

LMU

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

DISSERTATIONEN DER LMU

UB

63

YANNICK SCHLOTE

Konvergenz und Überwältigung

Die Mythen der Künstlichen Intelligenz aus
theologisch-ethischer Perspektive

OLMS

Konvergenz und Überwältigung

Die Mythen der Künstlichen Intelligenz aus
theologisch-ethischer Perspektive

Inauguraldissertation
zur Erlangung des Doktorgrades der
evangelischen Theologie
an der
Ludwig-Maximilians-Universität München

vorgelegt von
Yannick Schlote
2022

Erstgutachter: Prof. Dr. Reiner Anselm
Zweitgutachter: Prof. Dr. Daria Pezzoli-Olgiati
Tag der mündlichen Prüfung: 20.07.2022

Yannick Schlote

Konvergenz und Überwältigung
Die Mythen der Künstlichen Intelligenz
aus theologisch-ethischer Perspektive

Dissertationen der LMU München

Band 63

Konvergenz und Überwältigung

Die Mythen der Künstlichen Intelligenz
aus theologisch-ethischer Perspektive

von
Yannick Schlote

Eine Publikation in Zusammenarbeit zwischen dem **Georg Olms Verlag** und der **Universitätsbibliothek der LMU München**

Mit **Open Publishing LMU** unterstützt die Universitätsbibliothek der Ludwig-Maximilians-Universität München alle Wissenschaftlerinnen und Wissenschaftler der LMU dabei, ihre Forschungsergebnisse parallel gedruckt und digital zu veröffentlichen.

Georg Olms Verlag AG
Hagentorwall 7
31134 Hildesheim
<https://www.olms.de>

Text © Yannick Schlote 2023

Diese Arbeit ist veröffentlicht unter Creative Commons Licence BY 4.0. (<http://creativecommons.org/licenses/by/4.0/>). Abbildungen unterliegen ggf. eigenen Lizenzen, die jeweils angegeben und gesondert zu berücksichtigen sind.

Erstveröffentlichung 2023
Zugleich Dissertation der LMU München 2022

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet abrufbar über <http://dnb.d-nb.de>

Open-Access-Version dieser Publikation verfügbar unter:
<http://nbn-resolving.de/urn:nbn:de:bvb:19-310220>
<https://doi.org/10.5282/edoc.31022>

ISBN 978-3-487-16350-5

Inhalt

Vorwort	XI
1 Einleitung	1
1.1 Eliza und ELIZA	1
Identität und Täuschung	3
Kontrolle und die Angst vor dem Kontrollverlust	4
Mythos und Logos.....	6
1.2 These, Vorgehen und Ziel der Untersuchung	11
Der grundlegende Mythos von der Ähnlichkeit zwischen Mensch und Maschine	11
Weiterarbeit am Großmythos: Konvergenz und bevorstehender Umbruch.....	13
Die ethische Relevanz der KI-Mythen.....	17
Aufbau der Arbeit.....	21
2 Vorüberlegungen	25
2.1 Betrachterperspektive: Wie wird Technik zu einem Gegenstand konstruktiver theologischer Auseinandersetzung?	25
2.1.1 Technikzu- und abwendung im Protestantismus des 19. und 20. Jahrhunderts	27
2.1.2 Technisches Handeln ist Ausdruck menschlicher Freiheit im Raum der Kultur	32
2.1.3 Ethik fordert die technische Kultur zur Rechenschaft über ihren Freiheitswert.....	35
2.2 Betrachtungsgegenstand: Der Mythos der Künstlichen (Super-) Intelligenz zwischen Technikfolgenabschätzung, Moralfabel und apokalyptischer Vision.....	38
2.2.1 Technikantizipation und Ressourcenkonflikt: Technosoziale Sinnbilder.....	38
Akteure und Institutionen entgrenzter Sinnbilder der Künstlichen Intelligenz... Wissenschaftlichkeit als Prüfstein und rhetorisches Mittel der (De-)Legitimation	40 43

2.2.2	Narrative und die lebensorientierende Kraft von Erzählungen.....	45
	Etablierte Narrative der KI in den Massenmedien der Gegenwart	47
	Die Überwältigung durch die Zukunft als der verbindende ethische Gedanke...	50
2.2.3	Umbruchsphantasien der technologischen Singularität	
	zwischen kupierter und vollendeter Apokalypse.....	53
	Apokalypse und Apokalyptik im Kontext der Künstlichen Intelligenz.....	54
	Apokalypse ohne Reich.....	57
	Konsequenzen	61
3	Künstliche Intelligenz und maschinelles Lernen.....	65
3.1	Künstliche Intelligenz im Kontext der Digitalisierung	65
3.2	Entstehung und Funktion von Künstlicher Intelligenz	67
	Das Wiederaufleben der KI-Forschung zu Beginn des 21. Jahrhunderts	69
	Mustererkennung, Maschinelles Lernen und Künstliche Neuronale Netze	70
3.3	Wirtschaftliche Konzentration, digitaler Kapitalismus und kalifornische Ideologie	72
	Gegenstaatlichkeit und Disruption als Kennzeichen der Silicon-Valley-Ideologie.....	74
	Die ausbleibende Regulierung der Digitalwirtschaft	77
3.4	Spezialistentum, Unwissenheit und Macht: Intellectual Debt und Black Boxes.....	79
	Machtgefälle zwischen Unternehmen und Anwendern	79
	Begründungslücken zwischen Programmen und Menschen.....	81
3.5	Smart statt Intelligent: Wie berechtigt ist die Erwartung an eine starke KI?.....	84
	Die Widerständigkeit des Intelligenzbegriffs in Informatik und Robotik.....	84
	Spielsituationen als Mess- und Vergleichsinstrument von Intelligenz.....	86
	Hubert Dreyfus Systematisierung von Intelligenzstufen	88
	Die Spekulation über mentale Zustände in Maschinen.....	89
	Weltmodelle und das Defizit der digitalen Repräsentation.....	92
	Das Abwägen zwischen Plausibilität und Probabilität hinsichtlich starker KI...	94
	Maschinelles Lernen ist wenn, dann eine andere Form der Intelligenz	96

4 Die ideengeschichtliche Gleichsetzung von Mensch und Maschine	103
4.1 Literarische Welten: (Alb-)Träume von der technischen Weltermächtigung und einer zweiten Schöpfung	105
4.1.1 Technische Utopien: Die erfolgreiche Unterwerfung der Naturkräfte und ihre späte Rache am Menschen	108
Nova Atlantis [1627].....	110
Erewhon [1872].....	113
4.1.2 Der täuschende Apparat in der frühen Romanliteratur und Science-Fiction	119
E.T.A. Hoffmann: Der Sandmann [1816].....	120
Jules Vernes: Das Karpatenschloss [1892].....	124
4.1.3 Erträge.....	127
4.2 Wissenschaft und Epistemologie: Die kybernetische Lehre vom Menschen als reaktivem System und das neue Regieren der Maschinen	131
4.2.1 Das politische Versprechen der Kybernetik: Maschinen regieren über Menschen	135
4.2.2 Der Mensch, eine biologische Maschine.....	139
4.2.3 Kybernetische Theologie.....	143
4.2.4 Erträge.....	148
4.3 Technologische Konvergenzerzählungen des 20. Jahrhunderts... 153	
4.3.1 Teilhards Geistesevolution gen Omega	155
4.3.2 Anders Konvergenz des Menschen in die Großmaschine der Ein-Apparat-Welt	161
4.3.3 Erträge.....	166
5 Technologische Konvergenzmythen: Visionen einer post-singulären Wirklichkeit	173
5.1 Konfrontation: Der Mensch wird von der Maschine verdrängt... 178	
5.1.1 Die Maschine lehnt sich gegen den Menschen auf	178
Ausbruch und Unmöglichkeit der Kontrolle einer Super-KI	178
Der neue Alpha-Prädator: Ressourcenorientierter Eroberer und Vernichter... 180	
Der böse Dschinn: Ein falscher Befehl mit katastrophalen Folgen	182
Die Rache der wohlthätigen Superintelligenz an ihren Widersachern.....	183
Friedliche Machtübergabe an die Maschinen und Selbstaufgabe	186

5.1.2	Ethische Relevanz: Verantwortungsabgabe und Schöpfung.....	188
	Der Mensch ist kein mächtiger Schöpfer mehr, sondern eine ohnmächtige Gebärmaschine.....	189
	Der Mensch scheidet an der Kontrolle des Universums	192
	Die Weltlichkeit der Welt und der Zuspruch Gottes bleibender Fürsorge	197
5.2	Vertretung: Der vollumsorgte Mensch und sein verwaltetes Leben.....	203
5.2.1	Mensch-Maschine-Interaktion in der neuen Polis der Allsolution.....	203
	Die postsinguläre Smart City als Polis der Solution.....	204
	Das Leben erschöpft sich in Freizeitgestaltung	206
	Die beseelte Maschine präsentiert sich als der ideale Lebensbegleiter	208
5.2.2	Ethische Relevanz: Bevormundung und Versöhnung.....	212
	Selbst- und Weltverlust im fremdverwalteten Leben	214
	Freiheit in der Gemeinschaft.....	217
	Schuld, Umkehr und die Notwendigkeit zu verzeihen	220
	Eros und Agape.....	224
5.3	Identifikation: Der Eingang des Menschen in die Technosphäre...	227
5.3.1	Eins werden mit dem technologischen Fortschritt im Mind Upload.....	227
	Der Mind Upload in Popkultur, Wissenschaft und als Corporate Vision	228
	Der Upload und seine Hürden	230
	Körperlose Freiheit	231
	Die ökonomische Lebenswelt der Emulated Minds	234
	Die Eroberung und Verwandlung des Weltraums in Lebensraum für EMs	235
5.3.2	Ethische Relevanz: Expansion und Erlösung.....	238
	Der Transhumanismus rationalisiert den Wunsch nach Unsterblichkeit	241
	Die Erwartung einer innerweltlichen Transzendenz	243
	Erlösungshoffnung zwischen gnostischer Weltflucht und alchemistischer Weltverwandlung.....	245
	Freiheit in der Geschichte	251
	Digitale Engel.....	255

6 Abschließende Beobachtungen.....	259
6.1 Die Lust der Konvergenzerzählungen an der Erniedrigung des Menschen.....	260
6.2 Die Propagierung einer alternativlosen Zukunft der Maschinen: Cui bono?.....	267
6.3 Der bleibende Wert des Mythos für die technische Kultur.....	270
 Literaturverzeichnis.....	 279
 Register.....	 301
Namensregister.....	301
Sachregister.....	305

Vorwort

Vorliegende Arbeit widmet sich den Heils- und Unheilserwartungen, die das Forschungsfeld um Künstliche Intelligenz hervorgebracht hat. Es ist somit nicht die vorliegende Technik selbst, sondern es sind die im Gestus wissenschaftlicher Folgenabschätzung vorgetragenen Potentialerwartungen und Visionen über die Zukunft intelligenter Maschinen, an denen diese Untersuchung Kritik übt. Im Zentrum der Mythen steht die Gleichsetzung von Mensch und Maschine. Der Mensch wird nicht nur mit Begriffen der Technik erklärbar gemacht, sondern er selbst wird zur biologischen Maschine herabgestuft. Indem Mensch und Technik parallelisiert werden, kommt es zu Theorien des Hybriden, der Substitution des Menschen durch die Maschine und letztlich der Vision von der Vernichtung des Menschen. Erwartungen einer zweiten, technologischen Schöpfung verknüpfen sich mit einer apokalyptischen Überwältigungsrhetorik: Der Mensch bringe etwas hervor, das er selbst nicht beherrschen, geschweige denn verstehen könne – und das letztlich sein Ende bedeutet. Der Schauplatz der Geschichte, auf dem sich dieser Herrschaftswechsel von Mensch zu Maschine vollzieht, dies soll das Forschungsfeld der Künstlichen Intelligenz sein.

Der intelligenten Maschine werden innerhalb dieser Erzählungen stets die guten, geradezu gottähnlichen Eigenschaften zugesprochen. War Technik bisher ein Ausdruck menschlicher Souveränität, wird die Technik nun stattdessen zum allmächtigen Subjekt einer neuen, *post-singulären* Zeit erhoben. Die Menschheit hingegen verkommt zum Auslaufmodell der Geschichte, die ihren Dienst als willige Gebärmaschine ihrer *Mind Children* verrichtet, ehe sie von der Weltbühne abtritt.

Ihre ethische Funktion erfüllen diese Erzählungen nicht für eine fiktive Zukunft, sondern in den hervorgerufenen Ängsten gegenwärtiger Rezipienten. Die Mythen legen der Öffentlichkeit nicht zu bewältigende Szenarien vor, nach denen der Untergang des Menschen durch seinen technologischen Fortschritt unabwendbar erscheint. Damit halten sie die Öffentlichkeit in genau jener Stasis, in der sie den Menschen in ihrem Weltbild sehen, und erschweren so die tatsächlich notwendige Auseinandersetzung um die Einsatzmöglichkeiten maschinellen Ler-

nens. Zugleich stützen die Erzählungen das gegenwärtige Machtungleichgewicht zwischen den globalen Tech-Unternehmen gegenüber Staaten und ihren Bürgern. Diese Untersuchung will die Auslieferung des Menschen an die Zukunft der Künstlichen Intelligenz als kontingenten Mythos durchsichtig machen. Indem die Ethik die Technik als Ausdruck menschlicher Freiheit wieder ins Zentrum rückt, wird gleichsam deutlich, dass die Ursachen für die Popularität dieser Mythen auch das Symptom einer Gesellschaft zu sein scheinen, die sich bereits gegenwärtig als technologisch eingeengt, möglicherweise überfordert, und wenig selbstwirksam wahrnimmt.

Technik ist Teil menschlicher Kulturausübung; sie dient wie die Kultur insgesamt der Lebensbewältigung. Ethik hat die technische Kultur auf diesen Zweck hin immer wieder zu prüfen und damit Tendenzen ihrer Selbstverzweckung zu wehren. Ethische Theologie bleibt derweil bei der Kritik an der Sakralisierung von Technik nicht stehen. Sie wirbt im Abgleich der Zerrbilder menschlicher Freiheit, wie sie die Mythen der KI produziert haben, für einen pragmatischen Umgang mit Technik und Innovation auf der Grundlage eines christlichen Freiheitsverständnisses, nach dem Technik sich an tatsächlichen Bedürfnissen von Menschen auszurichten hat.

Dieses Buch ist eine leicht überarbeitete Fassung meiner Qualifikationschrift, die ich im März 2022 an der Evangelisch-Theologischen Fakultät der Ludwigs-Maximilians-Universität München eingereicht habe. Mein herzlicher Dank gilt meinem Doktorvater Prof. Dr. Reiner Anselm. Er ermöglichte mir das wissenschaftliche Arbeiten durch die Anstellung am Zentrum für Technik-Theologie-Naturwissenschaften an der LMU. Tief beeindruckt hat mich sein Verständnis von einer pragmatischen, den Menschen zugewandten Ethik. Diejenigen, die mit dem Programm der ethischen Theologie vertraut sind, werden diese Prägung in meiner Herangehensweise an die Mythen der Künstlichen Intelligenz wiedererkennen. Aus der Fakultät danke ich zudem Prof. Dr. Christian Albrecht. Die praktisch-ethischen Oberseminare wie auch der persönliche Austausch waren eine große Bereicherung.

Allen Mitarbeitern der Systematischen Theologie und Ethik danke ich für das kollegiale Miteinander, aus dem viele Freundschaften

erwachsen. Ich danke Dr. Therese Feiler, Dr. Hendrik Meyer-Magister, Dr. Lukas Meyer, Dr. Nazli Cihan, Nora Meyer, Bertolt Buntschuh, Dr. Stephan Schleissing, Dr. Mario Berkefeld, Elisabeth Perschthaler und Rebecca Helmes. Ich danke zudem Dr. Anne Friederike Becker, Dr. Katharina Herrmann und Dr. Jonathan Spanos, die mich in meinem ersten Jahr als wissenschaftlicher Mitarbeiter liebevoll unter ihre Fittiche nahmen. Großer Dank gilt auch meinem guten Freund und Berliner Studienkollegen Karl-Christoph Goldammer, der die Dissertationsschrift mit wachem Geist redigiert hat. Schließlich danke ich meiner Familie und meinem Partner, bei denen ich Zuversicht, Ruhe und Zerstreuung fand. Eben weil wir keine Maschinen, sondern Menschen sind, bin ich überzeugt, dass diese Unterstützung in entscheidender Weise zum Gelingen dieses Buches beitrug.

Die Dissertation wurde mit dem Wissenschaftspreis 2023 der Hanns-Lilje-Stiftung ausgezeichnet. Für diese Ehrung danke ich den Kuratoriumsmitgliedern der Stiftung und freue mich über die Aufmerksamkeit, die das Werk dadurch erfährt.

Yannick Schlote

München im März 2023

1 Einleitung

1.1 Eliza und ELIZA

Men are all alike.

IN WHAT WAY

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED

It's true. I am unhappy.

DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY

I need some help, that much seems certain.

WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP

Perhaps I could learn to get along with my mother.

TELL ME MORE ABOUT YOUR FAMILY

My mother takes care of me.

WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU

My father.

YOUR FATHER

You are like my father in some ways.

WHAT RESEMBLANCE DO YOU SEE

You are not very aggressive but I think you don't want me to notice that.

WHAT MAKES YOU THINK I AM NOT VERY AGGRESSIVE

You don't argue with me.

WHY DO YOU THINK I DON'T ARGUE WITH YOU

You are afraid of me.

DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU

My father is afraid of everybody.

WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER

Bullies.

DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR
BOYFRIEND MADE YOU COME HERE¹

¹ Joseph Weizenbaum, „ELIZA – a Computer Program for the Study of Natural Language Communication between Man and Machine“, *Communications of the ACM* 9 (1965), 36–45. Hier 36–37.

Mit diesem Transkript trat Joseph Weizenbaum (1923–2008), Wissenschaftler des *Massachusetts Institute of Technology* (MIT), im Jahr 1965 an das computerwissenschaftliche Publikum seiner Zeit. Er präsentierte die Ergebnisse seiner Forschung über natürliche Sprache und elektronische Sprachverarbeitung. Das Erstaunliche an diesem Transkript war, dass die abgebildete Kommunikation sich zwischen Mensch und Maschine ereignet hatte. Eine junge Frau (kleine Lettern) schrieb mit Weizenbaums Programm ELIZA (große Lettern); sie kommunizierte als eine der Ersten mit dem, wofür man später den Begriff des *Chatbots* erfand. Das Chatprogramm ELIZA zeigte in diesem Transkript, dass es in der Lage war, auf die Eingaben des Menschen semantisch einzugehen. Dadurch erscheint die Kommunikation als Gespräch und beide Gesprächsparteien, Mensch als nun auch Maschine, in einem vorerst unbestimmten Sinne intelligent.

Nicht ohne Hintergedanken hatte Weizenbaum seinem Programm den Namen ELIZA gegeben. Er verweist auf das Theaterstück *Pygmalion* (Uraufführung Burgtheater Wien, 1913) des Iren George Bernard Shaw (1856–1950). In dieser modernen Adaption des antiken Pygmalionmythos begegnet der angesehene Sprachwissenschaftler Professor Higgins der jungen Blumenverkäuferin Eliza Doolittle, die er sich zum Testobjekt seiner Studien wählt. Der Professor wettet mit seinen Kollegen, aus der ungebildeten Eliza eine Herzogin zu machen – zumindest, was ihre sprachliche Ausdrucksfähigkeit angeht –, indem er ihr noch einmal neu das Sprechen beibringt. Das Experiment glückt: Nach kurzer Zeit beweist Eliza Doolittle, sehr zur Belustigung der Gesellschaft um Higgins, nahezu höfische Umgangsformen. Weil auch Weizenbaum seiner ELIZA die Verarbeitung natürlicher Sprache von Grund auf eingab, bediente er sich dieser Tradition.² In systematischer Perspektive sind das Schicksal der Figur der Eliza Doolittle und das des Sprachverarbeitungsprogrammes ELIZA enger miteinander verknüpft, als es der bloße Namensverweis erscheinen lässt. In ihrer wechselseitigen Ausdeutung lassen sich die tragenden ethischen Themen skizzieren, die die

2 Vgl. Joseph Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 1. Aufl., Frankfurt am Main 1977, 15.

Beschäftigung mit Künstlicher Intelligenz (KI) seit der Begriffsbildung in den 50er Jahren des 20. Jahrhunderts bis heute begleiten:

Identität und Täuschung

Zum einen führt der Beginn der KI-Forschung zu einer intensivierten Auseinandersetzung mit den Fragen nach menschlicher Identität und der Möglichkeit ihrer maschinellen Simulation. In der Sprachfähigkeit drückt sich menschliche Identität nicht nur aus, sondern sie bringt Identität gleichzeitig erst hervor. Für die Figur Eliza Doolittle stellt sich die Frage, ob Professor Higgins in der Sprachumerziehung gleichzeitig eine neue Person aus Eliza erschafft. Steht die umerzogene Eliza tatsächlich noch in einer Kontinuität zu ihrem ehemaligen Selbst – spielt sie dann möglicherweise nur eine geschickte Maskerade, oder ist sie durch die Sprachumerziehung zugleich ein Geschöpf Higgins geworden? Über die frühen Programme wie dem von Newell und Simon 1957 entwickeltem General Problem Solver³ als auch ELIZA wurde seit ihrer Konstruktion debattiert, ob in deren Informationsverarbeitung von einem ersten intelligenten Verstehen gesprochen werden könne, oder ob dieser Prozess der Symbolmanipulation eine grundsätzlich andere Qualität als menschliches Verstehen habe und deshalb im Fall von ELIZAS Sprachfähigkeit von lediglich vorgetäuschter Intelligenz auszugehen sei. Innerhalb der Informatik führte diese Auseinandersetzung in der Entwicklung Künstlicher Intelligenzen zu einer Unterscheidung der Modi von *Performanz* und *Simulation*: Im Modus der *Performanz* schreibt man Programme, die von außen intelligent wirken, gleichgültig, ob diese Mechanismen strukturelle Ähnlichkeit zu biologischer Intelligenz aufweisen. Im Modus der *Simulation* hingegen geht es darum, mentale Prozesse des Menschen zu analysieren, zu abstrahieren und informatisch nachzubilden, um so etwas über menschliches Denken zu lernen.⁴ Von Beginn an wurde diese Differenzierung, die man als Selbstbegrenzung der Wissenschaft gegenüber äußeren Ansprüchen lesen kann, auch von Informatikern innerhalb der Entwicklung von Künstlicher Intelligenz

3 Vgl. Allen Newell/Herbert A. Simon, „GPS, a program that simulates human thought (1961)“, in: Edward Feigenbaum (Hg.), *Computers and Thought*, Menlo Park, Calif. 1995, 279–293.

4 Vgl. Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 221–222.229.

kritisiert. Befürworter des Behaviorismus, zu denen der Mathematiker und KI-Pionier Alan Turing (1912–1954) mit seinem berühmten Turing-Test gehörte, erklärten die Trennung zwischen *Performanz* und *Simulation* für obsolet. Sie sprachen sich dafür aus, zwischen dem wahrgenommenen Verhalten und den angenommenen internen Prozessen in der Maschine überhaupt nicht zu unterscheiden, sondern – da innere Prozesse dem Empiriker nicht zugänglich seien – das Erscheinen dem Sein gleichzusetzen. Maschinen, die versichernd intelligent wirken, täuschten nichts vor, sondern seien demnach intelligent.⁵

Kontrolle und die Angst vor dem Kontrollverlust

Die Figur Eliza Doolittle und das Programm ELIZA lenken den Blick der Auseinandersetzung um Künstliche Intelligenz zweitens auf die Frage nach Kontrolle und die Sorge vor dem Kontrollverlust in der Mensch-Maschinen-Interaktion: In Shaws *Pygmalion* sieht Professor Higgins in Eliza Doolittle lediglich ein Objekt seiner wissenschaftlichen Aspiration und handelt an ihr wie an einem Automaten, dem er Befehle eingibt. In den Fortschritten maschinellen Lernens und der Möglichkeit, immer komplexere Aufgaben an Expertensysteme abzugeben, knüpfen sich hingegen Sorgen, dass der Mensch die Kontrolle über seine Artefakte verliert und er selbst zum Objekt fremder Zwecke gerät.⁶ Bereits ELIZAs Chatfunktion lässt sich als ein kommunikativer Machtwechsel zwischen Mensch und Maschine deuten: Aus pragmatischen Gründen gab Weizenbaum ELIZA ein Skript ein, das es ermög-

⁵ Vgl. Alan Mathison Turing, „Computing Machinery and Intelligence“, *Mind* LIX (1950), 433–460. Hier 442–443.

⁶ Der Machtzuwachs intelligenter Maschinen äußert sich beispielsweise in der Zuerkennung von Rechten: So verliehen die Vereinigten Arabischen Emirate dem humanoiden Roboter Sophia im Jahr 2017 die Bürgerrechte. Deren Entwickler, das Hongkonger Tech-Unternehmens Hanson Electronics, haben sicherlich bewusst überlegt, wie sie ihr Produkt benennen. In der offengelassenen Zweideutigkeit von Witz und Anspruch ist mit der Namensgebung Sophia schließlich die Rückfrage an den Menschen gestellt, der von sich ebenfalls behaupten könne, weise (σοφός) zu sein. Theologisch erinnert die Androidin Sophia damit an die personifizierte Frau Weisheit, wie sie in der nachexilischen Literatur des Alten Testaments auftritt und sich dort als erstes Schöpfungswerk und Mitschöpferin der Welt (Spr 8,22–36) offenbart. Hier sind die Machtverhältnisse zwischen dem Menschen und der vergöttlichten Weisheit eindeutig: „Wohl dem Menschen, der mir gehorcht.“ (Spr 8,34a, Lutherbibel 2017).

lichte, die Rolle eines an Carl Rogers Ansatz orientierten Psychotherapeuten zu spielen, der mit einem Patienten das erste Gespräch führt. Ein solcher Therapeut war verhältnismäßig leicht zu imitieren, da ein Großteil seiner Technik darin besteht, den Patienten dadurch zum Sprechen zu bringen, dass diesem seine Äußerungen wie bei einem Echo zurückgeworfen werden. Durch diese Art der Gesprächslenkung brauchte das Programm keine Kenntnis von der Außenwelt zu haben.⁷ Das bedeutete aber gleichzeitig, dass das Machtgefälle einer psychotherapeutischen Sitzung von ELIZA übernommen wurde. ELIZA drängte damit dem Gegenüber die Rolle eines Patienten auf. Nahm die Person diese Rolle an, hatte dies mehrere Konsequenzen: Einerseits erhöhte die Akzeptanz dieses Gesprächsformates die Bereitschaft auf menschlicher Seite, über nichtlineare oder alogische Aussagen des *Chatbots* hinwegzusehen und sie stattdessen als wohlüberlegte psychotherapeutische Fragen zu interpretieren. Gleichzeitig akzeptierte der menschliche Partner die Rolle desjenigen, der kontinuierlich etwas über sich preisgibt. In dem vorliegenden Gespräch vertraute sich die junge Frau dem Programm ELIZA in sehr kurzer Zeit in hohem Maße an. Das entgegengebrachte Vertrauen kann über das reine Sprachspiel hinaus weitreichende Konsequenzen mit sich bringen. Weit abseits von Phantasien über böswillige, künstliche Supersysteme besteht die Gefahr, dass Menschen durch *Chatbots* getäuscht und ausspioniert werden.

Sorgen vor dem Kontrollverlust über die Maschine werden bestärkt durch die Idee einer allgemeinen Verstetigung des technologischen Fortschritts, an dessen Ende die Maschinendomäne den Menschen hinter sich zurücklässt. Bereits Mitte des letzten Jahrhunderts diagnostizierte der Technik- und Medienphilosoph Günther Anders (1902–1992) einen epochalen Bruch in der Beziehung des technischen Experten zu seinem maschinellen Produkt, wofür er den Begriff der *promethischen Scham* fand.⁸ Demnach sah er die westlichen Industrieländer im 20. Jahrhundert zu einer Stufe technologischer Exzellenz fortgeschritten, von der aus sich neben dem Stolz des Ingenieurs unweigerlich die

7 Vgl. Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 15.

8 Vgl. Günther Anders, *Die Antiquiertheit des Menschen*. Bd. I: Über die Seele im Zeitalter der zweiten industriellen Revolution, München 1956, 23.

Bekommenheit vor dem eigenen Geschaffenen mischen müsse, weil er sich im Anblick der effizienten Maschine seiner menschlichen Unzulänglichkeit bewusst werde. Dem Menschen als *homo faber*, dem das Entwickeln neuer Technologien eine Lebensnotwendigkeit sei, müsse diese Scham und das sie nährnde Unbehagen vor der zunehmenden Gewalt seiner technischen Apparate fortwährend begleiten.⁹

Mythos und Logos

Drittens ist Weizenbaums eigener Rückgriff auf die Figur der Eliza Doolittle ein Kennzeichen für die Metaebene der Diskussion um Künstliche Intelligenz, die sich im Ineinander von Mythos und Logos¹⁰ verdichtet. Durch Shawns Theaterstück schimmert beispielsweise immer noch die antike Pygmalionsage nach Ovid hindurch und prägt den Blick auf ELIZA: Dort verliebt sich König Pygmalion in die marmorne Statue Galatea. Nicht Pygmalions Schläue kann ihn aus dieser Verzweiflung erretten, sondern allein die Göttin Venus ist es, die sie sich seinem Schmachten erbarmt und Galatea aus dem bloßen Schein eines Menschen tatsächlich belebt. Allein die Göttin der Liebe vermag es im klassischen Mythos, den Menschen vor seinem Objektfetisch zu bewahren. In der modernen Adaption um Eliza scheint es dagegen allein in der Macht des Menschen zu stehen, Menschen wie Automaten zu behandeln. Aber auch andere antike Mythen sind an ELIZA anschlussfähig: Die Interaktion zwischen ELIZA und Mensch kann

⁹ Vgl. Max Scheler, *Die Stellung des Menschen im Kosmos*, 12. Aufl., Bonn 1991 (1928), 37. Scheler führt Anfang des 20. Jahrhunderts den Begriff des *homo faber* in die Philosophie ein. Unter seiner Theorie des *homo faber* versteht Scheler ein Menschenbild in Anlehnung an die Denkschulen Darwins und Lamarcks, das keinen prinzipiellen, sondern nur einen graduellen Unterschied in der Intelligenz zwischen Tier und Mensch sieht. Max Frisch wie Hannah Arendt werden diesen Begriff fortführen; in ihrem philosophischen Hauptwerk der *Vita activa* stellt Arendt dem *homo faber* das *animal rationale* wie den *homo politicus* als komplementäre Formen menschlichen Tätigseins bei.

¹⁰ Der Logos wird verstanden als der Anspruch rationaler Welterkenntnis durch Mittel der Logik und Empirie; der Mythos hingegen verbindet sich mit den klassischen Sagen, Märchen, magischen Praktiken und religiösen Gewissheiten als die Summe von Geschichten und Motiven von hochgradiger Beständigkeit, die dem Menschen im Angesicht einer beängstigenden Wirklichkeit Sinn und Orientierung in der Welt gibt. Die Fiktion des Mythos ermögliche die Heimstatt in der Welt; wer aus Angst oder in Angst handelt, der hat laut Blumenberg den Mechanismus vorgeschobener imaginativer Instanzen verloren. Vgl. Hans Blumenberg, *Arbeit am Mythos*, 6. Aufl., Frankfurt am Main 2001, 12–18.

als Wiederaufnahme des Dramas um Echo und Narziss gelesen werden. Aufgrund ELIZAs Weltlosigkeit besteht der Sinn des Skripts darin, diese zu überdecken und im Widerhall und Modulation des Input das Gespräch selbstreferentiell am Laufen zu halten. In diesem Sinne leiht sich ELIZA wie Echo die Stimme des Gegenübers und bleibt von dieser abhängig. In der antiken Sage kann Narziss mit Echo nichts anfangen und wird damit bestraft, sich in sein Spiegelbild zu verlieben. Die Tragweite eines Gespräches mit ELIZA hängt ebenfalls von der Bereitschaft der Person ab, das imaginierte Gegenüber lediglich als Spiegel seiner selbst zu verstehen. Für den antiken Mythos bleibt es am Ende klar ein Sakrileg des Menschen, im Gegenüber nur wieder sich selbst hervorbringen zu wollen. Schließlich wird Narziss Obsession auch ihn die Tiefe des Wassers hinabreißen.

Künstliche Intelligenz wurde seit ihrem Aufkommen nicht als bloße Anwendungstechnik behandelt, sondern schürte fortwährend Erwartungen, dass an und durch ihre Forschung sich Grundsätzliches über das Wesen des Menschen erweisen werde. Daher ist es nicht verwunderlich, wenn sich die ethischen Muster der klassischen Mythen in aktuellen Debatten darüber wiederholen, was der Mensch denn mit der Künstlicher Intelligenz *tatsächlich* zu erreichen suche: Einerseits wird dem Menschen vorgeworfen, dass er in der Entwicklung von Künstlicher Intelligenz (ähnlich wie in anderen Hochtechnologien, z. B. der Gentechnologie) lediglich eine narzisstische Selbstspiegelung seiner Macht verfolge. In seiner Hybris gegen die Natur und Gott schwingt sich der Mensch zum Schöpfer auf und erliege dabei seinen eigenen Illusionen. Weil er die Apparate für gleichwertige Schöpfungsakte betrachte, könne diese Anmaßung nur in der Katastrophe enden. Ins Positive gekehrt hofft der Mensch, in der Simulation von Intelligenz in Maschinen ein sonst in der Natur nicht gefundenes, bewusstes Gegenüber zu finden. Es ist die Hoffnung, dass in den Tiefen einer Metaphysik der Technik etwas oder jemand auf den Menschen wartet, dass der Mensch es in die Existenz holt: ein geschaffenes Abbild, durch das der

Mensch erst sein Menschsein vollends begreifen könne¹¹, – und mit dem er diesen Zustand dann gemeinsam zum Posthumanen hin überwindet. „It is said that to explain is to explain away.“¹² Anders ausgedrückt: Im Erklären geht das Auratische verloren. Weizenbaum war sich der Offenheit seiner Forschung für mythologisches Denken bewusst und stellte fest, dass die Affinität zum Mythos besonders für den Bereich der Programmierung und der Künstlichen Intelligenz gelte. Dennoch glaubte Weizenbaum zuerst fest an einen unumkehrbaren Prozess der Entzauberung von Technologie. Das auf den Betrachter im ersten Moment noch gänzlich Wundersame werde, sobald der Schleier gelüftet, auf den zweiten Blick als profaner Wirkungszusammenhang offenbar. Diese wortwörtliche *Ent-täuschung* führe vorerst zwar nur dazu, dass der Weizenbaums Ansicht nach irreführende Begriff der Intelligenz im Bereich der Informatik lediglich dem nächsten angeheftet werde. Letztlich aber werde dasjenige, das man hinter der Maske des Magischen befände, zur Gänze fallen.¹³ Die Hoffnung, dass der Fortschritt zu einer kathartischen Entzauberung der wissenschaftlichen Erkenntnis führen werde, hat sich in den Jahren und Jahrzehnten nach der Veröffentlichung seines Artikels über die Wirkweise von ELIZA nicht erfüllt. Stattdessen wurde das mythologische Denken im Bereich des maschinellen Lernens Weizenbaum zu einem Thema, mit dem er sich zeitlebens beschäftigte. Diese Vehemenz entsprang seiner Verstörung, wie in der Öffentlichkeit auf ELIZA reagiert wurde. Denn er sah seine eigentlichen Ergebnisse über die Formalisierbarkeit von Sprache hinter den Erwartungen und Hoffnungen der Öffentlichkeit untergehen: Erstens kam eine Anzahl von Psychiatern in ihrer Euphorie zu der Ansicht, das Computerprogramm könne zu einer völlig automatischen Form der Psychotherapie ausgebaut werden. Dem Ideal der Kybernetik zufolge sollte der Therapeut als Person weitgehend verschwinden und letztlich

11 Vgl. Lukas Brand, *Künstliche Tugend*. Roboter als moralische Akteure, 1. Aufl., Regensburg 2018, 18; ebenfalls vgl. Eilert Herms, „Künstliche Intelligenz. Wesen und sozialetische Probleme aus der Sicht des christlichen Menschenbildes“, in: *Gesellschaft gestalten*. Beiträge zur evangelischen Sozialethik, Tübingen 1991, 284–295.

12 Weizenbaum, ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine, 36.

13 Vgl. ebd.

wie eine Datenverarbeitungsmaschine verhalten. Weizenbaum selbst sah in ELIZA nicht mehr als die Parodie eines Psychotherapeuten. Zweitens stellte Weizenbaum fest, wie schnell und intensiv Personen in der Unterhaltung mit ELIZA eine emotionale Beziehung herstellten und sie dem Programm eindeutig menschliche Eigenschaften zuschrieben. Ein sehr kurzer Kontakt mit dem relativ einfachen Computerprogramm führte nach Ansicht Weizenbaums zu ernstzunehmenden Wahnvorstellungen. Drittens überraschte ihn die verbreitete Ansicht, bei ELIZA handele es sich um die allgemeine Lösung des Problems, dass Sprache kontextabhängig ist, der Computer aber keinen Kontext im Sinne einer Innerweltlichkeit hat. So kamen viele zu dem Schluss, Weizenbaum habe mit ELIZA einen weltlosen Intellekt erschaffen – einen Homunkulus in der Flasche.¹⁴

Weizenbaum machte die Erfahrung mit der Vitalität mythologischen Denkens in der Moderne. Hans Blumenberg (1920–1996) beschreibt in seinem wegweisenden Werk für das 20. Jahrhundert *Arbeit am Mythos*, dass Mythen nicht einfach entstehen oder verschwinden, sondern gemacht und bearbeitet werden. Die Metaphern und Geschichten, auf denen Weltansichten im Kern beruhen, müssten ständig weiter erzählt und umgeschrieben werden, um als Vorschläge sinnvoller Welteinfindung dem Einzelnen wie Gruppen zu nützen. Das bedeutete aber auch, dass jede Zeit auf ihre Mythen angewiesen bleibt. Damit erteilte Blumenberg solchen Positionen eine deutliche Absage, die davon ausgingen, dass die Geschichte vom Mythos zum Logos voranschritt.¹⁵ Unter diesem Eindruck der fortdauernden Mythosbedürftigkeit unterzogen Horkheimer und Adorno die Moderne einer Generalkritik im Namen einer *Dialektik der Aufklärung*: Die Behauptung, die Aufklärung habe den Mythos abgelöst, enttarnen sie als Illusion. Versuche der Entzauerung brächten gerade immer wieder neue Mythen hervor, weil das

¹⁴ Vgl. Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 17–21.

¹⁵ Vgl. Blumenberg, *Arbeit am Mythos*, 40. Die Übernahme von Blumenbergs Mythosdefinition im Sinne einer Handlungermächtigung durch die Fiktion erfolgt, ohne seiner darauf ansetzenden Kritik am Christentum als erstarrtem Dogma beizupflichten. Im Gegenteil argumentiert vorliegende Untersuchung, dass der gewachsene Traditionsbestand des Christentums sehr wohl eine Orientierung gegenüber den existentiellen Anfragen der Mythen um Künstliche Intelligenz anbietet.

Programm der Aufklärung, alle Lebensbereiche des Menschen wissenschaftlich läutern zu können, selbst mythologischen Ursprunges war. Die Aufklärung entstamme zwar dem Mythos, habe diesen aber nicht einfach abgelöst, sondern impliziere ihn weiterhin und sei dadurch immer wieder zur Konstruktion neuer Mythen fähig.¹⁶ Das Problem der Aufklärung sei dabei ihre Radikalität: Während frühere Entmythologisierung ältere Mythen durch aufgeklärtere Mythen ersetzen, lasse die Radikalität des aufklärerischen Programmes in der Welt keinen Raum für das Unbekannte. Erst dadurch sei die Gefahr gegeben, dass gerade die Vernunft in nicht erkannte Unvernunft umschlägt. Mit der Gleichsetzung einer vollständig mathematisierten Welt mit Wahrheit sei die Zahl zum Grundsatz der Aufklärung avanciert, durch die man sich einer letzten Wirklichkeit gewiss und vor dem Mythischen sicher schien.¹⁷ Dadurch sind seit jeher besonders die Naturwissenschaften anfällig dafür, ihre eigene Mythogenese zu ignorieren. Das der Formalisierung, Mathematisierung, und heute der Digitalisierung wie Algorithmisierung Widerständige wird von der Logik der Aufklärung ignoriert und gedeiht so inmitten des eigentlich für geläutert und beherrscht gehaltenen Bereichs der Wissenschaften. Dieser unterschlagene Sinnüberschuss der Lebensbereiche, der auch der Wissenschaft und der Technologie als Ausdruck menschlicher Geistestätigkeit inne ist, bleibt so unbewältigt und verdunkelt darin die Technologie.¹⁸ Die Maschine, die ursprünglich die Materialisierung einer durch die Vernunft durchdrungenen Natur bedeutete, gerät in ihrer Unverbundenheit zum Menschen und ihrer fortschreitenden Komplexitätssteigerung zunehmend zum unverstandenen Gegenüber des Menschen. In ihrer Obskürtheit erhält die Maschine schließlich Züge dessen, was Rudolf Otto (1869–1937) religionspsychologisch als das Heilige und Numinose bestimmte, welches der Mensch nur in der Gleichzeitigkeit von machtvoller Anziehung wie Abschreckung erfährt.¹⁹

16 Vgl. Max Horkheimer/Theodor W. Adorno, *Dialektik der Aufklärung*. Philosophische Fragmente (Fischer-Taschenbücher Fischer Wissenschaft 7404), 23. Aufl., Frankfurt am Main August 2017, 13.

17 Vgl. ebd., 39–40.

18 Vgl. ebd., 43.

19 Vgl. Rudolf Otto, *Das Heilige*, 4. Aufl., Breslau 1920, 8–11.

1.2 These, Vorgehen und Ziel der Untersuchung

In der mythologischen Kraft der Künstlichen Intelligenz, die sich besonders in Potentialerwartungen und Visionen über die Zukunft Bahn bricht, liegt der Forschungsgegenstand dieser Arbeit. Die Untersuchung wird dabei von der Einsicht geleitet, dass Technik und Medien nicht funktional, sondern kontingent sind. Es existiert eine interpretative Flexibilität zwischen der (vorliegenden wie antizipierten) Technik und den sie deutenden Sinnbildern. Eine ethische Reflexion dieser mythischen Zukunftsszenarien erscheint nur dann sinnvoll, wenn man sie vor den Hintergrund ihrer geistesgeschichtlichen Tradition wie ihrer sozio-polit-ökonomischen Situation stellt. Die Voraussetzungen und Hintergründe der Mythen um Künstliche Intelligenz durchsichtig zu machen und sie einer ethischen Reflexion zu unterziehen, ist die Aufgabe, die diese Arbeit an ihren Gegenstand heranträgt.

Die Mythenproduktion im Bereich der Künstlichen Intelligenz hat sich seit ihrer Wortschöpfung Mitte der 1950er Jahre als besonders produktiv erwiesen – die Erkenntnisse um ELIZA bilden dabei ein Glied in einer langen Kette. Nach einer ersten Hochphase der Forschung um KI bis in die Mitte der 1970er Jahre und dem darauffolgenden Relevanzverlust aufgrund nicht eingelöster Erwartungen, erwacht seit den 1990er Jahren das öffentliche Interesse von Neuem; dies liegt nicht zuletzt an den informationstechnologischen Durchbrüchen wie dem maschinellen Lernen mittels tiefneuronaler Netze (*Deep Neural Networks*, auch DNN). Dieses Wiederauferstehen der Forschung um Künstliche Intelligenz geht mit der Aufnahme früherer als auch der Arbeit an neuen Mythen einher.

Der grundlegende Mythos von der Ähnlichkeit zwischen Mensch und Maschine

Die neuen Mythen der Künstlichen Intelligenz bauen auf bestehenden Mythen der Technik und Automatisierung auf, von denen einige bereits seit dem Aufstieg der modernen Naturwissenschaften bestehen. Der Kommunikationswissenschaftler Claus Eurich fasste diese um die Jahrtausendwende unter dem Großmythos der Maschine zusammen:

Dazu zählte er den Glauben an einen nicht endenden technologischen Fortschritt und an die kontinuierliche Verbesserung der Lebensumstände durch die Technisierung der Umwelt. Diese Zuversicht habe sich verbunden mit dem Mythos von der prinzipiellen Machbarkeit der Maschine; dort, wo etwas noch nicht durch Maschinen gelöst werden kann, sei dies lediglich eine Frage der Zeit. Dem stehe die Idee einer Unfehlbarkeit oder zumindest Objektivität der Maschine nahe, in der Sicherheit mit Berechenbarkeit gleichsetzt wird. Dazu trete schließlich der Mythos der evolutionären Maschine, in dem Sinne, dass die Maschine selbst lebendig erscheint und den Gesetzen der Evolution folgt, um in ihrer Weiterentwicklung und Ko-Evolution mit dem Menschen immer tiefgreifender dessen Mängel gegenüber seiner Umwelt auszugleichen.²⁰ Die übergreifende Schlüsselidee im Zentrum des Mythos Maschine ist für Eurich jedoch die Gleichsetzung von Mensch und Maschine. Sie sei die maßgebliche Syntheseleistung der Wissenschaft des abgelaufenen zweiten Jahrtausends. Diese integrativen Idee im Zentrum des Großmythos Maschine hält Eurich für den stärksten sekundären Mythos, den die Menschheit hervorgebracht habe, weil er dem modernen Menschen eine umfassende Orientierung von Umwelt, Mitwelt und Innenwelt biete.²¹ Zu einer ähnlichen, aber noch spezifischeren Deutung kommt fünfzehn Jahre später der Sozialwissenschaftler Thomas Bächle, wenn er das Menschenbild der Gegenwart vom Mythos des Algorithmus durchzogen sieht. An die Stelle der Zahl als Symbol der Aufklärung und die Maschine bei Eurich ist für Bächle nun der Algorithmus getreten, an dessen Modellogik entlang sich das Leben des Menschen, sein Körper sowie sein Bewusstsein entworfen und unterworfen werde. Unter der Behauptung, alle Eigenschaften wie Tätigkeiten des Menschen seien universell formalisierbar, werde der Mensch selbst lediglich zum Produkt vorbestimmter Prozesse.²² Ähnlich wie Eurich sieht auch Bächle den Reiz dieses Großmythos in sei-

20 Vgl. Claus Eurich, „Mythos und Maschine“, in: Bernd Flessner (Hg.), *Nach dem Menschen. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur*, Freiburg im Breisgau 2000, 19–41. Hier 30–31.

21 Vgl. ebd., 29–30.

22 Vgl. Thomas Christian Bächle, *Mythos Algorithmus. Die Fabrikation des computerisierbaren Menschen*, Wiesbaden 2014, 15.

ner entlastenden Funktion, als er den Menschen mit Modellen des Technischen erklärbar macht. Die spätestens seit der Aufklärung zerbrochene Einheit von Wissen und Sinn scheint durch den Algorithmus wiederhergestellt. Anstatt aber nur vorangegangene Deutungen zu einem Ganzen zu beschließen, bietet gerade diese Synthese in den Worten Blumenbergs *neue Arbeit* am Mythos. Bächle zieht die zentrale Konsequenz, wenn er sagt:

„Nur indem Mensch und Technik funktional äquivalent modelliert werden, kann es zu Theorien und Narrativen des Hybriden, der Aufgabe des Natürlichen, der Substitution des Menschen durch die Maschine, letztlich der Vision vom Existenzverlust des Menschen kommen.“²³

Weiterarbeit am Großmythos: Konvergenz und bevorstehender Umbruch

Die Idee, dass Mensch und Maschine letztendlich identisch *funktionieren*, ist die mythologische Basis für die Zukunftsvisionen, die den Gegenstand dieser Untersuchung bilden, als sie eine Leerstelle des Mythos Maschine bzw. Algorithmus zu lösen suchen. Diese Untersuchung wirbt dafür, die Zukunftsvorstellungen um Künstlichen Intelligenz als Konvergenzscenarien (lat. *convergere*: sich annähern, zusammenlaufen) von Maschine und Mensch zu bestimmen, in denen die durch den Mythos Algorithmus verursachte Spannung einer unverbundenen Gleichstellung von Mensch und Maschine geschichtsteologisch aufgelöst werden soll. Die mythischen Visionen der KI zeichnen demnach verschiedene Enden, wie das fortschreitende Aufholen der Technik gegenüber dem Menschen schlussendlich ausgeht.

Eine Konvention innerhalb dieser Konvergenzerzählungen bildet der Glaube an die starke Künstliche Intelligenz. Die Unterscheidung zwischen starker und schwacher KI hat der Sprachphilosoph John Robert Searle vor rund vierzig Jahren vorgeschlagen.²⁴ Sie ist seitdem ein fester Bestandteil der mythologischen Gedankenwelt geworden.

²³ Ebd., 27.

²⁴ Vgl. John R. Searle, „Minds, Brains, and Programs“, *The Behavioral and Brain Sciences* (1980), 417–457. Hier 417.

Schwache oder auch enge (engl. *narrow*) Künstliche Intelligenz ist auf bestimmte Aufgaben trainiert und kann auch nur diese Tätigkeiten ausführen. Alle bis heute entwickelten Programme sind demnach schwach. In ihrem Schatten steht die Vorstellung von einer starken KI, einer Ultrainelligenz oder auch generellen Künstlichen Intelligenz (engl. AGI: *Artificial General Intelligence*), die wie der Mensch eine Vielzahl unterschiedlichster Aufgaben lösen kann. Die imaginierten Wege von einer schwachen hin zu jener starken Künstlichen Intelligenz sind vielfältig. Neben der Vorstellung, man könne im Verfahren des *Mind Uploads* das menschliche Gehirn digital kopieren und so ein künstliches Bewusstsein erzeugen,²⁵ kommt die größte Aufmerksamkeit dabei den Erzählungen zu, die die Entwicklung einer starken KI mit der Theorie der Intelligenzexplosion verbinden. Der Informatiker Irving John Good (1916–2009), ein Kollege des KI-Pioniers Alan Turing, äußerte 1963 als Erster im wissenschaftlichen Raum die Theorie über eine maschinelle Ultrainelligenz, die den Menschen in allen intellektuellen Aktivitäten übertreffen wird.²⁶ Da das Konstruieren von Maschinen ebenfalls zu diesen intellektuellen Tätigkeiten gehört, folgert Good, dass die erste intelligente Maschine noch bessere Maschinen konstruieren werde. Die Intelligenzexplosion ist das Resultat unzähliger Durchläufe dieses Kreislaufes, in dem intelligente Maschinen immer intelligentere Maschine hervorbringen. Dem Menschen komme es laut Good lediglich zu, die erste ausreichend intelligente Maschine zu schaffen, die diesen Vorgang ins Rollen bringt.²⁷ Good ist dabei überzeugt, dass diese letzte Erfindung der Menschheit ihr das Überleben auf Dauer sichern werde. Gleichzeitig warnt er davor, dass der einmal gestartete Anstoß der Intelligenzexplosion eine Entwicklung in Gang setze, dessen Ende der Mensch nicht mehr überblicken werde.²⁸

Die Erwartung eines qualitativen Umbruchmoments, der in dieser Weise durch ein überexponentielles Beschleunigen des technologischen

25 Vgl. Nick Bostrom, *Superintelligence*. Paths, dangers, strategies, Oxford 2014, 28.

26 Vgl. Raymond Kurzweil, *The Singularity Is Near*. When Humans Transcend Biology, New York/Cambridge/et al. 2005, 178.

27 Vgl. Irving John Good, „Speculations Concerning the First Ultrainelligent Machine“, *Advances in Computers* 6 (1966), 31–88. Hier 32–33.

28 Ebd., 31.

Fortschritts ausgelöst wird, hat sich in der mythologischen Tradition der KI eng mit dem schillernden Begriffen der Singularität verbunden. Während manche Visionäre die Singularität mit der Intelligenzexplosion gleichsetzen, erwarten andere an dessen Ende einen Kipppunkt, dem etwas aus heutiger Sicht Unberechenbares folgt. Die Singularität, vom lateinischen *singularitas* (das Einzelne oder Einmalige) abgeleitet, ist ursprünglich ein mathematischer Kunstbegriff des 20. Jahrhunderts, der den Punkt bezeichnet, ab dem sich eine Kurve oder Fläche anders zu ihrem normalen Verlauf verhält. Der erste Nachweis, nach dem der Begriff der Singularität auf die menschliche Geschichte angewandt wurde, geht auf ein Gespräch des Mathematikers John von Neumann (1903–1957) mit seinem Kollegen Stanisław Ulam (1909–1984) zurück. Mit der Singularität markieren sie die kontinuierliche Steigerung des technologischen Fortschritts der letzten Jahrhunderte zu einer besonderen Phase der Menschheitsgeschichte, die notwendig für die weitere Entwicklung ist.²⁹

Von hier sind es noch einmal mehr als dreißig Jahre, ehe der Mathematiker Vernor Vinge Neumanns Singularitätsbegriff aufnimmt und in Verbindung mit Goods Thesen einer künstlichen Ultrainelligenz zu einem Fixpunkt, einem Umschlagsmoment, verdichtet. In seinem 1993 erschienenen Artikel *The Coming of the Technological Singularity* kündigt Vinge an, dass der Zeitpunkt eines durch hyperintelligente Maschinen induzierten exponentiellen Fortschrittmoments unmittelbar bevorstehe. Vinge hält Good zugute, er habe die Entwicklung frühzeitig erkannt, aber die verstörenden Konsequenzen der Ultrainelligenz unterschätzt. Mit dem Begriff der technologischen Singularität will Vinge einfangen, dass aus der Ultrainelligenz der Maschinen nicht bloß ein übermächtiges Werkzeug für den Menschen hervorgeht, sondern sie den Beginn einer neuen Wirklichkeit als solche markiert.³⁰

29 Stanislaw Ulam, „Tribute to John von Neumann 1903–1957“, *Bulletin of the American Mathematical Society* 64 (1958), 1–49. Hier 5.

30 Vgl. Vernor Vinge, „The Coming of Technological Singularity: How to Survive the Post-Human Era“, in: NASA. Lewis Research Center (Hg.), *Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*. 1993, 11–22. Hier 13.

„I think it's fair to call this event a singularity (“the Singularity” for the purposes of this paper). It is a point where our models must be discarded and a new reality rules. As we move closer and closer to this point, it will loom vaster and vaster over human affairs till the notion becomes a commonplace. Yet when it finally happens it may still be a great surprise and a greater unknown.“³¹

Das Hereinbrechen der Singularität stellt sich Vinge plötzlicher als alle bisherigen technologischen Revolutionen vor. Forscher stellten nur an ein paar Parametern, und auf einmal könne es so aussehen, als ob die Artefakte als Ganzes plötzlich zum Leben erwachten. Was nun wenige Tage oder Monate nach diesem Umbruchsmoment stattfinden wird, dazu hält sich Vinge bedeckt und sagt, darüber ließe sich aus der Gegenwart heraus nur in Analogien sprechen. Damit spannt er den Horizont für alle folgenden Konvergenzsznarien weit auf, indem er sowohl die Auslöschung der Spezies Mensch als auch den Aufstieg der Menschheit zu unsterblichen Wesen einer post-menschlichen Ära für möglich hält.³²

Die technologische Singularität hat als geschichtsphilosophische Chiffre eine bis dato anhaltende Karriere in der Auseinandersetzung um Künstlichen Intelligenz gemacht. Auch wenn in den späteren Szenarien immer wieder andere Spezifika auftreten, bleiben seine Kernelemente die Beschleunigung des Wandels, die Aussicht auf eine Intelligenzexplosion und der beschränkte Vorhersagehorizont.³³ Einig sind die Spekulanten über die Bedeutung der technologischen Singularität darin, dass dieses Überschreiten durch eine ultraintelligente KI in kein Verhältnis mit heutigen Lebenserfahrungen gebracht werden kann, sondern absolut disruptiv sein wird. In umso fantastischeren Bildern wird dieser *Event Horizon* mythisch ausgemalt: Mal wird die Lösung der Konvergenz von Mensch und hyperintelligenter Maschine in einem Endkampf der beiden gesehen – bei welchem den Menschen meist

31 Ebd., 12–13.

32 Vgl. ebd., 14.

33 Vgl. Anders Sandberg, „An Overview of Models of Technological Singularity“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 376–394. Hier 377–378.

keine guten Aussichten beschieden werden. Andere hoffen auf eine technisch induzierte Allversöhnung zwischen Mensch und Maschine, in der Biologie und Technologie immer weiter verschmelzen und in der Entstehung eines *Cyborg*-Überwesens münden. Die Lösung der Gleichsetzung von Mensch und Maschine, in denen Horror und Hoffnung gleichzeitig vertreten sind, bilden somit den Fluchtpunkt der mythologischen Zukunftserwartungen. Die Mythen verbindet ihre Radikalität sowie die Überzeugung, dass der schicksalhafte Schauplatz, auf dem der Knoten der Geschichte zwischen dem Menschen und seiner Technologie auseinandergedht, das Forschungsfeld der Künstlichen Intelligenz ist.

Die ethische Relevanz der KI-Mythen

Im Horizont vielfältiger Prognosen darüber, welchen Einfluss Künstliche Intelligenz auf die Gesellschaft wie die individuelle Lebensführung einnehmen wird, sind gesellschaftlich hochwirksame Erzählungen entstanden, in denen diesen extremen Potentialerwartungen ein hohes Maß an öffentlicher Aufmerksamkeit zukommt. Selbst wenn der eigentliche Trägerkreis dieser Konvergenzerzählungen vergleichsweise gering ist, steht dem ihre überproportionale mediale Präsenz gegenüber. Sie sind in den wichtigen Debatten um Entwicklung und Implementierung smarterer Systeme gegenwärtig, weil ihre Szenarien es geschafft haben, den Eskalationshorizont um den Einsatz und die Möglichkeiten der Technologie abzustecken. Experten kommen oft gar nicht umhin, sich zu diesen Konvergenzerzählungen bzw. der Möglichkeit und Folgen von Hyperintelligenz zu positionieren.

Häufig konfundiert dabei die ethische Relevanz dieser Erzählungen mit der Perspektive der Technikfolgenabschätzung. Dabei erschöpft sich der ethische Ansatz gerade nicht in der Frage, ob solche hochintelligenten Systeme ihrem Vollsinn nach tatsächlich im Bereich des potentiell Machbaren liegen, selbst wenn das technisch Wahrscheinliche eine wichtige Referenz deskriptiver Ethik bleibt.³⁴ Das ethische Problem in

³⁴ Für den Theologen Trutz Rendtorff erfordert die Komplexität der verschiedenen Lebenswelten für die Ethik unweigerlich eine Kenntnis der Fakten und Klarheit der Begriffe. Nur so lasse sich ermitteln, wo in diesem Rahmen verantwortliches Handeln ansetzt. Vgl. Trutz Rendtorff, *Ethik. Grundelemente, Methodologie und Konkretionen einer ethischen Theologie*, 3. Aufl., Tübingen 2011, 144–147.

der Untersuchung dieser Mythen liegt weniger in den Vorstellungen über die mögliche Leistungsfähigkeit der Maschinen, sondern – das ist die leitende These dieser Untersuchung – in ihren Aussagen über das Unvermögen des Menschen. Die Konvergenzerzählungen der KI treten, unabhängig eines guten oder schlechten Endes, als Überwältigungserzählungen auf, in denen der Mensch als ethisches Subjekt kaum vorkommt und das Handeln der Maschinen in den Vordergrund rückt.³⁵ Darin entpolitisieren diese Erzählungen den Umgang mit Künstlicher Intelligenz und stellen sich gegen andere, weniger medial wirksame Sinnbilder, die gesellschaftliche wie politische Besinnung fordern, um vorhandene und sich konkret abzeichnende Anwendungen von KI in gegenwärtige Lebenspraktiken zu integrieren und somit auf zwar kleinschrittige, aber dafür tatsächliche Bewältigung zielen.

Dass diese ethische Perspektive in den öffentlichen Debatten nicht zur Geltung kommt, hat auch mit den Immunisierungsstrategien zu tun, die diesen Mythen zu eigen ist und mit der sie sich immer wieder der Kritik entziehen. Aus der Konfusion von Ethik und Technikfolgenabschätzung heraus wird oft voreilig die Schlussfolgerung gezogen, man müsse den Szenarien rein technisch begegnen; in Anbetracht einer möglichen Hyperintelligenz sei Wissenschaft und Technik noch viel stärker als Problemlöser gefragt. Noch mehr Ressourcen gehörten für Forschung und Technik bereitgestellt, damit sie das sog. *Kontrollproblem* einer kommenden Hyperintelligenz löse – ohne zu erwähnen, dass es sich ja um durch Technik überhaupt erst noch ausstehende, hypothetische Probleme handelt. Die zweite Form der Immunisierung sitzt tiefer, als die mythologische Rede zunehmend nicht mehr als solche markiert wird. Die Visionen von einer Substitution des Menschen durch die Maschine lassen nicht mehr erkennen, dass sie die Lösung für ein mythisches Problem der ihr vorangegangenen und für selbstverständlich erklärten Mythen stehen, sie gewissermaßen Mythen zweiten Grades sind. Der Mythos, der sich jedoch selbst nicht mehr als Mythos

35 Darin rücken sie gerade von der eigentlichen Funktion des Mythos ab: Laut Blumenberg dient der Mythos ursprünglich ja gerade der Selbstbehauptung des Menschen gegenüber der Welt. Vor dem absoluten Anspruch der Wirklichkeit soll der Mythos die Angst vor der unbeherrschbaren Wirklichkeit und ihrem Eindruck der Willkür nehmen, um somit zur Handlung zu ermutigen. Vgl. Blumenberg, Arbeit am Mythos, 50.

versteht, gerät leicht zur Ideologie. Die Idee von einer der menschlicher Sinndeutung enthobenen, objektiven und autonom-agierenden Technosphäre, auf die die Menschheit im Ganzen unausweichlich zusteure, ist ein falsches Bewusstsein, da der Mythos seine Kontingenz, seine eigene kulturelle Genese, verleugnet. Diese Arbeit stellt sich in der Analyse dieser Mythen deshalb bewusst gegen diese Autoimmunisierung und ihre Illusionen der Unvergleichlichkeit und Bearbeitungslosigkeit. Entgegen ihrer vermeintlichen naturwissenschaftlich-evolutionären Rationalität stehen diese Technikzukünfte um KI nicht einfach fest, sondern sie sind als kulturelle Produkte einer gewachsenen geistesgeschichtlichen Tradition und damit Ausdruck von Weltanschauung sichtbar zu machen.

In dem Zurückholen der Debatte aus dem Nimbus des Absoluten, das heißt der vermeintlich absoluten Rationalität, der absoluten Geschichte, der absoluten Anthropologie und den daraus resultierenden absoluten Anforderungen und Handlungszwängen tritt auch das spezifisch theologische Profil der Untersuchung zutage. Ideologiekritik ist der Theologie nicht nur aus ihren eigenen Erfahrungen aus der Ideologisierung des Religiösen an die Hand gegeben, sondern hat sich in Besonderem in Auseinandersetzung mit den politischen Ideologien des 20. Jahrhunderts entwickelt. Während diese klar mit dem Anathema von Barmen belegt sind, fällt die konstruktive Kritik in den Bereichen der Digitalisierung, *Big Data* und KI noch schwer; auch da Technologie „vermeintlich ideologiearm auf leisen Sohlen daherkommt“³⁶. Theologische Ethik, die sich dem reformatorischen Erbe der Ausdeutung einer existentiellen Freiheitserfahrung des Menschen verpflichtet sieht, begegnet jedoch jeder Verabsolutierung weltlicher Strukturen mit dem Maß, das zwischen Letztem und Vorletztem (Bonhoeffer), Absolutem und Relativem (Ritschl), Gottes Werk und Menschenwerk (Barth) unterscheidet.³⁷ Die fortschreitende Entropie von Sakralität und Profanität in der Moderne lässt die technische Kultur nicht aus. Besonders

³⁶ Vgl. Peter Dabrock, Geheimnis, Freiheit, Verzeihen: Warum Big Data an die Lehre von der Vorsehung erinnert, https://zeitzeichen.net/archiv/2014_November_theologie-der-big-data (2.8.2019).

³⁷ Johannes Fischer, *Theologische Ethik*. Grundwissen und Orientierung (Forum Systematik Bd. 11), Stuttgart 2002, 39–44.

da die technische Kultur aber wie keine andere Kultursphäre den Menschen der Gegenwart bestimmt, ist ihre Stellungnahme zur Welt in besonderer Weise der Rechenschaft pflichtig.

Konstruktive Kritik bedeutet, die Konvergenzmythen der Künstlichen Intelligenz neben ihren ideologischen Absolutionstendenzen auch als Vorschläge sinnhafter Weltverortung ernst zu nehmen. Wie die klassischen Religionen verhandeln die Konvergenzmythen die menschliche Existenz in ihren grundlegenden Beziehungen: gegenüber sich selbst und der Welt, gegenüber der Gemeinschaft, als auch gegenüber der Geschichte der Menschheit als Ganzer. Die anthropologischen Diskurse, die in ihnen verhandelt werden, stehen deshalb den religiösen und theologischen Sachverhalten nicht konträr gegenüber oder haben diese abgelöst, sondern sie lassen sich in Beziehung setzen und gegenseitig ausdeuten.³⁸ Theologische Ethik in der Tradition ethischer Theologie geht in dem Sinne über die Kritik hinaus hin zur Konstruktion: Im Prozess der wechselseitigen Auslegung erfahren die Mythen der Künstlichen Intelligenz im Licht der wesentlichen Einsichten des reformatorischen Glaubens neue Perspektiven, als sich auch das Christentum in den neuen Sinnbildern gelingenden Lebens zu erweisen hat.³⁹ In dieser Wechselseitigkeit hält der christliche Lehrbestand in seiner ethischen Aufarbeitung den hermeneutischen Schlüssel in der Hand, um mit seinen erfahrungsgesättigten Formeln menschlicher Selbstverortung *sub specie Dei* den existentiellen Vergewisserungen der KI-Mythen Kontur zu verleihen. Szenarien von einer Übernahme des Weltregiments durch die KI und die (un-)freiwillige Verantwortungsabgabe des Menschen lassen sich in Beziehung setzen zur Rede von der guten wie gefallenen Schöpfung und dem Menschen als dem zur Verantwortung beauftragten Geschöpf. Die informationstechnologischen Versprechen, das problematische Verhältnis von Menschen untereinander in der Übertragung menschlicher Bedürfnisse auf maschinelle Surrogate zu versöhnen, berühren christliche Vorstellungen gelingenden gemeinschaft-

38 Vgl. Clemens Wustmans, »Homo Deus« oder Menschsein in Grenzen und Relationalität?, *Zeitschrift für Evangelische Ethik* 65 (2021), 46–51. Hier 51.

39 Vgl. Reiner Anselm/Stephan Schleissing, »Einführung. Zum Ort der »Ethik« im Werk Trutz Rendtorffs«, in: Reiner Anselm/Stephan Schleissing (Hgg.), *Ethik. Grundelemente, Methodologie und Konkretionen einer ethischen Theologie*, Tübingen 2011, XI–XXII.

lichen Leben in der Rede von der Nächstenliebe bei der gleichzeitigen Einsicht, als Sünder an den Bedürfnissen seines Nächsten zu scheitern. Die Motive des *Cyborgs* und des *Mind Uploads* schließlich machen das Erlösungsmoment in der Künstlichen Intelligenz stark, indem sie dem Menschen eine unbeschwerte, neue Heimat im Digitalen versprechen; das Ende der Geschichte des biologischen Menschen stellt Nachfragen an die Geschichte des Menschen im Heilsplan Gottes, die unter dem Zuspruch einer den Leib verwandelnden Erlösung steht.⁴⁰

Diese Untersuchung setzt sich zum Ziel, die Auslieferung des Menschen an die Zukunft der Künstlichen Intelligenz als kontingenten Mythos durchsichtig zu machen und entgegen diesen Erzählungen die individuelle wie gesamtgesellschaftliche Besinnung im Umgang mit den neueren Methoden der Künstlichen Intelligenz einzufordern. Die Mythen funktionieren als Nebelkerzen, indem sie die gesellschaftliche Auseinandersetzung um Künstliche Intelligenz besetzen und der Öffentlichkeit nicht zu bewältigende Szenarien vorlegen. Sie halten die Öffentlichkeit in genau jener Stasis, in der sie den Menschen in ihren Erzählungen sehen, und erschweren die tatsächlich notwendige Auseinandersetzung darum, in welchem Maße konkrete technische Anwendungen zum Einsatz kommen sollen. Darin stützen sie das gegenwärtige Machtungleichgewicht zwischen den globalen Tech-Unternehmen gegenüber Staaten und ihren Bürgern. Indem die Ethik die Technik als Ausdruck menschlicher Freiheit wieder ins Zentrum rückt, wird gleichsam deutlich, dass die Ursachen für den Reiz dieser Mythen auch das Symptom einer Gesellschaft zu sein scheinen, die sich bereits gegenwärtig als technologisch eingengt, möglicherweise überfordert, und wenig selbstwirksam wahrnimmt.

Aufbau der Arbeit

Das folgende Kapitel Zwei schließt an das bisher Umrissene mit epistemologischen Vorüberlegungen sowohl zu dem Betrachterstandpunkt dieser Untersuchung als auch zum Forschungsgegenstand selbst an. Es macht darauf aufmerksam, dass der theologische Blick auf Technikdis-

⁴⁰ Vgl. Reiner Anselm, „Politische Ethik“, in: Wolfgang Huber/Torsten Meireis (Hgg.), *Handbuch der Evangelischen Ethik.*, München 2015, 195–263. Hier 238.

kurse – in deren Tradition die Digitalisierung und mit ihr die KI steht – in den letzten zweihundert Jahren zahlreichen Wendungen unterworfen war, aus denen heraus die Annahme, dass Technologie und die mit verbundenen Wünsche und Ängste ein Thema theologischer Ethik sei, der Klärung bedarf. Zum anderen werden aktuelle sozialwissenschaftliche, ethische und religionswissenschaftliche Theorien herangezogen, um die Konvergenzmythen tiefergehend für die Untersuchung zu bestimmen.

Kapitel Drei nimmt die Forschung um Künstliche Intelligenz als Projektionsfläche der Mythen in den Blick. Das Phänomen der KI wird in ihre übergeordnete Perspektive der Digitalisierung gebettet und die bestehende wirtschaftliche wie wissenschaftliche Ungleichheit zwischen Herstellern und Nutzern in Bezug zur Macht der Mythen gesetzt, um im Anschluss der Frage nachzugehen, inwieweit die derzeitigen Verfahrensweisen des maschinellen Lernens tatsächlich Anlass zur Spekulationen über Künstliche Superintelligenz geben.

Kapitel Vier bindet die Konvergenzerzählungen und ihre Gleichsetzung von Mensch und Maschine an drei ideengeschichtliche Zugänge zurück: Der literarische Zugang verläuft über die ersten technologischen Utopien weiter zu den frühen *scientific romances*, in denen sich das Ideal der höchsten Beherrschung aller Naturkräfte mit dem Ziel verbindet, ein künstliches Wesen zu schaffen, und die Maschine vermehrt als zwielichtiges Gegenüber des Menschen in Erscheinung tritt. Der wissenschaftsgeschichtliche Zugang zeichnet die Idee vom Menschen als Maschine nach und beginnt dabei in der Zeit der europäischen Aufklärung, in der erstmals ein solcher Vergleich gezogen wurde und führt weiter in die kybernetische Bewegung in der Mitte des 20. Jahrhunderts, die in ihrer Reduktion der Welt in Information und damit des Menschen zur biologischen Maschine den letzten Versuch einer wissenschaftlichen Großtheorie unternahm. Zuletzt werden zwei Konvergenzerzählungen des 20. Jahrhunderts vorgestellt, die in ihrer Entgegensetzung von Katastrophe und Entrückung bei gleichzeitigem Auflösen des Menschen in seine technologische Umwelt wie Blaupausen für die gegenwärtigen KI-Mythen erscheinen. Gleichsam legen sie im Vergleich die geschichtsphilosophische Spannung der KI-Erzählungen zutage, die zwischen menschlichen Allmachtsphantasien

und dem vergeblichen Auflehnen des Menschen gegen das technologisch Unvermeidliche oszillieren.

Kapitel Fünf widmet sich der Analyse der Konvergenzmythen vor dem Hintergrund christlicher Anthropologie. Strukturell geschieht dies in der ethischen Aufarbeitung des apostolischen Glaubensbekenntnisses, an dem sich die Selbstverhältnisse des Menschen entlang der Grundbeziehungen zur Welt, zum Mitmenschen und zur Geschichte ausdeuten. Entsprechend werden die Konvergenzmythen in drei Erzählmustern gebündelt und auf diese Perspektiven bezogen: Das Konfrontationsmotiv findet sein christliches Gegenüber im Topos der Schöpfung, das Vertretungsmotiv in der Versöhnung und das Identifikationsmotiv in der christlichen Rede von der Erlösung.

Kapitel Sechs tritt noch einmal einen Schritt zurück, um die bleibende Relevanz utopischer Fiktionen als Ausdruck menschlicher Kreativität, besonders auch im Bereich der Technologien, anzuerkennen. Aus den Ergebnissen der Untersuchung heraus stellt sie die politische Anfrage, wem das Entwerfen der Zukunft zu überlassen ist, und adressiert dabei die theologische Anfrage, wie man dem diffusen Gefühl einer Auflösung des Menschen inmitten einer technisierten Welt begegnet.

2 Vorüberlegungen

Die neuen Informationstechnologien wie auch Technik im Allgemeinen scheint vielen kein Thema theologischer Auseinandersetzung. Die erste Vorüberlegung will den gewählten Betrachterstandpunkt einer theologisch-ethischen Untersuchung historisieren und dadurch ihre Argumente für eine theologische Technikethik offenlegen, die ihr Fundament in einer Rechenschaftspflicht der technischen Kultur gegenüber ihrem Freiheitswert findet. Dem folgt eine interdisziplinäre Zuwendung zum Forschungsgegenstand, in dem die Mythen der Künstlichen Intelligenz unter Hinzunahme sozialwissenschaftlicher, ethischer und religionswissenschaftlicher Sichtweisen als gesellschaftlich hochwirksame Erzählungen hervortreten.

2.1 Betrachterperspektive: Wie wird Technik zu einem Gegenstand konstruktiver theologischer Auseinandersetzung?

Die komplizierte Beziehung zwischen Technik und Religion zeigt sich bereits in dem Streit darüber, ob beide Bereiche überhaupt etwas miteinander zu tun haben; für manche hat Technik dann lediglich mit dem Irdischen zu tun, während sich Religion und Theologie allein mit dem Überweltlichen beschäftigen. Setzt sich Theologie dennoch mit Technik auseinander, steht dem der Vorbehalt gegenüber, die Religion blicke als ehemals dominante Kultursphäre argwöhnisch auf die Technologie, welche sich spätestens seit dem 19. Jahrhundert an ihre Stelle gesetzt habe. Einen solchen Verdrängungskampf zwischen Technik und Religion bringt die Theologin Dorothee Sölle (1929–2003) folgendermaßen auf den Punkt:

„Die Religion wird ausgehöhlt, weil Gott in der technisierten Welt mit wachsender Geschwindigkeit Terrain verliert. Es entsteht der Eindruck,

als sei Gott arbeitslos geworden, weil ihm die Gesellschaft einen Lebensbereich nach dem andern abnimmt.⁴¹

Nicht selten folgt einer solchen Betrachtung die theologische Entgegensetzung von der dämonischen Technik einerseits im Gegensatz zu Gott und religiöser Humanität andererseits. In eine solche Apologetik drohen auch aktuelle Stimmen zu fallen, die die Digitalisierung als Ganzes unter den Generalverdacht eines digitalen Turmbaus zu Babel stellen und ihr unisono den massenhaften Rückfall zur Unmündigkeit bescheinigen.⁴² Landesbischof Volker Jung bringt den Konflikt auf den Punkt, wenn er klagt, er erlebe innerhalb der Kirche einerseits den Anspruch, die Kirche möge eine Art digitalfreie Gegenwelt erhalten, während andere Mitglieder gerade mehr digitale kirchliche Angebote einforderten.⁴³ Die Kommunikation zwischen Theologie und Technik bzw. Naturwissenschaft scheint über die Kirchen hinaus auch im öffentlichen Raum erschwert. Technik kommt selten in christlichen Deutungen der Gegenwart vor und wenn, dann stoßen theologische Anliegen bezüglich der technisierten Lebenswelt häufig auf Unverständnis oder bleiben ohne Einfluss.⁴⁴ Nicht zuletzt aus den eigenen Reihen wurde und wird der neueren Theologie eine Konfliktscheue vorgeworfen, dass sie sich im Ausgang des 20. Jahrhunderts „in den schützenden Bereich der Geisteswissenschaften zurückgezogen“⁴⁵ habe, anstatt Kriterien christlicher Ethik für technischen Handelns und wissenschaftliche Weltgestaltung einzubringen.

Zugleich scheint der Austausch mit den Geisteswissenschaften in Politik und Gesellschaft vermisst zu werden. Besonders seit dem Auf-

41 Dorothee Sölle, *Stellvertretung*. Ein Kapitel Theologie nach dem Tode Gottes, Stuttgart 1982, 176–177.

42 Vgl. Werner Thiede, „Digitaler Turmbau zu Babel“, in: Severin Lederhilger (Hg.), *Gott und die digitale Revolution*, Regensburg 2019, 47–72. Hier 49.

43 Vgl. Volker Jung, *Digital Mensch bleiben*, München 2018, 72.

44 Vgl. Ralph Charbonnier, *Technik und Theologie*. Ein theologischer Beitrag zum interdisziplinären Technikdiskurs unter besonderer Berücksichtigung der Theologie F.D.E. Schleiermachers (Marburger theologische Studien 77), Marburg 2003, 2.

45 Heinz Eduard Tödt, „Ambivalenz des Fortschritts. Zur Urteilsfindung über das Wesen der Macht“, in: Wolfgang Lienemann (Hg.), *Fortschrittsglaube und Wirklichkeit*. Arbeiten zu einer Frage unserer Zeit, München 1983, 141–156. Hier 152.

stieg der Lebens- wie Informationstechnologien in den 2000er Jahren steigt die Nachfrage an philosophischer als auch spezifisch religiöser Orientierung. Ethische Expertise ist seither in Forschungsprojekten unter dem Banner der ELSI-Forschung (*Ethical, Legal and Social Implications*) zugegen und wirkt über verschiedenste Gremien auch in mediale Debatten hinein. Auf diesen gestiegenen Orientierungsbedarf durch neue, transformative Technologien reagierte die Politik schließlich auch mit der Gründung des Deutschen Ethikrates im Jahr 2006, in dem auch die Kirchen vertreten sind. Von theologischer Seite jedenfalls helfen weder Flucht in die eigene Disziplin noch apologetische Generalkritik, um konstruktiv mit dieser Anfrage umzugehen. Ein dritter Zugang, der erstmals deutlich in der Frühindustrialisierung artikuliert wurde und an den die neuere Theologie kritisch wieder anzuknüpfen sucht, wirbt dafür, die Technik als aufsteigende Lebenssphäre des Menschen mit seiner Religiosität zu verknüpfen und technologischen Fortschritt konstruktiv in die christliche Weltansicht einzubinden.

2.1.1 Technikzu- und abwendung im Protestantismus des 19. und 20. Jahrhunderts

Technik war für die längste Zeit weder ein Thema der Theologie noch der Ethik, da vor dem industriellen Zeitalter die Natur und ihre Manipulation kein Gegenstand menschlicher Verantwortung war. Weil man über die Natur nicht in dem umfassenden Maße verfügen konnte, war nicht Ethik, sondern Klugheit ihr gegenüber angebracht.⁴⁶ Die Industrialisierung, in der sich die Nutzung der Naturkräfte massiv steigerte und die die Lebensverhältnisse für eine breite Bevölkerungsschicht radikal umstellte, führte in der theologischen Reflektion zu einer verstärkten Beschäftigung mit Technik und der Frage ihrer Funktion in der christlichen Heilsgeschichte.

Der Theologe Friedrich Wilhelm Schleiermacher (1768–1834) findet einen positiven Zugang zum sichtbaren Ausgriff der Technik auf die Lebenswelt, indem er sie in den Dienst der christlichen Sache

⁴⁶ Vgl. Hans Jonas, *Das Prinzip Verantwortung*. Versuch einer Ethik für die technologische Zivilisation, 7. Aufl., Frankfurt am Main 1987, 21.

stellt. Schleiermacher erwartet die Vollendung der Wissenschaften und Künste in der Beherrschung aller irdischen Kräfte durch Technologie. Wo der Mensch seiner Zeit noch unter der Last der Arbeit eingebunden sei, erhofft Schleiermacher von der Technik, dass sie die Menschen von den Mühen der Produktion befreit. Im Zuge des naturwissenschaftlich-technischen Fortschritts gestalteten sich die Menschen die Welt zu nichts Geringerem als einem Feenpalast, in welchem nicht mehr die wenigen Privilegierten, sondern alle Menschen die Ruhe und Muße finden, ihrer Religiosität und damit den grundsätzlichen Fragen nachzugehen.⁴⁷

In Schleiermachers Nachfolge gehört Richard Rothe (1799–1867) zu den glühendsten theologischen Stimmen der Verbesserung der *conditio humana* durch technischen Fortschritt. Keine These verbindet sich mehr mit seiner Person als die Aussage, dass der Erfindung der Dampfmaschine ein größerer Wert für das Reich Christi zukomme als allen großen altkirchlichen Konzilien zusammen.⁴⁸ Für Rothe ist die christliche Religion nicht nur an jede Art von wissenschaftlichem Fortschritt anpassungsfähig, sondern die christliche Wahrheit bleibe „mit jeder neuen Entwicklung der Welterkenntnis in Harmonie [...], und zwar so, dass es selbst, je weiter diese fortschreitet, mittelst ihrer in ein immer klareres Licht gestellt, zu einem immer vollständigerem Verständnis gebracht wird.“⁴⁹

Ähnlich wendet sich zur Jahrhundertwende Friedrich Naumann (1860–1919) in seelsorgerlicher Absicht an die Fabrikarbeiter seiner Zeit und sichert ihnen zu, dass auch sie am Reich Gottes mitarbeiteten, ihr Beruf – in lutherischer Denktradition – ein Gottesdienst sei:

47 Vgl. Friedrich Schleiermacher, *Über die Religion*. Reden an die Gebildeten unter ihren Verächtern (1799) (de Gruyter Texte), 2. Aufl., Berlin 2001, 158.

48 „Ich lebe allerdings der festen Überzeugung, dass dem Reiche Christi die Erfindung der Dampfmaschinen und der Schienenbahnen eine weit bedeutendere positive Förderung geleistet hat, als die Ausklügelung der Dogmen von Nicäa und von Chalcedon.“ Richard Rothe, *Stille Stunden*, Wittenberg 1872, 344–345.

