in einer Weise zu interagieren, die es durchaus erlaubt, den Begriff der Alterität zu verwenden. Meines Erachtens ist dies aber nur unter der Voraussetzung möglich, dass wir den Blick nicht mehr auf den agentiellen Status der Maschine richten, sondern uns das menschliche Verhalten gegenüber den Maschinen ansehen. Die Andersheit von Maschinen ist, wie mir scheint, grundlegend falsch verstanden, wenn wir fragen: können Maschinen ›als solche‹ in vergleichbarer Weise wie andere Menschen ›anders‹ sein? Sinnvoller (und philosophisch fruchtbarer) scheint mir folgende Frage zu sein: was macht das spezifische Anderssein von Maschinen mit uns ›alteritätsbedürftenden‹ (*sit venia verbo*) oder sogar ›alteritätsbedürftigen‹ Menschen?

Wir haben es in unseren Interaktionen selbstverständlich mit verschiedenen Arten von Maschinen zu tun. Fragen der Alterität scheinen vor allem bei humanoiden Robotern oder bei sehr komplexen maschinellen Systemen zu finden zu sein, die mit künstlich-intelligenten Algorithmen ausgestattet sind und die sich in einer Weise ›von selbst‹ verhalten, dass wir aufgrund ihrer ›Menschenähnlichkeit‹ in eine soziale Situation geraten. Weniger komplexe Maschinen scheinen Alteritätsfragen eher nicht aufzuwerfen (auch wenn wir in bestimmten Kontexten von einem sozialen Gewebe sprechen können, das wir mit ihnen eingehen). Auf eine mögliche Differenzierung von maschinellen Alteritätsformen will ich in diesem Text nicht eingehen. Mir geht es in diesem Beitrag vielmehr darum, verschiedene Momente einer alteritätstheoretischen Perspektive auf die Mensch-Maschine-Interaktion zu erörtern. An Paul Ricœurs Überlegungen in *Das Selbst als ein Anderer* anschließend, können wir sagen, dass Maschinen zur ›Wer-Frage‹ von Personen (im

Sinne der *ipse*-Identität) beitragen können und damit zum Thema einer *Hermeneutik des Selbst*, wie sie Ricœur vorschwebte,¹¹ gemacht werden können. Ricœur selbst hat die Alterität der Maschinen allerdings vernachlässigt. Das Feld der Vieldeutigkeit, das Ricœur mit der Dialektik zwischen Selbstheit und Andersheit erschlossen hat, kann aber durchaus um die verzwickten Erfahrungen mit maschineller Andersheit erweitert werden, wie in diesem Beitrag gezeigt werden soll.

Nach diesen Vorüberlegungen will ich im Folgenden dem Phänomen maschineller Alterität nachgehen. Dabei werde ich zunächst kurz auf ›alte‹ anthropologische Selbstverständigungsdiskurse verweisen, die wir im Blick auf die Maschinen schon seit der Antike zu verzeichnen haben (2.). Dann werde ich Gehlens Formulierung der Maschine als das ›Andere-als-Menschliche‹ als ein Moment der Mensch-Maschine-Begegnung herausstellen, in dem das Anderssein der Maschine in der philosophischen Anthropologie zum Thema wurde (3.), um dann auf einige aufschlussreiche Alteritätsmomente in der frühen KI-Entwicklung einzugehen (4.). In einem nächsten Schritt will ich mich mit dem von Wilhelm Schapp entwickelten phänomenologischen Ansatz auseinandersetzen, der davon ausgeht, dass wir uns selbst begegnen können, weil wir immer schon in Geschichten mit anderen und anderem verstrickt sind, und dies im Blick auf die Maschinengeschichten, die wir uns erzählen, diskutieren (5.). Daran anschließend werde ich schließlich mit der Responsivität (6.) und dem ›Anspruch des Anderen‹ (7.) zwei zentrale Theoriemomente der Phänomenologie der Alterität (insbesondere bei Bernhard Waldenfels und Lévinas) in Bezug auf die ›Maschinenfrage‹ fruchtbar zu machen suchen. Am Ende steht eine zusammenfassende Schlussbemerkung (8.).

2. Begegnungen mit (auffälligerweise häufig weiblichen) Maschinen

Maschinen wurden in der europäischen Kulturgeschichte als schillernde Gegenüber imaginiert, bei denen gerade die Kippmomente zwischen Ähnlichkeit und Andersheit ihren Reiz ausmachen.¹² Bereits in der griechischen Antike wurde die Maschine als etwas begriffen, was sich

11 Siehe Paul Ricœur: *Das Selbst als ein Anderer*. München 1996, S. 21ff, S. 26ff.

12 Siehe allgemein zu den Selbstverständigungsdiskursen in der abendländischen Kulturgeschichte Oliver Müller und Kevin Liggieri: »Mensch-Maschine-Interaktion seit der Antike: Imaginationenräume, Narrationen und Selbstverständnisdiskurse«. In: Kevin Liggieri, Oliver Müller (Hg.): *Mensch-Maschine-Interaktion. Geschichte-Kultur-Ethik* (Metzler-Handbuch). Stuttgart 2019, S. 3–14.

einerseits deutlich von der menschlichen Existenz unterscheidet, andererseits aber viele Züge menschlichen Verhaltens imitieren kann. Die Maschine wurde daher sowohl als etwas imaginiert, das für höchste technische Perfektion steht – weil sie sich am Menschen als dem ›Maß der Maschine‹ orientiert –, das aber gerade deshalb auch potentielle Täuschung sein kann: Was ein Mensch zu sein scheint, ist eigentlich (nur) eine Maschine. So erzählt der erste Erzähler Homer bereits von roboterähnlichen androiden, genauer gynoiden Statuen, die der antike Techniker Hephaistos hergestellt hat. Diese können sich nicht nur ganz von selbst bewegen, was im Griechischen »*automatoi*« heißt, sondern sind zudem in der Lage, zu sprechen und Arbeiten zu verrichten.¹³ Polybios wiederum berichtet von der künstlichen Frau Agepa, die so täuschend echt aussah, dass die sich ihr nähernden Männer erst zu spät erkannten, dass es sich nicht um eine Frau aus Fleisch und Blut, sondern um ein übles Folterwerkzeug handelte.¹⁴ Es ist bemerkenswert, dass ›das Andere‹, das an der Maschine fasziniert oder irritiert, seit Beginn des Nachdenkens über die Maschine oftmals mit Stereotypen von Weiblichkeit verschränkt zu sein scheint. Das Fremde der Maschine wird durch Feminisierung vertraut gemacht – und bleibt möglicherweise gerade deshalb fremd. Es scheint fast so, als sei das Andere-als-Männliche gleichermaßen von der Maschine und ›der Frau‹ verkörpert – so dass sich eine auffallende Überlagerung von Maschinenfiktionen und Geschlechterrollenvorstellungen bereits in den allerersten literarischen Dokumenten findet.

Die Faszination an Android:innen hat seither nicht abgenommen, wie nicht nur die Science Fiction beweist, denn Automatenvisionen und Maschinenimaginationen durchziehen die Kulturgeschichte in verschiedenen Formen.¹⁵ Eine der berühmtesten Maschinenfrauen ist Olimpia aus E.T.A. Hoffmanns Erzählung *Der Sandmann*. Hoffmann setzt ebenfalls bei der Erfahrung von der Ähnlichkeit und Andersheit in der (männlichen) Interaktion mit Olimpia an und spielt literarisch mit der Vertrautheit und Fremdheit, die in der Mensch-Roboter-Begegnung heute häufig als das *uncanny valley* bezeichnet wird. Auch Olimpia ist Menschen irgendwie ähnlich, irgendwie auf eine unheimliche Weise aber auch nicht. Hoffmann beschreibt seine maschinelle Protagonistin in einer Szene wie folgt: »Olimpia erschien sehr reich und geschmackvoll gekleidet. Man mußte ihr schöngeformtes Gesicht, ihren Wuchs bewundern. Der etwas

13 Homer: *Odysee*. In der Übertragung von Johann Heinrich Voß. Frankfurt am Main 1990, Gesang 18, V. 373–377.

14 Siehe Klaus Völker (Hg.): *Künstliche Menschen. Dichtungen und Dokumente über Golems, Homunculi, Androiden und liebende Statuen*. München 1976, S. 63.

