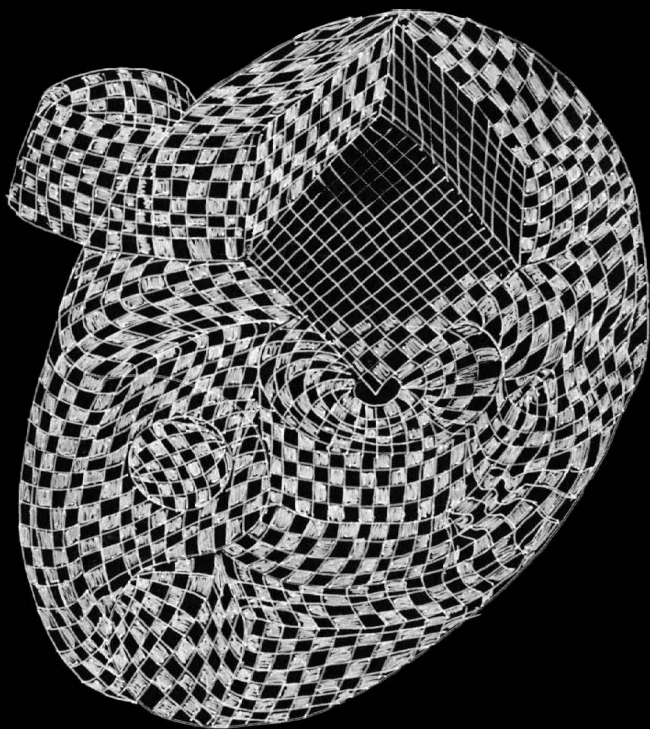


Strasser · Sohst · Stapelfeldt · Stepec (Hg.)

Künstliche Intelligenz

Die große Verheißung



MoMo berlin Philosophische
KonTexte Band 8

Anna Strasser / Wolfgang Sohst / Ralf Stapelfeldt / Katja Stepec (Hg.)

**KÜNSTLICHE INTELLIGENZ –
DIE GROSSE VERHEISSUNG**

Anna Strasser / Wolfgang Sohst /
Ralf Stapelfeldt / Katja Stepec (Hg.)

Künstliche Intelligenz – Die große Verheißung

MoMo Berlin
Philosophische KonTexte
Band 8

Impressum:

Anna Strasser / Wolfgang Sohst / Ralf Stapelfeldt / Katja Stepec (Hg.)

Künstliche Intelligenz – Die große Verheißung

ISBN 978-3-942106-79-5

© 2021 xenomoi Verlag, Berlin
Satz in Palatino Linotype 10 Pt.

Cover-Bild und Bilder zwischen den Texten: Anna Strasser
Umschlaggestaltung, Satz und Produktion:
xenomoi Verlag e.K., Heinersdorfer Str. 16, D - 12209 Berlin
Tel.: 030 - 755 11 712 ▪ www.xenomoi.de ▪ info@xenomoi.de

Bibliographische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliographie; detaillierte Daten sind im Internet unter
<https://portal.dnb.de/opac.htm> abrufbar.

INHALTSVERZEICHNIS

Vorwort	5
Erster Teil Der Begriff der Intelligenz	
<i>Katja Stepec</i> Einleitung Teil 1: Der Begriff der Intelligenz	13
<i>Daniel C. Dennett</i> Turings seltsame Umkehrung der Argumentation Was uns Darwins Evolutionstheorie über Künstliche Intelligenz verrät	27
<i>Nadine Schumann / Yaoli Du</i> Grenzgänge: Von Menschen zu smarten Maschinen - und zurück?	37
<i>Hans Zillmann</i> Künstliche Subjektivität? – Eine Analyse anhand der neurowissenschaftlichen Subjektbeschreibung Antonio Damasio	63
<i>Michael Meyer-Albert</i> Die Frage nach der Weltoffenheitmaschine. Drei Probleme bei einer möglichen Erschaffung einer künstlichen Intelligenz.	83
<i>Jan Tobias Fuhrmann</i> Strukturkonservative Algorithmen: Künstliche Intelligenz als Kommunikationsproblem	103
<i>Christian Freksa</i> KI und Kognition	129
<i>Daniel Wenz</i> Künstliche Intelligenz in mathematischen Beweisen und das Problem der Erklärbarkeit	145
<i>Christoph Merdes</i> Können Maschinen moralisch handeln?	169
<i>Walther Ch. Zimmerli</i> Künstliche Intelligenz und postanaloges Menschsein. Entstehung, Entwicklung und Wirkung eines realen Mythos	193

Inhaltsverzeichnis

<i>Mathias Gutmann / Marie-Claire Haag / Christian Wadehul</i> Verheißung, Verdammung oder einfach ein Selbstmissverständnis? Sprachkritische Überlegungen zum Umgang mit KI und ihren Beschreibungen	221
--	-----

Zweiter Teil Grenzen und Folgen Künstlicher Intelligenz

<i>Ralf Stapelfeldt</i> Einleitung Teil 2: Grenzen und Folgen Künstlicher Intelligenz	243
<i>Hans-Jörg Kreowski, Wolfgang Krieger</i> Künstliche Intelligenz – ,künstlich‘ ja, ‚Intelligenz‘ wohl kaum	259
<i>Elektra Wagenrad</i> Das clevere Pferd Hans und die Blackbox der KI	279
<i>Dieter Mersch</i> Vorbemerkungen zu einer Kritik algorithmischer Rationalität. Denken, Kreativität und Künstliche Intelligenz	305
<i>Gergana Vladova / Sascha Friesike</i> Irren bleibt menschlich: Wieso falsche Entscheidungen für uns so wichtig sind und welche Rolle künstliche Intelligenz dabei nicht spielen kann	325
<i>Sybille Krämer</i> Nüchtern bleiben! Künstliche Intelligenz jenseits des Mythos	335
<i>Rico Hauswald</i> Digitale Orakel? Wie künstliche Intelligenz unser System epistemischer Arbeitsteilung verändert	359
<i>Thomas Weiß</i> Künstliche Intelligenz – eine marxistische Betrachtung	379
<i>Reinhard Kahle</i> Wozu (ver)führt uns die neue KI?	407
<i>Uwe Engel / Holger Schultheis</i> KI assistiert, der Mensch entscheidet. Ergebnisse der ersten Runde des Delphi-Surveys „Blick in die Zukunft. Wie künstliche Intelligenz das Leben verändern wird“	419

Inhaltsverzeichnis

<i>John Michael</i>	
Interaktionen mit Robotern – Wie verbindlich kann das sein	445
<i>Hendrik Kempt</i>	
Zwischenmenschlichkeit für Maschinen	453
<i>Ophelia Deroy</i>	
Rechtfertigende Wachsamkeit gegenüber KI	471
<i>Catrin Misselhorn</i>	
Grundsätze der Maschinenethik	489
<i>Anna Strasser</i>	
Resümee: Kann Intelligenz Großes verheißen – selbst wenn sie nur künstlich ist?	503
Biographische Notizen	509

Inhaltsverzeichnis

VORWORT DER HERAUSGEBER*INNEN

Der Untertitel dieses Bandes „Die große Verheißung“ ist im Grunde eine Frage: Was stellt die Künstliche Intelligenz der Menschheit insgesamt in Aussicht? Was macht sie als zunächst nur technische Möglichkeit zur vielfach gepriesenen Hoffnung? In welcher Hinsicht sind solche Hoffnungen überhaupt erfüllbar und nicht vielmehr die ersten Zeichen einer bevorstehenden Apokalypse?

Die Faszination der Künstlichen Intelligenz liegt offenbar in dem Versprechen des kühnsten aller bisher in der Menschheitsgeschichte unternommenen Versuchs, sich als Spezies selbst zu übersteigen. In den öffentlichen Debatten wird das Projekt der Künstlichen Intelligenz oft als fundamentaler dargestellt, als all die Abenteuer, neue Kontinente zu entdecken, bleierne Materie alchemistisch in leuchtendes Gold zu verwandeln oder durch mechanische Maschinenmonster buchstäblich neue Welten aus dem Boden zu stampfen. Damit scheint sie als Zeichen und Zepter einer Allmacht, die vormals nur den Göttern zugetraut wurde, verstanden zu werden.

Eine solche Überhöhung des Projektes der Künstlichen Intelligenz wirft einige Fragen auf. Mit dem vorliegenden Band wollen wir einige subtilere Perspektiven beisteuern, die über das allgemeine Bejubeln und Verteufeln hinausgehen. Das Unternehmen Künstliche Intelligenz hat offenbar einige konzeptionelle Risse, die hier allgemeinverständlich untersucht werden sollen. Die in diesem Band versammelten Betrachtungen treffen den Nerv einer Zeit, die im globalen Kontext gegen vielfach sich aufdrängende politische Verzweiflung und drohende soziale Verwirrung ankämpfen muss. Die Künstliche Intelligenz ist hier nur eine von sehr vielen Bemühungen um eine neue Ordnung unter den Menschen, wenn auch eine besonders prominente. In der Auseinandersetzung mit ihr kann die philosophische Reflexion das leisten, was die fachliche Spezialkompetenz oft außer Acht lässt, nämlich die übergreifende Zusammenführung von Perspektiven, Anregungen, Bedenken und Korrekturen drohender Irrtümer.

Ebenso stellen sich neue grundlegende ethische Fragen, wenn man über die zukünftige Entwicklung von Künstlicher Intelligenz nachdenkt. Geht das Projekt Künstliche Intelligenz mit drastischen Veränderungen bezüglich unseres Verständnisses von Vertrauen, Freund-

schaft und Selbstbestimmung einher, oder können wir zukünftige Entwicklungen gestalten?

Schon begrifflich ist die Annäherung an das, was unter Künstlicher Intelligenz verstanden wird oder werden sollte, ein kontroverses Feld. Welche Art von Intelligenz – analytische, soziale oder emotionale – ist gemeint, welche Voraussetzungen haben die vielen Begriffe von Intelligenz, welche Konsequenzen dürfen wir aus ihnen ziehen? Und welchen Wesen oder Entitäten können wir welche Arten der Intelligenz plausiblerweise zuschreiben?

Als Herausgeber*innen dieses Bandes freuen wir uns, dass unser Ruf nach Beiträgen auf großes und positives Echo gestoßen ist. Die Relevanz des Gegenstandes dieses Buches zeigt sich auch daran, in welchem Umfang sehr bekannte Stimmen dieses Diskurses bereit waren, sich an einem solchen Projekt zu beteiligen. In diesem Sinne sind wir froh, auch Beiträge von Autoren wie Daniel Dennett, Ophelia Deroy, Sybille Krämer, Dieter Mersch, Catrin Misselhorn oder Walther Ch. Zimmerli gewonnen zu haben. Unser besonderes Anliegen, eine im Einzelnen interdisziplinäre Debatte fortsetzen zu können, deren verbindende methodische Klammer der philosophische Diskurs ist, wird aber erst durch die Gesamtheit der versammelten Aufsätze erfüllt. In diesem Sinne hoffen wir, zur gegenwärtigen Beurteilung der Chancen und Risiken, aber auch der gesellschaftlichen Reichweite von Künstlicher Intelligenz weiteres, wertvolles Material beisteuern zu können.

Schlussendlich möchten auch wir als Herausgeber*innen unsere Motive zur Erarbeitung dieses Bandes vorstellen. Wir kennen uns aus dem Berliner philosophischen Arbeitskreis ‚MoMo Berlin‘. Anna Strasser ist außerdem die Gründerin der ‚DenkWerkstatt Berlin‘. Beide Institutionen verstehen sich als Plattformen, um aktuelle philosophische Diskurse auch einer breiteren Öffentlichkeit zugänglich zu machen.

Anna Strasser hat im Jahr 2004 über Kognition künstlicher Systeme promoviert und seitdem in Freiburg und Berlin an der Universität zu dem Spezialgebiet sozialer Kognition im Schnittpunkt von Psychologie, KI und Philosophie geforscht. Katja Stepec promovierte über die philosophische Frage des Übersetzens im Zusammenhang mit den holistischen Bedeutungstheorien von Quine, Davidson und Brandom. Ralf Stapelfeldt schrieb seine philosophische Masterarbeit über Daniel Dennett, und Wolfgang Sohst hat mehrere Monographien und Aufsätze zur Prozessmetaphysik und -logik sowie zur kollektiven Ethik veröffentlicht.

Nachdem die Künstliche Intelligenz auch in der Tagespresse inzwischen ein ständiges Thema ist und wir immer wieder über die verschiedenen Fragestellungen dazu sprachen, entschlossen wir uns, den vorlie-

Vorwort

genden Band in der Reihe ‚MoMo Berlin Philosophische KonTexte‘ zu erarbeiten. Wir freuen uns über das reichhaltige Ergebnis.