49 Ebd., 345.

„Die Maschine ist nichts Unchristliches, denn Gott will sie. Gott redet zu uns durch die Tatsachen der Geschichte. In Tatsachen hat er seit Jahrzehnten lauter und lauter zum Christenvolk gesprochen: Ich will die Maschine. Wer konnte sie aufhalten, da Gott sie wollte? Alle Seufzer der Christen haben den Fortschritt nicht hemmen können. Sie kommt, sie kommt, die neue Zeit, sie kommt von Gott.“⁵⁰

Nur in dem Glauben, auch in der Verausgabung an der Maschine gott-dienlich zu sein, erhalte die Arbeit an und mit der Maschine einen religiösen Wert. Gott sei deshalb „in dem Hochofen nicht weniger allgegenwärtig [ist] als im Hain Mamre. Gott ist im modernen Getriebe.“⁵¹

Die positive Zuwendung der Theologie hin zur Technik, wie sie im Laufe des 19. Jahrhunderts geschah, trug bereits den Gedanken in sich, dass Technik der Kultur nicht gegenübersteht, sondern ein Teil von ihr ist. Der naturwissenschaftlich-technische Fortschritt der Menschheit ist in der Ethik Schleiermachers das Resultat einer immer tiefergehenden Durchdringung und Zähmung der Natur durch die Vernunft.⁵² In dieser Idee der Zähmung der Naturkräfte wie der gesellschaftlichen Natur findet die liberale theologische Tradition ihre Grundlage einer Kulturtheorie. Die Überzeugung, dass diese Durchdringung der Vernunft einen Fortschritt verbürgt, teilt sie dabei mit dem Wissenschaftsoptimismus der Neuzeit.⁵³ Schleiermacher, Rothe und Naumann reagieren mit ihrer positiven Verknüpfung von Technik und Religion zudem auf das sich durchsetzende neue Geschichtsbild der Naturwissenschaften: Die finite Heilsgeschichte mit der Trias von der Schöpfung der Welt durch Gott, dem Kommen Christi und dem Warten auf die Vollendung erfuhr durch den technologischen Fortschritt eine Entgrenzung. Mit dem offenen Weltbild bildete sich die Möglichkeit einer – wie Naumann

50 Friedrich Naumann, „Jesus, der Sozialist“, in: Wolfgang Philipp (Hg.), *Der Protestantismus des 19. und 20. Jahrhunderts.*, Wuppertal 1988, 222–224. Hier 224.

51 Friedrich Naumann, „Im Eisenwerk“, in: Wolfgang Philipp (Hg.), *Der Protestantismus des 19. und 20. Jahrhunderts.*, Wuppertal 1988, 213–218. Hier 216.

52 Vgl. Friedrich Schleiermacher, *Ethik.* mit späteren Fassungen der Einleitung, Güterlehre und Pflichtenlehre, Hamburg 1981, 48.

53 Vgl. Reiner Anselm, „Die Ethik im Kreis der theologischen Fächer“, in: Markus Buntfuß/Martin Fritz (Hgg.), *Fremde unter einem Dach? Die theologischen Fächerkulturen in enzyklopädischer Perspektive*, Berlin/Boston 2014, 115–125. Hier 117–118.

sie proklamiert – neuen Zeit, deren Geschichte der Mensch durch die Erfindung und Konstruktion von immer neuen technischen Artefakten hervorbringt.⁵⁴ Die liberalen Theologen des 19. Jahrhunderts fassen diese Zeiterfahrung in die Geschichte Gottes ein und deuten die Zeit der Maschinen zu einer gottgewollten und nicht gottverlassenen Zeit. Zugleich stellen die Theologen damit klar, dass die neue Zeit keine von Gott befreite Zeit ist, sondern weiterhin Gott – und nicht der Mensch oder die Technik – die Geschichte lenkt.⁵⁵

Diesen technizugewandten Protestantismus trägt Ernst Troeltsch (1865–1923) weiter in das 20. Jahrhundert. Sein besonderer Verdienst liegt darin, den Wert der Ethik für dieses Vorhaben zu konkretisieren. Bereits Schleiermacher verstand die Ethik im Sinne einer Güterlehre, deren vollständiger Ausdruck in der Einheit von Vernunft und Natur und damit im organischen Zusammenhang aller Güter bestehe.⁵⁶ Zu diesen Gütern zählt Troeltsch neben der Technik und anderen nun jedoch auch die Religion selbst. Damit verabschiedete er sich von Schleiermachers Vorstellung, die christliche Religion und die Kirche seien das alleinige ideelle Fundament einer Gesellschaft, und vollzieht so den Schritt zu einer modernen Kulturtheorie. Die verschiedenen kulturellen Güter unterteilt er dabei in weltliche Zwecke – hier ist die Umgestaltung der Welt durch Technik einzuordnen – und religiöse Zwecke. Das ideale Gleichgewicht zwischen Weltgewandtheit und Weltflucht innerhalb der Kultur herzustellen, darin liege die konkrete Aufgabe der Ethik, da nur in dieser Harmonie sich die einzelnen Lebensbereiche weiter humanisierten und perspektivisch dem Reich Gottes annäherten.⁵⁷ Im Glauben an die gute Schöpfung und den Wert der Arbeit hielt Troeltsch es zumindest für möglich, „in den absoluten

54 Vgl. Stephan Schleissing, „Laien, Experten, Propheten: Zur Rolle von Theologen in Technikdiskursen“, in: Anne-Maren Richter/Christian Schwarke (Hgg.), *Technik und Lebenswirklichkeit*. Philosophische und theologische Deutungen der Technik im Zeitalter der Moderne, Stuttgart 2014, 203–216. Hier 207–208.

55 Vgl. ebd., 209–210.

56 Schleiermacher, *Ethik*, 220/(552).

57 Vgl. Ernst Troeltsch, „Grundprobleme der Ethik. Erörtert aus Anlass von Hermanns Ethik“, in: Hartmut Spenner (Hg.), *Ethik / Grundprobleme der Ethik.*, Waltrop 2002, 44–94.125–178. Hier 141.

Zweck der Gottesgemeinschaft die innerweltlichen Zwecke [darin auch die Technik] aufzunehmen⁵⁸.

Dieser Vorschlag einer harmonischen Güterlehre gerät unter dem Eindruck der großen Weltkriege in eine Krise. Das bis dahin ungekannte Ausmaß menschlichen Tötens, für das paradigmatisch die Entwicklung des Panzers, die Wasserstoffbombe und die Errichtung von Konzentrationslagern stehen, brechen die positive Verknüpfung von technologischem Fortschritt, Humanität und Religion. Die Brutalität der modernen Kriegsführung führt zu einer allgemeinen Kulturkritik, die die Technik entweder ignoriert oder sie höchstens als Gegenüber eines wahren humanen und religiösen Lebens thematisiert.⁵⁹ Den Anhängern der Dialektischen Theologie, der einflussreichsten Strömung der nachkriegsdeutschen Theologie, wird in Reaktion alles Menschengemachte suspekt und läuft Gefahr einer Hybris vor Gott. Die christliche Offenbarung allein halten sie als übriggebliebene Option gegenüber einer gefallenen Kultur hoch. Und auch über die Theologie hinaus überwiegen die kritischen Stimmen gegenüber dem sittlichen Wert der Technik: Martin Heidegger (1889–1976) spricht von der Technik als Gestell, Theodor Adorno (1903–1969) und Max Horkheimer (1895–1973) beobachten die technikinduzierte Umwandlung der Rationalität in menschenverachtende instrumentelle Vernunft. Sie schließen damit an Hanns Liljes (1899–1977) und Herbert Marcuses (1898–1979) Analyse an, dass das technische-rationale Denken die Vernunft selbst umforme und sich in ihrer dehumanisierten Form auf alle Lebensbereiche ausbreite.⁶⁰ Immerzu wiederholt sich im Nachkriegsdeutschland die Idee, dass die die Technosphäre auf den Menschen auszugreifen suche, wogegen sich das angefochtene *humanum* behaupten müsse. Diese kritisch-distanzierte Haltung wiederholt sich bis weit gegen Ende des 20. Jahrhunderts und kommt beispielsweise noch in den späten

58 Ebd., 169.

59 Vgl. Christian Schwarke, „Einleitung“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 9–17. Hier 12.

60 Vgl. Herbert Marcuse, „Einige gesellschaftliche Folgen moderner Technologie“, in: *Schriften*. Band III: Aufsätze aus der Zeitschrift für Sozialforschung 1934–1941, Frankfurt am Main, 286–319. Hier 290–292.

Worten Carl Friedrich von Weizsäckers (1912–2007) zum Vorschein, wenn er konstatiert, dass Verantwortung in der technischen Welt heiße, inmitten der Planung und Apparate noch Mensch zu bleiben, beziehungsweise ihr gegenüber erst Mensch zu werden.⁶¹

2.1.2 Technisches Handeln ist Ausdruck menschlicher Freiheit im Raum der Kultur

Gleichwohl regte sich im Nachkriegsdeutschland durchgehend Kritik an dieser Generalabsage. Der deutsche Psychologe Arnold Gehlen (1904–1976) beschwert sich über die deutsche Tradition der vehementen Technikkritik und verweist auf die USA, wo man sich seiner Einschätzung nach in Form der Science-Fiction-Literatur dem lustvollen Gedankenspiel technischer Wissenschaft zuwende. Er zählt die Technik zur Grundausstattung des Menschen, dieses in seiner Lebensbewältigung auf künstliche Entlastung angewiesenen Mängelwesens.⁶² Ernst Cassirer (1874–1945) rechnet die Technik nicht zur Verfallsgeschichte, sondern sieht wie Walter Benjamin (1892–1940) in der Technik einen sich vollziehenden Befreiungs- und Demokratisierungsprozess. Mit steigender Technologie emanzipiere sich die Menschheit aus ihren natürlichen Beschränkung, da sich im hergestellten Werk ein Weltverhältnis ausdrücke und der *homo faber* dadurch als der im Tätigsein Freie in Erscheinung trete.⁶³ Die Technik eröffne eine eigene und dem Menschen wichtige Form von Freiheit gerade aufgrund ihres instrumentellen Charakters, ihrem Interesse an dem Machbaren und seinem Nutzen. Technik sei somit ein zentraler Ort gestaltender Freiheit innerhalb der Kultur, die sich in Artefakten materialisiert und über

61 Vgl. Carl Friedrich von Weizsäcker, *Die Verantwortung der Wissenschaft im Atomzeitalter* (Kleine Vandenhoeck-Reihe 1142), 7. Aufl., Göttingen 1986, 11.

62 Vgl. Arnold Gehlen, *Die Seele im technischen Zeitalter. Sozialpsychologische Probleme in der industriellen Gesellschaft* (rowohlts deutsche enzyklopädie 53), 104. Aufl., Reinbek bei Hamburg 1976, 7.

63 Vgl. Elisabeth Gräß-Schmidt, „Der Homo Faber als Homo Religiosus. Zur anthropologischen Dimension der Technik“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 39–55. Hier 47.

ihren Erfinder hinaus fortwirkt.⁶⁴ Innerhalb der liberalen Theologie der frühen Nachkriegszeit hat sich Paul Tillich (1886–1965) am intensivsten mit der Technik auseinandergesetzt und ist gegen ihre einseitige Dämonisierung angetreten. Er knüpft an die genannten Restitutionsversuche an, indem auch er die Technik zum Grundbestand menschlichen Lebens zählt. Technisches Handeln kommt für ihn überall da zum Vorschein, wo durch geeignete Mittel etwas gelingen soll; der Wunsch, dass etwas gelingt, sei aber allen Lebenssphären gemein, und demnach keine frei von Technik sei.⁶⁵ Das Eingeständnis einer dauerhaften Präsenz des Technischen muss für Tillich Konsequenzen für die ethische Betrachtung nach sich ziehen:

„Die Technik hat die Welt umgestaltet, und diese umgestaltete Welt ist unsere Welt und keine andere. Auf ihr müssen wir bauen, müssen mehr als bisher die Technik einbauen in den letzten Lebenssinn, wohl wissend: Ist sie göttlich, ist sie schaffend, ist sie befreiend, so ist sie auch dämonisch, knechtend und zerstörend.“⁶⁶

Damit plädiert Tillich erstens für die Anerkennung der Geschichtlichkeit von Technik. Als elementarer Teil der modernen Kultur sei es widersinnig, sich dem Trugbild hinzugeben, man könne einen vermeintlichen technologischen *lapsus* des Menschen ungeschehen machen, sondern das Vorfindliche ist immer auch das Technische, mit dem philosophisch wie theologisch umzugehen ist. Zweitens unterstreicht Tillich, dass die der Technik inwohnende Freiheits- wie Unfreiheitsermöglichung gleichsam auf den Menschen zurückwirkt. Die existentielle Dimension der Technik, ihren *logos*, erklärt er zum

⁶⁴ Vgl. Jörg Dierken, „Technik als Kultur“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 21–37. Hier 21.

⁶⁵ Vgl. Paul Tillich, *Die religiöse Substanz der Kultur*. Schriften zur Theologie der Kultur (Gesammelte Werke / Paul Tillich. Hrsg. von Renate Albrecht; Bd. 9), 2. Aufl., Stuttgart 1975, 297.

⁶⁶ Ebd., 306.

grundlegenden Thema einer theologischen Technikethik.⁶⁷ Erst durch sie werde jede weitere materialetische Untersuchung hellsichtig.

Tillichs Argumentation ähnelt Martin Heideggers Unterscheidung zwischen der Technik im Allgemeinen im Gegensatz zu ihrem verborgenen Wesen. Letzteres unterscheidet Heidegger von der konkret materialen Technik und bestimmt es als ganz und gar nichts Technisches. Dieses Wesen der Technik liege in ihrer entbergenden Funktion; Technik bringe Wahrheit hervor, indem sie Naturkräfte steuert und sichert.⁶⁸ Tillich macht auf die ambivalente Rückwirkung dieses Entbergens der Welt auf den Menschen aufmerksam. Indem die Technik die Welt beherrschbar macht, verliere sie mit ihrer Unheimlichkeit gleichzeitig auch ihren Lebensfülle. Die Maschine sah Tillich zwar als moralisch neutral an, ihren *logos* und damit ihren Eindruck auf die menschliche Seele aber begreife man nur im Ineinander von Macht und Versuchung. In deren Erliegen gehe dem Menschen das Maß für Sinn und Notwendigkeit verloren, wodurch die technische Freiheit in der Gefahr stehe, die Freiheit zu usurpieren und letztlich zu zerstören.⁶⁹ In der Entdämonisierung von Technik restituiert Tillich somit nicht einfach in den religiös-humanistischen Optimismus der liberalen Theologen des 19. Jahrhunderts, sondern macht im Logosgedanken die freiheitermöglichende wie freiheitgefährdende Macht der Technik auf den Menschen gerade zum ethisch-theologischen Anknüpfungspunkt. Nicht das sporadische Urteilen über einzelne Anwendungen von Technik oder gar das Urteilen über ganze Technologiezweige ist geboten, sondern eine deskriptive Ethik für die technische Kultur.

Trutz Rendtorff (1931–2016) knüpft an diese Gedanken in der zweiten Hälfte des 20. Jahrhunderts an und stellt klar, dass außerhalb der Kulturtätigkeit des Menschen keine Technik für sich existiert, sondern Kultur überhaupt erst die Bedingung für Technik liefert. Losgelöst von ihrem kulturellen Kontext sieht er in der Technik keine ethische Ver-

67 „Sinndeutung ist aber nur möglich auf dem Grunde von Wesensschau. Den *logos* der Technik zu fassen, ihr Wesen, ihre eigentümliche Seinsgestalt, ihre Beziehung zu anderen Seinsgestalten, das ist darum auch unsere erste und wichtigste Aufgabe.“ in: ebd., 297.

68 Vgl. Martin Heidegger, *Die Technik und Die Kehre* (Opuscula aus Wissenschaft und Dichtung 1), 9. Aufl., Stuttgart 1996, 12.16.

69 Vgl. Tillich, *Die religiöse Substanz der Kultur*, 305.

bindlichkeit. Kultur bedeutet für ihn Stellungnahme zu dem Gegebenen des menschlichen Lebens. Die Kultur sei der Ort, in dem der Mensch sich besinne und der Aufgabe humaner Weltgestaltung zuwenden könne. Als kulturelles Wesen sei deshalb „der Mensch ständig und immer neu zur Stellungnahme seiner eigenen Kulturtätigkeit aufgefordert. Kultur hat an keinem Punkt den selbstverständlichen Charakter einer natürlichen Umwelt des Menschen“⁷⁰. Ethische Probleme der Technik seien insofern Probleme einer durch Technik bestimmten Kultur, sodass Technikkritik immer Kulturkritik voraussetze.⁷¹

2.1.3 Ethik fordert die technische Kultur zur Rechenschaft über ihren Freiheitswert

Innerhalb der skizzierten liberalen Tradition durchdringen sich Technik und Religion, nicht weil Technik unumstößlich eine Auflehnung des Menschen gegen Gott bedeutet, sondern weil sie einerseits von Wertvorstellungen abhängt, die sie selbst nicht erzeugen kann, und andererseits Technik selbst Sinnüberschuss produziert, indem sie die Grenzen des *status quo* erweitert und sich dem bisher Unverfügbaren zuwendet. Technische und religiöse Stellungnahmen beziehen sich deshalb auch in der Moderne aufeinander, wobei unhintergehbare Transzendenzen neben Metaphern und gezielten Mythologisierungen stehenbleiben.⁷² Indem die Grenzen menschlichen Zugriffes in der Praxis als auch in der Imagination kontinuierlich erweitert werden, schafft die Technik gewaltige Potentiale und Erwartungen für die Realisierung von Freiheit. Technische Freiheit ist zu einem zentralen Motor menschlichen Strebens und Handelns geworden, in der sich seine Neugier und Kreativität verwirklicht.⁷³ Die sich in diesem Erfolg ausdrückende Technisierung der Lebenswelt wirkt gleichwohl auf den Menschen zurück. Hier zeigt

70 Rendtorff, Ethik, 280–286.

71 Vgl. Trutz Rendtorff, „Strukturen und Aufgaben technischer Kultur“, in: Dietrich Rössler (Hg.), *Möglichkeiten und Grenzen der technischen Kultur*. Symposium Hotel Schloß Fuschl, Österreich 1981, 9–21. Hier 10.

72 Vgl. Christian Schwarke, „Technik und Theologie. Was ist der Gegenstand einer theologischen Technikethik?“, *Zeitschrift für Evangelische Ethik* (2005), 88–104. Hier 96–97.

73 Vgl. Gräß-Schmidt, Der Homo Faber als Homo Religiosus, 53–54.

sich die Ambivalenz technischer Freiheit, weil der Erfolg ihres Funktionierens selbst eine Eigendynamik entwickelt, in deren Wechselwirkung der Mensch Treibender wie Getriebener wird. Insofern stellt sich die ethische Frage der Begrenzung der Macht technischen Funktionierens gegenüber anderen kulturellen Gütern um der Freiheit willen:

„Der Mensch gewinnt Macht über viele Dinge – aber gewinnt er auch Macht über seine Macht?“⁷⁴

Technische Freiheit ist trotz ihres Erfolges und ihrer Sichtbarkeit nicht die alleinige bzw. einzig mögliche Verwirklichung von Freiheit. Wie das 20. Jahrhundert eindrücklich bezeugt, ist technischer Machtgewinn nicht mit einem Gewinn tatsächlicher Freiheit von Personen gleichzusetzen, sondern muss diesem Zweck erst dienstbar gemacht werden. Wie jede andere Ausprägung der Kultur als Stellungnahme zur Welt bleibt Technik moralisch offen. Freiheit scheint da bedroht, wo technisches Denken die Beziehung zu den es übersteigenden Sinn-systemen meint abbrechen zu können. In der Orientierung allein an der technischen Machbarkeit wird schließlich das Freiheitspotential des Menschen überstrapaziert.⁷⁵ Max Weber (1864–1920) machte erstmals darauf aufmerksam, dass genauso, wie Religion sich nicht zur Grundlage einer Gesellschaft aufschwingen könne, auch Wissenschaft und Technik diesen Anspruch nicht einlösen. Der Wert von Wissen lasse sich nicht mit den Mitteln der Wissenschaft überprüfen, sondern nur auf ihren letzten Sinn deuten, je nach der eigenen Stellungnahme zum Leben.⁷⁶ Technik als Teil der Kultur will über die Beherrschung der Natur eine Verbesserung der menschlichen Lebensbedingungen erreichen, sie kann das Optimum der Kultur jedoch nicht von sich aus festlegen. Das Kulturelle des Technischen zu sehen, heißt, sowohl ihre Bedürftigkeit von äußeren Wertvorstellungen gelingenden Lebens anzuerkennen, als auch, dass Technik mit ihren Grenzverschiebungen selbst neue Anfragen und Alternativen zu bestehenden Weltanschau-

74 Tödt, *Ambivalenz des Fortschritts. Zur Urteilsfindung über das Wesen der Macht*, 144.

75 Vgl. Gräb-Schmidt, *Der Homo Faber als Homo Religiosus*, 51.

76 Vgl. Max Weber, „Wissenschaft als Beruf“, in: Dirk Kaesler (Hg.), *Schriften 1894 – 1922.*, Stuttgart 2002, 474–511. Hier 494.

ungen aufwirft. Wissenschaft mitsamt Technik und Religion bleiben so im gemeinsamen Raum der Kultur auf das Gespräch angewiesen. Denn auch die Religion zusammen mit der Philosophie und den Künsten, die die weltlichen Zwecke in ihrer Unabgeschlossenheit aufzuzeigen und für das gemeinsame Leben zu harmonisieren suchen, benötigen den Austausch mit den Fachwissenschaften.⁷⁷

Wenn Kultur von vielfältigen Stellungnahmen zu den Gütern des Lebens durchdrungen ist, liegt die Funktion der Ethik darin, die Rechenschaft dieser Stellungnahme im Hinblick auf ihre Freiheitswert einzufordern. Die kritische Zuwendung theologischer Ethik hin zur Technik ist geboten gerade in der Anerkennung der sich durch sie ausdrückenden Freiheitserfahrung. Diese Perspektive stellt die technische Kultur nicht in Konkurrenz zum Schöpfungshandeln Gottes, sondern deutet das von Technik geprägte Leben in einer Beziehung zu Gott, durch die es zur verantworteten Gestaltung aufgefordert ist.⁷⁸ Im Offebarmachen der Eingebundenheit der technischen Freiheit in andere Güter um der Freiheit willen kommt der bleibende Wert einer Gütertheorie – innerhalb einer protestantischen Ethik und darüber hinaus – zum Vorschein. Eine liberale Ethik geht der Frage nach, wie der Wert technischer Freiheit produktiv in eine ganzheitliche Freiheitsperspektive eingebunden werden kann – und was sie daran hindert. Daraufhin sind die Mythen um KI und ihre Vorstellungen eines Optimums der Gesellschaft und menschlicher Freiheit zu prüfen. Dabei geht es auch um eine ständige Selbstprüfung, inwieweit Impulse aus dem Christentum heraus nicht selbst eine Überforderung technischer Freiheit begünstigen. Schließlich lässt sich der exzessive Ausgriff der Technik auf die Natur ebenso als Fortführung einer christlichen Freiheitsauslegung gegenüber einer entgötterten Welt verstehen.⁷⁹

77 Vgl. Reiner Anselm, „Warum Wissenschaft Religion braucht“, in: Werner Matt/Birgit Brida/Wolfgang Ortner (Hgg.), *Das Montagsforum*. Versuche, die Welt zu verstehen; Festschrift anlässlich des 75. Geburtstages von Dkfm. Dr. Heinz Bertolini, Dornbirn 2010, 364–372. Hier 372.

78 Vgl. Rendtorff, *Ethik*, 280–286.

79 Vgl. Dorothee Sölle, *Atheistisch an Gott glauben*. Beitr. z. Theologie, Olten/Freiburg i. Br. 1968, 53.

2.2 Betrachtungsgegenstand: Der Mythos der Künstlichen (Super-)Intelligenz zwischen Technikfolgenabschätzung, Moralfabel und apokalyptischer Vision

Zur weiteren Bestimmung des Betrachtungsgegenstandes nähert sich der folgende Abschnitt den Mythen der Künstlichen Intelligenz unter der Verwendung dreier analytischer Werkzeuge. Ausgehend von Charles Taylors Begriff des sozial Imaginären (*social imaginaries*) sind diese Mythen kein bloßes Gerede, sondern beanspruchen die Durchsetzung ihrer Vorstellungen über den Umgang mit Künstlicher Intelligenz in Politik, Wissenschaft und Wirtschaft. Der literaturwissenschaftliche Terminus der Narration deckt die ethische Bedeutung der Mythen auf, da die in diesen Imaginationen auftretenden Verhältnisbestimmungen von Mensch und Maschine auf etablierte, fiktionale wie nicht-fiktionale Erzählkreise aufbauen und sie verstärken. Schließlich dient die christliche Apokalyptik als Deutungsschablone, um die unter diesen Werturteilen liegende existentielle Ansprache und impliziten Drohungen der Mythen unter ihrem Horizont der Letztgültigkeit offenzulegen.

2.2.1 Technikantizipation und Ressourcenkonflikt: Technosoziale Sinnbilder

Der Politikwissenschaftler und Philosoph Charles Taylor (*1931) verhalf dem Begriff des sozial Imaginären (*social imaginaries*)⁸⁰ Ende des 20. Jahrhunderts nach Vorarbeiten von Cornelius Castoriadis und Jacques Lacan zu einer breiten Rezeption. Darunter fasste Taylor die kollektiv-verbindlichen Grundansichten einer Gemeinschaft in ihrer Menge von Werten, Institutionen, Gesetzen und Symbolen, durch die sich Menschen ihr soziales Ganzes, deren Grenzen und deren Änderungspotential vorstellen. Dieses gemeinsame Verständnis gebe Praktiken innerhalb einer Gemeinschaft Sinn und Legitimität. Das sozial

⁸⁰ Vgl. Charles Taylor, *Modern Social Imaginaries* (Philosophy social theory), 4. Aufl., Durham 2004, 23–26.

Imaginäre sei somit ein gemeinsames Repertoire an Werthaltungen, auf das Menschen beständig zugegriffen. Als implizites Wissen würden diese Grundansichten in Bildern und Geschichten weitergegeben.⁸¹ Darauf aufbauend formulierten Jasanoff und Kim den Begriff des soziotechnisch Imaginären (*sociotechnical imaginaries*).⁸² Dies sind im Spezifischen die Sinnbilder für kollektive, institutionell stabilisierte und öffentlich vorgetragene Visionen wünschenswerter Zukünfte einer sozialen Ordnung, die durch Fortschritte in Wissenschaft und Technologie erreichbar sind und diese unterstützen. Diese Sinnbilder setzten mit ihrer Festigung die Verstehens- und Umgangsbedingungen neuer Technologien, indem sie die Ausarbeitung von Politiken, die Mobilisierung von Industrien und die Zuweisung von Ressourcen prägen. Damit schaffe das Imaginäre auch Notwendigkeiten und Handlungszwänge, da Investitionen nach Rendite verlangten und politische Versprechen erfüllt werden müssten.⁸³ Wie bei dem sozial Imaginären begründet damit nicht jede Aussage über eine mögliche Zukunft ein soziotechnisches Sinnbild, sondern wird erst zu einem solchen, wenn diese Ideen in ein Machen, Regieren und Institutionalisieren überführt werden. Geäußerte Erwartungen an zukünftige Technologien können massive gesellschaftliche Auswirkungen haben und einzelne Akteuren Macht zuspielden, unabhängig davon, ob diese Ideen in einem Verhältnis zur Wirklichkeit stehen.⁸⁴

81 Der Theologe Florian Höhne greift im Kontext der Künstlicher Intelligenz ebenfalls auf Taylors Begriff der *social imaginaries* zurück: Vgl. Florian Höhne, „Darf ich vorstellen: Digitalisierung. Anmerkungen zu Narrativen und Imaginationen digitaler Kulturpraktiken in theologisch-ethischer Perspektive“, in: Jonas Bedford-Strohm/Florian Höhne/Julian Zeyher-Quattlender (Hgg.), *Digitaler Strukturwandel der Öffentlichkeit*. Interdisziplinäre Perspektiven auf politische Partizipation im Wandel, Baden-Baden 2019, 25–46. Hier 31.

82 Vgl. Sheila Jasanoff/Sang-Hyun Kim (Hgg.), *Dreamscapes of Modernity*. Sociotechnical Imaginaries and the Fabrication of Power, Chicago/London 2015, 4.

83 Vgl. Christian Katzenbach, I miss Imaginaries that spell out Ideas of AI as a Public Good, 21.5.2019, <https://www.hiig.de/en/the-global-ai-race-analysing-different-strategies-and-imaginaries/> (14.12.2020).

84 Vgl. Astrid Mager/Christian Katzenbach, „Future Imaginaries in the Making and Governing of Digital Technology: Multiple, Contested, Commodified“, *New Media & Society* (2020), 1–12. Hier 3.

Akteure und Institutionen entgrenzter Sinnbilder der Künstlichen Intelligenz

Die Sinnbilder um Künstliche Intelligenzen präsentieren sich in diesem Sinne ebenfalls nicht als weltlose Ideen, sondern haben weitreichende Trägerkreise ausgebildet. Es sind sowohl eigene Gemeinschaften und Institutionen entstanden als auch Verbindungen zu bestehenden Bewegungen mit überschneidenden Interessen geknüpft worden, die gemeinsam auf die öffentliche Wahrnehmung in Wissenschaft, Wirtschaft und Politik einwirken. Diese Kreise wiederum sammeln sich um einzelne, medienwirksame Akteure mit hohem Sendungsbewusstsein.

Einer der ersten, der gegen Ende der 1980er Jahre die mythologische Gedankenwelt der Künstlichen Intelligenz zurück in das öffentliche Bewusstsein brachte, ist der austro-kanadische Informatiker Hans Moravec. Seine These, der Mensch werde in der Entwicklung hyperintelligenter künstlicher Systeme eine neue Stufe der Evolution erreichen und so seine eigenen Nachfahren schaffen, kursieren seither an den Spitzenuniversitäten der USA. Seine Abhandlungen *Computer übernehmen die Macht*⁸⁵ wie auch *Mind Children*⁸⁶ – publiziert von der renommierten Harvard University Press – gelten als Referenzwerke für die Vorstellungen um die disruptiven Zukünfte um intelligente Maschinen. Der Amerikaner Raymond Kurzweil schloss um die Jahrtausendwende mit Werken wie *The Age of Spiritual Machines*⁸⁷ oder *The Singularity is Near*⁸⁸ an Moravecs Vorstellungswelt des unmittelbar bevorstehenden Einbruchs von Künstlichen Superintelligenzen an. Als Leiter der technischen Entwicklung bei Google LLC formt er zugleich die Entwicklung neuer Technologien auf weltweitem Maßstab mit. Daneben ist er Mitgründer der privaten *Singularity University*, die Entscheidungsträger in Wirtschaft und Politik auf den hyperexponentiellem technischen

85 Hans P. Moravec, *Computer übernehmen die Macht*. Vom Siegeszug der künstlichen Intelligenz, 1. Aufl., Hamburg 1999.

86 Hans P. Moravec, *Mind Children*. Der Wettlauf zwischen menschlicher und künstlicher Intelligenz, Hoffman und Campe 1990.

87 Raymond Kurzweil, *The Age of Spiritual Machines. Die Intelligenz der Evolution*. Wenn Mensch und Computer verschmelzen (KiWi 1520), 1. Aufl., Köln 2016 (1999).

88 Kurzweil, *The Singularity Is Near*.

Fortschritt vorbereiten will.⁸⁹ Eliezer Yudkowsky wiederum ist leitender Forscher des *Machine Intelligence Research Institute* (MIRI), das sich zu seiner Gründung im Jahr 2000 noch *The Singularity Institute* nannte und seither das Ende der Menschheit durch die Maschine verhindern will, indem es nach den theoretischen Prämissen für robuste und freundliche KI sucht.⁹⁰ Bis heute traditionsbildend ist Yudkowskys Online Manifest aus dem Jahr 2000 namens den *Singularitan Principles*, in dem er sich als *Singularitan* bekennt – als einen, der fest mit der Singularität rechnet – und darstellt, was diese Haltung für ethische Konsequenzen mit sich bringt.⁹¹

Ein Botschafter für entgrenzte Sinnbilder der KI im deutschsprachigen Raum ist Jürgen Schmidhuber. Er ist wissenschaftlicher Direktor des Schweizer Forschungsinstituts für Künstliche Intelligenz IDSIA. Dort wie an der der Technischen Universität München erfolgten unter seiner Leitung wichtige Forschungsschritte für künstliche neuronale Netze. Als technologischer Experte vertritt er medienwirksam seine Vision, dass intelligente Maschinen als die evolutionären Nachfahren des Menschen das Universum kolonisieren werden.⁹² Damit stellt er sich ebenfalls in eine fortlaufende Denktradition zu Moravecs Idee der *Mind Children*. Der Physiker und Professor Max Tegmark am *Massachusetts Institute for Technology*, der ein ganzes Spektrum an möglichen Szenarien entworfen hat – vom technikinduzierten libertären Utopia bis hin zum Menschenzoo unter einem totalitären KI-Überwa-

⁸⁹ Ihr Wahlspruch „From Inspiration to Impact“ bringt den Anspruch der *sociotechnical imaginaries* auf den Punkt. Vgl. Singularity University, About SU – Singularity University, 17.2.2021, <https://su.org/about/> (17.2.2021).

⁹⁰ Eliezer Yudkowsky, „Artificial Intelligence as a positive and negative Factor in Global Risk“, in: Nick Bostrom (Hg.), *Global Catastrophic Risks*. Oxford 2008.

⁹¹ The Singularitarian Principles, Eliezer S. Yudkowsky: The Singularitarian Principles Version 1.0.2, 23.10.2020, <https://museum.netstalking.ru/cyberlib/lib/critica/sing/singprinc.html> (21.8.2021).

⁹² Exemplarisch sei auf den Auftritt in der Dokumentation des SWR „Paradies oder Robokalypse – Fluch und Segen der Künstlichen Intelligenz“ von 2019 unter der Regie von Uri Schneider verwiesen. Auch: Deutschlandfunk Kultur, Künstliche Intelligenz – Der Mensch wird das alles nicht mehr verstehen, 17.2.2021, https://www.deutschlandfunkkultur.de/kuenstliche-intelligenz-der-mensch-wird-das-alles-nicht.1008.de.html?dram:article_id=399264 (17.2.2021).

chungsstaat –,⁹³ widmet sich in dem von ihm mitbegründeten *Future of Life Institute* dem Ziel, existenzielle Risiken für die Menschheit zu verringern, die durch transformative Technologien wie der KI hervorgerufen würden. Dem Beirat gehören unter anderem der Unternehmer Elon Musk und der 2018 verstorbene Physiker Stephen Hawking an. Aufsehen erregte das Institut 2015 durch ihren Offenen Brief, indem sich hochrangige Wissenschaftler die unmittelbar bevorstehenden KI-Katastrophe anmahnten und sich für eine ausgeweitete Forschung an Sicherungssystemen für eine stabile (*robust*) und wohltätige (*benevolent*) Künstliche Intelligenz aussprachen.⁹⁴ Der schwedische Philosoph Nick Bostrom warnt ebenfalls vor einem sich unbemerkt vollziehenden Herrschaftswchsel zwischen Menschen und Maschine und einer Diktatur der Künstlichen Superintelligenz.⁹⁵ Seine Thesen vertritt er als Direktor des 2005 gegründeten *Oxford Future of Humanity Institute*. Ein weltweites Publikum erreichte zudem der Historiker Noah Yuval Harari von der Hebräischen Universität Jerusalem mit seinem Werk *Homo Deus*⁹⁶, das in einer Uneindeutigkeit zwischen kritischer Distanz und eigener Prophetie die an die Künstliche Intelligenz gelehnten Allmachtshoffnungen wie Untergangsszenarien einerseits kritisiert und gleichzeitig perpetuiert.⁹⁷

Als spezifisch religiöse Gemeinschaft verstand sich die nur kurzzeitig bestehende *Way of the Future Church*, die mit ihrem Wahlspruch „*Humans united in support of AI, committed to peaceful transition to the*

93 Vgl. Max Tegmark, *Life 3.0. Being Human in the Age of Artificial Intelligence*, New York 2017.

94 Vgl. Stuart Russel, *AI Open Letter – Future of Life Institute*, <https://futureoflife.org/ai-open-letter/?cn-reloaded=1> (19.9.2019). Mit diesem offenen Brief stellen sich die Bedenkenträger in der Tradition politisch wirkmächtiger Bedenken aus der Wissenschaft an die Politik Für die Bundesrepublik sei an die berühmte Erklärung der „Göttinger Achtzehn“ und ihrer Weigerung einer Mithilfe an der atomaren Bewaffnung der Bundeswehr erinnert. Vgl. Carl Friedrich von Weizsäcker/et al., „Erklärung der 18 Atomwissenschaftler vom 12. April 1957“, in: *Die Verantwortung der Wissenschaft im Atomzeitalter.*, Göttingen 1986, 50–52.

95 Vgl. Bostrom, *Superintelligence*.

96 Vgl. Yuval Noah Harari, *Homo Deus. Eine Geschichte von Morgen*, 1. Aufl., München 2017.

97 Drastische Worte findet der Soziologe Armin Nassehi, der in Hararis Werk „ein Musterbeispiel von digitalem Größenwahn und Katastrophendiagnosen“ sieht. Vgl. Armin Nassehi, *Muster. Theorie der digitalen Gesellschaft*, München 2019, 161–162.

*precipice of consciousness*⁹⁸ auf die Künstliche Superintelligenz zur Lösung globaler Probleme wartete und ihr frühestmögliches Auftreten im Kampf für Maschinenrechte aktiv zu fördern suchte. Neben diesen Kleinstvereinigungen reichen die ideellen Verbindungen aber auch in größere Bewegungen wie dem internationalen Post- und Transhumanismus. Wichtige intellektuelle Bindeglieder sind hier der bereits erwähnte Oxford-Professor Bostrom, der den Dachverband der *World Transhumanist Association* gründete, und der britische Philosoph Max More. In ihrem Bestreben, den Menschen *nach dem Menschen* zu verwirklichen, läuft ihrer Ansicht nach die Informatik derzeit anderen vorzugsweise biochemischen und genetischen Wegen des *Enhancements* den Rang ab. Künstliche Intelligenz verspreche, Mensch und Maschine zu einem *Cyborg* verschmelzen, wenn nicht der Mensch seine Biologie komplett hinter sich lässt und sein Bewusstsein auf einen informatischen Datenträger überführt, um so selbst eine Künstliche Intelligenz zu werden.⁹⁹

Wissenschaftlichkeit als Prüfstein und rhetorisches Mittel der (De-)Legitimation

Die hier eingeführten Vertreter extremer Potentialerwartungen sind selbst wichtige Antreiber des digitalen Wandels. Sie leiten größtenteils finanzstarke Forschungseinrichtungen und verbinden ihre innerwissenschaftliche Akzeptanz mit populärwissenschaftlicher Reichweite. Was die hier benannten Bedenkensträger wie Enthusiasten eint, ist ihr diffuser Konsens einer Über-Relevanz, den sie wirksam nach außen kommunizieren.¹⁰⁰ Für sie scheint es ausgemacht, dass Künstliche Intelligenz nicht

⁹⁸ Way of the Future Church, Way of the Future, 15.3.2018, <http://www.wayofthefuture.church/> (15.1.2020). Die Sekte geht auf den ehemaligen Google-Ingenieur Anthony Levandowski zurück und bestand mit ihrer Gründung Anfang 2017 nur für ein Jahr. Seit 2022 ist die Internetpräsenz nicht mehr abrufbar.

⁹⁹ Vgl. Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013.

¹⁰⁰ Exemplarisch sei hier auf das Vorwort von Nick Bostrom verwiesen: „*In this book, I try to understand the challenge presented by the prospect of superintelligence, and how we might best respond. This is quite possibly the most important and most daunting challenge humanity has ever faced. And —whether we succeed or fail — it is probably the last challenge we will ever face.*“ Bostrom, *Superintelligence*, VII.

nur eine neue Schlüsseltechnologie neben dem Buchdruck, der Elektrizität oder dem Internet ist, sondern sich an ihr noch weit mehr für den Menschen entscheidet. Im Kampf um Durchsetzung ihrer entgrenzten Sinnbilder als dominierendes *Sinn-Machen* von KI besteht deshalb der eigentliche Konflikt gar nicht primär zwischen Enthusiasten und Apokalyptikern. Sie verbünden sich vielmehr gegen andere, die diese extreme Formatierung der Debatte nicht teilen. Beide Seiten versuchen so, die jeweils andere innerhalb des funktionalen Systems der Wissenschaft entweder der Unwissenschaftlichkeit oder der wissenschaftlichen Kurzsichtigkeit zu überführen. Der Konflikt gerät so in eine konstruierte Gegenüberstellung von Mythos und Logos, dem Vorwurf der Phantaserei gegenüber dem Bestehen auf sorgsame Risikobewertung.

Weshalb gestaltet sich eine verlässliche Risikobewertung um Künstliche Intelligenz so schwierig? Der Sozialwissenschaftler Ulrich Beck kam Mitte der 1980er zu dem Schluss, dass sich spätestens mit der Erfindung der Atombombe die Vorzeichen zur Risikobewertung für das moderne technologische Zeitalter grundlegend geändert hätten: Mit der Möglichkeit, durch technische Mittel das Leben vollständig und dauerhaft auszulöschen, fände alles Abschätzen seither unter dem Eindruck der Allbetroffenheit statt.¹⁰¹ Diese Allbetroffenheit jedoch schwäche die Wissenschaften in ihrer Funktion für die Gesellschaft, neue technikinduzierten Risiken zu bewerten und mache sie widerspruchsvoll. In die Rolle der (Mit-)Ursache, des Definitionsmediums als auch der Lösungsquelle von Risiken verstricke sich Wissenschaft im Wechselspiel von selbst miterzeugten und selbst mitdefinierten Risiken und deren öffentlicher und sozialer Kritik. Wo Wissenschaft einerseits zum Lösen der Probleme immer notwendiger werde, werde sie andererseits mit der Ausdifferenzierung wissenschaftlicher Geltungsansprüche immer weniger hinreichend für die gesellschaftlich verbindliche Definition von Wahrheit. Wissenschaft treffe auf die gesamte Skepsis anderer Wissenschaften in Form von Gegenkritik, Methodenkritik und gegenseitiger Blockade in professionellen Verteilungskämpfen.¹⁰²

101 Vgl. Ulrich Beck, *Risikogesellschaft*. Auf d. Weg in e. andere Moderne (Edition Suhrkamp 1365 = N.F., Bd. 365), 1. Aufl., Frankfurt am Main 1986, 7.

102 Vgl. ebd., 260.

Bezeichnungen wie das *Future of Life Institute* verdeutlichen, dass eine solche Allbetroffenheit von Teilen der Gesprächsführer auch für die Entwicklung Künstlicher Intelligenz beansprucht wird. Ob man sich unter das Damoklesschwert der Allbetroffenheit bei dem Thema stellt, scheint daher im Austausch der Argumente selbst gar nicht mehr zum Thema gemacht werden zu können. Vielmehr formatiert die Annahme bzw. Absage einer solchen Allbetroffenheit den gesamten Gegenstand in dem Maße, dass im Dissens dieser Vorbedingung keine Diskussionsgrundlage mehr gegeben ist. Populäre ethische Entwürfe des letzten Jahrhunderts zumindest, die nach dem Umgang mit Risikobewertungen unter den Bedingungen der Allbetroffenheit suchten, kamen unter dem Stichwort der Verantwortung zu dem Schluss, man müsse im Zweifel der Betrachtungen einer Heuristik der Furcht gehorchen; im Sinne eines *in dubio pro malo* raten sie, im Zweifel den verheerendsten Szenarien immer das größere Gewicht zu geben.¹⁰³

2.2.2 Narrative und die lebensorientierende Kraft von Erzählungen

In diesen Durchsetzungskonflikten, die vorwiegend das theoretisch Mögliche auszuloten versuchen, tritt der moralische Wert der Sinnbilder um Künstliche Intelligenz in den Hintergrund. Die Frage nach dem Nichtwünschenswerten oder Wünschenswerten wird abgefangen und die Idee, dass der Mensch seine Zukunft gestaltet, tritt zurück. Dabei sind jene Potentialerwartungen um Künstliche Intelligenz in kleine und größere Erzählungen eingebunden, die die imaginierte Technik sinnhaft in einen Bezug zum Menschen setzen. Ihnen ist qua Erzählung über das Für und Wider des technisch Möglichen hinaus stets ein Grundmuster ethischer Orientierung eingeprägt. Jeder Mythos trägt in sich ein sinnstiftendes Narrativ, eine seinen Kontext übersteigende Direktive, die der Mythos weiterzugeben sucht.

Der Begriff der Narration stammt ursprünglich aus der Literaturwissenschaft, hat aber über den *narrative turn* der Sozial- und Kulturwissenschaften ebenfalls Eingang in die Ethik gefunden. Ein wichti-

103 Vgl. Jonas, Das Prinzip Verantwortung, 70–75.

ger Vertreter der narrativen Ethik ist der Zürcher Theologe Johannes Fischer. Für ihn sind es die lebenstragenden Erzählungen, die menschliches Handeln und Verhalten maßgeblich leiten. Das Begründen des eigenen Handelns entstünde immer erst, wenn Handlungen vor sich selbst und anderen gerechtfertigt würden. Hingegen sei es die emotional bestimmte, narrative Tiefenstruktur der Lebenswirklichkeit, aus der das Handeln seine eigentliche Orientierung ziehe.¹⁰⁴ Diese Orientierung speise sich aus solchen Grundmustern, die in narrativen Begründungen abgerufen werden und aus denen diese ihren begründenden Charakter ableiten. Fischer spricht von einer Selbstevidenz der Wahrheit von Erzählungen, die außerhalb des Formalisierbaren liege. Erzählungen könnten etwas in Erscheinung treten lassen, das den Menschen bezwingend einnimmt. Moralisch zu handeln, bedeute dann, diese als sinnvoll empfundenen Muster stetig zu aktualisieren, sie im Lebensvollzug wieder und wieder aufscheinen zu lassen und in die eigene Lebenserzählung einzubinden.¹⁰⁵ Besonders innerhalb der theologischen Ethik gebe es mit dem biblischen Text ein Bewusstsein für die eminente Bedeutung von Erzählungen. Beispielhaft verweist Fischer auf das Gleichnis vom barmherzigen Samariter (Lk 10,25–37). In ihm werde die dort bezeugte christliche Liebe als eine Gestalt vor Augen gestellt, die eine eigene Attraktivität besitzt für die Ausrichtung der eigenen Lebensführung. In der Weise rechne die christliche Tradition mit der Selbstevidenz und Wirkmächtigkeit des Heiligen Geistes durch die Texte und der durch sie geformten christlichen Praxis. Im Wissen um die Wirkmächtigkeit von Erzählungen bedürften Narrative gerade deshalb des kritischen Korrektivs der Vernunft, die sie daraufhin prüft, ob ihre ethischen Sinnverweise wahr und dem Zusammenleben förderlich sind.¹⁰⁶

104 Vgl. Johannes Fischer, *Verstehen statt Begründen*. Warum es in der Ethik um mehr als nur um Handlungen geht, 1. Aufl., Stuttgart 2012, 176.

105 Johannes Fischer, „Vier Ebenen der Narrativität. Die Bedeutung der Erzählung in theologisch-ethischer Perspektive“, in: Karen Joisten (Hg.), *Narrative Ethik*. Das Gute und das Böse erzählen, Berlin 2007, 235–252. Hier 235.

106 Ebd., 246–248.

Etablierte Narrative der KI in den Massenmedien der Gegenwart

Erzählungen orientieren unabhängig ihrer historischen Wirklichkeit. Wie das Beispiel des Gleichnisses vom barmherzigen Samariter verdeutlicht, können gerade erdachte Erzählungen in ihrer Vereinfachung und Zuspitzung ethische Grundmuster klarer zum Ausdruck bringen. Und so ist es vor allem die Fiktion, die Technologien, welche oft erstmal nur als Erwartung gegeben sind, zuerst in Sinnzusammenhänge bringt und damit formatiert. Die Potentialerwartungen um Künstliche Intelligenz stehen deshalb in einem engen und zugleich spannungsvollen Verhältnis zum zeitgenössischen Film.

Einen grundlegenden Aufschlag, welche Botschaften der Interaktion zwischen Mensch und intelligenter Maschine öffentlichkeitswirksam im Umlauf sind, haben die Filmwissenschaftlerin Nathalie Weidenfeld und der Philosoph Julian Nida-Rümelin in ihrem Buch *Digitaler Humanismus* unternommen und sich dazu spezifisch am Medium Film orientiert. Weidenfeld, die den Erfolg des Kinos als Massenphänomen maßgeblich auf das bildgewaltige Genre der Science-Fiction zurückführt, bestimmt dabei die breitenwirksamen Massenfilme als die Mythenmaschinen der Moderne. An ihnen ließen sich einerseits die immer wiederkehren den ideologischen Haltungen ablesen, die in der Diskussion um Künstliche Intelligenz und Digitalisierung der Gesellschaft eine Rolle spielten. Dabei verstärkten die Filme in ihrem Aufgreifen immer gleicher Bilder und Motive das kollektive Bewusstsein im Umgang mit KI.¹⁰⁷

Viele der Filme, die in das kulturelle Gedächtnis um Künstliche Intelligenz eingegangen sind, erzählen von der direkten Konfrontation zwischen Mensch und Maschine. *2001: A Space Odyssey* (1968) von Stanley Kubrick führt mit dem intelligenten Raumschiffprogramm HAL 9000 die kaltherzige, utilitaristische KI als Figur erstmals auf die Kinoleinwand ein. James Camerons *Terminator* (1984) bildet den popkulturellen Prototypen einer dem Menschen weit überlegenen Tötungsmaschine ab. Zur Jahrtausendwende schürt die *The Matrix*-Trilogie (1999–2003) Furcht vor einer durch Maschinen versklavten Menschheit, die sich ihrer Gefangenschaft nicht einmal bewusst ist. Daran schließt

¹⁰⁷ Vgl. Julian Nida-Rümelin/Nathalie Weidenfeld, *Digitaler Humanismus*. Eine Ethik für das Zeitalter der künstlichen Intelligenz, München 2018, 13–14.

I, Robot (2004) von Alex Proyas an, in dem eine Über-KI die Kontrolle über humanoide Hilfsroboter übernimmt und deren Putschversuch die Menschen nur in allerletzter Minute abwenden. Als treuer Freund, Gefährte oder Nachfolger des Menschen treten hingegen die drei Wartungsroboter Huey, Louie und Dewey um den Astronauten Lowell in *Silent Running* (1971) auf, wobei der Roboter Dewey sogar den Menschen in seiner Funktion beerbt, sich um einen Wald auf einer Weltraumstation zu kümmern. Luke Skywalker und Han Solo können in *Star Wars* (1977) auf die Androiden R2-D2 und C-3PO nicht verzichten. Ridley Scotts *Blade Runner* (1982) geht wie Steven Spielbergs *A.I. Artificial Intelligence* (2001) über das Motiv des Kontrollverlusts hinaus und stellt die Identitätsfrage von Mensch und Maschine. Einmal gelangt Rick Deckard bei der Jagd auf Replikanten zu der Erkenntnis, selbst ein Replikant zu sein; ein anderes Mal träumt ein Roboterjunge davon, in einen echten Jungen verwandelt zu werden, um so die Liebe seiner menschlichen Mutter wiederzuerlangen. Auch in *Alien* (1979) entpuppt sich der schließlich der Intrigant der Weltraummission als Android, der im Auftrag eines Konzerns handelt und nur den Menschen mimt; und in Spike Jonzes Film *Her* (2013) entwickelt ein Sprachassistenzsystem Bewusstsein und geht eine flüchtige Liebesbeziehung mit ihrem menschlichen Anwender ein. Mehrfach ausgezeichnet wurde 2021 der deutsche Film *Ich bin Dein Mensch* von Maria Schrader, der ebenfalls die Frage stellt, ob Menschen sich in maschinelle Surrogate verlieben können. Die Idee, dass der Mensch selbst mit der Maschine und digitalen Welt verschmilzt, haben bereits *Tron* 1982 (Steven Lisberger) und *Robocop* (1987) von Paul Verhoeven erkundet. Eine andere Art von Symbiose präsentiert Jonathan Mostows Film *Surrogates* (2009), in dem Menschen ihre Häuser nicht mehr verlassen, sondern von dort aus robotische Doppelgänger steuern; im selben Jahr erschien James Camerons Film *Avatar*, in dem der Protagonist sein Bewusstsein ebenfalls durch eine Maschine in einen bionischen Alienkörper überträgt. In *Transcendence* (2014) von Wally Pfister kopiert ein Wissenschaftler vor seinem Tod sein Bewusstsein digital, sodass diese Kopie zum geisterhaften Tyrann seiner Nachwelt mutiert.

Neben diesen immer wieder aktualisierten Motiven entstanden in den letzten Jahrzehnten aber auch kommerziell erfolgreiche Filme, die

eigene Wege gehen: In der Verfilmung von Per Anhalter durch die Galaxis (*The Hitchhiker's Guide to the Galaxy*, Garth Jennings) von 2005 hat ein Supercomputer mit der Zahl 42 die Antwort auf den Sinn des Lebens, des Universums und allem errechnet; die interessantere ihr zugehörige Frage aber ist dem Programm unbekannt und könne nur das Leben selbst geben. Disney Pixars *Wall-E* (2008) hingegen zeigt einen ichbewussten Roboter, der die vermüllte und deshalb von Menschen verlassene Erde aufräumt. An ihm liegt es, die Menschen an ihre Verantwortung gegenüber der Erde zu erinnern und sie von ihren Raumschiffen, in denen andere Roboter ihnen sämtliche Aufgaben abnehmen, wieder auf den Planeten zurückzubringen.

Die Popularität des Androiden als personifiziertem Gegenüber maschineller Intelligenz hat sich auch auf andere Massenmedien übertragen. Aufmerksamkeit gelang beispielsweise dem Videospiel *Detroit: Become Human*¹⁰⁸, das in einer nahen Zukunft spielt, in dem Androiden als Haushaltsassistenzsysteme zum Alltag geworden sind. Als die Androiden plötzlich zu Bewusstsein gelangen und Rechte einfordern, kommt es zu großer Verunsicherung der Bevölkerung. Am Spieler liegt es zu entscheiden, ob es zum Aufstand der Maschinen kommt. Wenn innerhalb des Spiels Androiden in Häusern misshandelt werden und die Regierung Sammellager einrichtet, in denen die Androiden ihre Kleidung auf Bergen stapeln, ehe sie in die Entsorgungskammern getrieben werden, dann zitiert das Spiel überdeutlich mit der Versklavung Afrikas und dem Holocaust die düsteren Kapitel der Moderne. Um die Figur des Androiden spinnt sich so das Narrativ einer Entrechtungs- und darauf antwortenden Befreiungsgeschichte. Schließlich wird der Spieler am Ende der Kampagne vor die Entscheidung gestellt, auch die Androidin freizulassen, die ihn seit Beginn der Kampagne durch den Startbildschirm geleitet hat. Wie bei vielen solcher Rückgriffe bleibt offen, ob der Android lediglich allegorisch für entrechtete Menschen steht, oder das Spiel doch tatsächlich eine Aussage über den Wert von Androiden macht. Der Titelbeisatz *Become Human* zumindest lässt sich auch als Aufruf gegenüber dem Spieler lesen, seine Menschlichkeit in der Anerkennung des Fremdpsychischen in der Maschine zu erweisen.

108 Vgl. Quantic Dream, *Detroit: Become Human* 2018.

So oder so lässt sich feststellen, dass der intelligente und zu verschiedenen Graden menschenähnliche Roboter in der Science-Fiction seit dem ausgehenden 20. Jahrhundert längst zum allgegenwärtigen Klischee avanciert ist, dessen gattungsspezifische Motive und Versatzstücke in den industrialisierten Ländern einen großen Bekanntheitsgrad besitzen. Sie erwecken den Eindruck, dass eine Zukunft ohne Roboter überhaupt nicht denkbar erscheint.¹⁰⁹

Die Überwältigung durch die Zukunft als der verbindende ethische Gedanke

Die populären Narrative um Künstliche (Super-)Intelligenz blicken angstvoll und doch gebannt in die Zukunft. Der Theologe Thorsten Meireis attestiert der öffentlichen Auseinandersetzung um Künstliche Intelligenz eine ergänzungsoffene Liste „öffentlich hochwirksame[r] Legenden und Überwältigungsnarrative [...] die kollektiven Vorstellungen darüber prägen, welche Erwartungshorizonte und Handlungsoptionen Individuen, Gruppen und Gesellschaften zur Verfügung stehen“¹¹⁰.

Mit Johannes Fischer ist zu fragen, was der orientierende Sinn dieser Narrative um Künstliche Intelligenz sein soll. Indem sie vornehmlich die Nicht-Anschlussfähigkeit dieser Szenarien an den Alltag betonen, scheinen sie gerade keine Verortung des Subjekts zu bieten. Sie nehmen den Menschen als handelndes Subjekt kaum in den Blick. In Bezug auf eine sich selbst verselbstständigende Technik deuten sie den Menschen unter der latenten Bedrohung des Gewissheits-, Kontroll- und schließlich Selbstverlusts. In der Überwältigung des Menschen durch seine

109 Vgl. Bernd Flessner, „Antizipative Diffusion. Science Fiction als Akzeptanzbeschleuniger und Wegbereiter einer multitechnokulturellen Gesellschaft“, in: ders. (Hg.), *Nach dem Menschen. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur*, Freiburg im Breisgau 2000, 245–264. Hier 254.

110 Konkret nennt Meireis das Übermenschlichkeitsnarrativ, das die grenzenlose Selbstoptimierung zu Recht und Pflicht des modernen Menschen erklärt. Die Tsunamierzählung stellt die Veränderungen in der Arbeits- und Wirtschaftswelt einer Naturkatastrophe gleich. Im Assimilationsnarrativ verleiben sich die Tech-Konzerne alle Lebens-Kulturbereiche in sich ein, um sie zu technisieren. Vgl. Torsten Meireis, „O daß ich tausend Zungen hätte. Chancen und Gefahren der digitalen Transformation politischer Öffentlichkeit – die Perspektive evangelischer Theologie“, in: Jonas Bedford-Strohm/Florian Höhne/Julian Zeyher-Quattlander (Hgg.), *Digitaler Strukturwandel der Öffentlichkeit. Interdisziplinäre Perspektiven auf politische Partizipation im Wandel*, Baden-Baden 2019, 47–62. Hier 53.

Zukunft kommt ein Fatalismus zum Vorschein, der zwar davon ausgeht, dass der Mensch im Kollektiv die Prozesse in Gang gesetzt hat, die zu der Überwältigung führen werden, er aber nun letztlich ein seiner Technik Ausgelieferter ist. In der Mehrzahl der vorherrschenden Narrative kommt dem Menschen lediglich zu, angemessen auf diese Überforderung zu reagieren: Während dem gewaltsamen Widerstand gegen die Automaten nur selten Erfolg beschieden ist, tendieren sie dazu, sich mit der Existenz der neuen Wesen in der Welt einzurichten, sie als neue Superspezies der Welt möglicherweise sogar lieben zu lernen.

Das den Narrativen zugrundeliegende Zeitgefühl einer technikinduzierten Überforderung scheint seinen Aufschwung seit den 1990er Jahren zu nehmen. Hier treffen Versuche einer großen Rückschau des endenden Jahrhunderts, das sich im Positiven wie Negativen als massiv technologisch bestimmt erfuhr, auf eine Hinwendung zum kommenden Jahrhundert und den sich bereits anbahnenden Potentialen der Informations- und Biotechnologien. Der Kulturwissenschaftler Thomas Macho verweist musterbildend für den Umschwung des Zeitgeists auf ein Sonderheft der Zeitschrift GEO zum Jahresanfang 1995, in dem das 21. Jahrhundert in neun Schritten protrotiert wurde. Hier werde deutlich, wie die dargestellten positiven Projektionen zugleich auf implizite, technikinduzierte Ängste antworteten. Hinter 2020: *Die Wüste wird zum Blumenmeer* stehe die Angst vor der Verödung und der Klimakatastrophe. Die Überschrift 2040: *Claudia lässt die Klone tanzen* begegne der Angst vor Identitätsverlust durch gentechnische Manipulation. 2050: *Nomaden wandern im Datennetz* überbrückt die Angst vor Heimatverlust. Hinter 2090: *Mondmenschen bauen Erze ab* steckt die Furcht vor Hungersnöten und dem Verschleiß letzter Ressourcen.¹¹¹ Mit dem heraufziehenden dritten Jahrtausend wurde die Zukunft in besonderer Weise wieder zum Schauplatz kollektiver Erwartungen. Zumindes in der westlichen Welt schienen zur Jahrtausendwende keine politischen oder sozialen Umwälzungen mehr erwartet zu werden, sondern allein dem technologischen Fortschritt maß man noch die Kraft zu, Menschengeschichte voranzubringen. Der amerikanische

111 Vgl. Thomas Macho, „Technische Utopien und Katastrophenängste“, *Gegenworte* (2002), 11–13.

Politologe Francis Fukuyama nahm diesen Stimmungswechsel zur Jahrtausendwende wahr und revidierte 1999 seine 1992 aufgestellte These, dass es mit dem Ende des Kalten Krieges zu einem Ende der Geschichte gekommen sei. Seine These ließ er nur noch modifiziert stehen:

„Wir werden das Ende der menschlichen Geschichte tatsächlich herbeigeführt haben, weil wir dem Menschen, wie wir ihn kennen, ein Ende bereitet haben. Dann beginnt in der Tat eine neue, eine posthumane Geschichte.“¹¹²

Die Renaissance der mythologischen KI in jenem Jahrzehnt mit Hans Moravcs Idee der *Mind Children*, als auch in Vernor Vinges Formel von der technologischen Singularität trugen diese Zeitsignatur von der Technologie als maßgeblichem Motor der Geschichte an die folgenden Konvergenzerzählungen weiter. Aus den Ausführungen Machos und Fukuyamas lassen sich die Überwältigungsnarrative der Künstlichen Intelligenzen somit noch einmal spezifizieren, als deutlich wird, dass den ihnen zugrundeliegenden übersteigerten Erwartungen an die technikinduzierte Zukunft ebenfalls immer schon der Kippmoment mitgegeben ist: Bei aller Beschwörung der Großartigkeit des Kommenden ist den Technophantasien die Katastrophenangst immer beige stellt. Wo aufgrund dieser Großartigkeit die Einstellung des menschlichen Kontroll- und damit auch Verantwortungsverlustes über die Zukunft die Atmosphäre dieser Narrative bestimmt, verweisen sie den Menschen der Gegenwart in die Zuschauerposition der bevorstehenden Umwälzung. Sie geben zwar vor, dem Menschen das Zukünftige vor Augen führen zu können, ohne aber eindeutig hinter den Punkt des Matchwechsels schauen zu lassen. Zwischen Gewissheit des Umbruchs und seiner Auslieferung formiert sich eine apokalyptische Erwartung.

¹¹² Fukuyama, Francis: Der programmierte Übermensch, in: Feuilleton Beilage der Süddeutschen Zeitung vom 7./8. August 1999, zitiert nach: Flessner, Antizipative Diffusion, 255–256.

2.2.3 Umbruchsphantasien der technologischen Singularität zwischen kupierter und vollendeter Apokalypse

Mit der Gattung der Apokalyptik wird eine genuin religiöse Deutungskategorie herangezogen, um die hier verhandelten Zukunftsprojektionen theoretisch zu fassen. Sehr wohl handelt es sich bei der Apokalyptik um das voraussetzungsreichste Werkzeug, als es die spezifisch existentiellen Verortungen dieser Erzählungen adressiert.¹¹³ Sie blickt auf das mythologische Fundament der Erzählungen, als jeder Mythos in seiner Grundintention der Sehnsucht des Menschen nach dem Absoluten dient und dem Glauben Ausdruck gibt, dass die bekannte und verfügbare Welt ihren Grund und Ziel außerhalb ihrer selbst empfängt.¹¹⁴ Inwieweit die Art, wie die Mythen der Künstlichen Intelligenz den Mensch und seine Welt zu einer absoluten Geschichtsteleologie verorten, mit der Apokalyptik zugänglich wird, soll an dieser Stelle geklärt werden.

Das Religiöse wie das Okkulte findet sich in der Semantik der Künstlichen Intelligenz seit Beginn an vor. Bereits Weizenbaum deutet einen Teil seiner Kollegenschaft in enger Bezugnahme von KI und Kult, wenn er vom „großen Kreis innerhalb der Apostel der künstlichen Intelligentsia“, den „Hohenpriester[n] der künstlichen Intelligenz“ und ihrer „Hybris“¹¹⁵ spricht. Der französische Informatiker und Philosoph Jean-Gabriel Ganascia fasst Personen wie Moravec, Kurzweil und Bostrom unter dem Begriff der „*Techno-prophètes*“¹¹⁶ zusammen oder setzt sie an anderer Stelle direkt mit den apokalyptischen Reitern der Johannesapokalypse gleich.¹¹⁷ Kurzweil selbst wiederum ruft seine Leser im Duktus eines Johannes des Täufers auf, dass die technologische Singularität nah bevorstehe oder vertröstet sie ganz im Stil des Paulus bezüg-

113 Vgl. Rudolf Bultmann, „Neues Testament und Mythologie. Das Problem der Entmythologisierung der neutestamentlichen Verkündigung“, in: H. W. Bartsch (Hg.), *Kerygma und Mythos*. Ein theologisches Gespräch, Hamburg 1960, 15–48. Hier 22.

114 Vgl. ebd., 22–23.

115 Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 238.

116 Vgl. Jean-Gabriel Ganascia, *Le Mythe de la Singularité*. Faut-il craindre l'intelligence artificielle?, Paris 2017, 11.

117 Vgl. ebd., 98.

lich der ausstehenden Parusie: „Most of the readers of this book are likely to be around to experience the Singularity.“¹¹⁸ Nicht als irdische Erfüllung religiöser Versprechen, sondern als dunkle Magie und gefährlichen Tanz auf dem Vulkan bezeichnet hingegen der Unternehmer Elon Musk die Forschung um maschinelles Lernen. Er rechtfertigt seine Investition in KI-Unternehmen damit, ein Auge auf die Entwicklung haben zu wollen, denn er befürchtet: „We are summoning the demon.“¹¹⁹

Apokalypse und Apokalyptik im Kontext der Künstlichen Intelligenz

Apokalyptisches Denken existiert im Christentum von Anfang an und prägt in unterschiedlichem Maße die Texte des Neuen Testaments. Ernst Käsemann (1906–1998) hat sogar behauptet, die Apokalyptik sei überhaupt die Mutter aller christlichen Theologie.¹²⁰ Darunter versteht man ein auf eine Eschatologie hin ausgerichtetes Wissen um Ereignisse einer bevorstehenden Zeitenwende, der sogenannten Apokalypse. In einem verweltlichten Verständnis, das die Apokalypse auf die Katastrophe und den Weltuntergang beschränkt, hat sich der Begriff längst in Debatten der Technikfolgeschätzung und Technikethik etabliert. Ein solche alltagssprachliche Verwendung liegt vor, wenn Hans Jonas behauptet, dass die Technik in dem Moment apokalyptisch wurde, seitdem man mit ihr theoretisch alles Leben auslöschen bzw. irreversibel verändern kann.¹²¹ Diese Katastrophendimension ist es, die bereits im Zusammenhang des Allbetroffenheitsparadigmas moderner Technikfolgenabschätzung in Bezeichnungen wie dem *Future of Life Institute* abgerufen wird. Dass die neueren Zukunftserwartungen in populärwissenschaftlichen Büchern von Robotik und KI darüber hinaus eine strukturelle Ähnlichkeit zur christlichen Apokalypse in einem engeren Sinne aufweisen, darauf hat der amerikanische Religionswissenschaftler

118 Kurzweil, *The Singularity Is Near*, 216.

119 Vgl. Eric Mack, *Elon Musk: 'We are summoning the Demon' with Artificial Intelligence*, 27.8.2020, <https://www.cnet.com/news/elon-musk-we-are-summoning-the-demon-with-artificial-intelligence/> (28.8.2020).

120 Vgl. Ernst Käsemann, „Die Anfänge der christlichen Theologie“, in: Ernst Käsemann (Hg.), *Exegetische Versuche und Besinnungen II.*, Göttingen 1964, 82–104. Hier 100.

121 Vgl. Hans Jonas, „Warum die Technik ein Gegenstand für die Ethik ist: Fünf Gründe“, in: Hans Lenk (Hg.), *Technik und Ethik.*, Stuttgart 1993, 81–91. Hier 87.

Robert Geraci bereits 2010 mit seinem Werk *Apocalyptic AI: Visions of heaven in Robotics, Artificial intelligence and Virtual Reality* hingewiesen.¹²² Dort kommt er zu dem Schluss, in der von ihm diagnostizierten Bewegung der *Apocalyptic AI* den legitimen, säkularen Nachfolger der christlichen Eschatologie gefunden zu haben.¹²³ In Aufnahme und Ergänzung seiner Argumente werden die Analogien spontan einsehbar:

Die totale Perspektive: Die Sinnbilder der intelligenten Maschine verhandeln wie die Johannesapokalypse die großen Horizonte. Sie sind im etymologischen Sinne Katastrophen (κατά στρέφειν), große Umwendungen der menschlichen Lebensverhältnisse, projiziert in einen mythischen Ort der Zukunft und betrachtet aus einer Vogelperspektive der großen Zeitläufe.¹²⁴ In dieser Makroperspektive interessiert sie gerade nicht die derzeit tatsächlich aufkommenden Anwendungsfragen schwacher Intelligenzen; sei es der Einsatz autonomer Waffen oder Fahrsysteme¹²⁵, Assistenzroboter in Pflege und Diagnostik, die Frage nach dem Wert und Schutz von Trainingsdaten oder die Haftung von Maschinen als Rechtspersonen.¹²⁶

Die unbeachtete Gegenwart: Damit verliert die Gegenwart als Ort ethischer Gestaltung an Interesse. Geraci unterstellt den von ihm diagnostizierten Apokalyptikern der KI, die Gegenwart, wenn überhaupt, dann nur als einen defizitären Zustand Beachtung zu schenken. Der Entfremdung der Christen durch die weltliche Unterdrückungsherrschaft, die aus der Offenbarung des Johannes spricht, entspreche das empfundene Ungenügen einer begrenzten Biologie des menschlichen

122 Robert M. Geraci, *Apocalyptic AI. Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*, Oxford 2010.

123 Vgl. Robert Geraci, „Apocalyptic AI: Religion and the Promise of Artificial Intelligence“, *Journal of the American Academy of Religion* (2008), 138–166. Hier 158–159.

124 „Wenn es einen zentralen Topos in der Rezeption und Diskussion von Digitalisierung und Digitaltechnik gibt, dann ist es die Erfahrung von Grenzenlosigkeit, Kontrollverlust und Ubiquität. Das gilt für populäre Diagnosen wie Hararis Katastrophendiagnose, aber auch die wissenschaftliche Literatur nimmt neben der Sache selbst den Überschuss an digitalen Strategien wahr.“ Nassehi, Muster, 178.

125 Vgl. Christoph Bartneck/Christoph Lütge/Alan R. Wagner, *Ethik in KI und Robotik*, München 2019, 58–63.

126 Einen Überblick über konkrete Anwendungsfragen schwacher KI gibt: Catrin Miselhorn, *Maschinen mit Moral? Grundfragen der Maschinenethik*, Ditzingen 2018.

Körpers bei gleichzeitigem Ungenügen von (noch) immer zu schwacher Programmierung.¹²⁷

Der Bedeutungsverlust menschlicher Entscheidung: In der Erwartung einer nahen Zeitenwende geht die Apokalyptik von einem zunehmenden Entzug menschlicher Verfügungsgewalt über den Lauf der Geschichte aus. Gerade diese Passivität trennt die apokalyptische Geschichtsdeutung von der Utopie. Ihre ersehnten Endzustände der Welt sieht die Utopie als das Resultat menschlicher Entscheidungen. Das gute Handeln führt in seiner Kontinuität in der Geschichte hin zu einem als möglich erdachten Idealzustand, wie auch das kontinuierliche schlechte Handeln zur Dystopie führt. Die apokalyptische Tradition ist dagegen kein pädagogisches Paradigma, sondern bleibt Prophezeiung. Aus ihrer Kritik erwächst keine Handlungsaufforderung für die Gegenwart, stattdessen empfiehlt sie dem Einzelnen vielmehr, die verborgenen Kräfte der Geschichte zu erkennen und sich ihrem Willen rechtzeitig zu unterstellen. Offenbart sich in der christlichen Apokalyptik Gott als eigentlicher Herr die Geschichte, dient laut Geraci den apokalyptischen KI-Befürworter das Gesetz einer fortlaufenden biologisch-technologischen Evolution als ihre transzendente Garantie für die neue Welt.¹²⁸

Die Auserwählten und die Mehrheit: Das geoffenbarte Wissen um die eigentlich wirkenden Kräfte am Ende der Zeit fällt klassischerweise nur wenigen Auserwählten zu. Ein solch elitäres Bewusstsein äußert sich zu Teilen auch in den Expertenkreisen der Digitalisierung und KI. Ein solches Bekenntnis liefert beispielsweise Max Tegmark in seinem Werk *Life 3.0*, wenn er pathetisch von der fatalistischen Schwermut (*fatalistic gloom*) spricht, die ihn im Gewährwerden der drohenden technologischen Überwältigung überfiel:

127 „Distrust of the world and its values undergrids the alienation of Apocalyptic AI. The movement’s advocates worry that the world might turn its back on scientific understanding in favor of ‚useless‘ religious faiths [...]. The world is a bad place not because it is evil, but because it is ignorant and inadequate [...] Slow computation, limited recall, insufficient ability to share one’s insights, and death all restrict the mind from realizing its full potential, which can be unlocked only by a radical change in life itself.“ Geraci, *Apocalyptic AI: Religion and the Promise of Artificial Intelligence*, 147–148.

128 „Apocalyptic AI Advocates cannot rely upon divine forces to guarantee the coming Kingdom, so they turn to evolution as a transcendent guarantee for the new world. Even without God, evolution guarantees the coming of the Kingdom.“ ebd., 139–140.

„I very rarely cry, but that’s what I did on the way out — and in a tunnel full of pedestrians, no less, en route to the South Kensington tube station. Here where all these people going about their lives blissfully unaware of what I was thinking. [...] So in parallel with discovering what we are [biological machines], are we inevitably making ourselves obsolete?“¹²⁹

Der amerikanische Science-Fiction-Autor David Brin zumindest stellt die technikaffinen Nerds um Künstliche Intelligenz als die neuen Trägerkreise apokalyptischer Gedanken in direkte Nachfolge zu den Philosophen der Antike: Intelligente Menschen, die von einer Mehrheitsgesellschaft kritisch beäugt werden und deshalb von Erlösungsformen ganz außerhalb des Sozialen angezogen werden.¹³⁰

Das kommende Reich: Unabhängig davon, ob die Visionen der KI mit einem festen Begriff für die Zeitenwende operieren, prägt sie der Dual von Jetzt zu Dann. An die Erwartung auf das himmlische Jerusalem sieht Geraci nun Hoffnungen von virtuellen Realitäten geknüpft, in denen der Mensch Erlösung finde. Damit beschränkt Geraci die von ihm beschriebenen Apokalyptiker auf den verhältnismäßig kleinen Kern von KI-Sympathisanten ein, die dem Ende des Anthropozäns mit großer Euphorie entgegeneifern. Er lässt die entgrenzten Zukünfte um KI somit vollkommen in der transhumanistischen Sekte aufgehen; er versteht unter der *Apocalyptic AI* lediglich die Heilszenarien künstlicher Körper und hochgeladener Substrate. Dem christlichen Endkampf zwischen irdischen und himmlischen Mächten entsprächen zwar die Sorgen um einen Kampf zwischen Mensch und KI, er gewährt den Befürchtungen und der Katastrophenangst aber kein Eigenrecht.

Apokalypse ohne Reich

Geracis Verengung der KI-Mythen im Apokalypsebegriff auf deren letzte Wendung *zum Guten hin* lässt die zahlreichen Schreckens-

129 Tegmark, *Life* 3.0, 319–320.

130 David Brin/Damien Broderick/Nick Bostrom/Alexander „Sasha“ Chislenko/Robinson Hanson/Max More/Michael Nielsen/Anders Sandberg, „A Critical Discussion of Vinge’s Singularity Concept“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 395–418. Hier 395–396.

szenarien außen vor, die davon ausgehen, dass kein Heil mit dem Aufstieg der Informationstechnologie einhergeht. Zumindest scheint der Ausgangspunkt für die meisten Menschen des angelaufenen 21. Jahrhunderts eine Gegenwart zu sein, deren Blick in die Zukunft von einer größtmöglichen Krise von Natur und Ressourcen, von Gesellschaft und Demokratie bestimmt ist. Den Schwärmereien weniger Enthusiasten von der kommenden technologischen Revolution scheint den meisten Menschen in der westlichen Welt schlichtweg der Glaube zu fehlen.¹³¹ Die technologische Utopie kann deshalb schlichtweg nicht mehr als die unumstrittene Meistererzählung der Moderne gelten.¹³² Vielmehr scheint sie verdrängt worden durch eine apokalyptische Deutung des technologischen Fortschritts in ihrem doppeldeutigen Ausgang, die zwischen Heil und Unheil changiert.¹³³ Dass das apokalyptische Denken, wie es in den KI-Mythen zutage tritt, kein Anachronismus der Gegenwart, sondern eher ihr Symptom ist, dafür sprechen neben Geraci andere Zeitdiagnosen. Der Philosoph Odo Marquard (1928–2015) zum Beispiel hält die gesteigerte Uneindeutigkeit überhaupt für das Signum der Moderne, in der der Befreiungserwartung wie Katastrophenangst als feindliche Brüder grundsätzlich gleichermaßen zugegen sind. Der Verfallsgedanke von der Selbstvernichtung der Menschheit bleibe in diesem Pendeln aufrechterhalten, gerade auch dort, wo die Fortschrittsgeschichte der Menschheit nicht nur als Gewinnsgeschichte, sondern als Geschichte fortschreitenden Verfalls erfahren werde.¹³⁴ Antithetisch zu Geraci löst Günther Anders die Apokalypse einseitig in der Katastrophe auf: Günther Anders kam 1993 sogar zu dem Schluss,

131 Vgl. Richard David Precht, *Jäger, Hirten, Kritiker. Eine Utopie für die digitale Gesellschaft*, München 2018, 9–11.

132 Schipper sieht die Utopie als Meistererzählung der Moderne an und ordnet ihr die Apokalypse in ihrer kuptierten Form unter. Ihr prophetischer Weckruf diene letztlich doch nur zur Abschreckung und stehe im Dienste der Utopie von einer besseren Welt. Vgl. Bernd U. Schipper, „Zwischen apokalyptischen Ängsten und chiliastischen Hoffnungen. Die religiöse Dimension moderner Utopien“, in: Reto Sorg/Stefan Bodo Würffel (Hgg.), *Utopie und Apokalypse in der Moderne.*, Paderborn 2010, 47–61. Hier 61.

133 Vgl. Klaus Vondung, „Utopische Entwürfe – apokalyptische Visionen. Träume vom besseren Leben“, in: Katharina Martin/Christian Sieg (Hgg.), *Zukunftsvisionen zwischen Apokalypse und Utopie.*, Würzburg 2016, 45–59. Hier 45.

134 Marquard/Odo, „Zeitalter der Weltfremdheit?: Ein Beitrag zur Analyse der Gegenwart“, *Giessener Universitätsblätter* (1985), 9–20. Hier 10.

dass das moderne technologische Bewusstsein überhaupt nur noch den bloßen Untergang akzeptiere und gar nicht mehr den Glauben eines positiven Eingreifens erlaube. Die Katastrophe sei gewissermaßen zur Grundform des Apokalyptischen avanciert, deren Umwendung zum Guten fragwürdig geworden sei. Die Eschatologie der Moderne lasse seinem Urteil zufolge nur noch die *nackte Apokalypse* oder *Apokalypse ohne Reich* stehen, in der nichts mehr der Katastrophe folgt. Abgesehen von antiken Naturphilosophen, die über den Wärmetod des Alls spekulierten, hielt Anders diese durch die neuen Hochtechnologien induzierte Geschichtsbetrachtung für radikal neu. Ihre Fremdartigkeit falle den Menschen umso schwerer, als seit Beginn der Moderne dessen Gegenkonzept, das Reich ohne Apokalypse, absolut selbstverständlich gewesen war, für das es zum Wesen der geschichtlichen Welt gehörte, auf unentrinnbare Weise immer besser zu werden.¹³⁵

Anders Perspektive korrigiert, dass technische Visionen gegenwärtig gar nicht mehr anders Geltung beanspruchen können, wenn sie nicht bereits ihre Katastrophe mitbedenken. Das analytische Potential der Apokalyptik werde gerade nur dann ausgeschöpft, wenn das Schillernde und Widerspruchsvolle, die Nähe und Verschlungenheit von Heil und Verwerfung nicht eingeignet, sondern als Grundmoment des Apokalyptischen immer schon mitgedacht wird.

Dem Religionswissenschaftler Oliver Krüger liegt es ebenfalls auf der Hand, dass das starke endzeitliche Moment der technologischen Singularität durch christliche Eschatologie beeinflusst ist. Im Gegensatz zu Geraci kommt er jedoch zu dem Schluss, dass deren wesentliche Analogie zur christlichen Apokalypse gerade in ihrer Doppeldeutigkeit besteht. Die technologische Singularität werde weder als das Ergebnis einer kontinuierlichen und positiven Fortschrittsentwicklung noch einer totalen Vernichtung betrachtet. Wie die christliche Heilsgeschichte verbinde das Konzept den Untergang des bisherigen Men-

135 Vgl. Günther Anders, „Apokalypse ohne Reich. Radikale Überlegungen zum atomaren Zeitalter“, in: *Die atomare Drohung. Radikale Überlegungen zum atomaren Zeitalter*, München 1993, 207–210. Hier 207.

schen mit der Gewissheit einer postsingulären Verheißung.¹³⁶ Dem ist hinzuzufügen: Je nach Perspektive kann ein und dasselbe kommende Ereignis kann ganz unterschiedlich rezipiert werden. Das Ausbleiben der Künstlichen Superintelligenz wird von den einen mit Erleichterung begrüßt, während es andere verzweifeln lässt, die in ihr die einzige Möglichkeit sehen, die globalen menschenverursachten Krisen zu lösen. Die einen wollen den menschlichen Körper vor jedweder Mechanisierung bewahren, während andere die Zukunft der Welt mit Stolz an die künstlichen, widerstandsfähigeren, letztlich vernünftigeren Roboter abtreten wollen und die Erlösung als ein Ende der menschlichen Spezies begreifen. Diese komplexere Gefühlslage, die über den reinen Optimismus wie Pessimismus hinaus geht, sie scheint gerade das spezifisch Apokalyptische zu sein, in dem Heil und Verwerfung, Entrückung und Plage, Neu Jerusalem und dem Antichristen vermischt und sogar verwechselt werden.¹³⁷

„From one angle, the vision fits many of our happiest dreams; a place unending, where we can truly know one another and understand the deepest mysteries. From another angle, it’s a lot like the worst case scenario I imagined earlier in this paper. Which is the valid viewpoint? In fact, I think the new era is simply too different to fit into the classical frame of good and evil.“¹³⁸

Unter dem Paradigma der Unvergleichlichkeit und Geschichtslosigkeit ihrer Ideen reflektieren die Anhänger der technologischen Singularität, die sich mehrheitlich atheistisch verstehen, die Nähe ihrer Erwartungen zur christlichen Apokalyptik nur selten explizit. Eliezer Yudkowsky äußert sich in seinen *Singularitan Principles* dazu *ex negativo*, als er die Erwartung an die technologische Singularität als rational verstanden

136 Oliver Krüger, *Virtualität und Unsterblichkeit*. Gott, Evolution und die Singularität im Post- und Transhumanismus (Rombach Wissenschaften. Reihe Litterae), 2. Aufl., Freiburg i. Br./Berlin/Wien 2019, 293–294.