15 Siehe Käte Meyer-Drawe: »Androiden«. In: Carlos Kölbl und Anna Sieben (Hg.): *Stichwörter zur Kulturpsychologie. Jürgen Straub zum 60. Geburtstag*. Gießen 2018, S. 15–20.

seltsam eingebogene Rücken, die wespenartige Dürre des Leibes schien von zu starkem Einschnüren bewirkt zu sein. In Schritt und Stellung hatte sie etwas Abgemessenes und Steifes, das manchem unangenehm auffiel; man schrieb es dem Zwange zu, den ihr die Gesellschaft auferlegte. Das Konzert begann. Olimpia spielte den Flügel mit großer Fertigkeit und trug ebenso eine Bravour-Arie mit heller beinahe schneidender Glasglockenstimme vor. «¹⁶ Hoffman flicht in seine Beschreibung der Olimpia die Irritationen ein, die diese in den Augen der Betrachter:innen auslöst, ohne dass von jenen erkannt wurde, dass es sich bei ihr nicht um ein lebendiges Wesen, sondern um eine Automatenfrau handelt. Auf der einen Seite entspricht ihr Auftritt der gesellschaftlichen Konvention und dem Frauenbild der Zeit, auf der anderen Seite hinterlässt sie aber den Eindruck des Fremdartigen, Unvertrauten, des unbestimmt von der Norm abweichenden. Hoffmann wählt dafür ein Vokabular, das das Mechanische insinuiert (»seltsam eingebogen«, »etwas Abgemessenes und Steifes«, »beinahe schneidende Glasglockenstimme«), auch wenn sie von der Festgesellschaft nicht als etwas Mechanisches, sondern nur als etwas leicht Irritierendes, nicht recht Einzuordnendes bemerkt wurde, das, wie er schreibt, »manchem unangenehm auffiel«.

Bekanntlich wird das Geheimnis der Olimpia am Ende doch noch gelüftet: Ein junger Mann namens Nathanael verliebt sich in die bezaubernde Maschinenfrau, obwohl sie in den intimen Zwiegesprächen lediglich »ach!« sagen kann,¹⁷ was dem romantischen Schwärmer aber nicht weiter auffällt. Doch schließlich muss Nathanael mit Schrecken feststellen: »Olimpias toderbleichtes Wachsgesicht hatte keine Augen, statt ihrer schwarzen Höhlen; sie war eine leblose Puppe.«¹⁸ Die Romantik hat das ›Andere der Vernunft‹ in vielfacher Weise zum Thema gemacht, sei es in Bezug auf das Triebhafte, Unterbewusste, Weibliche oder Wahnsinnige.¹⁹ Hoffmann verbindet die Themen des Wahnsinns und der als bedrohlich erfahrenen Weiblichkeit mit dem Maschinendiskurs der Neuzeit und wendet damit gewissermaßen die cartesianische Automatenbegeisterung in ihr Gegenteil. Die sich als bloße Apparatur entpuppende geliebte Frau wird zur Horrorvorstellung. Die Maschine steht nicht mehr für wissenschaftlich-technischen Triumph und Ingenieurskompetenz, sondern für eine elementare ontologische Verunsicherung.

Für das Verständnis des Maschinendiskurses ist die Figur der Olimpia interessant, weil die Konstruktion von Automaten mit den mechanisti-

16 E.T.A. Hoffmann: *Der Sandmann*. Stuttgart 1991, S. 29.

17 Ebd., S. 31.

18 Ebd., S. 36.

19 Siehe zur Gedankenfigur des ›Anderen der Vernunft‹ Hartmut Böhme und Gernot Böhme: *Das Andere der Vernunft. Zur Entwicklung von Rationalitätsstrukturen am Beispiel Kants*. Frankfurt am Main 1983.

schen Erklärungsmodellen der Natur zusammenhängen. Die Automaten zeigen, dass wir die Natur nachbauen können, gerade weil wir ihren Geheimnissen mechanistisch auf die Schliche gekommen sind. Bekanntlich ist es Julien Offray de LaMettrie, der mit seinem Buch *L'homme machine* den mechanistischen Diskurs der Neuzeit auf die Spitze treibt.²⁰ Während seine Mensch-Maschine-Gleichsetzung einen provokanten, aber letztlich ›klassisch‹ aufklärerischen Erklärungsrahmen bietet, in dem wir Menschen uns unserer mechanistisch vollständig erklärbaren ›Natur‹ selbst versichern sollen, führt Hoffmann mit Olympia das Horride vor, das die mechanistische Reduktion des Lebendigen auf seine Automatenhaftigkeit auslösen kann. Das maschinell Andere wird zum unheimlichen Doppelgänger unserer selbst.

3. Die Maschinen als das ›Andere-als-Menschliche‹

Das Horride ist aber nur ein Extremfall maschineller Alteritätserfahrung. Oftmals wird die Maschine als etwas wahrgenommen, was uns an die ›mechanischen‹ Vorgänge in unserem Organismus erinnert, ohne dass wir unsere menschliche Existenz darauf notwendigerweise reduzieren müssten. Diesbezüglich sind bei Arnold Gehlen einige Überlegungen zu finden, die das Mensch-Maschine-Verhältnis mit dem Begriff des ›Anderen‹ zu erfassen suchen, wiewohl sich sein Begriff des Anderen von den phänomenologischen Alteritätskonzepten unterscheidet (auf die unten näher eingegangen wird). Nichtsdestotrotz kann das von ihm beschriebene Alteritätsphänomen als eine der phänomenologischen Beschreibungen vorgelagerte Erfahrung verstanden werden, da in Gehlens Ansatz ›das Andere‹ gewissermaßen ›dem Anderen‹ vorausgeht.

Gehlen setzt dabei an, dass wir Menschen uns vor allem in dem erkennen können, was anders als wir selbst ist und uns gerade deshalb zur Selbstbefragung herausfordert: »Bedrängt von der Rätselhaftigkeit seines Daseins und seines eigenen Wesens ist der Mensch schlechthin darauf angewiesen, seine Selbstdeutung über ein Nicht-Ich heranzuholen, über ein Anderes-als-Menschliches. Sein Selbstbewußtsein ist indirekt, seine Bemühung um eine Eigenformel verläuft immer so, daß er sich mit einem Nichtmenschlichen gleichsetzt und in dieser Gleichsetzung wieder unterscheidet.«²¹ Dieses ›Anderes-als-Menschliche‹, das sich traditionell in den verschiedenen ›anthropologischen Differenzen‹ zeigen kann, umfasst nicht nur Pflanzen und Tiere (wie sie bekanntlich in den Anthropologien von Max Scheler und Helmuth Plessner ausgearbeitet wurden),

20 Julien Offray de LaMettrie: *Der Mensch eine Maschine*. Stuttgart 2001.

21 Arnold Gehlen: *Die Seele im technischen Zeitalter. Sozialpsychologische Probleme der industriellen Gesellschaft*. Hamburg 1957, S. 16.

sondern eben auch die Maschinen. Die Auseinandersetzung mit der Maschine kann damit Antwort(en) auf die anthropologische Grundfrage ›Was ist der Mensch?‹ bieten.²²

Der Maschine haftet damit eine Ambivalenz an: Sie ist bei Gehlen nicht nur jenes Andere-als-Menschliche, sondern wir erkennen uns in den Maschinen auch selbst wieder. Die Maschinen bleiben ›Kippapparaturen‹, wenn man so will. Als das Andere-als-Menschliche zeigen sie uns gerade auch zutiefst Menschliches – und zwar Menschliches, das wir oft zu übersehen neigen, nämlich z.B. die unbewussten physischen Prozesse. So notiert Gehlen – in einer Zeit, in der die Kybernetik als eine neue Leitwissenschaft in Anschlag gebracht wird –: »Nun ist aber der Mensch in der Tat in ganz zentralen Bereichen seiner Natur Automatismus, er ist Herzschlag und Atmung, er lebt geradezu in und von sinnvoll funktionierenden, rhythmischen Automatismen, wie sie in der Bewegung des Gehens, vor allem aber in den eigentlichen Hantierungen und Arbeitsgängen der Hand vorliegen, in dem ›Handlungskreis‹, der über Sache, Hand und Auge zur Sache zurücklaufend sich erschließt und dauernd wiederholt.«²³ Gehlen betrachtet die Maschine also als etwas gleichzeitig Menschliches und Nicht-Menschliches. Gerade dieses Changieren zwischen Identität und Differenz macht die Maschine zu einem anthropologischen Resonanzphänomen. Die Maschinen bringen also nicht bloß Entfremdungserfahrungen hervor, sondern bilden auch einen Resonanzraum, in dem wir uns selbst begegnen und Weltverhältnisse aufbauen können.²⁴ Die Erfahrung des ›maschinell Anderen‹ steht in einer ganz elementaren ›Resonanzbeziehung‹ zu uns Menschen.²⁵