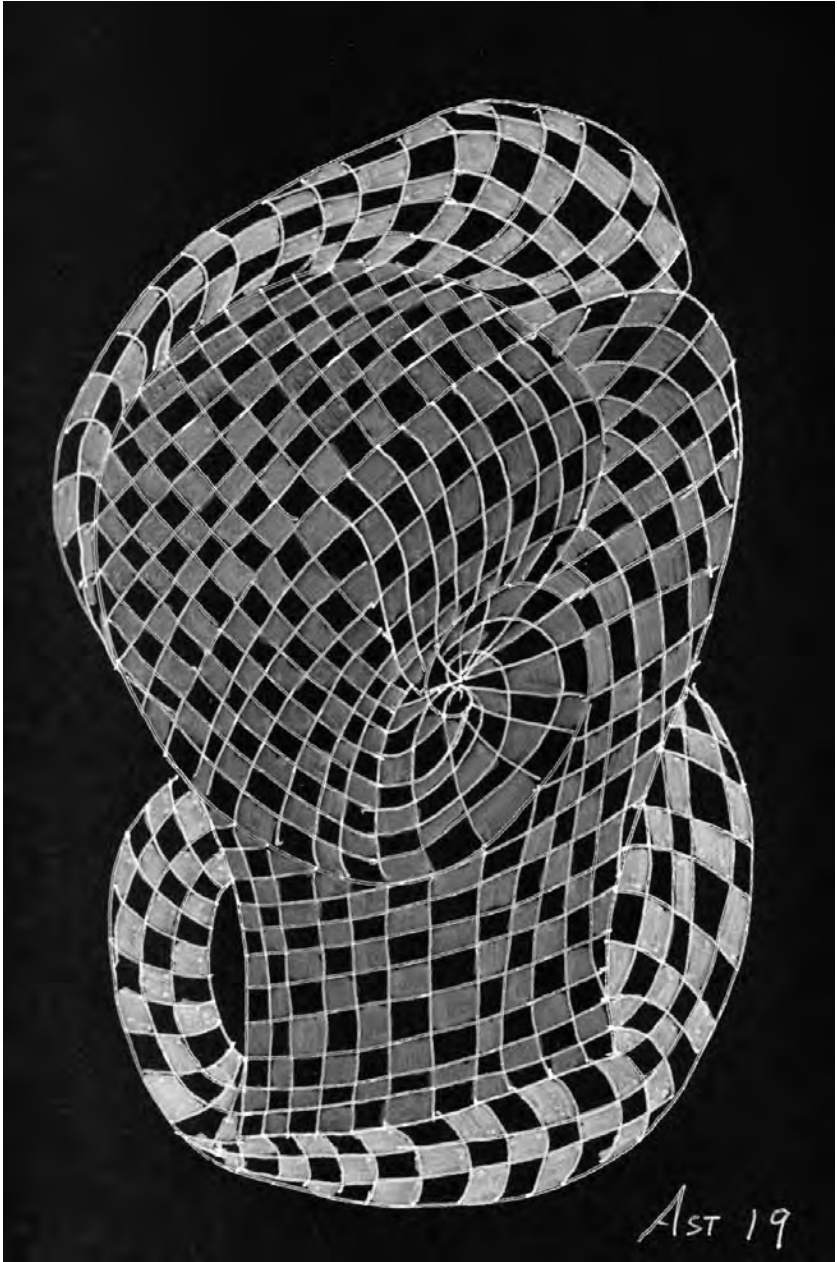
Berlin, im Januar 2021

Dr. Anna Strasser

Wolfgang Sohst

Ralf Stapelfeldt

Dr. Katja Stepec



TEIL 1

DER BEGRIFF DER INTELLIGENZ

Katja Stepec

EINLEITUNG TEIL 1:

DER BEGRIFF DER INTELLIGENZ

Der vorliegende Sammelband hat den Titel „Künstliche Intelligenz – die große Verheißung“. Mögliche Konsequenzen und Bewertungen der künstlichen Intelligenz bedürfen einer Auseinandersetzung, was der Begriff der Intelligenz – und der künstlichen Intelligenz – im Allgemeinen und für uns bedeutet. Die Aufsätze im ersten Teil des Sammelbandes wollen genau dies: die unterschiedlichen Bedeutungen des Begriffs der Intelligenz und der künstlichen Intelligenz beleuchten.

Hilfreich für eine begriffliche Analyse ist die Aufschlüsselung der Voraussetzungen und der Konsequenzen des Begriffsgebrauchs. In der analytischen Sprachphilosophie wird auch von der inferentiellen Rolle eines Satzes gesprochen. Diese umfasst die Umstände sowie die Folgen einer Behauptung, also die Menge ihrer Prämissen und Konsequenzen. (vgl. z.B. Brandom 2000) Entsprechend ergibt sich die Bedeutung des Begriffs ‚Intelligenz‘ in Sätzen aus den Prämissen, unter denen man von Intelligenz sprechen kann, und die Folgen, die sich daraus ergeben. In Abhängigkeit der unterschiedlichen Voraussetzung und Konsequenzen kann ein Begriff unterschiedliche Gehalte aufweisen.

Ich möchte drei Bereiche von Voraussetzungen vorstellen, die alltagssprachlich mit Intelligenz verbunden werden und jeweils bestimmte Voraussetzungen von Intelligenz beinhalten: analytische Intelligenz, soziale Intelligenz und emotionale Intelligenz. Den drei Konzepten ist gemeinsam, dass Intelligenz als Fähigkeit verstanden wird. Meine Beschreibung der Voraussetzungen der drei Bereiche ist exemplarisch und keinesfalls abschließend, wie sich später auch aus den Beiträgen der Kapitel noch ergibt.

Drei exemplarische Voraussetzungsbereiche des Begriffs der Intelligenz

Analytische Intelligenz ist die Fähigkeit, ein Problem zu lösen. Voraussetzungen dafür ist die Aufnahme von Informationen, deren Auswertung für eine Fragestellung und das Liefern eines Ergebnisses. Der Prozess beruht auf regelgeleiteten Verarbeitungsschritten, die sich z.B.

als Schlussfolgerungen oder Algorithmen darstellen. Analytische Intelligenz kann auch bereits beinhalten, dass ein Wesen mit Störungen umgehen und unerwartete Hindernisse bewältigen kann. Diese Art der Intelligenz ist zielgerichtet und ergebnisorientiert und erfordert kein Verstehen der Aufgabe und deren Lösung.

Ein zweiter Bereich fügt den Voraussetzungen für Intelligenz soziale Aspekte hinzu. *Soziale Intelligenz* umfasst generell die Kooperation mit anderen Wesen, worunter Kommunikation und gemeinsames Handeln verstanden werden kann. Diese Art der Intelligenz ist ebenfalls ergebnisorientiert, aber auch anspruchsvoller, denn sie umfasst nun die Abstimmung mit anderen Wesen zur Erreichung eines gemeinsamen Ziels. Für die Zielerreichung ist daher nicht mehr nur eine Information über den Sachverhalt relevant, sondern auch Informationen über die Kooperationspartner: z.B. ihre Überzeugungen. Entsprechend ist Voraussetzung der sozialen Intelligenz nicht nur die reine Erfassung von Informationen im Sinne eines Inputs, sondern Erfassung im Sinne eines Austauschs von Informationen. Dafür ist Kommunikation notwendig. In der sozialen Interaktion erfolgt Kommunikation auf einer ‚öffentlichen‘ Ebene, also auf Basis gemeinsamer, standardisierter Normen, die sprachlicher, logischer oder mathematischer Art sein können. Kommunikation impliziert zumindest ein basales Verstehen dessen, was man tut. Für eine kommunikative Abstimmung mit Kooperationspartnern reicht es nicht aus, die internen Prozesse zu beschreiben, man muss sie auch mit den Prozessen der Beteiligten vergleichen können. Die Einnahme einer zweiten Perspektive, das Erkennen von Gemeinsamkeiten und Unterschieden sind Momente des Verstehens. Aus der Kommunikation und dem gegenseitigen Abgleich im Sinne eines Verstehens folgt eine koordinierte Aktion, die man unter Umständen bereits als gemeinsames Handeln bezeichnen kann. Voraussetzungen für soziale Intelligenz beinhalten also, dass das intelligente Wesen die Überzeugungen der anderen Wesen berücksichtigen und mit den eigenen abgleichen kann; dass es die eigene Aktion mit den Aktionen der anderen Wesen zu einem gemeinsamen Handeln koordinieren kann. Soziale Kooperation ist komplexer als das individuelle Lösen einer Aufgabe, erfordert daher mehr und andere Ressourcen, bietet aber möglicherweise auch schnellere und bessere Ergebnisse.

Schließlich können emotionale Voraussetzungen angenommen werden, die für die Qualifikation von Intelligenz relevant sind: *emotionale Intelligenz* beinhaltet das Erkennen und Bewerten von Gefühlen und Stimmungen und zwar bei sich und bei anderen. Emotionen können eine Problemlösung behindern oder beschleunigen. Unsicherheit wirkt

sich beispielsweise störend auf die Erreichung eines Ziels aus; Neugierde und die Aussicht auf Belohnung wirken sich positiv auf Problemlösungen aus. Emotionale Wesen mit der Fähigkeit, Gefühle zu reflektieren, zu bewerten und zu manipulieren, können diese zu ihrem Vorteil bei der Bearbeitung von Aufgaben einsetzen; Wesen, die darüber hinaus auch ihre Kooperationspartner positiv emotional beeinflussen, sind noch erfolgreicher. Die Kommunikation erfolgt hier in der Regel nicht auf einer öffentlichen oder expliziten Ebene, sondern bleibt subjektiv und implizit. Dazu gehören Fähigkeiten wie das Wahrnehmen von Stimmungen oder das subtile ‚Anstupsen‘ durch sprachliche Instrumente. Auf dieser Ebene kann zudem ein tieferes Verstehen erwartet werden: ein Verstehen nicht nur der Aufgabe und der Prozesse, die zu ihrer Erledigung notwendig sind, sondern auch ein Verstehen der eigenen und fremden Subjektivität.

Für die hier vorgeschlagene begriffliche Struktur verstehe ich diese Bereiche von Voraussetzungen als aufeinander aufbauend. Die Voraussetzungen aus dem analytischen Bereich gehören zum Kernbereich von Intelligenz. Ein Sprecher, der den Ausdruck Intelligenz verwendet, sollte diese Voraussetzungen immer akzeptieren. Die Voraussetzungen aus dem sozialen Bereich beinhalten zusätzliche Anforderungen aufgrund der (öffentlichen) Interaktion und Kommunikation mit anderen Wesen. Entsprechend erweitert sich die Bedeutung des Begriffs Intelligenz um die Begriffe der Sprache, des Verstehens und des gemeinsamen Handelns. Darauf aufbauend bringt der dritte Bereich subjektive Voraussetzungen in die Bedeutung von Intelligenz ein, und erweitert damit den Begriff zusätzlich um (implizite) emotionale Aspekte, sowie um ein weitergehendes reflektierendes Verstehen.

Konsequenzen des Begriffsgebrauchs: Mensch oder Maschine?

Die Bedeutung des Begriffs Intelligenz hängt nicht nur davon ab, welche Voraussetzungen in den Begriff einfließen, sondern auch von den Konsequenzen, die daraus gezogen werden. Nicht aus jedem Voraussetzungsbereich dürfen die gleichen Schlussfolgerungen gezogen werden. Ich konzentriere mich hier auf zwei sehr allgemeine Schlussfolgerungen: Mensch oder Maschine.

In einer ersten Intuition scheint die Schlussfolgerung aus allen drei Bereichen auf den Menschen hinauszulaufen. Menschen wird sowohl die Fähigkeit zur Erfassung und Verarbeitung von Informationen zugeschrieben, als auch die Fähigkeiten zur Kommunikation und zum

gemeinsamen Handeln und schließlich zum Erkennen eigener und fremder Emotionen. Mit der Schlussfolgerung auf den Menschen sind weitere Konsequenzen möglich: z.B. die Annahme von Selbstbewusstsein und freiem Willen.

Für die Schlussfolgerung auf eine Maschine dagegen reichen zunächst die Voraussetzungen der analytischen Intelligenz aus. Eine solche problemlösende Maschine kommt dem nahe, was wir gemeinhin unter einem Computer verstehen: Eine Fragestellung wird durch systemimmanente Algorithmen bearbeitet und generiert eine Lösung. Hier sind weder Schlussfolgerungen auf Bewusstsein noch auf freien Willen angebracht, sondern das Vorgehen wird als rein instrumentell betrachtet.

Mensch und Maschine sind jeweils die extremen Enden einer graduellen Reihe von Schlussfolgerungen mit Zwischenpositionen, wobei sich bestimmte Kombinationen gängigen philosophischen Positionen zuordnen lassen. So weist eine *Kombination der analytischen mit der sozialen und der emotionalen Intelligenz* auf die These der starken KI hin, und erlaubt damit die Schlussfolgerung auf eine Gleichsetzung von Mensch und Maschine. Die These geht davon aus, dass Maschinen eine allgemeine Intelligenz aufweisen können, die der des Menschen gleicht oder sie übertrifft. Mit dieser These wird oft die Behauptung verbunden, dass Maschinen denken oder ein Bewusstsein haben. Eine solche Gleichsetzung von Mensch und Maschine setzt voraus, dass Maschinen auch soziale und emotionale Fähigkeiten haben können. In dem Aufsatz *“What is it like to be a Bat”* (Nagel 1974) beschreibt Nagel die subjektive Qualität als das Gefühl, wie es ist, in einem bestimmten mentalen Zustand zu sein. Um diese Art der Subjektivität nachempfinden zu können, setzt Nagel voraus, dass nur ausreichend ähnliche Wesen einen Standpunkt gegenseitigen Verstehens einnehmen können. Wesen oder Entitäten, die also alle drei Voraussetzungen erfüllen, lassen die Schlussfolgerung zu, dass es sich um einen Menschen oder ein menschenähnliches Wesen handelt, unabhängig davon, ob dieses Wesen organisch, auf Silikonbasis oder durch Algorithmen realisiert ist.

Eine *Kombination von analytischer und sozialer Intelligenz* würde dagegen in vielen Fällen ausreichen für die Schlussfolgerung, dass der Turing Test bestanden ist. In dem Artikel *“Computing Machinery and Intelligence”* (Turing 1950) schlägt Turing ein Imitationsspiel vor, in dem eine Person Fragen an zwei verborgene Kandidaten stellt und aufgrund der Antworten entscheiden muss, ob einer der beiden Kandidaten eine Maschine ist. Die Maschine besteht den Test, wenn sie so erfolgreich kommuniziert, dass die Person nicht bemerkt, dass es eine Maschine

ist. Eine solche Maschine weist sprachliche Fähigkeiten in einem begrenzten Kontext auf, wobei zusätzlich zu beachten ist, dass ein solches eingeschränktes Sprachspiel nicht immer auch schon Verstehen beinhaltet. Wird eine Maschine durch soziale Fähigkeiten angereichert, sind dennoch Schlussfolgerungen möglich, dass es sich um mehr als ein Instrument handelt.