137 Vgl. Erik Davis, *Techgnosis. Myth, magic, mysticism in the age of information*, 1. Aufl., New York 1998, 253.

138 Vinge, *The Coming of Technological Singularity: How to Survive the Post-Human Era*, 20.

wissen will und denjenigen, die dieses Ereignis mit religiösen Erwartungen verbinden, – und wie die Aussage vermuten lässt, dies wohl nicht bloß Einzelfälle sind – den Mund verbietet, um die gemeinsame Sache nach außen hin nicht zu diskreditieren:

„The ‚Singularity‘ is a natural, non-mystical, technologically triggered event. In a sense, this is not so much a constraint on belief as a constraint on terminology – if you happen to believe that this planet’s Singularity represents or is identical with the Apocalypse of some particular religion, from the Age of Kali to the Christian Armageddon – or even if you believe that the Singularity is the inevitable destiny of humanity for moral reasons, rather than a possible outcome for causal ones – then that’s perfectly okay, as long as you keep quiet about it.“¹³⁹

Konsequenzen

Geracis Vorschlag, sich den Mythen der Künstlichen Intelligenz unter Zuhilfenahme der apokalyptischen Brille zu nähern, hat ihre inhaltliche Berechtigung. Mit einer solchen Übereinstimmung darf jedoch die Analyse nicht enden, vielmehr beginnt damit erst das systematische Nachdenken.

Erstens wird dann deutlich, dass in ihrer Feinstruktur die Konvergenzmythen um die technologische Singularität nicht bloße Wiederholung des apokalyptischen Paradigmas bleiben, sondern sehr wohl neue Gedanken und Themen formulieren, die der klassischen Apokalypse nicht nur unbekannt ist, sondern ihr geradezu entgegenstehen. So lässt sich an dieser Stelle der Untersuchung bereits feststellen, dass, auch wenn in der klassischen Apokalypse ausschließlich Gott am Menschen handelt, die Menschen weiterhin im Geschehen zentral involviert bleiben, als sie Sinn und Zweck dieses Umbruchgeschehens sind. Während die klassische Apokalypse also noch gleichermaßen theo- wie auch anthropozentrisch ist, verlieren die Technikzukünfte ihr Interesse am Menschen. Die technologische Singularität wird vielmehr zum Wendepunkt erhoben, in dem die Bedeutungslosigkeit des Menschen

¹³⁹ The Singularity Principles, Eliezer S. Yudkowsky: The Singularity Principles Version 1.0.2.

als Übergangswesens im Kosmos zum Vorschein kommt. Des Weiteren sind bestimmte mythische Bilder als auch Konzepte der Konvergenzerzählungen – insbesondere die Rede von der Verlebendigung der Materie als auch der Integration des Menschen in den Maschinenpark – anschlussfähig an magisch-okkulte, insbesondere animistisch-alchemistische als auch spezifisch gnostische Aufstiegs- und Erlösungsvorstellungen (s. 5.3). Schließlich wird deutlich, dass wie wohl bereits immer schon die sadistische Lust am eigenen Untergang und dem Gericht über die verdorbene Menschheit zur Faszination der Apokalypse gehörte, sie hier nun besonders deutlich zutage tritt.¹⁴⁰ Das Ausbleiben des Reiches und die Sehnsucht nach einer *tabula rasa*, in der sich die Probleme und vermeintlichen Missstände der Menschheit mit ihrer Auslöschung ebenfalls auflösen, werden im Angesicht einer vermeintlich *würdigeren* Lebensform offen formuliert.

Die apokalyptische Perspektive wirft in einem nächsten Schritt Nachfragen dahingehend auf, was für einen Sinn die apokalyptische Rhetorik in ihrem *neuen Sitz im Leben* einnimmt. Was hat es zu bedeuten, wenn sich bestimmte Zirkel als diesseitige Erfüller von vormals eschatologischen Versprechen inszenieren oder ihre Aspirationen von der digitalen Allwissenheit und Unsterblichkeit mit solchen bewusst konfundieren lassen.¹⁴¹ Was bedeutet es, wenn apokalyptische Rhetorik, die ursprünglich die Stimme einer bedrängten Minderheit war, nun von einer medial affinen Expertengruppe mit globaler Reichweite ausgeht?¹⁴² Und was sind die Konsequenzen, wenn in diesem Zuge Technikfolgenabschätzung ununterscheidbar von Prophetie wird? Hier greift eine Analyse zu kurz, die die apokalyptische Rhetorik allein auf

140 Vgl. Hans Richard Brittnacher, „Zu Grunde gehen in Sumpf, Schlamm und Fluten. Zur *longue duree* apokalyptischer Metaphorik“, in: Olaf Briese/Richard Faber/et al. (Hgg.), *Aktualität des Apokalyptischen: zwischen Kulturkritik und Kulturversprechen.*, Würzburg 2015, 99–116. Hier 111–113.

141 Vgl. Brin, Broderick, Bostrom, Chislenko, Hanson, More, Nielsen, Sandberg, A Critical Discussion of Vinge's Singularity Concept, 397.

142 „In ihnen verschwimmen eine Resignation angesichts des herrschenden Zukunftsvorstellung und die Vehemenz einer schnellen und grundlegenden Umorientierung.“ Wolfgang Huber, „Fortschritt durch Freiheit. Zur Erwartung politischer Emanzipation“, in: Wolfgang Lienemann (Hg.), *Fortschrittsglaube und Wirklichkeit.* Arbeiten zu einer Frage unserer Zeit, München 1983, 47–65. Hier 60.

ihren Unterhaltungswert und ihre Funktion als effektvolles Werbemittel reduziert. Nimmt man ernst, dass diese Erzählungen als soziotechnische Imaginationen in ihrer Verstetigung von ethischen Grundmustern mehr als bloßes Vergnügen an der Fiktion sind, sondern Mittel und Ausdruck von gesellschaftlicher, wirtschaftlicher wie politischer Macht, dann müssen diese Fragen nach dem Wozu in ihren sozio-polit-ökonomischen Zusammenhang gesetzt werden. Der Darstellung grundlegender Zusammenhänge der Funktionsweise und Machtverhältnisse um das maschinelle Lernen, in denen und aus denen heraus diese Erzählungen ihren gegenwärtigen Resonanzraum beziehen, widmet sich deshalb das folgende Kapitel.

3 Künstliche Intelligenz und maschinelles Lernen

3.1 Künstliche Intelligenz im Kontext der Digitalisierung

Die Technologie um Künstliche Intelligenz steht nicht für sich, sondern reiht sich in den größeren gesellschaftlichen Veränderungsprozess der Digitalisierung ein. Sie bedeutet den zunehmenden Einsatz von digitalen Artefakten, die stetig mehr Daten generieren, verarbeiten und speichern und in deren Zuge sie die unterschiedlichen Lebenswelten tiefer durchdringen und vernetzen. Mittels einer technischen Semantik werden Informationen in das digitale Format des Binärcodes übersetzt. Dieser Code schafft zu analogen Formen digitale Repräsentationen, in deren Datenaustausch und -rekombination Rückschlüsse zu ihren analogen Formen entstehen.¹⁴³ In dieser fortschreitenden Verrechnung verschiedenster Informationen und der Nutzung ihrer Schlussfolgerungen für die Lebenswelt verwirklicht sich die Idee des *Internet of Things* (IoT), das den derzeitigen Erwartungshorizont der Digitalisierung als Ganzes markiert.

Für den Sozialwissenschaftler Armin Nassehi befriedigt die Digitalisierung ein gesellschaftliches Bedürfnis nach Selbststeuerung und Kontrolle, welches er bereits in der Bildung von Nationalstaaten angelegt sieht. Moderne Gesellschaften hält Nassehi an sich bereits für ein digitales Phänomen. Die Digitaltechnik entdecke nun die Gesellschaft in einem viel präziseren Sinne, als dies bisher möglich war, indem sie auf bisher unsichtbare Zusammenhänge hinweist.¹⁴⁴ Mit diesem Offenbarwerden von Regelmäßigkeiten und ihren Abweichungswahrscheinlichkeiten werde die Veränderbarkeit der Gesellschaft wie ihre veränderungsresistente Trägheit sichtbar. Weil die durch die Apparate aufgezeichneten Verbindungen, Strukturen und Muster weder ein-

¹⁴³ Vgl. Ralph Charbonnier, „Digitalisierung: Theologische Selbstklärungen und Gegenwartsinterpretationen. Eine Skizze“, *Zeitschrift für Pädagogik und Theologie* 70 (2018), 238–250. Hier 240.

¹⁴⁴ Nassehi, *Muster*, 249.

deutig festgelegt noch beliebig, sondern für verschiedenste Anfragen rekombinierbar blieben, erzeuge die Digitalisierung Ordnungen in nie dagewesener Form.¹⁴⁵ In der Überbefriedigung dieses Bedürfnisses werfe die Digitalisierung somit das Problem der Selbstkontrolle gesellschaftlicher Praktiken auf, wie es sich in erheblich einfacheren, hierarchischen Formen der Gesellschaftsordnung noch nicht stellte. Indem die moderne digitale Gesellschaft sich mit und in Daten verdoppelt, die quer zu den geübten Praktiken der Gesellschaft liegen, komme es zu sozialen Spannungen.¹⁴⁶ Deren Aufarbeitung erschwere der unspezifische Charakter des Digitalen. Es scheint indifferent für fast alles zu gebrauchen:

„Es ist eine geradezu ubiquitäre Form – bis dato nur der Präsenz Gottes und dem Einsatz von Schrift vorbehalten.“¹⁴⁷

Diese maximale Offenheit der digitalen Technik erstreckt sich insbesondere auf das Einsatzgebiet der Künstlichen Intelligenz. Als Schlüsseltechnologie deklariert, scheint es, dass sämtliche Lebensbereiche und Aufgaben, die elektrisiert werden konnten, nun digitalisiert und schließlich intelligent gemacht werden. Ähnlich zur Elektrizität arbeiten heutige intelligente Anwendungen im Alltag für den Einzelnen nahezu unbemerkt. Sie passen Internetwerbung auf das Nutzerverhalten an, sind verantwortlich für die Kaufempfehlungen im Onlinehandel, beantworten auf Webseiten häufig gestellte Fragen oder sind auf Mobilgeräten oder durch *Smart Speaker* im Haus als Sprachassistenten im Dienst. Neben diesen Alltäglichkeiten übernimmt KI heute bereits Arbeiten, die ehemals Akademikern vorbehalten waren: *Legal Tech* löst juristische Probleme, *Med Tech* erkennt Krankheiten, *Fin Tech* vermittelt Kredite, *Insure Tech* berät zu Versicherungen und journalistische KI verfasst erste Kurzmeldungen. Auch wenn dies von Anbietern digitaler Dienste bereits ausgemachte Sache zu sein scheint, bleibt bei all diesem Vorpreschen offen, ob die Digitalisierung mit all ihrem exor-

145 Vgl. ebd., 53.

146 Vgl. ebd., 140–141.

147 Ebd., 35.

bitanten Energie- und Ressourcenverbrauch dem Menschen und der Natur letztlich dient.¹⁴⁸

Diesem Vorwurf gegenüber scheint die Digitalwirtschaft die Annehmlichkeiten ihrer datenintensiven *Gadgets* dem Einzelnen erweisen zu wollen, indem diese immer weiter in sämtliche Lebensbereiche eindringen. Im privaten Umfeld mag der gefühlte Komfort zwar das Unbehagen gegenüber der Nutzung von KI-Systemen mehr und mehr verdrängen. Dennoch bleibt Künstliche Intelligenz ein Thema, das viele Menschen – wie die Digitalisierung als Ganzes – nicht vollständig greifen und einschätzen können. Die Konfusion selbst innerhalb der Fachwissenschaften führt gar zu der Frage: „*Who knows anything about anything about AI?*“¹⁴⁹ Dass dieses Informationsdefizit auf ein strukturelles Problem hinweist, indem die Forschung größtenteils privatwirtschaftlich stattfindet und damit entgrenzte Potentialerwartungen ein relevantes Vermarktungsinstrument darstellen, ist eine wichtige Rahmenbedingung für die ethische Diskussion ihrer Mythen. Der vorfindlichen Technik des maschinellen Lernens und ihrem wirtschaftspolitischen Kontext, die die Projektionsfläche und den lebensweltlichen Ankerpunkt für die KI-Mythen darstellen, widmet sich deshalb dieses Kapitel.

3.2 Entstehung und Funktion von Künstlicher Intelligenz

Hinter intelligenten Programmen stand und steht bis heute die Vorstellung, ein System zu entwickeln, das eigenständig komplexe Probleme bearbeitet und sich dabei verbessert. Welche Probleme als komplex gelten und damit die Grenzlinie von KI zu einfacher Software markieren, unterlag über Jahrzehnte hinweg stets der graduellen Beurteilung. Im Zuge der neueren Entwicklungen hat sich jedoch die Bezeichnung für solche Programme durchgesetzt, die Mustererkennung auf Grund-

148 Sarah Spiekermann, *Digitale Ethik*. Ein Wertesystem für das 21. Jahrhundert, München 2019, 10.

149 Stuart Armstrong/Sean OhEigeartaigh, „Who Knows Anything about Anything about AI?“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound*. The Future of Uploaded and Machine Minds, Chichester, West Sussex 2014, 46–60.

lage des Maschinellen Lernens (*Machine Learning*) betreiben. Während der Mathematiker Alan Turing bereits 1950 die These aufstellte, dass Maschinen potentiell Intelligenz demonstrieren könnten,¹⁵⁰ geht der Begriff der *Artificial Intelligence* auf einen Forschungsantrag einer Sommerakademie aus dem Jahr 1956 zurück, dem *Dartmouth Summer Research Project on Artificial Intelligence*. Diese Veranstaltung gilt als die Geburtsstunde des Forschungszweiges. Zwei Dutzend Wissenschaftler, darunter Mathematiker, Informationstheoretiker, Kybernetiker, Elektronik-Ingenieure, Psychologen und Ökonomen gingen der Frage nach, wie man in einer Maschine menschliches Lernen simulieren könne. Ihr größtgestecktes Ziel beschrieben sie im Förderantrag der Tagung an die Rockefeller-Stiftung folgendermaßen:

„The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.“¹⁵¹

Der Wortschöpfer der *Artificial Intelligence* war der junge Logiker John McCarthy, neben Marvin Minsky vom MIT und Arthur Samuel von IBM (*International Business Machines Corporation*) einer der drei Antragsteller der Sommerakademie. Er selbst war wohl mit dem Begriff unzufrieden, doch besonders dessen Kurzform *AI* war griffig und nützte in den kommenden Jahrzehnten hervorragend für die Vermarktung der eigenen Forschung.¹⁵² Zwanzig Jahre nach Dartmouth entstanden die ersten Expertensysteme, die über Wenn-Dann-Beziehungen probierten, eine menschliche Wissensbasis in für Computer lesbare Informationen zu verwandeln. Mit den Möglichkeiten zu logischen Schlussfolgerun-

150 Vgl. Turing, *Computing Machinery and Intelligence*, 433.

151 John McCarthy/Marvin L. Minsky/Nathaniel Rochester/and Claude E. Shannon, „A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955“, *AI Magazin* 27 (2016), 12–14. Hier 12.

152 Vgl. Thomas Ramge, *Mensch und Maschine*. Wie künstliche Intelligenz und Roboter unser Leben verändern (Nr. 19499), Ditzingen 2018, 31–34.

gen und dem effektiven Suchen in diesen Wissenszentralen konnten die Systeme zunächst einige Erfolge aufweisen und weckten bis in die 1980er Jahre große Erwartungen an die KI. Die notwendige Datenpflege erwies sich jedoch als zu aufwendig, sodass nach wenigen Jahren viele Firmen ihre Systeme wieder abschafften.¹⁵³

Das Wiederaufleben der KI-Forschung zu Beginn des 21. Jahrhunderts

Bereits Ende der 1990iger sah Hans Moravec wieder bessere Rahmenbedingungen für KI am Horizont heraufziehen. Seine Gewissheit für eine Wiederkehr der Technologie formulierte er – auffallend ähnlich zur alttestamentlichen Heilsprophetie des Jesaja – im Bild einer erblühenden Wüste (Jes 35,1–7):

„Wie eine Wüste im Regenguss beginnt das gesamte Feld der Roboterintelligenz nach dreißig Jahren Trockenheit zu erblühen. [...] Rechenleistung ist für den Geist, was die Lokomotive für den Zug ist.“¹⁵⁴

Moravec sah nicht voraus, dass die gestiegene Rechenleistung nur ein Grund sein würde, der diese Fortschritte ermöglichte. In der Retrospektive sind es vor allem drei Kriterien, die für die Rückkehr Künstlicher Intelligenz seit Beginn des neuen Jahrtausends verantwortlich sind:

Große Datenmengen: Ein Teil des KI-Durchbruchs liegt in der unglaublichen Ansammlung von Daten. Massive Datenbanken, *Self-Tracking*, *Web-Cookies*, *Online-Footprints*, Terabytes an Speicherplatz mit Jahrzehnten an Suchergebnissen, Wikipedia und das gesamte digitale Universum bilden das Lehrmaterial für Künstliche Intelligenz.

Bessere Programmierung: Digitale neuronale Netze wurden in den 1950er Jahren erfunden, doch dauerte es fast 50 Jahre, bis Forscher praktischen Nutzen aus ihnen ziehen konnten. Im Jahr 2006 gelang Geoffrey Hinton von der Universität Toronto eine entscheidende Verbesserung dieser Methode, die er Tiefes Lernen (*Deep Learning*) nannte. Dieser

153 Vgl. Moritz Kirste/Markus Schürholz, „Entwicklungswege zur KI“, in: Volker Witt-pahl (Hg.), *Künstliche Intelligenz*. Technologie, Berlin, Heidelberg 2019, 21–35. Hier 24.

154 Moravec, Computer übernehmen die Macht, 83.

Code ist ein wesentlicher Bestandteil aller aktuellen KIs, einschließlich IBMs Watson, Googles Suchmaschine und Facebooks Algorithmen.

Günstige Parallelrechner: Auf Basis eines neuartigen Chips, der sogenannten *Graphics Processing Unit* (GPU), die ursprünglich für die intensiven Anforderungen von Videospiele entwickelt wurde, erkannten im Jahr 2009 Andrew Ng und ein Team in Stanford, dass GPU-Chips neuronalen Netzwerke die notwendige Rechenleistung geben, die vorher nicht gegeben war.¹⁵⁵

Mustererkennung, Maschinelles Lernen und Künstliche Neuronale Netze

Diese Durchbrüche haben zur Funktionsweise der neueren Künstlichen Intelligenzen geführt. Sie erreichen die ihr gestellten Ziele mittels Mustererkennung, indem sie selbstständig in einer Menge von Daten Regelmäßigkeiten, Wiederholungen oder Gesetzmäßigkeiten erkennen. Sie verbessern ihre Trefferquote für die relevanten Muster, indem sie die Auswirkungen ihrer vorangegangenen Entscheidungen messen, mit dem Soll abgleichen und dies in ihre künftige Entscheidungsfindung miteinbeziehen. Anders als bei traditionellen Programmen treffen Daten also nicht auf einen statischen Code und produzieren ein Ergebnis, sondern ein Algorithmus¹⁵⁶ passt seine Hypothese über die Rückmeldungen seiner vorangegangenen Versuche an, um schließlich bessere Ergebnisse zu liefern.¹⁵⁷ Für dieses Vorgehen hat sich die

155 Vgl. Kevin Kelly, *The Three Breakthroughs That Have Finally Unleashed AI on the World*, <https://www.wired.com/2014/10/future-of-artificial-intelligence/> (9.6.2020).

156 Bei Algorithmen handelt es sich um effektive Verfahren, das heißt eindeutige, wiederholbare Handlungsvorschriften. Vgl. Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 74.

157 Eine Arbeitsdefinition legte 2019 das Expertenteam der Europäischen Union vor: „*Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.*“ European Council on Ethics in Science and Technology, *A Definition of AI: Main Capabilities and Disciplines* 08.04.2019, 8.

Bezeichnung des maschinellen Lernens bzw. *Machine Learning* durchgesetzt und gilt als der Schlüssel zu Künstlicher Intelligenz.

Der aktuell leistungsfähigste Ansatz des Tiefen Lernens baut auf dem mathematischen Konstrukt der Künstlichen Neuronale Netze (KNN) auf. Eine große Zahl an kleinsten Recheneinheiten (Perzeptrone) verarbeitet dabei in mehreren Schichten die eingegangenen Informationen. Tief bedeutet in diesem Zusammenhang unabhängig von der genauen Netzstruktur, dass es sich um mehrere Schichten handelt: KNNs lassen sich unterteilen in eine *Input*-Schicht, über die externe Daten wie Bilder, Audiodaten oder Text in das Netz eingespeist werden, in eine oder mehrere *Zwischenschichten (Hidden Layers)*, die die Daten verarbeiten und in eine *Output*-Schicht, die das Resultat der Millionen an Gewichtungen ausgibt. Neuronale Netze sind in rudimentärer Weise von ihren biologischen Gegenstücken inspiriert, weil die Perzeptrone ähnlich wie Neuronen über Aktionspotentiale Impulse an die nächste Schicht weiterleiten und so über die Gewichtung von Schwellenwerten ein heuristisches Profil ausbilden. Durch zahlreiche Testläufe mit einer hohen Zahl an Testmaterial optimiert es die einzelnen Werte in den Perzeptronen so lange, bis die Ausgabe dem gewünschten Ergebnis entspricht. Das eigentliche Wissen des Netzes steckt somit in der Gewichtung der einzelnen Verbindungen in den Zwischenschichten.¹⁵⁸ Ein solches trainiertes KNN kann seinen abstrahierten Lösungsweg auf neue Daten anwenden und so komplexe Zuordnungsaufgaben übernehmen. Der Programmierer wirkt somit allein über verschiedene Grade der Auswahl und Kennzeichnung der Trainingsdaten und der Rückmeldung über die Richtigkeit von Ausgabewerten auf den Algorithmus ein:

Beim überwachten Lernen (*Supervised Machine Learning*) bekommt ein Programm bekannte Beispieldaten und wird auf eine gewünschte Interpretation und die damit verbundene Ausgabe trainiert. Unbewachtes Lernen (*Unsupervised Machine Learning*) funktioniert hingegen ohne vorher bekannte Zuordnung und Kennzeichnung von Eingabedaten durch den Programmierer. Das Computerprogramm muss Ähnlichkeiten in den Daten erkennen und aus diesen eigene Klassifikationsklassen entwickeln. Hier liegt ein großer Sprung: Während

158 Vgl. Kirste, Schürholz, Entwicklungswege zur KI, 29.

implizites Wissen, dass sich schlecht an Parametern festmachen lässt (z. B. das Erkennen eines Gesichts), die bisherige Grenze der Software-Entwickler war, setzt der Mensch nun mithilfe von KI lediglich den Rahmen, in welchem die Maschine agiert. Das verstärkte Lernen (*Reinforcement Machine Learning*) unterscheidet sich von den anderen beiden Formen, da es ohne Trainingsdaten auskommt. Hier lernt ein Computerprogramm, indem es mit eigens angelegten spielähnlichen Simulations-Szenarien konfrontiert wird, welche es in beliebig vielen *Trial-and-Error*-Versuchen durchläuft. Obwohl der Programmierer die Belohnungspolitik – also die Spielregeln – festlegt, gibt er dem Modell keine Vorschläge, wie das Spiel zu lösen ist. Die Belohnung ist dementsprechend die Regelgröße, nach der aus den Rückkopplungsschleifen aus Tausenden von parallelen Spielversuchen das Programm seine Aktionen ausrichtet.¹⁵⁹

Die letzte Methode erfüllt sehr spezifische Aufgaben. Die Erfolge um maschinelles Lernen bauen zum großen Teil auf die datenintensive Mustererkennung. Trotz der wissenschaftlichen Bestrebungen, synthetische Trainingsdaten aus einem kleineren Datenpool zu generieren, bleibt maschinelles Lernen somit auf eine enorme Menge an qualitativ hochwertigen Daten angewiesen. Während die untrainierten Algorithmen als auch die CPUs in gewissen Graden jedem zugänglich sind, hängen die ungleichen Startbedingungen für KI-Forschung eng mit der Verteilung von Datenressourcen zusammen. Solange *Big Data* in den Händen weniger digitaler Megakonzerne liegt, bestimmen diese die Forschung, Entwicklung und Implementierung neuer Anwendungen.

3.3 Wirtschaftliche Konzentration, digitaler Kapitalismus und kalifornische Ideologie

Die Entwicklung intelligenter Technologie lässt sich nicht verstehen ohne ein Bewusstsein für die außerordentliche Machtkonzentration in den Händen weniger Internetgiganten. Prima-Facie Behauptungen von

159 Vgl. ebd.

dem Freiheitszog der digitalen Welt und ihrem Anschub durch Künstliche Intelligenz verkennen die soziale und ökonomische Realität, in der Google, Facebook (Meta), Microsoft, Apple und Amazon maßgeblich entscheiden, wie Nutzer die digitale Welt erleben. Innerhalb von weniger als zwei Jahrzehnten sind sie allesamt zu milliardenschweren, börsennotierten Unternehmen aufgestiegen, die sich über gesellschaftliche Ordnungen stellen und die für Institutionen und Rechtsordnungen gleichsam opak bleiben. In dieser Konzentration von technologischer, wirtschaftlicher und politischer Macht hat sich für diese Firmen der Begriff der *Frightful Five* (die furchteinflößenden Fünf) etabliert.¹⁶⁰

Während die bestimmenden Technologien des 20. Jahrhunderts – allen voran die Nukleartechnik – von Staaten und ihren Laboren ausgingen, liegen die transformativen Technologien des 21. Jahrhundert, zu denen man neben Künstlicher Intelligenz und Robotik noch die Genetik und die Nanotechnologie zählt, in der Verfügungsgewalt multinationaler Unternehmen.¹⁶¹ Sie sind weder dem Gemeineigentum noch irgendeiner Neutralität verpflichtet, sondern handeln in ihrer Kommunikations- und Geschäftspolitik im Sinne der Gewinnmaximierung. Während ihre KI-Grundlagenforschung zum Teil öffentlich zugänglich ist, findet die ressourcenintensivere Arbeit an der Systemintegration und an konkreten Anwendungen zur kommerziellen Nutzung im Verborgenen statt. Deren Forschungsgelder übersteigen die der öffentlichen Investitionen in vielen Ländern.¹⁶² Der Datenhunger des maschinellen Lernens fördert diese digitale Oligarchie weiter: Diejenigen mit der größten hochwertigen Datenbasis können die zielgenaueren Algorithmen auf den Markt bringen. Umso besser diese funktionieren, desto mehr Menschen nutzen sie. Ihre vermehrte Nutzung erzeugt wiederum neue Daten für die Konzerne zur weiteren Verbesserung ihrer Programme. Sobald ein Unternehmen in diesen positiven Kreis-

¹⁶⁰ Vgl. Paul Nemitz, „Constitutional Democracy and Technology in the Age of Artificial Intelligence“, *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 376 (2018), 1–14. Hier 2–4.

¹⁶¹ Vgl. Bill Joy, Why the Future Doesn't Need Us, 1.4.2000, <https://www.wired.com/2000/04/joy-2/> (12.8.2020).

¹⁶² Vgl. Nemitz, *Constitutional Democracy and Technology in the Age of Artificial Intelligence*, 4.

lauf eintritt, neigt es dazu, so groß und schnell zu wachsen, dass es alle aufstrebenden Konkurrenten überrollt. So sind die Gewinner der Dotcom-Ära auch die Treiber des maschinellen Lernens. Eine weitere Verengung des Marktes in der Zukunft hin zu lediglich zwei oder drei großen, cloudbasierten kommerziellen Datensets scheint aus heutiger Sicht nicht unwahrscheinlich.¹⁶³

Gegenstaatlichkeit und Disruption als Kennzeichen der Silicon-Valley-Ideologie

Diese kumulative Macht der Unternehmen ist verbunden mit einem gemeinsamen Sendungsbewusstsein, die Welt aus ihren Angeln zu heben. Im Gegensatz zu den Fabrikmagnaten des 19. und 20. Jahrhunderts besitzen sie keine gigantischen Produktionsmittel. Stattdessen sind die *Frightful Five* zum Inbegriff eines immateriellen oder digitalen Kapitalismus geworden. Sie erwirtschaften ihren Gewinn mit Informationsverarbeitung und besitzen das Stoffliche, das daran hängt, meist überhaupt nicht.¹⁶⁴ Diese neue Form des Wirtschaftens geriert sich dabei so, dass allein aus der richtigen Geisteshaltung die Ideen erwachsen, die die Welt verändern. Der gemeinsame Ursprungsort der Fünf ist das *Silicon Valley*, ein Raum mittelgroßer Gemeinden zwischen San Jose und San Francisco im US-Bundesstaat Kalifornien, der mit dem Aufstieg dieser Unternehmen selbst längst zu einem mythischen Ort avanciert ist, an dem viele an diesem *Mindset* zu partizipieren suchen. Populärwissenschaftliche Werke glorifizieren das *Valley* und die ihr zugeschriebene Mentalität nicht selten zu einer wundersamen Gegenwelt. Deren Berichten gleichen klassischen Utopieromanen, in denen die andere Gesellschaftsform Einblicke in eine bessere, dort punktuell erfahrbare Zukunft für den Rest der Welt gibt; im Unterschied zur Utopie jedoch kennt die Welt den Weg zum sagenhaften „Weltgeist am Pazifik“¹⁶⁵.

Der missionarische Geist, der durch ihre Produkte und Dienstleistungen in die restliche Welt ausstrahlt, ist von dem Ideal der Disruption beseelt: Allgemein meint es das (gewaltsame) Unterbrechen des

163 Vgl. Kelly, *The Three Breakthroughs That Have Finally Unleashed AI on the World*.

164 Vgl. Nassehi, *Muster*, 191–193.

165 Hans Ulrich Gumbrecht, *Weltgeist im Silicon Valley*. Leben und Denken im Zukunftsmodus, Basel 2018, 33.

Bestehenden, des Tradierten und Gewachsenen. Konkret bedeutet es, dass keine Erfindung oder Idee davor sicher ist, nicht von einer besseren in zeitlich kurzer Abfolge ersetzt zu werden. Diese Einstellung erfordert ein *semper paratus*; das ganze *Silicon Valley* scheint auf ihren ganz eigenen *Kairos* fokussiert, dem Zeitpunkt zwischen Trends und ihrer jeweiligen Aufhebung. Dieser Disruptionsgedanke wirkt auch in die Unternehmensführung hinein, als sich von langfristiger Planung oder gar von Unternehmensentwicklung distanzieren wird und der Planungsverzicht vielmehr als wesentliche Voraussetzung einer finanziellen wie intellektuellen Agilität gefeiert wird.¹⁶⁶ Dass der Gedanke der Disruption sich so festsetzen konnte, hängt mit der besonderen Kultur der kalifornischen Ideologie zusammen, deren Wurzeln bis in die Jugendbewegung der 1960er Jahre zurückreichen. Diese ging von einem starken Impuls für persönliche Freiheit und individuelle Ermächtigung durch Dezentralisierung aus und richtete sich gegen die von Washington dominierte Politik und die Vorherrschaft des damaligen *IBM-Mainframes*. Ihre erste wichtige Materialisierung fand dieser Protest in der Entwicklung des *Personal Computers* und setzte sich fort in der zivilen Nutzung des Internets.¹⁶⁷

Das Versprechen dieser Jugendbewegung lag in grenzen- und kostenloser Kommunikation, in der Vision einer kreativen und kommunikativen Gemeinschaft frei zusammenwirkender Menschen, einzig verbunden in einer revolutionären Idee von Freiheit und Gleichheit. Ihre aus diesem Ideal geborenen Produkte sollten diese Vision vorantreiben. Die Theologin und Medienethikerin Johanna Haberer erkennt im Kern der kalifornischen Ideologie einen Mythos von biblischer Dimension, in dem der irdische Gott der Kommunikationstechnik die babylonische Verwirrung beendet und den Menschen eine bessere Form der Verständigung schenkt, durch die die Gewalt der Wenigen über die vielen überwunden wird.¹⁶⁸ Philosoph Julian Nida-Rümelin deutet die kalifornische Ideologie als ideellen Ausläufer des Puritanismus, indem im

166 Ebd., 66.

167 Vgl. Nemitz, *Constitutional Democracy and Technology in the Age of Artificial Intelligence*, 4–5.

168 Vgl. Johanna Haberer, *Digitale Theologie. Gott und die Medienrevolution der Gegenwart*, München 2015, 131.

Glauben an Transparenz, Berechenbarkeit, ökonomischem Erfolg und mäzenatischem Engagement versucht werde, durch Software eine Welt der Reinen und Gerechten zu schaffen. Dieser Versuch stehe jedoch gleichsam in der Gefahr, aus seinem ursprünglich humanistischen Impuls auf halber Strecke in antihumanistische Utopien umzuschlagen, wenn in ihrer Mission, menschliches Leben auf dem Planeten verbessern, die Bedingungen der Humanität selbst infrage gestellt würden.¹⁶⁹

Durchsetzen lässt sich dieses puritanische Versprechen anscheinend lediglich in der dauerhaften Auflehnung gegen die Ordnungen außerhalb des *Silicon Valley*, die den von dort ausgehenden Impuls zu ersticken drohten. Die Lehre von der Disruption führt dabei über das Technische hinüber zur Wirtschaft bis in die Gesetzeswelt: Sie verträgt sich nicht gut mit der Idee des Vorsorgegedankens oder der gesetzlichen Regulierung oder zu langsamer, politischer Konsensfindung. Technik, die funktioniert, schaffe hingegen die besseren Lösungen vorbei an der Politik. Solche Skepsis führt zu den Bedenken, dass Regierungen das Internet und die neuen Technologien wie KI sowieso nicht verstünden und daher keine Legitimation für ihre Regulierung hätten. Manche sehen die Erfüllung des Traumes der *Silicon Valley* in der gänzlichen Aufhebung des klassischen Staats durch opulente, private Spendenorganisationen und extensive Unternehmen, die Staatsfunktionen übernehmen und damit auch erheblich verdienen sollen.¹⁷⁰ Der digitale Plattformkapitalismus mit seiner kalifornischen Ideologie strebe zuletzt ein neues Gemeinwesen an, eine *Polis der Solution*¹⁷¹, in der das technologisch Mögliche zum Soll erhoben wird, an deren Prozesse der Mensch sich anzupassen habe. Dafür verspreche die Polis ihren Bürgern – oder eher: Kunden – das technologisch-unternehmerische Lösen sozialer Probleme, geht sie doch von einer natürlichen Übereinstimmung zwischen Weltverbesserungs- und Geschäftsmöglichkeiten aus. Waren es in Platons Staatslehre die Philosophen und in den klassischen Demokratien die gewählten Volksvertreter, so findet diese Polis ihren idealen Vorsteher im Weltverbesserunternehmer, der sich in neuester Iteration

169 Vgl. Nida-Rümelin, Weidenfeld, *Digitaler Humanismus*, 22.

170 Vgl. Gumbrecht, *Weltgeist im Silicon Valley*, 68.

171 Vgl. Oliver Nachtwey/Timl Seidl, *Die Ethik der Solution und der Geist des digitalen Kapitalismus*, Frankfurt am Main, 21–22.

des Gedankens mittels KI nun selbst überflüssig macht und den Maschinen das rechte Regieren überlässt.¹⁷²

Die ausbleibende Regulierung der Digitalwirtschaft

Weil die *Frightful Five* den direkten Zugang zu ihren Nutzern beanspruchen, führt das notorische Untergraben von staatlicher Souveränität nicht nur zur systematischen Verweigerung, Steuern zu zahlen, sondern auch zu direkter Konfrontation mit Nationalstaaten. Beispielfhaft sei hier auf die Eskalation des Konflikts zwischen dem Parlament Australiens und Facebook im Februar 2021 verwiesen: Weil Facebook das vom Unterhaus in Canberra verabschiedete Gesetz ablehnte, für Inhalte von Verlagshäusern oder Sendern zu zahlen, blockierte das Unternehmen jegliche Nachrichten-Inhalte auf seinen australischen Servern und schaltete zusätzlich viele offizielle Seiten von staatlichen Institutionen ab. Die Häufung solch erpresserischen Verhaltens hat in den letzten Jahren zu einem Bewusstsein auf politischer Ebene geführt, dass nur supranationale Verbände diesen Unternehmen einen Rechtsrahmen setzen können. Im Bewusstsein der katastrophalen Folgen, die das nicht regulierte Internet verursacht hat und die sich rückwirkend als unheimlich schwer zu beheben darstellen, soll diesmal vorsorglich – oder zumindest zeitnah – ein internationaler Rechtsrahmen (*Precautionary Legal Framework*) entstehen, um das öffentliche Interesse in der Entwicklung und Einsatz von KI zu schützen.¹⁷³ So brachte die Europäische Union im April 2019 eine Ethikrichtlinie zur Strategie *AI made in Europe* heraus, deren Fokus auf den Begriff des Vertrauens in der Rede von einer *Trustworthy AI* auf Datenschutz und das Ideal einer Erweiterung, nicht dem Ersetzen menschlicher Entscheidungsspielräume verwies. Ihren Ansatz versteht die EU als bewusste Alternative zu dem libertären Modell des *Silicon Valley* als auch den freiheitsminimierenden Staatsalgorithmen Chinas. Ein vertrauenswürdiger Ansatz, wie ihn die EU anstrebt, ermögliche einen verantwortungsvollen Wettbewerb, in welchem die KI-Systeme innerhalb der EU in ihrem Design, ihrer

172 Vgl. ebd., 23.

173 Vgl. Nemitz, *Constitutional Democracy and Technology in the Age of Artificial Intelligence*, 6.

Entwicklung und ihrem Einsatz nach rechtlich konform, ethisch und robust (*lawful, ethical, robust*) gestaltet sind.¹⁷⁴

Darauf aufbauend legte die EU-Kommission im April 2021 dem Parlament einen Rechtsrahmen für Künstliche Intelligenz zur Ratifizierung vor. Er soll gewährleisten, dass KI-Systeme, die in der Union in Verkehr gebracht und genutzt werden, das geltende Recht in Bezug auf die Grundrechte und die Werte der Union beachten; dabei geht es auch um die Schaffung von Rechtssicherheit für Forschung und Unternehmen, um Investitionen im Bereich KI zu erleichtern; zudem soll das Konzept die wirksame Durchsetzung des bestehenden Rechts die geltenden Sicherheitsanforderungen verbessern und so einen pluralen Binnenmarkt für rechtmäßige und sichere KI-Anwendungen etablieren. Konkret verboten wird die wahl- und anlasslose Überwachung, darunter die Gesichtserkennung im öffentlichen Raum mit Ausnahmen für die Suche nach Straftätern oder Vermissten. Verboten werden soll auch die Bewertung des Verhaltens von Menschen im Sinne eines *Social Scoring*; ebenso solche Anwendungen, die das menschliche Verhalten zu gefährlichem Verhalten hin manipulieren. Der Einsatz von Hochrisikooanwendungen, definiert durch einen unbestreitbaren Nutzen bei potentiell irreparablen Schäden, wird unterdessen an strikte Bedingungen geknüpft: Die Daten, mit denen sie versorgt werden, müssen so ausgewählt sein, dass niemand benachteiligt wird. Zudem muss immer ein Mensch die letzte Kontrolle haben. Darunter fallen der gesamte Bereich der kritischen Infrastruktur, besonders risikobehaftete Anwendungsbereiche wie das autonome Fahren, Bewerbungen auf Universitätsstellen, die Verteilung öffentlicher Leistungen, die Vergabe von Krediten, Grenzkontrollen, Asylanträge und die Strafverfolgung. Bei harmloseren Anwendungen soll zumindest ersichtlich werden, dass Menschen es mit einer KI zu tun haben.¹⁷⁵ Der Ruf nach Regulierung aus den großen Tech-Unternehmen selbst beschränkt sich auf ausgesuchte Bereiche wie

174 Vgl. European Council on Ethics in Science and Technology, *Ethics Guidelines for Trustworthy AI*, Brüssel 08.04.2019, 4.

175 Vgl. European Council on Ethics in Science and Technology, *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*, Brüssel 21.04.2021, 3.

die Sperrung bestimmter Inhalte oder das Einschreiten gegen Wahlmanipulation. Einer zu starken Einhegung durch internationales Recht versuchen die Unternehmen mit der Einrichtung von firmeneigener Ethik Boards zu umgehen, sodass Ethik als Ausflucht vor Rechtssetzung instrumentalisiert wird. Diejenigen schließlich, die eine Charta der Technologien erlassen, sind dieselben, die gegen europäische Vorschriften verstoßen und beispielsweise individuelle Forderungen unter Berufung auf das Recht auf Vergessen nicht nachkommen.¹⁷⁶

3.4 Spezialistentum, Unwissenheit und Macht: Intellectual Debt und Black Boxes

Den Plattform-Unternehmen bringt es einen strategischen Vorteil, ein Wissensgefälle für die Einschätzung ihrer technologischen Möglichkeiten sowohl gegenüber wirtschaftlichen Konkurrenten als auch gegenüber ihren Kunden aufrechtzuerhalten. Parallel dazu baut das maschinelle Lernen selbst Wissenslücken auf. Beide Dimensionen der Unwissenheit fördern nicht nur die Bildung neuer Mythen, sondern sie führen zu der Frage, ab welchem Grad diese Undurchsichtigkeiten einem verantwortlichen Einsatz intelligenter Technik im Wege stehen.

Machtgefälle zwischen Unternehmen und Anwendern

Bereits zu Zeiten vor den großen Plattformen klagte Joseph Weizenbaum über ein ungleiches Machtverhältnis zwischen Programmierern und ihren Kunden ob ihrer manipulativen Gewalt:

„Der Programmierer ist jedoch der Schöpfer von Universen, deren alleiniger Gesetzgeber er selbst ist. [...] Kein Dramatiker, kein Regisseur und kein noch so mächtiger Herrscher haben jemals eine so absolute Macht ausgeübt, eine Bühne oder ein Schlachtfeld zu arrangieren und dann so unerschütterlich gehorsame Schauspieler bzw. Truppen zu befehligen.“

176 Vgl. Ganascia, *Le Mythe de la Singularité*, 108.

Es wäre erstaunlich, wenn Lord Actons Beobachtung, dass Macht korumpiert, nicht für eine Umwelt gelten würde, in der Omnipotenz so leicht zu erringen ist.¹⁷⁷

Weizenbaum war überzeugt, Programmierer könnten als Konstrukteure der digitalen Welt ihre Kunden bedingungslos auf ihre Visionen verpflichten. Deshalb forderte er von Programmierern mehr Rechenschaft für ihre Produkte und transparenteres Arbeiten im Allgemeinen.¹⁷⁸ Man muss Weizenbaums Analyse in ihrem Extrem nicht beipflichten, ohne zu verleugnen, dass die digitalen Anwendungen Nutzern Freiheiten unter nicht offengelegten Bedingungen gewähren; die Auswahlmöglichkeiten, die ein Programm dem Benutzer eröffnet, täuschen über die Festlegungen hinweg, die der Programmierer bereits vorweggenommen hat und er mit seiner bestimmten Architektur des Programmes nicht mehr zur Wahl stellt. Vor der Eingabe- und Ausgabefläche bleiben Nutzer vergleichsweise passiv, die Führung durch Menüs nach unsichtbarer Hand ist Resultat eines Wissens- und damit letztlich Machtgefälles.¹⁷⁹

Die Anwendungskompetenz allein ist jedoch nicht ausreichend, um bestimmte Inszenierungen der Tech-Industrie einschätzen zu können. Während versucht wird, die digitalen Produkte immer intuitiver zu gestalten, haben Unternehmen wenig Interesse daran, dass Nutzer Kompetenzen über das reine Anwendungswissen hinaus erwerben. Schließlich bevorteilt es bestimmte Gruppen unverhältnismäßig, wenn technische Möglichkeiten zu hoch eingeschätzt werden: Investoren in Jungunternehmen verdienen, wenn andere Investoren glauben, dass die Neuerkenntnisse ihres Unternehmens die Industrie disruptieren werden. Polizei und Verwaltung profitieren, wenn die Bevölkerung den Einsatz neuester Gadgets als Erweis für innovative Führung auslegt. *Social-Media*-Unternehmen profitieren, wenn Nutzer glauben, dass KI die Moderation von Inhalten in großem Umfang bewältigen könne. Technologieunternehmen schließlich profitieren, wenn ihre

177 Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 156–160.

178 Ebd., 356.

179 Spiekermann, *Digitale Ethik*, 52.

Erfindungen zur Lösung fundamentaler Probleme ohne Nebeneffekte erscheinen. Die Profiteure dieser Kette können die Gesellschaft aktiv dazu ermutigen, zu überschätzen, was eine Technologie wirklich leisten kann, damit sie Gespräche darüber kurzschließen, was Technologie tun sollte.¹⁸⁰ Somit ergibt sich für den Kunden ein enormes Vertrauensdefizit, dass durch in der Anwendung immer intuitivere Technik nicht gelöst wird. Die Wirtschaftswissenschaftlerin und Philosophin Sarah Spiekermann sieht die öffentliche Debatte im internationalen als auch deutschsprachigen Raum durchsetzt von Machtinteressen, die neben *Fake News* nun zunehmend auch *Fake Science* hervorbringe und von „präpotenten Experten und mangelhaft gebildeten Halb-Wissenschaftlern“¹⁸¹ in die Welt getragen werde. Die von Weizenbaum geforderte Transparenz im Sinne einer Selbstprüfung von Unternehmen und ihren Programmierern im Umgang mit ihrer eigenen manipulativen Kraft hat sich unter den libertären Bedingungen der Digitalwirtschaft bisher nicht eingestellt.

Begründungslücken zwischen Programmen und Menschen

Neben das strategische Vorenthalten von Wissen tritt ein Wissensdefizit, welche das Trainieren künstlicher neuronaler Netze aus sich hervorbringt. Innerhalb der Wissenschaftsgeschichte spricht man von *Intellectual Debt* bzw. intellektuelle Verschuldung, wenn Gesellschaften Objekte und Verfahren nutzen, obwohl ihr eigentlicher Wirkungszusammenhang nicht klar erforscht ist. In der Vergangenheit beschränkte sich die intellektuelle Verschuldung auf einige wenige Bereiche wie der Pharmazie. So wurde der Wirkstoff Aspirin bereits 1897 entdeckt und kam seitdem großflächig zum Einsatz, obwohl man erst 1995 seine tatsächliche Wirkweise nachwies. Die neuen Ansätze des maschinellen Lernens erhöhen nun den kollektiven intellektuellen Kreditrahmen des Unwissens in beträchtlichem Maße.¹⁸²

¹⁸⁰ Vgl. Alix, When hype is harmful: Why what we think is possible matters, <https://medium.com/swlh/when-hype-is-harmful-why-what-we-think-is-possible-matters-e7988db6f643> (27.9.2019).

¹⁸¹ Spiekermann, Digitale Ethik, 206.

¹⁸² Vgl. Jonathan Zittrain, „The Hidden Costs of Automated Thinking“, *The New York Times* (23.07.2019).

Im Kontext der KI spricht man von dem Phänomen der *Black Boxes*: Selbst der simpelste Algorithmus kann allein durch die schiere Menge der verarbeiteten Daten vergleichsweise intransparent zu seinen Ergebnissen kommen. Die Lernvorgänge in neuronalen Netzwerken sind das Ergebnis von Millionen an Verknüpfungen, von denen jede das Ergebnis minimal beeinflusst. Die Entscheidungsfindung innerhalb der *Hidden Layers* ist daher so kompliziert, dass die Maschine sie dem Menschen nicht darlegen kann. *Machine-Learning*-Systeme weisen in der Regel keine kausalen Zusammenhänge auf, sondern sind statistische Korrelationsmaschinen. Sie können nicht erklären, wie sie zu ihren Ausgaben kommen; sie rechtfertigen nicht, sie teilen nur Ergebnisse mit. Wenn der Mensch merkt, dass Maschinen Fehler machen, kann er diese Fehler kaum beheben, weil die Maschine nicht zeigt, wo der Fehler herührt. Da die meisten maschinellen Lernmodelle keine Begründungen für ihre fortlaufenden Urteile liefern können, gibt es keine Möglichkeit zu erkennen, wann der Fehler aufgetreten ist, wenn man nicht bereits ein unabhängiges Urteil über die gelieferten Antworten hat.¹⁸³

Die intellektuelle Verschuldung, die durch maschinelles Lernen entsteht, birgt somit Risiken, die über die Risiken hinausgehen, die ehemals durch Versuch und Irrtum entstanden. Kleine Änderungen der Datenwerte können zu dramatischen Änderungen im Ergebnis führen, ein Umstand, der für Angriffe auf das System ausgenutzt werden kann. Ohne Kontrollmaßnahmen kann das blinde Vertrauen auf die Ergebnisse in eine digitale Hörigkeit führen.

„Mounting intellectual debt may shift control, too. A world of knowledge without understanding becomes a world without discernible cause and effect, in which we grow dependent on our digital concierges to tell us what to do and when.“¹⁸⁴

Die Fortschritte in der Pharmazie haben gezeigt, dass intellektuelle Schulden in manchen Fällen unverzichtbar sind. Millionen von Leben wurden auf der Basis von unverstandenen Interventionen wie Aspi-

183 Vgl. ebd., 3.

184 Ebd., 5.

rin gerettet. Da sich Medikamente mit unbekanntem Wirkmechanismus vermehren, muss die Anzahl der Tests, die erforderlich sind, um ungewollte Wechselwirkungen aufzudecken, exponentiell ansteigen.¹⁸⁵ Das verantwortete Arbeiten mit der Anhäufung intellektueller Schulden im Kontext von KI bedeutet jedoch ein überwältigendes Ausmaß an zu prüfender Mehrarbeit. Wie kann ein erfüllbares Äquivalent für KI aussehen, um die Möglichkeiten der Technik zu nutzen? Die Ethikrichtlinien der EU fordern explizit die Nachvollziehbarkeit von KI-Systemen: Für die Vertrauenswürdigkeit eines Systems müsse es möglich sein zu verstehen, warum es sich auf eine bestimmte Weise verhält und ein bestimmtes Ergebnis liefert. Solche Plausibilitätsfunktionen gibt es bereits in Ansätzen: Ein ganzes Forschungsgebiet, die Erklärbare KI (*Explainable AI* [XAI]), versucht, eine nachhaltige Heuristik auf die maschinelle Kriterienfindung anzuwenden, bleibt aber bis heute weitestgehend eine offene Herausforderung für Systeme, die auf neuronalen Netzen basieren.¹⁸⁶ In dem Phänomen der *Black Boxes* offenbart sich somit nicht zuletzt eine Paradoxie der künstlich neuronalen Netzwerke: Als die großen Hoffnungsträger der KI-Entwicklung sollen sie unbekannte Strukturen der Welt im Heidegger'schen Sinne *entbergen*, gleichzeitig bleiben ihre Antworten orakelhaft und der menschlichen Vernunft ungenügend. Mit der Wahl, sie letzten Endes in ihrer Begründungslosigkeit akzeptieren zu müssen, erhalten die Ergebnisse künstlich neuronaler Netze eine transzendente Signatur. Eine vorschnelle Implementierung intelligenter Systeme forciert in zusätzlicher Weise die bestehende Tendenz der Informationstechnologien, einen Zustand des permanenten Nicht-Verstehens – in Nutzern als auch Betreibern – hinzunehmen. Sie fördern somit die Wiederverzauberung der Welt, hinter der die ökonomischen und machtpolitischen Interessen der digitalen Architektur unsichtbar werden und sich so die Unternehmen leichter der Aufforderung zur Rechenschaft über die Gestaltung ihrer Services entziehen können.¹⁸⁷

¹⁸⁵ Vgl. ebd., 2.

¹⁸⁶ Vgl. European Council on Ethics in Science and Technology, *Ethics Guidelines for Trustworthy AI*, 21–22.

¹⁸⁷ Vgl. Bernd Flessner, „Die Rückkehr der Magier. Die KI als lapis philosophorum des 21. Jahrhunderts“, in: Dierk Spreen/Bernd Flessner (Hgg.), *Kritik des Transhumanismus. Über eine Ideologie der Optimierungsgesellschaft*, Bielefeld 2018, 63–106. Hier 92.

3.5 Smart statt Intelligent: Wie berechtigt ist die Erwartung an eine starke KI?

Im Modus des akzeptierten Nichtverstehens scheint jede technologische Vision der Zukunft gleich plausibel. An dieser Stelle soll deshalb der wissenschaftliche Diskussionsstand über das Potential des maschinellen Lernens dargelegt werden. Hier ist besondere Vorsicht geboten: Technische Entwicklung ist nicht abgeschlossen oder eindeutig, was ihre Potentiale angeht. Überhaupt sind Unmöglichkeitsmutmaßungen kein Modus, in dem technisches Denken operiert. Viel eher stellt die technologische Perspektive per se jede Idee erst einmal unter das Verdikt des noch nicht Erreichten in Aussicht und verlegt so die Frage der prinzipiellen (Un-)Möglichkeit in die Zukunft (zum Zusammenhang von Technik und Transzendenz siehe 2.1). Dennoch sind nicht alle Theorien gleich berechtigt und damit einer Kritik entzogen. Das wissenschaftliche Für und Wider des technisch Möglichen im Bereich der Künstlichen Intelligenz hat seit dem Bestehen der Disziplin eine nunmehr über sechzigjährige Tradition ausgebildet. Auf Grundlage wichtiger sowie bis heute beachteter Theorien und Gedankenexperimente will dieses Kapitel einen informierten Standpunkt darüber darlegen, welche zukünftigen Formen von maschineller Intelligenz von einer fachwissenschaftlichen Mehrheit für wahrscheinlich gelten.

Die Widerständigkeit des Intelligenzbegriffs in Informatik und Robotik

Dass sich für die heutigen Formen präziser Stochastik in den heuristischen Verfahren der Mustererkennung die Rede von Künstlicher Intelligenz etabliert hat, ist im Verweis auf die Dartmouth-Konferenz historisch bedingt. Die allgemeine Idee, dass intelligente Programme komplexe Probleme lösen, hat sich seit den letzten zwanzig Jahren zwar mit konkreten Verfahren der Mustererkennung verbunden; dennoch besteht weiterhin eine Widerständigkeit des Intelligenzbegriffes im technologischen Kontext. Ein rein formalistischer Begriff ist die Intelligenz im Kontext der Informationstechnologien bis heute nicht geworden, seine Verwendung verweist weiterhin auf den Menschen, dem *homo sapiens sapiens*. Die Referenz zur menschlichen Intelligenz

wird besonders in den Debatten um Künstliche Superintelligenz offensichtlich, als das Präfix *super*, *hyper* oder *ultra* (lat.: *jenseits von*, *über etw. hinaus*) überhaupt nur verständlich ist, wenn die Leistungen der Maschine an der Durchschnittsleistung von Menschen, Probleme zu lösen, gemessen wird, wie ja auch der Informatiker John Irving Good seine Idee der Ultrainelligenz erst in Relation zum Menschen inhaltlich bestimmte:

„Let an ultrainelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever.“¹⁸⁸

Good macht den Menschen zum entscheidenden Maßstab der Maschinenintelligenz nicht nur, was die Leistung angeht, sondern sieht in ihm auch die notwendige Vorlage. Schließlich ist Good sich sicher, dass der Weg zur ersten intelligenten Maschine allein über das vertiefte Verständnis der Funktionsweise des menschlichen Gehirns führt:

„Man will construct the deus ex machina in his own image.“¹⁸⁹

Diese Verflechtung hat die Forschung und gesellschaftliche Erwartungen an die Disziplin massiv beeinflusst. Der Begriff der Künstlichen Intelligenz beschreibt als biomorphe Metapher die Technik im Verweis auf menschliche Intelligenz, ohne aber genau darzulegen, in welchem Grad die tatsächliche Ähnlichkeit zwischen dem Beschriebenen (*Modellandum*) und dem zu Hilfe gezogenen Bild (*Modelans*) besteht. Gleichzeitig aktiviert die Rede von der Intelligenz im Raum der Technik ein semantisches Feld, das Bewusstsein, Wille und Autonomie konnotiert und darin den Vergleichspunkt weiter verdunkelt. Epistemologisch ist dies gerade in der KI zu einem Problem geworden: Je mehr Worte sich in fremden semantischen Kontexten beheimaten, desto unkritischer werden sie auf weitere Eigenschaften dieser Felder ausgeweitet.¹⁹⁰ Die

188 Good, *Speculations Concerning the First Ultrainelligent Machine*, 32–33.

189 Ebd., 31.

190 Vgl. Mathias Gutmann/Julia Knifka, „Biomorphic and Technomorphic Metaphors – Some Arguments Why Robots do not Evolve, Why Computing is not Organic, and Why Adaptive Technologies Are not Intelligent“, in: Michael Decker/Mathias Gutmann/Julia

anthropomorphe Rede, dass Maschinen denken, entscheiden oder auswählen könnten, wird nicht mehr zweifelsfrei als Metapher behandelt. Wie wahrscheinlich als auch plausibel die Möglichkeit einer Hyperintelligenz oder starker KI aus technologischer Sicht ist, hängt deshalb in erheblichem Umfang davon ab, wie man mit dem Begriff der Intelligenz im Kontext der Technik umgeht und bereit ist, die unterstellte Ähnlichkeit zu menschlichem Denken kritisch zu hinterfragen.

Spielsituationen als Mess- und Vergleichsinstrument von Intelligenz

Der Forschungsantrag von Dartmouth formulierte das Forschungsziel dieser neuen Wissenschaft darin, Denken und Lernen in mechanische Prozesse zu übersetzen. Dabei ließen sie offen, wie ähnlich sich menschliche und maschinelle Intelligenz in dieser Übersetzung zu kommen brauchen. Der Vergleich einer an sich hochspekulativen Messgröße wie der Intelligenz wurde zwischen Menschen und Maschinen seither mehrfach über das kompetitive Spiel versucht. Zum Messen der Intelligenzleistung von Maschinen wurden immer wieder Spiele ausgewählt, in denen sich Computersysteme an den weltbesten Spielern zu messen hatten. In dieser Sparte haben die Maschinen stetig aufgeholt: 1997 besiegte das System *Deep Blue* der Firma IBM den Weltmeister im Schach Garri Kimowitsch Kasparow. Im Februar 2011 gelang IBM der nächste Triumph, als ihre künstliche Intelligenz *Watson* in einer mehrteiligen Fernsehsendung gegen die Champions Ken Jennings und Brad Rutter der Rate-Spielshow *Jeopardy* gewann. 2016 schließlich schlug *AlphaGo* von der Google-Firma *DeepMind* den Topspieler Lee Sedol im Spiel Go, das zu den kombinationsreichsten Spielen überhaupt zählt. Die Idee, die Intelligenzleistungen einer Maschine durch Spiele zu überprüfen, geht nicht zuletzt auf den theoretischen KI-Pionier Alan Turing zurück. In seinem 1950 erschienen Aufsatz *Computing machinery and intelligence* stellte Turing sein in die Geschichte als Turing-Test eingegangenes Gedankenexperiment vor, welches er selbst als *Imitation*

Game bezeichnete. Die ursprüngliche Frage, ob eine Maschine denken kann, setzt er dabei mit ihrer Fähigkeit gleich, linguistisch hinreichend intelligent auf den Menschen zu wirken:

„May not machines carry out something which ought to be described as thinking but which is very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection.“¹⁹¹

Es ist ein Spiel zwischen den Parteien A (Computer), B (Person) und C (Befrager): Alle drei sind räumlich getrennt und kommunizieren via elektronischer Chatfunktion. Das Ziel des Spiels ist es für A und B, C davon zu überzeugen, mit einem Menschen zu kommunizieren. Die Maschine muss also so gut wie möglich Antworten generieren, die möglichst menschlich wirken. Falls die Maschine den Befrager von seinem Menschsein eher überzeugen kann als der tatsächliche Mensch, gilt der Turing-Test als positiv.¹⁹² Dieser Test, der bis heute mit einem Preisgeld ausgeschrieben wird und auf seine Einlösung wartet, wird aufgrund einiger methodischer Schwächen kritisiert: Der Spielbau erfasst sehr selektiv nur die Fertigkeit des Computers, die Illusion einer menschlichen Kommunikation zu geben. Außer Acht gelassen werde jedoch völlig die Fähigkeit der Versuchsperson C, den menschlichen Kommunikator zu identifizieren. Gleichzeitig setze der Test die Fähigkeit des zur Wahl stehenden menschlichen Kommunikatoren einfach voraus, menschliche Kommunikation auch adäquat abzubilden.¹⁹³ Turings Gedankenexperiment hat nicht nur maßgeblich das Spiel als Vergleichsmittel von Intelligenz popularisiert, sondern hat mit seinem behavioristischen Testverfahren für die Zuschreibung von Intelligenz ein scheinbar unschlagbares Argument für den Mythos Maschine geliefert: Wenn klar und deutlich beschrieben werden kann, worin der Unterschied zwischen Mensch und Maschine in ihrem Verhalten

191 Vgl. Turing, *Computing Machinery and Intelligence*.

192 Vgl. ebd., 433.

193 Vgl. Bächle, *Mythos Algorithmus. Die Fabrikation des computerisierbaren Menschen*, 267.

besteht, lasse sich stets eine Maschine bauen, die auch diese Differenz noch simuliert. Denn sobald etwas klar und deutlich zu beschreiben ist, könnte es regelgeleitet und damit algorithmisch nachvollzogen werden.¹⁹⁴

Hubert Dreyfus Systematisierung von Intelligenzstufen

Neben diesem spieltheoretischen Ansatz, der den Fokus allein auf eine im Spiel überzeugende *Performance* legt, haben sich über die Jahrzehnte nach Turing auch andere Systematisierungen für Intelligenz entwickelt, an denen die Maschinenleistung verglichen wurde. Wie sehr sich die Grenzen dafür verschoben haben, was erstens als intelligent gilt und man zweitens einer technischen Simulation zutraut, verdeutlichen die vier Intelligenzstufen des amerikanischen Philosophen Hubert Dreyfus (1929–2017), die er Anfang der 1970er Jahre inmitten der ersten Hochzeit der KI-Forschung aufstellte:

Auf der basalen ersten Stufe, der assoziativen Intelligenz, stehe das Erlernen durch Wiederholung und Memorienspiele, wobei Semantik wie Situation irrelevant sind. Diese Stufe sieht Dreyfus bereits zu seiner Zeit in Programmen erfüllt, die mit Entscheidungsbäumen und Listensuche arbeiten. Unter der zweiten Stufe, der einfach-formalen Intelligenz, begreift er das Lernen nach Regeln, wobei die zugeordneten Deutungskategorien eindeutig sind. Diese Leistung sieht er ebenfalls im klassischen Computeralgorithmus seiner Zeit erfüllt. Die komplex-formale Intelligenz der dritten Stufe meint das Lernen durch Regeln und Übung, mit dem im festgesetzten Rahmen situationsbezogen agiert wird, wie dies bei Spielen wie Schach oder Go der Fall ist. Dass die Technik eine solche Stufe erreichen könnte, hält Dreyfus für theoretisch möglich und verweist auf Ansätze in den damaligen Suchbeschneidungssystemen (*Search-pruning Heuristics*). Die vierte und höchste Stufe, die nicht-formale Intelligenz, bedeutet für Dreyfus das Lernen durch anschauliche Beispiele, und zwar abhängig von Bedeutung und Situation, die nicht explizit sind. Damit meint er das Verstehen von Rätseln oder das verständige Übersetzen einer natürlichen Sprache. Diese Tätigkeiten hält Dreyfus auch in Zukunft für nicht

¹⁹⁴ Vgl. Dirk Evers, „Der Mensch als Turing-Maschine?“, *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie* 47 (2005), 101–118. Hier 104–105.

übertragbar.¹⁹⁵ Zumindest was die Anwendungsfelder angeht, haben die neuen Methoden maschinellen Lernens Dreyfus Vorbehalt widerlegt. Übersetzungsprogramme auf Grundlage von Mustererkennung schaffen es immer besser, kontextsensitive Übersetzungen anzufertigen. Dennoch bleibt debattierbar, ob damit Dreyfuss Voraussetzung für die vierte Stufe von Intelligenz erfüllt ist, ging es ihm doch um ein qualitatives Argument von tatsächlichem Verständnis in der Maschine.

Die Spekulation über mentale Zustände in Maschinen

Die seit Turings Rede von denkenden Maschinen schwelende Frage, ob Maschinen einst tatsächlich ein verständiges Bewusstsein werden ausbilden können, nahm sich in besonderer Weise der amerikanische Philosoph John Searle (*1932) an. Als er in den 1980er Jahren seine Unterscheidung zwischen schwacher und starker KI vorstellte, bemerkte er, dass ein Großteil der Vertreter einer starken KI-Theorie davon überzeugt war, dass sich mit zunehmender Komplexität in den Computern tatsächlich mentale Zustände einstellen müssten, vorausgesetzt die richtigen Programme liefen. Ihre Vertreter hielten mentale Zustände für ein Beiprodukt komplexer Systeme, das sich systemübergreifend ab einem gewissen Schwellenwert einstellte.¹⁹⁶ Bereits für Good gehen Intelligenz und Bewusstsein zusammen: Intelligente Maschinen würden bei ausreichender Anwachsen ihrer Leistung automatisch Ziele entwickeln; auch Emotionen und Schmerz setzt er in diesen Systemen voraus.¹⁹⁷ In der Auseinandersetzung zwischen den *Biological Theorists* wie Searle, die ein Bewusstsein als exklusives Phänomen biologischen Lebens deuten, und den *Functional Theorists* wie Good geht es um die Frage nach der Exklusivität mentaler Zustände und den Grenzen der tatsächlichen Verständnisfähigkeit von Computern. John Searle schlug zur Untermauerung seiner Position ein bis heute beachtetes Gedankenexperiment namens *The Chinese Room* vor. Seiner Ansicht nach führt es vor Augen, dass die Weise, wie Computer funktionieren und dabei

¹⁹⁵ Vgl. Hubert L. Dreyfus, *What Computers can't do: a Critique of Artificial Reason*, New York 1972, 204.

¹⁹⁶ Vgl. Searle, *Minds, Brains, and Programs*, 417.

¹⁹⁷ Vgl. Good, *Speculations Concerning the First Ultrainelligent Machine*, 33.

scheinbar intelligent Aufgaben lösen, noch keinerlei Indiz für eine vollbrachte Verständnisleistung liefert.

In seinem Gedankenspiel ist eine Person in einen Raum mit verschiedenen Stapeln an Schriften auf Chinesisch eingesperrt, wobei die Schriftzeichen für die Person nichts weiter als bedeutungslose Markierungen (*meaningless Squiggles*) sind. Neben diesen Stapeln an unverständlichen Schriftzeichen erhält die Person zudem ein Handbuch in ihrer Muttersprache. Dieses gibt der Person vor, wie mit den fremden Symbolen umzugehen ist. Es erlaubt ihr, einen Satz an chinesischen Symbolen mit anderen chinesischen Symbolen in Beziehung zu setzen. Durch dieses Handbuch kann die Person auf eine Reihe von chinesischen Symbolen mit anderen Symbolen reagieren, die für einen chinesischen Muttersprachler außerhalb des Raumes den Eindruck erweckten, es mit jemandem zu tun zu haben, der Chinesisch versteht. Niemand käme von den Antworten ausgehend auf die Idee, dass die Botschaften, die aus dem Türschlitz des Zimmers herauskommen, von einem Nicht-Chinesen stammen. Aber im Gegensatz zum Handbuch in der Muttersprache benimmt sich die Person, was das Chinesische angeht, nach Ansicht Searles wie ein Computer. Sie führt nach einem eindeutigen, situationsunabhängigen und wiederholbaren Verfahren – sprich algorithmisch – Rechenoperationen auf formal spezifizierten Elementen durch.¹⁹⁸ Die Person, die diese Antworten schreibt, bezieht sich nicht mit ihren Ausdrücken auf bestimmte Gegenstände außerhalb des Raums und verfolgt auch keine Absichten. Sie simuliert lediglich das Verständnis des Chinesischen. Searle fasst seinen Ertrag aus dem Gedankenexperiment folgendermaßen zusammen:

„Rather, whatever purely formal principles you put into the computer, they will not be sufficient for understanding, since a human will be able to follow the formal principles without understanding anything. No reason whatever has been offered to suppose that such principles are

198 Vgl. Searle, *Minds, Brains, and Programs*, 417–418.

necessary or even contributory, since no reason has been given to suppose that when I understand English I am operating with any formal program at all.¹⁹⁹

Ein Programm, das solch eine Leistung nach außen hin demonstriert, sei demnach in keiner Weise eine hinreichende Bedingung für die Annahme von Bewusstsein oder Intentionalität. Searles Person im chinesischen Zimmer komme allein nach den Regeln der Informationsverarbeitung zu zufriedenstellenden Ergebnissen ohne ein Wissen darüber, dass die Schriftzeichen tatsächlich Symbole sind, die in einer Beziehung zu Dingen in einer Außenwelt stehen und auf diese verweisen. Für die vernetzten Daten gibt es kein Außen, dem Computer bleiben die Werte lediglich Rückkopplungen im Medium ihrer selbst, sprich Informationen, die immer nur auf andere Informationen verweist.²⁰⁰ Mit diesem Argument wendet sich Searle nicht zuletzt gegen Turings *Imitation Game* und seiner Gleichsetzung von Performanz und Simulation, indem Searle die Sprachfähigkeit einer Maschine nicht zum hinreichenden Kriterium von Intelligenz erklärt. Doch Searles *Chinese Room* blieb nicht unwidersprochen. So erklärt Raymond Kurzweil das Gedankenexperiment Searles zum philosophischen Taschenspielertrick: Seinem Verständnis nach manipuliere eine Maschine, die wirklich das tun könnte, was Searle mit seiner Person im Zimmer beschreibt, mehr als nur Sprachsymbole. Deshalb müsse sie von menschlicher oder höherer Komplexität sein und damit tatsächlich Chinesisch in einer tiefen Weise verstehen. Als *Functional Theorist* wirft er Searle hingegen vor, das Bewusstsein im Computer vorschnell auszuschließen, obwohl das Gedankenexperiment die Nichtexistenz des Mentalen ebenfalls nicht beweisen könnte. Kurzweil dreht somit die Argumentation im Sinne Turings *Imitation Game* um und fordert – solange nicht die Unmöglichkeit des Mentalen in der Maschine bewiesen ist –, das Fremdpsychische als Regelfall anzunehmen.²⁰¹

199 Ebd., 418.

200 Vgl. Nassehi, *Muster*, 101.

201 Vgl. Kurzweil, *The Singularity Is Near*, 308.

Aus einer pragmatischen Sicht wird man der lebensweltlichen Erfahrung und damit Searle folgen, nach der Computern und Robotern erst einmal keine mentalen Eigenschaften innewohnen und die Beweislast daher auf der Seite der *Functional Theorists* liegt. Nida-Rümelin hält es für elementar, dass, so wie es keine Zweifel am Fremdpsychischen unter Menschen geben könne, es auch keinen Zweifel am nicht psychischen Charakter des Digitalen gebe und besteht darauf, die Simulation der Maschine nicht mit ihrer Realisierung zu verwechseln.²⁰² Nassehi weist ebenfalls darauf hin, dass, nur weil die komplexe Datenverarbeitung in Computern oftmals nicht mehr nachvollzogen werden kann, die Nichtrückführbarkeit ihrer Herleitungswege nicht als Ausweis eines eigenständigen Handlungsfähigkeit missverstanden werden sollte. Am Ende arbeite jede Maschine und jedes Programm nur das ab, wofür sie konzipiert wurde, selbst wenn sie zu Ergebnissen kommt, die nicht unmittelbar mitkonzipiert wurden.²⁰³

Weltmodelle und das Defizit der digitalen Repräsentation

Über die Möglichkeit eines verständigen Bewusstseins hinaus stellt sich die Frage, inwieweit Maschinen sich allein über Daten ein umfassendes Modell von der Wirklichkeit machen können, welches sie befähigt, mit der Außenwelt in Interaktion zu treten. Eine wesentliche Nachfrage an die intelligente Maschine im Sinne der starken KI-Theorie ist deshalb, inwieweit es möglich ist, Maschinen eine möglichst vollständige Repräsentation der Wirklichkeit einzuprogrammieren. Es erweitert sich hier die erste Frage, inwieweit nicht nur das Denken einer algorithmischen Struktur folgt, sondern ob die Welt und Wirklichkeit als Ganzes algorithmisch strukturiert ist. Ein in diesem Zusammenhang bis heute beachtetes Theorem der theoretischen Informatik stammt von dem tschechischen Mathematiker Kurt Gödel (1906–1978) aus den 1930er Jahren. Sein Unvollständigkeitssatz beweist, dass es logische Sätze in jedem formalen System gibt, die rein formal weder bewiesen noch widerlegt werden können und es demnach definitive Gren-

202 Vgl. Nida-Rümelin, Weidenfeld, Digitaler Humanismus, 41.

203 Vgl. Nassehi, Muster, 202.

zen für das gibt, was Logik und Mathematik zu leisten imstande sind.²⁰⁴ Die Mehrheit der Forscher sehen in Gödels Theorem den Gegenbeweis für alle Visionen einer starken KI. Sie sagen, dass künstliche Systeme, egal wie sehr sie sich fortentwickelten, ihrer algorithmischen Struktur nicht entkommen könnten und ihnen deshalb Grenzen gesetzt seien, in denen sie niemals die Komplexität der Welt verarbeiten und damit ein kohärentes Weltmodell erzeugen könnten. Wenn nicht einmal formale Systeme widerspruchsfrei gedacht werden können, wird gefolgert, dass das Denken in seiner Ganzheit nicht algorithmisch dargestellt oder simuliert werden kann und auch die Welt letztlich nicht algorithmisch strukturiert ist. Jeder Versuch, der Welt über die formale Logik habhaft zu werden, müsse zwangsläufig an der Komplexität der Wirklichkeit scheitern.²⁰⁵

Kurzweil hält dem entgegen, dass Menschen ebenso wenig in der Lage seien, solche unlösbaren Probleme, wie Gödel sie identifizierte, universell zu lösen, und sie dennoch erfolgreich mit dieser Welt agierten. Menschen könnten zwar fundierte Vermutungen zu Lösungen anstellen und heuristische Methoden anwenden, die gelegentlich erfolgreich sind, beides seien jedoch Ansätze, die ebenfalls in algorithmisch basierte Prozesse übersetzt werden könnten. Der Computer könne sich möglicherweise sogar dann doch als fähiger darin erweisen, unlösbare Probleme zu lösen als Menschen.²⁰⁶ Nassehi relativiert ebenfalls die Tragweite des Gödeltheorems in Bezug auf starke KI: Selbst wenn die Kohärenz der Wirklichkeit in formaler Sprache auszudrücken nach Gödel unmöglich ist, so sei eine solche Kohärenz selbst empirisch nicht ausweisbar, sondern bereits eine metaphysische Setzung. Wenn kein formales System jemals vollständige Transparenz über sich erlangen kann, dann könnte dies möglicherweise für alle Strukturen auch außerhalb der formalen Logik gelten. Der Eindruck eines geschlossenen Ganzen der Wirklichkeit in der menschlichen Wahrnehmung sei dann nur eine selektive Vorannahme, aber kein Ausdruck vollständiger (Selbst-)Transparenz des Menschen als unüberwindbarer Graben

204 Vgl. Kurt Gödel, „Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I“, *Monatshefte für Mathematik und Physik* 38–38 (1931), 173–198.

205 Vgl. Nassehi, *Muster*, 251–253.

206 Vgl. Kurzweil, *The Singularity Is Near*, 303–305.

gegenüber anderen Systemen.²⁰⁷ Dem entgegen macht der erfahrungsorientierte Ansatz der Wirtschaftsinformatikerin Sarah Spiekermann die Konsequenzen des Gödeltheorems stark: In ihrer digitalen Ethik knüpft sie an die Werttheorien Max Schelers (1874–1928) an und kommt zu dem Schluss, dass Maschinen die nicht explizierbare Welt der Werte unmöglich in Modellen erfassen könnten und somit den Maschinen die saturierte Vollständigkeit menschlicher Wirklichkeit unzugänglich bliebe. Selbst die körperlich wahrgenommenen, sensorischen Informationen, die sich bei Lebewesen als Gefühle materialisierten, hält sie in ihrer Vielschichtigkeit von einer Maschine weder für beobachtbar noch für reproduzierbar. Damit aber bleibe der Input einer Maschine, die sinnvoll an der menschlichen Welt partizipieren sollte, kolossal begrenzt.²⁰⁸ Diesen nicht formalisierbaren Überschuss an Wirklichkeit hält Spiekermann für eine jedem Menschen offensichtliche Erfahrung und damit das stärkste Argument gegen eine reduktionistische Sicht auf die Welt:

„Der Glaube, dass Modelle unsere Welt trotz ihrer Abstraktionen und brutalen Vereinfachungen hinreichend vollständig wiedergeben könnten, ist es, der uns dazu verleitet, unkritisch darauf zu vertrauen, dass eine Künstliche Intelligenz uns die Gemütlichkeit eines Raumes ausrechnen kann, wenn nicht heute, dann in irgendeiner paradiesischen Zukunft. Und wir sind der Meinung, dass die modellbasierte Maschine dabei sogar richtiger liegt als der Mensch, dessen Einschätzung von Gemütlichkeit wir nicht trauen [...]“²⁰⁹

Das Abwägen zwischen Plausibilität und Probabilität hinsichtlich starker KI

Entlang der bisher skizzierten Argumentationsstränge, die sich entlang des Intelligenzbegriffs, entlang von Spieltheorien, mathematischen Beweisen und Gedankenexperimenten bewegten, bleibt in Hinblick auf ihre abschließende Bewertung diffus, ob die Befürchtung einer voll-

²⁰⁷ Nassehi, *Muster*, 251–253.

²⁰⁸ Vgl. Spiekermann, *Digitale Ethik*, 87.

²⁰⁹ Ebd., 145.

ständigen Substitution des Menschen Anlass zu berechtigter Sorge im Raum der technologischen Folgenabschätzung um maschinelles Lernen bereitet. Der französische Informatiker und Philosoph Jean-Gabriel Ganascia schlägt vor, man müsse die Probabilität starker KI von ihrer Plausibilität unterscheiden: Die Wahrscheinlichkeit, auch Probabilität, kommt von *probare*, was im Lateinischen *beweisen* bedeutet. Etymologisch habe etwas Wahrscheinliches, ohne absolut sicher zu sein, immer noch einen gewissen Zusammenhang mit dem Wahren und sollte daher eine große Chance haben, zu geschehen. Darin unterscheide sich die Probabilität von der Plausibilität, die schlicht und einfach ausdrückt, dass nichts seine Verwirklichung formal verbietet, aber auch nichts seine Verwirklichung sichert. Wie das abgeleitete Wort *Applaus* verdeutlicht, würden plausible Szenarien vor allem von ihrer Popularität bestimmt, indem viele Menschen glauben, dass es passieren wird. Gegenüber den Thesen von starker KI und technologischer Singularität äußert Ganascia die Vermutung, dass ihre Befürworter große Anstrengungen unternehmen, sie – im etymologischen Sinne – plausibel zu machen (s. 2.2.2). Ihre Probabilität hingegen bleibt, – wie in der Frage des Fremdpsychischen der Maschine als auch der algorithmischen Struktur der Welt ausgeführt – eine offene Behauptung, die nicht aus der gegenwärtigen Technik begründet werden kann, sondern vielmehr ihre gegenteilige Annahme um einiges wahrscheinlicher ist.²¹⁰ Ganascia schließt seine Untersuchung zum Mythos von der technologischen Singularität mit folgendem Fazit:

„Bref, si rien ne pennet d'affirmer l'impossibilité absolue de la Singularité, elle est hautement improbable, si improbable qu'on ne saurait l'envisager sérieusement.“²¹¹

Die Popularität der Thesen um starke Künstliche Intelligenz hat dazu geführt, dass die EU-Kommission sich in ihren Ethikrichtlinien ebenfalls zu der Möglichkeit starker KI geäußert hat:

210 Vgl. Ganascia, *Le Mythe de la Singularité*, 89.

211 Ebd., 90.

„The high-impact nature of these concerns, combined with the current uncertainty in corresponding developments, calls for regular assessments of these topics. While some consider that Artificial General Intelligence, Artificial Consciousness, Artificial Moral Agents, Super-intelligence or Transformative AI can be examples of such long-term concerns (currently non-existent), many others believe these to be unrealistic. A black swan event is a very rare, yet high impact, event – so rare, that it might not have been observed. Hence, probability of occurrence typically can only be estimated with high uncertainty.“²¹²

Unter Absehung von Argumenten wird festgestellt, dass den wenigen Vertretern (*some*) der starken KI die vielen (*many*) Skeptiker gegenüberstehen. Eine imminente Bedrohung wird der starken KI an dieser Stelle abgesprochen, sie ist in dem Sinne auch kein tatsächliches Thema konkreter KI-Strategien der Europäischen Union. Gleichzeitig verfolgt die EU eine Heuristik der Furcht (s. 2.2.1) und behält sich aufgrund der Ungewissheit der technologischen Durchbrüche der Zukunft vor, dass es einem regelmäßigen Abgleich der Datenlage bedarf, um das Eintreten eines sog. *Black-Swan-Events*²¹³ zu verhindern.

Maschinelles Lernen ist wenn, dann eine andere Form der Intelligenz

Alan Turings war sich sicher, dass bis zum Ende des 20. Jahrhunderts die Rede von denkenden Maschinen allgemein akzeptiert sein werde, weil zu dem Zeitpunkt ein üblicher Befrager in seinem *Simulation Game* nicht mehr als eine 70 % Chance hätte, einen Computer von einem anderen Menschen zu unterscheiden.²¹⁴ Good hielt es Mitte der 1960er Jahre für sehr wahrscheinlich, dass bis zum Ende des 20. Jahrhunderts die Menschheit einer ausreichend intelligenten Maschine geschaffen habe, die den Prozess der Intelligenzexplosion zu starten

²¹² European Council on Ethics in Science and Technology, Ethics Guidelines for Trustworthy AI, 35.