Diese wenigen Stichworte mögen verdeutlichen, dass das Mensch-Maschine-Verhältnis von jeher kniffliger ist, als wir es üblicherweise denken. Daher ist es kein Wunder, dass dies auch auf die Entwicklungen der Künstlichen Intelligenz abgefärbt hat. Das berühmte *Dartmouth Summer Research Project on Artificial Intelligence* von 1956 ging davon aus »that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.«²⁶ In dem Antrag zu diesem Research Project findet sich

- 22 Dies ist die anthropologische Grundfrage in ihrer klassischen Formulierung. Es gibt aber gute Gründe die Frage anders zu stellen, um problematische Essentialisierungen zu vermeiden, wie etwa: ›Wer sind wir Menschen?‹.
- 23 Arnold Gehlen: *Die Seele im technischen Zeitalter. Sozialpsychologische Probleme der industriellen Gesellschaft*. Hamburg 1957, S. 16.
- 24 Siehe Oliver Müller: *Selbst, Welt und Technik. Eine anthropologische, geistesgeschichtliche und ethische Untersuchung*. Berlin, New York 2014.
- 25 Siehe zum Resonanzbegriff Hartmut Rosa: *Resonanz. Eine Soziologie der Weltbeziehung*. Berlin 2016.
- 26 John McCarthy et al.: »A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence« (August 31, 1955). Zitiert nach *AI Magazine* 27(4) (2006): 12–14, hier S. 12.

die programmatische Hypothese formuliert, dass die Maschine eigentlich nur eine andere Verkörperungsform als der menschliche Organismus darstellt, auf die aber letztlich alle Formen menschlicher Intelligenz übertragen werden können. Das Ziel dieses Projekts ist der Entwurf einer Maschine, die die Formen von Intelligenz auszuüben vermag, die bislang Menschen vorbehalten waren (»now reserved for humans«).²⁷ Abermals wird eine Maschine projiziert, die menschähnlich werden soll – und die idealerweise sogar so menschenähnlich wird, dass sie nicht mehr von einem Menschen unterschieden werden kann.

4. Imitation Games: Alterität in der frühen Entwicklung von Künstlicher Intelligenz

Wenn wir bei LaMettrie die Formel ›der Mensch als eine Maschine‹ finden, so können wir in dem Antragstext der frühen Pioniere der Künstlichen Intelligenz die Parole ›die Maschine als ein Mensch‹ extrahieren. Mechanisierung des Menschen und Vermenschlichung der Maschine gehen also Hand in Hand. Und so ist es vielleicht kein Wunder, dass Alterität bei den allerersten Vorstößen in Richtung künstlicher Intelligenz, die bekanntlich Alan Turing in Angriff genommen hatte, eine zentrale Rolle spielt. Der später so genannte Turing-Test ist inspiriert von einem damals beliebten Party-Spiel, das ›*imitation game*‹ heißt. Die Partygäste mussten erraten, ob es sich bei der Person, die ihnen, ohne sichtbar zu sein, mittels schriftlicher Notate auf Fragen antwortet, um eine Frau oder einen Mann handelt. Turing entwirft in seiner Variante ein Setting, in dem zu entschlüsseln ist, ob die Antworten von einem Menschen oder von einer Maschine gegeben werden.²⁸ Im Blick auf die erwähnten Androidinnen ist es, nebenbei gesagt, recht bemerkenswert, dass im Setting des Turing-Tests die Frau gewissermaßen durch die intelligente Maschine ersetzt wurde. Das rätselhaft Andere ist nunmehr die Maschine und nicht mehr das andere Geschlecht.

In der Entwicklung von Künstlicher Intelligenz, so können wir sagen, ging es von Beginn an um alteritätstheoretische Fragen. Ziel ist es nämlich nicht einfach, schnelle und effiziente Rechner zu bauen. Sondern bereits in den allerersten Anfängen werden Forschungsvorhaben entwickelt, die bei der Frage ansetzen, ob das uns begegnende Andere ein Mensch oder eine Maschine ist. Das liegt natürlich trivialerweise daran, dass die menschliche Intelligenz zum Vorbild für die maschinelle Intelligenz

²⁷ Ebd.

²⁸ Alan Turing: »Computing machinery and intelligence«. *Mind*, 49 (1950): 433–460.

genommen wird. Aber es bleibt trotzdem auffällig, dass sich die Entwicklung der Künstlichen Intelligenz in einer Imaginationstradition bewegt, die wir bei Homer und E.T.A Hoffmann bereits kennengelernt hatten. Der Reiz der künstlich-intelligenten Maschine liegt darin, täuschend echt menschlich zu sein. Und die Konstruktion einer solchen Maschine kann nur gelingen, wenn diese zu einem ›Gegenüber‹ wird, das mehr ist als ein bloßes technisches Werkzeug, das vielmehr eine Erfahrung von ›Andersheit‹ auslösen soll, das auf einen Menschen verweist. Dass auch hier die Maschinen weibliche Züge tragen können, zeigt das von Joseph Weizenbaum geschriebene Programm, mit dem er 1966 den Turing-Test weiterentwickelte, welches in seinem Setting eine Psychotherapeutin simulieren sollte. Dieses Programm hieß ELIZA,²⁹ nach Eliza Doolittle aus dem (den antiken Mythos aktualisierenden) Theaterstück *Pygmalion* von George Bernhard Shaw – ein gewissermaßen klassisches Produkt männlichen Verfügungswillens in der Anwendung von Wissenschaft.

So technikkritisch Weizenbaum dies letztlich auch gedacht haben mag, es bleibt bemerkenswert, dass auch hier wieder eine Maschinenfrau erdacht wurde, die die männliche Intelligenz zu täuschen versteht. Die Unruhe über die wahre Identität des Gegenübers, die der Turing-Test verkörpert, wurde auch von Philip K. Dick aufgegriffen, der in seinem Roman *Do Androids Dream of Electric Sheep?* (der bekanntlich die Vorlage für die *Blade Runner*-Filme von Ridley Scott darstellt) einer Maschine einen Auftritt gewährt, die zur Orientierung in einer Welt elementarer ontologischer Verunsicherung eingesetzt werden soll.³⁰ Denn der Roman spielt in einer Welt, in der zwischen dem Natürlichen und Künstlichen nicht mehr unterschieden werden kann. Die Voight-Kampff-Maschine ist eine Empathie-Test-Apparatur, mit der eruiert werden soll, ob es sich beim Gegenüber um einen Menschen oder eine/n Replikant:in handelt. Die Voight-Kampff-Maschine misst im Narrativ des Romas das Menschlichste an uns Menschen, die Empathiefähigkeit. Allerdings bedarf es in einer Art *contradictio in adjecto* eben gerade einer Maschine, die dieses Menschlichste exakt und verlässlich zu messen in der Lage ist. Auch in dem Film *Ex machina* wird die Idee des Turing-Tests aufgegriffen, und zwar in den Interaktionen zweier Programmierer mit der Androidin Ava, die einmal mehr die Erfahrung des Vertraut-Unvertrauten, das die intelligente Maschine verkörpert, feminisiert – und gerade dadurch zum unkontrollierbar Anderen macht. Denn der Film endet damit, dass sich Ava der männlichen Verfügungsmacht zu entziehen weiß.