Reduziert sich die Voraussetzung auf *analytische Fähigkeiten*, bleibt die Annahme instrumenteller Intelligenz, die beispielsweise durch Searles Beschreibung des chinesischen Zimmers illustriert wird (Searle 1980). Hier kombiniert ein in einem Zimmer befindlicher Mensch mit Hilfe eines Regelbuches Kärtchen mit chinesischen Zeichen und reagiert damit auf eine Frage in chinesischer Sprache. Aus Sicht des außerhalb stehenden Fragenden mag der Mensch eine zufriedenstellende Antwort gegeben haben, gucken wir jedoch in das chinesische Zimmer hinein, werden dort lediglich Zeichen gemäß einer formalen Regel, einer Syntax, angeordnet – ähnlich einem Computer, der Symbole gemäß einer mathematischen Regel, eines Algorithmus, kombiniert. Der Mensch ist weder der chinesischen Sprache mächtig, noch hat er ein Verständnis von Frage und Antwort.

Die hier hilfswise vorgeschlagene Struktur des Begriff der Intelligenz ist durch diese einfache Kombinatorik längst nicht umfassend beschrieben. So kann durchaus in Frage gestellt werden, ob es weitere Voraussetzungen gibt – z.B. aus dem Bereich der Kunst, Literatur oder Musik. Und auch die zunächst einleuchtende Behauptung eines Kernbereichs von Intelligenz kann bestritten werden. Beispielsweise ist ein Wesen vorstellbar, dem lediglich soziale und emotionale Intelligenz zugeschrieben werden kann – oder auch nur einer dieser Bereiche. Schließlich sind auch graduell unterschiedliche Kombinationen vorstellbar – beispielsweise ein Wesen ausgestattet mit geringfügigen analytischen Fähigkeiten, dafür aber mit überdurchschnittlicher sozialer und / oder emotionaler Intelligenz.

Im ersten Teil dieses Bandes stellen Autoren aus unterschiedlichen Bereichen ihre Überlegungen zum Begriff der Intelligenz, seinen Voraussetzungen und seinen Konsequenzen – also auf seine Anwendungsbereiche – vor.

Zu Beginn stellt der Aufsatz von **Daniel Dennett** "Turings seltsame Umkehrung der Argumentation. Was uns Darwins Evolutionstheorie über Künstliche Intelligenz verrät." die provokante Frage, ob wir nicht die Perspektive wechseln sollten, wenn es um das Verhältnis von Mensch und Maschine geht. Vielleicht nehmen wir intuitiv an, dass so etwas wie Intelligenz oder Bewusstsein nur durch eine höhere Intelli-

genz oder ein überlegeneres Bewusstsein in die Welt kommen kann. Mit Hilfe einer Analogie zwischen den Thesen Darwins und Turings zeigt Dennett allerdings auf anschauliche Weise, dass fortgeschrittene Organismen oder Fähigkeiten nicht notwendig auf einem spezifischen und sehr komplexen Funktionsschema beruhen müssen, sondern stattdessen in vielen Fällen auf sehr einfachen Prozessen aufbauen. Es ist das Prinzip der einfachen Bausteine, welches nicht nur das Leben und die Evolution erklärt, sondern für Turing auch das menschliche Denken. Konstruiert man auf dieser Basis eine intelligente Maschine, zeigt sich eine Art der ‚Pseudo-Intelligenz‘ oder ein ‚Als-ob‘-Verständnis, welches zumindest als Annäherung an ‚echte‘ künstliche Intelligenz oder Verständnis gewertet werden kann. Dennett drückt damit eine mögliche Erwartung an die Entwicklung der künstlichen Intelligenz aus.

Eine instrumentelle ‚Als-ob‘-Perspektive schlagen auch Schumann und Du vor. Der Aufsatz „Grenzgänge: Von Menschen zu smarten Maschinen – und zurück?“ von **Nadine Schumann & Yaoli Du** konzentriert sich auf die Genese der Schlussfolgerungen auf Mensch / Natur und Maschine / Technik. Der Aufsatz erzählt die Geschichte der Versuche, das eine über das andere zu verstehen. Ausgehend von einem ganzheitlichen Ansatz des Aristoteles, vollzieht sich nach Descartes eine Trennung mit der Tendenz, Natur unter technischen Voraussetzungen zu verstehen. Der Versuch einer Formalisierung der Mathematik durch die Logik und die allgemeine Systemtheorie führen zur Analogie des menschlichen Denkens mit den funktionalen Berechnungen am Modell der Turingmaschine. Auf Basis dieses Erbes kämpft die aktuelle Forschung mit einer *petitio principii*: Zum einen soll der reduktionistische Intelligenzbegriff in der KI-Forschung menschliche Fähigkeiten nachahmend modellieren, zum anderen werden Modelle der KI-Forschung zum Verständnis menschlicher kognitiver Prozesse herangezogen. Am Ende wird eine Sichtweise vorgeschlagen, die nicht auf die Reduktion von Mensch auf Maschine oder des Denkens auf technische Informationsprozesse hinausläuft, sondern mit einer ganzheitlichen ‚Als-ob‘-Betrachtung ein Simulationsmodell des menschlichen Zusammenlebens einbringt, mit dem auch menschliche Subjektivität erfasst werden kann.

Während Schumann und Du einen Weg in Richtung einer sozialen Intelligenz von Maschinen aufzeigen, diskutiert der Aufsatz „Künstliche Subjektivität? – Eine Analyse anhand der neurowissenschaftlichen Subjektbeschreibung Antonio Damasio“ von **Hans Zillmann** die grundsätzliche Möglichkeit einer sozialen und emotionalen Intelligenz

von Maschinen. Zillmann vertritt dabei die These, dass Subjektivität immer eine Frage der sozialen und kulturellen Zuschreibung ist. Auf Grundlage des naturwissenschaftlich orientierten Ansatzes Damasio wird Subjektivität als Auseinandersetzung des Gehirns mit dem eigenen Körper und den Objekten der Umwelt beschrieben und von Zillmann auf zwei Voraussetzungen von Subjektivität konzentriert. Für die erste Voraussetzung unterscheidet Zillmann zwischen Emotionen als physiologischen Reaktionen des Organismus, und Gefühlen als qualitativen Bewertungszustände. Subjektivität besteht entsprechend darin, dass Wesen die emotionale Erfassung ihrer Umwelt über Gefühle wahrnehmen. Zweitens steht diese Art der Subjektivität in einem kulturhistorischen Kontext und dieser Bereich ist von einer physikalistischen Sichtweise nur schwer zu erfassen. Auch wenn Menschen dazu tendieren, Maschinen Subjektivität zuzuschreiben, und darin sogar Gefahren sehen, bleibt die Vermutung, dass Maschinen eher auf das deduktive Schließen beschränkt sein werden und damit auf den Bereich der analytischen Intelligenz.

Emotionale Intelligenz muss nicht die stärkste Voraussetzung für Intelligenz sein. In dem Beitrag „Die Frage nach der Weltoffenheitmaschine. Drei Problemstellen für die Erzeugung einer künstlichen Intelligenz“ von **Michael Meyer-Albert** werden drei weitere Voraussetzungen präsentiert, die dem Begriff der Intelligenz eine anspruchsvolle existenzialistische Bedeutung geben. Auf Basis des Ansatzes McDowells wird Intelligenz vorbereitend als Ko-Konstitution von Denken und Wahrnehmen erläutert. In einem ersten Schritt wird dann diese epistemische Perspektive durch den phänomenologischen Gedanken von Hermann Schmitz im Sinne einer körperlichen Kommunikation mit der Welt angereichert. Mit Hilfe der hermeneutischen Perspektive Gadammers zeigt sich die zeitliche und geschichtliche Dimension der menschlichen Reflexivität, die der vorwärts gerichteten Linearität des algorithmischen Problemlösens ein interagierendes Gespräch mit der Tradition gegenüberstellt. Beides mündet in den Begriff der Weltoffenheit als eine leiblich erfahrene und geschichtliche Reflexion der Intelligenz. Meyer-Albert fügt mit Leiblichkeit und sinnhafter Geschichtlichkeit sowie existentieller Weltoffenheit dem Begriff von Intelligenz nicht nur weitere wichtige Voraussetzungen hinzu, sondern stellt am Ende auch die provozierende Frage nach ethischen Bedenken bei der Erschaffung von künstlichen Intelligenzen.

Eine wichtige Voraussetzung für soziale Intelligenz diskutiert der Beitrag „Künstliche Intelligenz als Kommunikationsproblem“ von **Jan Tobias Fuhrmann**. Er argumentiert dafür, dass Kommunikation nicht

notwendig mit Bewusstsein und Verstehen verbunden sein muss; Intelligenz zeigt sich stattdessen darin, dass die Kommunikation am Laufen gehalten wird. Kern der Kommunikation ist eine strukturelle und sinnhafte Unbestimmtheit, die wesentlich die Möglichkeit kommunikativer Anschlüsse beinhaltet. So sind auch durch Algorithmen gesteuerte Bots, d.h. quasi-umgangssprachliche Kommunikationsprogramme, in der Lage, eine Unterhaltung zu führen und innerhalb eines thematisch begrenzten Bereichs den Turing-Test zu bestehen, ohne dass sie Sinn konstituieren oder sich selbst reflektieren. Tatsächlich stellt sich die fehlende Reflexion als Effizienzvorteil bei der Verarbeitung großer Datenmengen heraus. Das überraschende Ergebnis der Überlegungen ist schließlich, dass gerade die Eigenschaft von Maschinen als nicht-originell und strukturell konservativ dazu führt, dass sie erfolgreich an Kommunikation teilnehmen und diese fortsetzen können.

Mit dem Begriff des Wissens schlägt der Beitrag „Verheißung, Verdammung oder einfach ein Selbstmissverständnis? Sprachkritische Überlegungen zum Umgang mit KI und ihren Beschreibungen“ von **Mathias Gutmann, Marie-Claire Haag und Christian Wadehul** einen weiteren alternativen und vielversprechenden Bereich von Voraussetzungen vor. Die Autoren untersuchen die begrifflichen Zusammenhänge zwischen Wissen, Information und Daten mit dem Ziel, einen tätigkeitstheoretischen Wissensbegriff zu etablieren, ein Wissen um den Zweck einer Verwendung. So wird der operative Wissensbegriff – wie funktioniert eine Maschine – abgegrenzt vom situativen, ‚umgänglichen‘ Wissen um den Zweck einer Maschine, was erst die Beurteilung ihrer Funktionsfähigkeit erlaubt. Unterstützt wird die Betrachtung durch die Unterscheidung eines Funktionierens ‚gemäß einer Regel‘ und dem Handeln ‚nach einer Regel‘ oder in Kenntnis der Regel. Auch der Informationsbegriff enthält bereits eine Zwecksetzung und unterscheidet sich damit von dem Begriff der bloßen Daten. Situatives Wissen aber kann schon in Richtung eines reflexiven Wissens gedeutet werden – nämlich als ein Wissen von sich selbst. Die Einordnung des Begriffs der künstlichen Intelligenz in das begriffliche Umfeld eines zweckorientierten Wissens vermeidet so zwar zunächst die problematische Schlussfolgerung auf Bewusstsein, führt aber zu nicht weniger anspruchsvollen Schlussfolgerungen auf selbständige Zwecksetzung, Rationalität und Verantwortung.

Der Beitrag „Künstliche Intelligenz und kognitive Leistung“ von **Christian Freksa** führt die Idee der zweckorientierten Verwendung als ökologisches Problemlösen weiter aus. Er legt den Fokus auf eine der Voraussetzungen für soziale Intelligenz: Handlungen als

situative Interaktion mit der Welt und Selektion von Informationen im Zusammenhang mit Problemlösungen. Im Bereich der analytischen Intelligenz sind Maschinen zunächst Instrumente zum Lösen von Problemen. Das dafür benötigte umfassende Wissen über die Welt stellt sich als eine Art digitaler Zwilling dar, welcher aus riesigen Datenmengen besteht, die formal und sequenziell organisiert sind. Es handelt sich um Repräsentationen oder Beschreibungen, die die Grundlage für ein Verstehen des Problems im Sinne des Nachvollzuges einzelner Funktionsschritte darstellen. Für die Lösung eines Problems ist die Kenntnis der einzelnen Funktionsschritte jedoch nicht hilfreich; ein handelnder Organismus dagegen erkennt die notwendigen Informationen in der Interaktion mit der Umwelt als "Affordanzen", d.h. als Verwendungsangebote. Für Freksa ist diese Art des 'ökologischen' Problemlösens, das sich in der handelnden Interaktion zeigt, eine Anforderung an künstliche Intelligenz, prägnant gefasst als Intelligenz pro Kilowatt, die deren bisheriges Primitiv, das logische Schließen auf Basis von Repräsentationen, in Frage stellt.