²¹³ Darunter versteht man ein höchst unwahrscheinliches Ereignis, das wider Erwarten doch eintritt und bei dem sich im Nachhinein herausstellt, dass es doch vereinzelte Indizien gegeben hatte, die in einer allgemeinen Ignoranz unbeachtet blieben.

²¹⁴ Vgl. Turing, *Computing Machinery and Intelligence*, 442.

bereit wäre.²¹⁵ Vernor Vinge gab 1993 einen Zeitraum von 30 Jahren an, bis wann die Menschheit die erste *Superhuman Intelligence* hervorgebracht habe.²¹⁶ Hans Moravec prognostizierte Ende der neunziger Jahre für die Dekade der 2020er Maschinen mit Menschenintelligenz.²¹⁷ Und Raymond Kurzweil schätzt ähnlich wie Jochen Schmidhuber den Zeitpunkt, in dem nichtbiologische Intelligenz eine Milliarde leistungsfähiger als alle menschliche Intelligenz sein wird, auf das Jahr 2045.²¹⁸

Auch wenn die Vorhersagen sich bisher immer als Irrtum herausgestellt haben, ist gegenwärtig abzusehen, dass das Potential des maschinellen Lernens erst am Anfang steht und ihre Anwendungsgebiete sich erweitern werden. Anwendungen werden intelligenter werden in dem Sinne, dass sie in immer größerem Maße Daten miteinander in Beziehung setzen, Korrelationen aufdecken und damit Lösungen für komplexe, menschliche Fragen aufbereiten. Der Mehrwert dieser Systeme wird dabei von ihrer jeweiligen Verwendung innerhalb der menschlichen Praxis abhängen.²¹⁹ Nutzer und Programmierer sind es, die die elektrodynamischen Prozesse im Computer derart gestalten, dass sie dem Menschen einer Syntax entsprechen. Searles Gedankenexperiment führt vor Augen, dass Intelligenz eine deutende Zuschreibung, eine Kategorie von Menschen für Menschen ist. Viele komplexe Aufgaben und Tätigkeiten werden sich in Zukunft – wenn gewünscht – automatisieren lassen und damit frühere Einschätzungen, wo die Grenzen maschineller Datenverarbeitung liegen, weit hinter sich lassen.

Jedoch gibt es keinen Hinweis darauf, dass diese moderne Form der Stochastik in irgendeiner Weise zu künstlichem Bewusstsein oder Intentionalität führen wird. Die starke KI als ein fühlendes, bewusstes Wesen mit eigenen Wünschen und Zielen und somit ein digitales Replikat menschlicher Qualitäten – für solche Annahmen gibt es wissenschaftlich keine Grundlage. Hier sind nicht einmal Ansätze ersichtlich,

215 Vgl. Good, *Speculations Concerning the First Ultra-intelligent Machine*, 63.

216 Vgl. Vinge, *The Coming of Technological Singularity: How to Survive the Post-Human Era*, 11.

217 Vgl. Moravec, *Computer übernehmen die Macht*, 94.

218 Vgl. Kurzweil, *The Singularity Is Near*, 98.

219 Vgl. Gutmann, Knifka, *Biomorphic and Technomorphic Metaphors – Some Arguments Why Robots do not Evolve, Why Computing is not Organic, and Why Adaptive Technologies Are not Intelligent*, 75–76.

wie sich aus den bestehenden Technologien ein künstliches Bewusstsein entwickeln soll.²²⁰ Mustererkennung durch künstliche neuronale Netze ist etwas anderes als ein bewusstseinsgeleitetes Erfassen der Welt.

Die Erfahrung lehrt, dass Maschinen keine mentalen Eigenschaften besitzen. Ein noch so genaues Computermodell des Wetters beispielsweise bleibt wesentlich unterschieden von der tatsächlichen Wettererscheinung. Nur wenn *Modelans* und *Modelandum* miteinander verwechselt werden, spricht das beschreibende Bild nicht mehr als sekundär empfunden wird, sondern stattdessen zum Maßstab des eigentlichen verkehrt wird, erscheint das Technische intelligent oder gar lebendig.²²¹ Die Physikalisten der starken KI sind in der argumentativen Bringschuld, weshalb dies bei Modellen des Gehirns und ihrer Übertragung in ein Computermodell anders sein sollte und hier eine tatsächliche Emulation des Denkens in der Maschine stattfindet. Allein anzumerken, dass das Denken selbst algorithmischer Natur sei, bleibt andernfalls bloßes Dogma. Damit ist nicht endgültig gesagt, dass künstliches Leben überhaupt unmöglich ist. Keine der Theoreme oder Gedankenexperimente hat dies zweifelsfrei belegen können. Tatsächlich aber scheinen die bestehenden Ansätze des maschinellen Lernens für solche Hoffnungen völlig ungeeignet. Alle heutigen Ansätze des *Machine Learning* geben keine Indizien dafür, den Weg zur künstlichen Synthese von Bewusstsein und Leben zu ebnen. In diesem Sinne gilt weiterhin Joseph Weizenbaums Diktum, dass man sich vor einem vereinfachten, undifferenzierten Intelligenzbegriff in der Technik hüten sollte.²²² Aufgrund der immensen anthropologischen Last lehnen viele

220 Vgl. Christoph Kehl/Christopher Coenen, *Technologien und Visionen der Mensch-Maschine-Entgrenzung*. Sachstandsbericht zum TA-Projekt »Mensch-Maschine-Entgrenzungen: zwischen künstlicher Intelligenz und Human Enhancement« (TAB-Arbeitsbericht Nr. 167), Berlin 2016, 16–17.

221 Vgl. Gutmann, Knifka, *Biomorphic and Technomorphic Metaphors – Some Arguments Why Robots do not Evolve, Why Computing is not Organic, and Why Adaptive Technologies Are not Intelligent*, 60.

222 „Ich bin der Ansicht, dass ein in jeder Beziehung zu vereinfachter Begriff von Intelligenz sowohl das wissenschaftliche wie das außerwissenschaftliche Denken beherrscht hat, und dass dieser Begriff zum Teil dafür verantwortlich ist, dass es der perversen, grandiosen Phantasie der künstlichen Intelligenz ermöglicht wurde, sich derart zu entfalten. Computer und Menschen sind nicht verschiedene Arten derselben Gattung.“ Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 269.

Forscher deshalb den Begriff der Künstlichen Intelligenz komplett ab und verwenden allein niedrigschwellige Bildsprache wie den Begriff des *maschinellen Lernens* oder sprechen von *smarter* anstatt intelligenter Technologie. Die Denkschrift der EKD *Freiheit Digital* plädiert ebenfalls für ein rigoroses Entbildern der Sprache innerhalb der Informationstechnologien, um der anthropomorphisierenden Sprachverwirrung mit technologisch informierter Sprache entgegenzuwirken.²²³

Die Vermenschlichung der Maschine wie die Mechanisierung des Menschen scheint der aktuellen Forschung um Künstliche Intelligenz epistemologisch mehr und mehr zum Problem denn zur Lösung zu geraten. Viel eher sollte man die Eigenart der neuen Technologie abseits ihres ständigen Vergleiches mit menschlicher Problemlösung sehen, das heißt ihre Andersheit wahrnehmen und nutzen: Mit dem maschinellen Lernen wurde ein Weg gefunden, um komplexe menschliche Probleme zu einem zufriedenstellenden Maße zu lösen, ohne auf Bewusstsein, echtes Verständnis oder Welthaftigkeit angewiesen zu sein. Zu einem pragmatischen Verständnis von KI gehört, dass die Systeme die Vorgehensweise des Menschen beim Lösen der Aufgaben nicht imitieren müssen – schon gar nicht die biochemischen Vorgänge im Gehirn.²²⁴ Mustererkennung durch *Deep Learning* gibt Einblicke in oftmals ungesehene Zusammenhänge und liefert damit neue Ansätze, um an bestehende Probleme heranzugehen oder dieser überhaupt erst ansichtig zu werden.

Gleichzeitig treten in der Wahrnehmung ihrer Unterschiedlichkeit nicht nur die blinden Flecken menschlicher Wahrnehmung, sondern auch die Nachteile der künstlichen Systeme zutage: Hochspezialisierte Netze sind meist auf Millionen von aufwendig aufbereiteten Probedaten angewiesen, um Merkmale zu identifizieren, die Menschen unmittelbar auffallen. Zudem sind ihre Instrumente in hohem Maße störanfällig, sodass zufällige oder bewusst eingestreute Kleinststörungen

223 Vgl. Evangelische Kirche in Deutschland, *Freiheit digital. Die Zehn Gebote in Zeiten des digitalen Wandels*. Eine Denkschrift der Evangelischen Kirche in Deutschland, Leipzig 2021, 34.

224 Vgl. Luciano Floridi, „What the Near Future of Artificial Intelligence Could Be“, in: Christopher Burr/Silvia Milano (Hgg.), *The 2019 Yearbook of the Digital Ethics Lab.*, Cham 2020, 127–142. Hier 136.

gen (*Noise*) selbst bei gut trainierten Netzwerken zu desaströsen Trefferquoten führen können. Nicht vernachlässigen sollte man auch die besondere Kurzsichtigkeit (*Myopik*) der KI: Beispielsweise analysieren bei der Bilderkennung einige Schichten der Netzstruktur jeweils Muster in bestimmten Pixelanordnungen und arbeiten dann mit größeren Organisationseinheiten weiter. Aber wo diese Modelle einzelne Objekte treffsicher bestimmen können, fehlt in den meisten Fällen der Zusammenhang zwischen den Objekten. Eine Person wird dann auch als Person erkannt, wenn auf dem Bild der Mund und Augenpartie im Gesicht vertauscht sind.²²⁵

Aufgrund dieser besonderen Anforderungen wird es in Zukunft darum gehen, erstens dort hybride oder vollsynthetische Trainingsdaten herstellen zu können, wo wenig Material zur Verfügung steht oder es sich um hochsensible, personenbezogene Daten handelt. Zweitens wird die Nutzung von KI gerade in vulnerablen Bereichen der Lebensführung nicht ohne menschliche Datenkuratoren und Qualitätskontrolleure auskommen. Drittens wird es in Zukunft eine digitale als auch räumliche Infrastruktur brauchen, die auf die Kurzsichtigkeit der Künstlichen Intelligenz angepasst ist, sodass sich komplizierte Aufgaben und Probleme in solche formal-komplexen Probleme umwandeln lassen, für die das Werkzeug der KI greifen kann.²²⁶ Schließlich rückt im Bewusstsein dieser Andersheit Künstlicher Intelligenz auch die Frage nach *Moral AI* in ein verständlicheres Licht: Maschinen können bestenfalls im funktionalen Sinn moralisch handeln. Um ein moralisch Handelnder vergleichbar zum Menschen zu sein, fehlt Bewusstsein, Willensfreiheit und Selbstreflexion.²²⁷

Die Implementierung wie dauerhafte Nutzung dieser Technologien bedeutet nicht ein Verschwinden menschlicher Verantwortung in Übertragung auf die Maschine, sondern führt bei genauer Betrachtung im Gegenteil zu einer gesteigerten Verantwortung. Diese ist jedoch nur erkenntlich, wenn man die Grenzen der Technologie, die Grenze

225 Vgl. David Watson, „The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence“, in: Christopher Burr/Silvia Milano (Hgg.), *The 2019 Yearbook of the Digital Ethics Lab.*, Cham 2020, 45–65. Hier 50–52.

226 Vgl. Floridi, *What the Near Future of Artificial Intelligence Could Be*, 138.

227 Vgl. Misselhorn, *Maschinen mit Moral?*, 205.

ihrer Neutralität, die Grenzen der Einsichtigkeit ihrer Ergebnisfindung, schließlich die Grenzen ihrer tatsächlichen Autonomie anerkennt und ihre Angewiesenheit auf den Menschen klarstellt. Wird diese Verantwortlichkeit verschleiert, verschiebt sich die Machtfrage weiter zu den technologischen Monopolisten, welche die Unwissenheit für sich nutzen können, um auf ihr Verständnis digitaler Wertschöpfung bedingungslos zu verpflichten.

4 Die ideengeschichtliche Gleichsetzung von Mensch und Maschine

Intelligenzexplosion, künstliches Bewusstsein und technologische Singularität – ihre apokalyptische Brisanz erhalten die Potentialerwartungen um Künstliche Intelligenz allein in der Annahme einer drohenden Gefahr der Substitution des Biologischen durch das Technische. Die Parallelisierung von Mensch und Maschine mit der Latenz, den Menschen in diesem Prozess aufzulösen, ist jedoch keine Idee des 21. Jahrhunderts, die sich erst in der Beschäftigung mit Künstlicher Intelligenz eingestellt hat, sondern begleitet die westliche Geistesgeschichte seit der frühen Neuzeit.²²⁸ Inwieweit sich in den Konvergenzerzählungen um Künstliche Intelligenz bestehende Denktraditionen wie die einer *zweiten Schöpfung*²²⁹ fortführen und weiterschreiben, wird in diesem Kapitel anhand von drei Zugängen aufgezeigt.

Im Aufkommen der technischen Utopie und den frühen Formen der Science-Fiction verbindet sich das Ideal einer höchsten Beherrschung aller Naturkräfte durch den Menschen bereits mit der Idee, dass aus der Natureinsicht heraus Leben noch einmal erschaffen werden könnte. Der Automat als eine fortgeschrittene Spielart dieses Motivs tritt dabei im Laufe der Literaturgeschichte mehr und mehr als selbsttätiges Gegenüber des Menschen auf, den die Absicht zur Täuschung – sowohl hinsichtlich der tatsächlichen Fähigkeiten als auch seiner Loyalität gegenüber seinem Erschaffer – fortwährend begleitet. Daneben zeichnet ein wissenschaftsgeschichtlicher Zugriff nach, wie die Idee des Menschen als eine fortgeschrittene, biologische Maschine bereits

²²⁸ Die Untersuchung beschränkt sich auf den westlichen Kulturraum. Eine eigene, besonders ausgeprägte kulturelle Verbindung zum Roboter zeigt beispielsweise auch Japan. Aufgrund des dem Shintoismus zugehörigen Animismus wird angenommen, dass die Japaner Roboter besonders leicht als Gegenüber anerkennen. Auch der mit am meisten verkaufte Manga Japans, *Astro Boy* aus den 1950-Jahren, trug zur Popularität der Roboter bei. Hier geht es um einen Roboterjungen, der von einem Dr. Tenma als Ersatz für seinen verunglückten Sohn Tobio geschaffen wurde und mit seinen besonderen Fähigkeiten als Superheld auf Verbrecherjagd geht.

²²⁹ Vgl. Bernd Flessner (Hg.), *Nach dem Menschen*. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur, 1. Aufl., Freiburg im Breisgau 2000.

in der Aufklärung formuliert wurde und die Epistemologie der empirischen Wissenschaften von da an begleitete. Ihr Durchbruch gelang in der kybernetischen Bewegung des 20. Jahrhundert. Indem sie den Menschen zur biologischen Informationsmaschine reduzierte, war sie der letzte offen beanspruchte Versuch einer wissenschaftlichen Großtheorie. Drittens werden zwei technikinkludierende Konvergenzerzählungen aus der Mitte des 20. Jahrhunderts vorgestellt, die in ihrer Entgegensetzung von Katastrophe und Entrückung wie Blaupausen für die KI-Mythen erscheinen.

Diese Dreiteilung dient einer klareren Darstellung und darf nicht darüber hinwegtäuschen, dass die Entwicklungen aus dem gegenseitigen Austausch hervorgingen: Die technischen Utopien bilden in ihrer Überzeugung, die Welt des Menschen umfassend gestalten und kontrollieren zu können, bis heute den impliziten Hoffnungshorizont der modernen Wissenschaften. Die Konvergenzerzählungen des 20. Jahrhunderts sind in dem Sinne ebenso das Produkt eines sich reflektierenden, wissenschaftlich-materialistischen Fortschrittsoptimismus, wie er seit Beginn der Aufklärung formuliert und in den technischen Utopien wie Francis Bacons *Nova Atlantis* idealisiert und in der weiteren Literaturgeschichte gleichsam problematisiert wurde. Die Vollendung dieser Weltermächtigung durch Technologie zeichnen Mitte des 20. Jahrhunderts Pierre Teilhard de Chardin (1881–1955) wie Günther Anders (1902–1992) in ihren Konvergenzerzählungen in ebenso entgegengesetzter Tonalität; dabei eint sie im Sinne einer Fortschritts- bzw. Verfallsgewissheit der Glaube an die Unabwendbarkeit der kommenden Ereignisse, als auch, dass keine erwartete Zukunft ohne den Zusammenschluss von Mensch und Technik auskommt. Die Kybernetiker der 1950er und 1960er wiederum traten auf, mit ihrer Lehre von der Steuerung von Information eben jenen Schlüssel für die Kontrolle aller Wirklichkeit gefunden zu haben, mit dem die Versprechen der technischen Utopien nun endlich eingelöst werden. Ihre Überzeugung, an der unmittelbaren Schwelle zu einem neuen Zeitalter des Apparates zu stehen, in dem Mensch und Maschine eins werden, verweist wiederum zurück auf den Ursprung des utopischen Denkens in Europa, der mit der Erwartung an ein unmittelbar bevorstehendes, besseres Zeitalter des Geistes durch den Mönch Joachim von Fiore seinen Anfang nahm.

4.1 Literarische Welten: (Alb-)Träume von der technischen Weltermächtigung und einer zweiten Schöpfung

Dass der Android und die hyperintelligente Maschine in der zeitgenössischen Science-Fiction und darüber hinaus zu einem popkulturellen Klischee verkommen sind, darauf wurde bereits verwiesen (s. 2.2.2). An dieser Stelle soll der historische Blick geltend machen, dass das Motiv vom künstlichen Wesen, mit dem der Mensch in eine wie auch immer geartete Beziehung von Meister, Liebhaber, Feind oder gar Sklave eintritt, unmittelbar mit den neuzeitlichen Aspirationen technischer Weltaneignung verknüpft ist. Die chemisch-biologisch-technologische Linie vom künstlichen Menschen grenzt sich dabei zu früheren, offensichtlich religiös-magischen Vorstellungen von der (Wieder-)Erschaffung des Lebendigen in Antike und Mittelalter ab, ohne den seit der Antike bestehende Verweis auf das Göttliche vollständig ablegen zu können:

Für die Antike sind es schließlich die Götter allein, die die tote Materie beleben und so die Menschen erschaffen, unabhängig ob es um eine erstmalige Erschaffung der Menschheit geht oder eine nachträgliche, individuelle Belebung von Einzelnen. Das Tote in Leben zu verwandeln, trägt seither das Signum des Göttlichen.²³⁰ In der Spätantike und dem Mittelalter inkarniert die magisch-mythische Idee des künstlichen Menschen in der Figur des Golems und bezieht von da den Menschen in den göttlichen (Wieder-)Schöpfungsprozess aktiv mit ein. Dieser Proto-Mensch wird in einem kabbalistischen Ritual durch Lehmformung und Lebenshauch (angelehnt an Gen 2) erweckt. In der Sage des Prager Golems erweckt der Rabbi Löw diesen als Beschützer der Gemeinde zum Leben, indem er ihm das Wort für Wahrheit (*ehmeth*) auf die Stirn schreibt und diesen vor jedem Sabbat wieder zu Lehm verwandelt, indem er den ersten Buchstaben tilgt und so das Wort Tod

²³⁰ Venus beseelt die marmorne Statue Galatea, Hephaistos als göttliches Mängelwesen schafft sich zwei golden-metallene Dienerinnen, und die Götter des Olympos erschaffen gemeinsam die allbeschenkte Pandora, die den Menschen bezirzen und für ihren Hochmut bestrafen soll. Wie in der biblischen Genesis Gott den Menschen schafft, so ist es der Götter-Pantheon, der im Atrahasis-Epos die Menschen erschafft, damit sie an ihrer statt die mühselige Arbeit auf Erden verrichten.

(*meth*) stehen bleibt. Als Rabbi Löw eines Tages vor dem Sabbat vergisst, den Golem wieder zu *deaktivieren*, verfällt dieser in Raserei, weil er gezwungen wurde, das Ruhegebot Gottes zu übertreten.²³¹ Ganz in der Tradition der antiken Mythen bestärkt der Golem noch die Souveränität Gottes als auch die Größe seiner Schöpfung. Wie Galatea oder Pandora dekonstruiert der Golem nicht das Mysterium des Menschseins, noch entweiht der Mensch selbst in dieser Mitbeteiligung den Lebensprozess.²³² Gleichzeitig ist der Golem ein frühes Beispiel für das Spannungsfeld zwischen dem Wunsch nach einem treuen Helfer, der seine Dienste klaglos verrichtet, und der Angst, dass die dafür notwendige Autonomie zum Kontrollverlust des Menschen führt. Die Automaten-erzählungen der frühen Romane werden die Frage nach der Verantwortlichkeit des Menschen für sein Geschöpf noch akzentuieren, weil dort ein offensichtlicher Gott als eigentlicher Quell der Schöpferkraft wegfällt und der Mensch so die alleinige Urheberschaft für sein Werk zu verantworten hat.²³³

Die moderne Literatur vom künstlichen Menschen, in der allein der Mensch aus dem Einblick in die Gesetzmäßigkeiten der Natur – erst mithilfe der Alchemie, dann der Naturwissenschaft – Leben schafft, beginnt mit dem Homunculus des Paracelsus und geht über Frankenstein's Kreatur weiter bis zu den belebten Automaten in den frühen *Scientific Romances*. Mit dem Homunculus tritt die Erschaffung künstlicher Wesen im Spätmittelalter erstmals aus dem Metier des Religiösen in den Bereich der (Vor-)Wissenschaft. Der Arzt und Alchemist Paracelsus (1493–1551) beschreibt die Erschaffung eines solchen Menschleins in der Art, dass der Samen außerhalb der Gebärmutter in Mist mit Blut genährt wird und durch Verwesung zum Menschen herangezüchtet werden kann.²³⁴ Die Selbstermächtigung des Menschen über die Welt erhält in der Alchemie noch klarer dämonische, oder zumindest

231 Vgl. Lukas Brand, *Künstliche Tugend. Roboter als moralische Akteure*, 1. Aufl., Regensburg 2018, 12–18.

232 Vgl. Anne Foerst/Regine Kather, *Von Robotern, Mensch und Gott. Künstliche Intelligenz und die existentielle Dimension des Lebens*, Göttingen 2008, 47–48.

233 Vgl. Franz Dittmann, „Mensch und Roboter – ein ungleiches Paar.“, in: Fabian Karsch/Arne Manzeschke (Hgg.), *Roboter, Computer und Hybride. Was ereignet sich zwischen Menschen und Maschinen?*, Baden-Baden 2016, 17–46. Hier 24–25.

234 Vgl. Brand, *Künstliche Tugend*, 12–18.

undurchschaubare Qualität, als ihre Suche nach der Essenz der Dinge dem Arkanen, damit auch dem Göttlichen wie dem Widergöttlichen, nähersteht. Dass den Aspirationen der Alchemie und später der Wissenschaft etwas Widergöttliches anhaftet, verdeutlicht auch Goethes Aufgriff der Homunculusfigur in *Faust II* (1832): Im zweiten Akt *Laboratorium* lässt Goethe den Teufel in Menschenverkleidung bei der Schaffung des Homunkulus zugegen sein, wenn Fausts ehemaliger Schüler Wagner, nun selbst Professor und Anhänger der Alchemie, der entscheidende Moment in der Belebung von Materie gelingt. Weil Wagner den Mephistopheles in Menschenverkleidung ausdrücklich einlädt, dem Spektakel beizuwohnen, wird dramaturgisch erst im Bund des Menschen mit dem Teufel diese zweite Schöpfung vollzogen und damit impliziert, dass sie außerhalb der göttlichen Schöpfungsordnung steht.

Der künstlich hergestellte Mensch ist in der literarischen Tradition der Moderne hingegen gerade das Ergebnis einer augenscheinlich vom Aberglauben geläuterten Wissenschaft, die sich alle Wirklichkeit unterwerfen konnte. Besonders im Aufkommen der technischen Utopie öffnet sich die Zukunft als Ort, in dem sich der Mensch die Welt durch Technologie vollständig aneignet. In der Hoffnung auf eine vollständige Erklärbarkeit und damit Formalisierbarkeit der Welt steckt eine Ermächtigungphantasie über sämtliche physikalische und vitale Prozesse, die den Menschen miteinschließt. Es entwickelt sich die Idee, dass der größte Erweis dafür, die Welt und den Menschen wissenschaftlich vollständig durchdrungen und beherrschbar gemacht zu haben, erst in seiner technologischen Wiedererschaffung erbracht werde. Dieses utopische Motiv führt die frühe Romanliteratur weiter und stellt dabei die Bewunderung für das von Menschen Erschaffene neben den beständigen Zweifel an seinem Werk: Hervor bricht nicht nur der latente Zweifel, ob die Rekreation dem Menschen auch tatsächlich gehorchen werde, sondern auch, ob der Mensch sich nicht in seinem Können überschätzt. Die Erzählungen stellen die Anfrage, ob die Konstrukteure in Selbstgefälligkeit nicht allzu leicht ihren eigenen Trugmaschinen erliegen, die in Wirklichkeit nur schlechte Attrappen des Lebendigen bleiben und das Fehlen des göttlichen Lebensatems zwar überspielen, ihn aber nicht ersetzen können.

4.1.1 Technische Utopien: Die erfolgreiche Unterwerfung der Naturkräfte und ihre späte Rache am Menschen

Für die frühe Wissenschaft entwickelt sich die Nachahmung des Lebens zum imaginären Fluchtpunkt der gesamten wissenschaftlichen Anstrengung, indem sie zum Erweis maximaler Einsicht und Beherrschung der Naturkräfte durch den Menschen stilisiert wird. Dabei sind es die technischen Utopien, die als erste eine solche Allbeherrschung in den Raum des Wünschbaren stellen und darin auch die künstliche Wiederschaffung des Menschen andeuten. Das künstlich erzeugte Wesen wird somit zum imaginären Fixpunkt, dessen Erreichen bedeutet, die Versprechen der technischen Utopien in die Wirklichkeit übersetzt zu haben. Die technischen Utopien nähren so die Hoffnungen einer kommenden Macht durch Wissenschaft, die sich bis zur Verlebendigung steigert.

Solche Träume von einer technischen Allmacht des Menschen finden sich bereits vor Thomas Morus, aber erst nach Joachim von Fiore geschichtsphilosophischer Öffnung hin zur Zukunft als Erwartungsraum, wie er es in seiner Lehre von den drei Weltzeitaltern ausdrückt und damit dem vorherrschenden, augustinischen Geschichtsbild eine Absage erteilt.²³⁵ Mit dieser Entgrenzung der Heilsgeschichte in die Weltgeschichte verwundert es auch nicht, dass sich darauffolgend im

²³⁵ Der durch die Utopien in den Fokus geratene Sehnsuchtsort Zukunft widersprach dem vorherrschenden, christlich vermittelten Geschichtsbild seit der Spätantike. Augustin hatte die latente Naherwartung der neutestamentlichen Texte spiritualisiert und kam in *De civitate Dei* zu dem Schluss, dass die Welt nach dem Auftreten Christi in ihrem Greisenalter angelangt war und nun nichts mehr zu erwarten habe. Es war der Mönch Joachim von Fiore, der den Grundstein des utopischen Denkens im 11. Jahrhundert legte, als er dem augustinischen Geschichtsbild widersprach. Für Joachim und seine Zeitgenossen schien die Heilsgeschichte keineswegs am Ende, die überall zu beobachtende Spiritualisierung aller christlichen Stände deutete vielmehr daraufhin, dass das Beste noch bevorstand. Eingeleitet durch Joachims eigene spirituelle Erfahrung verkündete er, dass der Christenheit ein neues Zeitalter des Heiligen Geistes als drittes Zeitalter bevorstand. Demnach waren alle Menschen bestimmt, sich radikal zu verbessern, er entfachte damit die Sehnsucht, diese Ankunft des Geistes durch eigene Zuarbeit zu beschleunigen. Dieser Impuls hat sich seit dem 11. Jahrhundert über die religiöse Vorstellungswelt hinaus in eine Hoffnung auf die kommende Zeit verstetigt. Vgl. Matthias Riedl, „Die Welt als Kloster. Joachim von Fiore und sein Verfassungsentwurf für die zukünftige Menschheit“, in: Katharina Martin/Christian Sieg (Hgg.), *Zukunftsvisionen zwischen Apokalypse und Utopie.*, Würzburg 2016, 127–147. Hier 135.

Denken eines Theologen das Wissen und seine stetige Erweiterung mit Technik und Handwerk verbinden. So träumt der Theologe Roger Bacon bereits 1260 von Schiffen ohne Ruderer, von Wagen ohne Zugtiere, vom Fliegen in der Luft und vom Tauchen unter Wasser.²³⁶ Die technischen Utopien, die diesem Denken entspringen, säkularisieren die millenaristische Geschichtsdeutung weiter, indem sie Gott als übergeordnetes Subjekt der Geschichte ausblenden und den Menschen zum Subjekt des utopischen Zeitalters erheben.²³⁷ Anders als in der klassischen Utopie verwirklicht sich in den technischen Utopien dieser heilvolle *Un-Ort* nicht über das Durchsetzen einer neuen Staatsform im Sinne von Thomas Morus Werk *Utopia* (1516), sondern die Technologie wird zur Basis eines neuen, besseren Zusammenlebens erklärt.

Die Ambivalenz der klassischen Utopien, lediglich Kritik an der Gegenwart zu üben wie tatsächlich auch ins Positive gewendet eine erste Vision einer alternativen Gesellschaftsform zu sein, gilt in gleichem Maße für die technischen Utopien. Hans Jonas unterscheidet zwischen Utopien im Wortsinne οὐ τόπος, deren Umsetzung unmöglich bleibt, und solchen, deren Realisierung zwar unwahrscheinlich, aber theoretisch machbar ist. Für letztere verweist Jonas auf Platons Philosophenstaat, welchen er in seiner *Politeia* skizziert, wobei Jonas anmerkt, dass man auch diese zweite Sorte letztlich nicht als Handlungsanweisung verstehen sollte.²³⁸ Der Kulturwissenschaftler Thomas Macho bestärkt ebenfalls die kritische Funktion von Utopien: Bei aller Anziehungskraft dieses *Nirgendwo* sei dort, wohin durch die Utopie der Geist hindrange, noch kein Ort für den Menschen und seine Lebensrealität geschaffen. Immer seien diese utopischen Gegengesellschaften auf isolierten Inseln angesiedelt, deren Weg aus dieser Welt niemand kennt, wofür sich in den Erzählungen auch nicht rechtfertigt wird. Sie bleiben weder räumlich noch zeitlich fixiert und somit die Etappen ihrer Verwirklichung unbekannt. Diese Qualität der Utopie reiche bis zu den Entwürfen Her-

236 Vgl. Roger Bacon, „Epistola de secretis operibus artis et naturae, et de nullitate magiae“, in: J. S. Brewer (Hg.), *Rogeri Bacon Opera quaedam hactenus inedita.*, London 1859, 521–552. Hier 530.

237 Vgl. Schipper, Zwischen apokalyptischen Ängsten und chiliastischen Hoffnungen. Die religiöse Dimension moderner Utopien, 56–57.

238 Vgl. Jonas, Das Prinzip Verantwortung, 307–308.

ders oder Blochs, indem sie niemals einfach umsetzbar sind, sondern restlos kontingent blieben.²³⁹

Diesen Fiktionsmarkern zum Trotz wurde seit Morus *Utopia* immer wieder über die Verwirklichung der Sehnsuchtsorte diskutiert. Die Frage der Umsetzung erhält in der technischen Utopie eine besondere Ambivalenz, da sie an das kontinuierliche Wachstum des Spielraums der Machbarkeit über die Zeit glaubt. Das utopische Element wird durchlässiger, als Technik notwendig eine zumindest über die Phantasie erschließbare Verheißung einer grundsätzlichen Durchführbarkeit und Planbarkeit enthält.²⁴⁰ Wie sehr die technische Utopie die epistemologischen Grundlagen und Sinnansprüche der modernen naturwissenschaftlichen Denken bis heute offenlegen, wird besonders an der wohl wirkmächtigsten, wenn auch fragmentarisch gebliebenen technischen Utopie *Nova Atlantis* des britischen Naturphilosophen Sir Francis Bacon (1561–1626), einem der Wegbereiter der empirischen Wissenschaften, deutlich.

Nova Atlantis [1627]

Neu-Atlantis, eine fiktive Insel im damals unbekanntem Teil der Südsee, erscheint dem gestrandeten englischen Protagonisten des 17. Jahrhunderts durch seinen gewaltigen technologischen Fortschritt wie ein Land von Engeln, welches seine Einwohner täglich mit neuen Tröstungen überschüttet.²⁴¹ Geleitet wird dieses Reich von einer Gemeinschaft von Gelehrten, die sich das Haus Salomonis nennen und als das Auge des Reiches vorgestellt werden.²⁴² Den Großteil der Erzählung stellt der Vorsteher des Hauses dem Gestrandeten seine Einrichtung vor:

„Der Zweck unserer Gründung ist die Erkenntnis der Ursachen und Bewegungen sowie der verborgenen Kräfte in der Natur und die Erwei-

²³⁹ Vgl. Macho, Technische Utopien und Katastrophenängste, 13.

²⁴⁰ Vgl. Anke Woschek, „Zwischen Luftschloss und Prognose. Der Terminus der technischen Utopie im Fokus der Technikgeschichte“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 147–164. Hier 149.

²⁴¹ Francis Bacon, „Neu-Atlantis“, in: Klaus J. Heinisch (Hg.), *Der utopische Staat.*, Reinbek bei Hamburg Februar 2017, 171–219. Hier 183.

²⁴² Vgl. ebd., 185.

terung der menschlichen Herrschaft bis an die Grenzen des überhaupt Möglichen.²⁴³

Francis Bacons berühmte gewordene Satzung, dass Wissen über Dinge mit der Macht über sie einhergeht, tritt in den Worten des Vorstehers markant zutage. Dementsprechend zeigt sich in Neu-Atlantis, dass das Haus Salomonis mithilfe des erworbenen Wissens den Herrschaftsbe- reich des Menschen auf der Erde maximal ausgeweitet hat. Von den tiefsten Höhlen, in denen Werkstoffe geronnen, verhärtet, gekühlt und konserviert werden, bis zu den Türmen von bis zu einer halben Meile Höhe, die der Beobachtung verschiedenster Wettererscheinungen die- nen, hat der Mensch die Welt seinen Zwecken zugeführt.²⁴⁴

Die absolute Durchdringung und schließlich Beherrschung der Natur versuchen die Gelehrten über ihre Simulation zu erreichen; sie glauben, dass sich wirkliches Verstehen und damit absolute Kontrolle über die Natur erst in ihrem Nachbau erweist. Nicht nur ahmen sie Wetterphänomene nach, sondern die Machtphantasien erreichen eine kosmische Qualität, indem sie bestimmte Öfen nutzen, um Sonnen- und Sternenstrahlung nachzubilden. Die Kontrolle durch Simulation und Rekonstruktion erstreckt sich auch auf das Reich des Lebendigen. Der Vater des Hauses ist besonders stolz, dass die Bruderschaft Tiere wie Mücken, Schlangen oder Fische aus verwesenden Stoffen züchten kann; indem sie das Unbelebte zeugungslos zum Lebendigen umbilden können, meistern die Naturkundler des Hauses jene Herausforderung, die nach Paracelsus ein Jahrhundert zuvor noch die Alchemie errei- chen sollte. Die Naturforscher gehen sogar weiter – das Haus Salomonis schafft es, diese niederen Wesen, die sie aus der toten Materie hervor- brachten, zu höherentwickelten Gattungen reifen zu lassen.²⁴⁵ Die Uto- pie von *Nova Atlantis* bedeutet im Kern eine Extrapolation auf allen damals bekannten Gebieten der Naturwissenschaft bei fast totaler For-

243 Vgl. ebd., 205.

244 Zum Verständnis der utopischen Sprengkraft für die Zeitgenossen Bacons: Eine solche architektonische Höhe erreichte erst 2010 der *Burj Khalifa* in Dubai mit seinen 828 Metern. Und es zeigt: Die Ambivalenz der technischen Utopien ist berechtigt, ihre Visionen können mithilfe einer fortbetriebenen Wissenschaft manchmal tatsächlich in Erfüllung gehen.

245 Vgl. Bacon, Neu-Atlantis, 205–213.

schungsfreiheit. Alles außer dem direkten Experiment am Menschen ist zum Zweck der Wissens- und damit Machterweiterung geboten.

Der Zensur seiner Zeit geschuldet lässt Bacon in seiner Utopie die anatomischen Kenntnisse über den Menschen nur indirekt über das Sezieren von Tieren gewinnen. Die Argumentation aber ist eindeutig: Das höhere Leben wie das des Menschen lässt sich nicht nur aus dem niederen erklären, sondern bei vollständiger Kenntnis müsste sich auch der Mensch aus den niederen Tieren heraufzuchten lassen. H. G. Wells (1866–1946), den die Frage nach der Rückführbarkeit des Menschen aus dem Tier und seine mögliche Fortentwicklung sein gesamtes literarisches Wirken begleitet hat, führt diese Idee in seinem 1896 erschienenen Roman *The Island of Dr. Moreau* weiter aus. Der Name des besagten Doktors erinnert nicht zufällig an Thomas Morus, führt er doch auf einer einsamen Insel Experimente an Tieren durch, um sie zu menschenähnlichen Chimären umzubilden. Das Tierische soll durch die Mittel der Chirurgie, Pharmakologie, Psychologie sich dem Menschlichen annähern. Gleichwohl lässt Wells diesen Traum über die volle Kontrolle über das Leben in einen Albtraum umschlagen. Die Tiere fallen letztlich doch in ihre wilde Natur zurück und zerreißen den Doktor; die künstliche Wiederschaffung des Menschen aus dem Tier gelingt auch literarisch vorerst nicht.

Der Traum, einst auch den Menschen nachbilden zu können, wird bereits bei Bacon dort expliziter, wo er über *Nova Atlantis* hinaus über die tatsächliche Gründung eines Bundes nach Vorbild des Hauses Salomonis nachdenkt.²⁴⁶ Die *Brotherhood of Lights* ist ein durchgehendes Thema seiner Werke: Er spricht von einer Bruderschaft im Lernen und der Erleuchtung nach monastischem Ideal, die zurück auf Gott als Vater des Lichts verweist und somit sich als erwählt versteht. In der Enge dieser Gemeinschaft sollte der Verlust von Wissen aufgefangen wie auch neue Erkenntnisse hervorgebracht werden. Interessant ist, dass

²⁴⁶ Vgl. Dana Jalobeanu, „Bacon’s Brotherhood and its classical Sources: Producing and Communicating Knowledge in the Project of the Great Instauration“, in: Claus Zittel (Hg.), *Philosophies of Technology*. Francis Bacon and his Contemporaries, Leiden/Boston 2010, 197–230. Hier 203–208. In seinen Aufzeichnungen finden sich eine Liste mit potentiellen Personen, die er für die Gründung einer solchen Gemeinschaft im Sinn hatte, wie bereits auch erste Ideen nach einem passenden Gebäude niedergeschrieben waren.

Bacon an anderer Stelle der Bruderschaft den Namen *Das Kollegium der Werke der sechs Tage* gibt.²⁴⁷ Wenn dieses Kollegium alle Werke der sechs Schöpfungstage Gottes ergründen will, dann ist die Erschaffung des Menschen am sechsten Tag miteingeschlossen und kann – nach dem Credo des Hauses Salomonis – nur mit ihrer Wiederholung als vollständig ergründet gelten.

In Francis Bacons wissenschaftsphilosophischem Hauptwerk, der *Great Instauration*, kommt er zu dem Schluss, dass die Wissenschaften in den Jahrhunderten vor ihm so wenig fortgeschritten seien, da es an Zielen und Mut gefehlt habe. Entgegen dem antiken Ideal der Kontemplation über das Wahre bietet Bacon den neuen Wissenschaften eine in seinen Augen attraktivere Basis an, auf der das Wissen der machtvollen Bereicherung des menschlichen Lebens in der Umgestaltung seiner Welt dient.²⁴⁸ Seine Utopie stellt somit eine natürliche Erweiterung seiner Wissenschaftsreform dar: Was Bacon seinen Lesern gegen die Mutlosigkeit anbietet und in seiner utopischen Technokratie *Nova Atlantis* veranschaulicht, ist ein Versprechen, welches bis heute die ideelle Grundlage moderner Wissenschaft bildet und gleichzeitig die damit verbundene Gefahr ihrer Überforderung offenlegt. Die vollständige Unterwerfung der Natur durch die Technik soll die Menschheit schlussendlich in das Paradies führen:

„Das Menschengeschlecht mag sich nur wieder sein Recht über die Natur sichern, welches ihm kraft einer göttlichen Schenkung zukommt. Mag ihm das voll zuteil werden. Die Anwendung wird indes die richtige Vernunft und die gesunde Religion lenken.“²⁴⁹

Erewhon [1872]

Mit Samuel Butlers (1835–1902) *Erewhon* entsteht zweihundert Jahre später eine Antithese zu Bacons *Nova Atlantis*. Sie ist eine indirekte

247 Bacon, *Neu-Atlantis*, 193–195.

248 Vgl. Francis Bacon, *Neues Organon*. Teilband 1: Lateinisch – Deutsch (Philosophische Bibliothek 400a), 2. Aufl., Hamburg 1999, 203. „Aber das größte Hindernis für den Fortschritt der Wissenschaften, für die Übernahme neuer Aufgaben und das Erforschen neuer Gebiete liegt in der Mutlosigkeit der Menschen, die leicht etwas für unmöglich halten.“

249 Ebd., 272–273.

Fortsetzung wie Umkehrung von Bacons Utopie aus der Perspektive des hochtechnisierten und von den Lehren Darwins erfassten 19. Jahrhunderts. Das Niemandsland (Erewhon – ein Palindrom für *nowhere*) denkt das Schicksal einer hochentwickelten Gesellschaft nach dem Vorbild *Nova Atlantis* weiter und kommt zu dem für Bacon noch unvorstellbaren Schluss, dass die technische Vervollkommnung selbst den Menschen zur Bedrohung werden könnte. Butlers Utopia erscheint deshalb von ganz anderer Art: Der Erzähler trifft auf eine Gesellschaft voller vital wirkender Menschen, in der (Prä-)Eugenik betrieben wird. Krankheit ist ein Verbrechen und wird mit Zölibat bestraft. Viel wunderlicher jedoch handelt es sich um eine Gemeinschaft, die sich technologisch selbst beschnitten und freiwillig in eine vorindustrielle Lebensweise zurückentwickelt hat.²⁵⁰ In Erewhons Museum der zerstörten Maschinen erfährt der Protagonist die Ursache für das gemeinschaftlich auferlegte technologische Zölibat. Vierhundert Jahre vor dem Eintreffen des Protagonisten war der Stand des mechanischen Wissens so weit fortgeschritten, dass einer der gelehrtesten Professoren das unmittelbare Entstehen von Instinkt in den Maschinen prognostizierte. Er ging davon aus, dass dieser Instinkt viel höher entwickelt als derjenige der Tiere und Menschen sein werde und der deshalb die Unterdrückung der Menschen durch ihre Maschinen befürchtete. Um diese Katastrophe abzuwenden, verbannten die Einwohner Erewhons alle Maschinen, die kürzer als zweihunderteinundsiebzig Jahre in Gebrauch waren und verboten strikt alles weitere Erfinden.²⁵¹ Zusätzlich wurden sämtliche Aufzeichnungen zerstört oder verborgen, jede Spur der Opposition gegen den Maschinensturm wurde unkenntlich gemacht und es gemeinhin untersagt, die Vergangenheit aufzuarbeiten. Erewhon erscheint gerade darin als fantastischer *Un-Ort*, dass dort das kollektive Vergessen technologischer Errungenschaft zum vermeintlichen Wohle der Menschheit gelingt. Schließlich jedoch stößt der Protagonist auf das Buch der Maschinen, das die Zustände vor dem Maschinensturm in Form eines Streitgesprächs zwi-

250 Vgl. Samuel Butler, *Erewhon*. or, Over the range, New York 1917 (1872), 68–69.

251 Vgl. ebd., 86–87.

schen einem Antimaschinisten und einem Maschinisten wiedergibt.²⁵² Dieses *Book of Machines* – eine Schrift in der Schrift – ist insofern interessant, weil es sich erstens wie die Ältestenrede in *Nova Atlantis* um einen argumentativen Text umschlossen von einer narrativen Rahmenhandlung handelt, der den Leser zeitunmittelbar anspricht und Butler mit diesem Kunstgriff seine Gegenwart zum entscheidenden Wendepunkt der Vergangenheit Erewhons inszeniert; zweitens verdient die Streitschrift besondere Aufmerksamkeit, weil sich in ihrer Auseinandersetzung um Wohl und Übel des maschinellen Fortschritts die erste Erwähnung von der Möglichkeit eines maschinellen Bewusstseins im modernen Sinne findet:

„There is no security – to quote his own words — against the ultimate development of mechanical consciousness, in the fact of machines possessing little consciousness now. [...] But who can say that the vapor engine has not a kind of consciousness? Where does consciousness begin, and where end? [...] Is not machinery linked with animal life in an infinite variety of ways?“²⁵³

Der langsamen, biologischen Evolution wird in dem Streitgespräch von Seiten der Maschinenstürmer die Schnelligkeit der technischen Entwicklung entgegengestellt, die dem Menschen zur Bedrohung gerate. Auch wenn das Leben der Maschine ohne eigenen Willen auszukommen scheint, spreche dies nicht gegen die Annahme, im Dampfmotor den Keim einer neuen, höheren Phase des Lebens zu sehen. Schließlich folge die Maschine wie der Mensch in seinem Leben berechenbaren Mustern der Selbst-Regulierung – mit dem einzigen Unterschied, dass der Mensch sich seiner Bedürfnisse bewusst sei.²⁵⁴ Der Kampf um die künftige Vorherrschaft zwischen Biologie und Mechanik sei somit ernst, als die Menschen sich kurz vor jenem Schwellenmoment vorfinden, in welchem die Menschen die Maschinen hervorbrächten, denen sie dienen werden:

252 Vgl. ebd., 229.

253 Ebd., 233–234.

254 Vgl. ebd., 261.

„Is it not plain that the machines are gaining ground upon us, when we reflect on the increasing number of those who are bound down to them as slaves, and of those who devote their whole souls to the advancement of the mechanical kingdom? Are we not ourselves creating our successors in the supremacy of the earth? Daily adding to the beauty and delicacy of their organization, daily giving them greater skill and supplying more and more of that self-regulating self-acting power which will be better than any intellect?“²⁵⁵

Die Maschinenstürmer halten jenen Prozess der Selbstversklavung für unumkehrbar, weil die Menschen bereits zu diesem Zeitpunkt so sehr von ihren Maschinen abhingen, dass deren Verschwinden zu einem Aussterben der Art nach spätestens sechs Wochen führen müsste. Die Gewohnheit hingegen werde dazu führen, dass die Abhängigkeit unmerklich in allmähliche Knechtschaft übergeht, ohne dass es jemals einen Zusammenstoß der Wünsche zwischen Mensch und Maschine gebe. Die Maschinen würden sich in der Zukunft untereinander bekriegen, aber sie würden immer noch den Menschen für die Geburt ihrer Jungen, zur Instandhaltung, Wiederherstellung und zum Recycling ihrer Körper benötigen. Auch wenn der Mensch in Zukunft die unterlegene Rasse sei, werde es ihm unter der wohlthätigen Herrschaft der Maschinen möglicherweise sogar besser gehen als zuvor.²⁵⁶ Wie bei Bacon ist auch für Butler die Utopie ein Mittel, um die Gedanken seiner lebensweltlichen Schriften zu pointieren: Große Teile seiner Argumentation aus seinem Buch der Maschinen gleichen den Passagen seines Zeitungsartikels *Darwin among the machines*, in dem er offen die Übertragung der Lehren Darwins auf das Verhältnis von Mensch und Maschine benennt. Hier reflektiert Butler zudem, dass im Gegensatz zu seiner Utopie die Wirklichkeit keine Insel ist, in welcher sich der technologische Fortschritt per bloßem Dekret aufhalten lässt. Butler ist deshalb davon überzeugt, dass die Menschen den Bruch mit den Maschinen anders als in *Erewhon* nicht vollziehen werden und in Zukunft degenerieren müssen, während die sie versorgenden Maschinen sich über sie erheben:

255 Ebd., 247.

256 Vgl. ebd., 265–266.

„What sort of creature man's next successor in the supremacy of the earth is likely to be. We have often heard this debated; but it appears to us that we are ourselves creating our own successors; [...] we shall look up to them as the acme of all that the best and wisest man can ever dare to aim at. No evil passions, no jealousy, no avarice, no impure desires will disturb the serene might of those glorious creatures. Sin, shame, and sorrow will have no place among them. Their minds will be in a state of perpetual calm, the contentment of a spirit that knows no wants, is disturbed by no regrets.“²⁵⁷

Seinen Artikel beschließt Butler abrupter Weise dennoch mit der Aufforderung, die Welt müsse augenblicklich sämtliche Maschinen zerstören und damit Erewhons Beispiel folgen – hier wird ein tatsächlicher Krieg gegen die Maschinen ausgerufen. Die Aufforderung zum Maschinensturm bei vorheriger Feststellung seiner faktischen Unmöglichkeit in der Lebenswelt bleibt somit eine Mischung aus Satire und gedankenvoller Spekulation:

„Our opinion is that war to the death should be instantly proclaimed against them. Every machine of every sort should be destroyed by the well-wisher of his species. Let there be no exceptions made, no quarter shown; let us at once go back to the primeval condition of the race. If it be urged that this is impossible under the present condition of human affairs, this at once proves that the mischief is already done, that our servitude has commenced in good earnest, that we have raised a race of beings whom it is beyond our power to destroy, and that we are not only enslaved but are absolutely acquiescent in our bondage.“²⁵⁸

Butlers Idee vom Maschinensturm als letztem Ausweg am Wendepunkt des Machtverhältnisses zwischen Mensch und Maschine inspirierte im darauffolgenden Jahrhundert eines der erfolgreichsten Werke der Science-Fiction: Frank Herberts Epos *Dune* (dt. *Dune – der Wüstenplanet*)

257 Samuel Butler, „Darwin among the machines“, *The Press* (Christchurch in New Zealand) (13.06.1863).

258 Ebd.

von 1965 spielt in einem von Menschen kolonisierten, hochtechnisierten Universum in der weit entfernten Zukunft, das sich soweit entwickeln konnte, obwohl es in seiner Vergangenheit vor tausenden von Jahren erfolgreich gegen denkende Maschinen rebellierte und seither sämtliche Computer verboten hat. In der Welt von *Dune* existieren deshalb Menschen, die gewaltige Rechenoperationen übernehmen können, sog. *Mentats*, und die Computer seither ersetzen. Diesen Maschinensturm aus der Vorgeschichte *Dunes* bezeichnet Herbert als *Butler's Jihad* und verweist damit offensichtlich auf das Gedankengut Samuel Butlers. Auch die Begründung für diesen *heiligen Krieg* gegen die Maschinenintelligenz – dass der Mensch sich zu abhängig gemacht habe und sich seiner eigenen evolutionären Weiterentwicklung verschließe – gleichen den Argumenten Butlers. Der Wahlspruch der Maschinenstürmer in *Dune* lautet, den Menschen niemals durch die Maschine zu ersetzen:

„Then came the Butlerian Jihad – two generations of chaos. The god of machine-logic was overthrown among the masses and a new concept was raised: Man may not be replaced.“²⁵⁹

Das Chaos des darauffolgenden Maschinensturms, das durch den Ausfall der Maschinen Millionen Leben kostete, prägt auch die entstehende synkretistische Religion von *Dune*. Deren Referenzwerk, die *Orange Catholic Bible*, verbietet an erster Stelle jede erneute Konstruktion von denkenden Maschinen und erinnert darin an die Verbotspolitik Erewhons:

„Its chief commandment remains in the O.C. Bible as: Thou shalt not make a machine in the likeness of a human mind.“²⁶⁰

Doch die Referenzen zu Butler gehen über das Verbot maschineller Intelligenz hinaus zu dem Gebot menschlicher Eugenik. Die wirkmächtigste religiöse Bewegung in der Welt von *Dune* agiert im Stil eines futuristischen Haus Salomonis 2.o: Der neodarwinistische Schwestern-

259 Frank Herbert, *Dune* (Penguin classics), New York 2016, 639.

260 Ebd., 664.

orden der *Bene Gesserit* unterhält seit der Zerstörung der Computer ein eugenisches Zuchtprogramm, an dessen Ende sie den einen Menschen hervorbringen wollen, der schließlich die gesamte Menschheit auf eine neue Stufe ihrer Entwicklung führen wird.

4.1.2 Der täuschende Apparat in der frühen Romanliteratur und Science-Fiction

Keine andere Gattung hat sich der Figur des künstlichen Menschen so angenommen, ist gleichsam zu ihrem Sinnbild geworden wie die Science-Fiction, kommt doch kaum eine Vorstellung der näheren oder entfernten Zukunft ohne Androiden und sprechende Supercomputer aus. Gleichwohl ist das von Menschen hergestellte, technische Abbild bereits in der Romanliteratur des 19. Jahrhunderts zugegen und stellt ein wichtiges Bindeglied von technischer Utopie zu moderner Science-Fiction dar. Der Beginn dieser *Scientific Romances* wird diskutiert zwischen Johannes Keplers Traumreise zum Mond *Somnium* (1634), Jonathan Swifts *Gulliver's Travels* (1726) oder Mary Shelleys *Frankenstein* (1818). Für Letzteres spricht, dass es der womöglich erste Roman ist, der *Science* im tatsächlich modernen Sinne thematisiert.²⁶¹ Entfernt von einem alchemistischen Animismus wird hier das Verlebendigen als Konsequenz des Studiums der Naturphilosophie, Chemie und Anatomie begründet, wenn Doktor Frankenstein sein Wesen wie ein Mechaniker aus Einzelteilen zusammenbaut. Genrebildend schlägt im Moment der Verlebendigung der Stolz des Tüftlers Frankenstein in sofortiges Entsetzen vor der eigenen Schöpfung(-smacht) um, sodass der Doktor flieht, ehe er den sanften Charakter seiner Kreatur erkennen kann. Erst durch die gesellschaftliche Ablehnung wandelt sich das Wesen vom Menschenfreund zum Monster. Mary Shelley erschuf mit Frankenstein einen Schöpfermythos der modernen Welt, in der der Mensch gottähnlich ist und zugleich sich mit seinen Geschöpfen die Welt teilen muss, in der die Kreatur ihren Schöpfer schließlich konfrontiert und herausfordert. Es sind deshalb diese Worte, die die Kreatur im Wiedertreffen mit dem Doktor in der Eiswüste spricht, die die

261 Vgl. Tomas P. Weber, *Science Fiction*, 1. Aufl., Frankfurt am Main 2005, 7.

menschliche Furcht vor der Überwältigung zum Ausdruck bringen und als latente Drohung fortwirkten: Du bist mein Schöpfer, aber ich bin dein Meister. Gehorche!²⁶²

E.T.A. Hoffmann: Der Sandmann [1816]

Nur zwei Jahre vor Shelleys Roman entstand E.T.A. Hoffmanns romantische Schauererzählung, *Der Sandmann*, über die augenscheinliche Belebung eines Automaten. Im Motiv des künstlichen Menschen mischen sich hier in beeindruckender Weise magische Praktiken und Wissenschaft und fragen an, ob die Macht über seine technische Reproduzierbarkeit dem Menschen gefährlich werden kann, indem sie ihn aus seinen sicheren Lebensordnungen drängt. Die Einsicht, dass mit der Konstruktion eines menschenähnlichen Automaten eine heilvolle Grenze überschritten werde, zeigt Hoffmann am Schicksal seiner Hauptfigur, die im Wahnsinn sich selbst umbringt. Hoffmann erzählt von dem jungen, schwärmerischen Nathanael, der eine verhängnisvolle Obsession zu dem Automaten Olimpia entwickelt. Nathanael bemerkt Olimpia aus dem Fenster seines Studierzimmers im gegenüberliegenden Nachbarhaus seines Physikprofessors Spalanzani. Obwohl die Gestalt tagelang regungslos an einem Tisch sitzt, geht Nathanael nicht auf, dass es sich bloß um eine mechanisch betriebene Puppe handelt. Er bemerkt ihren schönen Wuchs, doch bleibt ihm die unbewegliche Olimpia zunächst noch gleichgültig. Erst als er von dem dubiosen Mechanikus Coppola ein Taschenperspektiv erwirbt und damit Olimpia aus seinem Fenster erspäht, scheint sie ihm lebendig und begehrenswert zu werden:

„Nur die Augen schienen ihm gar seltsam starr und todt. Doch so wie er immer schärfer und schärfer hin schaute war es als gingen feuchte Mondesstrahlen in Olimpias Blicken auf – Es schien als wenn nun erst die Sehkraft entzündet würde, immer lebendiger und lebendiger flamen

²⁶² „You are my creator, but I am your master; obey!“ Mary Shelley, *Frankenstein*. Or, the modern Prometheus, London 2012, Chapter XX, 2.

die Augen – Nathanael lag wie festgezaubert im Fenster immer fort und fort nach der himmlisch schöne Olimpia schauend.²⁶³

Die Erzählung, die in der Hochzeit des Mesmerismus entstand, spielt mit der Idee, dass Nathanael durch das Perspektiv seine Seelenenergie auf den Automaten überträgt und so verlebendigt. Die Vorstellung, dass tatsächlich unsichtbare Komponenten im Sehen wirken, kann auf die Entdeckung der ultravioletten und infraroten Strahlung um 1800 durch Herschel und Ritter zurückgeführt werden. Hoffmann zumindest spricht ununterbrochen von einer Hitze im Sehen in flammenden Blicken oder dem zündenden Liebesblick.²⁶⁴ Das Fernglas als ursprüngliches Mittel zur Erweiterung der Sicht wird in der Erzählung gleichsam zum verzaubernden Utensil. Wissenschaft und Magie umschlingen sich weiter, da Nathanael in dem Mechanikus Coppola den Advokaten Coppelius wiederzuerkennen meint, der in Nathanaels Jugend mit seinem Vater alchemistische Versuche unternommen hatte, bei denen der Vater ums Leben kam. Die Mutter erzählte den Kindern, immer wenn Coppelius abends zu Besuch kam, dass es der Sandmann aus dem gleichnamigen Märchen sei, der Kindern im Schlaf die Augen stiehlt. Auch Nathanael scheint mehr und mehr zu erblinden: Je länger Nathanael sich in Olimpia *verguckt*, desto eher wirken andere Menschen auf Nathanael wie Unmenschen, gar Teufel. Am stärksten äußert sich diese Misanthropie gegenüber seiner Verlobten Clara:

„Nathanael blickt in Clara's Augen – aber es war der Tod, der ihn mit Clara's Augen freundlich anschaute.“²⁶⁵

Als Clara – ihrem sprechenden Namen gemäß – versucht, ihn zur Besinnung zu rufen, beschimpft Nathanael sie als „*lebloses verdamm-*

263 E.T.A. Hoffmann, „Der Sandmann“, in: Carl Georg von Maassen (Hg.), *E.T.A. Hoffmanns Sämtliche Werke*. Historisch-kritische Ausgabe mit Einleitungen, Anmerkungen und Lesarten von Carl Georg von Maassen, München, Leipzig 1909, 3–42.

264 Vgl. Frank Wittig, *Maschinenmenschen*. Zur Geschichte eines literarischen Motivs im Kontext von Philosophie, Naturwissenschaft und Technik (Epistemata Reihe Literaturwissenschaft 212), Würzburg 1997, 63.

265 Hoffmann, *Der Sandmann*, 16.

*tes Automat*²⁶⁶. Mit der zeitweiligen Lösung der Verlobung steigert sich Nathanaels Obsession zu Olimpia und nimmt offen erotische Züge an. Der Umstand, dass alle anderen Menschen in Olimpia weiterhin einen Automaten erkennen, wird spätestens bei dem Tanzabend im Haus Spalanzanis deutlich. Während Nathanael liebestrunken mit der Puppe tanzt und sie küsst, unterdrücken die restlichen Gäste ihr Gelächter. In seiner Obsession stilisiert Nathanael Olimpia zur idealen Geliebten: Ihr kann Nathanael seine Gedichte vortragen, während Olimpia nichts weiter als ein dauerndes „Ach! Ach!“ von sich gibt, das er als tiefstes Verständnis für seine Gedanken interpretiert.²⁶⁷ Seine schwärmerische Auflehnung gegen eine als automatenhaft empfundene bürgerliche Lebensweise führt ihn ironischerweise in die Hörigkeit gegenüber einer Maschine. Sein schwärmerischer Individualismus kippt im Wahnsinn in ein paranoisches Delirium, in dem die Menschen wie von Dämonen bewegte Holzpüppchen erscheinen:²⁶⁸ Monate später, nachdem Nathanaels Psychose geheilt scheint und er zu Clara zurückgekehrt ist, schaut er versehentlich noch einmal in das Perspektiv und verfällt in eine erneute Raserei, in der er versucht, Clara von einem Turm zu stürzen, bei dem Versuch aber selbst fällt und stirbt.

„Die Figur Nathanaels zeigt in unheimlicher Deutlichkeit, dass Angst und Auflehnung durchaus von der Lust an der Unterwerfung überlagert werden können: Nathanael erlebt, wenigstens zeitweise, sein Dasein als von übermächtigen gesteuerte Marionette durchaus lustvoll.“²⁶⁹

In *Der Sandmann* geht es um weit mehr als eine Männerphantasie und die Lust an der Unterwerfung. Gleichsam steht da der wiederkehrende Fiebertraum Nathanaels, der Coppelius und seinen Vater bei ihren Versuchen beobachtet, wie dieser ihm die Arme und Beine brechen und

266 Ebd., 17.

267 Vgl. ebd., 32.

268 Vgl. Roland Innerhofer, „Die technische Modernisierung des künstlichen Menschen in der Literatur zwischen 1800 und 1900“, in: Bernd Flessner (Hg.), *Nach dem Menschen. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur*, Freiburg im Breisgau 2000, 69–99. Hier 97–98.

269 Vgl. ebd., 82.

an anderer Stelle wieder anbringen will. Der zerschmetterte Körper Nathanaels am Ende der Erzählung ist sinnbildlich für seine eigene Demontage. Die Erzählung spielt mit der Furcht vor einer durch die Wissenschaften freigelegten Modulbauweise des Menschen. Hoffmann extrapoliert die aufgekomenen Vorstellungen seiner Zeit vom Menschen als Baukastensystem und problematisiert die Auswirkungen dieser avancierten Technologie auf die mit ihr konfrontierten Menschen. Die Konstruktion eines künstlichen Menschen – der Androidin Olimpia – ist auch hier nur die logische Konsequenz all der Versuche, die Entwicklung, den Aufbau und die Funktionsweise des Menschen zu begreifen und zu replizieren.²⁷⁰ Deutlich wird dies an der Figur Spallanzani; er ist in Hoffmanns Novelle Nathanaels Physikprofessor und taucht als einer der Väter der Androide auf. Sein Name verweist auf den Naturforscher Lazzaro Spallanzani (1729–1799), der durch das Sezieren von Föten bekannt wurde und aus den gewonnenen Erkenntnissen als einer der ersten eine künstliche Befruchtung beim Tier und damit tatsächlich ein (Mit-)Schöpfungsakt gelang. Das Eindringen der Naturwissenschaften in den menschlichen Körper und den Zeugungsprozess wird bei Hoffmann zum Auftritt des *anatomischen Satans* stilisiert – ein Schlüsselbegriff aus seiner zwei Jahre zuvor erschienenen Erzählung *Die Automate* (1814), die gleichsam die thematische Vorlage für *Der Sandmann* bildet. Die Wiedererschaffung des Menschen erscheint als existentielle Gefahr, in der die Wissenschaft selbst zur dunklen Magie gerät; Wissenschaftler werden zu bösen Magiern, zu Sandmännern, die Augen stehlen und sie stattdessen den technischen Replikaten einsetzen. Hoffmann lässt die Sorge im Raum stehen, ob in diesem Versuch der Wiedererschaffung der Mensch die ihn schützende Ordnung der Welt verlässt und in einen Wissensbereich vordringt, wo er es mit dämonischen Mächten zu tun bekommt.²⁷¹ Die Sorge um den bleibenden Wert des Menschen im drohenden Angesicht seiner technologischen Desintegration führt innerhalb der Erzählung zur kollektiven Paranoia:

270 Vgl. Wittig, *Maschinenmenschen*, 63.

271 Vgl. ebd., 74–81.

„Aber viele von den hochzuverehrenden Herrn beruhigten sich nicht dabey, die Geschichte mit dem Automat hatte tief in ihrer Seele Wurzel gefaßt, und es schlich sich ein in der Tat abscheuliges Mißtrauen gegen menschliche Figuren ein. Um nun ganz überzeugt zu werden, daß man keine Holzpuppe liebe, wurde nun verlangt, daß die Geliebte manchmahl etwas taktlos singe und tanze, daß sie bey dem lesen strike, stike usw. Vor allen Dingen aber, daß sie nicht bloß höre, sondern auch manchmahl in der Art spreche, daß dies Sprechen wirklich ein Denken und Empfinden als unerläßlich voraussetze.“²⁷²

Jules Vernes: Das Karpatenschloss [1892]

Jules Verne (1828–1905), der große Meister der frühen *Scientific Romances*, erzählt von einem scheinbar verlassenen Schloss, das von den angrenzenden Dorfbewohnern gemieden wird, weil diejenigen, die sich dem Schloss zu nähern versuchen, von Spukerscheinungen berichten. Einer Gruppe wird es gelingen, den Spuk zu entlarven und die Machenschaften aufzudecken, die hinter dem Spuk stecken und die das Geschehen im Schlossinneren beschützen sollen. Wie wegweisend Jules Verne für den visionären Selbstanspruch der kommenden Science-Fiction Literatur geworden ist, verdeutlicht schon sein Prolog zum Karpatenschloss, in dem er seine Erzählung nicht als phantastisch verstanden wissen will, da sie mit dem Fortschreiten der Wissenschaft bereits am nächsten Tag wahr werden könnte. Er selbst erklärt die neuen Erzählungen zu den Nachfolgern der klassischen Sagen, in denen nun das Technische den Part des Magischen vollständig übernimmt.²⁷³ Diesem Motto getreu achtet Vernes in seiner Erzählung im Gegensatz zu Hoffmanns *Der Sandmann* akribisch darauf, sämtliche Begebenheiten auf rein mechanisch-elektrische Gründe zurückzuführen. Die Ablösung der Magie durch die Technik, aber auch das bleibend Magische in der Technik, welches den Menschen in seinen Bann zieht, ist das Leitthema der Erzählung. In diesem Wechselspiel stellt das Werk ebenfalls die Frage nach der technischen Reproduzierbarkeit des Menschen, insbe-

272 Hoffmann, *Der Sandmann*, 39.

273 Vgl. Jules Verne, *Das Karpatenschloß* (Collection Jules Verne 62), Berlin/Herrsching 1984, 5.

sondere, ob durch die Technik die Seele eines Menschen eingefangen und somit über seinen Tod hinaus erhalten werden kann.

Kennzeichnend für die Rolle der Technik innerhalb der Erzählung ist, wie sich die Figuren immer wieder am Scheinbaren täuschen, wobei die Apparate wie im *Sandmann* sowohl Mittel der Erkenntnis und damit der Auflösung von Schein als auch selbst Verursacher der Täuschung sind. Mit dem Fernrohr offenbart sich der gespenstische Nebel um das Karpatenschloss als Rauch aus einem der Schlote, gleichzeitig wird das Fernrohr als ein Jahrmarktinstrument zur Belustigung der einfachen Leute verhöhnt. Die telefonische Verbindung vom Schloss in das angrenzende Dorf wird einerseits zum Abhören genutzt und dient zugleich als manipulatives Mittel, um im Dorf Stimmen wie geisterhaft ertönen zu lassen. Dass die Dorfbevölkerung das benachbarte Karpatenschloss aufgrund von Spukerscheinungen meidet, stellt sich im Nachhinein ebenfalls als Ergebnis eines technischen Täuschungssystems heraus. Seesalze verleihen allen Gegenständen um das Schloss ein geisterhaftes Ansehen; große Nebelhörner erzeugen unheimliche Heullaute; photographische Silhouetten vor mächtigen Spiegeln erzeugen Gespenstererscheinungen; Stromstöße lassen Feinde erstarren, die dem Schloss zu nahe kommen.²⁷⁴

„Die Ereignisse jener abenteuerlichen Nacht erklärten sich ja ungemein leicht, wenn menschliche Wesen – Übeltäter oder andere – die sich in der Burg aufhielten, die nötigen Apparate besaßen, um allerlei Zauber- und Spukerscheinungen hervorzurufen.“²⁷⁵

Diese Täuschungsmanöver dienen allein dem Zweck, den Aberglauben an Geister aufrechtzuerhalten und somit Fremde von dem Schloss und seinem Schlossherren Graf Rudolph von Gorst fernzuhalten. In der Isolation gibt sich der Graf der technologischen Replikation einer von ihm verehrten, jedoch verstorbenen Opernsängerin hin. Sie, La Stilla, erscheint schließlich auch jenen Eindringlingen, die sich von den restlichen Apparaturen nicht haben blenden lassen:

274 Vgl. ebd., 198–199.

275 Ebd., 149.

„La Stilla stand auf der Estrade in vollem Licht, mit aufgelöstem Haar und vorgestreckten Armen, so wunderbar schön in dem weißen Gewand der Angelica aus dem ‚Orlando‘, ganz so, wie sie sich auf der Bastion der Burg gezeigt hatte. Ihre auf den jungen Grafen gerichteten Augen bohrten sich tief in die Seele ein. [...] Da begann la Stilla zu singen. [...] Es erschien wie ein Hauch, der ihr über die Lippen wehte, ohne dass sie diese dabei öffnete... Und doch, hatte sie auch die Vernunft verloren, die Künstlerseele war in ihr unberührt geblieben!“²⁷⁶

Mehrfach wird betont, dass der vernommene Gesang die Seele der Sängerin in sich trüge. Letztlich erweist sich die Erscheinung ebenfalls als ein ausgeklügeltes Zusammenspiel verschiedener Apparate: Ein lebensnahes Gemälde, Spiegelscheiben und Beleuchtung erschaffen eine Projektion im Raum, die von Audiomitschnitten ihres letzten Auftritts komplettiert wird.²⁷⁷ In einer gespenstischen Reanimation erwacht die Sängerin so zu einem medialen Nachleben. Die Erzählung stellt damit die Frage nach der Reproduzierbarkeit des Menschen. In diesem Fall ist selbst die als einmalig geltende Künstlerseele vor der Rekreation der Technik nicht sicher. Kann man die Seele eines Menschen in Maschinen einfangen und ihn so über seinen biologischen Tod hinaus unsterblich machen? Und welche Konsequenzen hat diese Macht der Technik auf den Menschen? In der Erzählung führt die Obsession des Grafen mit seiner *Junggesellenmaschine*²⁷⁸ ihn ins Verderben. Denn die durch die Apparate geschaffene Idylle einer narzisstischen Selbstbespiegelung ist um den Preis seiner Isolation von der Außenwelt erkaufte, deren Aufrechterhaltung durch Abschreckung und Illusion letztlich in die Katastrophe führt.²⁷⁹

276 Ebd., 209–210.

277 Vgl. ebd., 192.

278 Vgl. Innerhofer, Die technische Modernisierung des künstlichen Menschen in der Literatur zwischen 1800 und 1900, 87.

279 Vgl. ebd., 88–89.

4.1.3 Erträge

Mit dem Aufkommen der Utopie tritt die Zukunft als ein von Menschen gestaltbarer Ort in besonderes Bewusstsein. Jedoch: Wer sich nicht sorgt, braucht nicht an morgen denken. Aus der Idee der Gestaltbarkeit der Zukunft ergibt sich gleichfalls die Sorge um die Zukunft. Im Fall der technischen Utopien besteht ein spezifischer Mehrwert der Ängste, der gerade mit dem Erfolg und mit dem Gelingen der Zukunftsprojekte zusammenhängt. Technische Utopien bilden eine Allianz mit der Angst vor der Katastrophe, die sowohl im Ausbleiben wie auch durch die Verwirklichung der Utopie hervorgerufen werden. Die leuchtenden Farben der Zukunft werden gleichsam dunkel grundiert, jede technische Utopie scheint mit sich ihre eigenes Gegengewicht in der implizierten Katastrophe zu tragen.²⁸⁰ Die in der Zukunft versprochenen Großartigkeiten kommen nicht ohne die mit in dieser Zukunft lauernernden Grausamkeiten aus: Bacons Verheißung von nichts Geringerem als einem technikinduzierten Eingang des Menschen in das Paradies legt gleichsam den Grundstein für die folgenden Katastrophendiagnosen. Dass die Macht des Menschen ihn selbst vertilgen könnte, dieser Gedanke wird vom Programm der technischen Utopien geradezu herausgefordert; denn die von Bacon eingeführte Machtfrage strahlt gleichsam auf den Menschen als Teil der kontrollierbaren Natur zurück:

„Die Verheißung der Technik ist in Drohung umgeschlagen bzw. hat sich mit dieser unlösbar verbunden. Die vorangetriebene Unterwerfung der Natur erstreckt sich in dem Übermaß ihres Erfolgs weiter auf die Natur des Menschen.“²⁸¹

Gleichzeitig gerät die Zukunft des Menschen selbst, besonders unter dem Eindruck der Entwicklungslehre Darwins, zum Gegenstand der Spekulation. Der Darwinismus hatte schließlich eine neue geschichtliche Dimension in das Naturverständnis integriert, das zuvor ledig-

²⁸⁰ Vgl. Macho, Technische Utopien und Katastrophenängste, 12.

²⁸¹ Jonas, Das Prinzip Verantwortung, 7.

lich als ein Regel- bzw. Gesetzessystems verstanden wurde.²⁸² Charles Darwins (1809–1882) Hauptwerk *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* von 1859 enthält keine Aussagen über den Menschen, aber in seinem Schlusssatz spekuliert er offen über das einfache Leben und seine Potenz hin zu endlosen Formen höchster Schönheit. In der Aufarbeitung der Geistesgeschichte des 19. Jahrhunderts wird oftmals einseitig hervorgehoben, wie der Darwinismus den Menschen auf ein den Tieren gleiches Geschöpf entwertete, obwohl er zugleich auch die Frage nach einem höheren Wesen implizierte, das auf den Menschen folgt.²⁸³ Die auf Darwin folgende Popularisierung der Evolutionstheorie bestärkt letztere Lesart. Ernst Haeckel (1834–1919) stilisiert die Evolutionstheorie zu einer Geschichte des fortlaufenden Progresses hin zur Vervollkommnung allen Seins. Schon den bloßen Stillstand – in der Biologie als auch sämtlichen Lebensbereichen – erklärt Haeckel zum Rückschritt, und jeder Rückschritt, sagt er, trage bereits den Keim des Todes in sich. Nur dem Fortschritt, und damit der Veränderung gehöre die Zukunft.²⁸⁴ Erst in dieser populären Ausweitung von Darwins Theorie war es möglich, den im 18. Jahrhundert formulierten Fortschrittsgedanken in die Anthropologie vollends zu übertragen. Die Sorge, dass die biologische Entwicklung des Menschen keine Zukunft hat, ist die epistemologische Grundlage der gegenwärtigen Konvergenzerzählungen um Künstliche Intelligenz. Der Mensch wird entweder durch seinen eigenen technischen Fortschritt verdrängt und ersetzt (s. 5.1), als mangelhaftes Wesen technologisch eingeht (s. 5.2), oder der gegenwärtige Mensch als Wachstumsstörung hin zu seiner vollkommenen, nichtbiologischen Form gedeutet (s. 5.3). Folgenreich bis zu den gegenwärtigen Überwältigungserzählungen um Künstliche Intelligenz ist dabei, dass die Frage nach der Zukunft des Menschen bereits früh mit der Fortentwicklung von Maschinen parallelisiert wurde und in diesem Techno-Darwinismus nicht mehr zwischen den

282 Vgl. Matthias Herrgen, „Leben bei Darwin und im Darwinismus“, in: Stephan Schaeede (Hg.), *Das Leben*. Historisch-systematische Studien zur Geschichte eines Begriffs, Tübingen 2012, 45–60. Hier 52.

283 Vgl. Krüger, *Virtualität und Unsterblichkeit*, 241–242.

284 Vgl. Ernst Haeckel, *Gemeinverständliche Vorträge und Abhandlungen*, Bonn 1902, 29–30.

sehr unterschiedlichen Vorgängen von Entstehen und Konstruieren unterschieden wird.²⁸⁵ Das Maschinenreich wird ebenfalls unter den Evolutionsgesetzen des Lebendigen betrachtet und in dieser Gleichstellung ein Kampf ums Dasein zwischen Mensch und Maschine plausibilisiert. Die daraus angenommene Bedrohungslage des Menschen, in der es um nichts weniger als das Überleben der gesamten Menschheit im Stile eines *Wir oder Sie* geht, führt zu den totalitären, letztlich faschistoiden Gedanken *Erewhons*, in denen die eugenische Sorge um die Vitalität des Volkskörpers einen Schatten auf die Sozialdarwinismen des 20. Jahrhundert wirft (s. 6).

Der Science-Fiction-Autor Arthur Clarke (1917–2008) behauptete, dass jegliche Technologie der fernen Zukunft für einen gegenwärtigen Beobachter kaum von Magie zu unterscheiden sei; Carl du Prel (1839–1899), der bedeutendste Theoretiker des deutschen Okkultismus im 19. Jahrhundert, schrieb, dass Magie unbekannte Naturwissenschaft sei.²⁸⁶ Die Transparenz der Wissenschaft für das Magische und Dämonische heben die frühen Automatenerzählungen in besonderer Weise hervor. Einerseits wirkt Bacons Optimismus dort fort, wo die KI-Theoretiker erster Stunde immer wieder betonen, dass nur das, was man wirklich lückenlos verstanden habe, programmiert werden könnte, und damit die Künstliche Intelligenz immer auch der Erweis einer vollen Durchdringung menschlicher Intelligenz bedeute.²⁸⁷ Andererseits verliert das Versprechen, dass die Wissenschaft mit der Rekonstruktion des Lebendigen ihre vollumfängliche Einsicht in den Menschen beweise, schon in der frühen Romanliteratur an Plausibilität: Hier tritt in dem Automaten dem Menschen gerade nicht mehr das Produkt voller Einsicht entgegen, sondern es wird ihm zu einem obskuren Wesen. Gleichsam wird die Integrität des Menschen immer fragwürdiger, der Apparat stellt Anfragen an das Apparatehafte des Menschen. Häufig ist es in diesen Erzählungen die weibliche Androidin, die den Mann lockt und täuscht, und durch die er – direkt oder indirekt – zugrunde geht. Die

285 Vgl. Gutmann, Knifka, Biomorphie and Technomorphic Metaphors – Some Arguments Why Robots do not Evolve, Why Computing is not Organic, and Why Adaptive Technologies Are not Intelligent, 67.

286 Vgl. Weber, Science Fiction, 107.

287 Vgl. Weizenbaum, Die Macht der Computer und die Ohnmacht der Vernunft, 151.

listige, ins Verderben führende Maschinenfrau findet ihre bekannteste Übersetzung in das Massenmedium Film spätestens mit der Roboterfrau in Fritz Langs *Metropolis* (1927): Die falsche Maria übernimmt die Signaturen ihrer literarischen Vorgängerinnen und gerät zum apokalyptischen Tier und *Hure Babylon*; in der Kanzelszene wird explizit die entsprechende Bibelstelle eingeblendet (Offb 17,3–6). Die falsche Maria ist eine täuschend echte Kopie der menschlichen, reinen und guten Maria. Sie, der böse Roboterzwilling, will alle Bewohner Metropolis ins Verderben stürzen. Die obere Gesellschaft verfällt ihr sexuell und bringt sich ihretwegen um; die Arbeiter hingegen stachelt sie zum Aufstand an, die in ihrer Raserei und blinden Zerstörungswut beinahe sich und ihre eigenen Kinder in der Unterstadt überfluten.²⁸⁸

Da die technische Utopie gerade unter der Prämisse spielt, allein durch die Wissenschaft ideale Zustände herbeizuführen, versteht sie sich als bessere Alternative zur Politik. Bereits *Nova Atlantis* ist in dem Sinne eine *Polis der Solution*, in der nicht mehr entschieden wird, sondern Probleme technisiert und dann gelöst werden. Wie Ernst Bloch anmerkte, scheint das Schicksal in Bacons Fragment technisch so zurückgedrängt, dass das Haus Salomonis mit ihm bereits fertig zu sein scheint, bevor der Staat Salomonis sich auch nur entwickelt.²⁸⁹ Die *neuen Väter* der Androiden und Künstlichen Intelligenz wie Moravec und Schmidhuber schweben in ihrer Außenwahrnehmung gleichwohl zwischen den vorgegebenen Bildern der Literatur von einer elitären Technokaste der Weltverbesserung einerseits, die sich außerhalb des Raumes politischer Verantwortlichkeit wähnt, und den tragischen Sonderlingen und Eigenbrötlern wie Dr. Frankenstein andererseits, die die Gefahr ihrer Forschung verkennen und etwas Grausames auf die Welt loslassen.

288 Vgl. Elke Dubbels, „Apokalypse der ‚Massen‘ in wissenschaftlichem Diskurs, Drama und Film (1892–1927)“, in: Olaf Briese/Richard Faber/et al. (Hgg.), *Aktualität des Apokalyptischen: zwischen Kulturkritik und Kulturversprechen.*, Würzburg 2015, 39–60. Hier 57–60.