29 Joseph Weizenbaum: »ELIZA-A Computer Program For the Study of Natural Language Communication Between Man and Machine«. *Communications of the ACM* 9 (1966): 36–35.

30 Philip K. Dick: *Do Androids Dream of Electric Sheep?* Weidenfeld & Nicolson, London 2012.

5. In Maschinengeschichten verstrickt

Der Verweis auf die Erzählungen von E.T.A. Hoffmann und Philip K. Dick erfolgte nicht ohne Grund. Denn wir erfahren ›das Andere‹ oftmals gerade über Narrative. Nicht nur in Ricœurs Philosophie der Alterität spielt die ›narrative Identität‹ eine zentrale Rolle, weil sie die Brüchigkeit, die Widersprüche und Dissonanzen einer Person als ›Lebensgeschichte in Lebensgeschichten‹ verstehbar machen kann, sondern auch in Wilhelm Schapps Buch mit dem schönen Titel *In Geschichten verstrickt*.³¹ In den Geschichten, mit denen wir unsere Identität erzählerisch formen, sind wir immer schon mit Tieren, Dingen und anderen Menschen verstrickt, einer lebensweltlichen Verstrickung, der wir gar nicht entkommen können: »Die Geschichte jedenfalls läßt sich nicht als Gegenstand untersuchen, weil etwas Geschichte nur insoweit ist, als ich in die Geschichte verstrickt bin. Dies Verstricktsein läßt sich nicht so aus der Geschichte lösen, daß auf der einen Seite die Geschichte übrigbliebe und auf der anderen Seite mein Verstricktsein oder so, daß die Geschichte überhaupt noch etwas wäre ohne den Verstrickten, oder der Verstrickte noch irgendetwas wäre ohne die Geschichte.«³² Das heißt, dass wir unser Selbstsein in besonderer Weise beschreiben und verstehen können, wenn wir uns die Geschichten, in die wir unentrinnbar verwickelt sind, näher anschauen und zu deuten beginnen.

Naturgemäß sind zunächst die Mitmenschen privilegiert, was die narrativen Verstrickungen angeht. Aber der ganze erste Teil von Schapps Buch widmet sich dem »Wozuding in der Außenwelt und seine[r] Wahrnehmung«³³ und ist eine Fortführung von Heideggers Analysen des ›Vorhandenen‹ in *Sein und Zeit*.³⁴ Auch bei Schapp geht es meistens um einfache handwerkliche Techniken und Werkzeuge, aber immerhin kommt das Automobil vor, das am ehesten zu den Maschinen zählen kann, die uns hier interessieren.³⁵ Da Schapp andere Fragen adressiert, hat er den gemeinsamen Geschichten von Menschen und Maschinen leider kaum Aufmerksamkeit geschenkt. Aber man darf aus seinem Ansatz durchaus folgern, dass Menschen und Maschinen ebenso in Geschichten verstrickt sein können wie Menschen und Menschen oder Menschen und Tiere. Auch Plessner hat die unauflösbare Verstrickung von Menschen und Maschinen unterstrichen. Bei ihm spielt die narrative Verstrickung

31 Wilhelm Schapp: *In Geschichten verstrickt. Zum Sein von Mensch und Ding*. Frankfurt am Main 1985.

32 Ebd., S. 85f.

33 Ebd., S. 11–82.

34 Martin Heidegger: *Sein und Zeit*. Niemeyer: Tübingen 1993, S. 66ff.

35 Wilhelm Schapp: *In Geschichten verstrickt. Zum Sein von Mensch und Ding*. Frankfurt am Main 1985, S. 56ff.

keine zentrale Rolle, trotzdem bringt er eine elementare Dynamik pointiert zum Ausdruck, die die Geschichten von Menschen und Maschinen generiert: »Von den Maschinen fortlaufen und auf den Acker zurückkehren, ist unmöglich. Sie geben uns nicht frei und wir geben sie nicht frei. Mit rätselhafter Gewalt sind sie in uns, wir in ihnen.«³⁶

Es ist unsere narrative Identität, die uns mit unseren Maschinen verbindet, mit denen wir gemeinsame Geschichten erleben. Und in diesen Geschichten begegnen wir Maschinen, denen wir vertrauen; Maschinen, die widerborstig sind; Maschinen, die wir verfluchen; Maschinen, die wir bewundern; Maschinen, die uns träumen lassen; Maschinen, die uns in tragische Situationen bringen können.³⁷ Wenn das Selbst nach Ricœurs Formel ›ein Anderer‹ ist, dann können wir uns nun erlauben, seine Formel zu erweitern und konstatieren: das Selbst ist (auch) ›ein Anderes‹. Denn wir betrachten uns nicht nur im Spiegel unserer Maschinen, wie die Phänomenologin Käte Meyer-Drawe umfänglich gezeigt hat,³⁸ sondern wir ›sind‹ im emphatischen Sinne ›wir selbst‹ durch die Maschinen, mit denen wir in gemeinsame Geschichten verwickelt sind und über die wir unsere (*ipse*-)Identität ausbilden. In *Das Selbst als ein Anderer* erwähnt Ricœur übrigens technologische Fiktionen ganz explizit.³⁹ Er betrachtet technische Träume und Fiktionen allerdings sehr sorgenvoll in einem recht engen Blickwinkel, weil er sich in diesem Kontext mit den *puzzling cases* der analytischen Philosophie auseinandersetzt, in denen es vor allem um Gehirntransplantationen bei Fragen personaler Identität geht. Durch diese Verengung der Perspektive in Bezug auf die technischen Fiktionen verspielt Ricœur einiges narratives Potential, das wir zur Erschließung von maschinellen Alteritätserfahrungen fruchtbar machen können.

Denn es lohnt sich, die Maschinengeschichten alteritätstheoretisch ernst zu nehmen. Um dies erneut mit einigen Verweisen auf Filme zu verdeutlichen: Wir können beispielsweise die so innige wie dystopische Beziehung der sehr schlauen und durchtriebenen und auf ihre Weise bemitleidenswerten künstlichen Intelligenz HAL 9000 mit dem Astronauten Dave Bowman verfolgen, die Stanley Kubrick in *Odyssee im Weltraum* erzählt. Oder wir können in der Miniserie *Maniac* einen intelligenten

36 Helmuth Plessner: *Die Utopie in der Maschine*. In: Ders.: *Gesammelte Schriften*. Bd. 10, hg. v. Günter Dux, Odo Marquard und Elisabeth Ströker. Suhrkamp: Frankfurt am Main 1985, S. 34.

37 Vielleicht spielen sich derzeit die wirklich tragischen Tragödien in unserer modernen technischen Zivilisation in erster Linie zwischen Menschen und Maschinen ab, etwa, wenn wir lernen, die Bombe oder andere Technologien zu lieben...

38 Siehe Käte Meyer-Drawe: *Menschen im Spiegel ihrer Maschinen*. München 1996.

39 Paul Ricœur: *Das Selbst als ein Anderer*. München 1996, S. 184.

Supercomputer kennenlernen, der beginnt, Emotionen zu entwickeln und Beziehungen zu den Menschen in seinem Umfeld aufzubauen. Im Film *Her* begleiten wir wiederum einen melancholischen Mann, der sich in die weibliche Stimme seines Computers verliebt, der im Kontakt zu den einfühlsam programmierten Algorithmen die Brüchigkeit seiner Identität erfährt und schließlich wieder zu sich selbst (zurück) findet. Die Tech-Industrie hat die identitätsstiftende Bedeutung der klugen Maschinen längst erkannt. In einer aktuellen Werbung, in der kurze Szenen verschiedener ›einzigartiger‹ Lebensmodelle gezeigt werden, wird ein neues Smartphone folgendermaßen angepriesen: »We all are unique. And now there's a phone that adapts to you. A phone like you.«⁴⁰ Das Versprechen, dass die Maschinen sich so anpassen, dass sie ›wie wir‹ werden, verdeckt die Kehrseite, dass wir uns selbst wiederum an die Maschinen anpassen müssen. Die Maschinen können nur wie wir werden, wenn wir werden wie unsere Maschinen.