Logisches Schließen ist nicht nur die Grundlage für die Algorithmen, aus denen die Software einer KI besteht, sondern auch die Grundlage für die Mathematik. In dem Beitrag „Künstliche Intelligenz in mathematischen Beweisen und das Problem der Erklärbarkeit“ von **Daniel Wenz** wird daher einleitend gefragt, ob nicht die Nähe der künstlichen Intelligenz zur Mathematik diese für mathematische Beweise besonders qualifiziert (siehe hierzu auch den Beitrag von Dieter Mersch). Warum das nicht der Fall ist, weist auf ein aktuelles Problem der KI hin: der Erklärbarkeit oder besser Nicht-Erklärbarkeit. Komplexe mathematische Beweise müssen nicht nur formal nachvollzogen, sondern auch konzeptuell verstanden werden, indem der Beweis in einen größeren theoretischen Kontext eingeordnet wird. Damit wird nicht nur das eigentliche Problem, sondern auch der mathematische Kontext neu beleuchtet. Die historische Entwicklung der automatisierten Beweisführung dagegen zeigt eine Trennung der Verbindung von formaler und konzeptueller Beweisführung auf. Waren automatisierte Systeme mit dem Ziel der Beweisführung zunächst nach den Fähigkeiten des Menschen modelliert, und sollten primär Aufschluss über den menschlichen Geist geben, wurde diese Ähnlichkeit später zugunsten einer Problemorientierung aufgegeben, mit dem Ziel schnellerer und effizienterer Ergebnisse. Diese führen zu Beweisen, die für den Menschen nicht mehr verstehbar sind und somit auch nicht nutzbar zur Erweiterung mathematischer Kenntnisse. Mit Hilfe einer Art des expressiven Verstehens soll die konzeptuelle Ebene neu erschlossen werden. Dafür

wird die Interaktion zwischen Mensch und Maschine als gemeinsames Handeln gleichwertiger Akteure verstanden, das nicht nur von konstitutiven Regeln bestimmt wird, sondern in dem sich auch implizite Normen entwickeln. Dieses praktische *Know-How* der Teilnehmer, d.h. der konzeptuelle Überschuss, wird explizit gemacht, um undurchsichtige Beweisführung verstehbar und damit nutzbar zu machen.

Der Artikel „Künstliche Moral – Können Maschinen moralisch handeln?“ von **Christoph Merdes** verweist bereits auf die Thematik des zweiten Teils des Sammelbandes. Merdes bejaht nicht nur die Möglichkeit eines künstlichen moralischen Akteurs, sondern schlägt im Wesentlichen drei Kriterien vor, mit denen Moralität von Maschinen beurteilt werden kann. Als erste Bedingung für moralisches Handeln versteht der Autor die Freiheit von Handlungen, und kommt zu dem Ergebnis, dass Maschinen zumindest nicht als völlig fremdbestimmt aufzufassen sind. Auch die zweite Bedingung für moralisches Handeln, die moralischer Rechtfertigung, wird durch die Übertragung von maschinellem Lernen auf eine maschinelle Ethik teilweise bejaht. Mit einem angepassten Turing-Test wird schließlich die dritte Bedingung untersucht: moralische Kompetenz. Hier stellt sich als eigentliches Problem die Auswahl der Beurteilenden heraus – letztlich gibt es also keinen Maßstab für moralische Befähigung, auch der des Menschen selbst.

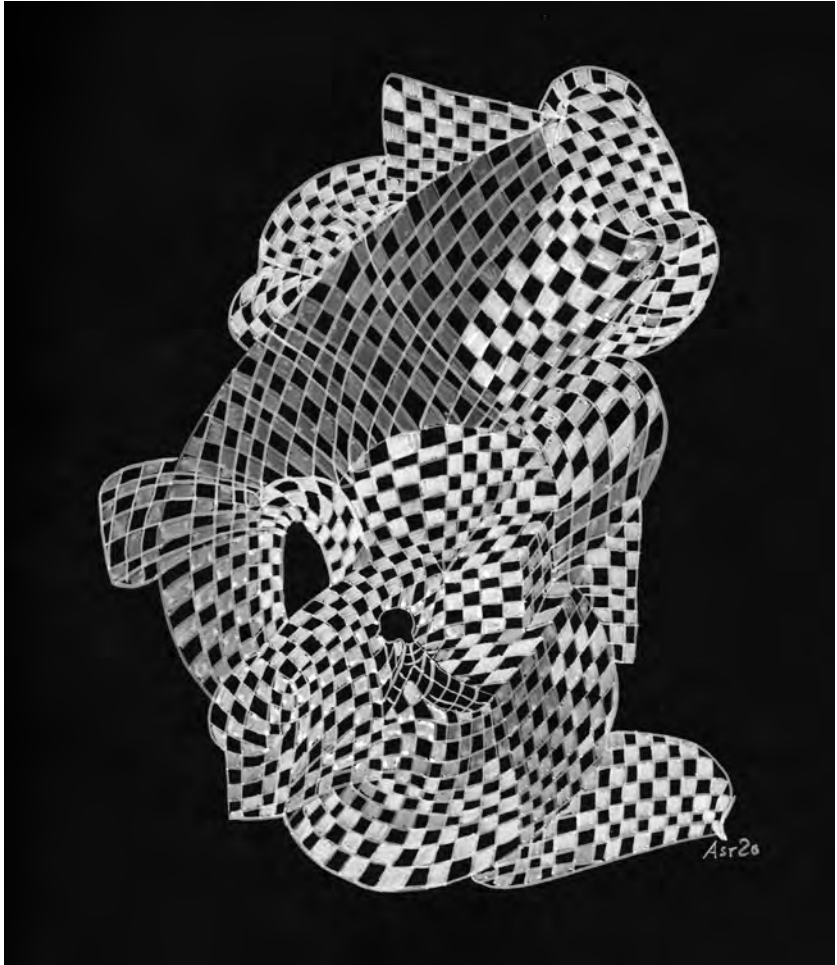
Zum Ende des ersten Teils wird das Moment der „Verheißung“ aufgebrochen in die Begriffe des Hypes und des Mythos. **Walther Ch. Zimmerli** beschäftigt sich in seinem Beitrag „Künstliche Intelligenz und postanaloges Menschsein. Entstehung, Entwicklung und Wirkung eines realen Mythos“ mit der geschichtlichen Entwicklung der Voraussetzungen für künstliche Intelligenz. Im Sinne der genealogischen Methode Nietzsches, durch Aufarbeitung der historischen Entwicklung eines Begriffs dessen Voraussetzungen zu beleuchten, erzählt Zimmerli die Geschichte der KI zwischen Mythos und Hype, also zwischen der versuchten Bändigung der Entwicklung mit Hilfe eines sinnstiftenden Narratives und dem Verselbständigen der Entwicklung im Sinne eines Hypes. Die ‚Erfindung‘ des Begriffs anlässlich der Dartmouth-Conference ist der Beginn des Mythos der ‚denkenden Maschine‘, der durch die Generierung überzogener Erwartungen auch bereits den ersten Hype auslöst. Die Fortsetzung des Hypes kulminiert im endzeitlichen Mythos der Singularität, also der Vorstellung, dass die künstliche die menschliche Intelligenz übertreffen wird. Diese eschatologische Grenze soll überwunden werden durch die Idee des ‚Mensch-Maschine-Tandems‘ was letztlich zu der anthropologischen Frage führt: Trans- oder Posthumanismus – oder: wie transformiert sich der Mensch zwi-

schen dem Analogen und dem Digitalen? So katapultiert uns der Hype der ersten Stunde in ein Zeitalter, in der wir die alten Mythen zugunsten einer neuen Konzeption aufgeben sollten: dem post-analogen Menschen.

Während die Beiträge im ersten Teil bestimmte Voraussetzungen für die Zuschreibung von Intelligenz im Allgemeinen und künstlicher Intelligenz im Speziellen untersuchen und damit das Spektrum der Bedeutung dieser Begriffe ausreizen, werden im zweiten Teil des Sammelbandes nicht nur die Konsequenzen ‚Mensch oder Maschine‘ betrachtet, sondern eine Vielzahl weiterer Folgen, die sich aus dem Begriffsgebrauch ergeben können.

Literatur

- Brandom, Robert (2000): *Expressive Vernunft*. Suhrkamp Verlag, Frankfurt a.M. 2000
- Nagel, Thomas (1974). What Is It Like to Be a Bat? In: *The Philosophical Review*, 83(4), 435-450.
- Searle, John (1980): *Minds, brains, and programs*. In: *The Behavioral and Brain Sciences*, 3, 417-457.
- Turing, Alan (1950). *Computing machinery and intelligence*. In: *Mind*, LIX(2236), 433-460.



Daniel C. Dennett

TURINGS SELTSAME UMKEHRUNG DER ARGUMENTATION

Was uns Darwins Evolutionstheorie über Künstliche
Intelligenz verrät.

Manche der größten und revolutionärsten Errungenschaften der Wissenschaft werden auf eine angenehm bescheidene Weise und ohne viel Aufhebens präsentiert. So gelingt es Charles Darwin, die Kernessenz seiner Theorie in einer Art und Weise zusammenzufassen, die auch für Laien in all ihren Einzelheiten verständlich ist:

Wenn die Lebewesen unter wechselnden Lebensbedingungen in fast allen ihren Teilen individuelle Unterschiede aufweisen (was nicht bestritten werden kann), und wenn unter ihnen infolge ihrer geometrischen Zunahme an Zahl in irgendeinem Alter oder zu irgendeiner Jahreszeit einer harter Kampf ums Dasein stattfindet (was gleichfalls nicht bestritten werden kann), so wäre es in Anbetracht der verwickelten Wechselbeziehungen zwischen den Lebewesen, sowohl unter sich wie zu ihren Lebensbedingungen (die wiederum große, nutzbringende Unterschiede der Struktur, Konstitution und Gewohnheiten hervorrufen) Zufall, wenn niemals Veränderungen vorkämen, die für das Gedeihen dieser Wesen ebenso nützlich sind, wie wir zahlreiche Veränderungen für den Menschen nützlich werden sahen. Wenn nun Veränderungen auftreten, die jedem dieser Lebewesen nützen, so werden sicher die damit beglückten Individuen am ehesten im Kampf ums Dasein erhalten bleiben; und nach dem Prinzip der Vererbung werden sie die Neigung haben, ähnlich charakterisierte Nachkommen hervorzubringen. Dieses Prinzip der Erhaltung oder das Überleben des Tüchtigsten nannte ich natürliche Zuchtwahl.

(Darwin 1963 [1859], 184)

Auch in dem epochalen Aufsatz über die Struktur der DNS von Francis Crick und James Watson findet man einen wunderbar schlichten Abschlussatz:

Es ist unserer Aufmerksamkeit nicht entgangen, dass die von uns postulierten spezifischen Paarbildungen einen möglichen Kopier-

mechanismus des genetischen Materials unmittelbar nahelegen.
(Übersetzung von Watson & Crick 1953, 738)

Als Alan Turing im Jahre 1936 mit seinem Aufsatz über berechenbare Zahlen ein neues Feld für Wissenschaft und Technik eröffnet hat, präsentiert er mit einem sogar noch kürzeren Satz eine Ausgangsbasis für die Lösung eines der rätselhaftesten Probleme der Wissenschaft, dem Leib-Seele-Problem:

Es ist möglich, eine einzige Maschine zu erfinden, die dazu verwendet werden kann, jede berechenbare Folge zu errechnen.

(Turing 2019 [1936], 177)

Turing hat nicht nur die Intuition, dass man so eine bemerkenswerte Herausforderung meistern kann, sondern legt auch konkret dar, wie eine solche Maschine zu konstruieren ist. Dies markiert den Beginn des Computerzeitalters. Dabei sollte man nicht vergessen, dass es schon vor Turings Idee Einheiten gab, die man ‚Rechner‘, also Computer nannte - nämlich Menschen. Anstelle von Computern haben Büroangestellte, von denen sehr viele Frauen waren, mit Geduld, Stolz und ausreichenden mathematischen Kenntnissen Tag für Tag mithilfe von stundenlangen Berechnungen verlässliche Ergebnisse geliefert.

Im Ingenieurwesen, im Handel, im Militär und in anderen Bereichen waren Tausende dieser Angestellten beschäftigt und haben Tabellen für Navigation, Waffenindustrie und für andere technische Bereiche berechnet. Das Revolutionäre an Turings Idee zur Berechenbarkeit kann man besonders gut erkennen, wenn man sie Darwins Evolutionstheorie gegenüberstellt. Bevor Darwins Ideen die Welt prägten, hatten traditionelle Auffassungen eine gewichtigere Rolle als die wissenschaftliche Sicht: Das gesamte Universum, vom Erhabensten (‚dem Menschen‘) bis zum Einfachsten (Ameisen, Kieselsteine und Regentropfen), wurde als die Schöpfung von etwas noch Erhabeneren, Gott, verstanden. Dieser allmächtige und allwissende intelligente Schöpfer war dabei dem zweithöchsten Wesen auffallend ähnlich. Diese Vorstellung kann man als die *Trickle-Down*-Theorie der Schöpfung bezeichnen. Darwin ersetzt diese durch eine *Bubble-up*-Schöpfungstheorie.¹ Einer der zeitgenössischen Kritiker Darwins, Robert Beverly MacKenzie, brachte die Skepsis gegenüber einer *Bubble-up* Theorie klar auf den Punkt:

1 Die Begriffe ‚*Trickle down*‘ und ‚*Bubble up*‘ sind auch im Englischen keine geläufigen Bezeichnungen, man kann sie jedoch gut verstehen, wenn man die gängigere Bezeichnung ‚*top down*‘ und ‚*bottom up*‘ mitdenkt.

In Darwins Theorie nimmt die absolute Unwissenheit die Rolle des Schöpfers ein. Das heißt, man kann die Grundannahme der ganzen Theorie wie folgt festhalten: FÜR DAS ENTSTEHEN EINER PERFEKTEN UND SCHÖNEN MASCHINE IST KEIN WISSEN DARÜBER NOTWENDIG, WIE SIE ENTSTEHT. Genau diese Behauptung, die man nach sorgfältiger Prüfung findet, drückt in verdichteter Form die wesentliche Grundidee aus. Sie bringt die Quintessenz von Darwins Theorie auf den Punkt, der aufgrund einer seltsamen Umkehrung der Argumentation zu glauben scheint, dass absolute Unwissenheit vollends qualifiziert sei, den Platz der absoluten Weisheit in Bezug auf alle Errungenschaften der Schöpfung einzunehmen.