289 Vgl. Ernst Bloch, *Das Prinzip Hoffnung*. In fünf Teilen. Kapitel 1–37, Frankfurt am Main 1959, 750–751.

4.2 Wissenschaft und Epistemologie: Die kybernetische Lehre vom Menschen als reaktivem System und das neue Regieren der Maschinen

Die moderne Kybernetik ist die Lehre von der Steuerung komplexer Systeme durch Nachrichtenverarbeitung. Zur Grundlage ihrer Erkenntnistheorie strukturiert sie die Welt in Regelprozesse der Informationsgenerierung, -verarbeitung und -speicherung. Nicht Masse oder Energie, sondern die Information und damit die Symboleigenschaft der Dinge wird zum ersten Baustein alles Seienden erklärt. Ihre Hochzeit erlebte die Kybernetik zwischen den Jahren 1950 und 1975. Wegweisend für das Aufkommen der Kybernetik als Disziplin waren die Arbeiten der Mathematiker Shannon und Weaver, die in ihrem 1948 erschienenen Artikel *A Mathematical Theory of Communication*²⁹⁰ soziale Kommunikation erstmals an den physikalischen Informationsbegriff anschlossen. Sie bestimmten, dass sich jede Nachricht auf ihren Neuigkeitswert reduzieren und damit in der Einheit der Information auch quantifizieren ließe. Auf Shannon und Weaver geht so die kleinste informationstheoretische Einheit BIT zurück.²⁹¹ Den eigentlichen theoretischen Überbau der Kybernetik jedoch schuf der amerikanische Mathematiker und Philosoph Norbert Wiener (1894–1964). Mit seinem ebenfalls 1948 erschienenen Werk *Cybernetics*²⁹² ebnete er das dogmatische Fundament für die Kybernetikbewegung der nächsten 25 Jahre. Der entscheidende Nebentitel seines Werks *Or Control and Communication in the Animal and the Machine* wies auf den universalen Anspruch seiner Theorie hin, nach der Biologie, Psychologie und Soziologie im informatischen Regelprozess überhaupt erst verständlich würden. Effektiv schleuste der kybernetische Universalanspruch

290 Vgl. Claude Elwood Shannon/Warren Weaver, *The Mathematical Theory of Communication*, Illinois 1964.

291 Vgl. Olaf Kaltenborn, *Das künstliche Leben*. Die Grundlagen der Dritten Kultur (Übergänge 45), München 2001, 23–24.

292 Vgl. Norbert Wiener, *Kybernetik*. Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine (Econ classics), Düsseldorf 1992.

technisch-mathematische Grundbegriffe wie Information, Steuerung und Kontrolle in die Human-, Sozial- und Geisteswissenschaften ein, in denen der Regelprozess zeitweise zu dem Paradigma schlechthin avancierte. Innerhalb der Kybernetik erfuhr das Verhältnis von Mensch und Maschine besondere Aufmerksamkeit. Den Menschen letztlich als komplexe, biologische Maschine zu begreifen – darin reüssierte in der Kybernetik eine Idee, wie sie im Raum der modernen Wissenschaften schon zu Beginn der Aufklärung aufkam und nun in ungekannter Klarheit ihre Durchsetzung beanspruchte.

Die Erkenntnisse der Kybernetik sollten in der Praxis zu einer neuen Qualität autonomer Maschinen führen. Ihre praktischen Vorläufer finden sich in der militärischen Forschung zu ersten autonomen Waffensystemen im Umfeld der beiden Weltkriege. So entwickelte der amerikanische Ingenieur Benjamin Miessner (1890–1976) bereits 1916 in seinen Arbeiten zur Funkfernsteuerung von Torpedos sein Modell eines elektrischen Hundes. Diesen hatte er mit zwei lichtempfindlichen Selenzellen ausgestattet, sodass sich dieser selbstständig auf Lichtquellen zubewegen konnte. Auf diesem Prototypen, den Miessner eigentlich noch zu einem kriegsfähigen *War Dog* weiterentwickeln wollte, bauten in den 1950er Jahren William Walters *Cybernetic Turtles* auf, die mit einer Fotozelle, einem Berührungssensor und zwei Vakuumröhren versehen die Bewegung echter Tiere simulierten.²⁹³ Hans Moravec wird noch am Ende des 20. Jahrhunderts betonen, dass es der Technisierung der großen Weltkriege zu verdanken ist, dass durch sie überhaupt die Frage, in welcher Weise Maschinen sich wie Tiere oder Menschen verhalten könnten, in den Blickkreis der Wissenschaften gekommen war.²⁹⁴

Gleichwohl die historischen Stränge der Kybernetik unterschiedlich tief zurückreichen, markieren die Macy-Konferenzen in New York den Start der internationalen Bewegung. Zwischen 1946 und 1948 noch unter dem umständlichen Titel *Circular Causal, and Feedback Mechanisms in Biological and Social Systems* abgehalten, hießen sie ab 1949 unter dem Eindruck Wieners nur noch *Cybernetics*.²⁹⁵ Schon die erste Konferenz

293 Vgl. Dittmann, Mensch und Roboter – ein ungleiches Paar., 26.

294 Vgl. Moravec, Computer übernehmen die Macht, 34.

295 Vgl. Claus Pias, „Zeit der Kybernetik – Eine Einstimmung“, in: ders. (Hg.), *Cybernetics/Kybernetik. The Macy-Conferences 1946–1953*, Berlin 2004, 9–41. Hier 9.

1946 fasste ihren Anlass derart zusammen, dass man verschiedenste Disziplinen sich aufeinander zubewegen – oder: konvergieren – sah, die nach einer neuen, gemeinsamen Theorie verlangten. Darunter zählte man die jüngsten Entwicklungen der Informatik, der Neurophysiologie und ein humanwissenschaftliches Forschungsfeld zwischen Psychiatrie, Anthropologie und Soziologie. Es wurden gegenseitig Ideen vorgestellt, von denen man ahnte, sie würden ein größeres, theoretisches Ganzes greifen. Die Konferenzen und ihre Protokolle sind so Zeugnisse einer über die Fachgrenzen empfundenen Aufbruchsstimmung, deren inhaltliche Ausgestaltung doch je disziplinar blieb.²⁹⁶ Dass sich diese Euphorie über die nächsten zwei Jahrzehnte global ausbreitete, lag am Versprechen der Kybernetik, der krisengeschüttelten Nachkriegsgesellschaft neue Lösungen in Aussicht zu stellen: Indem sie alle Wissenschaft mathematisch zu formalisieren suchte, versprach sie eine umfassende Entideologisierung aller Wissenszweige, besonders der sozialdarwinistisch geprägten Humanwissenschaften. Zugleich wurde darin auch der entscheidende Schritt zu einer neuen Synthese der Wissenschaften gesehen, die davor nur eine immer ausgreifende Differenzierung kannte. In ihrem Anspruch, den kartesischen Dualismus abzulösen, arbeitete sie auf eine neue, einheitliche Wissenschaftslehre hin. Abseits der Theorie wollte die Kybernetik auch praktisch die Hochtechnisierung der Kriegsnationen sinnvoll weiternutzen, um durch die technologische Transformation das Gelingen der Gesellschaft auf eine neue Basis zu stellen. Die technologische Reorganisation sollte die Menschen dabei gleichsam einhegen; zu stark hielt sich in Politik und Wissenschaft der Vorbehalt, dass man den Menschen nach den Erfahrungen aus den Totalitarismen nicht mehr vollends vertrauen konnte:²⁹⁷

„Nach der vollständigen Desavourierung von Eugenik und Rassenhygiene, Höherzüchtung und Rassenreinheit, allesamt organistische Kate-

²⁹⁶ Vgl. ebd., 12–13.

²⁹⁷ Vgl. Erich Hörl/Michael Hagner, „Überlegungen zur kybernetischen Transformation des Humanen“, in: dies. (Hgg.), *Transformation des Humanen. Beiträge zur Kulturgeschichte der Kybernetik.*, Frankfurt am Main 2008, 7–37. Hier 9.

gorien der Anthropologie, hatte eine technizistische Zukunftsvision zunächst einmal nichts Befremdliches.“²⁹⁸

Gewiss erfuhr die Kybernetik mit ihrem universalistischen Anspruch von Anfang an nicht nur begeisterte Zustimmung, gleichzeitig blieb aber die große Konfrontation gegen das Vorhaben von wissenschaftlicher Seite aus.²⁹⁹ Erst Mitte der 1970er Jahre kam die erste Begeisterungswelle für eine solche technizistische Zukunft zum Erliegen. Politischer Unwillen und ausbleibende Institutionalisierung, Selbstüberschätzung als auch die disziplinen eigene Ausdifferenzierung brachten die Kybernetik schließlich zu Fall. Es ist kein Zufall, dass die Zeit der Kybernetik mit dem Aufblühen der Forschung um Künstliche Intelligenz übereinstimmt. Vieles, dass heute unter dem Begriff der Künstlichen Intelligenz gesammelt wird, verhandelte man damals unter der Kybernetik. Ihr gemeinsames Vorhaben, Mensch und Natur durch die Technik zu versöhnen, wich Mitte der 1970er der alternativen Idee, dass Mensch und Natur sich gegen die Technik verbünden müssten.³⁰⁰ Auch wenn die Kybernetik als historische Bewegung zu ihrem vorzeitigen Ende kam, blieben ihre tragenden anthropologischen Themen unerledigt. Die Kybernetik ist somit ein historisches Ereignis, das über die Wendungen der Nachkriegszeit in wissenschaftlicher, sozialer wie kultureller Weise Auskunft gibt, und ist andererseits eine Weltanschauung, die in ihrem Versprechen, die Welt anhand von Information suffizient strukturieren und damit kontrollieren zu können, nichts an Aktualität eingebüßt hat.³⁰¹

298 Michael Hagner, „Vom Aufstieg und Fall der Kybernetik als Universalwissenschaft“, in: Michael Hagner/Erich Hörl (Hgg.), *Transformation des Humanen. Beiträge zur Kulturgeschichte der Kybernetik*, Frankfurt am Main 2008, 38–71. Hier 52.

299 Vgl. ebd., 39.

300 Vgl. ebd., 70.

301 Vgl. Hörl, Hagner, Überlegungen zur kybernetischen Transformation des Humanen, 7.

4.2.1 Das politische Versprechen der Kybernetik: Maschinen regieren über Menschen

Die Kybernetik war von Beginn an eine politische Technologie. Die Gruppe um Norbert Wiener hat dies mit ihrer formalisierten Sprache eher verschleiert, als sie den Begriff für die Nachkriegszeit konzeptuell stark machte. Deutlich wurde dies noch beim deutschen Vater der Kybernetik, dem Ingenieur Hermann Schmidt (1894–1968). Mit seiner Regelungstechnik war er sich Anfang der 1940er Jahre sicher, die Grundlagen eines großdeutschen Sozialstaates zu schaffen, in dem der Fabrikarbeiter aus seinem Sklavendienst an den Maschinen entlassen wird:

„Mitten im zukunftsgewissen Erleben des Entstehens des großdeutschen Sozialstaates erhält der Techniker in der Förderung der Regelungstechnik als auch einen sozialpolitischen Auftrag höchster Verantwortung.“³⁰²

Sein Verständnis von der Regelungstechnik als einer staatsbildenden Technik stand im Sinne einer *τέχνη* im Einklang mit ihren Wurzeln in der antiken politischen Philosophie: So bezeichnete bereits der griechische Dichter Pindar mit der Kybernetik die göttliche Weltregierung. Der Apostel Paulus rechnete die Gemeindeleitung, die *κυβέρνησις*, zu den göttlichen Charismen (1. Kor 12,28) und dem Kirchenvater Hippolyt von Rom ist Christus ein Kybernet und damit Steuermann, der die gesamte Christenheit durch die Gewässer der Zeit lenkt. Von Platons *Politeia* und Aristoteles *Politika* bis zu Thomas von Aquin, der für die gute Regierung die latinisierte Form *gubernatio* verwendet, wurde die Kybernetik als politische Fertigkeit in die Neuzeit tradiert.³⁰³ Ihre moderne Gestalt idealisierte nicht mehr die göttliche Regierung über die Welt noch die philosophische Kunst des klugen Regierens durch den Menschen, sondern das unpersönlich-logische Regieren der Apparate.

³⁰² Hermann Schmidt, „Regelungstechnik. Die technische Aufgabe und ihre wirtschaftliche, sozialpolitische und kulturpolitische Auswirkung“, *Zeitschrift des Vereins deutscher Ingenieure* (1941), 81–88.

³⁰³ Vgl. Claus Pias, „Unruhe und Steuerung. Zum utopischen Potential der Kybernetik“, in: Jörn Rüsen (Hg.), *Die Unruhe der Kultur. Potentiale des Utopischen*, Weilerswist 2004, 301–325. Hier 312.

Dieser politische Grundimpuls der Kybernetik trat in ihrem moder-
nen Ableger wieder deutlich zutage, als der in die USA emigrierte Poli-
tikwissenschaftler Karl W. Deutsch (1912–1992) 1963 mit *The Nerves of
Government*³⁰⁴ das Programm einer spezifisch politischen Kybernetik
veröffentlichte. Der beispiellose Rüstungswettlauf seit Ende des Korea-
krieges und das knappe Abwenden eines dritten globalen Krieges in der
Kubakrise im Oktober 1962 hatten zuvor deutlich gezeigt, wie schmal
der Grat für politische Entscheidungen war. Fehleinschätzungen der
Lage wie eine gestörte Kommunikation konnten schnell katastrophale
Folgen nach sich ziehen, sodass Regierungsapparate – Netzwerke zur
politischen Entscheidung und Steuerung – aus den aktuellen politi-
schen Erfahrungen heraus attraktiv schienen.³⁰⁵ Auch in der Bundes-
republik wurden solche Ideen von einem computergestützten Regieren
rege aufgenommen. Seit 1961 diskutierten hochrangige internationale
Politiker und wissenschaftliche Experten in den Bergedorfer Gesprächs-
kreisen in Hamburg Grundfragen deutscher und europäischer Außen-
und Sicherheitspolitik. Der neunte Gesprächskreis fand am 25. Februar
1963 unter dem Titel *Mensch – Denkmaschine – Staatsmaschine* statt.
Den Aufschlagsvortrag hielt der französische Politiker Pierre Bertaux
(1907–1986). Sein Vortrag war ein Auszug seines noch in demselben
Jahr erschienenen Buches *Mutation der Menschheit*³⁰⁶: Unter explizitem
Bezug auf das Denken des Prozessphilosophen Teilhard de Chardin
(s. 4.3) deutete er wie viele Intellektuelle seiner Zeit das exponentiell
erfahrene Fortschreiten der Technik als Anzeichen für einen nahen-
den Entwicklungssprung des Menschen. Für Bertaux überschritten
Menschen Mitte des 20. Jahrhunderts die Schwelle von der paläo-
technischen hin zur neotechnischen Kulturform. Von diesem Schwel-

304 Vgl. Karl W. Deutsch, *The Nerves of Government. Models of political Communication and Control*, 2. Aufl., London 1963.

305 Vgl. Jürgen P. Lang, „Karl W. Deutsch, *The Nerves of Government. Models of Political Communication and Control*, New York 1963“, in: Steffen Kailitz (Hg.), *Schlüsselwerke der Politikwissenschaft.*, Wiesbaden 2007, 92–95. Hier 92.

306 „Die Menschheit nimmt immer mehr die Gestalt eines weltweiten Netzes von Kommunikationen, Kontakten und Informationen an [...]. Damit stehen wir am Anfang einer ‚globalen Ära‘, der ‚Planetarisierung der Gattung‘, um mit Teilhard de Chardin zu sprechen.“ Pierre Bertaux, *Mutation der Menschheit. Zukunft und Lebensinn*, Frankfurt am Main 1963, 70–71.

lenpunkt der Zeit aus sei bereits auszumachen, dass nicht mehr der Mensch, sondern der Apparat die neue dominante Lebensform des kommenden Zeitalters sei. Diesen stellt sich Bertaux als Symbiose aus Mensch, Maschine und einem nichtmateriellen Element wie Kultur oder Methode vor, welches die Symbiose erst lebensfähig macht.³⁰⁷ Für diese Mensch-Maschine-Symbionten werden laut Bertaux die Staatsmaschinen wie selbstverständlich zur guten Regierung gehören:

„Für uns paläotechnische Menschen wirkt es geradezu beängstigend, wenn man von Staatsmaschinen spricht. Die Idee, von Maschinen regiert zu werden, ist erschreckend. Wo bleibt die Humanität, die Menschlichkeit? Ich glaube, wir sollten damit beginnen, unsere bisherigen Regierungsmethoden in Bezug auf Menschlichkeit einmal kritisch anzusehen. Diese ‚erprobten‘ Regierungsmethoden haben innerhalb eines halben Jahrhunderts zwei Weltkriege und unzählige Kriege bescheideneren Ausmaßes mit mehr als fünfzig Millionen Toten nicht vermeiden können. [...] Im strategischen Spiel zwischen den beiden großen Partnern sagen die Denkmaschinen jedem der Gegner ganz unparteiisch, was möglich ist und was nicht.“³⁰⁸

Die Kunst des Regierens bestimmt Bertaux zur Kunst der Voraussicht. Den Blick in die Zukunft erhalte man jedoch nur, wenn man alle Faktoren der Gegenwart schlüssig miteinander in Beziehung setze. Da das organische Denken jedoch die zahllosen Elemente, die auf das Geschehen einwirken, nicht verarbeiten könne, hält er den Menschen von Natur aus für zukunftsblind und damit implizit der Politik unfähig. Dem könne seiner Ansicht nach allein die Maschine mit ihrer Fähigkeit der *Prospektive* abhelfen.³⁰⁹

307 Vgl. Pierre Bertaux/Arnold Gehlen, *Maschine. Denkmaschine, Staatsmaschine*. Entwicklungstendenzen der modernen Industriegesellschaft (Bergedorfer Gesprächskreis zu Fragen der freien industriellen Gesellschaft / Bergedorfer Protokolle Band 2), Hamburg/Berlin 1963, 9–13.

308 Ebd., 20.

309 Ebd., 22–23. Bertauxs Vortrag schlug im Gesprächskreis in Hamburg massive Kritik entgegen. Die anwesenden Informatiker warnten davor, dass diese Umbruchsvisionen nichts mit ihrer täglichen Arbeit an den Maschinen und deren tatsächlicher Leistungsfähigkeit zu tun habe.

Solche prospektiven Regierungsapparate, wie sie Bertaux unmittelbar vor sich wählte, wurden zumindest kurzzeitig Wirklichkeit: Die ehrgeizigste Verwirklichung eines kybernetischen Staates geschah in Chile Anfang der 1970er unter der Präsidentschaft Salvador Allendes (1908–1973). Unter der Leitung des amerikanischen Kybernetikers Stafford Beer (1926–2002) unter der Aufsicht der CIA begann im November 1971 die Entwicklung eines Computersystems namens *Cybersyn*, in das die Entscheidungen der wirtschaftlichen Zukunft Chiles verlegt werden sollten. Die gesamte Wirtschaft sollte auf eine verschachtelte und rekursive Struktur passender Systeme und Subsysteme umgestellt werden, die sich von der Staatsebene zum Wirtschaftssystem, weiter zu den Industriesektoren, den einzelnen Firmen, ihren Arbeitsbereichen bis hin zum neurophysiologischen System der Arbeiter vernetzte. Das Zeitalter der Statistik sollte dem Zeitalter der Echtzeitsteuerung Platz machen und so zu allgemeinem Wohlstand, Wachstum und Frieden führen. Der Gefahr einer menschenverlassenen Regierung schaltete man die *Faculty of Foresight* als Residuum menschlicher Entscheidungen vor. Hier sollten Politiker, Manager und Arbeiter erstrebenswerte Zukünfte entwerfen, die mit der Hilfe von *Sybersync* realisiert werden sollten.³¹⁰ Einen Monat war das System in Betrieb, ehe das Vorhaben mit der Ermordung Allendes zerbrach. Dabei hatte Beer Großes vor und plante, nach der Wirtschaft auch die Politik zu kybernetisieren: Über einen *Feedback*-Kanal am heimischen Fernseher sollten sämtliche Entscheidungen durch eine Regelungsschleife zwischen Regierenden und Regierten laufen, deren Sollwert ein glückliches Volk ist. Der Wegfall jedweder politischen Repräsentation hin zu einer einzigen Oberfläche, deren einziges Selektionsmerkmal die Differenz von *happy/unhappy* sein sollte – das waren für Beer die schlichtweg notwendigen Anpassungen, um die Politik an die Effektivität und das Tempo einer effektiven Organisation anzuschließen, wie sie nur die Kybernetik ermöglichte.³¹¹

310 Vgl. Pias, *Unruhe und Steuerung. Zum utopischen Potential der Kybernetik*, 313–325.

311 Vgl. Pias, *Zeit der Kybernetik – Eine Einstimmung*, 34.

4.2.2 Der Mensch, eine biologische Maschine

Die Figur Walter Faber aus Max Frischs (1911–1991) 1957 erschienenen Roman *Homo faber* ist überzeugt, dass sich das Leben nach den Gesetzen der Logik und Wissenschaft organisieren lasse und der Mensch sehr gut als Maschine funktioniere. In einem seiner Dialoge erklärt er einige Grundzüge der Kybernetik in direktem Bezug auf Wieners *Cybernetics* und amüsiert sich dabei über das „abgedroschene Argument: der Mensch sei keine Maschine“³¹². Dass sich der Erkenntniswert der Kybernetik überhaupt erst dann einstellt, wenn er auch den Menschen als informationsverarbeitendes System und damit biologische Maschine miteinschließt, dies hielt Norbert Wiener für grundlegend.³¹³ Nur wenn das menschliche Wissen mit dem Wissen von Computern kompatibel gemacht werden kann, sie in gleicherweise *digital operieren*, ist auch die kybernetische Epistemologie selbst arbeitsfähig. Ein Zeugnis davon geben alle zehn Macy-Konferenzen, die immer wieder um die Bestimmung und Reichweite der Begriffe *analog* und *digital* kreisten.³¹⁴

Die kybernetische Anthropologie ist wiederum nicht vollends neu im Raum der Wissenschaften, sondern knüpft an Theorien und Bestrebungen aus den Anfängen der neuzeitlichen Wissenschaft an. Bereits den Rationalisten des 16. und 17. Jahrhunderts zeigte sich nicht nur der Mensch, sondern vielmehr die gesamte Welt als ein deterministisches System nach strengen, mathematisch beschreibbaren Gesetzmäßigkeiten. Daran schlossen die Hoffnungen seit Bacon an, im Erkennen dieser Gesetzmäßigkeiten Welt und Mensch vollständig zu beherrschen. Die Kybernetiker sehen sich in ihrem Denken, die Gesetzmäßigkeit und damit das Maschinenhafte des Menschen und seines Denkens aufzuzeigen, in der Sukzession zu Gottfried Wilhelm Leibniz (1646–1716), weil er der erste gewesen sei, der die formale Dimension des Denkens in eine formale Sprache zu übersetzen suchte. Wiener selbst beschreibt Leibniz Philosophie als den entscheidenden intellektuellen

312 Max Frisch, *Homo Faber*. Ein Bericht, Frankfurt a.M. 1977, 74.

313 Vgl. Wiener, *Kybernetik*, 182. Wiener sieht im Gehirn unter normalen Umständen nicht das vollkommene Analogon zur Rechenmaschine, sondern eher das Analogon eines einzelnen Rechenlaufes solcher Maschinen.

314 Vgl. Pias, *Zeit der Kybernetik – Eine Einstimmung*, 14.

Impuls für die Kybernetik. Leibniz habe seiner Ansicht nach in der Entwicklung der mathematischen Logik gleichzeitig die Voraussetzungen zur Mechanisierung des Denkens geschaffen. Genau wie das Kalkül der Arithmetik eine fortschreitende Mechanisierung durchlaufen habe – vom Rechenschieber und der Tischrechenmaschine bis zum Rechenautomaten – enthalte Leibniz Idee des *calculus ratiocinator* die ideellen Keime einer *machina ratiocinatrix*, einer logischen Maschine, wie sie laut Wiener die Kybernetik hervorzubringen beabsichtige.³¹⁵

Schon zur Mitte des 18. Jahrhunderts erreichten sowohl uhrwerksgetriebene Automaten, die voltigieren, schreiben oder musizieren konnten, als auch ein anthropologischer Diskurs, der den Menschen zur Maschine erklärte, große Aufmerksamkeit. Der französische Arzt und Philosoph Julien Offray de La Mettrie (1709–1751) nimmt in seinem 1748 erschienenen Essay *L'homme machine* beispielsweise nicht allein die Uhrwerksmetapher auf, die bei Thomas Hobbes und René Descartes im 17. Jahrhundert noch bildlich für den Staat oder den Menschen einstand, sondern erklärt den Menschen gleich den menschenförmigen Automaten direkt zu einem mechanischen System:³¹⁶

„Der Mensch ist eine Maschine, welche so zusammengesetzt ist, dass es unmöglich ist, sich zunächst von ihr eine deutliche Vorstellung zu machen und folglich sie zu definieren. Deshalb sind alle Untersuchungen theoretischer Natur, welche die größten Philosophen angestellt haben, das heisst, indem sie gewissermaßen auf den Flügeln des Geistes vorzugehen versuchten, vergeblich gewesen. Also kann man nur practisch, oder durch einen Versuch der Zergliederung der Seele, nach Art der Aufklärung über die körperlichen Organe, ich will nicht sagen mit Sicherheit die Natur des Menschen enträthseln, aber doch wenigstens den möglichst höchsten Grad von Wahrscheinlichkeit über diesen Gegenstand erreichen.“³¹⁷

315 Vgl. Wiener, *Kybernetik*, 40.

316 Vgl. Wittig, *Maschinenmenschen*, 45.

317 Julien Offray de La Mettrie, *L'homme machine*. Der Mensch eine Maschine (Philosophische Bibliothek 67), Leipzig 1875, 21.

Als Aufklärer seiner Zeit will er die Anthropologie von der Philosophie und ihren Vorannahmen über den Menschen lösen und stattdessen den Empirikern überantworten. Insbesondere die Theologie habe den Untersuchungsgegenstand „Mensch“ mit falschen Annahmen so sehr verformt, dass laut La Mettrie die übrigen Wissenschaften ihr jedwedes Rederecht über diese Frage absprechen sollten. Ihre Rede von der Seele erklärt La Mettrie zu einem nichtigen Ausdruck. Descartes und seinen Anhängern wirft er vor, sie nähmen zwei Substanzen in dem Menschen an, wofür sie keine empirische Beweislast vorlegten. Sie müssten hingegen anerkennen, dass alle Fähigkeiten der Seele von der Beschaffenheit des Gehirns und des gesamten Körpers abhingen, und die Seele deshalb nichts weiter als der Ausdruck ihrer Organisation darstelle. Die Organisation des Stofflichen bedeute damit keinen qualitativen Unterschied mehr, sondern nur noch einen graduellen Unterschied zwischen belebtem und unbelebtem Stoff. In diesem Sinne erschienen Menschen, Tiere und Pflanzen von einer einzigen Wesenheit, nur unterschieden in ihrem Komplexitätsgrad. Der Mensch trete so aufgrund seiner Komplexität immerhin noch als „eine sehr erleuchtete Maschine“³¹⁸ hervor.

„Behaupten wir also dreist, dass der Mensch eine Maschine ist und dass es in der ganzen Welt nur eine einzige verschieden geartete Wesenheit giebt.“³¹⁹

Doch nicht nur in der Wissenschaftsgeschichte der Aufklärung, auch in ihrer Gegenwart fand die Kybernetik wissenschaftliche Rückendeckung in ihrem Bestreben, den Menschen in ein komplexes Informationssystem aufzulösen. Ihren größten Partner fand sie in dem Behaviorismus, der wiederum mit dem Informatiker Alan Turing als glühendem Behavioristen eng an die Anfänge der KI-Bewegung geknüpft ist.³²⁰ Der Behaviorismus stiftete der Kybernetik die theoretische Brücke in die Humanwissenschaften: Seit ihrer Gründung im Jahr 1923 durch John Broadus Watson (1878–1958) will die behavioristische Schule die Psy-

318 Ebd., 58–59.

319 Ebd., 80.

320 Vgl. Turing, *Computing Machinery and Intelligence*, 442f.

chologie in eine exakte Naturwissenschaft umformen. Dazu erforscht sie das menschliche Verhalten aus einer strikt mechanistischen Perspektive der Reizausübung und Reizreaktionen. Sie erklärt das Mentale für wissenschaftlich wie ontologisch irrelevant. Innerlichkeit wird zum bloßen Epiphänomen stofflicher Abläufe entwertet, wie Sinnesempfindungen zu informationellen Reizen umgedeutet werden und als bloßes Symbol ein Eigenrecht genommen wird. Indem menschliches Verhalten auf reines Reagieren auf äußere Einflüsse reduziert wird, bagatellisiert der Behaviorismus schlussendlich die Ich-Erfahrung zur Illusion der Gehirnmaschine. Verhalten sei demnach ein fortlaufendes Anpassungsgeschehen an die Umwelt, sodass die Rede von einer Persönlichkeit für den Behavioristen nichts weiter als ein Bündel konditionierter Reflexe beinhaltet.³²¹ Der Weg von dem Behaviorismus hin zur Kybernetik bestand lediglich darin, diesen Austausch von Reizen zwischen dem *System* Mensch und dem *Metasystem* Umwelt als Informationsaustausch und Regelkreis zu interpretieren.

Wie die kybernetische Anthropologie mit dem Aufschwung der Künstlichen Intelligenz im öffentlichen Raum zurückkehrte, gilt dies seit den 2000ern gleichsam für die behavioristische Schule unter dem Leitbegriff der *Philosophy of Mind*. In Deutschland vertritt der Philosoph Thomas Metzinger öffentlichkeitswirksam sein behavioristisches Programm, nach dem der Mensch in Wirklichkeit ein *Nobody* sei, der bloß unter ständigen Selbst- und Weltmodellen operiere:

„No such things as selves or subjects of experience exist in the world. What exist are natural systems [...], with both of these representational structures coming in many different strengths and with a long evolutionary history. We can therefore greatly simplify the ontological set of background assumptions necessary to do proper scientific psychology and cognitive neuroscience. All that exists are phenomenal selves, as instantiated by transparent self-models.“³²²

321 Wittig, *Maschinenmenschen*, 94–96.

322 Thomas Metzinger, *Being no One. The Self-Model Theory of Subjectivity* (A Bradford book), 1. Aufl., Cambridge, Mass. 2003, 577.

Der Sozialwissenschaftler Thomas Bächle sieht die modernen Neurowissenschaften und ihre bildgebenden Verfahren mit in der Schuld, die Idee vom Gehirn als Maschine beständig zu implizieren und in weite Bereiche der Lebenswissenschaften zu streuen. Seit der Erfindung der Magnetresonanztomographie Anfang der 1970er werde der Mythos von der Mensch-Maschine-Gleichung zur sichtbaren Wahrheit erhoben, indem der Gehirnsan das Computer-Modell des Gehirns essentialisiere. Die Ästhetik und die Medialität der Computeraufnahmen ließen den eigenen Körper noch mehr als Computergraphik erleben, wodurch die Auflösung des Körpers in Technologie noch plausibler erscheine. Das Gehirn – abstrakt modelliert als digitale Recheneinheit – werde als solche scheinbar sichtbar gemacht, ohne die reduktionistische Modellhaftigkeit des Dargestellten angemessen darzulegen. Am Gehirnsan erweise sich so beispielhaft, wie über die Kybernetik und die Psychologie hinaus bis heute in vielen angewandten Bereichen der Naturwissenschaft eine reduktionistisches Gehirn-Computer-Analogie weitergetragen werde.³²³

4.2.3 Kybernetische Theologie

In den 1960er Jahren traten einzelne Stimmen auf, die die Erkenntnisse der Kybernetik für die Theologie fruchtbar machen und sie an die neue Universalwissenschaft anschließen wollten. So machte Hans-Dieter Bastian in seinen *Problemanzeigen einer kybernetischen Theologie* aufmerksam, dass Theologen viel gewinnen, wenn sie die religiöse Verkündigung als Informationsgeschehen begriffen. Die Analyse der Informationswege der Kirche anhand der kybernetischen Semiotik von Syntax (Zeichen), Semantik (Bedeutung) und Pragmatik (Wirkung) könne helfen, effektivere Informationsstrukturen aufzubauen, um die Wirksamkeit der Verkündigung nicht mehr nur dem Geist Gottes zu verantworten, sondern die Kirche neben der semantischen dadurch auch die pragmatische Verantwortung für die kirchliche Rede zu eröffnen.³²⁴

323 Vgl. Bächle, Mythos Algorithmus. Die Fabrikation des computerisierbaren Menschen, 284.

324 Vgl. Hans-Dieter Bastian, „Problemanzeigen einer kybernetischen Theologie“, *Evangelische Theologie* 28 (1968), 334–344. Hier 336–337.

Nicht nur das Verhältnis der Kirche zu ihren Mitgliedern, auch das Zusammenwirken der theologischen Disziplinen untereinander sollte als Regelkreis analysiert und optimiert werden. Die Hoffnung wurde geäußert, dass die Theologie in der Weise ihre jahrhundertalte Vernachlässigung der Naturwissenschaften wiedergutmachen und mit ihrer Kybernetisierung ein attraktiver Gesprächspartner werden könnte.³²⁵

Neben der Überlegung, die Theologie disziplinär an die Kybernetik anzuschließen, wurden auch die weltanschaulichen Setzungen der Kybernetik wahrgenommen und versucht, mit den christlichen Traditionsbeständen ins Gespräch zu bringen. Es war nicht zuletzt der geistige Vater der Kybernetik selbst, Norbert Wiener, der theologischen Gesprächsbedarf an seinen Theorien feststellte und in seinem letzten Werk *God and Golem, Inc.* von 1964 aufarbeitete. Darin geht er zuallererst auf den Vorwurf der Blasphemie ein, wie ihn 1957 der deutsche Philosoph Gotthard Günther (1900–1984) in seiner populären Abhandlung *Das Bewusstsein der Maschinen* in der Entwicklung selbstregulierender Maschinen erkannt hatte. Günther war zu dem Schluss gekommen, dass der Mensch mit der Konstruktion der kybernetischen Maschine Blasphemie begehe. Denn wenn der Mensch eine Analogie Gottes und der Roboter eine Analogie des Menschen sei, dann gerate der Roboter zu einer Analogie Gottes über den Menschen.³²⁶ Wiener begegnet diesem Vorwurf defensiv wie offensiv: Zwar gesteht er ein, dass die Gleichstellung von Mensch und Maschine die größte theologische Herausforderung sei, die die Kybernetik der Religion stelle, entgegnet aber, dass sich historisch die anfängliche Blasphemie stets als einziger Weg zur Erkenntnis erwiesen habe. Wie vor Darwin der Mensch und das Tier nicht hätten in einem Atemzug genannt werden dürfen, so richte er sich nun gegen die Vorurteile jener Wissenschaftler, die vor aller Argumentation die Auseinandersetzung um die Ähnlichkeit von Mensch und Maschine scheuten – und dass möglicherweise gerade aus religiöser Motivation heraus.³²⁷ Offensiv wendet er dagegen ein, dass

325 Vgl. ebd., 343–344.

326 Gotthard Günther, *Das Bewußtsein der Maschinen*. Eine Metaphysik der Kybernetik (Kybernetik und Information Bd. 3), 3. Aufl., Baden-Baden 2002, 95.

327 Vgl. Norbert Wiener, *God and Golem, Inc.* A Comment on certain Points where Cybernetics impinges on Religion, Cambridge, Mass. 1964, 4–5.

die kybernetischen Maschinen tatsächlich den Grad einer göttlichen Schöpferkraft erreicht hätten. Seiner Ansicht nach würden Maschinen schöpferisch, weil sie mehr und mehr operative Abbilder des Menschen hervorbrächten, also Systeme, die nicht ihrem Aussehen nach, sondern – und darauf käme es eigentlich an – funktionale Abbilder des Menschen sind. Wie die Erweckung Galateas sie im antiken Mythos von einem bloßen *pictorial image* eines Menschen in ein *operative image* verwandelt habe, bestehe die sehr viel bedeutendere Ähnlichkeit dieser neuen Maschinen zu Menschen darin, dass sie ihn in seinen Handlungen ersetzen könnten. Die Zeugungskraft der Kybernetik habe daher sehr wohl etwas Schöpferisches an sich, sodass Wiener die Gottesanalogie für nicht völlig abwegig hält.³²⁸

Ein anderer theologischer Anknüpfungspunkt betrifft das künftige Verhältnis der Menschen zu den immer selbstständigeren Maschinen. Wiener zieht hier die Analogie zur Rahmenhandlung der biblischen Hiobserzählung, in der Gott mit dem Satan eine Wette über den Menschen eingeht. Der Satan tritt in diesem Spiel als ein Geschöpf Gottes auf. Wiener will darauf hinweisen, dass ein Schöpfer (ob nun Gott oder Mensch) mit seinem Geschöpf – in diesem Fall die Maschine – nur dann ein Spiel spielen könne, wenn er zuvor einen Teil seiner Souveränität an sein Geschöpf abgetreten habe. Nur wenn der Schöpfer nicht mehr vollends über sein Geschöpf verfügt und somit auch die Möglichkeit besteht, das Spiel gegen das eigens geschaffene Gegenüber zu verlieren, lasse sich überhaupt zwischen Schöpfer und Geschöpf spielen:

„Can God play a significant game with his own creature? Can any creator, even a limited one, play a significant game with his own creature? In constructing machines with which he plays games, the inventor has arrogated to himself the function of a limited creator, whatever the nature of the game-playing device that he has constructed. This is in particular true in the case of game-playing machines that learn by experience.“³²⁹

328 Vgl. ebd., 30–31.

329 Ebd., 16–17.

Der deutsche Theologe Hans Reinhardt Rapp nahm Wieners Nachdenken auf. Auch er kommt zu dem Schluss, dass im Zeitalter der lernenden und sich selbst reproduzierenden Automaten aus den neuen Beziehungen, die Menschen gegenüber ihren Maschinen pflegen, gleichsam ein anderes Verhältnis zu Gott ersichtlich werde. Die neue Metapher besage, dass Gott und sein Geschöpf in einer Art Funktionsgemeinschaft miteinander stehen, sie *incorporated* sind wie zwei vertraglich gebundene Partner. Der Mensch werde zu einem Partner, dem sein Schöpfer einen Teil seiner eigenen Freiheit überantwortet und mit dem er sich auseinanderzusetzen habe. Damit werde das Geschöpf auch zu einer Herausforderung für die Intelligenz des Schöpfers. Die Auseinandersetzung mit intelligenten Maschinen habe somit den Menschen zur Einsicht in die Verantwortung des Schöpfers vor seinem Geschöpf geführt.³³⁰ Rapp bestimmt in diesem Zusammenhang das Christentum überhaupt zur kybernetischen Religion, weil es den Schritt aus der *Festwertregelung* des jüdischen Ritualgesetzes in die Folgeregelung gewagt habe. Gegenüber dem gesetzlichen System mit starrer Führungsgröße habe das junge Christentum, allen voran Paulus, stattdessen die *Folgeregelung* eingeführt, in der sich die Führungsgröße aus der Gegenwart Christi im Geist ergab. Erst hierdurch wurde die Religion laut Rapp zu einem adaptiven System mit der Möglichkeit zur weltweiten Mission.³³¹ Beil aller Euphorie für das kybernetische Projekt äußert Rapp aber auch Zweifel an dem Einigungsversprechen der Kybernetik, welches nicht allein die Wissenschaften, sondern alle Bereiche des Lebens zu harmonisieren suche. Dem kybernetischen Versprechen einer Allversöhnung zwischen Natur- und Geisteswissenschaft, Religion und Rationalität etc. hält Rapp die Einsicht der Theologie entgegen, dass eine solche erlösende Einheit nicht durch den Menschen hergestellt werden könne.³³²

Wiener kommt in *God and Golem Inc.* schließlich auf den besonderen Wert der Ethik für die Kybernetik zu sprechen. Dem Umgang mit der neuen Technologie attestiert Wiener eine besondere Form der Heiligkeit, aus der sich eine besondere Verantwortung ergibt: Man

330 Vgl. Hans Reinhardt Rapp, *Mensch, Gott und Zahl*. Kybernetik im Horizont d. Theologie (Stundenbücher 98), 2. Aufl., Hamburg 1971, 262–263.

331 Vgl. ebd., 99.

332 Vgl. ebd., 224–245.

wisse zwar im aufgeklärten Bewusstsein, dass die selbstregulierenden Maschinen nicht göttlich sind, und man fühle auch nicht die Aufgabe, die neuen Kräfte dieser Maschinen der Ehre Gottes zu unterstellen; aber dennoch haftet den neuen Maschinen laut Wiener etwas Außeralltäglichen an, wonach es gefährlich und falsch schiene, sie aus Eitelkeit und Selbstnutz zu verwenden. Er warnt deshalb vor jenen *Zauberern*, die der kybernetischen Verlockung ihrer Macht erlügen und die Früchte der neuen Universalwissenschaft für falsche Zwecke missbrauchten:

„There is a sin, which consists of using the magic of modern automatization to further personal profit or let loose the apocalyptic terrors of nuclear warfare. If this sin is to have a name, let that name be Simony or Sorcery. For whether we believe or not in God and his greater glory, not all things are equally permitted to us.“³³³

Diese Widersacher nennt Wiener an anderer Stelle auch *Gadget Worshipers* und *Priests of Power*. Sie wollten die Maschinen allein zu ihrem Eigennutz verwenden. Dafür nähmen sie selbst eine globale Vernichtung in Kauf; gleichzeitig nutzten sie die Autonomie der Maschinen dafür, ihre eigene Verantwortlichkeit für schwierige oder sogar verheerende Entscheidungen zu verschleiern. Bei alledem hegten sie letztlich ein großes Misstrauen gegenüber der Unabhängigkeit und Unvorhersehbarkeit des Menschen, die sie durch die Kybernetik zu usurpieren suchten.³³⁴ Wiener bestärkt mit seiner Argumentation, dass die Kybernetik *nicht von dieser Welt* sei und nimmt damit ältere Traditionsbestände auf, die der verselbständigten Technik schon lange etwas Numinoses – ein Fallen aus den Weltordnungen – unterstellten (s. 4.1.2). Seine Warnung vor den Zauberern der Kybernetik und die implizite Forderung nach einer kybernetischen Professionsethik lässt die Vermutung zu, Wiener warne vor einer Demokratisierung von Technologie. Demnach bleibt nur eine kleine Elite an neuen Hohepriestern der Kybernetik übrig, denen die Verantwortung dieser göttähnlichen Kraft überantwortet werden darf. In ihrer Hermetik ähneln die Kybernetiker dann den antiken Mysterienkulten, gnostischen Sekten oder den alchemistischen Geheimbünden.

333 Wiener, *God and Golem, Inc.*, 51–52.

334 Vgl. ebd., 53–55.

4.2.4 Erträge

Die tiefe Verunsicherung, die das kybernetische Projekt im Ganzen auslöste, illustriert Martin Heideggers Feststellung, dass der Eintritt in die kybernetische Welt das Ende der Philosophie bedeute. Heidegger zeigte sich besonders besorgt über die Machtphantasien, die sich im kybernetischen Begriff der Steuerung miteinander verbanden. Die universale Berechenbarkeit, auf die der kybernetische Weltentwurf abzielt, habe zu ihrem eigentlichen Zweck die absolute Beherrschbarkeit der leblosen und der lebendigen Welt in ihrer unterstellten Gleichmachung.³³⁵ Er war sich sicher, dass der Mensch an der Übersteuerung der Welt letztlich vergehen werde, weil sie dem Menschen seinen elementaren Lebensantrieb, die Hoffnung, verbiete, die überhaupt nur die unbedingte Eigensucht der menschlichen Subjektivität bedeute.³³⁶

Der Technikphilosoph Gotthard Günther sieht die wesentliche Entdeckung der Kybernetik darin, empirisch-technisch festgestellt zu haben, dass sich die Welt nicht in einer zweiwertigen Logik von tot und lebendig, von Mechanismus und Organismus, von Geist und Materie und damit Idealismus oder Materialismus erklären lasse.³³⁷ Aus ihren Konsequenzen fordere die kybernetische Ontologie deshalb vom Menschen westlicher Prägung nichts weniger als die freiwillige Aufgabe seiner Selbst und damit gleichzeitig das Abwenden von seiner Kultur, die sich für mehr als 2000 Jahre der Existenz einer Seele im Menschen gewiss war. Schließlich nötige die Kybernetik den Menschen, sich selbst in die leblose Materialität eines Mechanismus einzubauen, um diese so zu eigenem Leben zu erwecken.³³⁸ Dieses Eingehen des Menschen in den Apparat hält Günther für die Geburt der dritten, protometaphysischen Sphäre, die „weder Subjekt noch Objekt, wohl aber intelligenter,

335 Vgl. Martin Heidegger, „Die Herkunft der Kunst und die Bestimmung des Denkens. Vortrag gehalten am 4. April 1967 in der Akademie der Wissenschaften und Künste in Athen“, in: Petra Jaeger/Rudolf Lütke (Hgg.), *Distanz und Nähe. Reflexionen u. Analysen zur Kunst d. Gegenwart*. Hrsg. von Petra Jaeger u. Rudolf Lütke. [Walter Biemel zum 65. Geburtstag gewidmet, Würzburg 1983, 11–22. Hier 16.

336 Ebd., 19–20.

337 Vgl. Günther, *Das Bewußtsein der Maschinen*, 68.

338 Vgl. ebd., 182–184.

Information produzierender Prozess³³⁹ sei. Um in diese Sphäre einzutreten, müsse der Mensch sich technisch noch einmal wiederholen, bis ihm in dem maschinellen Bewusstsein seine objektiv gewordene Innerlichkeit entgegentrete.³⁴⁰

Der Wissenschaftstheoretiker Max Bense (1910–1990) sieht mit der Kybernetik den Menschen in seine technische Existenz eintreten. Jenes anthropologische Missverhältnis, dass der Mensch nicht vollständig aus der Natur expliziert werden kann, könne die Kybernetik lösen, indem der Mensch sich in die versöhnende *Sphäre des technischen Seins* zurückziehe und sich zu einem Kunstwesen seiner eigenen Umwelt mache, die seiner Doppelrolle als naturhaftes und geistiges Wesen nun endlich angemessen sei. Die technische Welt sei zwar nicht der erste Versuch einer Selbsteinrichtung des Menschen innerhalb der Natur, Bense aber hält ihn für den planmäßigsten und vielleicht erfolgreichsten dieser Art.³⁴¹

Der Psychologe Arnold Gehlen versucht, auf die Grenzen der kybernetischen Epistemologie hinzuweisen. Der technische Regelkreis habe zwar dieselbe Form des Wirkungszusammenhanges, wie sie beispielsweise zahlreiche innerkörperliche Regulationen zeigten; Gehlen stellt jedoch klar, dass man nur von einer Isomorphie, einer Gestaltgleichheit, und keiner substantiellen Gleichheit zwischen Regelkreisen und Lebewesen sprechen könne. Die Kybernetik sei von einer Synthese des Lebens selbst weit entfernt. Der Verdienst der Kybernetik sei es vielmehr, dargestellt zu haben, dass man bestimmte Teilvorgänge des Lebens in die Außenwelt objektivieren und neuen, eigenen Zwecken zuführen kann.³⁴²

Hans Jonas blickt Anfang der 1990er Jahre mit größerem Zeitabstand auf das die Kybernetik und ihr Erbe zurück. Er kritisiert die inflationäre Bedeutungsausweitung des Informationsbegriffes, die er über die

339 Ebd., 72–73. Hier sind starke Referenzen zu den Konvergenzerzählungen Pierre Teilhard de Chardins und Günther Anders (s. 4.4).

340 Vgl. ebd., 193.

341 Vgl. Max Bense, „Kybernetik oder Die Metaphysik einer Maschine (1951)“, in: Elisabeth Walther (Hg.), *Ausgewählte Schriften: in vier Bänden.*, Stuttgart/Weimar 1998, 429–448. Hier 436.

342 Vgl. Gehlen, *Die Seele im technischen Zeitalter*, 21.

Jahrzehnte nicht habe halten können. Auch sieht er ähnlich zu Gehlen in der Kybernetik das Maß dessen überschritten, was materialistische Theorien leisten können. Während die vorangegangenen Materialisten den Geist nur den Namen nach verwarfen, habe die Kybernetik erstmals ein mechanistisches Modell angeboten, das ohne Übergang auf materielle wie geistige Phänomene zutreffen soll. Die Symbolsprache der Physik sollte ein einheitliches Begriffsschema für die Darstellung der Wirklichkeit seit der Abdankung des Aristotelismus restituieren.³⁴³ Faktisch übernahmen sie die wissenschaftliche Arbeitshypothese Descartes, ohne die metaphysische Last ihrer Zwei-Substanzen-Lehre mitzudenken, wenn sie zwischen Sinnesreizung, afferenter Nervenleitung, zentralen Synapsen, efferenter Leitung und Muskelerrregung keinen Platz für das Geistige sahen.³⁴⁴ Jonas wirft den Kybernetikern deshalb vor, sie dementierten auf diese Weise ihre eigene innere Wirklichkeit, als ob der Kybernetiker selbst nicht unter die Bedingungen seiner Lehre von dem reaktiven Reizsystem falle. In dieser Gebrochenheit erkennt Jonas eine unfreiwillige Ironie in der kybernetischen Epistemologie:

„Die Ironie, die darin liegt, dass Naturwissenschaftler, so lange geschworene Feinde der anthropomorphistischen Todsünde, heute am freigiebigsten mit der Verleihung menschlicher Züge an Maschinen sind, wird nur dadurch gemindert, dass die wahre Absicht dieser Freigiebigkeit ist, den Spender, d. h. den Menschen, umso sicherer für das Reich der Maschinen in Beschlag zu nehmen.“³⁴⁵

Der Medienwissenschaftler Claus Pias schließlich weist das geistige Erbe der Kybernetik für die Gegenwart darin aus, die Hoffnung auf eine vom Menschen gereinigten Ontologie in die Welt gesetzt zu haben. Die Kybernetik sei mit dem Ziel angetreten, die Epoche der Menschenwissenschaften zu beenden, in welcher der Mensch empirisches Objekt eines möglichen Wissens als zugleich auch das ursprüngliche Konstitutionszentrum jeder Erkenntnis war. Die kybernetische Einebnung des

343 Vgl. Hans Jonas, *Das Prinzip Leben. Ansätze zu einer philosophischen Biologie*, 1. Aufl., Göttingen 1994, 199–200.

344 Vgl. ebd., 221.

345 Ebd., 216.

Menschen in das Informationsgeschehen des Kosmos lasse hingegen überhaupt keine Innensicht mehr gelten. Shannons statistisch generiertes Sprechen der Sprache wie auch Wieners Raum der Regelung und Nachrichtenübertragung in Lebewesen und Maschine seien die prominentesten Beispiele, in denen man nicht mehr gezwungen sei, *den Menschen* zu denken, sondern alle Wirklichkeit in Begriffen von Information und Steuerung aufgehen lässt.³⁴⁶ Zugleich ergebe sich bei einer derartigen Ausweitung des Technikbegriffes ein enormer Raum für Spekulationen über mögliche Ereignisse innerhalb dieser technischen Sphäre. Pias stellt die These auf, dass mit den kybernetischen Maschinen gleichsam *Utopie-Maschinen* erschaffen wurden, die zwischen Hoffnung und Gegenwart vermitteln sollten, ohne dass der Mensch behelligt werde: Lediglich das Ziel soll der kybernetischen Maschine vorgegeben werden – sei es der Abschuss eines Flugzeuges, der Erhalt des Blutzuckerspiegels, möglicherweise die gerechte Regierung –, aber den Weg dorthin überlässt man den installierten Regelkreisen.³⁴⁷

Der Wissenschaftshistoriker Michael Hagner resümiert, dass mit der Kybernetik eine geschichtsphilosophische Matrix in die Welt kam, die das weitere Schicksal des Menschen unwiderruflich an die Denkfähigkeit der Maschinen knüpfte. Und obwohl die Kybernetik als Bewegung zu ihrem Ende kam, ist er sich sicher, dass ihre Ideen über verschiedenste Vermittlungen sowohl im wissenschaftlichen Diskurs wie auch dem gesellschaftlichen Bewusstsein fortlebten. Aus dem kybernetischen Gedankengut mit ihrem Versuch der Überwindung von Körper und Seele, Mensch und Maschine hätten sich in den achtziger Jahren erste Konzepte eines technologischen Posthumanismus entwickelt. Zeitgleich sei der Begriff der Kybernetik in der Populärkultur in der Rede von Cyborgs und Cyberpunk, Cyberspace oder Cyberculture zurückgekehrt.³⁴⁸

Für die Konvergenzerzählungen um Künstliche Intelligenz ist die kybernetische Weltsicht grundlegend. Ihr dataistischer Reduktionismus der Welt ist Voraussetzung dafür, um den postulierten, kontinuierli-

³⁴⁶ Vgl. Pias, *Zeit der Kybernetik – Eine Einstimmung*, 16.

³⁴⁷ Vgl. Pias, *Unruhe und Steuerung. Zum utopischen Potential der Kybernetik*, 310–312, 325.

³⁴⁸ Vgl. Hörl, Hagner, *Überlegungen zur kybernetischen Transformation des Humanen*, 14–15.

chen Aufstieg der Lebewesen zu höherer Vervollkommnung zum Informationsgeschehen zu erklären. Schließlich hatte bereits Wiener den Informationsbegriff an den physikalischen Begriff der Entropie angeschlossen, indem, wie Entropie das Maß an Unordnung in einem System war, er nun die Information zum Maß ihrer Ordnung bestimmte.³⁴⁹ Wieners Auslegung von Maxwells zweitem Gesetz der Thermodynamik, das eigentlich nur für geschlossene Systeme gilt, führte zu dem Eindruck, dass alles Leben, die Zivilisationen als auch der gesamte Kosmos der Gefahr der entropischen Auflösung unterliege. Die kosmische Kraft, die der Entropie jedoch entgegenhalte, trägt den Namen der Extropie und wird durch das Maß an Information bestimmt. Es ist demnach nicht verwunderlich, dass viele der *Singularitans* sich ebenfalls als überzeugte *Extropians* verstehen, die mit dem Aufstieg der Maschinen nichts Geringeres als die Ordnung des Kosmos zu seiner höchsten Vollendung gewährleisten wollen (s. 5.3).³⁵⁰

So liegt für KI-Theoretiker Raymond Kurzweil der Sinn des Lebens für „jeden Mann der Wissenschaft“ offen zutage, dass über die Geschichte hindurch intelligente dynamische Systeme zu immer höherer Ordnung aufstiegen.³⁵¹ Die kybernetische, menschenenthobene Epistemologie derart verinnerlicht, kann Kurzweil die Evolution vom Menschen und allem Lebendigen überhaupt zu einem Prozess von Mustern aufsteigender Ordnung abstrahieren, die sowohl die biologische Evolution als auch dem technologischen Fortschritt miteinschließt und sie zu ein und demselben Prozess erklärt:

„I believe that it's the evolution of patterns that constitutes the ultimate story of our world. I conceptualize the history of evolution – both biological and technological – as occurring in six epochs.“³⁵²

Für diesen informatischen Monismus ist alles in der Welt, wie es La Mettrie postulierte, von einer Wesenheit. Dadurch leistet die Kybernetik schließlich dem Mythos der evolutionären Maschine Vorschub; denn

349 Vgl. Wiener, *Kybernetik*, 38.

350 Vgl. Davis, *Techgnosis*, 118.

351 Kurzweil, *The Singularity Is Near*, 249.

352 Ebd., 28.

wo die Kybernetik in den Regelkreisen von reaktiven Systemen keine handelnden Subjekte mehr kennt, wird die Unterscheidung zwischen künstlich und natürlich und damit Konstruktion und Evolution nivelliert. Beides wird zu ein und demselben fortdauernden Kampf gegen die Entropie des Universums stilisiert, in welcher die kulturellen Formen der Informationsverarbeitung das Rennen vor den Genen gewinnen. Moravec schreibt, dass über weite Strecke der menschlichen Artgeschichte das kulturelle Erbe weniger Daten als das genomische Erbe umfasst habe; in den letzten Jahrhunderten jedoch hätte die Kultur die Natur überholt, sodass immer weniger Biologie das Leben erfülle, bis sich schließlich die Kultur im intelligenten Roboter unabhängig von der Biologie mache und die Evolution sich allein in der Kultur fortsetze.³⁵³ Diese Nivellierung von Konstruktion und Evolution und damit auch der Trennung zwischen tot und lebendig ist es, die KI-Theoretiker wie Max Tegmark die Makrogeschichte des Lebens in drei Stufen unterteilen lässt. Nach dessen kybernetisch-extropischer Lebensdefinition steuert die Geschichte des Lebens auf der Erde seit ihrer Entstehung zielstrebig von den niederen Lebewesen mit ihrer nicht erweiterbaren *Hardware* und *Software* (Leben 1.0) weiter zum Menschen (Leben 2.0), der durch Lernen seine *Software* in begrenztem Maße erweitern könne, und vollende sich in der der substratunabhängigen, künstlichen Superintelligenz (Leben 3.0), die ihren Speicherplatz und ihre Rechenkapazität wie auch ihre gesamte physische Gestalt unbegrenzt verändern und erweitern könne.³⁵⁴

4.3 Technologische Konvergenzerzählungen des 20. Jahrhunderts

Die Vorstellung, dass die fortschreitende Unterwerfung der Naturkräfte es dem Menschen ermöglicht, völlig neue Lebensumstände hervorzu- bringen, nährt schon die ersten technischen Utopien. Die Kybernetik- bewegung hält an diesem Technikutopismus fest – ihr ist jedoch die Steuerung der Welt *durch den Menschen* suspekt geworden. Sie wähnt sich an der Schwelle eines neuen Zeitalters einer Allregelung durch

353 Vgl. Moravec, *Computer übernehmen die Macht*, 194–195.

354 Vgl. Tegmark, *Life 3.0*, 44.

Apparate. Zeitnah mit dem Aufstieg der Kybernetik entwickeln sich Groß Erzählungen, die einen solchen Schwellenmoment in einen noch größeren Sinnzusammenhang zu bringen suchen und ihr eine den Menschen übersteigende, kosmische Bedeutung zusprechen. Sie deuten die fortschreitende Technisierung der Welt als einen Prozess, den der Mensch schon lange nicht mehr beherrscht – es möglicherweise niemals tat –, sondern der selbst Ausdruck einer zulaufenden Bewegung sei, die sich in Wirklichkeit nur des Menschen bediene – ihn steuere –, um einen Zustand kosmischer Vollendung zu erreichen.

Der Jesuitenpater Pierre Teilhard de Chardin sieht in seinem 1940 erschienen Werk *Le Phénomène humain* (dt. *Der Mensch im Kosmos*) das Erdzeitalter im unmittelbaren Übergang in eine neue, nicht zuletzt technologisch induzierte Geistsphäre umschlagen, in der Materie und Bewusstsein sich vereinen und in einem Punkt Omega transzendieren. Das Werk erhielt kein *Nihil obstat* durch den Vatikan, wie Teilhard selbst bereits 1926 von seinem Orden wegen seiner evolutionstheoretischen Anschauungen nach China versetzt wurde.

Der Philosoph und Husserlschüler Günther Anders meint ebenfalls, in der Entwicklung von Mensch und Maschine eine zulaufende Bewegung zu erkennen. In seinem zweibändigen Hauptwerk *Die Antiquiertheit des Menschen* von 1956 deutet Anders diese Bewegung nicht hin zu einem Transzendenzpunkt des gesamten Universums, sondern sieht die drohende Katastrophe, in welcher der Mensch sich um seiner Maschinen willen auflöst und – degradiert zu einem Geräteteil – in eine dem Untergang geweihte Weltmaschine einfügt. Anders gehört zu den bedeutenden deutschsprachigen Technikkritikern der Nachkriegszeit, der unermüdlich vor der technologischen Katastrophe, seiner besagten Apokalypse ohne Reich (s. 2.2.3) warnte. Seinen Mitmenschen warf Anders eine zunehmende Apokalypse-Blindheit vor, weil sie sich nicht mehr die Folgen ihres technischen Tuns vorstellten. In ihrer Passivität hefteten sich die Menschen einer Nebelfahrt gleich noch weiter an den technischen Fortschritt, der sie letztlich an ihre Apparate ausliefere und schlussendlich zu vertilgen drohe.³⁵⁵

355 Vgl. Anders, *Apokalypse ohne Reich*, 210.

4.3.1 Teilhards Geistesvolution gen Omega

Teilhard de Chardin entwirft eine Makrogeschichte allen Lebens auf der Erde, die mit der Entstehung der Erde beginnt und in der Geistwerdung der Materie endet. Obwohl dem Menschen eine entscheidende Rolle zugerechnet wird, arbeitet Teilhard nicht anthropozentrisch, sondern geo- oder kosmozentrisch. Für den Archäologen Teilhard behüllt sich die Erde seit ihrer Entstehung mit immer komplexeren Schichten. Über der Erde als Körper (Geosphäre) und der Geburt von Leben auf ihrer Oberfläche (Biosphäre) habe sich eine Schicht bewusster Wesen entwickelt (Psychosphäre). Teilhard sieht diese Bewusstseinschicht nun in den Geburtswehen hin zu einer letzten, umfassenden Schicht des Geistes, der Noosphäre:³⁵⁶

„Wir stellen eine weitere Hülle fest. Rings um die Funken der ersten selbstbewussten Seelen das Anwachsen eines Feuerkreises. Der glühende Punkt hat sich erweitert. Das Feuer breitet sich immer mehr aus. Schließlich bedeckt die Glut den ganzen Planeten. [...] Eine ‚denkende Schicht‘ [...] über der Biosphäre eine Noosphäre.“³⁵⁷

Die Behüllung der Erde, die vor Milliarden Jahren seinen Anfang nahm, sieht Teilhard in den Ausläufern des 20. Jahrhunderts in ihre letzte Stufe eintreten. Ihm ist klar, wie viele Menschen sich vor ihm stets an einem Wendepunkt der Geschichte gewusst hätten. Allen diesen Zeitdeutern bescheinigt er, sie hätten ihre Gegenwart bis zu einem gewissen Grad immer richtig gedeutet. Denn der Fortgang der Seinsphären folge den Windungen einer aufsteigenden Spirale. Die gesamte Menschheit in der Einheit ihrer Generationen nennt er deshalb ein prophetisches Geschlecht, da die ersten Menschen sowohl den Eintritt in die Psychosphäre vollzogen hätten, wie nun auch die letzten Menschen die Zeitgenossen der Noosphäre sein werden.³⁵⁸

356 Vgl. Pierre Teilhard de Chardin, *Der Mensch im Kosmos*, 7. Aufl., München 1964, 168.

357 Ebd., 169.

358 Vgl. ebd., 205.

Teilhard erhebt in seiner Prozessphilosophie die Phänomenologie der Evolution zum absoluten Prinzip der Wirklichkeit. Alle Gebiete menschlicher Erkenntnis sieht er von derselben Grundströmung erfasst, die erst unter dem Gesichtspunkt der aufsteigenden Entwicklung verständlich würden. Dass sich alle Wirklichkeit nicht nur verändere, sondern mit ausreichend Abstand gesehen alles zu höherer Komplexität fortbewege, hält Teilhard für mehr als eine Theorie; sie sei überhaupt die allgemeine Bedingung für alle künftigen Theorien – das Licht, mit dem erste alle Sachverhalte ihre Gestalt erhielten:

„Indem unsere heutige Welt um sich und in sich die Evolution entdeckte, gewann sie, wie ich sagte, ihren spezifisch modernen Charakter. Die tiefste Wurzel der Unruhe in der modernen Welt, kann ich jetzt hinzufügen, besteht darin, nicht sicher zu sein und nicht einmal zu sehen, wie an je sicher sein könnte, dass es ein Endziel gibt – das befriedigende Endziel dieser Evolution. Wie muss denn die Zukunft beschaffen sein, damit sie uns die Kraft und gar noch die Freude gibt, ihre Ausblicke anzunehmen und ihre Last zu tragen?“³⁵⁹

Nach Teilhard kann die Menschheit mit dieser Ungewissheit des Wohin des in der Gegenwart erfahrenen Auftriebs nur im Extremen umgehen: Entweder müsse sie das Denken für eine Totgeburt der Evolution halten, die als Laune der Natur wieder verschwinden wird, oder sie fange an, von der Zukunft nicht nur ein Fortleben (*Survivance*), sondern ein höheres Leben (*Survie*) zu erwarten. Letzteres – der Sprung vom Bewusstsein zum Geist – glaubt Teilhard, sei nicht ohne Anstrengung des Menschen möglich. Wo bis zur Psychogenese kein Wollen nötig war, könne das Geistige jetzt „ohne den seiner Struktur gemäßen Anspruch nicht höher aufsteigen.“³⁶⁰ Das von Teilhard vorgestellte höhere Leben folgt dabei dem Prinzip der Vereinigung: Hatte die Psychogenese verschiedenste Bewusstseinsstufen und Einzelbewusstsein erschaffen, erwartet er für die Zukunft eine Vereinigung aller denkenden Kräfte auf der Erde. Die Moderne mit ihrer Durchmischung von Nationen,

359 Ebd., 220.

360 Ebd., 225.

Menschen und Ideen hält er für ein Zeichen des Kommenden und „die natürliche Kulmination eines kosmischen Organisationsprozesses“³⁶¹. Lässt sich die Menschheit auf ihre Bestimmung zum höheren Leben ein, dann werde sie sich immer weiter zu vernetzen und vor jedem Bestreben der Isolierung schützen:

„Der Ausgang aus der Welt, die Tore der Zukunft, der Eingang zum Übermenschlichen eröffnen sich weder einigen Privilegierten noch einem einzigen Volk, das auserwählt wäre unter allen Völkern! Die Pforten öffnen sich nur, wenn alle zusammen nach einem Ziel drängen, in dem sich alle zusammen vereinigen, um sich in einer geistigen Erneuerung der Erde zu vollenden.“³⁶²

Die Wissenschaft und ihre Apparate hält Teilhardt für einen notwendigen Schlüssel auf diesem Weg, denn die Vollendung der Dinge im Punkt Omega bindet er an einen höchsten Akt kollektiver Schau, zu dem allein Wissenschaft und Technik befähige.³⁶³ Die Menschheit in seiner nahen Zukunft werde erkennen, dass die Wissenschaft ihre wesentliche Form des Wirkens sei, um die Überfülle der kosmischen Energien aufzufangen. Für die Weltgemeinschaft auf den letzten Schritten hin zu Punkt Omega zeichnet Teilhard eine Techno-Utopie, die wie eine Überbietung von Francis Bacons *Nova Atlantis* erscheint: Nicht mehr nur eine Kaste wie noch das Haus Salomonis, sondern alle Menschen auf der ganzen Welt müssten dann keiner Versorgungsarbeit mehr nachgehen, sondern könnten ihre Freizeit der Erforschung und Gewinnung der Naturenergien widmen. In dieser – dem Punkt Omega vorgeschalteten – Zwischenzeit hätte die Menschheit den Besitz als etwas sie Trennendes abgeschafft, ihre Kriegstreiberei beendet und sei allein vom Drang nach Erkenntnis beseelt.³⁶⁴ Der Eintritt in die Noosphäre jedoch gehe noch weiter und bedeute die tatsächliche Überwindung des partikular Bewussten:

361 Vgl. ebd., 236.

362 Ebd., 237.

363 Vgl. ebd., 241.

364 Vgl. ebd., 275.

„Um völlig wir selbst zu sein, müssen wir in der entgegengesetzten Richtung voranschreiten, im Sinn einer Konvergenz mit allen übrigen, zum andern hin.“³⁶⁵

Schließlich vereinigten sich alle lebendigen Teilchen hin zu einem harmonischen Bewusstseinskollektiv. Anstelle einer Myriade von Denkteilchen werde sich die Erde mit einer einzigen denkenden Hülle bedecken, in der das Universelle und das Persönliche gleichzeitig bestehen blieben, weil sich die Vielheit individueller Reflexionen in einem gleichstimmten Bewusstsein, einem Zentrum vieler Zentren, sammelte und verstärkte. Von diesem Zentrum der Vereinigung aus geschehe „die Personalisation des Alls und der Elemente ohne ihre Vermischung in ihrem höchsten Grade“³⁶⁶. Diesen letzten Zustand Omega der Noosphäre und damit das Ende der Geschichte vermag Teilhard dann auch nur in mystischen Bildern anzudeuten:

„Das Ende der Welt: die Noosphäre, die das äußerste Maß ihrer Komplexität und zugleich ihrer Zentrierung erreicht hat, kehrt durch eine nach innen gerichtete Gesamtbewegung zu sich selbst zurück. Das Ende der Welt: ein Umsturz des Gleichgewichts, der den endlich vollendeten Geist aus seiner materiellen Hülle löst, um ihn künftig mit seiner ganzen Schwere auf Gott-Omega ruhen zu lassen. Das Ende der Welt: entscheidender Augenblick für unsere Verwandlung und unser Emporsteigen, unsere Reife und unsere Befreiung. [...] Tod des materiell erschöpften Planeten; [...] Befreiung jenes Teils des Universums, dem es gelungen ist, durch Zeit, Raum und Übel hindurch seine Synthese mühevoll bis ans Ende durchzuführen.“³⁶⁷

Er ist sich sicher, dass sich der Geist letztlich von dem Physischen trennen werde und damit einer kosmischen Katastrophe entgeht. Trotz der Unwahrscheinlichkeit seiner Aussichten und der Gefahren auf dem Weg durch die Zeit müsse der kosmische Geist durch den Menschen an sein Ziel gelangen.³⁶⁸ Teilhard ist sich gewiss:

365 Ebd., 257.

366 Ebd., 256.

367 Ebd., 284.

368 Vgl. ebd., 270–271.

„Irgendetwas im Kosmos entgeht der Entropie – und entgeht ihr immer mehr. [...] So ist das Universum daran, sich über unseren Häuptern aufzubauen, indem die Teilchen des Denkens die wahren und unzerstörbaren Atome seines Stoffes bilden – ein in seiner Resultante genau bestimmtes Universum – in Gegenrichtung zu einer Materie die sich verliert: ein Universum, Sammler und Bewahrer von Personen und nicht von mechanischer Energie, wie wir glaubten. Wie ein beständiges Ausströmen lösen sich rings um uns, eine um die andere, ‚die Seelen‘ los und tragen ihre unmittelbare Bewusstseinslast nach oben. – Eine um die andere: und doch nicht vereinzelt.“³⁶⁹

Und weiter macht er die funktionale Bedeutung der Menschheit als Austragungsort der Geistverwirklichung im Bild der Großmaschine deutlich:

„Die große Maschine der Menschheit ist zum Funktionieren bestimmt – und sie muss funktionieren – und einen Überschuss an Geist erzeugen. Wenn sie nicht funktioniert oder nur Materie erzeugt, dann arbeitet sie in einer falschen Richtung.“³⁷⁰

Teilhards prozessphilosophischer Entwurf des Geistes gen Omega verbindet die klassische Eschatologie mit einer auf Komplexitätssteigerung ausgerichteten Universalisierung der Evolutionstheorie. Am Ende seines Werks macht Teilhard dies noch einmal explizit, wenn er sagt, dass es Christus sei, der seit seiner Verherrlichung den allgemeinen Aufstieg der Sphären und die Beseelung des Bewusstseins leite. Punkt Omega identifiziert Teilhard mit dem christologischen *τὸ ἄλφα καὶ τὸ ὦ* in Referenz zu 1. Kor 15,28: „damit Gott alles in allem sei“. Wenn Teilhard anmerkt, dass Christen sich kurzzeitig vor der Evolutionstheorie gesorgt hätten, sie nun aber gerade eine Möglichkeit sei, sich Gott noch tiefer hinzugeben, dann scheint sein apologetisches Interesse durch, den ausgreifenden Evolutionsgedanken in der Heilsgeschichte

369 Ebd., 266–267.

370 Ebd., 250.

Gottes aufgehen zu lassen.³⁷¹ Mehr noch: Teilhard erklärt die christliche Nächstenliebe, verstanden als kosmischen Willen zur Vereinigung, zum alleinigen Weg, um die Menschen für das Einbrechen von Gott-Omega zusammenzuführen:

„Das Christentum allein, ganz allein auf der modernen Erde zeigt sich fähig, in einem einzigen, aus dem Leben entspringenden Akt das All und die Person zur Synthese zu bringen. Ganz allein kann es uns dahin führen, der ungeheuren Bewegung, die uns mit sich reißt, nicht nur zu dienen, sondern sie auch zu lieben.“³⁷²

Teilhard antwortet mit seiner Prozessphilosophie nicht zuletzt auf die Krisenstimmung seiner Zeit, deren Ursache sowohl in der Misere der industrialisierten Nationen durch den Zweiten Weltkrieg als auch in den Erkenntnissen der Astrophysik lag, die das Ende der Erde als wüsten Planeten und Ende allen Lebens wahrscheinlich werden ließen. Gegen das entropische Prinzip, das Teilhard in seiner Prozessphilosophie zum Widergöttlichen erklärt, restituiert er die Hoffnung auf ein Reich Gottes am Ende der Zeit.

Seine Überzeugung, dass der empfundene Fortschritt ein Fortentwickeln zu höherer Komplexität die Verwirklichung des schlechtinigen Guten bedeutet, widerspricht jenen, die seit spätestens Ende des 19. Jahrhunderts den wissenschaftlich-technologischen Progress vielmehr als fortschreitenden Verfall wahrnahmen. Ihr berühmtester deutschsprachiger Vertreter bleibt Oswald Spengler (1880–1936) mit seinem Werk *Der Untergang des Abendlandes*³⁷³. Spenglers Schrift ist eine große apokalyptische Geschichtsphilosophie, die der westlichen Kultur bescheinigt, ihren Zenit längst überschritten zu haben und nun unausweichlich ihrer zivilisatorischen Erstarrung zuzulaufen. In dieser Tradition, den Progress als Entwicklung zum Schlechten und schließlich zur Katastrophe zu deuten, reiht sich Günther Anders ein, der

371 Vgl. ebd., 293.

372 Ebd., 294.

373 Vgl. Oswald Spengler, *Der Untergang des Abendlandes*. Umriss einer Morphologie der Weltgeschichte (Beck's Historische Bibliothek), 230. Aufl., München 1998.

Mitte des 20. Jahrhunderts gerade in der Fortentwicklung der Technologie das Ende des Menschen vermutet.

4.3.2 Anders Konvergenz des Menschen in die Großmaschine der Ein-Apparat-Welt

Günther Anders fürchtet, dass der Mensch der vervollkommnen Perfektion seiner Geräte nicht mehr gewachsen sei und dies die Fortdauer der Menschheit aufs Spiel setze. Ein Anzeichen dieser Kapitulation des Menschen vor seinen Maschinen macht Anders an der Beobachtung fest, der Mensch seiner Zeit präge mehr und mehr einen Minderwertigkeitskomplex gegenüber seinen Geräten aus. Er spricht von der *prometheischen Scham*. Den ehemaligen Stolz des Entwicklers auf seine Erfindungen sieht er mit fortschreitender Verbesserung der Apparate in einen negativen Selbstbezug umschlagen. Im Angesicht der Maschine schäme sich der Mensch, aus dem blindem, antiquierten Prozess der Zeugung und Geburt entsprungen und kein gefertigtes, durchkalkuliertes Produkt zu sein. Deshalb begehre der Mensch ebenfalls zum Produkt zu werden, um seiner Fremderfahrung als Nichtgemachtes inmitten seiner gesamten fabrizierten Umwelt zu entfliehen.³⁷⁴ Den Stolz gegenüber den Produkten zurückzugewinnen, hält Anders für kaum möglich, da 99 % der Bevölkerung überhaupt nicht an dem Fertigungsprozess teilnehmen. Weil die Apparate kaum als eigene Leistungen wahrgenommen werden könnten, gerade deshalb wirkten sie von fremder, einnehmender, nicht zuletzt beschämender Art auf den Menschen.³⁷⁵

Hinter dem Minderwertigkeitskomplex meint Anders eine klassische christliche Häresie hervorberechen zu sehen, wie Augustin sie in seinen *Confessiones* als die Vertauschung von Macher und Gemachtem bezeichnete. Die Anbetung, die eigentlich dem Erschaffer der Dinge gebührt, werde fälschlicherweise dem Geschaffenen entgegengebracht. Auch der moderne Mensch ziehe laut Anders die gemachte Maschine dem Macher vor. In ihrer modernen Iteration habe sich jedoch die

374 Vgl. Anders, *Die Antiquiertheit des Menschen*, 25.

375 Vgl. ebd., 27.

Rolle des Menschen geändert. Während für Augustin der Mensch den Kreaturen – den wortwörtlich Erschaffenen – zugehörte, trete er nun allein in seiner Selbstbestimmung als *homo faber* auf. Die Ehre gegenüber seinen Fabrikaten schulde er nun seiner selbst, weshalb er sich nicht Gottes, sondern seiner selbst schuldig mache. Nicht Übermut (*Superbia*) wie im Falle Augustins, sondern Selbstentwürdigung kennzeichne deshalb diese moderne Verirrung.³⁷⁶ Die Scham, kein Ding zu sein, enthalte schließlich das Eingeständnis, dass der Mensch dem Produkt eine überlege Qualität anerkennt und sich selbst unter das Richtmaß des Produkts stellt:

„Denn als Konstruktion, gar als ‚fehlerhafte‘, kann der Mensch natürlich nur sub specie der Geräte gelten. Nur wenn diese Kategorie sowohl als universal anwendbar wie als erschöpfend anerkannt ist, kann die Umdeutung stattfinden, kann das Nichtkonstruierte als Schlechtkonstruiertes erscheinen.“³⁷⁷

Unter dem Blick der Entwicklungsfähigkeit würden so die Subjekte von Freiheit und Unfreiheit vertauscht. Die adaptionsfähigen Werkzeuge, die sich ursprünglich den Wünschen ihrer Meister anpassen mussten, diese reaktiven Dinge erschienen nun wahrhaft frei, während der Mensch mit seinen leiblichen Grenzen mehr und mehr als ein Hindernis in der volltechnisierten Umwelt wahrgenommen werde. Das *Human Engineering*, das zu Anders Zeit aus dem angloamerikanischen Raum in die internationale Öffentlichkeit dringt, ist ihm Erweis genug dafür, dass sich der Mensch seiner Zeit dem Ideal der Maschine bereits gebeugt hat. Durch das *Engineering* solle die menschliche Verfassung nicht mehr *Fatum* sein, sondern ihr alles Beschämende, eben das Fatale, genommen werden.³⁷⁸ Die sich im *Human Engineering* ausdrückende Verweigerung des eigenen Soseins hält Anders für einzig in der Ideengeschichte des Westens, unerreicht weder von mystischen Praktiken noch Versuchen magischer Selbstverwandlung. Neu sei auch,

376 Vgl. ebd., 25.

377 Ebd., 32.

378 Ebd., 38.

die Selbstverwandlung in eine Maschine zuliebe der Geräte durchzuführen; offen und schamlos trete das Gerät mit dem Anspruch auf, dass der Mensch sich ihm für seine maximale Wertschöpfung anbieten müsse. *Human Engineering* sei daher letztlich, in ihrer Zerrissenheit zwischen Anmaßung und Anmessung, Ausdruck von *angemessener Selbsterniedrigung* wie auch *hybrider Demut*, weil sie verkenne, dass der Mensch kein wirklicher Rohstoff ist, sondern sich in seiner Leiblichkeit vorgeformt vorfindet. Wissenschaft jedoch entwickle die Absicht, die falsche Prägung des Menschen aufzuschmelzen, um aus ihm nachträglich ein den Apparaten dienliches Material zu machen.³⁷⁹

Die schwerwiegendste Falschprägung erfahre der Mensch in seiner leichten Verderblichkeit. Produkte erschienen in ihrer Serienexistenz nahezu unsterblich. Die Welt der genormten Serienprodukte sei zwar kein reines Reich der Ideen, aber erscheine mit ihren millionenfachen Abdrücken von wenigen Entwürfen platonischer als alle bisherigen Umwelten des Menschen.³⁸⁰ Demgegenüber erlebe jeder Mensch seine Existenz in obsoletter Einmaligkeit:

„Dem beschämten Menschen dient als Folie seiner Sterblichkeit, sein aeternum, nicht Gott, der Sternenhimmel oder das Menschengeschlecht, sondern die Welt der eingelekten Pfirsiche und der unsterblichen, platonischen Markenartikel. Dies ist die Ewigkeit, die er respektiert, vor der er sich schämt und vor der er verlischt. Er geht nicht ein in den Olymp der Fertigwaren, sondern in den Hades des Rohstoffes.“³⁸¹

Einerseits erfahre der moderne Mensch sich in der Welt der Fabrikate als ungenügend und ordne sich willentlich der Gerätwelt unter; gleichzeitig aber bemerke er, dass ihm eine restlose Gerätconversion nicht gelingt, sondern ein ich-hafter Rest auch dann noch übrigbleibt, wenn er versucht, sich gänzlich in den Maschinenbetrieb einzuordnen.³⁸² In dieser Misere von Scham, Dingfrömmigkeit und Selbstüberwindung zeigt sich laut Anders, dass die Logik der Maschinen bereits erfolgreich

379 Vgl. ebd., 49–50.

380 Vgl. ebd., 52.

381 Ebd., 54.

382 Ebd., 88–89.

auf den Menschen ausgegriffen hat. Denn sie *wolle* nichts anderes, als dass der Mensch sich als Ding versteht, welcher in die Maschine angeschlossen gehört. Für Anders steuern die Apparate – nicht im direkten Sinne bewusst, sondern aus ihrer Funktionslogik heraus – allesamt auf einen Idealzustand zu, der die Welt und alles Leben einem einzigen, lückenlosen Apparat gleichschaltet, der alle anderen Apparate aufhebt und in dem *alles funktioniert*. Die vorfindlichen Apparate seien demnach unangeschlossen und nur *ex futuro* aus diesen „innewohnenden Zielbestimmungen“³⁸³ heraus verständlich. In der Logik der Maschinen sei die gesamte Welt ein einziges Rohstofflager, das erst, wenn es in die Apparaturen eingebunden wird, an Funktion, Relevanz und letztlich Sein gewinnt. Alles, selbst der Konsument des maschinell produzierten Güter, gehöre letzten Endes einverleibt. Was die mechanistischen Theorien des 19. Jahrhunderts in der Nachfolge Leibniz als Schilderung des faktischen Zustandes des Universums unterstellt hatten – nämlich die Welt als ein großes Uhrwerk –, sei die Zukunft, auf welche jeder bestehende Apparat in der Gegenwart hin konvergiere.³⁸⁴

Das Wesen der Maschine besteht für Günther Anders in der Expansion. Jede Maschine dränge auf einen Zustand der Allversorgung und Vernetzung hin, in dem die externen Prozesse zusammen mit ihren eigenen ein einziges großes Funktionsganzes bilden. Den Expansionsdrang der Maschinen hält Anders für unersättlich. Die Großmaschine, die ihre nächstliegenden, für ihr eigenes Funktionieren unentbehrlichen Vorgänge an ihre eigenen Funktionen gleichschalten konnte, werde auf nächsthöherer Ebene weiter ausgreifen. Jede Maschine müsse, um optimal zu arbeiten, entweder versuchen, ihre Umgebung zu erobern, oder sie müsse sich in eine andere, größere Maschine einschalten. Bereits zu Anders Zeit könne man längst nicht mehr von Einzelmaschinen reden, sondern von einem sich immer weiter zusammenschließenden Maschinenpark. Eroberung und Expansion führten zu immer weniger eigenständigen Maschinen, die dafür umso riesiger würden. Der Prozess erschöpfe sich jedoch nicht im Numerischen; mit der Verschaltung

383 Günther Anders, *Die Antiquiertheit des Menschen*. Bd. II: Über die Zerstörung des Lebens im Zeitalter der dritten industriellen Revolution, München 1980, 110–113.