Das Interessante an derartigen Filmen ist nicht (bzw. nicht nur), dass sie uns anregen können, über nähere und fernere Technikzukünfte nachzudenken, sondern dass sie Geschichten darüber erzählen, wie tief und unauflöslich wir mit unseren Maschinen bereits in gemeinsame Geschichten verstrickt sind – und wie wir uns in den Begegnungen mit den Maschinen als ›Selbste‹ erfahren können. Einer der Gründe hierfür ist darin zu sehen, dass wir auf Maschinen nicht nur reagieren, sondern auf sie auch antworten können.

6. Maschinen antworten

Das Antwortenkönnen, die Responsivität, ist ein zentrales Moment in Alteritätstheorien, das insbesondere Bernhard Waldenfels im Anschluss an Kurt Goldstein umfänglich untersucht hat.⁴¹ Die Bedingung der Möglichkeit des Selbstseins wird von ihm in der Verschränkung von »Pathos« und »Response« erfasst, also grundlegend alteritätstheoretisch gedacht.⁴² Wir Selbst erfahren, erleiden, erleben etwas durch andere und anderes und antworten dann auf dieses Erleiden. Das Antwortenkönnen ist höchst vielschichtig und umfasst mindestens leibliche, sprachliche, soziale, kulturelle und politische Dimensionen. Anthropologisch

40 Siehe <https://www.youtube.com/watch?v=VxERxFXPw3M> (letzter Zugriff 20.1.22).

41 Siehe vor allem Bernhard Waldenfels: *Antwortregister*. Frankfurt am Main 1994.

42 Siehe etwa Bernhard Waldenfels: *Erfahrung, die zur Sprache drängt. Studien zur Psychoanalyse und Psychotherapie aus phänomenologischer Sicht*. Berlin 2019, S. 254ff.

hat der ebenfalls an Goldstein anschließende Ernst Cassirer diesen Zusammenhang in seinem *Essay on Man* mit folgender Formel erfasst: *From animal Reactions to Human Responses*.⁴³ Das Antwortenkönnen ist damit eine menschliche Fähigkeit, die die Reiz-Reaktions-Mechanismen durchbricht. Das Antworten erfolgt verzögert und ändert das antwortende Selbst und ›den anderen‹ oder ›das andere‹ (welches bei Waldenfels auch als das Fremde firmiert) im Prozess des Antwortens. Dies ist als eine Dynamik zu verstehen, in der wir uns im sinngebenden Antworten auf ein unvorhergesehenes Angesprochenwerden sowohl uns selbst als auch ›dem anderen‹ in einer Weise begegnen, dass wir weder das Eigene zum Maßstab für die Erfahrung und Beurteilung des anderen machen, noch dass wir uns den oder das andere unzugänglich machen, indem wir es als etwas gänzlich Fremdes aus unserem Erfahrungshorizont ausschließen. Über das Zusammenspiel von Pathos und Response wird der andere in vielschichtiger Weise als ›der andere‹ erfahren – mit der Konsequenz, dass wir im Antworten auch unser Selbstsein immer wieder neu entdecken.

Vor diesem Hintergrund stellt sich nun die Frage, ob und wie wir auch auf Maschinen und insbesondere künstlich-intelligente Maschinen in diesem gehaltvollen Sinne antworten können. Nach der oben genannten Fassung des Antwortenkönnens in der Verschränkung von Pathos und Response können wir durchaus nicht nur Menschen, sondern auch Maschinen antworten. Nehmen wir zum Beispiel die von klugen Algorithmen empfohlenen Bücher oder Filme: In den seltensten Fällen reagieren wir einfach reflexhaft und erwerben die empfohlenen Produkte sofort – wie es sich vielleicht mancher Digitalkonzern erhofft –, sondern wir antworten auf diese Empfehlungen, indem wir uns beispielsweise über diese Produktempfehlungszumutungen ärgern oder indem wir die Vorschläge doch, ein wenig neugierig geworden, mit unseren Interessen abgleichen. Das heißt, wir antworten auf diese Algorithmen, weil wir diese vor dem Hintergrund unserer Sinnhorizonte wahrnehmen und bewerten – was nicht ausschließen muss, dass wir auf diese Empfehlungsübergriffigkeit auch recht barsch antworten können, indem wir die Vorschläge der Streaming-Dienste in den Wind schlagen.

In ähnlicher Weise begegnen wir auch Robotern verschiedener Couleur. Einen Pflege- oder Service-Roboter, mit dem wir interagieren, nehmen wir vor dem Hintergrund unserer sozialen Erfahrung und unseren sozialen Erwartungen wahr. Wir reagieren auch hier nicht nur, sondern wir antworten ebenfalls im oben genannten Sinne, wenn wir beispielsweise die derzeitige Unbeholfenheit des Roboters belächeln, wenn wir die

43 Ernst Cassirer: *An Essay on Man. An Introduction to a Philosophy of Human Culture*. In: Ders.: *Gesammelte Werke*. Bd. 23, hg. v. Birgit Recki, Meiner: Hamburg 2006, S. 32ff.

Interaktion als ›Begegnungsspiel‹ mitspielen oder wenn wir in Gedanken vielleicht bereits eine Geschichte skizzieren, die wir den Freund:innen über diese seltsame Begegnung erzählen können. Auch wenn der Roboter vielleicht nicht zu einem ›Freund‹ wird, der uns zum Bücher- und Juwelendiebstahl animiert wie in dem Film *Robot & Frank*, kann die pathische Erfahrung durch die Maschine zur Ausbildung eines vielschichtigen Antwortgeschehens führen.

Vermutlich wird es dabei bleiben, dass das Antwortgeschehen in der Interaktion mit einer Maschine ein anderes ist als das mit einem anderen Menschen. Eine sehr komplexe künstlich-intelligente Maschine kann menschenähnliches Antworten aber vielleicht in einer Weise simulieren, dass wir mitunter gar nicht sagen können, wie sich die maschinelle Antwort von der möglichen menschlichen Antwort unterscheidet. Nichtsdestotrotz dürften wir angesichts der Brüchigkeit, Vieldeutigkeit und Unberechenbarkeit menschlichen Antwortkönnens wohl immer nur zeitweise vergessen, dass es sich bei unserem Gegenüber um eine künstlich-intelligente Maschine handelt, die sich so verhält, weil Programmierer:innen wollten, dass sie sich menschenähnlich verhält. Doch dies ändert nichts daran, dass wir diese Maschinen nicht nur als ein nützliches Werkzeug betrachten und entsprechend verwenden, sondern dass wir ein alteritätstheoretisches Surplus vorliegen haben, das sich phänomenologisch und anthropologisch aus der Struktur des menschlichen Antwortkönnens in der Verschränkung von Pathos und Response erklären lässt.

Ein interessanter Grenzfall in diesem Kontext ist das Phänomen des Vertrauens. Vermutlich stimmt die Beschreibung, dass wir aufgrund der sozialen Situationen, die in Mensch-Maschine-Interaktionen entstehen, auch Formen von Vertrauen in die Maschine ausbilden, die über die bloße Verlässlichkeit, die wir von technischen Geräten erhoffen und erwarten können, hinausgehen.⁴⁴ Bemerkenswert ist hier, dass die vertrauensvolle Sozialität der Mensch-Maschine-Begegnung zu den Phänomenen von *overtrust* oder *automation bias* führen kann, wie man dies in der Debatte nennt.⁴⁵ Das heißt, pointiert gesagt, dass wir der Maschine in ihrer vermeintlichen Perfektion in einer Weise vertrauen, die eigentlich nicht angebracht ist. *Overtrust* und *automation bias* aktualisieren in gewisser Weise die ›Antiquiertheit‹ des Menschen angesichts der Perfektion der Maschinen, wie sie Günther Anders beschrieben hat.⁴⁶ Wir unter-

44 Siehe hierzu noch einmal Mark Coeckelbergh: »Can we trust robots?« In: *Ethics Inf Technol* (2012) 14:53–60.

45 Siehe Alan R. Wagner, Jason Borenstein, Ayanna Howard: »Overtrust in the Robotic Age«, *Communications of the ACM*, 61(9) (2018): 22–24; Alex Leveringhaus: *Ethics and Autonomous Weapons*. Palgrave-Macmillan: London 2016, 99.