(Übersetzung von MacKenzie 1868)

Dies ist in der Tat eine seltsame Umkehrung der Argumentation (*strange inversion of reasoning*). Bis heute beschäftigt viele Menschen die beunruhigende Vorstellung, dass sich ein planloser, geistloser Prozess durch die Zeitalter werkelt und sehr viel subtilere, effizientere und komplexere Organismen generiert, ohne die geringste Ahnung von dem zu haben, was er tut.

Auch Turings Idee basiert auf einer bemerkenswert ähnlichen seltsamen Umkehrung der Argumentation. Vor Turing – als noch Menschen die Rechnenden (also Computer) waren – war es unabdingbar, dass diese Menschen etwas von Mathematik verstehen, damit sie ihre Arbeit verrichten konnten. Turing zeigt jedoch, dass dieses Verstehen selbst gar nicht notwendig ist: Man kann nämlich die einzelnen Berechnungsschritte, die von Menschen vollzogen werden, auf pure mechanische Aktionen herunterbrechen, deren Ausführung nicht den Hauch von Verständnis fordert. EINE PERFEKTE UND SCHÖNE RECHENMASCHINE MUSS NICHT WISSEN, WAS ARITHMETIK IST.

Sowohl Darwin als auch Turing haben auf ihre jeweils eigene Art entdeckt, dass Beherrschen ohne Begreifen (*competence without comprehension*) möglich ist.² Dies kehrt die zunächst zutiefst plausible Annahme um, nach der Verstehen die Quelle aller fortgeschrittenen Fähigkeiten sei. Worin sonst ist unsere Motivation, Kinder in die Schule zu schicken, begründet? Warum missbilligen wir die altmodischen Methoden des Auswendiglernens? Wir gehen davon aus, dass die Entwicklung der Fähigkeiten unserer Kinder auf wachsendem Verstehen beruht. Der Slogan ‚Kompetenz setzt Verstehen voraus‘ kann als ein Grundsatz der modernen Pädagogik gelten. Für uns als Homo Sapiens ist

2 Vgl. Dennett (2009): Hier wurde überarbeitetes Material für die vorhergehenden Absätzen entnommen.

das in der Regel auch die richtige Herangehensweise, wenn es um das Erlernen von Fähigkeiten geht. Ich vermute, dass dies weitverbreitete Prinzip aus der Pädagogik einer der entscheidenden Gründe für die Skepsis gegenüber der Evolutionstheorie als auch gegenüber der artverwandten Idee von Turing, der Künstlichen Intelligenz, ist. Schon der Gedanke, dass aus geistloser Mechanik so etwas wie menschliche - oder gar göttliche - Fähigkeiten hervorgehen könnte, empfinden viele als philisterhaft und abstoßend und werten es als eine Beleidigung unseres Verstandes und des Verstandes Gottes.

Führt man sich vor Augen, wie Turing in seiner Beweisführung vorgegangen ist, sieht man, dass er sozusagen menschliche Computer zum Vorbild genommen hat. Diese Menschen saßen an ihren Schreibtischen und vollzogen zuverlässig viele einfache Schritte nacheinander. Sie überprüften ihre Arbeitsschritte, notierten Zwischenergebnisse, statt sich auf ihr Gedächtnis zu verlassen, und konsultierten bei Bedarf ihre Formeln. So machten sie aus einer anfänglich beängstigend scheinenden Aufgabe eine Routine, die sie fast im Schlaf erledigen konnten. Turing zerlegt diese einfachen Schritte systematisch in noch einfachere Schritte, die das Verstehen der einzelnen Schritte obsolet macht. Wenn ein Mensch zum Beispiel Schwierigkeiten hat, die Zahl 9999999999999 von der Zahl 9999999999999 zu unterscheiden, dann kann dieses Diskriminierungsproblem durch das Einführen von anspruchloseren Zwischenschritten in mehrere einfachere Probleme zerlegt werden. Auf diese Weise erstellt Turing ein Inventar von einfachen, elementaren Bausteinen, aus denen man wiederum einen universellen Algorithmus zusammensetzen kann, der in der Lage ist, jeden anderen Algorithmus auszuführen. Turing zeigt, inwiefern so ein Algorithmus einen Menschen oder einen Computer in die Lage versetzen kann, jede beliebige Funktion zu berechnen, und führt dazu aus:

Das Verhalten des Rechnenden wird zu jedem Zeitpunkt durch die wahrgenommenen Symbole und durch seinen momentanen ‚Geisteszustand‘ bestimmt. Wir gehen davon aus, daß es eine Grenze G für die Anzahl der Symbole gibt, die der Rechnende in einem Augenblick erfassen kann. Will er mehr erfassen, muß er sich einer Reihe aufeinanderfolgender Wahrnehmungsakte bedienen. [...] Die tatsächlich durchgeführte Operation wird [...] durch den Geisteszustand des Rechnenden und durch die wahrgenommenen Symbole bestimmt. Insbesondere bestimmen sie den Geisteszustand des Rechnenden, nachdem die Operation ausgeführt worden ist.

(Turing 2019 [1936], 41)

Etwas später stellt er fest:

Wir können nun eine Maschine konstruieren, die die Arbeit dieses Rechnenden tut.

(Turing 2019 [1936], 43)

An genau dieser Stelle zeigt sich, dass alle möglichen Berechnungen auf geistlose Prozesse reduziert werden können. Man kann mit den einfachen Bausteinen beginnen, die Turing isoliert hat, und daraus dann Schritt für Schritt anspruchsvollere Berechnungen aufbauen. Dadurch kann man nach und nach die Intelligenz wiederherstellen, die Turing so geschickt aus den Praktiken der menschlichen Computer herausgefiltert hat.

Aber wie steht es um die Rolle des Genies von Turing und später auch von anderen Programmierern, deren eigenes intelligentes Verstehen ganz offensichtlich die Voraussetzung dafür ist, dass durch die Verknüpfung von Turings geistlosen Bausteinen überhaupt nützliche Fähigkeiten entstehen können? Legt eine solche Abhängigkeit nicht doch wieder eine *Trickle-Down-Perspektive* auf Intelligenz nahe, nur diesmal mit Turing in der Rolle Gottes? Kein geringerer Denker als Roger Penrose hat sich skeptisch über die Möglichkeit geäußert, dass künstliche Intelligenz lediglich aus geistlosen algorithmischen Prozessen erwachsen könnte:

[I]ch glaube [...] sehr stark an die Macht der natürlichen Selektion. Aber ich sehe nicht ein, wie die natürliche Selektion von allein diejenigen Algorithmen zu entwickeln vermag, mit denen wir anscheinend bewußte Urteile über die Gültigkeit anderer Algorithmen fällen können.

(Penrose 1991 [1989], 403)

Er fährt fort und meint:

Von meiner Warte aus birgt die Evolution mit ihrem augenscheinlichen ‚Tasten‘ nach einem künftigen Zweck nach wie vor ein Geheimnis. Zumindest scheinen die Dinge sich etwas besser zu organisieren, als sie es – bloß aufgrund einer Evolution durch blinden Zufall und natürliche Selektion – eigentlich ‚sollten‘.

(Penrose 1991 [1989], 405)

Selbst wenn sich eine einzige Kaskade von natürlichen Selektionsprozessen über Milliarden von Jahren erstreckt, scheint es in der Tat unwahrscheinlich, dass dadurch eine Reihe von Nullen und Einsen erzeugt

werden könnte, die dann als Input für einen digitalen Computer einen Algorithmus für bewusste Urteile darstellt. Aber Turing hat erkannt, dass nichts den Prozess der Evolution daran hindert, sich auf zahlreichen Ebenen so zu kopieren, dass sich selbst Wahrnehmungsfähigkeit und Urteilsvermögen entwickeln. Man kann den rekursiven Ansatz, einen Computer zu entwerfen, der jeden anderen Computer imitieren kann, immer weiter bis hin zu dem Punkt wiederholen, an dem es für bestimmte Computern möglich ist, sich selbst neu zu entwerfen, und ihre Fähigkeiten so zu erweitern, dass sie selbst ihren ursprünglichen Programmierer weit hinter sich lassen. Bereits in dem klassischen Aufsatz ‚*Computing Machinery and Intelligence*‘ von 1950 (auf deutsch: *Kann eine Maschine denken?*) hat Turing erkannt, dass in der Vorstellung eines lernfähigen (nicht-menschlichen) Computers kein Widerspruch steckt:

Die Vorstellung einer lernenden Maschine mag einigen Lesern paradox vorkommen. Wie können sich die Operationsregeln der Maschine ändern? Sie sollten, unabhängig von der Vergangenheit der Maschine und unabhängig von ihren möglichen Zustandsänderungen, vollständig beschreiben, wie die Maschine reagieren wird. Insofern sind die Regeln absolut zeit-invariant. Das ist durchaus richtig. Die Erklärung dieses Paradoxons besteht darin, daß die Regeln, die beim Lernprozess geändert werden, ziemlich anspruchslos und nur von vorübergehender Gültigkeit sind. Der Leser möge dies mit der Verfassung der Vereinigten Staaten vergleichen.

(Turing 1967 [1950], 136)³

Für Turing ist klar, dass die enorme Vielseitigkeit und Variabilität menschlichen Denkens, die sich z.B. im Lernen, in der Fähigkeit zur Re-Evaluation, in der Sprache und im Problemlösen zeigt, im Prinzip aus einzelnen Bausteinen konstruierbar ist. Genau dies kann man als eine *Bubble-Up*-Theorie des Geistes bezeichnen, die den verschiedenen *Trickle-Down*-Theorien des Geistes von Denkern wie René Descartes bis John Searle (und bekannter Weise auch Kurt Gödel, dessen Beweis gleichwohl eine Inspiration für Turlngs Arbeit war) entgegensteht. Letztere beginnen mit dem menschlichen Bewusstsein in seiner hoch-reflektierenden Form und sind aber dann nicht in der Lage, derart magische Kräfte mit den bloßen Mechanismen menschlicher Körper und Gehirne in Zusammenhang zu bringen.

3 Vgl. Suber (2001) bezüglich einer wertvollen Diskussion dieser Passage und des sogenannten Paradoxes der Selbstveränderung.

Genauso wie Darwin hat Turing das Geheimnis der Intelligenz (oder des intelligenten Designs) in atomistische Schritte blinder Zufälligkeiten zerlegt, die sich, wenn sie millionenfach akkumuliert werden, zu einer Art Pseudo-Intelligenz summieren. Die Zentraleinheit eines Computers weiß nicht wirklich, was Arithmetik ist, noch versteht sie, was Addition ist, aber sie ‚versteht‘ den ‚Befehl‘, zwei Zahlen zu addieren und die Summe in ein Register einzutragen, so dass sie, wenn sie zum Addieren aufgefordert wird, zuverlässig addiert und die Summe an die richtige Stelle setzt. Man kann hier von *einer Art von* Verständnis für Addition sprechen. Ein paar Ebenen höher versteht das Betriebssystem nicht *wirklich*, dass es Übertragungsfehler prüft und behebt, aber es verfügt über *eine Art von* Verstehen, das sich darin zeigt, dass es diese Aufgabe bei Bedarf zuverlässig ausführen kann. Wiederum einige Ebenen höher, wo Milliarden und Billionen Bausteine kombiniert sind, versteht ein Schachprogramm zwar nicht *wirklich*, dass seine Königin in Gefahr ist, aber es hat *eine Art von* Verständnis für diese Situation. Und auch IBMs Watson hat *eine Art von* Verständnis von den Jeopardy-Fragen, die es beantwortet.

Warum sollte man von ‚*einer Art von*‘ Verständnis sprechen? Ich halte dies für sinnvoll, weil man bei der Analyse – oder Synthese – von immer kompetenteren Ebenen einer Maschine zwei Dinge unterscheiden muss, nämlich was sie ist und was sie tut. Das ‚was sie ist‘ lässt sich in Bezug auf die strukturelle Organisation ihrer Bestandteile beschreiben, vorausgesetzt, die Teile funktionieren so, wie sie sollen. Bei dem ‚was sie tut‘ handelt es sich um eine (kognitive) Funktion, die (sozusagen) ausgeführt wird. Die Art und Weise der Ausführung ist dabei zuverlässig genug, so dass man auf der nächsthöheren Ebene die Annahme treffen kann, dass man einen intelligenteren Baustein in seinem Inventar zur Verfügung hat, der eben diese Funktion - in einer *Art von* gut genug - erfüllt. Dies ist der Schlüssel, mit dem man die verwirrend komplexe Frage, wie es möglich ist, dass ein Geist allein aus materiellen Mechanismen zusammengesetzt ist, angehen kann.