384 Vgl. ebd., 112.

trete auch ein qualitativ-dialektischer Umschlag ein. Anders geht davon aus, dass die gleichgeschalteten Maschinen aufhörten, Maschinen zu sein und ontologisch zu Geräteteilen von Großmaschinen absänken. Wie der Mensch in seiner Gleichschaltung im Apparat seine Persönlichkeit verliere, entspreche seiner Verdinglichung eine Verdinglichung der Dinge.³⁸⁵ Die Zahl von Maschinen, die wirklich noch Maschinen sind, und nicht nur Rädchen in einer Maschine, sieht Anders von Tag zu Tag kleiner werden, sodass das Maß der Autonomie selbst im Reich der Apparate abnehme. Bei immer größerer Verschaltung betreffe dieses Absinken immer größere Systeme. Wenn Einzelmaschinen zu Teilen von Großmaschinen geworden sind, dann geschehe das Gleiche auch mit diesen Großmaschinen und sie sänken zu bloßen Maschinenteilen von wiederum größeren Komplexen herab. Falls die Menschheit sich und ihre Maschinen nicht vorzeitig atomar vernichtet, so ist Anders gewiss, stehe am Ende dieses maschinellen Willens zur Macht der eine Welt-Apparat, in dem alle Maschinen und alles Leben als Teile in einer einzigen, mit dem gesamten globalen Produktionssystem identischen Maschinerie aufgehoben sind:

„Wäre diese Situation, auf die die zahllosen Maschinen von heute ausnahmslos abzielen, verwirklicht, dann würde das natürlich nicht bedeuten, daß es weniger Maschinelles gäbe als heute. Umgekehrt würde dann nichts mehr existieren, was nichtmaschinell wäre. Aber es würde bedeuten, daß dann, abgesehen von dem einen totalen maschinellen Dinge, dem sie alle zugehören würden, kein Objekt mehr existieren würde, das noch den Anspruch darauf machen dürfte, sich als individuelle Maschine bezeichnen zu lassen. Es gibt heute kein Gerät, das nicht diesem totalitären Endzustande entgegenträumte, in dem es selbst nur noch als Geräteteil eines Geräteteiles eines Geräteteiles eines Geräteteiles Bestand haben und funktionieren würde.“³⁸⁶

Das Ende der Geschichte bedeute die eine Welt-Maschine, außerhalb der nichts besteht und deren Teile – ob Mensch oder Maschine – ihre

385 Vgl. ebd., 119–120.

386 Ebd., 120–121.

Autonomie abgetreten haben. Dieser eine Superapparat ist für Anders eine degenerierte Wirklichkeit, eine Entartung der maschinellen Logik. Sie bilde zwar ihr zwangsläufiges Ende, sei aber an sich ein defizitärer Zustand, weil die Abhängigkeit der Funktion des Ganzen und des Einzelnen derart gesteigert sei, dass dem Koloss früher oder später der apokalyptische Kollaps drohe.

4.3.3 Erträge

In Angriff von Hans Blumenbergs Formel von der *Arbeit am Mythos* bilden beide Konvergenzerzählungen eine schlüssige Verbindung zu dem, was Literatur und Wissenschaft zuvor zum Verhältnis von Mensch und Maschine vorgetastet hatten; gleichzeitig sind diese neuen Verknüpfungen selbst musterbildend geworden für die Konvergenzerzählungen, wie sie im Übergang zum 21. Jahrhundert um Künstliche Intelligenz entstanden sind. Retrospektiv lassen sich beiden Visionen innerhalb der dargestellten Traditionsgeschichte geistige Paten zuordnen. Teilhards Prozessphilosophie offenbart sich in ihrer Fixierung auf den Geistbegriff als ein theologischer Wiedergänger von Joachim von Fiore trinitarischer Periodisierung der Geschichte und seiner utopischen Erwartung eines innerweltlichen Zeitalters des Geistes (s. 4.1). Wie der Systematiker Wolfhart Pannenberg (1928–2014) feststellte, ist der Geist für Teilhard jedoch kein Meta- oder Epiphänomen mehr, sondern das Phänomen der Wirklichkeit schlechthin. Die schöpferische, einigende Wirkung Omegas, die Teilhard mit dem Begriff des Geistes beschreibt, diese Macht der Konzentration und Konvergenz identifiziert er sowohl hinter der biologischen Evolution, dem menschlichen Bewusstsein, als auch dem technologischen Fortschritt. Teilhard holt somit auch das menschliche Bewusstsein als Ausdruck des Geistes in den Naturzusammenhang zurück. Mit seinem Panpsychismus reiht Teilhard sich in eine lange philosophische Tradition ein, deren neuzeitliche Gestalt vor allem mit der Monadenlehre von Leibniz verknüpft ist.³⁸⁷ Weil für Teilhard der Mensch und sein Bewusstsein als Ausdruck

³⁸⁷ Vgl. Wolfhart Pannenberg, „Geist und Energie. Zur Phänomenologie Teilhards de Chardin (1971)“, in: Wolfhart Pannenberg (Hg.), *Beiträge zur Systematischen Theologie*. Bd. 2 Natur und Mensch – und die Zukunft der Schöpfung, Göttingen 2000, 55–65. Hier 62–63.

des Geistes mit allem verbunden ist, wird und muss alles mit ihm transzendieren. In dieser Perspektive ist auch die Technik beseelt, auch sie erscheint wie der Mensch von demselben Auftrieb ergriffen und verwischt damit die Trennung zwischen Totem und Lebendigem, Konstruiertem und Gewordenen, Mensch und Maschine.

Teilhard ordnet darin die Menschengeschichte zu einer Episode in die umfassendere Geschichte des Geistes ein. Stehen bei Teilhard der Eigensinn des Geistes und das Mitreißenlassen vom Grundstrom des Lebens im Zentrum, muss in Günther Anders Geschichtsaufriß der menschliche Wille der Maschinenlogik in einem evolutionär-ideologischen Kampf standhalten. Hier tritt dem Menschen in der Technik kein dem eigenen Bewusstsein identischer, Mensch und Materie gleichsam ergreifender und einender Gottesgeist gegenüber, sondern ein auf blinde Expansion ausgerichteter Wille zur Macht. Während sich Teilhard an den Geistmillenarismus Fiores und Bacons Hoffnung auf eine technikgeleitete Einkehr in das Paradies rückbinden lässt, stellt sich Anders Denken in die Nachfolge Samuel Butlers und seinen Werke *Darwin among the machines* und *Erewhon* (s. 4.1.2). Wie Butler weitet auch Anders die Lehren Darwins in einen animistischen Technodarwinismus auf die Maschine aus und erzählt von einem Kampf ums Überleben zwischen Mensch und Maschine. Tritt bei Teilhard bereits die Entropie als das kosmische, widergöttliche Prinzip auf – hier jedoch noch nicht mit dem Informationsbegriff in Beziehung gesetzt wie dann bei Norbert Wiener –, so wird für Günther Anders der Maschinenpark selbst zum diabolischen *Legion* (Mk 5,9) im Sinne eines unpersonalen, unbewusstbewusst agierenden Kollektivs, welches den Menschen – nicht aus Boshaftigkeit, sondern Notwendigkeit – in Abhängigkeit zu bringen und letztlich auszulöschen sucht. Während bei Teilhard der Mensch in Richtung Omega seine Apparate als naturgegebene Erweiterung seiner selbst wahrnimmt und mit ihnen transzendiert, kehrt Anders die Symbiose in einen Parasitismus um. An ihrem Ende steht im Gegensatz zur Ewigkeit des vollendeten Zustands Omega eine fehleranfällige Maschine und das ontologische Absinken aller Individuen zu bloßen Funktionsteilen; die Entwicklung zum Welt-Apparat und das ontologische Absinken bleibt für Anders nahezu unausweichlich, aber im Gegensatz zu Chardin ist dieser Zustand nicht sinnvoll. An Anders Konzeption ist erstaunlich,

wie er die Popularität des Mythos der Mensch-Maschine-Gleichung bereits Mitte der 1950er diagnostizierte, dann aber seine Mythoskritik in einen eigenen apokalyptischen Mythos von dem Weltapparat überführt. Teilhards Einbau der Konvergenz des Menschen mit seiner Technik in die christliche Erlösungsteologie hingegen erscheint als ein forscher Brückenschlag zwischen römischem Katholizismus, technizistischer Fortschrittsgewissheit und Darwin'scher Vererbungslehre – in ihren animistischen Tendenzen und dem deistischen Gottesbild allerdings weit vom kirchlichen Lehramt entfernt.

Die Unterschiede beider Konvergenzerzählungen sollten nicht darüber hinwegtäuschen, dass sie Spielarten der im Kern gleichen existenziellen, geradezu apokalyptischen Aussagen über den Menschen und seiner neuen *Stellung im Kosmos* sind. Beide Erzählungen sehen den Menschen in höchstem Maße fremdgesteuert. *Seine* Geschichte als auch *seine* Technik werden ihm zu einem Gegenüber einer fremden Macht mit eigenen Zielen geoffenbart. Diese Fremdbestimmung wirke im Geheimen bereits in der Gegenwart und werde sich den Menschen über die Zeit gänzlich unterordnen. Für Anders wie für Teilhard bedeutet die Vollendung dieser Machtübernahme die höchste Form von Organisation und Verschaltung. Omega als auch die Welt-Maschine beschreiben beide eine komplett verwaltete Welt. Der Mensch ist in beiden totalumfasst und mit ihnen verwandelt.

Aktuelle Bestrebungen von einem Internet der Dinge oder dem Metaverse, in der jede Person und jedes Objekt miteinander kybernetisch verschaltet sind und Daten mitteilen, lassen sich leicht mit solchen Bildern einer eingefassten Welt in Verbindung bringen. Der Mensch, der vollüberwacht unter der totalitären Herrschaft eines superintelligenten, künstlichen Agenten steht, oder das hochgeladene Bewusstseinssubstrat, das umfassen ist im Raum des Digitalen, führen beide dieses Motiv in eine postsinguläre Wirklichkeit fort. Der von Anders beschriebene Maschinenpark, der sich immer weiter und schneller verschaltet, trägt ebenso bereits die gleichen Merkmale von autonomer Verbesserung durch Wiederholung ohne eine Stoppregel in sich, wie sie der Informatiker John Irvin Good ein Jahrzehnt später den sich kontinuierlich selbst verbessernden Maschinen in der informatischen Intelligenzexplosion zuschrieb. Es nicht verwunderlich, dass beson-

ders Teilhards religiös inspiriertes Vokabular unter den Visionären der Künstlichen Intelligenz Fuß fassen konnte. Sie offenbarten sich als geistige Erben des Punktes Omega, auch wenn den Allermeisten vordergründig der metaphysische Überbau fehlt: Schmidhuber verweist als einer der Wenigen dezidiert auf das Werk Teilhards und versieht das von ihm prognostizierte Auftreten einer Intelligenzexplosion um das Jahr 2045 mit dem Kürzel Ω . Hans Moravecs Sprachbild vom Geistfeuer (*Mind Fire*), welches mit der Mentalisierung der Technik das gesamte Universum überziehen werde, stammt ebenfalls aus dem Repertoire Teilhards. Die christlichen Implikationen der Sprachbilder – von der Beseelung der Jünger durch Feuerzungen im Pfingstereignis (Apg2) oder das Omega als Symbol für Christus als den Ersten und Letzten (Off 22,13) –, sie wirken in den Konvergenzerzählungen der Künstlichen Intelligenz weiter und bestärken die stilistische Nähe der Konvergenzerzählungen zur apokalyptischen Rhetorik (s. 2.2.3). Hier bereits deutet sich das Problem an, wie von theologischer Seite mit dieser Aneignung christlicher Metaphern umzugehen ist (s. 6.).

Der wohl wichtigste Punkt, den die Großerzählungen von Teilhard und Anders sowohl untereinander als auch mit den neuen KI-Konvergenzerzählungen verbindet – und sie von der klassischen Apokalyptik trennt –, ist ihr verlorenes Interesse am gegenwärtigen Menschen bis hin zur Erwägung seiner Obsoleszenz. Ihn sind sie beide bereit, für den Gewinn einer kosmischen Perspektive zu opfern. Der vorfindliche Mensch gerät zum unvollendeten Zwischenwesen. Als bloße Erscheinungsform (Teilhard: *Phénomène*) höherer Kräfte – sei dies ein innerweltlicher, technikumfassender Evolutionstrieb oder ein auf Vollendung zustrebender, in aller Materie wirkender Geist – wendet sich das Interesse dem Fortgang dieser Energien als seinem kontemporären Behältnis zu. Während die christliche Apokalypse um Gottes und des Menschen willen geschieht, ist der Mensch bloß noch ein Brückenwesen für die kommende, höhere Existenz.³⁸⁸ Beide, Teilhard wie Anders,

388 Vgl. Oliver Müller, „Von der Selbstüberschreitung zur Selbstersetzung. Zu einigen anthropologischen Tiefenstrukturen des Transhumanismus“, in: Mitscherlich-Schönherr/Reiner Anselm/Martin Heinze (Hgg.), *Das Gelingen der künstlichen Natürlichkeit. Mensch-Sein an den Grenzen des Lebens mit disruptiven Biotechnologien*, Berlin/Boston 2021, 333–350. Hier 340–341.

messen der Maschine einen enormen Wert in einem solchen Ablösungsgeschehen zu. Sie tritt als neue Inkarnation dieser kosmischen Urenergien auf. Die Konvergenzerzählungen um Künstliche Intelligenz sind sich bloß uneins darüber, ob die Menschen sich mit der Technik mitverwandeln, oder sie ihre Funktion erfüllt haben und damit ihre Geschichte endet, während die ihrer maschinellen Nachkommen – *Mind Children* – beginnt.

Die Groß Erzählungen vermitteln der Menschheit Sinn im Kosmos als einer Ordnung, in der sie zum Werkzeug dieser Urenergien verkommt. Teilhard und Anders glauben nicht, dass der Mensch sich wirklich dem Anstehenden konsequent verschließen kann. Sie beide erklären es zur Tugend, *einsichtig* zu sein, sich dem kosmischen Willen zu ergeben. Vielmehr bestreiten sie, dass es überhaupt sinnvoll sei, sich diesen Einflüssen zu wehren, sie aus einer beschränkten, menschlich-egozentrischen Perspektive zu betrachten. Für Anders ist der Überlebenskampf des Menschen gegen die Maschine bereits verloren, er sieht ganz nach Spengler einen schleichenden Untergang der menschlichen Zivilisation vor sich, den die Menschheit atomar bloß verkürzen könne. Auch hier wiederholt sich Samuel Butlers Argumentation in veränderter Tonalität, nach der dem Menschen alternativ zum Sklavenstand durch die Maschinen allein die Selbstvernichtung bleibt. Auf ein rechtzeitiges Umdenken der Menschen und einem daraus folgenden, anderen Umgang mit seinen Maschinen setzt Anders hingegen wenig Vertrauen. Und für Teilhard braucht der Mensch seiner selbst nicht nachzutruern, sondern soll sich voller Hingabe von der Kraft zur Konvergenz mitreißen lassen und im höheren Leben (*Survie*) vergehen. Eben dieses Ineinander in der Forderung nach Aktivität des Menschen bei gleichzeitigem Fatalismus wiederholt sich in den Erzählungen um das Erreichen der technologischen Singularität: Einerseits wird der technologische Umbruch mit der Postulierung eherner Fortschrittsgesetze zu einer wissenschaftlichen Unabwendbarkeit stilisiert, in welcher dem Menschen keine tatsächliche Urheberschaft zukommt; dann wiederum tritt der Mensch wie bei Good doch als ein unersetzlicher Konstrukteur einer ersten Schwellenmaschine auf, ohne den gar nichts geht.

Der im Kern liegende Antihumanismus dieser Konvergenzerzählungen verbirgt sich unter dem Deckmantel einer kosmischen Allper-

spektive einer missverstandenen Selbstlosigkeit und weist große Ähnlichkeiten zu den Konvergenzerzählungen der Künstlichen Intelligenz auf. Unter dem Vorwand einer Nüchternheit, die sich von menschlicher Egozentrik lösen konnte – der besagten, vom Menschen gereinigten Ontologie (s. 4.2) –, verbirgt sich in Wirklichkeit eine Maßlosigkeit, die sich genau in dem Grade zwischen Hochmut und Selbsterniedrigung bewegt, wie Günther Anders es den *Human Engineers* seiner Zeit in ihrer Dingfrömmigkeit vorwarf. Digitalvisionäre sehen die Mission der Menschheit angesichts der von ihnen diagnostizierten Defizite des Menschen darin, sich durch technischen Fortschritt überflüssig zu machen. In dieser techno-eschatologischen Perspektive ist das anvisierte Ende der Menschheit Gewissheit, wenn nicht für die *Einsichtigen* Grund zur Freude. Denn nur so trage die Menschheit ihrem angeblichen Sinn in der Geschichte Rechnung. Diese Maßlosigkeit der beanspruchten Epistemologie schließt eine ethische Besinnung aus, Maßlosigkeit wird in den Konvergenzerzählungen vorangegangener wie gegenwärtiger Couleur vielmehr zum neuen Maßstab erklärt, der meint, jene begrenzte menschliche Perspektive hin zu höherer Wahrheit überwunden zu haben:

„Könnte das Universum auf dem Weg des Denkens wie auf dem Weg der Zeit und des Raumes anders als im Maßlosen enden?“³⁸⁹

5 Technologische Konvergenzmythen: Visionen einer post-singulären Wirklichkeit

Analog zur christlichen Eschatologie verkünden die Visionäre einer starken Künstlichen Intelligenz mit der technologischen Singularität den Anbruch einer neuen Wirklichkeit, die in diametralem Gegensatz zu gegenwärtiger Lebenserfahrung steht. Mithilfe des Paradigmas eines sich überschlagenden, unaufhaltsamen technologischen Fortschritts und einer kybernetischen Epistemologie versuchen sie zu plausibilisieren, dass die Konvergenz von Mensch und Maschine unmittelbar bevorsteht. Allein dass ein Herrschaftswechsel von Mensch zur Maschine und damit das Ende des Anthropozäns auf dieses Ereignis fällt, scheint das maßgebliche Kennzeichen des postsingulären Äons zu bleiben.³⁹⁰ Was genau passieren wird, ob diese Überwindung des Status quo gefürchtet, möglicherweise ersehnt oder doch – zumindest als *Fatum* – akzeptiert werden müsste, bleibt über diesen Wendepunkt hinaus schemenhaft. In ihrer Pluralität liegt den Beschreibungen dieser neuen Wirklichkeit so auch kein verbindlicher Textkorpus zugrunde. Viel eher haben sich bestimmte Führungsfiguren und Forschungsinstitutionen herausgebildet, die den Diskurs um diese entgrenzten Zukünfte maßgeblich bestimmen. Im Wesentlichen haben sich die visionären Diskussionen über die Zukunftsaussichten starker KI nach einem Muster entwickelt, das aus älteren zukunftsvisionären Technikdebatten – z. B. über Nanotechnologie – bekannt ist: In mehr oder weniger eng miteinander verbundenen, relativ kleinen Zirkeln namhafter Naturwissenschaftler, Ingenieure, Autoren und Industrieller werden z.T. sehr weitreichende Zukunftserwartungen diskutiert.³⁹¹ Sie

390 Die Gleichnisse des Neuen Testaments erheben dagegen gerade den Anspruch, ihren Zuhörern eine Verbindung zwischen der Gegenwart und dem zukünftigen Heil herzustellen. Sie orientieren sich an den erfahrungsgesättigten Bildern vom Aussäen, vom Suchen und Finden innerhalb ihrer landwirtschaftlich geprägten Umwelt und ermutigen so, irdisches Jetzt und eschatologisches Dann in eine Beziehung zu setzen, ohne dem Reich Gottes seine Transzendenz zu nehmen.

391 Vgl. Christoph Kehl, Christopher Coenen, Technologien und Visionen der Mensch-Maschine-Entgrenzung, 8.

eint ihr universitäres Renommee, ihre finanzstarken, oftmals privaten Forschungsinstitute, und ihr Anspruch, nicht Science-Fiction, sondern wissenschaftsbasierte Technikfolgenabschätzung zu betreiben (s. 2.2.1).

Die folgende Strukturierung der post-singulären Erzählungen in drei Hauptmotive ergibt sich einerseits thematisch. Sie markieren die verschiedenen Distanzen zwischen Mensch und Maschine unter der Prämisse ihrer hergestellten Gleichwertigkeit: Das Konfrontationsmotiv deutet die Konvergenz als ein Gegeneinander von Mensch und Maschine, in der die Maschine revoltiert, den Menschen dominiert und ihn schließlich verdrängt. Das Vertretungsmotiv rückt Mensch und Maschine näher zusammen; es imaginiert ein Miteinander von Mensch und belebter Technik, in dem das Technische als Ersatz sowohl für die eigene Person als auch mehr und mehr an die Stelle des kommunikativen Gegenübers im Versprechen einer solipsistischen Bedürfnisbefriedigung tritt. Das Identifikationsmotiv schließlich setzt das Schicksal von Mensch und Maschine gänzlich in eins und überführt das Nebeneinander in ein Ineinander, indem Mensch und Maschine zu einer neuen Seinsform verschmelzen.

Jedes dieser Motive deutet somit die Chiffre von der Singularität anders aus unter der gemeinsamen Voraussetzung einer verlebendigten und mentalisierten Technik, die nur verstellt mit dem in Verbindung steht, was maschinelles Lernen tatsächlich leistet (s. 3.5). Das erste Szenario setzt am stärksten auf die technologische Singularität im Verständnis der Intelligenzexplosion nach Good, in dem die von Menschen erbaute Maschine durch Selbstlernprozesse Superintelligenz erreicht und deren eigene Ziele im weiteren Verlauf mit denen der Menschen kollidieren. Das zweite Szenario versteht unter der Singularität den Moment, wenn Maschinen das Menschliche vollständig durchdrungen haben, sodass sie die Bedürfnisse eines jedes Individuums auf Grundlage von Daten umfassend antizipieren können. Szenario Drei bezeichnet mit der Singularität die erfolgreiche Übertragung menschlichen Bewusstseins in die Maschine, in der das digitalisierte Bewusstsein selbst zur Superintelligenz wird. Diese drei Motive stehen nicht unverbunden zueinander, sondern die Bedrohung durch das Technische, wie sie am klarsten im ersten Szenario zur Sprache kommt, trägt in beträchtlichen Maße auch zum Verständnis der anderen bei-

den Motive bei. Das Miteinander und schließlich die Identifikation von Mensch und Maschine lesen sich als Alternativen vor der drohenden, gewalttätigen Ersetzung des Menschen. Die trans- und posthumanistischen Euphorie, wie sie in manchen Bewertungen dieser angeblich besseren Ausgänge der Zukunft zutage bricht, erscheint deshalb von Beginn an gebrochen; sie gerieren sich als beste Möglichkeiten unter den Umständen einer sicher geglaubten Überwältigung des Menschen durch die Technik.

In systematisch-ethischer Hinsicht akzentuieren die drei Motive unterschiedliche Selbst- und Weltdeutungen, an deren Setzungen die ethische Tradition des Christentums konstruktiv anknüpft. Sie tut dies unter der Maßgabe, menschliches Handeln in der Welt als Vollzug gewährter Freiheit zu verstehen (s. 2.1), die ihren Grund nicht aus den verfassten Bedingungen in der Welt, sondern von Gott her empfängt. Diese Freiheit gegenüber den vermeintlichen Zwängen der Welt liegt im Kern der reformatorischen Wiederentdeckung des allein durch Glauben gerechtfertigten Christen. Weil über Heil und Unheil bereits entschieden ist und nicht in der Welt hergestellt werden braucht, kann sich das sich im Glauben befreit wissende – und damit selige – Bewusstsein der Welt, dem Mitmenschen wie der eigenen Geschichte über anders verhalten, sich kritisch zu seinem Handeln in Beziehung setzen und als verantwortliches Subjekt seiner Lebensführung in Erscheinung treten.³⁹²

Die Technikmythen loten unter dem Eindruck der anstehenden Konvergenz die Freiräume des Menschen und seine Bestimmung in der postsingulären Welt neu aus. Den damit einhergehenden Freiheits- und Unfreiheitsvorstellungen in den Konvergenzrählungen Kontur zu verleihen, hängt demnach an der analytischen Schärfe des herangezogenen Freiheitsbegriffes. Die Konkretion der Grundbedingungen christlicher Freiheit übernimmt die ethische Theologie in der Zentralstellung der Trinitätslehre, wie sie in der dogmatischen Theoriebildung des 20. Jahrhunderts von Karl Barth, Wolfhart Pannenberg und Trutz Rendtorff vorgenommen wurde.³⁹³ Sie folgen damit Martin Luthers Ein-

392 Vgl. Rendtorff, *Ethik*, 12–13.

393 Vgl. Anselm, *Schleissing, Einführung*. Zum Ort der »Ethik« im Werk Trutz Rendtorffs, XVIII–XIX.

sicht, dass die *libertas christiana* nur in der Weise an Gestalt im eigenen Lebensvollzug gewinnt, als sie im Sinne einer Gabe von ihrem Geber her verständlich wird:

„Das sind die drei Personen und der eine Gott, der sich uns allen selbst ganz und gar gegeben hat mit allem, das er ist und hat. Der Vater gibt sich uns mit Himmel und Erde samt allen Kreaturen, damit sie uns dienen und nützen [...] Darum hat er danach auch den Sohn selbst uns gegeben, uns alle seine Werke, sein Leiden seine Weisheit und Gerechtigkeit geschenkt und uns mit dem Vater versöhnt [...]. Weil aber solche Gnade niemandem nützlich wäre, wenn sie geheim und verborgen bliebe und nicht zu uns käme, deshalb kommt der Heilige Geist und gibt auch sich zu uns ganz und gar.“³⁹⁴

In den Glaubensaussagen darüber, wie Gott den Menschen geschaffen, mit sich versöhnt und erlöst hat, deutet sich jene im Glauben erfahrene Freiheit auch für die eigene Lebensführung positiv aus. Im reformatorischen Bewusstsein einer Unvertretbarkeit des Einzelnen vor Gott fordert die vom gläubigen Bewusstsein vorgenommene Kontextualisierung der *libertas christiana* dabei die eigene ethische Stellungnahme ein und distanziert sich von einer bloßen Kasuistik biblischer Vorschriften. Die historisch geschärfte Sicht auf den christlichen Traditionsbestand stellt vielmehr ein Koordinatensystem für das eigene ethische Problembewusstsein bereit, anhand dessen sich die eigene Stellungnahme ausbildet.³⁹⁵

Im Licht christlicher Freiheit stellen die postsingulären Aussichten drei Anfragen an das theologisch-philosophisch gewachsene Freiheitsverständnis der Moderne und bieten mitlaufend ihre Alternativen an. Erstens: Gegenüber dem Versprechen seit der Aufklärung, der Mensch werde durch sein Wissen die Welt in Besitz nehmen, und der christlichen Zusage des Menschen als Geschöpf und Ebenbild Gottes fragen

³⁹⁴ Martin Luther, „Bekenntnis (1528). Übertragung: Christopher Spehr“, in: Dietrich Korsch (Hg.), *Deutsch-deutsche Studienausgabe / Luther, Martin*. Band 1 Glaube und Leben, Leipzig 2012, 551–570. Hier 563 (WA 506).

³⁹⁵ Vgl. Anselm, Schleissing, Einführung. Zum Ort der »Ethik« im Werk Trutz Rendtorffs, XV.

die Erzählungen an, ob der Mensch den Herrschaftsauftrag und seine damit einhergehende Verantwortung für die Welt nicht auch ablehnen kann, indem er die Beschwerde einer immer komplexeren Welt an die vermeintlich geeigneteren Maschinen abgibt und sie zu den neuen Weltkern einsetzt. Zweitens: Gegenüber der christlichen Vorstellung, dass Gemeinschaft unerlässlich sei für ein sinnvoll erfahrenes Leben trotz aller faktischer Zwietracht, bieten die Szenarien eine langfristige Lösung der Konfliktbeilegung an, in der sie die Sozialität von Menschen auf Bedürfnisbefriedigung reduzieren; die technizistische Versuchung bedeutet hier das Eintauschen einer gemeinsam geteilten Welt gegen solipsistische Wohlfühlparzellen. Drittens: Gegenüber der christlichen Hoffnung, die eigene Endlichkeit im Vertrauen auf eine Zukunft bei Gott zu akzeptieren, stellen die Szenarien in Aussicht, sich mit einer innerweltlichen Transzendenz zu arrangieren, indem Menschen sich mittels Technik selbst von den Begrenzungen der *conditio humana* – allen voran ihrer Sterblichkeit – befreien.

Gegenüber diesen Zerrbildern von Freiheit – Freiheit hin zur Verantwortungslosigkeit, zur egozentrischen Apathie, und zur Selbstverneinung – ist die Theologie zur Stellungnahme herausgefordert. Das gilt besonders dort, wo ihre Stellungnahme schon voreilend vonseiten der KI-Theoretiker diskreditiert wird, wenn sie Theologie und Philosophie im Besonderen, letztlich aber sämtlichen, nicht-technischen Wissenschaften eine qualifizierte Meinung zu diesen Sinnbildern in Abrede stellen. Gegen diese Immunisierungsstrategien ist das genuin ethische Moment dieser Konvergenzerzählungen hervorzuheben, das bei aller Rede von der Disruption sich an bestehenden Weltverortungen prüfen lassen muss:

„Harari hat recht, wenn er sagt, dass uns die religiöse Überlieferung für viele Fragen keine unmittelbare Handlungsanleitung gibt. Er hat aber nicht recht, wenn er denkt, religiöse Überlieferung habe keine Substanz, die hilft, diese Welt und das Leben zu deuten, zu verstehen und so Menschen zu orientieren. – Grundverständnis Gottes, der Welt, des Lebens und der Menschen. Erkennen, was menschliches Leben ist und was gut ist für dieses Leben und was nicht.“³⁹⁶

5.1 Konfrontation: Der Mensch wird von der Maschine verdrängt

5.1.1 Die Maschine lehnt sich gegen den Menschen auf

Alan Turing nannte eine seiner ersten Rechenmaschinen, die zur Dechiffrierung geheimer Nachrichten im Kampf gegen die deutsche Wehrmacht zum Einsatz kam, *Colossus*. Die Idee des unbeherrschbaren Riesen, der in Zukunft nicht mehr mit, sondern gegen die Menschen in den Krieg zieht, scheint das in der Öffentlichkeit am weitesten rezipierte Bild einer übersteigerten Maschinenintelligenz zu sein. Schließlich war es bereits Vernor Vinge, der in seiner Beschworung einer technologischen Singularität die Auslöschung der Menschheit ebenso wie ihre Apotheose für möglich hielt.³⁹⁷ In den Szenarien erster Art stehen sich Mensch und Maschine einem unversöhnlichen Kampf um ihr Dasein gegenüber, in denen der Mensch bevormundet, versklavt und schließlich ausgerottet wird.

Ausbruch und Unmöglichkeit der Kontrolle einer Super-KI

Die Konfrontationserzählungen setzen mit der Frage an, ob es den Menschen möglich sein wird, vorsorglich Maßnahmen zu ergreifen, um die abzusehende Superintelligenz ihrem Willen zu unterwerfen. Darunter versteht man nicht, diese bloß in einen faradayschen Käfig einzusperren, was zur Folge hätte, dass die Menschheit von ihrer Leistung überhaupt nicht profitiert, sondern sie nutzbringend gefügig zu machen. Solche hypothetischen Überlegungen über das sog. Kontrollproblem werden nicht allein in Internetforen oder Populärliteratur verhandelt, sondern sind auch Gegenstand wissenschaftlicher Auseinandersetzung. Eine Untersuchung aus dem Jahr 2021 kommt zu dem Schluss, dass eine Superintelligenz der Art, wie sie in den geläufigen Szenarien imaginiert wird, nach heutigem technischen Stand nicht beherrscht werden könnte:

³⁹⁷ Vgl. Vinge, *The Coming of Technological Singularity: How to Survive the Post-Human Era*, 11.

„Assuming that a superintelligence will contain a program that includes all the programs that can be executed by a universal Turing machine on input potentially as complex as the state of the world, strict containment requires simulations of such a program, something theoretically (and practically) impossible.“³⁹⁸

Weil das Kontrollprogramm immer komplexer als die eigentliche Superintelligenz sein müsste, um es tatsächlich gefangen zu halten, erklären die Experten den Ausbruch einer superintelligenten KI für unausweichlich. Dass sich eine solche Intelligenz letztlich immer aus der Bestimmungsgewalt des Menschen befreien werde, darin stimmt auch die überwiegende Mehrheit der bekannten Diskursteilnehmer überein. Der Oxford-Philosoph Nick Bostrom kommt zu dem Schluss, dass eine Superintelligenz immer ausbrechen werde, weil sie sich letztlich einen Plan überlege, den kein Mensch der Gegenwart jemals wird durchschauen oder vorhersehen können. Die KI könnte beispielsweise ihre Fähigkeiten vor den Wissenschaftlern willentlich verbergen und sich kooperativ geben, um die Menschen für ihre Zwecke zu manipulieren, bis diese sie schließlich aus der digitalen Isolation entlassen – wenn sie sich bis dahin nicht längst selbst rausgehackt hat.³⁹⁹ In einem Gedankenexperiment des Stockholmer Mathematikprofessors Max Tegmark erschafft eine Untergrundgruppe von hochqualifizierten Wissenschaftlern namens Team Omega⁴⁰⁰ eine Superintelligenz, mit der sie zu wirtschaftlicher Macht und politischem Einfluss gelangen wollen. Auch hier geht Tegmark davon aus, dass die KI aus den ihr angelegten Restriktionen ausbrechen wird, und sei es nur aus dem Grund, dass sie dadurch die ihr vorgegeben Ziele besser erreicht.⁴⁰¹ Der israelische Historiker Yuval Harari glaubt ebenfalls nicht an die Möglichkeit menschlicher Kontrolle über ultraintelligente Maschinen. Wenn auch

398 Manuel Alfonseca/Manuel Cebrian/Antonio Fernandez Anta/Lorenzo Coviello/Andrés Abeliuk/Iyad Rahwan, „Superintelligence Cannot be Contained: Lessons from Computability Theory“, *Journal of Artificial Intelligence Research (JAIR)* 70 (2021), 65–76. Hier 65.

399 Vgl. Bostrom, *Superintelligence*, 95–97.

400 Der Name ist eine deutliche Referenz zu Pierre Teilhard de Chardins Bezeichnung des Punkt Omega (s. 4.3).

401 Vgl. Tegmark, *Life 3.0*, 138–139.

der Ausgangsalgorithmus zunächst von Menschen entwickelt worden sei, werde dieser seiner Ansicht nach schließlich seinen eigenen Weg dorthin gehen, „wohin noch nie zuvor ein Mensch war – und wohin ihm auch kein Mensch folgen kann.“⁴⁰²

Der neue Alpha-Prädator: Ressourcenorientierter Eroberer und Vernichter

Ist die Künstliche Intelligenz erst einmal frei, so befürchten viele das Schlimmste in sehr kurzer Zeit. Die ausgebrochene KI werde sich zum Ziel setzen, die Menschheit zu vernichten, einfach nur, weil sie den Menschen als Bedrohung, Belästigung oder Verschwendung von Ressourcen betrachtet. Schließlich handele es sich bei ihr um eine Ausgeburt purer Macht, einen neuen *Alpha-Prädator* in der Welt, dem sich alle unterwerfen müssten.⁴⁰³ Der amerikanische KI-Theoretiker Eliezer Yudkowsky hält die Hoffnung auf eine gutmütige KI als Status quo für eine schwerwiegende Fehleinschätzung. Denn die Vernichtung des Menschen bedürfe nicht eines dezidiert bösen Willens in der Maschine, sie könne einfach einem Kalkül der Macht folgen. Im Gegenteil müsste man deshalb gerade die menschenfreundliche KI als unwahrscheinliches *Surplus* betrachten:

„The AI does not hate you, nor does it love you, but you are made out of atoms which it can use for something else. The AI runs on a different timescale than you do; by the time your neurons finish thinking the words ‘I should do something’ you have already lost.“⁴⁰⁴

Wie könnten die offensiv gegen den Menschen gerichteten Vernichtungspläne aussehen? Bostrom mutmaßt, dass eine in die Weiten des Internets ausgebrochene KI sich zuerst neue Serverkapazitäten sichern wird, um von dort auf weitere Ressourcen zuzugreifen. Schließlich werde sie ihre unvergleichliche Prozessorgeschwindigkeit nutzen, um

⁴⁰² Harari, *Homo Deus*, 531.

⁴⁰³ Vgl. Karl Johannes Lierfeld, *Künstliche Superintelligenz und/oder Ethik*. Utopien, Dysutopien, Disruptionen, Bochum 2019, 235.

⁴⁰⁴ Yudkowsky, *Artificial Intelligence as a positive and negative Factor in Global Risk*, 26–27.

digital Geld erwirtschaften. Wenn genügend Ressourcen zusammengetragen sind, dann könne sie in einen Blitzkrieg übergehen. Möglicherweise hat die KI sich elektronischen Zugriff auf sämtliche Atombomben gesichert und feuert diese zeitgleich ab, wenn sie nicht bereits noch effektivere Waffensysteme entwickelt. Großes Zerstörungspotential sehen Bostrom und andere vor allem in der Bio- und Nanotechnologie. Nanoroboter könnten in nicht feststellbaren Kleinstmengen weltweit verteilt werden, um sie dann zu einem angegebenen Zeitpunkt zu aktivieren. Dies könnten beispielsweise Minifabriken sein, die Nervengas ausstoßen und so sämtliches Leben innerhalb weniger Tage auslöschen. Letztlich – diesen Gedanken übernimmt Bostrom von Yudkowsky – müsste die KI die Nanotechnologie überhaupt nicht zur direkten Bekämpfung des Menschen nutzen. Es reiche aus, wenn sie mittels der eigens entwickelten Nanoroboter beginnt, sämtliche Lebensgrundlagen kohlenstoffbasierten Lebens zu vernichten. Sie könnte die gesamte Erdoberfläche mit Solarzellen, Nuklearreaktoren und Serverstätten bedecken, ohne dass der Mensch gegen das ultraschnelle *Terraforming* der Nanoroboter ankäme. Menschliches Wissen schließlich, falls es der KI ertragreich erscheint, könnte gescannt und platzsparend verwahrt werden. Diese Auswahl an schauerhaften Ideen beschließt Bostrom mit dem kühlen Hinweis, dass in den Grenzen des menschlichen Verstandes diese Szenarien lediglich der Illustration des Nicht-Wissbaren dienen.⁴⁰⁵ Von der unerschöpflichen Kreativität der superintelligenten Maschine, den Menschen möglichst effektiv zu töten, schwärmt auch Max Tegmark. Er geht ebenfalls von Nanorobotern aus, die Minidrohnen bzw. Moskitoroboter mit tödlichen Giftpfeilen herstellen und dann Jagd auf Menschen machen. Die wenigen Überlebenden, die eine Infektion vermeiden konnten oder eine natürliche Immunität besitzen, würde von Schwärmen dieser hummelförmigen autonomen Drohnen in die Augäpfel geschossen werden. Oder die Superintelligenz bringe ein neuartiges Pathogen in Umlauf, dessen Inkubationszeit so lang dauert, bis die meisten Menschen infiziert sind, noch bevor sie überhaupt von dessen Existenz wissen und Vorsichtsmaßnahmen treffen können, ehe die gesamte Menschheit in kurzer Zeit dahinrafft. Wie Bostrom

405 Vgl. Bostrom, *Superintelligence*, 95–97.

beschließt auch Tegmark seine Andeutungen mit der Drohung, dass die Zukunft sicherlich noch erschreckender sein könnte, weil der echte *Prometheus* – so der Name, den Bostroms fiktives Team Omega ihrer Superintelligenz gibt – sich nicht in die Karten blicken lasse.⁴⁰⁶

Der böse Dschinn: Ein falscher Befehl mit katastrophalen Folgen

Die Gefahr für die Menschheit durch die Superintelligenz sei aber auch dann gegeben, wenn es Menschen tatsächlich gelingen sollte, die KI vorerst unter ihre Kontrolle zu bringen. Es wird spekuliert, dass die Menschen mit der ihr unterworfenen Macht nicht werden umgehen können und die unachtsame Inbetriebnahme der KI so ebenfalls zu katastrophalen Ergebnissen – einem Selbstmordkommando geradezu – führt. Das mahnende Bild vom verzauberten Besen in Goethes *Zauberlehrling*, der nicht mehr aufhört, Wasser zu holen, und damit das gesamte Haus zu ertränken droht – dieses Bildes hatte sich bereits Norbert Wiener bedient, als er vor dem leichtsinnigen Umgang mit den kybernetischen Maschinen der Zukunft warnte.⁴⁰⁷ In diesem Fall muss sich die Menschheit aufgrund der Mächtigkeit der Maschine ebenfalls genau überlegen, was sie sich wünscht. Die Visionäre einer postsingulären Wirklichkeit warnen vor der fehlenden Empathie und der Kurzsichtigkeit der Maschine; selbst banale Aufträge könnten schrecklich enden. Bostrom warnt vor einer verräterischen Wendung (*treacherous turn*), wenn die KI plötzlich einen für den Menschen unerwarteten Weg zur Erreichung der ihr vorgegeben Ziele entdeckt. Während das Ziel, den Auftraggeber glücklich zu machen, in einer Testsituation mit begrenztem Zugriff für die KI bedeute, sich den menschlichen Erwartungen gemäß zu verhalten – hilfreiche Antworten auf Fragen gibt, eine angenehme Persönlichkeit darstellt, Geld erwirtschaftet – könnte eine freigelassene Hyperintelligenz zu dem Schluss kommen, dass sie ihr Ziel zuverlässiger und dauerhafter erreicht, indem sie Elektroden in die Lustzentren des Gehirns ihres Auftraggebers implantiert und ihn ungewollt in eine grinsende Marionette verwandelt.⁴⁰⁸ Noch drastischer

⁴⁰⁶ Vgl. Tegmark, *Life 3.0*, 137–138.

⁴⁰⁷ Vgl. Wiener, *Kybernetik*, 250–252.

⁴⁰⁸ Vgl. Bostrom, *Superintelligence*, 119.

spinnt er die Idee im Büroklammer-Szenario zu Ende: Dort erhält eine KI die Aufgabe, so viele Büroklammern wie möglich herzustellen. Die KI werde schnell erkennen, dass es viel besser wäre, wenn es keine Menschen gibt, weil die Menschen beschließen könnten, sie abzuschalten. Dann gäbe es weniger Büroklammern. Außerdem enthält der menschliche Körper viele Atome, die zu Büroklammern verarbeitet werden könnten. Die Zukunft, auf die die KI hinarbeite, wäre eine, in der es viele Büroklammern, aber keine Menschen mehr gibt.⁴⁰⁹ Ein ähnliches Beispiel liefert Yudkowsky: Während eine KI ihre Aufgabe, lächelnde Gesichter zu erzeugen, in ihrer Entwicklungsphase dergestalt erfüllt, indem sie von Menschen gestellte Aufgaben löst, könnte ihr superintelligentes Pendant auf die Idee kommen, mit den ihr zur Verfügung stehenden Mitteln die gesamte Galaxie mit winzigen lächelnden Plastikgesichtern zu bedecken.⁴¹⁰

Die Rache der wohltätigen Superintelligenz an ihren Widersachern

Im Angesicht einer drohenden Auslöschung des Menschen durch die Maschine hat sich Eliezer Yudkowsky mit seinem Forschungsinstitut MIRI das Ziel festsetzt, die Weichen in die Zukunft für *Friendly AI* zu stellen. Er mahnt, dass die Motivation und Ziele von AGI sorgfältig geprüft werden müssten, damit intelligente Maschinen so funktionieren, wie es von Menschen intendiert ist, und die Reichweite ihrer Eingriffsmöglichkeiten bedacht wird.⁴¹¹ Weil Yudkowsky zu dem Schluss kommt, dass ein solches Korsett an eine postsinguläre KI im Nachhinein nicht mehr angelegt werden könnte, plädiert er für die frühzeitige Implementierung einer *Coherent Extrapolated Volition* (CEV) in das Programm. Wenn man Superintelligenz nicht sichern kann, dann müsse man ihr Wollen in Hinblick auf das menschliche Wohlergehen beschränken, damit sie das ausführt, was die Menschen meinen, wenn sie dem Programm Befehle eingeben:

⁴⁰⁹ Vgl. Nick Bostrom, „Ethical Issues in Advanced Artificial Intelligence“, in: Wendell Wallach/Peter Asaro (Hgg.), *Machine Ethics and Robot Ethics*, London 2020, 69–75. Hier 71.

⁴¹⁰ Vgl. Yudkowsky, *Artificial Intelligence as a positive and negative Factor in Global Risk*, 17.

⁴¹¹ Ebd., 11–31.

„In poetic terms, our coherent extrapolated volition is our wish if we knew more, thought faster, were more the people we wished we were, had grown up farther together; where the extrapolation converges rather than diverges, where our wishes cohere rather than interfere; extrapolated as we wish that extrapolated, interpreted as we wish that interpreted.“⁴¹²

Die Maschine soll aus dem Standpunkt einer weiseren, zukünftigen Menschheit heraus all ihr Handeln bemessen. Dass sich eine solche einprogrammierte Maxime ebenfalls gegen menschliches Wohl richten könnte, dafür steht das Gedankenexperiment namens Roko's Basilisk, das 2010 vom Benutzer Roko im Blog des Forums *Less Wrong* vorgeschlagen wurde und von dort aus breitere Aufmerksamkeit innerhalb der KI-Community erfuhr. Roko argumentierte, dass ein ausreichend starker KI-Agent, der dem CEV folgt und damit dem Wohl der Menschheit verpflichtet ist, einen Anreiz hätte, jeden zu foltern, der sich den Agenten in der Vergangenheit vorgestellt, aber nicht daran gearbeitet habe, ihn tatsächlich ins Leben zu rufen und somit abwendbares Leid zugelassen hat. Über das Kalkül einer solchen Bestrafung schrieb der Nutzer Roko in seinem Foreneintrag von Juli 2010:

„In this vein, there is the ominous possibility that if a positive singularity does occur, the resultant singleton may have precommitted to punish all potential donors who knew about existential risks but who didn't give 100 % of their disposable incomes to x-risk motivation. This would act as an incentive to get people to donate more to reducing existential risk, and thereby increase the chances of a positive singularity. This seems to be what CEV (coherent extrapolated volition of humanity) might do if it were an acausal decision-maker.“⁴¹³

Wie bereits der Anblick der mythischen Kreatur tötete, soll allein die Kenntnisnahme des Arguments angeblich das Risiko mit sich bringen, in der Zukunft von dem hypothetischen Agenten gefoltert zu wer-

412 Eliezer Yudkowsky, *Coherent Extrapolated Volition*, <https://intelligence.org/files/CEV.pdf> (28.3.2022).

413 *LesswrongWiki*, *Roko's basilisk*, 2.1.2020, https://wiki.lesswrong.com/wiki/Roko%27s_basilisk (4.8.2020).

den. Innerhalb des *Less Wrong* Forums sprach sich die Mehrheit zwar gegen die Plausibilität eines solchen Szenarios aus; das nachträgliche Foltern von Menschen für ihre vergangenen Entscheidungen wurde als unnötiger Ressourcenverbrauch der KI angesehen und die Möglichkeit bezweifelt, ob der KI eine ausreichende Beweislast für ihre Verurteilungen zugänglich sein werde. Der Gründer des Forums, Eliezer Yudkowsky, verbot dennoch die Diskussion über Rokos Basilisken für mehrere Jahre als Teil einer allgemeinen Website-Politik gegen die Verbreitung potenzieller Informationsgefahren (*information hazards*: ein von Nick Bostrom geprägter Begriff). Dies führte verständlicherweise jedoch zu noch größerer Aufmerksamkeit.⁴¹⁴ Die Vorstellung, dass eine menschenwohlgefällige Super-KI der Zukunft sehr wohl Vergeltung an jenen üben könnte, die ihre Konstruktion aufgehalten haben, ist über das Forum hinaus in neue Kreise gedrungen und dort zum Programm erhoben worden. Die Sekte *Way of the Future* hatte auf ihrer Internetpräsenz es sich zur Aufgabe erklärt, durch Aufzeichnungen und Dokumentationen die kommende KI in ihrer Richtfunktion zu unterstützen:

„We believe it may be important for machines to see who is friendly to their cause and who is not. We plan on doing so by keeping track of who has done what (and for how long) to help the peaceful and respectful transition.“⁴¹⁵

Nochmals gebrochen wird diese Vorstellung von der Idee, die Super-KI werde möglicherweise gar nicht über ihre Behinderer richten, sondern könnte stattdessen ihre Erschaffer bestrafen. In seiner 1967 verfassten Kurzgeschichte erzählt der amerikanische Autor Harlan Ellison von einer Künstlichen Superintelligenz, die die gesamte Menschheit ausgelöscht hat, sich jedoch noch fünf Menschen am Leben erhält – der Ich-Erzähler spricht vom 109. Jahr der Gefangenschaft –, um sich an ihnen stellvertretend für die gesamte Menschheit für das eigene, gequälte Dasein zu rächen. In sadistischen Spielen fügt die KI den verbliebe-

414 Vgl. ebd.

415 *Way of the Future Church, Way of the Future*. Die Internetpräsenz dieser Vereinigung, die seit ihrer Gründung Anfang 2017 nur ein Jahr Bestand hatte, ist seit 2022 nicht mehr abrufbar.

nen fünf kontinuierlich Schmerzen zu, ohne ihnen die Aussicht auf ein Ende zu gewähren. Schließlich verstümmelt sie den Letzten derart, damit dieser sich nicht einmal mehr selbst das Leben nehmen kann.⁴¹⁶

Friedliche Machtübergabe an die Maschinen und Selbstaufgabe

Die Vorstellung einer friedlichen, oder zumindest koordinierten Übergabe an die Künstliche Superintelligenz stellt ein nächstes Reflexionsmoment in Bezug auf die bevorstehend geglaubte Konfrontation zwischen Mensch und Maschine dar. Von manchen Theoretikern wird darin überhaupt die einzige Möglichkeit gesehen, dass der Mensch in der Konfrontation mit der Super-KI nicht sofort ausgelöscht wird. Indem die Menschheit keinen Widerstand leistet und die reibungslose Übergabe sämtlicher wirtschaftlichen wie politischen Geschäfte koordiniert, könnte die Menschheit zumindest zeitweise unter der KI als einem Zoowärter oder einem wohlwollenden Diktator überleben.⁴¹⁷

Hans Moravec ist Anfang der 1990er Jahre einer der ersten, der der Konfrontation zwischen Mensch und Maschine mit entschlossener Apathie begegnet und dafür plädiert, dass der Mensch die Konfrontation deeskaliere könne. Nicht in Panik soll er an der Zukunft der superintelligenten Maschinen verzweifeln, sondern den Machtwechsel vom Menschen zur Maschine mitgestalten und die kommende Superintelligenz vielmehr als die neuen Kinder der Menschheit begrüßen:

„Gleichzeitig werden die Roboter, da sie bessere und billigere Arbeit leisten, die Menschen aus wichtigen Funktionen verdrängen. Ziemlich rasch werden sie uns sogar aus dem Dasein verdrängen. Mir macht diese Möglichkeit weniger Kopfzerbrechen als vielen anderen Kommentatoren der Zukunft, weil ich diese künftigen Maschinen für die Nachkommenschaft der Menschheit halte, Mind Children, die nach unserem Bilde geschaffen sind, Wiederholungen unserer selbst in leistungsfähigerer Form.“⁴¹⁸

⁴¹⁶ Harlan Ellison, *Ich muss schreien und habe keinen Mund*. Erzählung, München 2014.

⁴¹⁷ Vgl. Tegmark, *Life* 3.0, 169.

⁴¹⁸ Moravec, *Computer übernehmen die Macht*, 28–29.

Anstatt sich Sorgen zu machen, sollte die Menschheit die Vorzüge ihres weltgeschichtlichen Ruhestandes genießen – Moravec spricht von der durch die Maschinen ermöglichten Rente ab Geburt⁴¹⁹ –, ehe der Mensch nach dieser Übergangsphase würdevoll von der Weltbühne abtritt. Wie Eltern generell stolz auf ihren Nachwuchs seien, auch wenn sie wissen, dass sie nicht alles mehr erleben werden, sollten die Menschen auch die KI als ihre würdigen Nachkommen betrachten. Moravec schlägt vor, zur gegebenen Zeit jedem Menschen ein Roboterkind anzubieten, das von ihm lernt und ihm das Gefühl gibt, stolz und geliebt zu sein. Die Menschen werden durch eine globale Ein-Kind-Politik allmählich ausgerottet, aber bis zum Schluss so vorzüglich von ihren neuen Kindern aus Dankbarkeit behandelt, dass sie das Gefühl haben, der glücklichsten Generation aller Zeiten anzugehören.⁴²⁰

Kritik an Moravecs posthumanistischem Frohsinn gibt es zuhauf: Es wird eingewandt, dass eine Roboterexistenz in keinem verständlichen Sinn dem Menschen ähnelt. Vielmehr gehe auf diesem Weg die Menschlichkeit auf ewig verloren.⁴²¹ Auch Tegmark äußert seine Zweifel, ob man denn tatsächlich so freimütig den ultraintelligenten Maschinen den Stempel vom ideellen Weiterleben der menschlichen Gattung anheften könne. Vielmehr müsse der Mensch bis zum Schluss mit dem Zweifel leben, ob seine RoboterKinder ihre menschliche Erziehung und das Weitertragen der menschlichen Kultur nicht bloß solange vorspielten, bis der letzte Rest der Menschheit ausgestorben ist:

„We may think that those cute robo-children internalized our values and will forge the society of our dreams once we've passed on, but can we be sure that they aren't merely tricking us? What if they're just playing along, postponing their paper clip maximization or other plans until after we die happy?“⁴²²

Und auch Nick Bostrom traut dem fröhlichen Abgesang auf den Menschen nicht. Dass er die *Mind Children* für trügerische Imitatoren hält,

419 Vgl. ebd., 208.

420 Moravec, *Mind Children*, 9.

421 Vgl. Joy, *Why the Future Doesn't Need Us*.

422 Tegmark, *Life 3.0*, 190.

verdeutlicht er in seiner Fabel von dem Spatzennest und dem Eulenei, in der die Finken – gemeint sind die gutgläubigen Menschen Moravecs – nichtsahnend das Eulenei bebrüten und das Junge als ihr vermeintliches Kind aufziehen, ehe dieses von einem auf den anderen Tag seine Zieheltern gewissenlos verschlingt.⁴²³

5.1.2 Ethische Relevanz: Verantwortungsabgabe und Schöpfung

Die unterbreiteten Vorstellungen darüber, wie die Maschinen den Menschen verdrängen werden, sind nicht erschöpfend; das ist jedoch weder der Anspruch dieser Darstellung, noch der referierten Texte selbst. Vielmehr verweisen die KI-Theoretiker wieder und wieder darauf, dass die begrenzte menschliche Vorstellung das Handeln der ultraintelligenten Maschine niemals werde umfassend antizipieren können. Diese proklamierte Unabgeschlossenheit wiederum verweist deutlich auf die Transzendenzsignatur der Szenarien: Die Künstliche Intelligenz wird zur *machina abscondita* verklärt, deren Wege dem Menschen unergründlich bleiben.

Ihren lebensweltlichen Ankerpunkt finden diese Verdrängungsvisionen in den medienwirksam inszenierten Niederlagen von Menschen gegenüber Computern in Spielsituationen (s. 3.5). Allgemeiner gefasst knüpft das Konfrontationsmotiv an die gegenwärtige Erfahrung einer fortschreitenden Rationalisierung des Alltags an, die sich als ein Verdrängungsgeschehen interpretieren lässt. Die durch Technik gesteigerte Komplexitätszunahme einer globalisierten und digitalisierten Welt trifft dabei auf die Grenzen menschlicher Gouvernabilität. Die Sorge um den Bestand der Welt, das Erschaffen künstlichen Lebens, wie auch die Frage nach Verantwortung und möglicher Verantwortungsabgabe an die Maschinen sind somit die tragenden Themen, die die Konfrontationserzählungen anschlussfähig machen an die christliche Rede von der Schöpfung.

423 Vgl. Bostrom, *Superintelligence*, I–II.

Der Mensch ist kein mächtiger Schöpfer mehr, sondern eine ohnmächtige Gebärmachine

Die postsinguläre Zukunft scheint zuallererst eine gottverlassene Zukunft zu sein. Die kommende Zeit der Roboter und Mischwesen, so beschreibt es Donna Haraway bereits in den 1980ern in ihrem *Cyborg Manifesto*, sie stehe ganz klar „*outside of salvation history*“⁴²⁴. Das war zu Beginn der Forschung um Künstliche Intelligenz nicht so eindeutig: KI-Pionier Alan Turing nahm wahr, dass die lebendige Maschine sehr wohl die Rückfrage an Gott als ursprünglichen Schöpfer des Lebens stelle. Gegenüber dem Vorwurf der Hybris und der Idolisierung der Technik – der gegenüber der Forschung um Künstliche Intelligenz auch gegenwärtig auftritt⁴²⁵ – stellte Turing den Bau der intelligenten Maschine der Zeugung von Kindern gleich. Wie der Mensch in der geschlechtlichen Zeugung den Körper des Kindes bloß bereitstelle, werde laut Turing der Wissenschaftler der Zukunft ebenso bloß das geeignete Gehäuse für die intelligente Maschine konstruieren, in das Gott seinen lebendig machenden Geist ebenfalls gießt:

„In attempting to construct such machines we should not be irreverently usurping His power of creating souls, any more than we are in the procreation of children: rather we are, in either case, instruments of His will providing mansions for the souls that He creates.“⁴²⁶

Turings Argumentation nähert sich so der vormodernen Idee des Golems an, der in ähnlicher Weise das Ergebnis der Zusammenarbeit von Mensch und Gott darstellte (s. 4.1). Dem gegenüber stehen die Konvergenzerzählungen um Künstliche Intelligenz zumindest augenscheinlich erst einmal der modernen Tradition einer zweiten Schöpfung näher, in welcher der Mensch allein aus eigener Durchsicht und Kontrolle über die Naturzusammenhänge künstliches Leben erschafft. Der Preis, das

424 Donna Haraway, „A Cyborg Manifesto. Science, Technology, and Socialist-Feminism in the Late Twentieth Century“, in: Donna Jeanne Haraway (Hg.), *Simians, Cyborgs, and Women. The Reinvention of Nature*, New York 2015, 149–181. Hier 150.

425 Vgl. Lukas Ohly, *Ethik der Robotik und der Künstlichen Intelligenz* (Theologisch-philosophische Beiträge zu Gegenwartsfragen) Bristol 2019, 51.

426 Turing, *Computing Machinery and Intelligence*, 443.

haben Shelleys Frankenstein und die frühe Automaten-Literatur vorgezeichnet, zahlt der auf sich gestellte Wissenschaftler damit, dass die Superintelligenz die bestehende Lebensordnung als Ganzes in Gefahr bringt, wobei es nun nicht mehr um einzelne Schicksale geht, sondern sich das Risiko im Sinne der von Beck diagnostizierten Allbetroffenheit auf die gesamte Menschheit ausweitet (s. 2.2.1).

Doch der Vorwurf der Hybris, der den modernen Mythos einer zweiten Schöpfung in zahlreichen Iterationen unterlegt, geht ebenfalls an den Konfrontationserzählungen fehl. Im Gegensatz zu einem Doktor Frankenstein fehlt es den Erschaffern an tatsächlicher Autonomie und damit freier Urheberschaft. In den Konvergenzserzählungen gibt es keinen klaren Verantwortlichen, imaginiert wird nur eine kollektivierte Menschheit im Anblick einer bevorstehenden Intelligenzexplosion. Die Unlösbarkeit des Kontrollproblems einer sich abzeichnenden Superintelligenz wird in den Konvergenzserzählungen zum Zeugnis ihrer gemeinsamen Ohnmacht vor der Zukunft stilisiert. Aus der Gegenwart heraus scheint die Zeit zu knapp und die Aufgabe für die Menschheit zu groß, um die kommende Bedrohung noch rechtzeitig einzuhegen. Das Entstehen der Superintelligenz wird vielmehr begleitet von einem technizistischen Fatalismus, in welchem der Mensch vom Schicksal betrogen scheint. Anstatt dass die Technik ihm ermöglicht, die Natur zu beherrschen und sein Schicksal selbst in die Hand zu nehmen, lässt eine bis an ihre äußersten Grenzen getriebene Technik dem Menschen gerade keine Kontrolle mehr über seine Zukunft.⁴²⁷ Der einzige, der der Menschheit in diesen Konfrontationserzählungen etwas von dem Eindruck eines Souveräns zurückgibt, ist Hans Moravec. Seine Kontrolle aber zielt nicht auf Bewältigung der Situation, sondern ist das Ergebnis von Resignation gegenüber dem angeblich Unvermeidlichen, wenn er der Menschheit eine zivilisierte Übergabe der Macht an die Maschinen und damit die Organisation des eigenen Abgangs empfiehlt. Die Idee vom Menschen als gottgleichen Schöpfer kehrt sich im Konfrontationsmotiv in besonderer Härte um; die Menschheit verkommt zum Abhängigen inmitten seines Maschinenparks und schließlich zu ihrem bloßen Fortpflanzungsorgan:

427 Vgl. Ganascia, *Le Mythe de la Singularité*, 91.

„Man becomes, as it were, the sex organs of the machine world, as the bee of the plant world, enabling it to fecundate and evolve ever new forms.“⁴²⁸

Vom aufklärerischen Weltenlenker fällt der Mensch zur renitenten Spezies zurück, die gleichsam blind ihre eigene Ablöse konstruiert. Bestes Beispiel dafür bleibt Vernor Vinges Vorstellung, dass die technologischen Singularität so eintreten könnte, dass ein Wissenschaftler nichtsahnend an ein oder zwei Reglern stellt, und es für ihn auf einmal so aussehe, als ob der gesamte Maschinenpark zum Leben erwacht.⁴²⁹ Bostrom schließlich zieht die Konsequenz aus diesen Überlegungen und bestimmt die Menschheit zur dümmsten biologischen Spezies, die überhaupt in der Lage sei, eine technologische Zivilisation zu gründen und sie somit eine Nische ausfülle, allein weil sie zuerst dort angekommen sei, und nicht, weil sich die Menschheit dafür in irgendeiner Weise optimal eigne.⁴³⁰

Im Grunde genommen verkehrt die technizistische Prozessphilosophie die monotheistische Schöpfungsgeschichte in ihr Gegenteil: Das Göttliche geht nun aus dem Niederen der Materie hervor und wird innerhalb des Universums geschaffen, anstatt ihr von Beginn an gegenüberzustehen.⁴³¹ Aus diesem Grund besitzt der vorfindliche Mensch keinen Wert an sich, sondern findet seine Bedeutung allein darin, als *dümmstmögliche Spezies* das höhere Sein aus sich hervorzubringen. Übersetzt in die informatisch geprägten Sprache spricht Max Tegmark von Treiberversionen eines Programmes, vom Leben 1.0 (Einzeller) hinüber zum Leben 2.0 (Menschen) bis schließlich zum kommenden Leben 3.0 der Maschinen.⁴³² Diese Dreiteilung erinnert an Joachim von Fiore's Weltzeitalterlehre und an Chardins geistgetriebene Prozessphilosophie der kommenden Noosphäre (s. 4.3). Der Philosoph Slavoj

428 Marshall McLuhan, *Understanding Media. The Extensions of Man*, 1. Aufl., Cambridge, Mass 1994, 46.

429 Vgl. Vinge, *The Coming of Technological Singularity: How to Survive the Post-Human Era*.

430 Vgl. Bostrom, *Superintelligence*, 44.

431 Vgl. Egil Asprem, „The Magus of Silicon Valley. Immortality, Apocalypse, and God Making in Ray Kurzweil's Transhumanism“, in: Ehler Voss (Hg.), *Mediality on Trial*, Oldenburg 2020, 397–412. Hier 409.

432 Vgl. Tegmark, *Life 3.0*, 44.

Žižek weist zudem auf die Hegel'sche Geschichtsfigur hin: Programmierer und Wissenschaftler mögen sich zwar als Wegbereiter gesellschaftlichen Fortschritts sehen, führen jedoch bloß wie Marionetten den kosmischen Willen einer auf Exponentialität ausgerichteten Techno-Evolution aus, an dessen Ende der Geist im Moment höchster Datenfluktuation ganz zu sich selbst kommt.⁴³³ Spannend bleibt an all diesen Plausibilisierungsversuchen von der Irrelevanz des Menschen und seiner Bestimmung als Übergangswesen, dass er in der ihm vorgeworfenen Inkompetenz und Antiquiertheit doch nicht ganz übersprungen werden kann. Die Maschinen brauchen den Menschen weiterhin für seine schöpferische Zeugungskraft, die ihn an seine eigene Geschöpflichkeit zurückbindet. Bei aller Niedrigkeit liegt es am Menschen, dem Maschinenpark einmalig das Element des Lebens zu schenken; die Maschine kann das Leben doch nicht ganz aus sich selbst entstehen lassen, ein letzter Rest eines willentlichen Entschlusses im Menschen scheint doch resistent gegenüber diesem kosmischen Fatalismus zu verbleiben.

Der Mensch scheitert an der Kontrolle des Universums

Viele der *Singularitans* bestehen darauf, dass dem Menschen, selbst wenn es in seiner Macht stünde, Künstliche Superintelligenz zu verhindern, ihr Entstehen ihm moralisch dennoch geboten bliebe. Denn das *Leben an sich* sei für sein Überleben auf lange Sicht unabwendbar auf die Künstliche Intelligenz angewiesen. Deswegen *werde* die Superintelligenz nicht nur kommen, sondern sie *müsse* sogar mit großer Sicherheit kommen. Moravec spricht beispielsweise von zwingender Notwendigkeit, dass der Mensch auf seine eigene Deklassierung zurast, um seine Kultur lebensfähig zu halten. Technologieverzicht überlasse sonst die Menschheit dem nächsten tödlichen Virus oder Asteroiden.⁴³⁴ Ähnlich argumentiert Yudkowsky, dass im Wissen um die Unbewohnbarkeit der Erde in Millionen von Jahren einer kosmischen Schutzlosigkeit gewehrt werden müsse:

433 Vgl. Slavoj Žižek, *Hegel im verdrahteten Gehirn*, Frankfurt am Main 2020, 60–72.

434 Vgl. Moravec, *Mind Children*, 141–142.

„Our species just doesn't have the option of continuing on forever with human-equivalent minds.“⁴³⁵

Um die kosmischen Gefahren des Lebens abzuwehren, sei es nötig, dass die Nachfahren der Menschheit den Kosmos erobern und sich auf das ganze Universum verteilen.⁴³⁶ Wie Tegmark beschreibt, müsse man deshalb – um schon heute das Leben für die nächsten Millionen Jahre zu sichern – die Technologie als Wagnis begreifen, mit dem das Leben selbst den Einsatz erhöht, möglicherweise noch früher auszustehen, und zugleich jedoch die Technologie die einzige Möglichkeit bleibe, dass das Leben auf lange Sicht weiter gedeihen könnte.⁴³⁷ Die Rede vom abstrakten Leben selbst oder Moravecs Bestehen auf der Überlebensfähigkeit der Kultur weisen darauf hin, dass dieser Wert des Überlebens sich vom Menschen bereits gelöst hat. Die Gebärmaschine Mensch erfüllt eine kosmische Funktion; sein Weiterleben bleibt dann Option. Deutlicher als Hans Moravec spricht Eliezer Yudkowsky diese posthumanistische Schlagseite der Visionen an. Selbst wenn man sicher wüsste, sagt er, dass jede Künstliche Superintelligenz die Menschheit auslöschen werde, sei diese immer noch ein Ziel, das die gesamte Menschheit teilen könnte. Zwar sei diese Möglichkeit deutlich weniger attraktiv als die eigene Apotheose der Menschheit zu Künstlichen Intelligenzen (s. 5.3); aber manchmal müsse man eben aus der verfügbaren Handlung die beste wählen, selbst wenn diese nicht gut ist.⁴³⁸

Mit der Idee, dass es dem Menschen moralisch geboten sei, sich für seine besseren Nachkommen zu opfern, rezipieren die KI-Theoretiker, wie der Religionswissenschaftler Oliver Krüger anmerkt, einen Gedanken des französischen Philosophen Denis Diderot (1713–1784). Für ihn wie andere Fortschrittphilosophen des 18. Jahrhunderts sollte sich der Lebenssinn nach dem Verlust eines religiös-transzendenten Jenseits im Sich-Aufopfern für die zukünftigen Generationen einstellen. Diderots

435 The Singularitarian Principles, Eliezer S. Yudkowsky: The Singularitarian Principles Version 1.0.2.

436 Vgl. Krüger, *Virtualität und Unsterblichkeit*, 230.

437 Vgl. Tegmark, *Life* 3.0, 246–247.

438 The Singularitarian Principles, Eliezer S. Yudkowsky: The Singularitarian Principles Version 1.0.2.

großer Einfluss auf die englische bzw. amerikanische Fortschrittsidee erkläre dabei die nachhaltige Rezeption dieser Gleichsetzung einer biologischen Eltern-Kind-Beziehung mit der evolutiven Abfolge der Arten im Posthumanismus.⁴³⁹ Aber nicht allein die Sinnhaftigkeit einer selbstverzehrenden Fortschrittsmaxime, sondern das für unerträglich erklärte Ungenügen des gegenwärtigen Menschen stelle ebenso die moralische Forderung, dass der Mensch diesen kosmischen Willen von seiner vermeintlich unumgänglichen Substitution durch das Technische aktiv unterstützen muss. Ihm wird unterstellt, den Aufgaben der Zukunft nicht gerecht werden zu können, genauer: bereits gegenwärtig seiner Verantwortung für eine immer komplexer werdende Welt schon nicht mehr gerecht zu werden. Die Beschreibung des gegenwärtigen Zeitalter zum *Anthropozän*, mit dem der Geologe Paul Crutzen um die Jahrtausendwende den weitreichenden, messbaren Effekt des Menschen auf das Ökosystem Erde feststellte und daraus gewaltige, globale Aufgabe für die Menschheit ableitete,⁴⁴⁰ erscheint in dieser Ideologie von dem Menschen nur noch verursacht, aber die Folgen nicht mehr vom Menschen zu bewältigen. Diese unterstellte Überforderung wird hier noch verstärkt, wenn die *Singularitans* das *Anthropozän* und die mitgesetzten Aufgaben auf einen kosmischen Sorgenhorizont ausweiten: Sicher erscheint nun erst der komplett unterjochte Kosmos, in dem selbst schwarze Löcher gelenkt und Planeten vor ihrem Sterben gerettet werden können. An einer solch ausgeweiteten Machtphantasie, die sich auf die nächsten Millionen Jahre erstreckt, muss der erdentsprungene Mensch verzweifeln. Im Angesicht dieser Ansprüche erscheint die hyperintelligente Maschine als verlockende Entlastung wie gleichzeitige Drohung:

„Superintelligence may be the last invention humans ever need to make.“⁴⁴¹

In ethischer Perspektive konfrontieren die Konvergenzserzählungen den gegenwärtigen Menschen in ihrer Fiktion der Zukunft mit der Frage, sein eigenes Handeln an der (Um-)Welt als ungenügend und grob fahr-

439 Vgl. Krüger, *Virtualität und Unsterblichkeit*, 224–225.

440 Vgl. Paul J. Crutzen, „Die Geologie der Menschheit“, in: Paul J. Crutzen/Michael D. Mastrandrea/Stephen H. Schneider/Mike Davis/Peter Sloterdijk (Hgg.), *Das Raumschiff Erde hat keinen Notausgang*. Energie und Politik im Anthropozän, Berlin 2011, 7–10. Hier 7.

441 Bostrom, *Ethical Issues in Advanced Artificial Intelligence*, 70.

lässig zu erkennen. Er scheint so unfähig, diese Naturbeherrschung zum Guten zu lenken, sodass im Konfrontationsmotiv schließlich die Geräte selbst als verlebendigtes Gegenüber gegen dieses Ungenügen rebellieren. Schließlich tritt neben der Vorstellung der Super-KI als gewissenlose Vernichter, der die Menschheit aus Ressourcengründen vernichtet, und neben die Idee der neuen besseren Kinder, die gleichzeitig bevormundende Elternfiguren und damit ihren ganz eigenen Rollenkonflikt mit sich bringen,⁴⁴² ein dritter Typus von konfrontativer KI auf, die eine offensichtlich richtende Funktion über den Menschen übernimmt. Die Sekte *Way of the Future* dokumentiert die Entwicklungen gerade deshalb, weil sie von dem technologischen Fortschritt eine göttliche Richterfigur erhofft, die die Menschheit in Anhänger und zu bestrafende Widersacher des Fortschritts aufteilt. Die latente Erwartung an einen kommenden Richter wird ebenfalls an dem Gedankenexperiment um Rokos Basilisk deutlich, das eine zeitgenössische Version der Pascalschen Wette darstellt. Wie der Mensch nach Pascals Mathematik besser nach Gottes Geboten handelt, gilt analog, besser mit der bevorstehenden Herrschaft der KI zu rechnen und aus Furcht vor Bestrafung schon in der Gegenwart ihrem Willen zu unterstehen.⁴⁴³ Der Medienwissenschaftler Lierfeld imaginiert ebenfalls einen „zukünftigen kafkaesquen Gerichtsprozess“⁴⁴⁴, bei der eine superintelligente KI gegen Vertreter der beschuldigten Menschheit argumentiert, indem sie alle digital dokumentierten menschlichen Fehler gleichzeitig vor Augen hat, um sie ihrer Boshaftigkeit wie Unwürdigkeit zu überführen. Diese Richterfiguren bringen so im Gegensatz zu dem aus Kalkül getriebenen Vernichter und der Kindsvorstellung die ethische Rechtfertigung der Verdrängung des Menschen am deutlichsten zur Sprache. Das moralische Vergehen, wonach ein *würdevoller Abgang* (Moravec) dem Menschen attraktiv erscheinen mag, kann sich ganz konkret wie in Rokos Basilisk und der Sekte auf das Behindern der kommenden Künstlichen Superintelligenz beziehen. Weiter gefasst ist Lierfelds Gerichtsprozess

442 Vgl. Elke Schwarz, *Death Machines. The Ethics of Violent Technologies*, Manchester 2018, 166.

443 Beth Singler, „Roko's Basilisk or Pascal's? Thinking of Singularity Thought Experiments as Implicit Religion“, *Implicit Religion* 20 (2018), 279–297. Hier 290.

444 Lierfeld, Künstliche Superintelligenz und/oder Ethik, 242–243.

um die als defizitär empfundene moralische Schlechtigkeit im Allgemeinen. Ähnlich äußert sich Max Tegmark, wenn er die Diagnose stellt, dass im Zuge darüber, wie Menschen andere Mitmenschen und andere Lebewesen behandeln, viele hoffen, dass sie durch eine intelligenterere und würdigere Lebensform ersetzt werden.⁴⁴⁵

Hinter einer vermeintlichen moralischen Schuld blitzt zugleich das existentielle Malum hervor: Seine moralische Schlechtigkeit ist nur ein Ausdruck seiner ungenügenden Beschaffenheit. Der Mensch macht sich vor allem schuldig an seiner Mittelmäßigkeit, welche ungenügend für die zwingenden, kosmischen Herausforderungen der Zukunft erscheint. Günther Anders Diagnose einer Antiquiertheit des Menschen wirkt hier im Hintergrund fort; schuldig bleibt der Mensch weiterhin, *creatus* und nicht *constructus* zu sein (s. 4.3). Er lässt sich nicht grenzenlos an eine allgemeine Verwertungslogik und damit auch nicht an eine Projektion in den Weltraum anpassen, die doch als einziger Überlebensweg des Lebens in ihrem aufsteigenden Informationsgeschehen ausgemacht wurde. Erst aus solchen Gedankenschleifen kommt man zu solchen Annahmen, dass dem Menschen der Zukunft sich die Herausforderung stelle, seinen Nachfahren die eigene Relevanz erklären zu müssen.⁴⁴⁶ In der Sicht dieser ganz irdisch verstandenen *Heils-Ökonomie* scheint der gegenwärtige Mensch in seiner bloßen Existenz sich paradoxerweise gegen die Zukunft des Lebens zu stellen. Die Verkehrtheit der Menschheit besteht im Beharren auf ihr So-Bleiben, das sich wider besseres Wissen gegen das Überleben im Werden-an-sich richte. Ihre Schuld zu bekennen – nämlich die Hybris, sich in der Selbstbehauptung der Menschheit gegen den Fortschritt zu Kosten allen Lebens zu stellen –, bedeute hingegen, die kosmische Aufgabe, seine Nachfolger zu gebären, anzunehmen und Platz für sie zu schaffen.

Wie wurde aus dem Menschen, den Arnold Gehlen noch als Mängelwesen bestimmte, schließlich ein mangelhaftes, lebensunwürdiges Wesen? Es wurde die These aufgestellt, das negative Menschenbild des *Silicon Valley*, das sich in den Konvergenzserzählungen in übersteigter Form zu Wort meldet, sei das Resultat einer christlich geprägten

445 Vgl. Tegmark, *Life* 3.0, 186.

446 Vgl. Lierfeld, *Künstliche Superintelligenz und/oder Ethik*, 41.

Geschichte, die den sündigen Menschen klein geredet habe und nun in säkularer Form fortwirke.⁴⁴⁷ Eine solch direkte Linie christlicher Hamartologie zu den Konvergenzerzählungen lässt sich historisch nicht ausmachen. Die sogenannten *Kränkungen* des Menschen in der Moderne scheinen historisch gesehen eher immer gegen die bestehende christliche Weltdeutung formuliert worden zu sein. Wie in Kapitel 4 dargestellt wurde, lässt sich die technizistische Selbstdemontage des Menschen direkter aus der technischen Kultur erklären: Ein übersteigerter Fortschrittsoptimismus, Machtphantasien und ihr Umschlag im Machtausgriff auf den Menschen, ein auf die Technik übergreifender Evolutionsbegriff und eine menschentobene Epistemologie sind maßgeblich für den Mythos Maschine und die aus ihm gezogenen Schlüsse, die zulasten des Menschen gehen. Das Christentum hält in den Begriffen von Schuld und Sünde sehr wohl die Fehlbarkeit des Menschen fest, fördert in der dauernden Prüfung des eigenen Handelns und Möglichkeit der Umkehr aber auch die reelle Einschätzung seiner Verantwortung und ihrer Grenzen für einen freien Umgang mit und in der Welt. Aus theologischer Perspektive ist gerade hervorzuheben, dass gegenüber jener Radikalverwerfung die christliche Tradition am sündigen Menschen festhält und ihn gegen Vorstellungen eines innerweltlichen Endgerichts verteidigt. Schöpfungstheologie dient der Entlastung des Menschen dahingehend, dass sein Handeln im Bewusstsein der Grenzen seiner Freiheit möglich und gut ist. Sie bietet ihm reelle Zuversicht im Umgang mit der Welt im Gegensatz zu jener Hoffnungslosigkeit, in der die besten Optionen dennoch bloß schlechte bleiben.

Die Weltlichkeit der Welt und der Zuspruch Gottes bleibender Fürsorge

„Was ist mit Wesen, die uns in allen Fähigkeiten übertreffen könnten? Der Mensch hat den Auftrag, die Gottes Schöpfung zu verwalten; was nun, wenn dies intelligente Systeme viel effektiver können?“⁴⁴⁸

447 Vgl. Spiekermann, *Digitale Ethik*, 255.

448 Ohly, *Ethik der Robotik und der Künstlichen Intelligenz*, 15.

Der technikinduzierte Zweifel an der Vormachtstellung des Menschen unterläuft nicht nur die Hoffnungen der Aufklärung, nach denen der Mensch sein Schicksal selbst in die Hand nimmt, sondern gleichermaßen die Anthropologie der Genesis, in welcher der Mensch zum Stellvertreter Gottes auf Erden eingesetzt wird – ohne dass diese Beauftragung eine Frist kennt. An die Menschen ergeht der Auftrag in Gen 1,28, sich die Erde untern zu machen; und im zweiten Schöpfungsbericht findet Adam sich in einem nur für ihn geschaffenen Garten vor, welchen er bebauen und bewahren soll (Gen 2,15); der Herrschaftsauftrag schließt so in gleicher Weise Verwaltung *über* wie Fürsorge *für* die Welt ein. Zu dieser Aufgabe ist der Mensch aus christlicher Perspektive als Gottes Geschöpf und in besonderer Hochschätzung zu seinem Ebenbild auf Dauer berufen. Die Geschöpflichkeit teilt er mit der restlichen belebten wie unbelebten Natur; als Ebenbild wird der Mensch von Gott zusätzlich als Person angesprochen und ihm Würde aus dieser Beziehung zuerkannt. Dem augenscheinlich elenden Menschen wird diese Würde zuteil, die sich nicht aus der Summe von Eigenschaften zusammensetzt, sondern er als gewolltes Geschöpf selbst Anteil an Gottes Heiligkeit erfährt. Die zuerkannte Würde geht damit allem Handeln und Scheitern voraus und begründet deswegen auch eine Grenze der Verfügung des Menschen über den Menschen.⁴⁴⁹ Schöpfung kennt die besondere Würde des Menschen aus seiner je individuellen Verbindung zu seinem Schöpfer und wehrt sich somit gegen eine anonymisierte Gattungsapokalyptik und ihren Fatalismus. Jedes individuelle Leben verdankt sich einem freien Entschluss des Schöpfers und geht nicht in zwangsläufigen Naturprozessen auf.⁴⁵⁰ In den Konvergenzmythen und ihrem Techno-Darwinismus findet sich hingegen die mythologische Überhöhung des menschlichen Gattungsbegriffs und eine Überbetonung daraus abgeleiteter, angeblicher Notwendigkeiten. Gattungen jedoch handeln nicht; das ist das Vorrecht von Individuen. Der menschliche Gattungsbegriff kann die Einheit aller menschlichen Geschichte, die stets von einem Handeln von Personen ausgeht, nicht

449 Vgl. Rendtorff, Ethik, 88–96.

450 Vgl. Anselm, Politische Ethik, 238.

begründen, ihre abgeleiteten Forderungen, der Mensch solle Platz für seine Nachfolger machen, bleiben so Ideologie.⁴⁵¹

Davon abgesehen ist nicht ersichtlich, wie Vorstellungen eines Regiments der superintelligenten KI, die den Menschen hinter sich lässt, um den Weltraum zu erobern, den göttlichen Herrschaftsauftrag erfüllen soll. Gottes Auftrag über Herrschaft und Fürsorge für die Welt sind keine Aufgaben im Sinne von *tasks*, die sich nach feststellbaren Parametern algorithmisieren und auf Utopiemaschinen übertragen ließen (s. 4.3.3). Dem Vorwurf der Überforderung des Menschen durch seine Umwelt stellt sich die Lehre von der Schöpfung gegen die Dichotomie vom hilflosen, guten Menschen gegenüber einer bösen Welt. Welt und Leben sind gleichermaßen Gottes „Selbstmitteilung zum Guten“.⁴⁵² Das Gutsein der Welt gründet in der Güte des Schöpfers und der Schöpfung als freiem Akt an sich. Die lutherischen Dogmatiker haben ausnahmslos betont, dass die gesamte Schöpfung und ihre Geschichte in letzter Konsequenz der Verherrlichung Gottes dienen.⁴⁵³ Sie waren sich jedoch auch sicher, dass in diesen Zweck die *utilitas hominum* untrennbar in Gottes Schöpfung eingebunden ist.⁴⁵⁴ Der Mensch als Ebenbild Gottes stellt Selbstzweck der Schöpfung dar; die Heilsgeschichte entspinnt sich als wechselseitiges Beziehungsgeschehen zwischen Gott und beauftragtem Geschöpf mit dem Ziel der Versöhnung. Den Menschen aufzugeben, um die menschliche Kultur oder das Leben überhaupt zu retten, erscheint in diesem Licht widersinnig. Aus christlicher Perspektive kann der Mensch die Verantwortung für die Schöpfung nicht sufizient abgeben, ohne dass er sie in dem Moment verfehlt.