46 Günther Anders: *Die Antiquiertheit des Menschen. Bd. 1: Über die Seele im Zeitalter der zweiten industriellen Revolution*. Beck: München 1956.

werfen uns mitunter recht bereitwillig unseren tadellos funktionierenden Maschinen – mit dem Effekt, dass wir dann auf die Maschinen gar nicht mehr antworten, sondern nur noch reagieren, uns gewissermaßen selbst mechanisieren, weil wir die funktionelle ›Norm‹ der Maschine internalisiert haben. Eine Technikkritik könnte bei dem menschlichen Antwortkönnen ansetzen, um einen Ausweg aus der ›Selbstversklavung‹ durch einen bloß reflexhaft reagierenden Umgang mit den Maschinen zu finden (um einen beliebten Topos zu bemühen).⁴⁷

7. Der Anspruch der Maschine

In der Phänomenologie der Alterität ist die Frage zentral, welcher (ethische) Anspruch von ›dem anderen‹ ausgeht. Dies kann sich als eine Phänomenologie des Fremden ausbuchstabieren, wie bei Waldenfels, oder in einer radikalen Subjektkritik formuliert werden, wie wir es bekanntlich bei Lévinas vorliegen haben.⁴⁸ Lévinas geht von einem Primat des Anderen aus, der aller Subjektkonstitution vorausgeht und diese sogar grundsätzlich in Frage stellt. Im Kern seiner Philosophie steht, kurz gesagt, die Gedankenfigur, dass die »*Bewegung des Selben zum Anderen*« in einer Weise erfolgt, in der sie »*niemals zum Selben zurückkehrt*«. ⁴⁹ Im Blick auf die Diskussion einer maschinellen Alterität dürfte es philosophisch keine kleine Herausforderung sein, den Anspruch des Anderen, wie er von Lévinas formuliert wird, auch in der Maschine als ›das Andere‹ ausmachen zu können (dies könnte vielleicht der Grund sein, warum Ihde die vorsichtige Wendung der *quasi-otherness* wählt).

Es stellt sich also die Frage, in welcher Weise wir von dem Anspruch der Maschine als ›das andere‹ sprechen können. Lévinas schreibt in *Die Spur des Anderen*, dass die »heteronome Erfahrung« eine »Bewegung« ermöglicht, »die nicht zu ihrem Ausgangspunkt zurückkehrt«. ⁵⁰ Das heißt, dass ›der Andere‹ uns in dem Sinne ›bestimmt‹, dass er ein unvorhersehbares Antwortgeschehen in Gang setzt. Diese Art heteronomer Erfahrung könnte sich auch auf die Interaktion mit Maschinen übertragen lassen: Dann würden wir uns in der Interaktion einer Maschine in einer Weise selbst erfahren, die weder auf der bloßen Festlegung dessen, was die Maschine ist und leisten soll, noch auf unser sich an

47 Siehe etwa Georg Simmel: *Philosophie des Geldes*. In: *Gesamtausgabe*. Bd. 6, hg. von David P. Frisby, Klaus Christian Köhnke, Frankfurt am Main 1989, 673.

48 Siehe Emmanuel Lévinas: *Die Spur des Anderen. Untersuchungen zur Phänomenologie und Sozialphilosophie*. Freiburg 1983.

49 Ebd., 215.

50 Ebd., 215.

die Funktionsweise der Maschine anpassendes Verhalten zu reduzieren ist. Durch den eigentümlichen *default mode* des sozialen Gewebes und unser Antwortenkönnen nehmen wir die Maschine (vielleicht nicht nur, aber auch) als ›etwas anderes‹ wahr, was nicht mit der bloßen Aneignung durch unsere üblichen epistemischen Kategorien zu erklären ist.

Bemerkenswert bleibt dabei, dass die Maschine etwas von Menschen Hergestelltes ist, ihre Existenz also nur deshalb hat, weil wir die entsprechenden Kategorien zur Verfügung haben, mit der die Maschine konstruiert, gebaut und genutzt wird. Das heißt, wir interagieren mit einer Entität, die erst durch unser Verfügungswissen existiert und die wir doch in der Begegnung als etwas Unverfügbares erfahren – weil das Antwortgeschehen in der Pathos-und-Response-Verschränkung nicht völlig vorhersehbar verläuft und uns in unberechenbarer Weise verändern kann. Dies kann grundsätzlich auf alle Technologien zutreffen, aber in verstärktem Maße können wir dies sicherlich im Blick auf künstlich-intelligente Systeme beobachten, weil wir es hier auch mit einem sich entziehenden *blackboxing* zu tun haben (auf dieses epistemologisch interessante ›Problem‹ in der Mensch-Maschine-Interaktion soll an dieser Stelle nur pauschal verwiesen werden).⁵¹

Auch wenn die Begegnung mit anderen Menschen eine andere Qualität hat, ist es erlaubt zu sagen, dass sich mit der Maschine die oben genannten heteronomen Erfahrungen verbinden können. Denn möglicherweise stellt uns die Maschine in ähnlicher Weise in Frage, wie es Lévinas für die Begegnung mit dem Anderen beschreibt: »Die Beziehung zum Anderen stellt mich in Frage, sie leert mich von mir selbst; sie leert mich unaufhörlich, indem sie mir so unaufhörlich neue Quellen entdeckt.«⁵² Eine derartige radikal gedachte ›Selbstentleerung‹ mag in Bezug auf die Begegnung mit einer Maschine etwas hyperbolisch wirken, aber bestimmte Formen des Sich-selbst-in-Frage-stellens kann man auch in der Interaktion mit Maschinen erfahren. Das kann von Zweifeln angesichts der eigenen Kompetenzen im Umgang mit Maschinen bis hin zur Selbstbefragung bei komplizierten Entscheidungssituationen reichen, die wir in Teilen an die Maschinen delegieren müssen, etwa im klinischen Kontext oder im Militärsektor.

51 Philipp von Hilgers: »Ursprünge der Black Box«. In: Anna Ofak und Philipp von Hilgers (Hg.): *Rekursionen. Von Faltungen des Wissens*. München 2010, S. 135–153; Andreas Kaminski (2018): »Der Erfolg der Modellierung und das Ende der Modelle. Epistemische Opazität in der Computersimulation«. In: Andreas Brenneis et al. (Hg.) (2018): *Technik – Macht – Raum. Das Topologische Manifest im Kontext interdisziplinärer Studien*. Wiesbaden: Springer VS, 317–333.

52 Lévinas: *Die Spur des Anderen. Untersuchungen zur Phänomenologie und Sozialphilosophie*. Freiburg 1983, S. 221.

Im Blick auf den Begriff des »Antlitzes«, der in Lévinas' Theorie der Alterität eine zentrale Rolle einnimmt, dürften sich aber gewisse Zweifel einstellen, ob und inwiefern dieser Begriff auf die Begegnung mit einer Maschine zu übertragen ist. Lévinas beschreibt das Antlitz ›des anderen‹ und den damit verbundenen Aufforderungscharakter in eindringlichen Worten: das Antlitz sei durch seine »Nacktheit« und »Not« charakterisiert und in dieser seiner Nacktheit ist es ein »inständiges Flehen«.⁵³ Gleichzeitig »fordert« dieses Flehen aber auch: »So bedeutet die Anwesenheit des Antlitzes eine nicht abzulehnende Anordnung, ein Gebot, das die Verfügungsgewalt des Bewußtseins einschränkt. Das Bewußtsein wird durch das Antlitz in Frage gestellt [...]. Seine Gegenwart ist eine Aufforderung zu Antwort.«⁵⁴ Angesichts des historischen Hintergrunds, vor dem Lévinas seine Philosophie entwickelt hat, mag die Übertragung des Antlitz-Begriffes auf die Begegnung mit der Maschine regelrecht frivol erscheinen. Denn auch wenn wir die Aufforderung zur Antwort im oben genannten Sinne auch in der Interaktion mit Maschinen wiederfinden können, dürften sich Nacktheit, Not und Flehen (oder andere Begriffe in Lévinas' Vokabular) kaum sinnvoll mit der Maschine verbinden lassen.