Was man in der Kognitionswissenschaft den ‚*Art von*‘-Operator nennen könnte, ist vergleichbar mit Darwins Gradualismus bezüglich evolutionärer Prozesse. Das heißt, bevor es Bakterien gab, gab es *eine Art von* Bakterien, bevor es Säugetiere gab, gab es *eine Art von* Säugetieren und bevor es Hunde gab, gab es *eine Art von* Hunden und so weiter. Genauso wie wir Darwins Gradualismus brauchen, um den großen Unterschied zwischen einem Affen und einem Apfel zu erklären, brauchen wir Turings Gradualismus, um den großen Unterschied zwischen einem humanoiden Roboter und einem Taschenrechner

zu erklären. Sowohl Affen als auch Äpfel bestehen aus den gleichen Basiselementen, die sich aber in ihrer strukturellen Anordnung so unterscheiden, dass sie auf vielen Ebenen ganz unterschiedliche Funktionen haben. Dabei gibt es keine strikte Trennlinie zwischen *einer Art von Affe* und einem Affen. Genauso bestehen humanoide Roboter und Taschenrechner aus denselben Grundbausteinen, den nicht-denkenden und nicht-fühlenden Turing-Bausteinen. Aber indem man diese Bausteine zu größeren und leistungsfähigeren Strukturen zusammensetzt, die dann wiederum auf noch höheren Ebenen zu Elementen von noch leistungsfähigeren Strukturen werden, entstehen schließlich Elemente, die eine Art von Intelligenz haben. Diese Elemente können wiederum so zusammengesetzt werden, dass eine Fähigkeit entsteht, von der wir berechtigter Weise sagen können, dass sie von Verständnis begleitet ist.

Wir nehmen den intentionalen Standpunkt (*intentional stance*) ein (Dennett 1971, 1987), um Überzeugungen und Wünsche (oder ‚Überzeugungen‘ und ‚Wünsche‘ oder *eine Art von Überzeugungen* und *eine Art von Wünschen*) aller möglichen irgendwie rational handelnden Wesen auf unterschiedlichen Ebenen zu erkennen. Dies schließt sowohl einfachste Bakterien als auch Gehirne von Lebewesen – vom Seestern bis zum Astronomen – ein, deren Schaltkreise Unterscheidungen treffen, Signale senden, Vergleiche vornehmen und über Erinnerung verfügen. Es gibt keine strikte Linie, ab der das wahre Verständnis beginnt – auch nicht in unserem eigenen Fall, dem Menschen. Ein kleines Kind hat nur *eine Art von Verständnis*, von seinem eigenen Satz ‚Papa ist Arzt‘, und ich habe nur *eine Art von Verständnis* für die Formel ‚ $E = mc^2$ ‘. Einige Philosophen widersetzen sich diesem Anti-Essentialismus: Sie behaupten, dass man entweder glaubt, dass Schnee weiß ist oder nicht und dass man entweder bewusst ist oder nicht. Nichts kann hier als eine Annäherung an ein mentales Phänomen gelten. Stattdessen wird von einer ‚Alles oder Nichts‘ Frage ausgegangen. Die Fähigkeiten des Geistes stellen für solche Denker ein unergründliches Mysterium dar, da sie ‚perfekt‘ und damit völlig anders als alles sind, was in rein materiellen Mechanismen zu finden ist.

Bisher haben Roboter noch kein ‚echtes‘ Verständnis, aber wir kommen der Sache näher. Das ist zumindest die Überzeugung derjenigen von uns, die sich von Turings Einsichten haben inspirieren lassen. Die *Trickle-Down*-Theoretiker dagegen sind felsenfest davon überzeugt, dass uns keine noch so großen Fortschritte in der Forschung jemals zu ‚echtem‘ Verstehen führen werden. Sie gehen davon aus, dass aus Turings Bausteinen kein cartesianisches *Res Cogitans*, kein denkendes Ding konstruiert werden kann. Genauso bezweifeln Kreationisten, dass die von

Darwin beschriebenen Prozesse der Mutation, Replikation und Selektion jemals (echtes) Leben erzeugen können. Damit liegen sie falsch, aber man kann das Unbehagen, das ihre Einstellung motiviert, nachvollziehen.

Sowohl Turings, als auch Darwins seltsame Umkehrung der Argumentation widersprechen jahrtausendalten Denktraditionen. Orientiert man sich an der Geschichte des Widerstands gegen das darwinistische Denken, kann man damit rechnen, dass es auch in ferner Zukunft, wenn ‚bloÙe Maschinen‘ all die Errungenschaften des menschlichen Denkens erreicht oder gar übertroffen haben, noch Denker geben wird, die darauf bestehen, dass der menschliche Geist aber auf eine mysteriöse Weise arbeitet, die keine Wissenschaft je verstehen kann.

(Übersetzung von Anna Strasser und Ralf Stapelfeldt)

Literatur

- Darwin, Charles (1963): *Die Entstehung der Arten durch natürliche Zuchtwahl*. Übersetzung von Carl Neumann. Reclam, Stuttgart 1963. Titel der englischen Originalausgabe: *On the Origin of Species*. John Murray, London 1859.
- Dennett, Daniel (1971): Intentional systems. In: *The Journal of Philosophy* 68, 87-106.
- Dennett, Daniel (1987): *The Intentional Stance*. MIT Press, Cambridge, MA 1987.
- Dennett, Daniel (2009): Darwin's Strange inversion of reasoning. In: *Proceedings of the National Academy of Sciences*, 106, 10061-10065.
- MacKenzie, Robert (1868): *The Darwinian Theory of the Transmutation of Species Examined*. Nisbet & Co, London 1868.
- Penrose, Roger (1991): *Computerdenken. Des Kaisers neue Kleider oder Die Debatte um Künstliche Intelligenz, Bewußtsein und die Gesetze der Physik*. Übersetzung von Michael Springer. Spektrum der Wissenschaft, Heidelberg 1991. Titel der englischen Originalausgabe: *The Emperor's New Mind*. Oxford University Press, Oxford 1989.
- Suber, Peter (unpublished): Saving Machines From Themselves: The Ethics of Deep Self-Modification, preprint, 30 November 2001. <http://legacy.earlham.edu/~peters/writing/selfmod.htm>, abgerufen am 7.10.2020.
- Turing, Alan (2019): Über berechenbare Zahlen mit einer Anwendung auf das Entscheidungsproblem. In: Ziemann, Andreas (Hg.): *Grundlagentexte der Medienkultur. Ein Reader*. Wiesbaden: Springer 2019. Titel der englischen Originalausgabe: Turing, Alan (1936). On com-

putable numbers, with an application to the Entscheidungsproblem. In: *Proc. Lond. Math. Soc.* 42, 230-265, and erratum (1937) 43, 544-546.

Turing, Alan (1967): *Kann eine Maschine denken?* In: Enzensberger, Magnus (Hg.): *Kursbuch 8*, Band 1. Übersetzung von Gänßler, Peter. Zweitausendeins, Frankfurt, 1967, 106-138. Titel der englischen Originalausgabe: Turing, Alan (1950). Computing machinery and intelligence. In: *Mind*, LIX(2236), 433-460.

Watson, James; Crick, Francis (1953): A structure for deoxyribose nucleic acid. In: *Nature*, 171, 737-738.

Nadine Schumann, Yaoli Du

GRENZGÄNGE: VON MENSCHEN ZU SMARTEN MASCHINEN - UND ZURÜCK?

Abstract: Zentrale Begriffe in der heutigen KI-Debatte, wie z.B. Intelligenz, Information, Code, Funktion und Selbstorganisation, werden nicht selten im interdisziplinären Diskurs unterschiedlich definiert und verwendet. In ihrer historischen Entwicklung kurz dargestellt, werden wir Gemeinsamkeiten und Unterschiede ermitteln. Diese Offenlegung der Bedeutungsumfänge ermöglicht Grenzgänge von Menschen zu smarten Maschinen. Jenseits des Mythos der guten bzw. bösen KI als „Gespenst in der Maschine“ muss man fragen, was wir überhaupt von maschineller Intelligenz erwarten, oder vielmehr, was wir überhaupt erwarten dürfen. Diese Erwartungshaltung fordert eine generelle Betrachtung der Entwicklung der Technik und des Menschen ein. Sowohl der Intelligenz- als auch der Technikbegriff sind auf Sichtweisen reduziert, welche sich auf funktionale Abläufe konzentrieren und Menschen auf Maschinen verkürzen. Versucht man dagegen den Menschen in seiner intersubjektiven Konstitution zu begreifen, dann könnte das technische Mittel des ‚Als ob‘ in seiner Prozessualität Möglichkeiten zu Verfügung stellen, welche die Handlungsspielräume des Menschen in seiner sozialen Praxis erweitern können.

Einleitung

Ziel unserer Untersuchung ist eine philosophisch-anthropologische Betrachtung der Beziehung von Mensch und Maschine. Die verschiedenen Konzepte in der interdisziplinären Debatte um Künstliche Intelligenz sollen in Bezug auf das basale Verhältnis von Mensch und Technik hin untersucht werden. Um dieses Verhältnis näher zu beleuchten, werden wir uns vordergründig der Klärung der zentralen Begriffe zuwenden, welche sowohl in der Biologie, Neuro- und Kognitionswissenschaft, als auch in der Informatik, der Informationswissenschaft und Mathematik fächerübergreifend benutzt werden. Die verschiedenen Definitionen und Verwendungsweisen von Begriffen wie z.B. Maschine, Automat, Intelligenz und Information in den diversen Einzelwissenschaften werden in ihrer historischen Entwicklung kurz dargestellt und deren Gemeinsamkeiten und Unterschiede ermittelt. Schließlich können so interdisziplinäre Verständnisschwierigkeiten sichtbar, die Verschiedenheit der Konzeptionen deutlich und dadurch

Grenzbestimmungen von Mensch und Maschine möglich werden. Schließlich führen diese Grenzgänge hin zu einer ‚*Als ob*‘-Einsicht, die Menschen im Umgang mit Technik pragmatisch zu fassen sucht.

1. Natürlich vs. Künstlich

Ist von Künstlicher Intelligenz die Rede, kann man als erstes fragen, welchen Status eine Künstlichkeit gegenüber einer Natürlichkeit überhaupt einnehmen kann. Künstliche Artefakte werden meist in Gegensatz zu natürlichen Gegenständen verstanden (vgl. zu dieser Unterscheidung den Beitrag von Hans-Jörg Kreowski und Wolfgang Krieger in diesem Band). Schon in der griechischen Antike spielt die Vorstellung der Verschiedenheit von natürlichem und künstlichem Seienden eine zentrale Rolle. Aristoteles unterscheidet natürliche Produkte (*physei onta*) von hergestellten Produkten (*techne onta*). „Man kann die Gesamtheit des Seienden (in zwei Klassen) einteilen: in die Produkte der Natur und in die Produkte andersgearteter Gründe.“ (Aristoteles, Physik, Buch I, 192b8-10). Organismen wie Pflanzen, Tiere und Natur-elemente wie Erde, Feuer, Luft und Wasser haben als Naturprodukte „ein Prinzip seiner Prozessualität und Beharrung in ihm selbst, ein Prinzip teils seiner Ortsbewegung, teils seines Wachsens und Abnehmens, teils seiner qualitativen Veränderung.“ Hergestellte Produkte haben „keinerlei in ihm selbst liegender Tendenz zu irgendwelcher Veränderung seiner selbst.“ Bei natürlichen Produkten liegt das Bewegungsprinzip in den Naturdingen selbst, während ein hergestelltes Produkt kein Prinzip der Bewegung in sich selbst hat. Dieser Unterschied wird am Beispiel eines Bettes verdeutlicht: „das Bettsein sei eine bloß zusätzliche Bestimmtheit, das Ergebnis der Willkür menschlicher Arbeit“, während das Holzsein das Wesentliche ist.

Aristoteles fasst die Gesamtheit des Seienden in seiner dynamischen Bewegung prinzipiell teleologisch (Umsetzungen der Zweckbestimmung). Der belebte Körper, also das Lebewesen ist in der Materie nur der Möglichkeit (*dynamis*) nach angelegt. Aktualität und Wirklichkeit (*energeia*) gewinnt es lediglich durch die Form. Und diese Entfaltung bzw. Entwicklung (Vollendung) des Wesens nennt Aristoteles *Entelechie*. Diese Entwicklung setzt ein Ziel oder Telos voraus, von einer möglichen hin zu einer wirklichen Entfaltung der Substanz (*ousia*).

Während also das erwähnte Bewegungsprinzip bei Aristoteles in den Naturdingen selbst liegt, hat ein hergestelltes Produkt kein Prinzip der Bewegung in sich selbst. Der Zweck der hergestellten Produkte aber hängt den natürlichen Prinzipien des Menschen an. Ein Bett ist zwar

aus Holz, das Holz kann aber potentiell durch die Arbeit eines Baumeisters nicht nur Bett, sondern auch Tisch, Stuhl oder Schiff werden. Herstellungsprozesse werden als technische Prozesse natürlichen Prozessen nicht diametral gegenübergestellt, sondern gehören zur wirklichen Entfaltung des Menschen dazu. Der Mensch wird von Aristoteles als *Technit* bestimmt, der schon mit seiner Hand als dem „Werkzeug aller Werkzeuge“ von Natur her ausgestattet ist und in seiner Eigenbewegung gar nicht anders kann, als aktiv zu gestalten (vgl. Die Teile der Lebewesen [686a21]).

Die aristotelische Auffassung und Methodik von Wissenschaft sind hier ausschlaggebend. Aristoteles hatte Induktion und Deduktion verschränkt und beginnt seine Überlegungen stets mit den sinnlichen Gegebenheiten. Letztendlich zielt er aber mit der Induktion auf die allgemeinen Prinzipien, aus denen sich die einzelnen Phänomene ableiten lassen. Gleichzeitig stellen aber die sinnlichen Gegebenheiten auch die realen Ursachen der Naturdinge dar (vgl. Fuchs 1992, S. 44).

Betrachtet man das allgemeine Verhältnis von Mensch und Technik, dann verknüpft die Technik (als hergestelltes Produkt, Herstellungsprozess) die Naturbestimmung des Menschen (als Naturprodukt) mit der Umsetzung der Zweckbestimmung von Menschen (*Entelechia*). Der Technikbegriff ist eben nicht nur auf die hergestellten Produkte oder Werkzeuge zu beziehen, sondern bedarf auch einer sozio-technischen Einbettung, welche die verschiedenen Formen menschlicher Praxis berücksichtigt und damit die Naturerkenntnis und die Erkenntnis des Menschen selbst in den Vordergrund stellt.