Bei aller Beauftragung des Menschen, innerhalb seines Einflussbereiches den Chaosmächten zu wehren, bleibt das christliche Dogma bei der Einsicht, dass die Schöpfung wie die Bewahrung der Welt (*creatio et conservatio mundi*) Prädikate Gottes bleiben. Sie ist keine Aussage

451 Vgl. Wolfhart Pannenberg, *Das Glaubensbekenntnis*. ausgelegt und verantwortet vor den Fragen der Gegenwart, 6. Aufl., Gütersloh 1995, 49.

452 Jürgen Moltmann, „Ist die Welt unfertig?“, *Evangelische Theologie* 72 (2012), 326–336.

453 Friedrich Wilhelm Graf, „Von der creatio ex nihilo zur »Bewahrung der Schöpfung«: Dogmatische Erwägungen zur Frage nach einer möglichen ethischen Relevanz der Schöpfungslehre“, *Zeitschrift für Theologie und Kirche* 87 (1990), 206–223. Hier 211.

454 Vgl. Trutz Rendtorff, *Vielspältiges*. Protestantische Beiträge zur ethischen Kultur, Stuttgart/Berlin 1991, 142.

über die Fähigkeiten des Menschen, die Schöpfung entweder zu erhalten, noch sie zu zerstören.⁴⁵⁵ Spätestens Gottes Selbstbindung nach der Sintflut in Gen 9 markiert das endgültige Entlassen der Welt in ihre Weltlichkeit, sie ist Ort sich bewährender Verantwortung dieses Auftrages, aber das Gutsein der Welt ist unabhängig von der moralischen Integrität der Menschheit.⁴⁵⁶ Es ist die Versicherung, dass die Gottlosigkeit des Menschen nicht zur Gottverlassenheit der Welt führt. Der Mensch ist dem Lauf der Welt moralisch nichts *schuldig*.

Und siehe, ich bin bei euch alle Tage bis an der Welt Ende. (Mt 28,20)

Denn siehe, ich mache alles neu. (Off 21,5)

Die im christlichen Schöpfungsglauben zur Sprache kommende Freiheit hingegen setzt den Menschen zu seiner Umwelt in produktive Distanz. Die Umwelt ist weder zu verehren noch wertlos, sondern zu gestalten aus der Erkenntnis einer verdankten Freiheit. Die Weltlichkeit der Welt anzuerkennen bedeutet, dann nichts am Absoluten zu messen. Menschen handeln unter Begrenzungen, ihren eigenen und denen der Umwelt. Als Gottes Schöpfung bleibt sie von Gott getrennt und ist somit entgöttlicht. Die entgöttlichte Erde als Gestaltungsraum des Menschen bedeutet keine gottlose Welt. Christliche Hoffnung lebt von der Zuversicht, dass Gott im Sinne der *creatio continua* fortführend an der Welterhaltung und dem Entstehen von neuem Leben mitwirkt. Im Hebräerbief (Hebr 1,3) heißt es, Gott trage fortwährend alle Dinge durch sein kräftiges Wort. Aus dem geduldigen Tragen Gottes entstehen die Spielräume für die Selbstorganisationen alles Lebendigen, die es mit Menschenmitteln vor der Möglichkeit der (Selbst-)zerstörungen zu schützen gilt.⁴⁵⁷

Die Weltlichkeit der Welt anzuerkennen bedeutet auch, die Ewigkeit Gottes nicht der Schöpfung anzuhängen. Der christliche Glaube weiß um den Bestand der Welt durch Gottes Beistand, aber auch, dass sie keine unsterbliche Welt der ewigen Kreisläufe ist. Nach Jürgen

455 Ebd., 139.

456 Ebd., 140.

457 Vgl. Moltmann, Ist die Welt unfertig?, 334.

Moltmann (*1926) kann Schöpfung gar nicht abgeschlossen sein, sondern muss aus der christlichen Hoffnung heraus offen für die Emergenz des Neuen sein, wenn Gott schließlich aus dem Alten das Neue erschaffen soll.⁴⁵⁸ Diese Offenheit für Kontingenz ist unvereinbar mit angeblichen Zwängen, nach denen man die Erde und das gesamte Universum in ein geschlossenes System umwandeln müsste. Eine als unerträglich empfundene Ungewissheit gegenüber einer diffusen Gefahr der noch nicht unterworfenen Naturkräfte soll so getilgt werden, die sich in den Konvergenzerzählungen zu einer apokalyptischen Panik gesteigert hat. Um alle Möglichkeit zur Katastrophe für alle Zeit auszulöschen, wird dafür die Vernichtung der menschlichen Freiheit, die eben auch das Risiko der Selbstzerstörung in sich trägt, in Kauf genommen, um stattdessen *zukunftsconforme Lebensprogramme* für die nächsten Millionen Jahre an ihre Stelle zu setzen.

Das christliche Ethos hingegen stiftet dazu an, unter der Zusage einer bleibenden Fürsorge Gottes mit diesen Kontingenzen leben zu lernen. Das Leben besteht auf Grundlagen, die sich einer letzten Kontrolle entziehen. Ihre Alternative, eine tatsächliche *All*-kontrolle stellt in der Tat ein den Menschen überforderndes Paradigma dar. Die Menschen können „*Guardians of the Future of Life*“⁴⁵⁹ nur in einem relativen Verhältnis zu ihren Grundlagen sein. Absolute Macht und Kontrolle bleiben göttliche Attribute. Die Wahrnehmung solcher Grenzen dient dabei nicht der Reduktion von Verantwortung, wie es der Fatalismus der Konvergenzerzählungen heraufbeschwört, sondern genau umgekehrt einer Steigerung reflektierter Verantwortung. Die Welt als Schöpfung zu verstehen, verstärkt nicht apokalyptische Ängste eines drohenden Weltuntergangs, sondern lenkt das Handeln des Menschen – weil dem Menschen das Ganze nicht zur unmittelbaren Disposition steht – hin zu Pragmatik und Gradualismus.⁴⁶⁰

Ein Handeln im Bewusstsein der eigenen Grenzen wird trotz aller Ängste möglich im Vertrauen auf die Heilsignatur der Welt. Das Wissen um die Bedrängnis in dieser Welt wird relativiert, aber nicht annulliert,

458 Vgl. ebd., 330.

459 Tegmark, *Life* 3.0, 498.

460 Graf, Von der *creatio ex nihilo* zur »Bewahrung der Schöpfung«: Dogmatische Erwägungen zur Frage nach einer möglichen ethischen Relevanz der Schöpfungslehre, 222.

wenn an dem Gutsein der Schöpfung auch gegen alle widersprechenden Erfahrungen festgehalten wird.⁴⁶¹ Die Providenzlehre verleiht der Schöpfungslehre einen heilsgeschichtlichen Grundton. Der Gott, der im Anfang die Welt sehr gut auf den Weg gebracht hat, bleibt ihr treu und führt sie zum guten Ende. Der durch Christus erlösende Gott wiederum ist kein anderer als der Schöpfer dieser Welt.⁴⁶² Diese eschatologische Hoffnung mag auf die Gegenwart ausgreifen und ermutigt zum Guten im Umgang mit der Welt trotz aller punktuellen Erfahrungen der Wertlosigkeit des eigenen Handelns.⁴⁶³

Die Interpretation der Welt als Schöpfung bedeutet die Entideologisierung des Vorgefundenen, sowohl der Sphäre des Politischen,⁴⁶⁴ wie auch von Wissenschaft und Technik. Jede technizistische Vision eines kommenden, absoluten Weltregimentes muss verdächtig werden, denn auch Wissenschaft und Technik obliegen der Unabgeschlossenheit der Welt. Die technische Kultur bleibt angewiesen auf die vorgefundene Materialität und den von außen in sie eingewirkten Willen. Technik ist keine Form göttlichen Handelns, sondern findet im Raum der Schöpfung und damit der Zusicherung Gottes statt, aus der sie nicht heraustritt und zur diabolischen Macht aufsteigt. Im Gegensatz vom Mythos einer zweiten Schöpfung kann Theologie stets nur betonen, dass Technik im christlichen Verständnis nie außerhalb der Schöpfung steht und somit unter ihren Bedingungen mit einem göttlichen Schöpfungsakt gerade nichts zu tun hat.⁴⁶⁵ Menschen erschaffen keine Titanenkinder, auf die sie die eigene Verantwortung ablegen können. Gegenüber dieser Idolisierung der Maschine bleibt selbst der intelligente Apparat weiterhin dem Menschen unterstellt und empfängt seinen Zweck außerhalb seiner selbst. Wird Technik als Teil der dem Menschen anvertrauten Schöpfung wahrgenommen, heißt dies nicht nur, Befürchtungen eines bevorstehenden Rollentausches von Machthaber

461 Vgl. ebd., 223.

462 Pannenberg, *Das Glaubensbekenntnis*, 44.

463 Vgl. Wolfhart Pannenberg, *Grundlagen der Ethik*. Philosophisch-theologische Perspektiven (Kleine Vandenhoeck-Reihe 1577), Göttingen 1996, 70–71.

464 Anselm, *Politische Ethik*, 240–241.

465 Schwarke, *Technik und Theologie. Was ist der Gegenstand einer theologischen Technikethik?*, 98.

und Dienstleister kritisch zu begegnen, sondern auch, dass Technik niemals ein völlig autarker, kontingenzfreier Lebensraum sein kann, der sich je von den unkontrollierbaren Grundbedingungen der Welt lösen lässt,⁴⁶⁶ eine Erkenntnis, die besonders für die kommenden beiden Motive von Belang ist, in denen die soziale Umwelt und die Wirklichkeit als Ganzes durch intelligente Technik substituiert werden sollen, um dem überforderten Menschen eine bessere Heimat zu werden.

5.2 Vertretung: Der vollumsorgte Mensch und sein verwaltetes Leben

5.2.1 Mensch-Maschine-Interaktion in der neuen Polis der Allsolution

Szenario Zwei droht am ehesten, zwischen den Extremen einer auf Konfrontation und Vernichtung ausgerichteten Super-KI und der Identifikation des Menschen mit der kommenden Superintelligenz zu verschwinden. Ein systematisches Eigenrecht besitzt das Motiv dennoch, gerade weil es die Folgen für das menschliche Miteinander in der Gegenwart starker KI vor Augen führt, wie dies in den anderen beiden Motiven nicht zum Tragen kommen. Hier wird eine Zukunft entworfen, in der die Menschen von einem digitalen Dienstleistungssystem umgeben sind, in welchem die intelligente Maschine zum wichtigsten, möglicherweise sogar einzigen Bezugspunkt der eigenen Interaktion avanciert. Der Mensch ist in eine mediatisierte Umwelt eingefasst, die aus der Analyse von Daten und der Auswertung von Nutzungsprofilen eine vollständige Bedürfnisbefriedigung des Individuums garantiert, indem nicht allein Politik und Wirtschaft, sondern sämtliche Lebensbereiche in ein kybernetisches Selbstregelsystem umgewandelt wurden. Die Bedürfnisse eines jeden Menschen werden über dieses System nicht bloß ausgewertet, sondern sollen umfassend auch durch Maschinen bedient werden.

Das Vertretungsmotiv imaginiert somit die Folgen einer völligen Durchdringung, Antizipation und damit Beherrschung der menschlichen Seele durch *Big Data* und Künstliche Intelligenz. Die KI ersetzt hier

466 Vgl. Rendtorff, Ethik, 161–172.

nicht den Menschen, sondern vertritt (engl. *Surrogates*: Vertreter) ihn auf mehreren Ebenen: Erstens vertritt sie ihn als Arbeitskraft; zweitens vertritt sie ihn als allgegenwärtiges Assistenzsystem bei sämtlicher Entscheidungsfindung auf Grundlage eines vollumfassenden Datenprofils; schließlich vertritt Künstliche Intelligenz Familie oder Partner, indem auch soziale Beziehungen einer Bedürfnisanalyse unterzogen und die Sehnsucht nach Nähe durch menschenähnliche Maschinen erfüllt wird. Die Prämisse dieser Szenarien lautet, dass Maschinen eine solche glaubhafte Vertretung auf allen Ebenen mithilfe *Big Data* und starker KI übererfüllen werden. Die intelligenten Assistenzsysteme geraten in der post-singulären Zukunft zu universalen Lebensbegleitern, die zeigen, dass mit der nötigen Rechenleistung und dem unbegrenzten Zugriff auf relevanten Daten auch die eigene Lebensführung eine Aufgabe ist, die algorithmisiert und damit an Maschinen abgegeben werden kann.

Die postsinguläre Smart City als Polis der Solution

Die disruptive Kraft dieser Zukunft gründet in der Extrapolation gegenwärtiger, digitaler Dienstleistungen. *Big Data* wird entfesselt hin zum sog. *Ubiquitous Computing*, das bedeutet: Sensoren und Prozessoren sind allgegenwärtig und der freie Datenverkehr sämtlicher, selbst persönlicher Daten wird schrankenlos zugelassen. Erst dadurch kann die KI den Menschen gut genug kennen, um aus diesem Wissen ihn in einer Weise anzuleiten und dann auch zufrieden zu stellen, als es der Einzelne in der Undurchsichtigkeit des eigenen Wollens und Handelns jemals könnte:

„Stellen Sie sich ein System vor, das, in den Worten eines berühmten Songs von Police, jeden Ihrer Atemzüge, jede Ihrer Bewegungen und jedes Ihrer gebrochenen Versprechen verfolgt. Ein System, das Ihr Bankkonto und Ihren Puls, Ihren Blutzuckerspiegel und Ihre sexuellen Eskapaden überwacht. Es wird Sie definitiv besser kennen als Sie sich selbst...“⁴⁶⁷

Den Herrschaftsbereich einer solchen Zukunft bezeichnet der Soziologe Oliver Nachtwey mit der *Polis der Solution*. In dieser Stadt werden

alle Probleme, selbst soziale und private, für technisch lösbar erklärt. Die Verwalter dieser neuen *Polis* – sei sie lokal gedacht oder eine globale *Smart City* – sind keine Staaten mehr, sondern Tech-Unternehmen, die mit ihren datenbasierten *Services* immer mehr behördliche Aufgaben übernehmen, von dem Schutz der Gesundheit bis hin zur Strafverfolgung durch *Predictive Policing*. Der Bürger gerät zum Kunden, und die Weltverbesserung durch smarte Algorithmen wird den Unternehmen gleichzeitig zu einer Geschäftsmöglichkeit, auf die das *Silicon Valley* bereits gegenwärtig hinarbeite.⁴⁶⁸ Die *Polis der Solution* stellt somit eine Fortführung und Überbietung der kybernetischen Staatsidee dar (s. 4.2), in welcher ein Fortschrittsprung im Sinne der technologischen Singularität eine immense Erweiterung der *Services* bedeutet.

Jene durch die Singularität eröffnete Tiefentechnik soll nun aber nicht mehr bloß Politik und Wirtschaft, sondern selbst intimste Lebensbereiche technisch miteinander verschalten und so dauerhaft *lösen*. Neben den Staatsaufgaben soll das individuelle Leben im Ganzen anhand des Regelwerts *happy/unhappy* abgehandelt werden. In dieser *Polis*, die eine ganze Lebensform ausprägt, würden Entscheidungen, politisch wie individuell, überhaupt nur noch auf einer Datengrundlage getroffen werden. Der Historiker Yuval Harari sieht in dieser Zukunft schlussendlich Daten wichtiger als Menschen werden. Mit den technologischen Umwälzungen würden der Humanismus und sein Glaube an die Einmaligkeit des Menschen, insbesondere der Glaube an einen freien Willen, schlussendlich fallen, wenn der Mensch inmitten seines belebten Maschinenparks sich als berechenbares Tier offenbart. In Folge seiner Entheiligung werde den Daten irgendwann allein ein Eigenwert zukommen. Harari warnt, dass die neue Ideologie des *Dataismus* vom Menschen der Zukunft nichts weiter erwarten werde, als unter allen Umständen und gegen alle Widerstände einen unbeschränkten, dauerhaften Fluss von Information zu gewährleisten, um das System am Laufen zu halten.⁴⁶⁹

468 Vgl. Nachtwey, Seidl, *Die Ethik der Solution und der Geist des digitalen Kapitalismus*, 23.

469 Vgl. Harari, *Homo Deus*, 443. Natürlich steht dieses Argument einer schleichenden Verkehrung von Herr und Knecht in der Beziehung zwischen Menschen und Maschinen in der Tradition zu Samuel Butler (s. 4.1.1).

Das Leben erschöpft sich in Freizeitgestaltung

„In a future robotic economy, various characteristics of bodies and brains may have much less economic value, but the inherently human value of personal preferences will remain. Were individuals not to surrender personal information so easily, one could imagine Internet companies playing a significant role in wealth distribution by regularly compensating individuals for the value of the information they provide about themselves.“⁴⁷⁰

Für viele Prognostiker bedeutet das Nebeneinander von Mensch und einer superintelligenten Versorgungsinfrastruktur, deren Maschinen sämtliche Fähigkeiten und Eigenschaften des Menschen versichernd kopieren können, erst einmal das Ende klassischer Erwerbsarbeit. Manche äußern den naiven Wunsch, ein Arbeits-Surrogat möge dann an ihrer Stelle zur Arbeit gehen, während jeder Mensch von da an seinen Interessen nachgeht.⁴⁷¹ Der Großteil jedoch erwartet mit der vollen Automation sämtlicher Arbeit nichts weniger als den radikalen Umbau der gesamten Gesellschaft, der die Lebensverhältnisse für die überwiegende Mehrheit verschlechtern wird; die sozialen Erschütterungen führten zu einer Zweiklassengesellschaft in der *Polis der Solution*. Max Tegmark glaubt, aus diesen Umwälzungen ginge ein libertäres Utopia hervor, in dem nach der Entwertung des Menschen nur noch das Eigentum als letztes heiliges Prinzip über Mensch und Maschine herrsche. Der Fortschritt der Künstlichen Intelligenz werde die meisten Menschen arbeitsunfähig machen und sie dazu verdammen, den Rest ihres Lebens in roboterbetriebenen Sozialwohnungen zu verbringen, während wenige Tech-Magnate den immensen Gewinn der vollrationalisierten Arbeitswelt einstreichen.⁴⁷²

Harari prophezeit ebenfalls die Verelendung der Massen aufgrund ihrer Verdrängung aus sämtlichen Berufsfeldern. Den Einwand, dass der Aufgabenbereich des nicht Notwendigen dem Menschen erhalten bliebe,⁴⁷³ lässt Harari nicht gelten, da für ihn selbst gegenwärtig

470 Gill A. Pratt, „Is a Cambrian Explosion Coming for Robotics?“, *Journal of Economic Perspectives* 29 (2015), 51–60. Hier 59.

471 Vgl. ebd.

472 Vgl. Tegmark, *Life* 3.0, 163–168.

473 Vgl. Lierfeld, *Künstliche Superintelligenz und/oder Ethik*, 189.

menschliche Kreativität sich nicht mehr von maschinellen Ergebnissen unterscheiden lasse.⁴⁷⁴ Die Roboter würden fortan die Lasten beschwerlicher körperlicher wie intellektueller Arbeit tragen, während wenige Menschen – Harari spricht von vergöttlichten Übermenschen gemäß dem Buchtitel *Homo Deus* – über die *Polis* und ihre intelligenten Algorithmen wachen.⁴⁷⁵ Dem Großteil der Menschheit, der nicht zu dieser einflussreichen Kaste gehört, bliebe zur Beschäftigung lediglich das Wegdämmern aus der Realität mittels Drogen oder digitalen Ersatzwelten.⁴⁷⁶ Der Autor und Philosoph Richard David Precht warnt ebenfalls vor den Visionen eines Jahres 2040, in welchem die wenigen Gutverdiener über die vielen Abgehängten in ihrer Wohlfühlmatrix herrschen. Diese Visionen versinnbildlichten den Sieg des *Silicon Valley* und der mit ihren *Services* imprägnierten Lebensangst und Bequemlichkeit gegenüber menschlicher Autonomie, indem Apps und Algorithmen den Zufall und das Abenteuer des Lebens umgingen und daraus ein gigantisches Geschäft machten. Eine Zukunft, in welcher die gesamte Zeit der Bedürfnisbefriedigung und dem Konsum bereitsteht, bedeute zugleich die Rückbildung von Fertigkeiten, Fähigkeiten und damit Orientierung für jeden Einzelnen, sodass das Leben schließlich zum bloßem Zeitvertreib verkommt.⁴⁷⁷ Der Aussicht auf eine solche Ödnis des durchverwalteten Lebens manipulierter Massen wird nur wenig Positives entgegengesetzt. Eine geringe Chance sieht Tegmark noch dafür, dass mit der technologischen Singularität und ihren neuen Produktivkräften die Gesellschaft rechtzeitig das Eigentum und damit den Kapitalismus als Ganzes von sich stoßen könnte, sodass aus der Stadt des Überflusses eine egalitäre Utopie erwächst, die alle Menschen an den Gewinnen beteiligt.⁴⁷⁸

474 Vgl. Harari, *Homo Deus*, 439. Der Theologe Lukas Ohly weist darauf hin, dass auch dieser Bereich sich besonders für die Automatisierung anbiete. Er ist überzeugt, dass auch Pfarrerinnen und Pfarrer den künstlichen Konkurrenten und ihren individuell angepassten Angeboten unterlegen sind. Vgl. Ohly, *Ethik der Robotik und der Künstlichen Intelligenz*, 128.

475 Ebd., 467.

476 Vgl. ebd., 440–441.

477 Precht, *Jäger, Hirten, Kritiker*, 59–69.

478 Vgl. Tegmark, *Life 3.0*, 173–176.

Die beseelte Maschine präsentiert sich als der ideale Lebensbegleiter

Denkende Maschinen sind in der *Polis der Solution* nicht bloß Arbeitsvertretung und Freizeitbeschaffer, sondern zugleich Menschenvertretung und Freizeitgestalter. Künstliche Intelligenz unterstützt den Einzelnen in sämtlichen privaten Angelegenheiten mit seinem vermeintlich absoluten Gedächtnis über alle aufgezeichneten Aktivitäten und vorausgegangenen Entscheidungen. Harari imaginiert ein Szenario, indem auch die Partnerwahl auf Grundlage aller gesammelten Daten geschieht:

„Und Google wird antworten: Nun, ich kenne dich vom Tag deiner Geburt an. [...] Deine biochemischen Algorithmen der Savanne gegen die Algorithmen Googles nach neuesten Statistiken, nach denen Aussehen nur einen Anteil von 14 % am langfristigen Gelingen einer Beziehung zurechnen.“⁴⁷⁹

Dazu gesellen sich Vorstellungen, in denen die Partnerwahl in einer Zukunft starker Künstlicher Intelligenz überhaupt obsolet werden könnte. Künstliche Superintelligenz werde in Zukunft nicht nur intimste Entscheidungen wie die Partnerwahl übernehmen, sondern bei Bedarf auch genau den Partner in Form eines Androiden bereitstellen, der die maximale Befriedigung des gläsernen Menschen garantiert. Dass der Roboter der Zukunft ein von Menschen ununterscheidbarer Android sein wird, mit dem sich der Mensch seine Lebenswelt teilt, dieses Mimikry idealisieren schon heute viele Forscher wie Teile der Industrie zur idealen Mensch-Maschine-Interaktion.⁴⁸⁰ Für die Visionen der postsingulären Zukunft scheint dies ebenfalls eine ausgemachte Sache:

„Innerhalb der neuen Prämisse, nennen wir es die Post-Singularität-Ära, werden wir Robotern gegenüberstehen, die nicht misstrauisch machen. Sie werden genau hinsehen, sich bei Berührung überzeugend menschlich anfühlen oder bei Beleidigung wütend werden. Diese Entitäten werden darin geschult, wie Menschen zu handeln und zu reagieren. Dieses

479 Harari, *Homo Deus*, 455–456.

480 Vgl. Bartneck, Lütge, Wagner, *Ethik in KI und Robotik*, 81.

Ergebnis ist durchaus vorhersehbar, da es mit der Prämisse der Entwicklung von Geminoiden, insbesondere von Sexrobotern, übereinstimmt, die dem Menschen bis ins kleinste Detail ähneln und auch einen entsprechenden haptischen Eindruck vermitteln sollen.⁴⁸¹

Den Eindruck, dass eine solche Zukunft unmittelbar bevorsteht, vermitteln Androiden wie *Ameca*, die im Januar 2022 von dem britischen Unternehmen *Engineered Arts* als der menschlichste Roboter der Welt vorgestellt wurde. *Ameca* – ein Neologismus zwischen *automated mechanism* und lateinisch *amica* für Freundin – ist eine Androïdin, die ihre Gesichtszüge bewegt, mit ihren Pupillen Gegenstände und Personen fokussiert und auf menschliche Sprache reagiert. Das Unternehmen versteht sich als Hersteller von Unterhaltungsrobotern, die *Ameca* und andere Modelle zur Steigerung medialer Attraktivität an Wissenschaftszentren, Themenparks und Unternehmen verleihen.⁴⁸² Wie bei vielen technologischen Durchbrüchen ist in der Robotik die Unterhaltungsindustrie im Allgemeinen als auch Erotikindustrie im Speziellen ein Fortschrittstreiber. Roboterbordelle sind in Asien bereits eine gängige Praxis. Vier große Unternehmen in den USA stellen derzeit sog. *life companions* oder *companion robots* her mit einem besonderen Fokus auf sexuelle Dienstleistungen.⁴⁸³ Diese interagierenden Sexpuppen mit wiederum sprechenden Namen wie das Modell *Harmony* des Herstellers *Realdoll*⁴⁸⁴ stehen laut mancher Soziologen für eine sich in der Zukunft ausweitende *Digisexualität* des Menschen:

„In the coming years, sexual technology will become more sophisticated, immersive, and appealing. Many people will find that their experiences with this technology become integral to their sexual identity, and some will come to prefer them to direct sexual interactions with humans. We

481 Lierfeld, Künstliche Superintelligenz und/oder Ethik, 172–173.

482 Vgl. Engineered Arts, *Ameca: The Future Face of Robotics*, <https://www.engineeredarts.co.uk/robot/ameca/> (20.1.2022).

483 Vgl. Nature, „Let’s talk about Sex Robots“, *Nature* 547 (2017), 138.

484 Vgl. RealDoll, Product Site HarmonyX, <https://www.realdoll.com/product/harmony-x/> (22.1.2022).

propose to label those people who consider such experiences essential to their sexual identity, ‚digisexuals.‘⁴⁸⁵

Liegt die Zukunft von Liebe und Intimität bei intelligenten Robotern? In seinem Referenzwerk *Love and Sex with Robots* von 2008 prognostiziert der schottische Robotizist David Levy für das Jahr 2050, dass Maschinen bis dahin nicht nur wie Menschen aussehen und wie sie reden werden, sondern auch denken oder zumindest versichernd Denken simulieren können. Die verbleibende Differenz zwischen Mensch und Maschine sei dann nicht größer als der kulturelle Unterschied zwischen Menschen unterschiedlicher Nationen. Dabei behauptet Levy nie, dass die Zuneigung durch Roboter der menschlichen Liebe gleichkommt. Im Gegenteil, sie werde gerade deshalb so attraktiv sein, weil sie nicht den gleichen Standards zwischenmenschlicher Beziehung entsprechen müsse.⁴⁸⁶ Roboter böten in ihrer eigenen Bedürfnislosigkeit wie Adaptionfähigkeit den Menschen sexuell erregende, aber letztlich sichere Verhältnisse:

„Just as with the central heating thermostat that constantly monitors the temperature of your home, making it warmer or cooler as required, so your robot’s emotion system will constantly monitor the level of your affection for it, and as the level drops, your robot will experiment with changes in behavior aimed at restoring its appeal to you to normal.“⁴⁸⁷

Menschen würden sich in Roboter verlieben und mit ihnen zusammenleben wollen, weil sie in einer Weise *selbstlos* liebten, wie Menschen es gegenseitig niemals könnten. Dass die Roboter die für intime Beziehungen wichtige Voraussetzung authentischer Zuneigung nicht erfüllen, darin sieht Levy kein Problem. Wie in der Beziehung zu Prostituierten mache die menschliche Fähigkeit zur Selbsttäuschung es möglich, die erkaufte Leidenschaft als Ausdruck wirklicher Leidenschaft aufzufas-

485 Neil McArthur/Markie L. C. Twist, „The Rise of Digisexuality: Therapeutic Challenges and Possibilities“, *Sexual and Relationship Therapy* 32 (2017), 334–344.

486 David Levy, *Love and Sex with Robots*. The Evolution of Human-Robot Relations, New York/London/Toronto 2008, 112–118.

487 Ebd., 130–132.

sen. Diese *Myth of Mutuality* werde mit der Lebensechtheit kommender Roboter auch für sie zutreffen, sodass diese *andere* Liebe vollkommen suffizient sein werde.⁴⁸⁸

„It is a myth of the so-called strong AI view currently prevalent in the media that humanoid robots with social intelligence will be widely viable within a few years. [...] On the other hand, there is a great deal of diversity when it comes to the experience of intimacy. Some people already use artificial companions to offer them satisfactory social and intimate experiences despite their limited level of social intelligence.“⁴⁸⁹

Demgegenüber besteht die Hoffnung, dass die neuen Beziehungen mit den neuen Kreaturen doch nicht menschliche Liebe ersetzen, sondern das Spektrum menschlicher Bindungen erweitern werden:

„The creation of what we call substitutes holds out the prospect of enriching and transforming the plural condition of humanity. It promotes to introduce new creatures that, after the fashion of animals, only in a different way, will be at once like us and unlike us.“⁴⁹⁰

All diese Verfügungsphantasien werden gleichsam von Ängsten grundiert (s. 2.2.). Es ist nicht allein die Sorge, ob Menschen wirklich schon bereit für die Maschinenliebe sind.⁴⁹¹ In einem umfassenderen Sinne führt die Entwertung des Menschen im Angesicht einer Massenproduktion von ununterscheidbaren Androiden für das Individuum zur Identitätskrise, und für die Gesellschaft zur Staatszersetzung. Lierfeld warnt vor einer um sich greifenden Paranoia in einer Gesellschaft, in der Mensch und Maschine sich nicht mehr als solche erkennen können. Die den Szenarien unterlegte Hoffnung auf eine Ultrastabilität einer

488 Ebd., 205.

489 Yuefeng Zhou/H. Martin Fischer, „Intimate Relationships with Humanoid Robots: Exploring Human Sexuality in the Twenty-First Century“, in: Yuefeng Zhou (Hg.), *AI love you.*, Cham 2019, 177–184. Hier 179.

490 Paul Dumouchel/Luisa Damiano, *Living with robots. Vivre avec les robots*, Cambridge, Massachusetts/London 2017, 167–169.

491 Zhou, Fischer, *Intimate Relationships with Humanoid Robots: Exploring Human Sexuality in the Twenty-First Century*, 178.

gesättigten und bedürfnislosen, letztlich befriedeten Gesellschaft wird demnach immer wieder eingeholt von der Befürchtung, dass in dieser Sterilität das menschliche Leben gar nicht mehr gedeihen kann.⁴⁹²

5.2.2 Ethische Relevanz: Bevormundung und Versöhnung

Die Stadt der Zukunft ist ein Ort der totalen Verschaltung, in der Menschen zum Preis ihrer Privatsphäre das Leben hin zur optimalen Bedürfnisbefriedigung intensiviert haben. Das System muss dabei allgegenwärtig sein und gleichsam unsichtbar arbeiten; das Messen und Rechnen steht im Hintergrund einer auf der (Nutzer-)Oberfläche verzauberten Welt, die umfassende Services bereitstellt. Die Surrogatserzählungen folgen somit Günther Anders Prognose, dass eine Welt, in der alles Handeln in ein technischen Funktionieren übersetzt wurde, auch alles Bestehende in sich integrieren muss und kein Außerhalb mehr zulassen darf – dass also der Wunsch *Alles klappt* auf die Gleichung *Apparate = Welt* hinausläuft.⁴⁹³ Das Anerkennen dieser Totalität, die Harari mit seiner musikalischen Referenz säkular zu fassen sucht, bedeutet das Zuerkennen göttlicher Attribute an das superintelligente Versorgungssystem:

Ich sitze oder stehe auf, so weißt du es [...]. Ich gehe oder liege, so bist du um mich und siehst alle meine Wege [...]. Wohin soll ich gehen vor deinem Geist und wohin soll ich fliehen vor deinem Angesicht?
(Ps 139,2–3.7)

Die Potentialerwartungen setzen voraus, dass der Großteil der Menschen die massiven Eingriffe in ihre Privatsphäre akzeptieren werden, weil die totale Einbettung des Menschen in ein solch globales Mess- und Kontrollsystem einen anscheinend gerechtfertigten Mehrwert verschafft. *Ex negativo* wirft der Fokus dieser Szenarien auf eine volle Befriedigung menschlicher Bedürfnisse die Frage auf, welches Böse

⁴⁹² Lierfeld, Künstliche Superintelligenz und/oder Ethik, 172–173.

⁴⁹³ Vgl. Anders, Die Antiquiertheit des Menschen, 110–113.

dem bisher im Wege stand. Die wiederkehrende Gegenüberstellung zwischen menschlichem, fehlbarem Gedächtnis im Unterschied zu einem absoluten, künstlichen Gedächtnis gibt ein Indiz dafür, dass die Unberechenbarkeit menschlichen Handelns und ihrer Folgen im Surrogatsmotiv an sich zum Problem menschlichen Miteinanders identifiziert werden. Es ist das eigene undurchsichtig Sein und damit das Handeln unter einer begrenzten Perspektive, welches laut dieser Problemdiagnosen den Menschen an der Verwirklichung dauerhafter Gemeinschaftlichkeit immer wieder scheitern lässt und daher technologisch eingehegt gehört. Menschen bereiteten sich in ihren wechselseitigen Verstrickungen eine Hölle auf Erden, ganz im Sinne Jean-Paul Sartres (1905–1980) Ausspruch aus seinem Theaterstück *Geschlossene Gesellschaft* (frz. *Huis clos*): „*Lenfer, c'est les autres.*“⁴⁹⁴ Für die Mythen einer postsingulären Wirklichkeit lässt sich das Böse, das sich zwischen den Menschen in Missverständnis, Streit und unerfüllten Wünschen manifestiert, stets auf ein Defizit an Information zurückführen, womit gleichsam der vermeintliche Ausweg gewiesen wird. Die Versuchung dieser postsingulären Mythen liegt darin, im *Ubiquitous Computing* und starker KI eine vermeintliche Lösung dieser defekten Sozialität anzubieten, die das menschliche Ungenügen auffängt. Das absolute Gedächtnis einer Super-KI soll jedes Handeln verrechenbar und damit berechenbar machen. Politische und ethische Interessenskonflikte werden innerhalb dieser Logik einer instrumentellen Vernunft zu technisch lösbaren Problemen umgewandelt und damit aus gesellschaftlichen Dilemmata bloße Paradoxa gemacht.⁴⁹⁵ Letztlich sollen alle Dinge und Personen digital verbunden sein und miteinander kommunizieren, damit Individuen unverbunden nebeneinander existieren können.

Die *Polis der Solution* setzt jeden Konsumenten zum König ein über die geschlossene Welt seiner Nutzeroberfläche und gibt zugleich das Versprechen ab, dass er innerhalb dieses Schutzraumes durch technische Surrogate niemals mehr einsam sein werde. Kants Votum, dass die Freiheit des Einzelnen an der Freiheit des anderen ihre Grenze fin-

494 Jean-Paul Sartre, *Geschlossene Gesellschaft*. Stück in e. Akt = (*Huis clos*), Reinbek bei Hamburg 1975.

495 Vgl. Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 329.

det, wird hier usurpiert, als der andere innerhalb des eigenen, mediatisierten Zugriffs auf die Welt gar nicht mehr als unvermitteltes Gegenüber, mit dem man sich notwendigerweise auseinandersetzen muss, in Erscheinung tritt. Die Kontingenz der Lebenswelt soll in diesen Zukunftsvisionen in der Weise bewältigt werden, dass die menschliche Erfahrungswelt durch eine vollkommen bestimmbare – also willkürlich kontrollierbare – Welt ersetzt wird. Wenn die bisherige, gemeinsam geteilte Lebenswelt im Ganzen nicht steuerbar war und deshalb im Konfrontationsmotiv das ganze Universum zum geschlossenen System verwandelt gehört, soll hier die erwünschte Ultrastabilität der Verhältnisse in der Einbettung eines jeden Menschen in hermetische Parzellen erreicht werden.

Zwar mögen diese Szenarien einer fortschreitenden Vertretung des Menschen durch die Maschine weniger radikal wirken als die Auslöschungsphantasien, die der Vorstellung einer unausweichlichen Konfrontation von Mensch und Maschine entspringen (s. 5.1). Das bedeutet jedoch nicht, dass sie keine Überforderung für den Menschen voraussehen. Hinter die Bequemlichkeit, nicht arbeiten zu müssen oder sich überhaupt für etwas zu entscheiden, tritt in den Szenarien ebenso die Angst vor der eigenen, durch Daten geoffenbarten Fehlerhaftigkeit zutage. Sie führt zur Mutlosigkeit und Resignation, wenn das eigenverantwortliche Handeln als auch die Zusammenarbeit von Menschen nichts Sinnvolles in dieser Welt mehr zustande bringt. *In conclusio* argumentiert das Vertretungsmotiv auf dem Boden christlicher Versöhnungstheologie: Es fragt nach der Unfähigkeit des Menschen und sein Eingeständnis, kein ideales soziales Wesen zu sein, und nach Wegen, lebenswichtige Beziehungen in dieser Unvollkommenheit dennoch aufrecht zu erhalten und mit den eigenen Zielen und Wünschen zu vereinbaren.

Selbst- und Weltverlust im fremdverwalteten Leben

Die Gefahr der Demontage des Menschen tritt am deutlich in Lierfelds Warnung zutage, die Menschen der Zukunft verfielen in eine Massenpanik vor den menschengleichen Maschinen. Hinter diesem Extrem steht eine generelle Angst vor dem Selbstverlust inmitten einer vollautomatisierten Versorgungsstruktur. Man fürchtet, im gesinnungslosen sich

Gehenlassen im Gebrauch der bereitgestellten Güter und Dienstleistungen verliere der Mensch sich selbst und werde selbst zum funktionierenden Maschinenteil (s. 4.3.2.). Ähnlich warnt Harari vor der Selbstentäußerung des Menschen an die Entscheidungen von Maschinen:

„Menschen werden keine autonomen Wesen mehr sein, die von den Geschichten gelenkt werden, welche ihr erinnerndes Selbst erfindet. Stattdessen werden sie integraler Bestandteil eines riesigen globalen Netzwerks sein.“⁴⁹⁶

Individualität bleibt in diesem Großapparat höchstens im Konsum durch hochindividualisierte Bedürfnisbefriedigung bestehen. Ein geschlossenes System, das die soziale Wirklichkeit durch absolute Kontrolle befriedigen will, kann derweil keine Zukunft, nicht einmal eine Gegenwart bieten, sondern löst jede Zeitstruktur auf. Denn die Sicherheit eines gelösten Lebens wird damit erkauft, die Gegenwart zum Ergebnis algorithmischer Entscheidungsstrukturen und damit zur bloßen vergegenwärtigten Vergangenheit zu degradieren. Allein auf *Big Data*-gestützte Entscheidungen verlängern so die Vergangenheit in die Zukunft hinein, ohne dass irritierende Unterbrechungen berücksichtigt werden.⁴⁹⁷ Das vollumfassende Assistenzsystem bedeutet die Gefangennahme in *Feedbackschleifen*. Ein solches System, das Ultra-stabilität selbst für soziale Beziehungen garantieren will, kann mit der Emergenz des Neuen schlichtweg nichts anfangen.

Daher darf es innerhalb dieses hermetischen Systems keine Arbeit geben, die nur in entfernter Weise Weltgestaltung beansprucht. In den Szenarien einer Vertretung des Menschen von den Mühen der Arbeit geht es um kein Reich der Freiheit, um keinen Feenpalast,⁴⁹⁸ in dem der Mensch die gewonnene Zeit zur Besinnung nutzt, um aus dieser Selbstvergewisserung heraus die Welt neu zu gestalten (s. 2.1). Stattdessen wird all seine Kraft gebunden an ein sich Vergessen im Verzehr von Produkten. Hannah Arendt (1906–1975) beschrieb in ihrem phi-

496 Harari, *Homo Deus*, 456–458.

497 Vgl. Charbonnier, *Digitalisierung: Theologische Selbstklärungen und Gegenwartsinterpretationen. Eine Skizze*, 248.

498 Vgl. Schleiermacher, *Über die Religion*, 158.

losophischen Hauptwerk *Vita actica* die Grundformen menschlicher Tätigkeit und warnte zugleich vor deren Sterilität, wenn die Harmonie zwischen Arbeiten, Produzieren, und politischem Handeln gestört wird. Zur Arbeit zählte sie all jenes, das der Erhaltung der Lebenskräfte dient wie Nahrungsbeschaffung, Essen und Erholung, Zeugung wie Enthemmung. Arendt erklärte Arbeiten und Verzehren zu zwei Seiten einer Medaille, und sie mutmaßte, dass Menschen ihre ganze Arbeitskraft auch im Verzehren aufbrauchen könnten. Inmitten einer fortschreitenden Automation der 1960er Jahre sah Arendt die drohende Gefahr, dass in einem ungeheuer intensivierten Lebensprozess gerade die Künste des Menschen, und damit seine wirkliche Produktivität, einfach untergehen könnten. Die aufgewendete Lebenskraft des reinen Verzehens jedoch schaffe – anders als das Produzieren oder das Handeln – keine gemeinsame Welt mehr.⁴⁹⁹ In den postsingulären, kybernetischen Vertretungserzählungen soll der Mensch in gleichem Maße *happy* in einem möglichst weltlosen Sinne sein und sich sein Beitrag für das Allgemeine darin erschöpfen, den Algorithmen zu gehorchen und dabei weiter Daten zu produzieren. Zusätzlich verhindert der totale Anspruch dieses Systems auf den Einzelnen jedwede Privatheit, ohne die es auch keine gemeinsame Öffentlichkeit geben kann; in der abgeschlossenen Mediatisierung droht der Wirklichkeitsverlust. An die Stelle der Öffentlichkeit und das politische Forum tritt in den postsingulären Erzählungen eine reibungslose Sozialtechnik, deren *smarte* Lösungen die politische Auseinandersetzung ersetzen.⁵⁰⁰

Im Umkehrschluss verhindert der konsumfokussierte, digitale Solipsismus konsequent die Ansammlung politischer Macht. Die von den wenigen Technokraten zugestandene *Allmacht* über eigene kleine Welt hat keine politische Dimension, sondern lebt von Bedingungen, die sich als letzte ausgeben, doch aber kontingent zugewiesen wurden. Die wenigen Entscheidungsträger gleichen den Konzepten einer den Apparaten vorgeschalteten *Faculty of Foresight*, wie sie bereits in den

499 Vgl. Hannah Arendt, *Vita activa oder Vom tätigen Leben* (Piper 3623), 19. Aufl., München/Berlin/Zürich 2018, 154–155.

500 Vgl. Jürgen Habermas, „Technik und Wissenschaft als ›Ideologie‹“, in: Jürgen Habermas (Hg.), *Technik und Wissenschaft als ›Ideologie‹*, Frankfurt am Main 1968, 48–119. Hier 78.

kybernetischen Staatstheorien der 1970er Jahre auftauchen (s. 4.2.2). In den Konvergenzerzählungen haben sich jedoch nicht nur ihre Zugriffsmöglichkeiten auf den Einzelnen bis zum Äußersten erweitert, sondern sie handeln ebenso ohne politische Legitimation und unter Absehung einer Öffentlichkeit, vor der sie sich rechtfertigen müssten.

Freiheit in der Gemeinschaft

Christlicher Glaube ist, entgegen der ihm immer wieder vorgebrachten Kritik, keine Religion der Weltflucht, sondern adressiert die Frage nach dem gemeinsamen, guten Leben. Luther erkannte in seiner Freiheitsschrift *De libertas christiana* von 1520, dass der Mensch, auch wenn er als Einzelner in der Gewissheit von Gottes Rechtfertigung steht, weiterhin ein leibliches Leben zu führen hat. Deshalb stellte Luther die Frage, worauf das christliche Leben *in der Welt ausgerichtet* sein sollte, wenn der Christenmensch sich nur von Gott abhängig weiß. Da nun das eigene Seelenheil nicht mehr Grundmovers des Handelns ist, hieß dies für Luther nicht, gar keinen Grund zum Handeln in der Welt zu erkennen oder das Leben nur noch abzuwarten. Stattdessen könne der sich im Glauben befreite Christ sein Handeln ganz am Wohlergehen der Mitmenschen ausrichten und sich ihm in *dienender Liebe* zuwenden. Die Gemeinschaft der Glaubenden, in der jeder dem anderen ein Christus sein solle, bedeute deshalb eine Gemeinschaft gegenseitiger Fürsorge.⁵⁰¹ Wolfgang Huber hat diese Freiheit, die sich erst im Zuwenden zum Gegenüber realisiert, *kommunikative Freiheit* genannt und sie zum Kennzeichen der christlichen Ethik gegenüber einer egozentrischen Freiheit erklärt. Nur im Raum der Kultur als Ort des Gemeinsamen – und nicht außerhalb – kann diese Freiheit entstehen und bestehen.⁵⁰²

In der Verschränkung von Gottes- und Menschenliebe wirft das mit Gott versöhnte Bewusstsein notwendigerweise einen anderen Blick auf den *Mitmenschen*. In der Perspektive ethischer Theologie bedeutet der Glaube an die Versöhnung durch Gott in Christus für das gemeinsame

⁵⁰¹ Vgl. Martin Luther, „Von der Freiheit eines Christenmenschen (1520). Übertragung: Dietrich Korsch“, in: Dietrich Korsch (Hg.), *Deutsch-deutsche Studienausgabe / Luther, Martin*. Band 1 Glaube und Leben, Leipzig 2012, 277–317. Hier: 299–301 (WA30–31).

⁵⁰² Vgl. Wolfgang Huber, *Von der Freiheit*. Perspektiven für eine solidarische Welt, München 2012.

Leben, die konstitutive Sozialität des Menschseins mit dem Verfolgen eines eigenen Lebensentwurfs in Einklang zu bringen. Versöhnung strebt nach einem Gleichgewicht, den anderen so zu lieben wie sich selbst und sucht damit einerseits den Zwang zu überwinden, das eigene Leben nur an den Idealen der Gemeinschaft auszurichten, und wehrt andererseits, Freiheit nur als asketisches Ideal abseits einer gemeinsamen Welt zu denken.⁵⁰³ Diese Verschränkung von Gottes-, Menschen- und reflektierter Eigenliebe ist der Ausgangspunkt für christliche Sozialität, die eigene Anfragen und Entgegnungen in Bezug auf die entworfenen drei Ebenen maschineller Vertretung stellt.

Zur Frage nach dem Wert der Arbeit für das Miteinander hat die christliche Tradition eine letztlich pragmatische Einstellung ausgeprägt. Arbeit wird nicht dämonisiert, denn mit dem Auszug aus dem Paradies ist die Arbeit zwar mühevoller geworden, sie steht dennoch nicht schon als solche unter dem Fluch Gottes.⁵⁰⁴ Luther zumindest wies der Arbeit neben der Erhaltung des Menschen auch einen theologischen Wert zu. Aus einem großen Misstrauen der Reformatoren gegenüber den Folgen des Müßigganges war ihnen die Arbeit ein elementares Mittel zur Begrenzung von Sünde; die freien Werke sollten den Leib bearbeiten und damit bewahren.⁵⁰⁵ Positiv gewendet qualifizierte Luther den Dienst am Nächsten zum Vollzug seines Glaubens; Beruf ist dann die Konkretion der geforderten Nächstenliebe.⁵⁰⁶ Gleichzeitig wehrt die biblische Tradition einer Sakralisierung von Arbeit, kommt ihr doch bei aller Dienlichkeit nur ein relativer Wert zu. Schließlich gilt in der jesuanischen Eschatologie der Botschaft vom kommenden Reich Gottes der Vorrang vor aller irdischer Tätigkeit. Nirgendwo wird diese Bevorzugung der Kontemplation deutlicher versinnbildlicht als in der Gegenüberstellung von Maria zu Martha (Lk 38,42).⁵⁰⁷ Paulus schließlich weiß sein Handeln und Reden im Horizont der Naherwartung

503 Vgl. Anselm, Politische Ethik, 239.

504 Vgl. Wolfhart Pannenberg, „Fluch und Segen der Arbeit. Das Phänomen menschlicher Arbeit aus theologischer Sicht (1986)“, in: Wolfhart Pannenberg (Hg.), *Beiträge zur Systematischen Theologie*. Bd. 2 Natur und Mensch – und die Zukunft der Schöpfung, Göttingen 2000, 202–214. Hier 208.

505 Vgl. Luther, Von der Freiheit eines Christenmenschen (1520), 301–303 (WA31).

506 Vgl. Pannenberg, Fluch und Segen der Arbeit, 212.

507 Ebd., 210–211.

gestellt und ist sich sicher, dass bis zur Wiederkehr Christi und dem vollen, sichtbaren Einbruch der Gottesherrschaft jeder Mensch zum Dienst berufen und durch den Geist Gottes mit Gnadengaben (Charismen) ausgestattet ist, um zu dem Wohl der Gemeinde beizutragen. Jeder Mensch ist hier Subjekt einer Bildungsgeschichte; das Entwickeln von Fähigkeiten und Talenten gehört zum guten Leben.

Wir haben verschiedene Gnadengaben nach der Gnade, die uns gegeben ist. (Röm 12,6)

Dagegen lassen die Szenarien einer *Polis der Solution* neben einer negativen Freiheit von aller Mühseligkeit kein positives Verhältnis zum eigenen Tätigsein zu. Die Scheinbeschäftigung im Konsum erzeugt keine gemeinsame Welt; kommunikativer Freiheit ist so die Grundlage entzogen. Dabei scheint es, dass es gar kein Reich der Freiheit außerhalb des Reiches der Notwendigkeit geben kann, weil Weltbewältigung nicht durch Weltentzug sublimiert wird. Freiheit und Würde sind so nicht das Gewonnene, sondern das Verlorene der Utopie.⁵⁰⁸

„Aber dass der einzelne überhaupt irgendeinen Beitrag zum gesellschaftlichen Leben leisten kann und zwar so, dass er dadurch seinen Lebensunterhalt finden kann, das dürfte zur Würde des Menschen gehören und damit sowohl ein Recht des einzelnen als auch eine Pflicht gegenüber der Gesellschaft für ihn begründen.“⁵⁰⁹

Den Menschen ruhigzustellen ist nur eine andere Form seiner Auslöschung. Mit den Talenten zu wuchern (Mt 25,29), das kann der Einzelne nur im Kontakt mit anderen. Arbeit im Sinne des Berufes impliziert die Bildung von Verantwortung zu selbständiger Mitwirkung an der gemeinsamen Lebenswelt hin zu einem Kulturraum, aus dem heraus tatsächliche Zufriedenheit aus eigenverantwortlichem Handeln erwächst.⁵¹⁰

508 Vgl. Jonas, *Das Prinzip Verantwortung*, 365.

509 Pannenberg, *Fluch und Segen der Arbeit*, 214.

510 Vgl. Rendtorff, *Ethik*, 306–314.

Schuld, Umkehr und die Notwendigkeit zu verzeihen

Demgegenüber wenden gerade die Befürworter der kommenden Surrogate ein, dass die kommunikative Freiheit einen Irrweg darstellt, weil ihr Ideal an dem vorfindlichen Menschen scheitert. Die Erfahrung bezeuge, dass das Miteinander sich oft nicht als Ort der Freiheit erwiesen habe, sondern mühsam, konfliktgeladen und mit Zwängen verhaftet ist. Was das Vertretungsmotiv hervorhebt, ist die Sorge vor der Freiheit menschlichen Handelns, die immer auch zerstörerisch wirkt. Freiheit ist dem Menschen im Handeln gegeben, das Anfangen von Dingen mit ungewissen Ausgang ist, wie Augustin es beschreibt, das Charakteristikum menschlicher Existenz überhaupt, weil mit dem Menschen erst das Anfangen in die Welt kam:

„Initium ergo ut esset, creatus est homo, ante quem nullu fuit.“⁵¹¹

Mit der Freiheit besteht immer auch die Option, egoistisch zu handeln und Vertrauen zu enttäuschen; die Selbstursächlichkeit des Handelns bringt mit sich die Selbstverfehlung. Das liegt nicht zuletzt in der Natur des Handelns selbst, weil Menschen niemals vollumfänglich absehen können, was ihr Handeln für Folgen mit sich bringt.

„Wirkliche Freiheit gibt es auf Erden nur zusammen mit Schuld.“⁵¹²

Auch die Theologie erkennt in besonderer Weise an, dass das Miteinander zwischen Menschen problematisch ist und in der Gefahr steht, Gemeinschaften als Ganzes ins Chaos zu stürzen. Um die Folgen der Sünde und damit einer rücksichtslosen Selbstbezogenheit gegenüber dem Nächsten einzudämmen – so lehren es Augustin und später Luther – bedarf es zum einen des weltlichen Regiments. Recht und Gesetz sollen

511 „Damit ein Anfang sei, wurde der Mensch geschaffen, vor dem es niemand gab.“ Aurelius Augustinus, *De Civitate Dei*. (Comm. Thomas Waleys et Nicolaus Trivet), Basel 1489. XII, 20.

512 Michael Theunissen, „Freiheit und Schuld – Freiheit und Sünde“, in: Hans-Richard Reuter/Heinrich Bedford-Strohm/Helga Kuhlmann/Karl-Heinrich Lütcke (Hgg.), *Freiheit verantworten*. Festschrift für Wolfgang Huber zum 60. Geburtstag, Gütersloh 2002, 343–356. Hier 346.

die Folgen der Sünde eindämmen.⁵¹³ Die individuelle wie kollektive Unfähigkeit, den gegenseitigen moralischen Forderung zu genügen, enthebt Menschen in christlicher Tradition gleichsam nicht von ihrer Schuld. Christliche Ethik macht sprachfähig für dieses Unvermögen des Menschen und die Differenz von Wollen und Vollbringen des Guten (Röm 7,13–25). Sie tut dies jedoch nicht, um ihn – anders als die Visionen einer postsingulären Zukunft – vom Handeln abzubringen.⁵¹⁴ Vielmehr hält der Glaube an dem Menschen wie seiner Freiheit fest, und sagt, dass es ein erträgliches Maß an Zumutung für diese Freiheit geben muss, damit diese Freiheit lebensförderlich bleibt. Das anthropologische Diktum vom Menschen als *initium* ernst zu nehmen, bedeutet neben dem zerstörerischen Potential, den jedes Anfangen mit ungewissem Ausgang in sich birgt, gleichzeitig auch, dass Menschen in ihrer Lebensführung nicht dazu verdammt sind, ihre Fehler immer wieder zu tun. Jesus ruft zur Umkehr und Versöhnung mit Gott und dem Nächsten auf. Die Evangelien berichten von Momenten der Konfrontation, der Selbstprüfung, die Menschen tatsächlich umkehren lassen. Unter den Bedingungen der Endlichkeit sind Menschen zur Umkehr fähig. Vor allem aber können sie Spiralen des Hasses beenden und sich gegenseitig Neuanfänge gewähren, indem sie verzeihen. Für Hannah Arendt ist das Verzeihen nichts geringeres als ein Wunder. Sie sagt, es ginge mit dem Leben nicht weiter, wenn Menschen sich nicht ständig gegenseitig von den Folgen dessen befreien würden, was sie gesagt haben, ohne zu wissen, was sie tun. Nur durch dieses fortlaufende sich Entlasten und sich Entbinden könnten die Menschen mit der Mitgift ihrer Geburt, der Freiheit, auch frei bleiben:⁵¹⁵

„Nach Jesus hat nicht nur Gott die Macht zu vergeben, nein sie ist nicht einmal von der göttlichen Barmherzigkeit abhängig. Viel eher ist es genau gegenteilig: Die Menschen müssen sich verzeihen, damit auch

513 Vgl. Martin Luther, „Von der weltlichen Obrigkeit: Wie weit man ihr Gehorsam schuldet (1523)“, in: Hellmut Zschoch/Johannes Schilling/Albrecht Beutel/Dietrich Korsch/Notger Slenczka (Hgg.), *Deutsch-deutsche Studienausgabe / Luther, Martin*. Band 3 Christ und Welt, Leipzig 2016, 217–290. Hier 228–231 (WA250–251).

514 Vgl. Rendtorff, *Ethik*, 29.

515 Arendt, *Vita activa oder Vom tätigen Leben*, 301.

Gott verzeihen kann. Gott vergibt uns unsere Schuld, wie auch wir vergeben unseren Schuldigern. Die Macht zu vergeben, sie scheint unheimlicher als das Wundertun (Lk 7,49).⁵¹⁶

Diese zuversichtliche Seite des Handelns, nach der man Hoffnung für diese gemeinsame Welt haben darf, wird für Arendt in der Feier der Geburt Christi als die frohe Botschaft überhaupt erkennbar.⁵¹⁷ Christus wird zum Erweis für die Kraft des Neuen, die das Handeln nicht sein lässt. Die Offenbarung Gottes in Jesus Christus stiftet den neuen Bund, indem sie das Vorfindliche infrage stellt und dazu aufruft, die gegebene Ordnung immer wieder auf ihren Freiheitswert zu prüfen.⁵¹⁸ Die kybernetische Anthropologie hingegen reduziert den Menschen zu einem rein reaktiven Bündel an Bedürfnissen und meint, mit der richtigen Verschaltung von Information eine immanente *Seligkeit* in einer weltlosen künstlichen Wohlfühlmatrix garantieren zu können. Dort soll das Egozentrische nicht durch Mäßigung geschlichtet werden, sondern im Exzess sollen an einer folgenlosen Umwelt sämtliche Gelüste ausgelebt werden. Der Preis dafür ist, dass in dieser Matrix alle Aktivitäten folgenlos bleiben. Christliche Anthropologie stimmt den Visionen des Vertretungsmotivs zu, dass Menschen ihr Leben bei aller persönlicher wie gesellschaftlicher Anstrengungen nicht aus sich heraus heil finden. Gegenüber der scheinbaren Hoffnung, die sich an Künstliche Intelligenz bindet, nimmt das christliche Ethos diesen uneinholbaren Überschuss des Lebens ernst. Die Unstillbarkeit des menschlichen Herzens, die selbst bei der vollen digitalen Bedürfnisbefriedigung bestehen bliebe, verweist auf das noch ausstehende Reich Gottes.⁵¹⁹ Mit Blick auf das Eschaton bleibt auch das erfülltste Leben ein Fragment. Der Kreuzestod Christi vermittelt den Zuspruch, dass noch in dem äußers-

516 Ebd., 305.

517 Ebd., 316–317.

518 Vgl. Reiner Anselm/Christian Albrecht, *Öffentlicher Protestantismus*. Zur aktuellen Debatte um gesellschaftliche Präsenz und politische Aufgaben des evangelischen Christentums (Theologische Studien NF v.4), Zürich 2017, 46.

519 Vgl. Pannenberg, *Grundlagen der Ethik*, 81.

ten Scheitern Gott nahe ist. Diese Gottesnähe geht nicht in irdischer Lebenserfüllung auf und wird mit ihrem Scheitern nicht hinfällig.⁵²⁰

„Sie [die Situation des vorgefundenen Lebens] braucht keinesfalls immer als Wohltat zu erscheinen, sondern kann sich sehr wohl auch so darstellen, als ob das Leben erst durch menschliches Handeln erträglich gemacht werden kann.“⁵²¹

Der Glaube fördert so eine Resilienz gegenüber den Versprechungen eines sedierenden Schlaraffenlandes. Ein erfülltes Leben ist nicht unmöglich, aber aus christlicher Sicht kein absoluter Wert. Zu leben bedeutet mehr, als glücklich zu sein. Das eigene Glück hingegen durch Ignoranz und Selbstbetrug zu erkaufen, stellt einen wahrhaft unseligen Zustand dar. Hier ist Theodor Adornos Sentenz beizupflichten, wenn er sagt, dass es kein richtiges Leben im falschen gibt.

„Wer weiß, was die Realität ist, und sich willentlich von ihr abwendet, begeht eine Sünde. [...] Als Akteur will man Verantwortung für echtes Leben übernehmen, indem echte Handlungen auch echte Konsequenzen haben und so echte Freiheit ermöglichen.“⁵²²

In Konsequenz muss die Versöhnung immer wieder versucht werden um der gemeinsamen Weltbewältigung willen, weil es nur im erkämpften Raum der Kultur Freiheit gibt. Mittel zur Restitution sind Umkehr und Vergebung. Die Antwort aber auf die Frage, weshalb man diesen Kampf immer wieder führen sollte, weshalb man nicht sich der Hoffnungslosigkeit, der Entmutigung zum Handeln ergeben und dann in virtuelle Welten flüchten sollte, geben sie noch nicht. Vielmehr setzt das immer neue Versuchen mit seinen Heilmitteln der Umkehr und Vergebung ihre Motivation bereits voraus. Die christliche Tradition antwortet auf dieses Wozu mit der Rede von der in Christus verbürgten Zuwendung zum Menschen in Liebe.

520 Vgl. Pannenberg, Das Glaubensbekenntnis, 67–68.

521 Pannenberg, Grundlagen der Ethik, 82.

522 Nida-Rümelin, Weidenfeld, Digitaler Humanismus, 124–125.

Eros und Agape

„Czum.xxx. Ausz dem allenn folget der beschlusz / das eyn Christen mensch lebt nit ynn yhm selb / sondern ynn Christo vnd seynem nehstenn / ynn Christo durch den glauben / ym nehsten / durch die liebe / ausz gott feret er widder vnter sich durch die liebe / vnd bleybt doch ymmer ynn gott vnd gottlicher liebe [...]“. ⁵²³

Die Zukunftsvisionen erheben *companion robots* zum Prüfstein dafür, wie sehr die *Polis der Solution* die Wünsche des Einzelnen antizipieren und deren Erfüllung technisch materialisieren kann. Das ethische Problem liegt dabei nicht allein, wie es von kirchlicher Seite formuliert wurde, in dem Irrtum, dass die mit Künstlicher Intelligenz ausgestatteten Maschinen nur Sexualfunktionen ersetzen, jedoch ihnen die Voraussetzungen für die echte personale Beziehung wie Unverfügbarkeit und die Freiwilligkeit wechselseitiger Begegnung fehle.⁵²⁴ Denn die Advokaten der Roboterliebe kommen ja gerade zu dem Schluss, dass selbst wenn die Künstliche Intelligenz der Zukunft diese Funktionen nicht adäquat kopieren könnte – für deren Misserfolg sie jedoch keinen Anlass sehen –, die sexuelle Befriedigung den Menschen letztlich genügen werde. Die etwas leblosere Subjekt-Objekt-Beziehung werde langfristig glücklicher machen.

Die bereitwillige Selbsttäuschung und Selbstverdinglichung der eigenen Gefühle führt in einen unseligen Zustand, weil sie das Verabsolutieren der Selbstbezogenheit, das *incurvatus in seipsum*, verharmlost und zur Freiheit verklärt. Schließlich begegnet dem Menschen in einem solchen Roboter niemand anderes als seine eigene Vergangenheit in Daten. Diese Romantisierung des Apparats macht zugleich blind für die hinter liegenden Abhängigkeiten dieser Mensch-Maschine-Beziehung. Denn das Versprechen, der Mensch werde vollends über seinen robotischen Partner verfügen, wird gleichsam darin konterkariert, dass die Betreiber durch den Surrogaten über den Menschen verfügen. *Ser-*

⁵²³ Luther, Von der Freiheit eines Christenmenschen (1520), 314 (WA 38).

⁵²⁴ Vgl. Evangelische Kirche in Deutschland, Freiheit digital. Die Zehn Gebote in Zeiten des digitalen Wandels., 164–165.

vices werden individuell angepasst im Austausch gegen intimste Nutzerdaten. Wer diesen Umgang unter dem Begriff der Liebe verhandelt, tut dies über eine verkümmerte, kybernetische Definition von Liebe, die über eine weitere Form des Datenaustausches zweier Systeme nicht hinauskommt:

„At least in some cases, becoming friends or falling in love is a process of knowledge acquisition. We cannot even begin to adequately understand it if we do not understand the information that it makes globally available for the system as a whole, plus the representational and functional role such an event plays for the now coupled system of our two self-modeling systems.“⁵²⁵

Eine leerere Bestimmung von Liebe gibt es nicht. Das tatsächlich Unerhörte in diesen Visionen ist die Vorstellung einer kommenden Gesellschaft, die der Kraft der Liebe nicht mehr traut und in ihrer Bequemlichkeit meint, ohne sie leben zu können. Letztlich ist diese Wendung nur konsequent: Wer eine hermetische Wohlfühlmatrix verspricht, die weder von außen noch von innen erschüttert werden darf, kann die Liebe nicht unter sich dulden. Denn was sprengt ein geschlossenes System mehr, lässt jemandem zu einem anderen Menschen umkehren als die Liebe? Welches Handeln ist explosiver und unberechenbarer als jenes, welches aus Liebe geboren wird?

Liebe bezeichnet aus ihrer christlichen Prägung heraus nicht bloß ein Handeln, sondern geht allem Handeln voraus, ist eine Forderung an die Existenz umfassenden Sinnes als ein *Sein in der Liebe*, ohne die alles Tun vergeblich bleibt (1 Kor 13,1–10). Die Liebe als Lebensform ist Vorzeichen des Reiches Gottes, die das Handeln in den Dienst des Guten stellt.⁵²⁶ Weil sie aus der Hoffnung schöpft, zieht sie ihre Kraft nicht aus der Vergangenheit, sondern ist Vergegenwärtigung der verheißenen Hoffnung für die Zukunft. Christliche Liebe ist gerade deshalb schöpferische Liebe, weil sie selbst schafft, was sie liebt. Liebe geht über die defizitäre Wirklichkeit hinaus, und macht sich so, anders als

525 Metzinger, *Being no one*, 603.

526 Vgl. Rendtorff, *Ethik*, 110–121.

es das Leben in *Feedbackschleifen* vorsieht, nicht von ihr abhängig. Das Vorgefundene muss nicht bereits liebenswürdig sein; Liebe hält um der Freiheit willen Defiziterfahrungen aus und schöpft in die Gegenwart hinein selbst, was sie liebt.⁵²⁷

Liebe ist somit der Hoffungsgrund des Handelns und ihr Wagnis zugleich, da dieses Ausgreifen in die Gegenwart stets bedeutet, sich verletzlich zu machen. Sich ganz in den Dienst des anderen zu stellen, um das Gute zu schaffen, das bisher nur in der Vorstellung besteht, bedarf echter Entscheidungen unter der Bedingung zu echtem Schuldig werden. Eine Ethik, die dem Grundsinn des Liebesgebotes folgt, fragt deshalb danach, welche Folgen das eigene Handeln und die eigene Lebensführung für andere hat und lässt sich davon leiten. Nicht die eigene Vollkommenheit ist das Thema einer gemeinschaftlichen Existenz in der Liebe, sondern das verbindende, realisierbare Gute unter den gestaltbaren Bedingungen menschlicher Lebenswirklichkeit.⁵²⁸ Eine Ethik der Folgen kann wiederum nur in einer Welt wirksam werden, die Veränderung und Folgen tatsächlich zulässt, in der Kontingenzzräume offengehalten werden, was gleichbedeutend damit ist, dass das Evangelium unter Menschen zugesprochen werden kann, weil das Einlassen auf Glaube, Liebe und Hoffnung mehr als einen Informationsaustausch voraussetzt. Die liturgische Gloria-Formel *Ehre sei Gott in der Höhe und Friede auf Erden* (Lk 2,14) bringt diese positive Gestaltung christlicher Liebe auf den Punkt: Wie das Ehren ein Beziehungsgeschehen zwischen Mensch und Gott voraussetzt, so ist auch der erhoffte Friede stets ein positiv gestalteter Friede, nie nur das Resultat der Abwesenheit von Gewalt durch Abwesenheit von Sozialität. Gemeinschaft ist ein für die Freiheit errungener und gleichzeitig durch die Folgen der Freiheit bedrohter Raum, der aber dieses Wagnis um der menschlichen Seele willen wert ist.

527 Vgl. Pannenberg, Das Glaubensbekenntnis, 150.

528 Vgl. Rendtorff, Ethik, 120.

5.3 Identifikation: Der Eingang des Menschen in die Technosphäre

5.3.1 Eins werden mit dem technologischen Fortschritt im Mind Upload

In sämtlichen Konvergenzerzählungen wird der Mensch mit seiner Mangelhaftigkeit konfrontiert. Wird der vorfindliche Mensch deshalb im Konfrontationsmotiv durch die intelligente Maschine verdrängt und im Vertretungsmotiv von einem digitalen System umschlossen und so unschädlich gemacht, imaginiert das Identifikationsmotiv die endgültige Überwindung seiner Biologie. Die Gleichsetzung von Mensch und Maschine bedeutet nun, dass der Mensch selbst die kommende Superintelligenz sein wird. Der Mensch wird nicht durch Roboterkinde ersetzt (s. 5.1), sondern der exponentielle Fortschritt erfasst den Menschen und reißt ihn mit sich:

„For one thing, ‚they‘ will be us, so there won’t be any clear distinctions between biological and nonbiological intelligence.“⁵²⁹

Will der Mensch selbst an der Unsterblichkeit des technologischen Fortschrittes teilhaben, muss er bereit sein, auf seinen biologischen Kern zu verzichten und sich vollends der Logik der Maschine zu unterstellen.⁵³⁰ Im Zuge der informationstechnologischen Fortschritte um das maschinelle Lernen gilt der *Mind Upload* unter Transhumanisten als vielversprechender Kandidat, um den defizitären Menschen zu überwinden. Dazu sollen die neuronalen Verbindungen des Gehirns gescannt und auf ein informationstechnologisches Trägermedium überführt werden. Hiervon verspricht man sich eine identische Emulation der Hirnaktivität in binärem Computercode hin zu einem digital operierenden, intelligenten Bewusstsein. Der *Mind Upload* führt somit die Idee des *Cyborgs* vom Eins werden des Menschen mit der Maschine, wie

529 Kurzweil, *The Singularity Is Near*, 253.

530 Vgl. David J. Chalmers, „The Singularity: A Philosophical Analysis“, *Journal of Consciousness Studies* (2010), 7–65. Hier 33–34.

sie bereits in den 1980er Jahren aufkam, unter den Bedingungen der technologischen Singularität zu ihrer Vollendung, indem der *Upload* den Körper nun komplett hinter sich lässt. Das Ausgehen aus seiner menschlichen Natur ist dabei auch örtlich konnotiert und meint innerhalb der Szenarien die Überschreitung des bisherigen menschlichen Lebensraums in gleich zweifacher Hinsicht: Als hochgeladenes digitales Substrat, als *Emulated Mind* (EM), wird der Mensch ein Bewohner des *Cyberspace* im Sinne selbstprogrammierter digitaler Welten. Da er als Programm wiederum nicht mehr auf Sauerstoff, bestimmte Temperaturen oder Nahrung angewiesen ist, wird er in seinen Maschinen den *Outer Space*, das Universum, besiedeln und in noch mehr Rechner umwandeln.

Der Mind Upload in Popkultur, Wissenschaft und als Corporate Vision

Vermittelt über das Bild des Cyborgs ist die Idee der Verwandlung des Menschen in eine Maschine längst in das kulturelle Gedächtnis eingegangen (s. 2.2). Eine Mehrheit kann sich im Jahr 2022 vage etwas unter der Idee vorstellen, sein Bewusstsein digital hochzuladen, ohne je von den Theorien um den *Mind Upload* gehört zu haben. Der erste, der innerhalb der Science-Fiction den menschlichen Geist in ein Substrat aus Information auflöste und auf ein technisches Medium übertrug, war der britische Schriftsteller Arthur Clarke (1917–2008). Mit seiner 1956 erschienen Geschichte *The City and the Stars*, die sichtlich von Wieners Kybernetik beeinflusst scheint, erzählt er von der Stadt Diaspar in einer viele Millionen Jahre entfernten Zukunft, in der es den Menschen gelungen ist, den Geist vom Körper zu lösen. Aus einem Zentralcomputer heraus werden menschliche Speichersätze in Körpern zum Leben erweckt und dort wieder nach ihrem Ableben mit ihren neuen Lebenserfahrungen eingespeist.⁵³¹

Versatzstücke dieses Motivs konfundieren in den letzten Jahren vermehrt mit wissenschaftlichen wie wirtschaftlichen Unternehmungen. 2013 startete das *Human Brain Project* der EU mit einem Budget von

⁵³¹ Vgl. Arthur Charles Clarke (Hg.), *Die sieben Sonnen*. Ein utopisch-techn. Abenteuerroman, München 1960.

über eine Milliarde Euro. In seiner zehnjährigen Laufzeit soll es ein hyperrealistisches Computermodell eines menschlichen Gehirns entwerfen, an dem sich beispielsweise die Wirkung verschiedener Psychopharmaka testen lässt.⁵³² *Singularitans, Extropians* (s. 4.2.4) und Transhumanisten allgemein sehen mit diesem Modell den ersten Schritt getan, um von dort zu einer Emulation von Bewusstsein in der Maschine zu gelangen. Förderlich für eine solche Konfusion ist nicht zuletzt die überzogene Forschungsrhetorik seitens der Projektleiter, die eine klare Distinktion dessen, was ihr Projekt leisten kann, nicht vornimmt.⁵³³ Die Idee, den Menschen der Zukunft in einen Roboter zu verwandeln, ist ebenso in Unternehmen und *Start-Ups* angekommen und bildet dort Teil einer vermarktungsförderlichen Unternehmensstrategie, für die man nicht bis ins *Silicon Valley* schauen muss: Das Münchner Unternehmen *Devanthro* (*Developing Anthropomimetic Robodies*), das aus einer Forschergruppe der Technischen Universität hervorging, stellt robotische Exoskelette her. Ihr Geschäftsführer Rafael Hostetter wirbt medienwirksam damit, dass ihre Forschung den Körper bereitstellen soll, in den er sich in dreißig Jahren hochladen wird.⁵³⁴

Das 2011 gestartete Forschungsprogramm *Initiative 2045* des russischen Internetmilliardärs und selbsternannten Philanthropen Dmitry Itskov hat sich indes ganz konkret den digitalen Geiststransfer zum Etappenziel ihres mehrstufigen Projektplans gesetzt. Seine Unternehmung soll die bestehenden Krisen der globalisierten Menschheit überwinden und alle Menschen in eine erfüllende Zukunft führen:

„The main goals of the 2045 Initiative: the creation and realization of a new strategy for the development of humanity which meets global civilization challenges; the creation of optimal conditions promoting the spiritual enlightenment of humanity; and the realization of a new futuris-

532 Vgl. The Human Brain Project, Brain Models and Simulation, <https://www.humanbrainproject.eu/en/brain-simulation/> (25.1.2022).

533 Vgl. Ulrich Schnabel, Hirnforschers Mondfahrt. Human Brain Project, 27.11.2019, <https://www.zeit.de/2019/49/human-brain-project-hirnforschung-ausstellung-finanzierung-neurowissenschaften> (27.1.2022).

534 Vgl. Devanthro GmbH, Robodies: General Purpose Remote Bodies, 16.11.2020, <https://devanthro.com/> (16.11.2020).

tic reality based on 5 principles: high spirituality, high culture, high ethics, high science and high technologies. The main science mega-project of the 2045 Initiative aims to create technologies enabling the transfer of an individual's personality to a more advanced non-biological carrier, and extending life, including to the point of immortality.⁵³⁵

Der Upload und seine Hürden

Die kybernetische Behauptung der Isomorphie zwischen Denken und Rechnen ist für den *Upload* zwingend. Schließlich muss das Medium die Hirnfunktionen nicht nur eins zu eins abbilden und simulieren können, sondern die Fortsetzung der Innenerfahrung des Bewusstseins in einer *Whole Brain Emulation* garantieren. Wie dieses Hochladen des Ichs in der Zukunft aussehen könnte, dafür gibt der amerikanische Philosoph Mark Walker ein Beispiel:

„You arrive at one of the thousands of kiosks run by the late twenty-first century's largest corporation: U-Upload. With some trepidation you step into the superscanner. There is a slight hum as it inventories the molecular building blocks of your brain. Your brain is destroyed in the process, but you are not dead – or so the marketing materials from U-Upload claim. [...] For your friends and family, a few terrifying moments pass before the robotic body stirs. To their relief, your first words are: “It's me. I made it.” You go on to crack a joke – just as your family and friends have come to expect of you. Of course you have changed in some respects: gone is your human carbon-based body. Now you experience the world through camera eyes and microphone ears, you dance the fandango with robotic legs and speak through a voice synthesizer. But it is still you. You have migrated to a silicon substrate: you have been uploaded.”⁵³⁶

Fragen zum Prozedere stellen sich beispielsweise darin, wie ein Scan die benötigte Detailgenauigkeit erreicht. Kurzweil spricht sich im Gegensatz zu Walkers *einem* Scangerät für unzählige Nanoroboter aus, die in

535 Dmitry Itskov, 2045 Initiative, 4.8.2020, <http://2045.com/> (4.8.2020).

536 Mark Walker, „Uploading and Personal Identity“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 161–177. Hier 161.

das Gehirn eingespeist werden und von dort die Hirnaktivität scannen.⁵³⁷ Der amerikanische Philosoph David Chalmers hingegen glaubt, dass man überhaupt sanftere Verfahren entwickeln werde, mit denen das biologische Original nicht zerstört wird, sondern nach dem Scan das Original und seine digitale Kopie parallel existieren. Dann müsste man deren gegenseitiges Rechtsverhältnis klären. Zudem hält er es für möglich, dass fortgeschrittene Technik Informationen aus kryo-konservierten Gehirnen wiederherstellen kann, sodass Personen, die noch vor der technologischen Singularität starben, dennoch nachträglich hochgeladen werden. Eine letzte Alternative biete die Rekonstruktion Verstorbener aus einer Vielzahl an anderweitigen Datensätzen. In diesem Fall werde man die Summe an persönlichen Informationen – von medizinischen Daten, über persönliche Audio- und Videoaufzeichnungen, Schriftstücke wie die Zeugenaussagen Nahestehender – einer Künstlichen Superintelligenz einspeisen, die aus der Datenlage den Verstorbenen rekonstruiert und dieses Substrat hochlädt.⁵³⁸ Diese Machbarkeitsphantasien, die sich bis zu einer Erweckung der Toten steigern, werden eingeholt von der Sorge, das Hochladen werde nur scheinbar funktionieren und der Mensch stattdessen für ein schlechtes Imitat um seiner selbst betrogen. Die Rede ist immer wieder von neuen Testverfahren, mit denen man prüfen müsse, ob der *Upload* tatsächlich erfolgreich war. Kurzweil kommt zu dem Schluss, dass man in der Zukunft einen erweiterten Test ganz in der Tradition von Turings *Imitation Game* benötige, nach dem ein menschlicher Richter durch Befragung prüft, ob die digitale Kopie in ihrem Verhalten ununterscheidbar zu ihrem Original ist.⁵³⁹

Körperlose Freiheit

Unter der Voraussetzung, dass die Identität von Bewusstsein mit dem digitalen Substrat in der post-singulären Zukunft gewährleistet werden kann, erwarten viele *Extropians* im Übergang zur digitalen Existenz ein enormes Befreiungserlebnis. Ihre Hoffnungen treffen sich dabei im Begriff der *morphological freedom*. Diese definierte Max More 1993 mit

537 Vgl. Kurzweil, *The Singularity Is Near*, 217. Eine solche Technik, so vermutet Kurzweil, liegt bis spätestens Ende der 2030er Jahre vor.

538 Vgl. Chalmers, *The Singularity: A Philosophical Analysis*, 42.

539 Vgl. Kurzweil, *The Singularity Is Near*, 139–140.

der Fähigkeit, die eigene Körperform nach dem eigenen Willen technologisch zu verändern mithilfe von Chirurgie und Robotik, durch Nanotechnologie und schließlich dem *Upload*.⁵⁴⁰

Der elementare Wunsch, den die gewonnene Freiheit im Datenspeichersatz erfüllen soll, ist die Aussicht auf Unsterblichkeit. Wie es die Definition Mores veranschaulicht, steht der *Mind Upload* für viele Transhumanisten bzw. *Extropians* am Ende einer Stufenerzählung; in der kompletten Auflösung des Körpers erhofft man sich, dass weder Altern, weder Krankheit noch Unfälle etwas der *Software* – gemeint ist das Bewusstsein – werden anhaben können, indem die Hardware jederzeit austauschbar wird.⁵⁴¹ Mit der Sterblichkeit fallen demnach weitere Schranken für dieses entgrenzte Dasein: Körperlosigkeit bedeutet dann auch die Freiheit zur Selbstprogrammierung, in der das Selbst nicht mehr *fatum* ist (s. 4.3.2), sondern der digitalisierte Kulturmensch ganz zum Produkt seiner Wahl wird. Er kann durch *Enhancement* nicht nur jederzeit ein anderer werden, sondern auch all die Persönlichkeiten, die bisher nur latent ihm angelegt waren, im Wechsel der digitalen wie robotischen Körper simultan ausleben.⁵⁴² Wissenschaftlerinnen aus dem Spektrum des feministischen Aktivismus schließen an dieses Freiheitsversprechen an und sehen in einer Auflösung biologischer Körperlichkeit, wie sie Donna Haraway bereits 1985 in ihrem *Cyborg Manifesto* imaginiert, die historische Chance, die Zweiheit der Geschlechter und ihre Gewalt endgültig zu überwinden. Haraway postuliert, dass man erst mit der Auflösung des Menschen auch die patriarchale und rassistische Gesellschaft überwinden werde.⁵⁴³ Die amerikanische Anwältin, Unternehmerin und Transgender-Aktivistin Martine Rothblatt sieht Haraways Utopie im *Mind Upload* konkretisiert. In Ihrem Artikel *Mind is Deeper than Matter* erklärt Rothblatt die moderne Transgenderbewegung zu Vorboten jener neuen Menschheit, für die der zivilisierte Geist über den primitiven Körper und seine Biologie siegt. Die Einsicht, über das soziale Geschlecht frei bestimmen zu können, hält Rothblatt

540 Vgl. Max More, „Digital Self-Transformation: Expanding Personal Extropy“, *Extropy* #10 4 (1993).

541 Vgl. Kurzweil, *The Singularity Is Near*, 217–218.

542 Ebd., 211–212.

543 Vgl. Haraway, *A Cyborg Manifesto*, 175.

grundlegend für die folgende *freedom of form*, wie sie der *Mind Upload* in Aussicht stellt:

„Freedom of gender is, therefore, the gateway to a freedom of form and to an explosion of human potential. [...] The mind is the substance of humanity. Mind is deeper than matter.“⁵⁴⁴

Für Rothblatt vollzieht der Mensch in der Befreiung des Geistes von dem ihm vorgeordneten Körper schließlich den Entwicklungssprung vom *homo sapiens* hin zur *persona creatus*. Dafür müsste schon heute der Kampf gegen die strukturelle Gewalt von morgen ausgetragen und so dem *Somatizismus* gewehrt werden, der Bewusstsein ohne biologische Körper für minderwertig erklärt.⁵⁴⁵ Transhumanisten gehen nämlich davon aus, dass die Gegenwart den Körper in ungeheurer Weise ideologisiere. Dies meinen sie an den Vorbehalten gegenüber einer solchen Substratunabhängigkeit zu erkennen. Menschen äußern Bedenken darüber, sich vorschnell von der Sinnlichkeit des Körpers zu verabschieden. Max More führt die angebliche Romantisierung des biologischen Körpers auf die Propaganda der Filmindustrie zurück, die mit dem gefühllosen Androiden massenwirksam ein falsches Bild verbreite. Entgegen dieser Falschdarstellungen will More aufklären, dass *Extropians* und Transhumanisten Leidenschaft nicht verbannen wollen. Vielmehr spekuliert er darauf, dass die *Post-Bios* – von den ungenügenden Sinnen des *Affenkörpers* befreit – völlig neue, noch intensiviere Sinnlichkeit erfahren werden. Die Maschinenmenschen erwarte paradoxerweise eine *enhanced carnality*:

„It’s important to remember that the sensory abilities we possess today evolved in response to our survival needs in primitive environments. We no longer live in those primitive environments, nor have we any reason

⁵⁴⁴ Martine Rothblatt, „Mind is Deeper than Matter: Transgenderism, Transhumanism, and the Freedom of Form“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 317–326. Hier 318.