Doch trotz dieser Vorbehalte kann man Lévinas' Alteritätstheorie für das Verständnis von Mensch-Maschine-Begegnungen fruchtbar machen, wie mir scheint. Gerade durch Lévinas' ungewöhnliche Terminologie können wir uns dem Nützlichkeitsdenken und der technomorphen Einpassung in die mechanisierte Welt (etwa als das unsere Denk- und Handlungsformen prägenden »Ge-stell«)⁵⁵ entziehen.⁵⁶ Lévinas kann uns daran erinnern, dass wir den Maschinen jenseits des funktionalen Zusammenhangs, den wir mit ihnen bilden, begegnen (können) und dass wir möglicherweise neue Vokabulare brauchen, um zu erfassen, wie die Maschinen unser Selbstverhältnis ändern. Lévinas' Pointe war ja, dass wir in der Bewegung zum anderen hin nicht wieder zu dem ›selben Selbst‹ zurückkehren. Und so können wir auch im Antworten auf die Maschine selbst andere werden.

Lévinas' Terminologie aufgreifend und weiterentwickelnd, könnte man die Maschine als das ›Antlitzlose‹ beschreiben. Mit der Maschine tritt uns etwas entgegen, das uns gerade in seiner Antlitzlosigkeit herausfordert. Die Rede von Antlitz und Antlitzlosigkeit bietet sich vor allem in Bezug auf die Interaktion mit künstlich-intelligenten Robotersystemen

53 Ebd., 222.

54 Ebd., 223f.

55 Martin Heidegger: *Die Frage nach der Technik*. In: Ders.: *Vorträge und Aufsätze. Gesamtausgabe*. Bd. 7, hg. v. Friedrich-Wilhelm v. Herrmann, Frankfurt a.M. 2000, S. 5–36.

56 Siehe Christoph Hubig: *Die Kunst des Möglichen*. Bd. 1: *Grundlinien einer dialektischen Philosophie der Technik*. Transcript: Bielefeld 2008.

an, die in ihren humanoiden Varianten ein an Menschen orientiertes ›Gesicht‹ bekommen, das aber in den eingeschränkten mimischen Funktionen antlitzlos bleibt bzw. zwischen Antlitzanmutung und Antlitzenttäuschung wechselt. ›Antlitzanmutung‹ will ich hier behelfsmäßig, in Ermangelung eines besseren Begriffs für diese Phänomenlage, wählen, um anzuzeigen, dass sich diese Antlitzerfahrung grundlegend von Lévinas' Antlitz-Konzept unterscheidet, wiewohl wir hier eben eine Erfahrung maschineller Alterität vorliegen haben. Die Begegnung mit einem humanoiden Roboter, der Affekte zeigen kann, wird daher eine besondere Herausforderung bleiben. Denn wir dürften einerseits dazu tendieren, ›menschlich‹ zu reagieren, indem wir auf Mimik und Gestik des Roboters antworten. Andererseits dürften uns die Antlitzlosigkeit der Maschine auch irritieren, vielleicht sogar zurückschrecken lassen. Da die Maschine in der oben gewählten paradoxen Formulierung als das verfügt Unverfügte bzw. das unverfügt Verfügte beschrieben werden kann, könnte sich mit der Maschine auch eine besondere Art von Zerbrechlichkeit verbinden, die auf die menschlichen Nutzer zurückspiegelt. Hannah Arendt hatte das technische Herstellen in *Vita activa* bekanntlich mit einer besonderen Macht verknüpft, die darin besteht, dass wir das selbst hergestellte jederzeit auch selbst wieder zerstören können.⁵⁷ Und die Interaktion mit einem Gegenüber, das wir im Prinzip jederzeit zerstören könnten, macht die Maschine zu etwas Fragilem – das uns als Interagierende in Frage stellen kann.

Im Fall von Sexrobotern dürfte dies in besonderem Maße augenfällig sein.⁵⁸ Sexroboter sind gezielt nach unseren Vorstellungen gebaut und insofern haben wir maximale Macht über sie. Gleichzeitig antworten wir diesen obskuren Objekten unserer Begierde auch in einer nicht vorhersagbaren Weise durch die oben genannte Antlitzanmutung. Damit entziehen sie sich bis zu einem gewissen Grad unserer Verfügbarkeit, auch wenn sie zur Bedürfnisbefriedigung produziert wurden. In unserem Begehren geben wir dem Roboter eine stark sexualisierte Antlitzanmutung (wie stereotyp diese auch immer sein mag) und haben gleichzeitig die Macht, diesen jederzeit zur Antlitzlosigkeit zu degradieren. Der erotische Imaginationsraum verbindet sich mit der Macht, den Sexroboter zu ersetzen oder zu zerstören, wenn dieser unsere Bedürfnisse nicht (mehr) befriedigt. Das ›Ebenbild‹, das wir uns in dem Sexroboter schaffen, fußt auf einem alteritätstheoretischen Surplus, das wir bereits im Fall der Androidinnen kennengelernt haben: wir begegnen hier ebenfalls häufig feminisierten

57 Hannah Arendt: *Vita activa oder Vom tätigen Leben*. Piper: München 1981, S. 165ff.

58 Siehe zum Überblick über die Sexroboter-Debatte Catrin Misselhorn: *Künstliche Intelligenz und Empathie. Vom Leben mit Emotionserkennung, Sexrobotern und Co.* Stuttgart 2021, S. 110–134.

Maschinen, die einen erotischen Imaginationsraum für die Erfahrung ›der Anderen‹ eröffnet. Da dies aber mit der Macht verbunden ist, ›die Andere‹ selbst herstellen (und damit prinzipiell jederzeit zerstören) zu können, kann die Antlitzanmutung verschwinden, und die im Sexroboter imaginierte ›sie‹ wird auf ihre Antlitzlosigkeit reduziert – was wiederum die Nutzer von Sexrobotern in gewisser Weise selbst antlitzlos macht.

In dem Film *Ich bin dein Mensch* wird dies in der Begegnung zwischen einer Frau und einem (männlich gestalteten) Roboter durchgespielt: Die Wahrnehmung des Roboters schwankt im Laufe der Geschichte zwischen menschlichem Partner und verfügbarem Objekt und zeigt, wie die Hauptfigur sich in der Interaktion mit dem Roboter immer wieder selbst in Zweifel bringt. So könnte der ›Anspruch der Maschine‹ paradoxerweise in dem bestehen, dass sie uns in ihrer Antlitzlosigkeit daran erinnert, dass wir Menschen der Begegnung mit dem Antlitz anderer bedürfen, um wir selbst zu werden. Insofern spiegelt uns die Antlitzlosigkeit der Maschine unsere eigene Vulnerabilität wider, die wir in der Interaktion mit dieser Maschine erfahren können.

8. Schlussbemerkung

In meinem Beitrag bin ich dem Phänomen maschineller Alterität aus verschiedenen Perspektiven nachgegangen. Nach ersten Überlegungen zum ›sozialen Gewebe‹, das wir mit (nicht nur, aber auch und insbesondere künstlich-intelligenten) Maschinen bilden können, habe ich im Verweis auf die Kulturgeschichte einige Hinweise gegeben, dass die Maschine schon früh als ›Andere‹ verstanden wurden, was sich in Imaginationen über Android:innen niederschlug, wobei auffallend häufig eine Feminisierung der Maschine zu beobachten ist. Die Erfahrung der Andersheit der Maschine überlagert sich in bemerkenswerter Weise mit Geschlechtervorstellungen und entsprechen Stereotypisierungen. Im Blick auf die Anthropologie Gehlens und die frühe Programmatik der KI-Forscher habe ich herausgearbeitet, wie sich hier alteritätstheoretische Momente unterschiedlicher Art finden. Schließlich habe ich drei zentrale Konzepte der Phänomenologie für meine Überlegungen zur maschinellen Alterität zugrunde gelegt: Das ›In-Geschichten-verstrickt sein‹ als Erfahrung ›des Anderen‹; die Responsivität (und die Verschränkung von Pathos und Response); sowie den ›Anspruch des Anderen‹, wie er sich in der Philosophie von Lévinas findet. Ich hoffe, dass ich Folgendes zeigen konnte: Auch wenn die Erfahrung von menschlich Anderen anders gelagert ist, kann eine Phänomenologie der Alterität dazu beitragen, dass wir unsere Begegnungen mit den Maschinen genauer beschreiben und besser verstehen können. Für wertvolle Hinweise zu meinem Beitrag danke ich Katharina Kohler.