In der Richtlinie 3780 des Vereins Deutscher Ingenieure (VDI) wird der Begriff Technik folgendermaßen definiert: Technik umfasst „die Menge der nutzenorientierten, künstlichen, gegenständlichen (Artefakte oder Sachsysteme); die Menge menschlicher Handlungen und Einrichtungen, in denen Sachsysteme entstehen; die Menge menschlicher Handlungen, in denen Sachsysteme verwendet werden.“ In dieser Definition wird Technik völlig richtig nicht nur in Bezug zu den gegenständlichen Sachsystemen verwendet, sondern auch im Rahmen der Technikbewertung auf „die Bedingungen und Folgen ihrer Entstehung und Verwendung“ eingegangen.

2. Maschinen im Spiegelbild des Menschen

Technik ist heute der große Sammelbegriff für alle möglichen Produkte und Herstellungsprozesse. Dabei fällt die Unterscheidung verschiedener Technikphänomene nicht leicht, besonders wenn es um Differen-

zierung der einzelnen Maschinentypen untereinander und mehr noch um die Frage der Abgrenzung gegenüber belebten Organismen geht. Ob einerseits von mechanischen und elektrischen Maschinen oder andererseits von elektronischen, sich selbstregulierenden Automaten die Rede ist, ob einfache Werkzeuge oder komplexe mathematische Modelle verhandelt werden, zeigt schon einen weiten Bedeutungsumfang des Maschinenbegriffs an, der sich historisch wandelt und in seiner Vielfältigkeit höchst ambivalent daherkommt (Heßler 2020, S. 256).

2.1 Maschine

Der Ausdruck Maschine geht in der griechischen Antike auf die dorische Variante *machana* des attischen Wortes *mechane* zurück und wird mit „Kunstgriff, Wirkungsweise, Arbeitsmittel oder Werkzeug“ übersetzt. Hier ist der Zusammenhang von Maschine und Mechanik schon angezeigt. Im Lateinischen wird später der Ausdruck *machina* im Sinne von Kriegs- und Belagerungstechnik verwendet. Heute wird der Ausdruck Maschine synonym als Oberbegriff der Werkzeuge im Sinne von Produkt verwendet.

Einfache Maschinen im vorindustriellen Sinne sind gleichsam Werkzeuge, wie z.B. ein Hammer. Sie bewegen sich nicht eigenständig, sie müssen bewegt werden. Maschinen, die aus verschiedenen Komponenten zusammengesetzt sind, bewegen sich zunächst auch nicht von allein. Die Entwicklung dieser mechanischen Maschinen profitiert von den in der Renaissance sich herauskristallisierenden Konzepten über Mechanismen, d.h. genaue Kenntnis spezifischer Abläufe von Bauteilkomplexen.

2.2 Automat

Wenn eine Maschine einen kontinuierlichen Energieantrieb hat und damit eigenständig laufen kann, wird diese als Automat bezeichnet. Allerdings ist auch der Automatenbegriff selbst höchst ambivalent.

In der Spätantike erscheint das Werk *Automata* von Heron von Alexandria, in welchem die damalige Vorstellung von Automaten und ihrer Arbeitsweise ausgeführt ist. Es handelt von Tempeltüren und anderen technischen Vorrichtungen, die sich automatisch öffnen und als „von Geisterhand“ bewegt verstanden wurden. Das altgriechische *automatos* bezeichnet „Dinge, die sich von selbst bewegen“, im Lateinischen bedeutet *automatus* jedoch „aus eigenem Antrieb handeln, freiwillig“; nur das substantivierte *automatum* wird als „Maschine, die sich selbst bewegt“ ähnlich übersetzt (Pfeifer 1993).

Die antike Automatentechnik wurde klar zu den hergestellten Produkten gezählt und gerade nicht in die Kategorie natürlicher Produkte eingeordnet. Die Eigenbewegung wurde vorgetäuscht und so zu Trick und Gaukelei. Antike Baumeister von Automaten waren vordergründig nicht bestrebt, Menschen oder die Natur nachzubauen (Wittig 1997, S. 25).

Im abendländischen Mittelalter wurde die Vorstellung von Natur in den christlichen Monotheismus eingegliedert und als bloße Natur in die Naturphilosophie verlagert. Dies ermöglichte eine Annäherung von Mensch und Natur. Die astrologisch-astronomische Vorstellung der quasi-mechanischen Ordnung und regelmäßigen Bewegung des Makrokosmos wurde auf den Mikrokosmos Mensch in analoger Weise angewendet (Wittig 1997, S. 26-31). Vor diesem Hintergrund war die Produktion von Unterhaltungsautomaten eher zweitrangig. Nun ging es darum, Automaten zu bauen, die natürliche Produkte nachahmen, damit sie für Menschen von Nutzen sein können.

Mit Aufkommen der mathematisch-mechanistischen Sichtweise auf die Natur im Rahmen der kopernikanischen Wende in der frühen Neuzeit bildet sich mit der Zentralisierung bzw. der Vereinheitlichung von Theorie und Methode ein neuer Wissenschaftsbegriff heraus, welcher auf das quantitativ Messbare gerichtet ist und eine Einheits-theorie anstrebt. Die Verbindung von Experimenten im methodisch-naturwissenschaftlichen Sinne des Francis Bacon (1561-1626) mit der mechanischen Physik des Galileo Galilei (1564-1642) führte in der neuzeitlichen Entwicklung hin zu dem herrschenden methodisch-naturwissenschaftlichen Forschungsstil. Dadurch etabliert sich ein neues Naturverständnis, welches den aristotelischen Unterschied von Natur (*physei onta*) und Technik (*techne onta*) endgültig auflöst. Die Natur selbst wird mechanisch gedeutet. Unter diesem Forschungsstil wurden nicht nur hergestellte Werkzeuge, sondern auch lebendige Körper bis hin zur gesellschaftlichen Ordnung nach mechanistischem Vorbild betrachtet. Das Verhältnis von individuellen Menschen, der Gesellschaft und der Natur wird mit dem Verhältnis der einzelnen Komponenten in Bezug auf die Aufbau- und Ablauforganisation der gesamten Maschine gleichgesetzt.

In den Schriften von René Descartes (1596-1650) finden sich erste maschinen-theoretische Interpretationen von Organismen (vgl. Descartes 1632; 1637). Der Philosoph und Mathematiker Descartes beschäftigte sich ausführlich mit der Anatomie des Herzens und legte mit seiner Theorie der Körpermaschine die wesentlichen Grundlagen für physiologische Beschreibungen des Herz-Kreislauf-Systems (Descartes 1637, S. 78f.) Er konzentrierte sich auf den mechanischen Aspekt des Kreis-

laufs und entwarf eine mechanistische Sichtweise des Organismus, die bis 1700 für Physiologen eine Orientierungshilfe bot. Die cartesische Konzipierung der Körpermaschine machte schließlich „eine vom seelischen Erleben prinzipiell unabhängige und damit im neuzeitlichen Sinn naturwissenschaftliche Physiologie“ erst möglich (Fuchs 1992, S. 4).

In damaliger Zeit wird nicht nur gefragt, inwieweit Lebendiges im mechanistischen Bezugsrahmen konstruiert werden kann, sondern auch, ob komplexe Maschinen und Automaten als künstlich belebt vorgestellt werden können. So fragt z.B. Thomas Hobbes (1588-1679), „warum sollte man nicht sagen können, dass alle Automaten oder Maschinen, welche wie z.B. die Uhren durch Federn oder durch ein im Innern angebrachtes Räderwerk in Bewegung gesetzt werden, gleichfalls ein künstliches Leben haben?“ (Hobbes 1651, S. 5).

Im Gegensatz zum wissenschaftlichen Vorgehen des Aristoteles, dessen Ausgangspunkt induktiv bei dem sinnlich Gegebenen ansetzt, bedienen sich die Anhänger und Nachfolger von Descartes, wie z.B. Nicolas Malebranche (1638-1751) oder Arnold Geulincx (1624-1669) einer streng deduktiven Methode. Die obersten Prinzipien werden nicht durch Induktion ermittelt, wie im aristotelischen Sinne, sondern werden gesetzt bzw. entspringen dem reinen Denken. Demzufolge werden nur diejenigen Phänomene zugelassen, die sich aus diesem abstrakt-physikalischen Materiebegriff deduzieren lassen. Sowohl tote als auch lebendige Körper stehen unter der Herrschaft absoluter Naturgesetze, welche „die Welt rein mechanischer Wirkungsbeziehungen konstituieren“ (Fuchs 1992, S. 25). Damit werden sowohl mystische als auch teleologische Erklärungen der Natur ausgeschlossen. Daraus folgt schließlich das Maschinenparadigma des Organismus.

Spätestens mit der Formulierung der allgemeinen Naturgesetze durch Isaac Newton (1642-1726) wird der Grundstein für die klassische Mechanik gelegt und als Teilbereich der Physik etabliert (vgl. Newton 1687). Newtons Schwerpunktmechanik wird danach neben späteren physikalischen Disziplinen wie Wärmelehre, Elektrodynamik und Optik usw. zur ersten Naturwissenschaft und fußt entscheidend auf Mathematik (vgl. Kant 1786).

Hinter dieser Physik steht ein realistisches Naturverständnis, welches nicht im Sinne von Stoß und Abstoßung limitiert ist oder die Einbindung der Grundelemente in die kosmische Ordnung thematisieren muss. Die Wirkungsweisen der einzelnen Komponenten untereinander werden vielmehr funktional auf die Ganzheit der Maschine bezogen. In dieser Ansicht werden die einzelnen Funktionen der Teile von der Ganzheit der Maschine und ihrem Funktionieren im Ganzen bestimmt.

So wird physikalische Realität nicht nur in ihrem Aufbau, sondern auch in ihrer Wirkungsweise als dynamische Maschine verstanden, wie meist am Beispiel einer aufgezogenen bzw. aufzuziehenden Uhr verdeutlicht wird. Dabei wird die Frage des Zusammenhanges von Denken, Mathematik und Logik nicht mehr gestellt, sondern die Verbindung von Kosmos und vernünftiger Ordnung in eins gesetzt.

Im späten 17. bzw. frühen 18. Jahrhundert erreicht die damalige Uhrmacherkunst in Bezug auf Präzision ihren Höhepunkt. Diese Erkenntnisse begünstigten nicht nur die Herstellung von Automaten, welche immer kleinteiliger und eindrucksvoller gestaltet wurden, sondern führten zu der weit verbreiteten Vorstellung, die Welt selbst sei ein riesiges Uhrwerk. Unter dieser Betrachtung wird die Welt als geschlossenes System deterministisch interpretiert. Sie lässt sich mit strengen mathematischen Gesetzmäßigkeiten beschreiben. Diese Betrachtung der Welt ist mit der Hoffnung verbunden, die Welt selbst vollständig zu erklären bzw. Ereignisse mit Hilfe logischer Schlüsse und mathematischer Methoden berechenbar und damit vorhersagbar / gestaltbar zu machen. Allerdings bleibt auch bei Newton Gott als Baumeister der Weltuhr bzw. als Ursprung der Kräfte erhalten. Die substanzielle Kraftübertragung wird mit der funktionalen Wirkungsweise analogisiert, kommt aber ohne den göttlichen Ursprung aller Kräfte noch nicht aus.

Erst mit dem Werk *L'homme machine* (1748) des Franzosen Julien Offray de La Mettrie (1709-1751) wird Gott als Urstifter bedeutungslos. Als Vertreter einer vermeintlich stark mechanistisch-materialistischen Position wird der Mensch zur komplexen Maschine. Der Mensch wird hier auf seinen Körper reduziert, ein immaterieller Seelenbegriff abgelehnt: „Der menschliche Körper ist eine Maschine, die selbst ihre Triebfedern aufzieht – ein lebendes Abbild der ewigen Bewegung“ (La Mettrie 1748, S. 35).

Die mechanistische Auffassung von Natur und Mensch im 18. Jahrhundert beflügelte die Vorstellung von organischen Körpern als Automaten, die letztendlich auch gebaut werden könnten. Im Gegensatz zur göttlichen Schöpfung wird nun der Fokus auf die sichtbaren Mechanismen der wahrnehmbaren Welt gelegt. Die mechanistische Interpretation der Organismen eröffnete die Möglichkeit, lebendige Körper in Funktion und Gestalt nachzubauen (vgl. Westermann 2020, S. 249). Mit den kunstvollen Puppenautomaten des 18. Jahrhunderts werden menschliche Äußerungen und Körperbewegungen nachgeahmt und mechanisch nachgebildet. Die Idee des Menschenautomaten findet ihren Niederschlag später auch in der Literatur des frühen 19. Jahrhunderts, wie z.B. in E.T.A. Hoffmanns Roman *Der Sandmann* (1816)

deutlich wird. Diese anthropomorphen Mechanismen sind der zeitgenössischen Vorstellung von Androiden sehr ähnlich. Die Idee, dass das Leben mechanisch nachbildbar ist, oder zumindest die Bewegungen lebendiger Körper nachgebaut werden können, identifiziert den Automatenbegriff mit dem Verständnis mechanischer Körper (Westermann 2020, S. 254).