Literatur

- Anders, Günther (1956): *Die Antiquiertheit des Menschen. Bd. 1: Über die Seele im Zeitalter der zweiten industriellen Revolution*. München: Beck.
- Anderson, Michael and Susan Leigh Anderson (Hg.) (2007): *Machine Ethics*, Cambridge: Cambridge University Press.
- Arendt, Hannah (1981, 2002): *Vita activa oder vom tätigen Leben*, Piper: München.
- Böhme, Hartmut, Gernot Böhme (1983): *Das Andere der Vernunft. Zur Entwicklung von Rationalitätsstrukturen am Beispiel Kants*, Frankfurt am Main: Suhrkamp.
- Cassirer, Ernst (2006): *An Essay on Man. An Introduction to a Philosophy of Human Culture*, in: ders.: *Gesammelte Werke*, Bd. 23, hg. v. Birgit Recki, Hamburg: Meiner.
- Coeckelbergh, Mark (2012): »Can we trust robots?«, *Ethics Inf Technol*, 14: 53–60.
- Crawford, Kate (2021): *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*, New Haven.
- Dick, Philip K. (2012): *Do Androids Dream of Electric Sheep?* London: Weidenfeld & Nicolson.
- Floridi, Luciano, J.W. Saunders (2004): »On the Morality of Artificial Agents«, *Minds and Machine*, 14: 349–379.
- Floridi, Luciano, Josh Cows, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, Effy Vayena (2018): »AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations«, *Minds and Machines*, 28(4): 689–707.
- Gehlen, Arnold (1957): *Die Seele im technischen Zeitalter. Sozialpsychologische Probleme der industriellen Gesellschaft*, Hamburg: Rowohlt.
- Gunkel, David J., Joanna J. Bryson (eds.) (2014): »Machine Morality: The Machine as Moral Agent and Patient«, special issue of *Philosophy & Technology*, 27(1): 1–142.
- Gunkel, David J. (2017): *The machine question. Critical perspectives on AI, robotics, and ethics*, Massachusetts: The MIT Press.
- Heidegger, Martin (1993): *Sein und Zeit*, Tübingen: Niemeyer.
- Heidegger, Martin (2000): *Die Frage nach der Technik*, in: ders.: *Vorträge und Aufsätze. Gesamtausgabe*, Bd. 7, hg. v. Friedrich-Wilhelm v. Herrmann, Frankfurt am Main: Klostermann.
- Hilgers, Philipp von (2010): »Ursprünge der Black Box«, in: Anna Ofak und Philipp von Hilgers (Hg.): *Rekursionen. Von Faltungen des Wissens*, München: Wilhelm Fink.
- Hoffmann, E.T.A. (1991): *Der Sandmann*, Stuttgart: Reclam.
- Homer (1990): *Odyssee*, in der Übertragung von Johann Heinrich Voß. Frankfurt am Main: Insel.

- Hubig, Christoph (2008): *Die Kunst des Möglichen, Bd. 1: Grundlinien einer dialektischen Philosophie der Technik*, Transcript: Bielefeld.
- Ihde, Don (1990): *Technology and the Lifeworld. From Garden to Earth*, Bloomington: Indiana University Press.
- Kaminski, Andreas (2018): »Der Erfolg der Modellierung und das Ende der Modelle. Epistemische Opazität in der Computersimulation«, in: Andreas Brenneis, Oliver Honer, Sina Keesser, Annette Ripper, Silke Vetter-Schultheiß (Hg.) (2018): *Technik – Macht – Raum. Das Topologische Manifest im Kontext interdisziplinärer Studien*. Wiesbaden: Springer VS.
- LaMettrie, Julien Offray de (2001): *Der Mensch eine Maschine*, Stuttgart.
- Leveringhaus, Alex (2016): *Ethics and Autonomous Weapons*. London: Palgrave-Macmillan.
- Lévinas, Emmanuel (1983): *Die Spur des Anderen. Untersuchungen zur Phänomenologie und Sozialphilosophie*. Freiburg: Karl Alber.
- Liggieri, Kevin, Oliver Müller (2019): »Mensch-Maschine-Interaktion seit der Antike: Imaginationsräume, Narrationen und Selbstverständnisdiskurse«, in: Kevin Liggieri, Oliver Müller (Hg.): *Mensch-Maschine-Interaktion. Geschichte-Kultur-Ethik* (Metzler-Handbuch), Stuttgart: J.B. Metzler.
- Lin, Patrick, Keith Abney, Ryan Jenkins (eds.) (2017): *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, New York: Oxford University Press.
- McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, Claude E. Shannon (1955): »A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence«, (August 31, 1955). Zitiert nach AI Magazine, 27(4) (2006): 12–14.
- Meyer-Drawe, Käte (1996): *Menschen im Spiegel ihrer Maschinen*, München: Wilhelm Fink.
- Meyer-Drawe, Käte (2018): »Androiden«, in: Carlos Kölbl, Anna Sieben (Hg.): *Stichwörter zur Kulturpsychologie. Jürgen Straub zum 60. Geburtstag*, Gießen: Psychosozial Verlag.
- Misselhorn, Catrin (2018): *Grundfragen der Maschinenethik*, Stuttgart: Reclam.
- Misselhorn, Catrin (2021): *Künstliche Intelligenz und Empathie. Vom Leben mit Emotionserkennung, Sexrobotern und Co*, Stuttgart: Reclam.
- Müller, Oliver (2014): *Selbst, Welt und Technik. Eine anthropologische, geistesgeschichtliche und ethische Untersuchung*, Berlin, New York: de Gruyter.
- Nørskov, Marco (ed.) (2017): *Social Robots*, London: Routledge.
- Plessner, Helmuth (1985): *Die Utopie in der Maschine*, in: ders.: *Gesammelte Schriften*, Bd. 10, hg. v. Günter Dux, Odo Marquard und Elisabeth Ströker. Frankfurt am Main: Suhrkamp.
- Ricœur, Paul (1996): *Das Selbst als ein Anderer*, München: Wilhelm Fink.
- Rosa, Hartmut (2016): *Resonanz. Eine Soziologie der Weltbeziehung*, Berlin: Suhrkamp.
- Schapp, Wilhelm (1985): *In Geschichten verstrickt. Zum Sein von Mensch und Ding*, Frankfurt am Main: Klostermann.

- Simmel Georg (1989): *Philosophie des Geldes*, in: *Gesamtausgabe*, Bd. 6, hg. von David P. Frisby, Klaus Christian Köhnke, Frankfurt am Main: Suhrkamp.
- Spiekermann, Sarah (2019): *Digitale Ethik. Ein Wertesystem für das 21. Jahrhundert*, München.
- Turing, Alan (1950): »Computing machinery and intelligence«, *Mind*, 49: 433–460.
- Völker, Klaus (Hg.) (1976): *Künstliche Menschen. Dichtungen und Dokumente über Golems, Homunculi, Androiden und liebende Statuen*, München: dtv.
- Vöneky, Silja, Philipp Kellmeyer, Wolfram Burgard, Oliver Müller (2022, i. Ersch.): *The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives*, Cambridge: Cambridge University Press.
- Wagner, Alan R., Jason Borenstein, and Ayanna Howard (2018): »Overtrust in the robotic age«, *Communications of the ACM*, 61(9): 22–24.
- Wagner, Johanna (2020): *Künstliche Intelligenzen als moralisch verantwortliche Akteure*, Paderborn: mentis.
- Waldenfels, Bernhard (1994): *Antwortregister*, Frankfurt am Main: Suhrkamp.
- Waldenfels, Bernhard (2019): *Erfahrung, die zur Sprache drängt. Studien zur Psychoanalyse und Psychotherapie aus phänomenologischer Sicht*, Berlin: Suhrkamp.
- Weizenbaum, Joseph (1966): »ELIZA-A Computer Program For the Study of Natural Language Communication Between Man and Machine«, *Communications of the ACM*, 9:36–35.
- Zuboff, Shoshana (2019): *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York: Public Affairs.