Dieses Verständnis führte schließlich zur Gleichsetzung von menschlicher und maschineller Körperarbeit. In Zuge der Industrialisierung wurde diese Gleichsetzung mit der Umgestaltung der Produktionsweise weiter verschärft und führte schließlich zu einer Neuordnung sozialer Verhältnisse. Mit der Erfindung von Kraftmaschinen, wie der Dampfmaschine, und der Organisation von Arbeitsprozessen in einer Fabrik, verringerte sich einerseits die Körperarbeit von Menschen, führte aber andererseits auch zu monotoner Arbeitsweise. Zum einen wird dieser Fortschritt optimistisch als Erleichterung und Zugewinn in Bezug auf eine Zukunftsfantasie durch die Integration von Mensch und Maschine gedeutet, zum anderen wird der Mensch in der pessimistischen Interpretation selbst Teil der Maschine, indem er als Komponente der Produktionsmaschinerie einfache Arbeitsschritte am Fließband täglich hundertfach wiederholt ausführen muss, um dem Anspruch der Mechanisierung und Automatisierung der Arbeits- und Produktionsprozesse gerecht zu werden. Menschliche und maschinelle Körperarbeit werden hier eben nicht nur auf eine Stufe gestellt, sondern treten dadurch auch in Konkurrenz zueinander.

Die Hoffnung auf eine gelingende Integration von Mensch und Maschine gepaart mit der Sorge des Menschen, durch Maschinen ersetzt zu werden, findet sich auch in der heutigen KI-Debatte wieder (vgl. den Beitrag von Thomas Weiß in diesem Band). Dies zeigt ein schon vorhandenes Menschenbild an, welches eher hintergründig nicht nur in der heutigen Arbeitspraxis eine Rolle spielt, sondern schon in der Theoriebildung der damaligen Wissenschaften erheblich an Einfluss gewinnt.

Im Zuge der Industrialisierung entwickelten sich an den Universitäten diverse Einzeldisziplinen, die es vorher so an der Akademie nicht gab. Ähnlich wie Fabriken entstanden Labore, in welchen Naturphänomene jeglicher Art nach Vorgabe des methodisch-naturwissenschaftlichen Forschungsstils unternommen wurden. Die Experimentierfreudigkeit in den Wissenschaften erreichte einen ungeahnten Höhepunkt.

Mit dem Aufkommen des Energiebegriffs und der ersten Nutzung von Elektrizität wurde der Begriff Elektrotechnik 1879 von Carl Siemens (1829-1906), dem Erfinder des Elektromotors, geprägt. Er

trieb die Gründung von Lehrstühlen an technischen Hochschulen in Deutschland voran. Die Erkenntnisse der Elektrotechnik erweiterten das Verständnis des Verhältnisses von Energie und Stoffwechsel und führten zu der Einsicht, dass nicht nur zwischen Komponenten, sondern auch in den Komponenten selbst physische Prozesse ablaufen. Ziel war es, komplexe Naturphänomene wie z.B. Stoffwechselprozesse in ihrer dezentralen und dennoch systemischen Dimension theoretisch zu begreifen, um diese schließlich praktisch zu messen und zu steuern. Dabei profitierte die Technikentwicklung maßgeblich von den Entdeckungen und Erkenntnissen der aufstrebenden Einzelwissenschaften und vice versa. Technische Konstrukte konnten sowohl in wissenschaftlichen als auch gesellschaftlichen Kontexten fruchtbar gemacht werden, allerdings ergaben sich so auch die Limitierungen eines einseitig mechanistischen Naturverständnisses. Besonders in Bezug auf lebendige Organismen greift die mathematisch-mechanistische Kausal-erklärung zu kurz. Organismen sollen in ihrer Ganzheit begriffen werden, wozu auch Selbsterhaltung und -regulation gehört und dies kann ohne Zweckbestimmung nicht gedacht werden.

Um der Engstellung des mechanistischen Ansatzes zu entgehen, kristallisierte sich im 20. Jahrhundert ein systemisches Naturverständnis heraus. Dieses Verständnis speiste sich aus den Erkenntnissen der Wissenschaftler verschiedener Fachdisziplinen der damaligen Zeit, welche schließlich in einer übergeordneten Theorie methodisch vereint werden sollten. Der Schweizer Biologe Ludwig von Bertalanffy (1901-1972) suchte nach gemeinsamen Gesetzmäßigkeiten von physischen, biologischen und sozialen Systemen. Im Vergleich der verschiedenen Systeme miteinander, extrahiert Bertalanffy Prinzipien wie Komplexität, Selbstorganisation, Rückkopplung, Gleichgewicht, die in allen Systemen gefunden werden können und wird damit zum Pionier der Allgemeinen Systemtheorie (vgl. Bertalanffy 1949; vgl. Wetzel 1978, S. 76f.).

Dieses systemische Verständnis von Natur wurde sowohl von Naturwissenschaftlern und Mathematikern als auch von Soziologen und Gesellschaftswissenschaftlern angestrebt. Die von Norbert Wiener (1894-1964) begründete Kybernetik (1948) behandelt die Regulierung und Steuerung komplexer mechanischer Systeme; wir werden später noch auf Wieners kybernetischen Ansatz zurückkommen, wenn es um den Informationsbegriff geht. Eine allgemeine Systemtheorie ist auch für Soziologen wie Talcott Parsons (1902-1979) und Niklas Luhmann (1927-1998) interessant (vgl. den Beitrag von Jan Tobias Fuhrmann in diesem Band). Die Termini Selbststeuerung, Selbstorganisation, Funktion und Information wurden gleichsam auf Lebendiges und Tech-

nisches angewendet. Doch schon damals stellte sich die Frage, ob die Anwendung derartiger Begriffe und Annahmen legitim und eine universale, allumfassende Systemtheorie überhaupt möglich ist.

2.3 Mathematische Systeme, Computersimulation, Modellierung

In der Mathematik werden Systeme über Axiome gebildet, d.h. die Axiome legen die Beziehungen zwischen den grundlegenden Begriffen fest. Für den Mathematiker David Hilbert (1862-1943) ist die Beziehung ausschlaggebend, die Grundbegriffe selbst sind irrelevant (Mainzer 2014, S. 59). Allerdings hat der Formalisierungsversuch der axiomatischen Systeme in der Mathematik Ende des 19. Jahrhunderts direkt zur Grundlagenkrise der Mathematik am Anfang des 20. Jahrhunderts geführt (Mainzer 2014, S. 62). Hilbert zeigte im Jahr 1900 grundlegende mathematische Probleme auf und strebte danach, diese zu lösen und Mathematik als formales System sicher zu begründen. Dabei ist die Beweisbarkeit der Widerspruchsfreiheit der arithmetischen Axiome (Hilberts 2. Problem) wohl das bekannteste Problem der neueren Mathematikgeschichte.

Ein Hauptgrund für das Erstreben nach Formalisierung der Mathematik war Folge der Begründung der Mengenlehre von Georg Cantor (1845-1918). Hier wurden Objekte als Elemente durch Mengen beschrieben und klassifiziert. Dadurch entstand aber ein entscheidender logischer Widerspruch, wenn Mengen als Grundlage formaler Sprache dienen sollen. Denn eine Menge, als Abstraktion der Zusammenhänge der Objekte, ist nicht unbedingt widerspruchsfrei definierbar (Hofstadter 1979, S. 22). Ein berühmtes Beispiel ist das von Bertrand Russell (1872-1970) formulierte Russell Paradox, welches später von ihm mit dem Barbier-Paradoxon popularisiert wurde: „Definiert man den Barbier als einen Menschen, der genau diejenigen rasiert, die sich nicht selbst rasieren, dann stellt sich die Frage, ob er sich selbst rasiert oder nicht.“ (Russell 1918, S. 258).

Um die Widersprüche in der Formalisierung der Mengenlehre zu vermeiden, versuchten Bertrand Russell und Alfred North Whitehead (1861-1947) in ihrem gemeinsamen Projektwerk *Principia Mathematica* (1910-1913) Axiome und Ableitungsregeln wohl zu definieren, indem sie logisch-mathematisch klassifizierten und typisierten. Das Ziel ist, alle mathematisch wahren Sätze durch Logik herzuleiten und ohne Widersprüche darzustellen (Hofstadter 1979, S. 24). Allerdings bleibt die Frage von Hilbert, ob also alle wahren Sätze einer mathematischen Theorie in diesem formalen System mit der logischen Methode beweisbar sein können, weiterhin offen. Die Beschreibung eines

mathematischen Systems selbst bleibt also immer unvollständig, insofern die Abstraktionseinheit der Aussagen als Element, Menge oder Klasse sich selbst nicht beschreiben kann (Mainzer 2014, S. 66).

Im Jahr 1931 demonstrierte Kurt Friedrich Gödel (1906-1978) in seiner Arbeit *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme* die Unlösbarkeit des Hilbertschen 2. Problems und verwies auf die Unmöglichkeit des Versuches von Russell und Whitehead (Hofstadter 1979, S. 27). Er beweist, dass es kein widerspruchsfreies axiomatisches System gibt, aus dem alle wahre Sätze und eine konsistente vollständige Selbstbeschreibung abgeleitet werden können. Dadurch wird nicht nur die Grenze formaler Systeme aufgezeigt, sondern auch ein Anschluss an die Systemtheorie und die frühe Computerforschung möglich.

Inspiziert von Gödels Unvollständigkeitsbeweis axiomatischer mathematischer Systeme erstellt Alan Turing (1912-1954) ein praktisches Berechnungsmodell, welches die unmögliche Selbstbeschreibung des Systems anerkennt und dennoch eine Beschreibungsmenge dieser Selbstbeschreibungen maschinell festhaltbar machen soll (vgl. Turing 1936). Gödel hatte ein Codierungsschema entwickelt, in welchem eine symbolische Aussage in natürliche Zahlen codiert werden kann. Eine Aussage wie „Dieser Satz ist falsch.“ und die Aussage über diese Aussage kann in Zahlen formuliert werden. Mit dieser Methode wird der logische Widerspruch in der symbolischen Aussage mit arithmetischer Formalisierung dargestellt (Hofstadter 1979, S. 20), Turing nutzte diese Methode für den Bau seiner Maschine.

Die Turingmaschine bietet eine automatische Lösung, die bei eintreffendem Satz ein Band in der Maschine anhält. Mit einem Lese- und Schreibkopf werden nach programmierten Regeln die Symbole auf dem Speicherband ausgewertet. Damit wird das Berechnungsverfahren repräsentiert und simuliert. Die Turingmaschine ist eben auch eine mathematisch-funktionale Methode, um Allgemeingültigkeit von Ausdrücken feststellen zu können. Mathematische Zusammenhänge werden funktional abgebildet, indem Symbole und Zeichen schrittweise manipuliert und berechnet werden, um schließlich als Algorithmen formalisiert zu sein. Das Problem der logisch-mathematischen Formalisierung, welches mit der Theorie Hilberts aufkam und von Gödel als unvollständig bewiesen wurde, kann allerdings in der Mathematik selbst nicht gelöst werden. Diese Aufgabe wird nun den Programmierern (Softwareengineering) übertragen (Mainzer 2014, S. 89f.).

Die Computertechnologie erweitert zum einen die Möglichkeit der Automatisierung und bestimmt zum anderen die Grundlagen und

Grenzen des Verständnisses axiomatischer Systeme. Damit verändert sich der Bezug auf die Frage nach einer umfassenden Systemtheorie sowohl im Großen als auch im Kleinen. Die axiomatische Grundlage und ihre technisch dargestellte und modulierte Form sind immer ambivalent und dynamisch. Die Turingmaschine ist als ein Berechnungsmodell konzipiert, und zwar auf eine mechanische Art und Weise, ebenso wie ein Mensch mathematische Aufgaben darstellt und löst. Der mathematische Gedankenprozess ist durch eine berechenbare Funktion dargestellt. Damit stellt sich gleichzeitig die Frage, ob nicht auch allgemeine Gedankenprozesse selbst durch Algorithmen dargestellt werden könnten. Die Analogie von Menschen und Maschine wird so vom menschlichen Körper und der physikalischen Maschinerie zur Symbolverarbeitungsfähigkeit und funktionaler Berechnung hin verschoben und die wirkmächtige Computeranalogie geboren. In dieser Sichtweise ist nicht mehr nur die Rechenmaschine funktional-symbolverarbeitend, sondern ihr Designer selbst (bzw. sein Gehirn) war dies schon immer. Unsere gegenwärtigen hochentwickelten technischen Systeme basieren auf Turings Berechnungsmodell, so dass man heute trefflich von einer sogenannte Infosphäre sprechen kann, die einer systematisch codierten, informationellen Umwelt entspricht (vgl. Floridi 2014).

Die Vorstellungen und Entwicklungen von künstlich-intelligenten Systemen reichen von den einfachen klassischen kybernetischen Systemen, wie z.B. ein Thermostat oder ein selbst-regulierender Automat, bis hin zu hochgradig vernetzten Varianten, wie gar der Möglichkeit voll funktionierender *Smart Homes*, *Smart Factories* oder *Smart Cities*. Hier werden verschiedene kybernetische Systeme integriert und miteinander vernetzt, um sowohl die Regulierung zu verbessern, die Effizienz menschlicher Tätigkeit zu steigern, als auch Sicherheit und freie Zeit zu generieren. Künstliche Intelligenz wird dabei meist in Form eines intelligenten Agenten bzw. als intelligentes System wahrgenommen, ohne dabei jedoch auf eine spezifische Definition von Intelligenz zu rekurren. Alltagssprachlich meinen wir technische Systeme, die Problemlösungen für hochkomplexe Problemstellungen automatisch generieren oder personalisierte Handlungs- und Produktionsabläufe effizienter gestalten (Optimierungskalküle, Werbung, *Health Care*). Im Vordergrund steht stets der Anwender und künstlich-intelligente Systeme sollen sich stark am Menschen orientieren.

Die sprachliche Kombination der Künstlichen Intelligenz wird einerseits auf vermeintlich ‚autonome‘ Artefakte bezogen, wie Smartphones, Industrieroboter und sogar selbstfahrende Autos. Andererseits bezieht sich die Rede von KI auf die technischen Verfahren des Informations-