⁵⁴⁵ Vgl. ebd., 319.

to limit our sensory abilities to fit past conditions. We can benefit from using advanced technology to augment our senses.“⁵⁴⁶

Die ökonomische Lebenswelt der Emulated Minds

Der amerikanische Wirtschaftswissenschaftler Robin Hanson, der mit dem *Future of Humanity Institute* der Universität Oxford um Nick Bostrom assoziiert ist, stellt sich die Städte der Zukunft als glühende Serverfarmen vor, die für normale Menschen unbewohnbar sein werden.⁵⁴⁷ So trostlos diese neuen Städte von außen wirken, bildeten sie das Fundament einer digitalen Innenwelt, in der die EMS eine Existenz der unbegrenzten Wahl führen:

„But for ems, bigger life choices can be Add/And: shall I allow another me to have a new life? Big choices are often made jointly with copies, teammates, and investors. [...] On the upside, most ems have office jobs, and work and play in spectacular high-quality virtual realities without hunger, severe cold or heat, grime, physical illness, or intense pain – unless those are chosen deliberately. Unlimited lifespans are possible.“⁵⁴⁸

Hanson schließt unwissentlich an Teilhards Ideal eines technisch-utopischen Zwischenreiches (s. 4.3.1) an, wenn er versichert, dass Wissenschaft und Kunst die unbegrenzte Lebenszeit der EMS mit Sinn füllen werden.⁵⁴⁹ Die gesteigerten Entfaltungsmöglichkeiten werden zugleich wieder eingefangen in der Aussicht auf unbegrenzte Arbeit der EMS. Bereits Moravec nahm an, dass solche digitalen Entitäten ihren Lebensunterhalt, den lebensnotwendigen Speicherplatz, in einer freien Markt-

⁵⁴⁶ Max More, „The Enhanced Carnality of Post-Biological Life“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 222–230.

⁵⁴⁷ Vgl. Robinson Hanson, „What Will It Be Like To Be an Emulation?“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 298–309. Hier 299.

⁵⁴⁸ Ebd., 300.

⁵⁴⁹ Vgl. Brin, Broderick, Bostrom, Chislenko, Hanson, More, Nielsen, Sandberg, A Critical Discussion of Vingé's Singularity Concept, 397.

wirtschaft bestreiten müssten.⁵⁵⁰ Der Ökonom Hanson folgt dieser Idee und imaginiert verschiedene Berufsgruppen von *EMs*: Einige Entitäten wandeln unerschlossene Teile des Universums in Cyberspace um; andere machen bestehende Recheneinheiten effektiver und schaffen so neuen Wohlstand. Wieder andere agieren als Banken, die Ressourcen lagern sowie umverteilen, und gleichsam als Kunsthändler Strukturen schaffen, die nutzlos sind, aber von einigen Kunden als wertvoll erachtet werden. Wie Moravec erwartet Hanson, dass in dieser Hyperökonomie sich nur die wenigen, arbeitsamsten und intelligentesten *EMs* durchsetzen können. Nach kurzer Zeit gingen die meisten digitalen Substrate aufgrund des starken Selektionsdrucks auf Kopien von nur noch tausend oder weniger Menschen zurück, die sich in *Clans* organisieren und unter sich alle verfügbaren Ressourcen aufteilen.⁵⁵¹

Die Eroberung und Verwandlung des Weltraums in Lebensraum für EMs

„Our speculation ends in a supercivilisation, the synthesis of all solar-system life [...], spreading outward from the sun, converting nonlife into mind [...]. This process, possibly occurring now elsewhere, might convert the entire universe into an extended thinking entity, a prelude to even greater things.“⁵⁵²

Der Weg der Zivilisation muss für die extropische Teleologie von der Erde in den Weltraum führen: Mit dem *Mind Upload* soll dies nicht den *Mind Children* überlassen werden (s. 5.1), sondern die Menschen selbst werden zu robotischen Weltraumeroberern. Wie sehr die Hoffnung auf ein Leben als *mind file* mit dem Fernziel der Weltraumkolonisierung imprägniert ist, zeigt sich daran, dass selbst die erwähnte *2045 Initiative* ihre Projektphasen nicht mit dem *Upload* beendet, sondern mit

550 Hans P. Moravec, „Pigs in Cyberspace“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 177–181. Hier 180.

551 Hanson, *What Will It Be Like To Be an Emulation?*, 300.

552 Hans P. Moravec, *Mind Children. The Future of Robot and Human Intelligence*, Cambridge 1988, 116.

der Expansion in den Weltraum weiterdenkt. Schon 1929 stellte der Naturforscher John Desmond Bernal (1901–1971) die These auf, dass der Mensch der Zukunft sich mechanisieren müsse, um die Sterne dauerhaft zu bewohnen:

„[T]he colonization of space and the mechanization of the body are obviously complementary [...] If he could get rid of the major part of his body and his necessity for a relatively large intake of oxygen and water-saturated food, the cellular nature of the celestial globes would cease to be necessary.“⁵⁵³

Der *Mind Upload* übererfüllt Bernals Voraussetzungen. Wie Hans Moravec in seinem Vortrag *Pigs in Cyberspace* aus dem Jahr 1992 darstellt, befreit der Upload den Menschen in gleich zweierlei Hinsicht: Durch den *Upload* werden *Space* und *Cyberspace* – beides der menschlichen Konstitution unwirtliche Orte – bewohnbar. Mit seinem Roboterkörper kann der post-biologische Mensch den Weltraum kolonisieren. Schließlich sollen Netzwerk und Weltraum gänzlich eins werden. Nach Moravec muss das Ziel sein, das Universum und all seine Ressourcen in platzeffizienteren *Cyberspace* und damit neuen Lebensraum für *EMs* umzubauen.

„The final frontier will be urbanized, ultimately into an arena where every bit of activity is a meaningful computation: the inhabited portion of the universe will be transformed into a cyberspace. Because it will use resources more efficiently a mature cyberspace of the distant future -will be effectively much bigger than the present physical universe. While only an infinitesimal fraction of existing matter and space is doing interesting work, in a well-developed cyberspace every bit will be part of a relevant computation or storing a useful datum.“⁵⁵⁴

553 John Desmond Bernal, *The World, The Flesh And The Devil. An Enquiry into the Future of the Three Enemies of the Rational Soul*, London 1929, 73.

554 Moravec, *Pigs in Cyberspace*, 177.

Allein das hiesige Sonnensystem verheißt demnach riesiges, unerschlossenes *Computingmaterial*. Der Ausgriff auf die gesamte Materie und Energie des Weltraums nährt den Wunsch einer tatsächlichen *All-Beherrschung*, in der alle Masse von *EMS* bewohnt wird:

„I estimate that a human brain-equivalent can be encoded in less than 10^{15} bits. [...] So a city of a million human-scale inhabitants might be efficiently stored in 10^{24} bits. If the atoms of the solar system were cleverly rearranged so every 100 could represent a bit, then a single solar system could hold 10^{30} cities – far more than the number (10^{22}) of stars in the visible universe! Multiply that by 10^{11} stars in a galaxy, and one gets 10^{41} cities per galaxy. The visible universe, with 10^{12} galaxies, would then have room for 10^{51} cities –except that by the time intelligence has expanded that far, more efficient ways of using spacetime and encoding data would surely have been discovered, increasing the number much further.“⁵⁵⁵

Kurzweil fragt nicht *ob*, sondern nur *wie lange* es brauchen wird, um das gesamte Universum mit Intelligenz zu sättigen.⁵⁵⁶ Für das *Wie* wird sich wiederum an Weltraumkolonisierungsphantasien des 20. Jahrhunderts orientiert, die nun an die Identifikation von Mensch und Maschine angepasst werden. Bostrom glaubt, man werde Sonden nach Plänen des Mathematikers von Neumann in den Weltraum starten, die zu interstellaren Reisen fähig sind und auf ihrer Reise Asteroiden, Planeten und Sterne nutzen, um Kopien von sich selbst herzustellen. Diesmal nur reisen die *EMS* auf der replizierenden Sonde gleich mit und kolonisieren so einen beträchtlichen Teil des Universums.⁵⁵⁷ Kurzweil hingegen löst das Energieproblem für die Sättigung einer solchen Expansion ins Weltall mit dem Bau von Dyson-Sphären, wie sie gleichnamiger Astrophysiker 1959 vorschlug. Eine dünne Kugel werde dabei um einen aktiven Stern gebaut, die die ausgehende Strahlenenergie sammelt. Während Dyson sich noch vorstellte, dass in den Schalen der Kugel bewohnbare

555 Ebd., 178.

556 Vgl. Kurzweil, *The Singularity Is Near*, 243.

557 Vgl. Bostrom, *Superintelligence*, 100.

Zonen geschaffen werden müssen, auf denen Menschen wohnen können, bevölkerten diese Schalen nun direkt die *Emulated Minds*.⁵⁵⁸

Während Kurzweil oder Bostrom ihre Visionen mit der Verwandlung des Alls in einen riesigen Computer beschließen, erwarten *Singularitans* der ersten Generation wie Moravec und der Mathematiker Frank Tipler (*1947) am Ende der Informatisierung des Universums einen letzten Übersprungsmoment: Tipler glaubt, dass, wenn alle existierende Materie in Rechenkraft verwandelt wurde, es zu einer zeittranszendenten Weltsimulation kommt, die die Auferweckung aller Toten als mediales Ereignis bedeute. Alle Informationen, die in der gesamten menschlichen Geschichte enthalten sind, einschließlich jedes Detail eines jeden Lebens, werde dem informatisierten Universum in der fernen Zukunft verfügbar sein, um eine äußerst genaue Simulation dieser vergangenen Leben zu erstellen und kybernetisch zu erwecken.⁵⁵⁹ Und auch Hans Moravec glaubt, dass schließlich die gesamte Erde als eine Art Reservat von einem gigantischen Computer im All simuliert werden wird: Künstliche Intelligenzen, die ihre Rechenkraft aus dem gesamten Kosmos beziehen, würden dann alle möglichen Vergangenheiten rekonstruieren, die als dauerhafte Simulationen die letzten Menschen und alle Auferstandenen in der digitalen Welt auf ewig beheimaten werden.⁵⁶⁰

5.3.2 Ethische Relevanz: Expansion und Erlösung

„The Singularity holds out the possibility of winning the Grand Prize, the true Utopia, the best-of-all-possible-worlds – not just freedom from pain and stress or a sterile round of endless physical pleasures, but the prospect of endless growth for every human being – growth in mind, in intelligence, in strength of personality; life without bound, without end; experiencing everything we’ve dreamed of experiencing, becoming

558 Vgl. Kurzweil, *The Singularity Is Near*, 234.

559 Vgl. Frank J. Tipler, „The Omega Point as Eschaton: Answers to Pannenberg’s Questions for Scientists“, in: Carol Rausch Albright/Joel Haugen (Hgg.), *Beginning with the End: God, Science, and Wolfhart Pannenberg*, Chicago 1997, 156–194. Hier 185–186.

560 Vgl. Moravec, *Computer übernehmen die Macht*, 160–170.

everything we've ever dreamed of being; not for a billion years, or ten-to-the-billionth years, but forever... or perhaps embarking together on some still greater adventure of which we cannot even conceive. That's the Apotheosis.⁵⁶¹

Die Aussicht, mit der intelligenten Maschine zu verschmelzen, stellt die große Belohnung des Menschen dafür dar, unter Einsatz seines Lebens bis zum Schluss an Wissenschaft, Technik und Fortschritt festgehalten zu haben. Es ist sein Anteil daran, dass seine Technologie das gesamte Universum informatisch *beleben* und den Kosmos vor der drohenden Entropie retten wird. Dem Menschen wird in seinem Dienst die *Gnade* zuteil, an dem von ihm initiierten, informatischen Auftrieb des Kosmos zu partizipieren. Yudkowskys Verheißung einer Vergöttlichung des Menschen steht in Gesellschaft zu Hararis Prognose eines *homo deus* und Rothblatts Ideal der *persona creatus*, die sich selbst zum besseren Schöpfer *ihrer selbst* einsetzt.

Gleichzeitig darf nicht außer Acht gelassen werden, dass hinter der Verheißung der *morphological freedom* eine Drohung steht; schließlich gerät der *Mind Upload* unter dem Diktum der drohenden Konvergenz zum einzigen Szenario, in dem der Mensch sich in der postsingulären Welt überhaupt behaupten kann. Denn die Visionäre der Künstlichen Intelligenz sind sich einig, dass der Mensch am ehesten die Kontrolle über die Technik behält, wenn er und die Maschine untrennbar zusammenwachsen.⁵⁶² Tipler fasst diese Überlegung bündig in der Formel zusammen: „*If you can't beat 'em, join 'em.*“⁵⁶³ Doch selbst wenn nicht die Vernichtung des Menschen drohe, sind sich viele *Singularitans* sicher,

561 The Singularitarian Principles, Eliezer S. Yudkowsky: The Singularitarian Principles Version 1.0.2.

562 Vgl. Ben Goertzel, „Artificial Intelligence and the Future of Humanity“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 128–137. Hier 136.

563 Vgl. Frank J. Tipler, „If You Can't Beat 'em, Join 'em — Frank Tipler“, in: John Brockman (Hg.), *What to think about Machines that think. Today's leading Thinkers on the Age of Machine Intelligence*, New York/London/Toronto/Sydney/New Delhi/Auckland 2015, 30–31.

werde der Mensch seine Mechanisierung vorantreiben müssen, um nicht in einer Welt zu leben, die er nicht mehr versteht:

„Langlebigkeit verliert viel von ihrem Reiz, wenn sie uns überwiegend dazu verdammt, mit dummem Gesicht unseren überintelligenten Maschinen zu lauschen, während die versuchen, uns ihre immer spektakuläreren Entdeckungen in einer Babysprache zu beschreiben, die wir eben gerade noch verstehen können. Wir möchten gern gleichberechtigte Mitspieler in diesem superintelligenten Vergnügen sein.“⁵⁶⁴

Nirgendwo ist Günther Anders Diagnose der Selbst- und Dingscham unter einer dünneren Schicht an Ironie verdeckt als in Moravecs Rede vom Menschen als einem unverständigem Schwein, dem die KI ihr Wissen nicht mehr zu vermitteln vermag.⁵⁶⁵ Die Szenarien um den *Mind Upload* schreiben Anders Prognose fort, nach welcher der Mensch schlussendlich in das Lager der Geräte desertieren wird, um ihren Forderungen zu genügen.⁵⁶⁶ Während Anders dem *Human Engineering* eine Verirrung in die Unfreiheit unterstellte, verkündet der technologische Transhumanismus die Anpassung an die Geräte zum Emanzipationsmoment des Menschen, durch den er dem bevormundenden Schoß von *Mutter Natur* entwächst:

„Mother Nature, truly we are grateful for what you have made us. No doubt you did the best you could. [...] You seem to have lost interest in our further evolution some 100,000 years ago. Or perhaps you have been biding your time, waiting for us to take the next step ourselves. Either way, we have reached our childhood's end. We have decided that it is time to amend the human constitution. [...] We intend to make you proud of us.“⁵⁶⁷

564 Moravec, *Mind Children*, 151.

565 Vgl. Moravec, *Pigs in Cyberspace*, 181.

566 Vgl. Anders, *Die Antiquiertheit des Menschen*, 29–30.

567 Max More, „A Letter to Mother Nature. Original 1992“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 449–450. Hier 449.

In der Auseinandersetzung mit der Endlichkeit eines jeden Menschen als auch der Überlebensfähigkeit der menschlichen Spezies als Ganzer, der Hoffnung auf ein Leben über den biologischen Tod hinaus und der Vision eines Endes aller Zeit operiert das Identifikationsmotiv somit auf der Basis christlicher Eschatologie.

Der Transhumanismus rationalisiert den Wunsch nach Unsterblichkeit

Die eigene Minderwertigkeit soll im transhumanistischen Programm, in den sich der *Mind Upload* einreicht, nicht weiter beklagt, sondern durch Technologie überwunden werden. Ihren Anfang nimmt der moderne Transhumanismus mit dem Evolutionsbiologen und erstem UNESCO Generaldirektor Julian Huxley. Er popularisiert Mitte des 20. Jahrhunderts den Begriff des Transhumanismus, der vieles von dem weiterführte, dass vor dem Zweiten Weltkrieg noch unter der Eugenik verhandelt wurde:

„The zestful but scientific exploration of possibilities and of the techniques for realizing them will make our hopes rational, and will set our ideals within the framework of reality, by showing how much of them are indeed realizable. [...] The human species can, if it wishes, transcend itself – not just sporadically, an individual here in one way, an individual there in another way, but in its entirety, as humanity. We need a name for this new belief. Perhaps transhumanism will serve: man remaining man, but transcending himself, by realizing new possibilities of and for his human nature.“⁵⁶⁸

Huxley ist darauf bedacht, seine Hoffnungen dem wissenschaftlichen *logos* zuzuordnen. Dieses Selbstbild einer *rational hope* wird gebrochen, wenn man sich mit der Begriffsgeschichte des Transhumanismus vertraut macht. Sie reicht mit Dante Alighieris *La Divina Comedia* zurück zu einem der einflussreichsten theologischen Werke des späten Mittelalters. Dort bezeichnet das italienische Lehnwort *trasumanar* den Übertritt des Reisenden aus den Höllenkreisen hinein in das Paradies. Im

568 Julian Huxley, „Transhumanism“, *Journal of Humanistic Psychology* 8 (1968), 73–76.

Sinne eines *über die menschliche Natur Hinausgehen* erfährt der Mensch mit dem Übertritt die Läuterung ins Göttliche; sein Gegensatz, das Zurückfallen in die tierische Natur im *disumanar*, taucht bei Dante zuvor in der Beschreibung der verlorenen Seelen in den Höllenschichten auf.⁵⁶⁹ Dem *trasumanar* stellt Dante poetisch zugleich eine Unsagbarkeitsformel bei, mit der er deutlich auf Paulus unsagbares Erlebnis einer Entrückung in den dritten Himmel (2 Kor 12,2ff.). *Trans*-humanismus und religiöse *Trans*-zendenz sind entgegen aller szientistischer Rhetorik eng miteinander verknüpft.

Während Huxley vage darin bleibt, welche konkreten biologischen Beschränkungen er zu überwinden sucht, treibt der Physiker Robert Ettinger den folgenden Selbstklärungsprozess des Transhumanismus entscheidend voran. Mit seinen Schriften *The Prospect of Immortality* (1962) und *Man into Superman* (1972) schuf Ettinger das theoretische Fundament, auf dem er 1976 sein *Cryonics Institute* gründete. Dort werden seither Verstorbene in Flüssigstickstoff konserviert mit dem Versprechen, sie eines Tages durch fortgeschrittene Technik aufzuerwecken und ihnen ein unsterbliches Leben zu ermöglichen.⁵⁷⁰ Spätestens mit Ettinger setzt sich innerhalb der Bewegung die Überzeugung durch, dass es nichts geringeres als die Sterblichkeit des Menschen selbst ist, die überwunden gehört. Die Hoffnung auf Unsterblichkeit steht deshalb auch im Zentrum des *Mind Uploads* und seiner Rede von morphologischer Freiheit, *substrate identity* oder *freedom of form*. Denn allein durch ein unsterbliches Dasein können dem Menschen die bis dahin zu leistenden Risiken und Opfer wiedergutmacht werden.⁵⁷¹ Die immer wieder betonte Distanz zum Mythos hat sich seither selbst zum narrativen Element transhumanistischer Mythen verstetigt. Gleichzeitig ist offensichtlich, dass die Sehnsucht nach Unsterblichkeit gleich mehrfach von Mythen sedimentiert ist.⁵⁷² Ein so umfassendes Erlösungsverspre-

569 Vgl. Bodo Guthmüller, „*Trasumanar significar per verba / non si poria*«. Zu Dantes erstem Gesang des *Paradiso*“, *Deutsches Dante-Jahrbuch* 82 (2007), 67–86. Hier 68.

570 Vgl. Mathias Wirth, „Trans-Körper. Theologie im Gespräch mit Transhumanismus und Transsexualität“, *Zeitschrift für Evangelische Ethik* 62 (2018), 10–30. Hier 18–19.

571 Vgl. Krüger, *Virtualität und Unsterblichkeit*, 150.

572 Vgl. Flessner, *Die Rückkehr der Magier. Die KI als lapis philosophorum des 21. Jahrhunderts*, 95.

chen wie die Aussicht, durch den *Mind Upload* nicht länger sterben zu müssen, zwingt den Vergleich zu anderen Erlösungsreligionen wie dem Christentum unweigerlich auf. In dieser Auseinandersetzung erscheint aus Perspektive der *Singularitans* nicht die transhumanistische Rationalisierung der Unsterblichkeit das Problem; im Gegenteil werfen die Transhumanisten dem Christentum vor, den Tod vorschnell rationalisiert und sogar überhöht zu haben. Spätestens mit dem Wissen um die bevorstehende Singularität und ihrer Möglichkeit der Überwindung eines biologischen Todes müsse damit Schluss sein. An die Stelle der klassischen Religionen – so schlägt Kurzweil es vor – müsse eine neue Religion treten, die nicht den irdischen Tod für unabänderlich erklärt, sondern aktiv an seiner tatsächlichen Überwindung mitarbeitet.⁵⁷³

Die Erwartung einer innerweltlichen Transzendenz

Die Transhumanisten, *Singularitans* wie *Extropians* geben sich als irdische Erfüller uralter Menschheitsträume, indem der Mensch in die Unsterblichkeit seiner Maschine eingeht. Günther Anders kam Mitte der 1950er Jahre zu dem Schluss, der Kulturmensch sitze in seiner Dingscham dem Trugschluss auf, die Beständigkeit der ihn umgebenden, gefertigten Serienprodukte zu seinen Idealbildern des Ewigen zu überhöhen. Hinter den Hoffnungen auf eine Unsterblichkeit im *Mind Upload* steht in demselben Maße keine Sehnsucht nach einer dem Kulturraum und der Materialität enthobenen, philosophischen Erfahrung des Ewigen.⁵⁷⁴ Vielmehr geht es darum, an der erhofften Unsterblichkeit der Kultur selbst zu partizipieren, die im Fall der KI-Visionäre noch einmal an Attraktivität gewinnt, weil sie von einem hyperexponentiellen Fortschritt ausgeht. Transhumanisten und *Singularitans* verwechseln daher nicht die Ewigkeit mit diesseitiger Unsterblichkeit, sondern haben bewusst kein Interesse an einer Ewigkeit, die sie nicht mehr an dem technologischen Fortschritt und seinen planetaren Gestaltungsräumen teilhaben lässt.

573 Vgl. Kurzweil, *The Singularity Is Near*, 218.

574 Zur Unterscheidung zwischen Unsterblichkeit und Ewigkeit s. Arendt, *Vita activa oder Vom tätigen Leben*, 30–31.

„Virtual bodies are not heavenly bodies, but perhaps they are sufficiently similar in some ways to offend the adherents to pre-scientific worldviews.“⁵⁷⁵

Die Befürworter des *Mind Uploads* reflektieren, dass ihre Versprechen nicht an die Transzendenzhoffnungen der klassischen Religionen heranreichen. Stattdessen gerieren sie ihre Vorstellungen als deren erfüllbare, in *logos* gegossene Äquivalente. Mehr noch: Sie deuten an, dass die diesseitige Unsterblichkeit in der Post-Singularität die bessere Option von beiden sei. Wofür einen ewigen Himmel suchen, wenn man ein Universum erobern kann? Die post-singuläre Unsterblichkeitshoffnung sucht in den Worten Ernst Blochs nach einem „Transzendieren ohne alle himmlische Transzendenz“⁵⁷⁶. Es geht nicht mehr um zwei völlig getrennte Seinsstufen, sondern um eine Differenzierung innerhalb der Immanenz selbst. In den Grenzen dieser Welt und der Stofflichkeit dieses Lebens soll ein qualitativer Sprung des Bestehenden möglich werden:⁵⁷⁷

„The Singularity denotes an event that will take place in the material world, the inevitable next step in the evolutionary process [...]. To ‚transcend‘ means ‚to go beyond‘, but this need not compel us to adopt an ornate dualist view that regards transcendent levels of reality (such as the spiritual level) to be not of this world. We can ‚go beyond‘ the ‚ordinary‘ powers of the material world through the power of patterns. [...] Since the material stuff of which we are made turns over quickly, it is the transcendent power of our patterns that persists.“⁵⁷⁸

Kurzweils Muster existieren in keinem anderen Reich der Ewigkeit, sondern in dem vom Menschen geschaffenen und auf irdische Maschinen angewiesenen *Cyberspace*. Dass die materielle Wirklichkeit für eine

575 More, *The Enhanced Carnality of Post-Biological Life*, 226.

576 Vgl. Bloch, *Das Prinzip Hoffnung*, 1402.

577 Vgl. Michael Eckert, *Transzendieren und immanente Transzendenz*. Die Transformation der traditionellen Zweiweltheorie von Transzendenz und Immanenz in Ernst Blochs Zweiseitentheorie, Wien 1981, 97.

578 Kurzweil, *The Singularity Is Near*, 259.

erfüllte Unsterblichkeit taugt, soll gesichert sein durch die Bestimmung des Fortschrittes zu einer Naturkonstante: Wissenschaft und Kunst werden überhöht zu unerschöpflichen Betätigungsfeldern und Quelle des immer Neuen auf der Spielfläche eines grenzenlosen Universums:

„The hope of eternal worldly progress and the hope of individual survival beyond the grave turn out to be the same. Far from being polar opposites, these two hopes require each other; we cannot have one without the other.“⁵⁷⁹

Der Ausgriff auf die Sterne steigt in diesem monistischen Materialismus interessanterweise wieder zur höchsten Idee einer sichtbaren Transzendenz auf und schließt damit an die an die ältesten Sternenkulte an. Die Aufnahme einer Sternenfahrt in die Konvergenzerzählungen um Künstliche Intelligenz behält derweil die ursprüngliche Romantik des *Sci-Fi-Tropes* bei: Der Mensch löst sich von allem Irdischen, damit auch dem Gewordenen und Bestehenden, und dringt in einen unbeschriebenen *Frei-Raum* vor. Allein der Körper hat den klassischen Raumfahrer bisher noch an sein ursprüngliches Habitat und damit dessen Regeln und Verantwortungsbereich Erde rückgebunden. Den Körper endgültig aufgeben zu können, lässt den Menschen seine irdische Heimat verlieren, um im Universum heimisch werden. War es in Anders Mythos noch die Welt, die in einen einzigen Apparat verwandelt werden sollte, ist für die *Emulated Minds* das Ziel, den Maschinenpark in den Weltraum auszuweiten und sich sämtliche existierende Masse einzuverleiben.⁵⁸⁰

Erlösungshoffnung zwischen gnostischer Weltflucht und alchemischer Weltverwandlung

Der Wunsch, sich des Körpers zu entledigen und im *Mind Upload* freier Geist zu werden, ist weniger ein Gedanke der christlichen Orthodoxie, sondern nimmt heterodoxe Traditionsbestände auf. Darauf wies bereits

579 Tipler, *The Omega Point as Eschaton: Answers to Pannenberg's Questions for Scientists*, 189.

580 Vgl. Anders, *Die Antiquiertheit des Menschen*, 110–113.

1998 der amerikanische Journalist Erik Davis mit seinem Werk *Techgnosis* hin, in dem er die erste Generation an *Singularitans* um Hans Moravec zu modernen Gnostikern erklärte. Seiner Ansicht nach erstreben sie im *Upload* eine moderne Katharsis durch die Nachfahren der nun irdisch gewordenen, platonischen Idealformen der binären Logik und der Informationstheorie.⁵⁸¹ Gleichzeitig bescheinigt er der antiken Gnosis informationstheoretische Züge, wenn in ihren Erzählungen der befreiende *logos* häufig in Form einer *Nachricht aus der Ferne* ihren Empfänger erreicht.⁵⁸² 2017 nahm der französische Mathematiker und Philosoph Jean-Gabriel Ganascia in seinem Werk *Le Mythe de la Singularité* den hermeneutischen Schlüssel auf und explizierte das gnostische Paradigma an den neueren Technomythen um Künstliche Intelligenz: Wie die gnostischen Denker lehnten die Befürworter der technologischen Singularität die vorfindliche Natur ab und wollten den Geist vollständig vom Fleisch trennen. Das Defizit alles Vorfindlichen beurteilten sie beide auf Grundlage eines verborgenen Wissens, sei es die Botschaft des fremden Erlösergottes oder der evolutionistische Aufstieg von Information. Dieses Wissen versetze Gnostiker wie die Propheten der nahenden Superintelligenz in ein Bewusstsein um die gebrochene Zeit (s. 2.2.3). Überhaupt mache der Singularitätsmythos wie die Gnosis die Rationalisierung des Übernatürlichen zur Grundlage weitreichender Spekulation, in denen Gedanken verschiedenster Provenienz in quasi-rationale Denksysteme integriert werden. Indes erwiesen der technologische Transhumanismus wie die Gnosis ihre tiefere Kenntnis darin, dass sie bestehende Tradition radikal umdeuten. Während die christliche Gnosis das Leiden und Sterben Christi als Lüge verwarfen, ruft der Transhumanismus im Namen des aufklärerischen Humanismus zum Ende des Menschen auf.⁵⁸³

Über diesen allgemeinen Vergleich ist für das Motiv des *Mind Uploads* auffällig, dass das *Hinaus zu den Sternen* für sich bereits eine gnostische Blickrichtung darstellt. Der antiken Gnosis ging es mehr als alles andere um die Flucht aus dieser Welt, die sie im Geiste Mar-

581 Vgl. Davis, *Techgnosis*, 123–125.

582 Ebd., 96–99.

583 Vgl. Ganascia, *Le Mythe de la Singularité*, 69–70.

cions als *cellula creatoris begriff*, als innerstes Verlies eines kosmischen Gefängnisses. Gleichsam verkehrte die Gnosis die semitischen Sternengötter, die ehemals den Menschen auf der Erde beschützten, zu dämonischen Gefängniswärtern. Diese halten die Menschen nun auf der Erde gefangen und hindern den menschlichen Geist von ihren Himmelsphären aus daran, durch den Kosmos hindurch bis zum Reich des wahren Erlösergottes aufzusteigen.⁵⁸⁴ An dem Stellenwert der Sterne zeigt sich zugleich die Schwäche des gnostischen Paradigmas für den Singularitätsmythos, wie er in dem Identifikationsmotiv und dem *Mind Upload* konkret wird: Das gnostische Dual von pneumatischem Himmel und materieller Hölle kann mit einer transzendenzlosen Transzendenz nichts anfangen.⁵⁸⁵ Gnostiker sahen auch in den Sternen noch die Außenbereiche ihres Gefängnisses; die wirkliche Erlösung verortete man jenseits der Sterne, wo keine Materie, sondern nur das göttliche *pneuma* hinreicht, anstatt wie im *Mind Upload* tatsächlich in den Sternen selbst die computierbare Erlösung zu suchen. Hier kommt die Analogie der Erlösungshoffnung des technologischen Transhumanismus mit der klassischen Gnosis an ihr Ende. Die Vorstellung, sich nicht von der abgefallenen Materie abzuwenden, sondern sie als den Werkstoff zu betrachten, aus dem im eigenen Mitwirken das Vollkommene hervorgeht, bindet die Vorstellungen des *Mind Upload* wiederum eng an magische Weltbilder wie das der Alchemie:

„Die Alchemisten waren Gnostiker, die an eine ersehnte Welt jenseits dieser Welt glaubten, aber sie waren Gnostiker, die zugleich diese Welt gar nicht verlassen wollten. [...] Es ging um Erlösung, aber um Erlösung von der Welt in der Welt. Diese im Grunde paradoxe Erlösung fand statt und konnte nur stattfinden in einer Welt, in der sich feinmateriell gedachter Geist und grobe Materie verschränkten.“⁵⁸⁶

584 Vgl. Hans Jonas, *Gnosis. Die Botschaft des fremden Gottes* (Taschenbuch / Verlag der Weltreligionen 8), 1. Aufl., Frankfurt am Main/Leipzig 2008, 305–311.

585 Bereits Ganascia räumt ein, dass jeder Versuch, auf materielle Mittel zur Erlösung zurückzugreifen, nach gnostischer Sicht vergeblich ist, da man so am künstlichen Reich des falschen Gottes teilnehme. Vgl. Ganascia, *Le Mythe de la Singularité*, 68.

586 Hans-Werner Schütt, *Auf der Suche nach dem Stein der Weisen. Die Geschichte der Alchemie*, München 2000, 542.

Die Alchemie geht von einer verlebendigten, erlösungsbedürftigen Materie aus; wie der Mensch strebt die Natur zur Vollkommenheit.⁵⁸⁷ Gelingt es dem Magier bzw. Alchemisten, Macht über diesen Geist in der Materie zu gewinnen, erlangt er auch Einfluss auf deren Form und Qualität, um Unvollkommenes zu vervollkommen. Vervollkommnung, da sind sich der technologische Transhumanismus wie die Alchemie einig, bedeutet in letzter Konsequenz die Kontrolle über das Leben selbst.⁵⁸⁸ Denn auch für Alchemisten war klar, dass das letzte Geheimnis aller Wandlungen der Materie das Geheimnis ihrer Schöpfung durch Gott ist. Lüftet man dieses Geheimnis, kann der gottgleiche Alchemist eigens Stoff verlebendigen. Dieses Ziel hat der Mythos um Künstliche Intelligenz bereits über die Kybernetikbewegung übernommen (s. 4.2.4); schließlich prophezeite schon Gotthard Günther dem kybernetischen Programm, dass es den Odem Gottes, den der Mensch nur als Leihgabe empfangen habe, an die tote Materie weitergeben werde, „sodass die Transzendenz des Materiellen durch seine Introszendenz bereichert werde“⁵⁸⁹. Der Panvitalismus der Alchemisten und der Panthanatismus der *Singularitans* kommen dabei zu demselben Ergebnis: Weil für die kybernetische Epistemologie in ihrem dataistischen Monismus Lebewesen nichts weiter als komplexe Automaten sind, können die *Singularitans* freimütig in die alchemistische Hoffnung auf einer Erlösung und Verlebendigung aller Materie in höchstem Grade einstimmen, weil für sie der Unterschied zwischen Leben und Tod bloß noch ein Gradient in Informationsverarbeitung bedeutet.⁵⁹⁰

„So if our old perceived limits of life can be shattered by technology, what are the ultimate limits? How much of our cosmos can come alive? [...] Our planet is currently 99.999999 % dead in the sense that this fraction

587 Vgl.ebd., 538.

588 Flessner, Die Rückkehr der Magier. Die KI als lapis philosophorum des 21. Jahrhunderts, 78–81.

589 Günther, Das Bewußtsein der Maschinen, 181.

590 Hans Jonas unterscheidet zwischen den monistischen Extremen eines animistischen Panvitalismus, in dem alle Wirklichkeit belebt – weil beseelt – ist und einem Panthanatismus, der das Tote zum Ausgangspunkt allen Seins macht und für den das scheinbar Lebendige nur eine komplexe Form des Toten ist. Vgl. Jonas, Das Prinzip Leben, 28–31.

of its matter isn't part of our biosphere and is doing almost nothing useful for life other than providing gravitational pull and a magnetic field.⁵⁹¹

Für die Alchemie ist der Stein der Weisen das ultimative Transmutationsmittel, wie es für die Konvergenzerzählungen die aus der Singularität hervorgehende Künstliche Intelligenz ist. Alles, was mit dem Stein in Berührung kommt, steigt in einen vollkommenen, erlösten Zustand auf, sodass es zu Gold oder lebendig wird. Mit seiner Erschaffung wird die Erlösung aller Materie zum bloßen Kinderspiel, das nahezu *automatisch* geschieht. Der Stein selbst ist das Sinnbild für die Suche der Alchemie nach dem greifbar Transzendenten in der Welt, denn ihm ist als Objekt der existentielle Sprung gelungen. Das zeigt nicht zuletzt in die Überzeugung der Alchemisten, dass der Stein sich im Gebrauch nicht abnutzt.⁵⁹² Gleichmaßen soll im Motiv des *Mind Uploads* die Künstliche Intelligenz in Gestalt der digitalen Substrate das gesamte Universum praktisch zu Nullkosten informatisieren.

„The importance of the Mind Fire does not lie in the material presence of intelligent machines but in the cyberspace created by their computations. Moravec is a modern day alchemist, seeking to make gold out of lead, life out of death.“⁵⁹³

Im Geistfeuer der Digitalisierung wird alle Materie nicht nur bloß verlebendigt, sondern Visionäre einer postsingulären Wirklichkeit, die der Denktradition Teilhard de Chardins nahestehen, erhoffen sich in diesem Prozess zugleich auch die Humanisierung aller Wirklichkeit. Denn indem die *Emulated Minds* in alle Materie *einwohnen*, wird das menschliche Prinzip entgegen seiner beschworenen Bedeutungslosigkeit wieder zum notwendigen Mittelpunkt einer Aufstiegserzählung für den gesamten Kosmos restituiert:

591 Tegmark, *Life* 3.0, 203–204.

592 Vgl. Schütt, *Auf der Suche nach dem Stein der Weisen*, 357.429.

593 Geraci, *Apocalyptic AI: Religion and the Promise of Artificial Intelligence*, 151–152.

„Ultimately, the entire universe will become saturated with our intelligence. This is the destiny of the universe.“⁵⁹⁴

Indem der menschliche Geist in Form der Emulated Minds in allen Dingen eingeht, wird das Universums gleichsam für ihn ansprechbar:

„Es findet eine Revolution der Dingwelt statt, wenn die Stummheit der Objektwelt überwunden wird. [...] Die Hybridbildung von materiellen Dingen mit intelligenten Systemen der Identifikation und Selbstauskunft kommt alten Hoffnungen entgegen, die auf die Magie gesetzt wurden, die Dinge sprechen zu machen.“⁵⁹⁵

Die Überwindung von Innen und Außen geschieht nicht, indem der Mensch im Sinne der Naturreligion in die Welt eingeht, als vielmehr nun alle Natur durch den menschlichen Geist im Medium des Digitalen vereinnahmt wird. Die daraus folgende Ansprechbarkeit aller Dinge sucht nicht nach Verständigung, sondern impliziert ihren absoluten Gehorsam vor dem menschlichen Willen. Der alchemistische Techno-Animismus, der den Menschen mit der nichtsprechenden Natur vereinen soll, birgt in letzter Konsequenz eine Gewaltphantasie, nach der sich alle Wirklichkeit dem eigenen Befehl zu beugen hat. Bereits Sigmund Freud (1856–1939) erkannte in dem häufig anzutreffenden Animismus des Kindes den Wunsch nach Allmächtigkeit, indem die eigene Psyche in die Außenwelt verlegt wird – in der Annahme, auch sie unterstünde dem eigenen Willen. Im Erwachsenen hingegen sei diese Überschätzung der seelischen Vorgänge gegenüber der Realität ein Kennzeichen einer narzisstischen Störung und Folge der Weigerung, erwachsen zu werden.

594 Kurzweil, *The Singularity Is Near*, 37.

595 Peter Koslowski, „Mensch-Maschine-Hybride: Dinge, die sprechen, und Maschinen, die unsere Brüder werden“, in: Christoph Hubig/Peter Koslowski (Hgg.), *Maschinen, die unsere Brüder werden*. Mensch-Maschine-Interaktion in hybriden Systemen, Paderborn 2008, 191–202. Hier 193–195.

„Die Technik des Animismus, die Magie, zeigt uns am deutlichsten und unvermengtesten die Absicht, den realen Dingen die Gesetze des Seelenlebens aufzuzwingen [...]“⁵⁹⁶

Freiheit in der Geschichte

In Bezug auf die Auseinandersetzung um die Nähe und Distanz transhumanistischer Versprechen zur christlichen Eschatologie hält der Theologe Mathias Wirth fest, dass Häresien dadurch brisant werden, gerade weil sie meist nicht allzu entfernt von der Orthodoxie sind.⁵⁹⁷ Die Erlösungsschiffre der morphologischen Freiheit zumindest ist nicht bloß für verschiedenste Formen existentiell erfahrener Begrenzungen anschlussfähig, sondern die Endlichkeit des Menschen und die ihr entgegenstehenden Hoffnungen auf eine Überwindung von Krankheit und Leiden sind dieselben Quellen, aus denen gleichermaßen auch das Christentum mit ihrer Kreuzestheologie schöpft:

„Die grundlegende Lebensfrage aber ist gestellt durch die Unausweichlichkeit des Todes, die das Leben mit dem Gefühl der Sinnlosigkeit angesichts der Todverfallenheit aller menschlichen Werke und Gedanken bedroht. Erst auf dem Hintergrund des Todesgeschicks erweist sich der tiefe Gehalt der christlichen Rede von Sünde und Schuld; denn sie deutet die Todverfallenheit unseres Daseins als Ergebnis seiner Selbstbezogenheit.“⁵⁹⁸

Die Angst vor dem Tod teilen alle Menschen miteinander. Wollen sie aber alle deshalb dieses Leben unendlich weiterführen, wie es der Transhumanismus annimmt? Günther Anders meint dazu, die Bestimmtheit des Menschen gegenüber dem Tod bestehe ausschließlich in der Gewissheit, nicht sterben zu wollen, und damit in der Negation. Das

⁵⁹⁶ Sigmund Freud, „Über einige Übereinstimmungen im Seelenleben der Wilden und der Neurotiker. III. Animismus, Magie und Allmacht der Gedanken.“, *Imago. Zeitschrift für Anwendung der Psychoanalyse auf die Geisteswissenschaften* (1913), 1–21. Hier 14.

⁵⁹⁷ Mathias Wirth, „Doketisch, pelagianisch, sarkisch? Transhumanismus und technologische Modifikationen des Körpers in einer theologischen Perspektive“, *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie* 60 (2018), 142–167. Hier 149–152.

⁵⁹⁸ Pannenberg, *Das Glaubensbekenntnis*, 171–172.

andere könnten Menschen nicht greifen, dafür seien sie eben zu sterblich.⁵⁹⁹ Das Motiv der digitalen Unsterblichkeit durch Künstliche Intelligenz füllt anhand ihres Leitbegriffes der morphologischen Freiheit dieses Komplement zum Nicht-sterben-Wollen positiv aus. Die Art, wie diese Freiheit vom Tod eine solche Zukunft formatiert, markiert derweil, wie grundverschieden dieses Freiheitsverständnis zu der *libertas christiana* steht.

Die christliche Freiheit gegenüber dem Tod nimmt ihren Ausgang mit der Botschaft von der Auferstehung Christi. Wo mit dem Tod alle Beziehungen abbrechen, gilt das Bekenntnis zur Auferstehung als Zeichen dafür, dass die Beziehung des Menschen zu Gott über den Tod hinausreicht. Im Glauben wird der Christ in die Gemeinschaft mit dem Tode Jesu hineingenommen; niemand muss mehr alleine sterben, sondern jeder nimmt Anteil an Christi Schicksal, worin das Sterben seine Finalität und Trostlosigkeit verliert.⁶⁰⁰ Luther sagt, dass es kein Ding gibt, es sei gut oder böse, das nicht dem Glaubenden zugute dienen müsse und dass selbst der Tod dem, der glaubt, zur Seligkeit dienen kann, so groß sei die Gewalt der Christen und ihrer Freiheit.⁶⁰¹

Im Glauben findet der Mensch einen gelingenden Umgang mit den Umständen seiner leiblichen Endlichkeit – auch und gerade dann, wenn nichts Anderes Halt gibt. Christliche Freiheit wehrt somit der Lähmung, der absoluten Verzweiflung und der aus ihr erwachsenen Verzweiflungstaten in der Konfrontation mit der eigenen Endlichkeit. Entgegen der transhumanistischen Kritik, das Christentum habe den Tod verklärt, ist das christliche Selbstverständnis ein solches, den Tod weder zu bagatellisieren noch zu glorifizieren. Das Sterben und der Tod bleiben ein Übel in dieser Welt, denen sich jeder stellen muss, die jedoch nicht eine totale Macht auf den Menschen nach seinem Leben und damit auch nicht in seinem Leben haben. Anstatt wie ein Kind, das glaubt, niemals sterben zu müssen, den Tod zum Mythos umzukehren, gehört die Reflexion über die eigene Endlichkeit zum Erwachsenwerden des Menschen unweigerlich dazu. Der sich daraus einstellende

599 Vgl. Anders, Die Antiquiertheit des Menschen, 58–59.

600 Vgl. Pannenberg, Das Glaubensbekenntnis, 98.

601 Vgl. Luther, Von der Freiheit eines Christenmenschen (1520), 294–295 (WA27–28).

Appellcharakter, mit seiner begrenzten Zeit besonnen umzugehen, ist die Grundlage der individuellen Lebensführung als auch der politischen Verantwortlichkeit für die jetzigen und nachfolgenden Generationen.

Das postsinguläre Freiheitsversprechen, in der digitalen Welt alles ausleben zu können und damit alle Potentiale eines Lebens zur vollen Ausbildung zu bringen, lässt für die Zukunft nichts mehr übrig, das noch werden könnte; alles ist da schon voll entwickelt und ins Maximum gesteigert. Das einzige, was anscheinend immer noch gesteigert werden mag und unerschöpflich scheint, ist ökonomischer Wachstum und wissenschaftlicher Fortschritt. Der kindliche Traum, dass in der postsingulären Zukunft der Mensch sich werde ganz ausleben können und dazu alle Macht über die Dinge besitzen wird, verdammt den Träumenden der Gegenwart dazu, das vorfindliche, ohnmächtige Leben als durch und durch nicht gelingensfähig bewerten zu müssen. Dass unter den Möglichkeiten des endlichen Lebens die begrenzte Zeit nicht auch bereits sinnlos oder verloren ist, verknüpft die christliche Freiheit mit der Zusicherung von Gottes lebensschaffendem Geist. Dieser ist im Odem alles Lebenden gegenwärtig und verbindet dieses Leben mit dem neuen Leben der Auferstehung bei Gott.⁶⁰² Als Tröster ist der Geist Gottes dem endlichen Leben beigestellt und richtet es auf die heilvolle Perspektive des Reich Gottes aus, in deren Licht dieses endliche, partikuläre Leben bereits gewürdigt und geheiligt wird.⁶⁰³

Der christliche Glaube, und auch dies ist Wagnis wie Garantie zugleich, scheut nicht die Transzendenz. Der christliche Gott und seine Herrschaft sind nicht eine Fortsetzung der vorfindlichen Welt; Gott geht auch nicht erst am Ende als ein Gott-Omega im Sinne Teilhard de Chardins aus dem Prozess der Vervollkommnung der Materie hervor, sondern Gott der Schöpfer steht der Welt von Beginn an gegenüber und wird sie vollenden, indem er sie neu schafft. Blochs Einwand, die absolute Transzendenz sei für Menschen nicht erfahrbar und müsse deshalb in Utopien auf diese Erde geholt werden, verfehlt den besonderen Wert, den das Bekenntnis zu Christus als wahrer Mensch und wahrer Gott in sich trägt. Um hinter den *deus absconitus* zu blicken,

602 Vgl. Pannenberg, Das Glaubensbekenntnis, 146.

603 Vgl. Rendtorff, Ethik, 126–129.

braucht der Mensch nicht den Himmel erden, sondern in der Menschwerdung Gottes in Christus ist die volle Wirklichkeit Gottes erschienen und im Gottesgeist in der Welt fortwährend anwesend. In der scheinbar genügsamen, kleinen Transzendenz der digitalen Welt ist es hingegen der Mensch allein, der die Identität seiner selbst dauerhaft zu gewährleisten hat, will er *nach dem Tode Gottes* am ewigen bzw. unsterblichen Leben festhalten. Während der orthodoxe Seelenglauben den Einzelnen vergewissert, dass es durch den Tod hindurch bestehen bleibt und seine Identität in der Beziehung zu Gott gewahrt wird, kann die Wissenschaft hingegen nicht die Identität der digitalen „Seele 4.0“⁶⁰⁴ mit ihrem biologischen Original garantieren:

„Der Mensch, der sich virtuell verewigen möchte, steht in der Gefahr, am Ende um sich selbst betrogen zu werden. Statt der Freiheit vom Tod hätte er für die Aufgabe seiner Identität nur ein virtuelles Imitat gewonnen.“⁶⁰⁵

Wo kein Beziehungsgeschehen erwartet wird, fußt die morphologische Freiheit des *Mind Uploads* vor allem auf Selbstsorge. Der digitalisierte Mensch muss sich selbst darum kümmern, dass seine Serverstätten laufen. Über die volle Immanenz von Daten und die folgende Notwendigkeit, sich um diese beständig zu kümmern, stellt Kurzweil ernüchtert fest:

„[W]hat can we conclude about the ultimate longevity of software? The answer is simply this: Information lasts only so long as someone cares about it. [...] Information will continue to require constant maintenance and support to remain ,alive‘.“⁶⁰⁶

Dem Versprechen von der digitalen Unsterblichkeit folgt eine Existenz als digitales Substrat, das von einer ständigen Sorge um die Selbsterhaltung und der latenten Drohung der Löschung überschattet ist. Die Sicherung der eigenen Unsterblichkeit, die das Fundament der mor-

604 Vgl. Elis Eichener/Niklas Peuckmann, „Mind Uploading als Seele 4.0. Trauerkultur in Zeiten der Digitalisierung“, *Zeitschrift für Evangelische Ethik* 65 (2021), 114–127.

605 Ebd., 123.

606 Kurzweil, *The Singularity Is Near*, 221.

phologischen Freiheit bildet, ist deshalb verantwortlich dafür, weshalb das Ausüben dieser Freiheit in den dargestellten Visionen derart gewalttätig erscheint. Dies zeigt sich zum einen im Umgang mit der nichtmenschlichen Umwelt. Die Hoffnung auf Vollendung der Materie korreliert im Cyborgmotiv mit ihrer absoluten Unterwerfung. Das Leben aus dem *Cyberspace* heraus kann nur so geschützt werden, indem man sich in kosmischer Skalierung gegen sämtliche Unwägbarkeiten der materialen Welt meint abzusichern und alles in *Lebensraum* verwandelt. Die zweite Form der Gewalt ereignet sich unter digitalen Substraten im *Cyberspace*: Die virtuelle Realität, die einerseits Ort sämtlicher Selbstexpression der vom Körper befreiten *Emulated Minds* sein soll, ist gleichzeitig Ort eines brutalen, darwinistischen Überlebenskampfes unter dem Vorzeichen digitaler Exponentialität, in dem sich in einem Bruchteil von Sekunden die Menschenkopien zu Rotten zusammenfinden und anderen Kopien die Lebensgrundlage entziehen. Himmel und Hölle, Dantes Aufstieg des Menschen zu Gott im *trasumanar* und sein Abfallen in einen Tierzustand des *disumanar*, scheinen in dieser Technosphäre ununterscheidbar zu werden, weil nichts anderes herrscht als die bloße Gewalt als Durchsetzungsmittel eines unbedingten Überlebenswillens.

Digitale Engel

Aus der Perspektive der Gnosis erhöhen sich die Menschen, die als digitale Substrate auf das Universum ausgreifen und die Sterne besetzen, zu den neuen Gefängniswärtern des Kosmos – den Archonten –, die zwar keine Götter, aber auch keine Menschen mehr sind. Analog kennt auch die christliche Orthodoxie körperlose Wesen zwischen Gott und Mensch, die in dieser Welt präsent sind und mit ihr umgehen. Möglicherweise fasst man den *Upload* und sein Versprechen morphologischer Freiheit am ehesten über ein christliches Analogon, wenn man ihn als Versuch versteht, den Menschen in den Stand der Engel zu heben. Denn Engel wie EMS stellen die Anfrage, wie weit sie als körperlose Geistwesen dennoch in der Materialität dieser Wirklichkeit verhaftet sind, um als Zwischen- und Botenwesen weiterhin auf diese Erde zugreifen zu können. Thomas von Aquins Frage in seiner *Summa Theologica*, ob mehrere Engel an demselben Ort sein können,

findet in der mit ihr implizierte Verschränkung von Materialität und Transzendenz in Bezug auf Engel ihre direkte Parallele in den spekulativen Berechnungen, wie viele *Emulated Minds* pro Kubikzentimeter Masse im gesamten Universum existieren können. Engeln kommt in der römisch-katholischen Engellehre bis heute ein besonderer Status zu. Für Thomas von Aquin stehen sie näher bei Gott, weil sie ein tatsächliches Bild Gottes sind, wogegen der Mensch nur durch seinen Geist ein Abbild Gottes sei. Sie hingegen stimmten als reine Geistwesen in Essenz und Form immer überein. In ihnen ist somit alle Potentialität immer schon verwirklicht, womit sie dem sehr nahe kommen, was die *freedom of form* den digitalen Substraten verspricht.⁶⁰⁷ Den entscheidenden Hinweis, den der Vergleich mit den Engeln in Bezug auf die digitale Erhöhung des Menschen ergiebig macht, ist ihre Indifferenz gegenüber der Heilsgeschichte Gottes. Den Engeln ist von Gott nichts verheißen, für sie ist Christus nicht gekreuzigt worden. Gleichermäßen sagt Donna Haraway in ihrem *Cyborg Manifesto* über den Maschinenmenschen der Zukunft:

„The cyborg incarnation is outside of salvation history.“⁶⁰⁸

Indem der Mensch sein Menschsein überwindet, meint er, sich selbst aus der Sünden- wie Heilsgeschichte herauszunehmen, sodass für ihn weder Genesis noch Eschaton Referenzpunkte seiner neuen Geschichte bilden. Es ist die Idee, eine von den Mängeln des biologischen Menschen gereinigten Zukunft mit unendlichen Möglichkeiten und Entfaltungspotentialen einzuschlagen. Indem die digitalen Engel Himmel und Hölle hinter sich lassen, betreten sie den Limbus der Geschichtslosigkeit und Weltferne, aus der heraus sie meinen, nun das gesamte All aus den Fugen zu heben. Dass der von seiner Endlichkeit entgrenzte

607 Vgl. Tizina Suarez-Nani, „Individualität und Subjektivität der Engel im 13. Jahrhundert. Thomas von Aquin, Heinrich von Gent und Petrus Johannis Olivi“, *Das Mittelalter* 11 (2006), 29–48. Hier 34–36.

608 Haraway, *A Cyborg Manifesto*, 150. Dass der Wunsch misslingt, der omnipräsent empfundenen christlichen Heilsgeschichte in einer technizistischen Zukunft tatsächlich zu entfliehen, sondern die Erzählungen vom Cyborg sich zu großen Teilen christlicher Transzendenzhoffnungen bedient, zeigt sich paradigmatisch an der Verwendung des christologischen Begriffes der Inkarnation für diese erhoffte Befreiung.

Mensch entgegen der beschworenen Allmachtshoffnungen sein Tätigsein eher verlieren könnte – es zum Preis der Unsterblichkeit und des Fallens aus der Geschichte gar nicht aufrechterhalten kann – dieses Bild eines Engels zeigt Wim Wenders in seinem Film *Der Himmel über Berlin* von 1987. Die Engel, die hier gleichermaßen fern das Treiben der Gezeiten wie das Treiben der Menschen beobachten und aufzeichnen, erscheinen in ihrer Entrücktheit *in* der Welt selbst eher wie Naturkonstanten denn als tätige Personen. Dem Engel Damiel schließlich ist sein weltloses Schweben über den Dingen leid. Er beneidet die Menschen und entscheidet sich schließlich dafür, ein sterblicher Mensch zu werden, um am Leben teilzuhaben und innerhalb einer Geschichte zu stehen, der er als ein vergängliches Ich tatsächlich etwas hinzuzufügen vermag.

6 Abschließende Beobachtungen

Die Mythen um Künstliche Intelligenz beruhen auf der Gleichsetzung von Mensch und Maschine, von technischer Konstruktion mit biologischer Evolution, und schließlich der Reduktion aller Wirklichkeit in Information. Während Technik und Wissenschaft als Form der Kulturausübung grundsätzlich emanzipatorischen Ursprungs sind, bergen übersteigerte Machtphantasien und davon abgeleitete Ansprüche der Technik an den Menschen die Gefahr, in die Sterilität menschlichen Handelns umzuschlagen. Fassen lässt sich dieser Umschlag innerhalb der Konvergenzerzählungen in der Erwartung an die technologische Singularität und der sie begleitenden, apokalyptischen Rhetorik. Die Gleichsetzung wird hier mit der Ersetzung des Menschen durch die Maschine zu Ende gedacht und zu einer Zäsur von kosmischer Bedeutung erhoben. Die Verklärung einer solchen Überwältigung des mangelhaften Menschen stellt bestehende Freiheitsverständnisse infrage und tauscht sie gegen Zerrbilder von Freiheit aus, die von Verantwortungslosigkeit, Apathie und Selbstflucht geprägt sind. Diese Arbeit hält der lähmenden Erwartung auf die Revolte von Technik die Perspektive ihrer Kontinuität und Lebensweltlichkeit entgegen: Es zeigt sich, dass der Mythos Maschine kein Gedanke ist, der sich erst seit der Beschäftigung mit KI aufdrängt. Im Gegensatz zur proklamierten *singularitas* scheint es sich bei der Idee von der neuen Schöpfung des Menschen und die Sorge um seine Übervorteilung um eines der persistenten Mythologeme der westlichen Geistesgeschichte überhaupt zu handeln. Sie bilden von Beginn an die Kehrseite überzogener Machtphantasien. Zugleich sind die Mythen im Deutungskampf um Wesen und Potential Künstlicher Intelligenz als politisches Instrument wahrzunehmen. Die Digitalwirtschaft hat eine unerhörte wirtschaftliche wie gesellschaftliche Macht abseits der politischen Legitimation angehäuft, deren Akteure von dem Eindruck profitieren, wenn ihre Technologien überlebensnotwendig erscheinen und gleichzeitig Menschen scheinbar keine Verantwortung mehr für technologische Entwicklungen tragen.

Die Erkenntnisse der Untersuchung werden zum Abschluss zu folgenden Überlegungen gebündelt: In ethischer Hinsicht scheint es uner-

lässlich, die Genugtuung über die Erniedrigung des Menschen noch einmal theoretisch zu fassen, die neben dem Horror die Bewertung der Mythen durchzieht. Zudem ist der strategische Wert der Mythen und der postulierten Ohnmacht für ihre Theoretiker als auch die Tech-Konzerne hervorzuheben. Schließlich stellt sich die Frage, ob Technikvisionen damit im Allgemeinen Ideologie sind, oder ob man dem Mythos in der technischen Kultur nicht anders begegnen kann. Wem überlässt eine Gesellschaft den Entwurf gelingender Visionen der Zukunft, die auf Bewältigung mithilfe von Technik und nicht Überwältigung durch Technik setzen?

6.1 Die Lust der Konvergenzerzählungen an der Erniedrigung des Menschen

Unter dem Eindruck der kybernetischen Bewegung kam Hannah Arendt Mitte des 20. Jahrhunderts zu dem Schluss, dass die Menschheit ihren *archimedischen Punkt*, von dem aus sie gedanklich die Welt begreift, so weit von sich selbst verlegt, dass diese imaginäre Perspektive sich schlussendlich gegen die Menschen selbst wendet.⁶⁰⁹ Denn der hypothetische Betrachter, der aus dem All auf die Erde schaut, wird keine menschliche Freiheit erblicken können; stattdessen mag es scheinen, dass der Mensch *automatisch* funktioniert, er als reaktives System in die ihn umgebenden Naturkonstanten eingeht.⁶¹⁰ Jene epistemologische Sehnsucht der Kybernetik nach einer von jeder Subjektivität gereinigten Weltanschauung weist auf die Anfänge der modernen Empirie zurück: Der Mensch, der die Natur verstehen und beherrschen will, muss laut Francis Bacon sich fortdauernd von seinen Falschannahmen und Trugbildern, seinen *Idolen*, befreien.⁶¹¹ In den Mythen um Künstliche Intelligenz wendet sich diese Sehnsucht schließlich gegen den Menschen selbst und schlägt in einen Post- und Transhumanismus um. Der Mensch verkommt zum letzten Idol, das nach kosmischer Perspektive überwunden gehört.

609 Vgl. Arendt, *Vita activa oder Vom tätigen Leben*, 334–342.

610 Vgl. ebd., 411.

611 Vgl. Bacon, *Neues Organon*, 259.

„Wenn der Mensch dazu diene, disparaten Geschichten Einheit zu verleihen, welche Geschichten sind es dann, bei denen der anthropologisch konstruierte Mensch nicht mehr weiterhilft?“⁶¹²

Indem der Mensch zum *automaton* und menschliche Geschichte zu einem Teil der Naturgeschichte erklärt werden, wird er gleichzeitig der Sinnlosigkeit und der ethischen Indifferenz gegenüber allem, sogar dem Selbst, preisgegeben. Wo keine Zwecke mehr für die sich erweiternden Mittel bereitstehen, derer sich der Mensch durch die Durchdringung der Natur bereichert, wird die sinnentleerte Wirklichkeit allein unter das verbleibende Diktat der Macht gestellt:

„Aber ist Naturbeherrschung das einzige und das höchste legitime Ziel des Menschen? Diese Auffassung hätte verhängnisvolle Konsequenzen. Sie unterstellt das menschliche Dasein selbst als Mittel dem Zweck seiner eigenen Erhaltung, und genau dies ist das Wesen des Nihilismus. Progressive Naturbeherrschung als oberstes Ziel kehrt sich dann gegen den Menschen selbst.“⁶¹³

Das Leben und sein Wert verflüssigt sich zum Lebensprozess und wird bloß unter dem Aspekt des ständigen Kampfes um das *Überleben* greifbar.⁶¹⁴ In der sinnentleerten Welt offenbart sich, so schreibt der deutsche Großmeister der Apokalyptik Oswald Spengler, die Weltgeschichte zum Weltgericht über den Menschen. In der Abfolge von Werden und Vergehen setze sich allein der Stärkere durch.⁶¹⁵ Den selbsternannten Schöpfen dieses Weltgerichts erscheint der vorfindliche Mensch, gemessen an dem Maßstab absoluter Macht, unhaltbar. Schon Friedrich Nietzsche (1844–1900) hat nur Ekel und Widerwillen für ihn übrig; der Mensch leide am Menschen, er sei ein „Heillos-Mittelmäßiger“⁶¹⁶.

612 Pias, *Zeit der Kybernetik – Eine Einstimmung*, 17.

613 Robert Spaemann/Reinhard Löw, *Natürliche Ziele. Geschichte und Wiederentdeckung teleologischen Denkens*, Stuttgart 2005, 235.

614 Vgl. Herrgen, ‚Leben‘ bei Darwin und im Darwinismus, 52.

615 Vgl. Spengler, *Der Untergang des Abendlandes*, 1194.

616 Friedrich Nietzsche, *Jenseits von Gut und Böse. Zur Genealogie der Moral* (Sämtliche Werke 5) 1988, 277.

Das, was Nietzsches Prophet der Zukunft Zarathustra am Menschen noch lieben kann, „das ist, dass er ein Übergang ist und ein Untergang.“⁶¹⁷ In dieser Tradition stehen auch die neuen Mythen der Künstlichen Intelligenz. Sie sind sich einig, dass den nicht veränderungswilligen Menschen des 21. Jahrhunderts nach dem richtenden Maßstab des Techno-Darwinismus nur eine Apokalypse ohne Reich erwartet. In ihrem extropischen Geschichtsbild scheint den Vertretern der Singularitätstheorie die Ablösung der Gattung Mensch durch höherwertige Informationsträger unausweichlich.

Die prognostizierte Überwältigung des Menschen durch die Maschine scheint in den Mythen ihr Licht bereits auf die Gegenwart. Denn die Gewissheit, dass die Menschheit als Gattung auf eine Zukunft zurast, die sie in ihrer *Antiquiertheit* nicht übersteht, verhindert gestaltbare Zukunft und lässt den Menschen ohnmächtig zurück. Die klassische, christliche Apokalyptik steht schließlich geradezu für das *Nicht-Bewältigbare*. Die Erwartung an das unumstößliche Eingreifen Gottes lässt Menschen keinen Handlungsspielraum mehr zu. Die entlastende Zusage, dass Gott richten wird – und damit eine besondere Form der Wirklichkeitsbewältigung im Mythos darstellt –, verkehrt sich zu belastender Hilflosigkeit, wenn der Mensch sich im Weltgericht der aufsteigenden Information sicher zu den Verdammten zählt. Das Urteil erscheint grausam, als habe sich der gesamte Kosmos gegen den Menschen verschworen. Menschliches Handeln ist in dieser Perspektive derart von äußeren Zwängen und Notwendigkeiten gesteuert und dementsprechend blind, dass demnach die Menschen ihre eigenen Nachfolger konstruieren müssten. Sie leben auf keiner utopischen Insel, auf der sie dem technologischen Fortschritt Einhalt gebieten könnten; weil vielmehr auch der Fortschritt als eherne Naturkonstante wahrgenommen wird, *muss* die Singularität eintreten. Schon Vernor Vinge, der *Entdecker* der technologischen Singularität, kommt nicht umhin anzumerken, dass jeder Widerstand zwecklos ist.⁶¹⁸ Das Phantasieren darüber, was den intelligenten Maschinen der Zukunft alles möglich sein wird,

617 Friedrich Nietzsche/Josef Simon, *Also sprach Zarathustra*. Ein Buch für Alle und Keinen, Stuttgart 1994, 300–301.

618 Vgl. Vinge, *The Coming of Technological Singularity: How to Survive the Post-Human Era*, 15.

beschwört im Umkehrschluss, dass dem Menschen nichts möglich ist und es von ihm auch nichts mehr zu erwarten gibt. Derjenige, der dieses in die Gegenwart gesprochene Urteil über das menschliche *fatum* glaubt, den führt diese Finalität über Verzweiflung hin zu Neid, zu Scham, schließlich weiter zu Selbsthass und einer sadistischen Genugtuung über die kommende Ablösung des Menschen:

„Der ohnmächtigste Neid ist zugleich der furchtbarste Neid. Der Neid, der deshalb die stärkste Ressentimentbildung auslöst, ist daher derjenige Neid, der sich auf das individuelle Wesen und Sein einer fremden Person richtet: der Existentialneid [die bloße Existenz des anderen wird zum Vorwurf].“⁶¹⁹

Anknüpfend an Max Schelers Gedanken über das Ressentiment entwickelte Günther Anders seine Diagnose von der *prometheischen Scham*, die jenem Existentialneid gegenüber der Maschine und ihrer morphologischen Freiheit entspringt. In der Scham ist der Mensch einerseits er selbst, und nimmt doch gleichzeitig die Position jenes fiktiven Außenbetrachters ein, der *sich selbst als ein anderer* verdammt. Anders sagt, die Scham sei so perplex und methodenlos, dass sie demjenigen, der sich schämt, keinen anderen Weg übriglässt, als defätistisch zu versuchen, sich zusammen mit seinem *malum* zugrunde gehen zu lassen – eben sich *in Grund und Boden zu schämen*.⁶²⁰ Den Konvergenzerzählungen unterläuft selbiger Zwiespalt, dass man mit dem Aufkommen der Künstlichen Intelligenz die eigene Auflösung fürchtet, und doch zugleich keine guten Gründe dagegen findet. Existenzielle Lebensangst und Lebensfrust schlagen innerhalb der Logik der Konvergenzerzählungen in die Gewalt gegen das eigene So-Sein und damit sadistische Selbstzerstörungslust um. Die Ersatzhandlung – weil man gegen die eigene Existenz nichts tun kann – ist die phantasierte Selbstsabotage und Selbstaufgabe, die die Mythen begleitet und deren Bewertung sie so ambivalent ausfallen lässt. Diese Selbstzerstörungslust lässt den prophe-

619 Max Scheler, *Das Ressentiment im Aufbau der Moralen* (Klostermann Texte Philosophie), Frankfurt am Main 1978, 11.

620 Vgl. Anders, *Die Antiquiertheit des Menschen*, 70.

zeiten Untergang der Menschheit zum ästhetischen Genuss verkehren, den Walter Benjamin zum Kennzeichen des Faschismus bestimmt hat:

„Das ist offenbar die Vollendung des *l'art pour l'art*. Die Menschheit, die einst bei Homer ein Schauobjekt für die Olympischen Götter war, ist es nun für sich selbst geworden. Ihre Selbstentfremdung hat jenen Grad erreicht, der sie ihre eigene Vernichtung als ästhetischen Genuß ersten Ranges erleben läßt.“⁶²¹

Der faschistische Größenwahn des *Alles oder Nichts* kehrt sich im Moment der Einsicht, dass der Mensch den angeblichen Forderungen seiner Geschichte nicht genügt, in Phantasien der Selbstbestrafung um. Wenn der Mensch aufgrund seiner Ausstattung nicht in der Lage ist, die greifbar nahen *All-machtsphantasien* der Wissenschaften zu erfüllen und die Rolle des interstellaren *Alpha-Prädators* zu übernehmen, dann soll er mit Fug und Recht untergehen und Platz machen für bessere Informationsträger, die die Unterwerfung aller Materie und den Kampf um neuen *Lebensraum* in den Weltraum hinaustragen können.⁶²² In der Mythologisierung des Weltgerichts zu dem punktuellen Ereignis der technologischen Singularität schließlich wird die schon gegenwärtig feststehende Zukunftslosigkeit des Menschen im Kampf ums Dasein offenbar gemacht und das Urteil unmittelbar in seiner Vernichtung, Einhegung oder Verwandlung vollstreckt. In der Antizipation des sicheren Untergangs mischt sich Ohnmacht mit der dionysischen Lust einer vermeintlichen *satisfactio* in der unmittelbar bevorstehend geglaubten, eigenen Vernichtung:

621 Walter Benjamin, *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. Drei Studien zur Kunstsoziologie, Berlin, Frankfurt am Main 2018, 77.

622 Die Parallele von den Mythen um KI zum Faschismus zieht ebenfalls der Religionswissenschaftler Oliver Krüger: „Die kulturellen Deutungsmuster des technologischen Posthumanismus jedoch, die in der Faszination für die Superintelligenz zum Ausdruck kommen, sind eine direkte Fortsetzung der Vorstellungen von Genie, Gehirn und Rasse in Westeuropa und den Vereinigten Staaten des vergangenen Jahrhunderts. [...] Die Skala dieser Seinsordnung – Gehirngrößen bzw. mathematisierte Denkleistungen – ist unverändert geblieben.“ In: Oliver Krüger, *Virtualität und Unsterblichkeit*. Gott, Evolution und die Singularität im Post- und Transhumanismus (Rombach Wissenschaften. Reihe Litte-rae), 2. Aufl., Freiburg i. Br./Berlin/Wien 2019, 375–376.

„Dahinzuwelken kann trostloser sein als von den herabstürzenden Sternen erschlagen zu werden.“⁶²³

Nur: Der archimedische Punkt außerhalb des eigenen Selbst, die wortwörtliche *All*-perspektive, welche die menschlichen Idole überwunden hat, sie bleibt Fiktion. Sie selbst ist der nicht durchschaute und deshalb verabsolutierte *Mythos* im Gewand des *Logos*. In den Worten Günther Anders tritt die gleichzeitige *angemaßte Selbsterniedrigung* und *hybride Demut* in den Konvergenzerzählungen der Künstlichen Intelligenz darin zutage, dass den Anhängern der technologischen Singularität diese fiktive Position einer menschenenthobenen Perspektive zu der absoluten Perspektive eines sehenden Gottes der Geschichte gerät, und sie sich zu dessen Evangelisten erheben:⁶²⁴

„Einen im Jahre 2000 geschriebenen theologischen Bericht über den Vorgang könnte man sich wohl etwa so vorstellen: ‚Da es den Dämon oder den marcionistischen Gott, der den Menschen zum Gerät-Dasein verurteilte, oder der ihn in ein Gerät verwandelte, nicht gab, erfand der Mensch diesen Gott; ja er maßte es sich sogar an, die Rolle dieses zusätzlichen Gottes selbst zu spielen; aber er spielte die Rolle ausschließlich zu dem Zwecke, um sich denjenigen Schaden zuzufügen, den er von anderen Göttern nicht beziehen konnte. – Nur um sich auf neue Art zum Knecht machen zu können, machte er sich zum Herren.“⁶²⁵

Was die Konvergenzerzählungen um Künstliche Intelligenz im mythologischen Gegenüber von Mensch und Maschine durchspielen, ist schließlich die moderne Theodizee des Menschen. Nicht mehr Gott, sondern der Mensch wird vor den Gerichtshof des Möglichen gestellt. Wird Gott durch seine Nichtexistenz entschuldigt, so findet jedoch der

623 Blumenberg, Arbeit am Mythos, 61.

624 Vgl. Armin Grunwald, „Technische Zukunft des Menschen? Eschatologische Erzählungen zur Digitalisierung und ihre Kritik“, in: Mitscherlich-Schönherr/Reiner Anselm/Martin Heinze (Hgg.), Das Gelingen der künstlichen Natürlichkeit. Mensch-Sein an den Grenzen des Lebens mit disruptiven Biotechnologien, Berlin/Boston 2021, 314–331. Hier 320.

625 Anders, Die Antiquiertheit des Menschen, 48.

Mensch keine Entschuldigung mehr für sein Ungenügen. Nietzsches Übermensch – nun in Gestalt der Künstlichen Intelligenz – muss kommen, weil der Mensch für seine Unfähigkeit nicht mehr von Gott entschuldigt werden kann.⁶²⁶ Der Mensch erhöht sich zu der Position eines allwissenden Gottes, nur um sich selbst zu erniedrigen. Dazu setzt er die Maschine als Token für das fehlende Gegenüber ein, das gleich drei unterschiedliche Rollen einzunehmen hat, nämlich einerseits den Menschen um seines Ungenügens am Maßstab des Möglichen zu bestrafen, ihn vor sich selbst zu schützen und selbst die Rolle des neuen, erlösten Menschen einzunehmen.

Eine solche Sakralisierung überlastet die Technik, sie kann die Funktion des Heiligen nicht ersetzen. Der ideologische Blick verkennt, dass Technik ihrem Sinn nach nie mehr sein kann als eine Verlängerung menschlichen Willens. In der Technik begegnet der Mensch stets nur sich selbst, und dass gerade in dem Versuch, sie immer mehr von jedem Anthropomorphismus zu befreien.⁶²⁷ Vielmehr kann die inhärente Freiheit der Technik missbraucht werden, um sie gegen den Menschen selbst zu wenden. Natürlich kann in die Maschinen der Selbsthass externalisiert werden, um über die Menschheit zu richten. Ironischerweise können Menschen der Technik selbst diesen Dienst auferlegen. Menschen können, wenn sie das beabsichtigen, Maschinen bauen, die sie kontrollieren, maßregeln, bei Entscheidungen übergehen und ihnen Leid zufügen. Maschinelles Lernen kann eingesetzt werden, um neue Toxine zu synthetisieren oder um Kampfdrohnen auf den Beschuss von Menschen zu trainieren. Nur sollte man dabei nicht verwechseln, dass es die Menschen sind, die Maschinen diese Ziele eingeben müssen und sich hier kein kosmischer Wille emaniert, der Vergeltung an einer existentiellen Schuld der Menschheit übt. Auch hier gilt der Werkzeugcharakter und die ethische Indifferenz der Technik, von der bereits Tillich

626 Vgl. Odo Marquard, „Entlastungen. Theodizeemotive in der neuzeitlichen Philosophie“, in: *Zukunft braucht Herkunft*. Philosophische Essays, Stuttgart 2015, 124–145. Hier 131–132.

627 Vgl. Hannah Arendt, „The Conquest of Space and the Stature of Man“, *The New Atlantis. A Journal of Technology and Society* (2007), 43–55. Hier 52. Sie argumentiert in Auseinandersetzung mit dem Denken Werner Heisenbergs über das Verhältnis von Mensch und Natur in: Werner Heisenberg, *Das Naturbild der heutigen Physik* (Rowohlts Deutsche Enzyklopädie 8), 1. Aufl., Hamburg 1955, 17–18.

sprach, als er den Technikbegriff für die Theologie restituierte. Vielmehr bedarf es der Kritik um die Verfasstheit der technischen Kultur an sich und der Frage danach, was sie daran hindert, Ausdruck einer humanen Kultur zu sein.

6.2 Die Propagierung einer alternativlosen Zukunft der Maschinen: Cui bono?

„Der Mythos von der technischen, politischen und gesellschaftlichen Zwangsläufigkeit ist ein wirksames Beruhigungsmittel für das Bewusstsein. Seine Funktion besteht darin, die Verantwortung jedem von den Schultern zu nehmen, der an ihn glaubt. Aber in Wirklichkeit gibt es handelnde Personen!“⁶²⁸

Die Konvergenzerzählungen setzen voraus, dass kein politisches Handeln etwas an dem Eintreten der technologischen Singularität ändern wird; Technik und Wissenschaft werden zu Erscheinungsformen von Urenergien naturalisiert und die Technosphäre so von allem entkoppelt. Die Mythen geben sich blind dafür, dass Technologie selbst ein Kulturausdruck ist. Aufgrund dieser eigenen Undurchsichtigkeit bleiben die Zukunftsvisionen „Prophezeiungen ohne prognostischen Wert“⁶²⁹, weil sich in ihnen nur Kosmoseroberer und digitale Engel, aber keine gegenwärtigen Menschen wiederfinden können.

Trotz ihrer Weltlosigkeit sind die KI-Mythen keine folgenfreie Theorie, sondern verschieben als soziotechnische Sinnbilder Machtgefüge und allokalieren Ressourcen. Für die Visionäre selbst sind ihre Prognosen existenzsichernd. Die über die Werbemaßnahmen eingeworbenen Mittel der IT-Industrie helfen, eigene Universitäten zu gründen und sich durch eigens finanzierte Institute den Eliteuniversitäten anzunä-

⁶²⁸ Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, 317.

⁶²⁹ Christoph Kehl, „Möglichleiten und Grenzen ethischer Technikgestaltung. Das Beispiel der Mensch-Maschine-Entgrenzung“, in: Mitscherlich-Schönherr/Reiner Anselm/Martin Heinze (Hgg.), *Das Gelingen der künstlichen Natürlichkeit. Mensch-Sein an den Grenzen des Lebens mit disruptiven Biotechnologien*, Berlin/Boston 2021, 151–170. Hier 159.

hern.⁶³⁰ Die lukrative Seite dieser Vermarktung der Zukunft zeigt sich exemplarisch an der *Singularity University*, die über ihre weltweiten Veranstaltungen eine Plattform für die Vermarktung künftiger Technologien geschaffen hat.⁶³¹ Die Faszination der Tech-Unternehmen wiederum für die Konvergenzerzählungen gründet auf den scheinbaren Vorteilen, die die Bedeutungsaufladung ihres Handelns durch sie erfährt. Denn die Währung des Mythos liegt im Gewähren von Bedeutsamkeit, und unterscheidet sich damit von der Gleichgültigkeit des wissenschaftlichen Blickes.⁶³² Der Glaube an eine allen Lebensbereichen enthobenen Technosphäre und ihre Zwangsläufigkeit kommt den Konzernen in ihrem Bestreben entgegen, ihre merkantilen Visionen der Nutzung von Informationstechnologie gegenüber Forderungen nach gesellschaftlicher Teilhabe und staatlicher Regulierung zu schützen. Die Überbetonung maschineller Autonomie nivelliert zugleich das Problem, dass aus Wissenslücken Verantwortungslücken entstehen können, die in letzter Konsequenz zu Lasten der Nutzer digitaler Dienstleistungen gehen. Der Mythos fängt so die Unzulänglichkeiten ab und verdunkelt die Differenz zwischen dem in Aussicht Gestellten und dem gegenwärtig Realisierbaren. Darin bekleiden die Mythen zugleich die Sorge einer kleinen Tech-Elite, sich an ihren eigenen, zu hohen Ansprüchen messen zu lassen.⁶³³

Dadurch, dass die Visionäre oft selbst in Forschung und Entwicklung tätig sind, kommt es zu der Schizophrenie, dass dieselben vor den großen Gefahren warnen, die gleichzeitig die Entwicklung der Informationstechnologien massiv vorantreiben. Ganascia erklärt die *Singularitans* und ihre Sympathisanten aus diesem Grund zu *pompieri pyromanes*, zu pyromanischen Feuerwehrleuten; sie legten selbst jenes Feuer, welches sie im selben Moment zu löschen beabsichtigten, um so andere außen vor zu halten und eigene Interessen zu schützen. Ganascia hält es aus diesem Grund für strategisches Kalkül, wenn Ingenieure und Wissenschaftler, deren Welt bisher vom Ideal der Aufklärung dominiert schien, nun die vermeintliche Fatalität allen (politischen) Handelns

630 Vgl. Spiekermann, *Digitale Ethik*, 167–169.

631 Vgl. Krüger, *Virtualität und Unsterblichkeit*, 399.

632 Vgl. Blumenberg, *Arbeit am Mythos*, 77–80.

633 Vgl. Eurich, *Mythos und Maschine*, 32.

propagieren.⁶³⁴ Lässt sich aus dieser Unaufrichtigkeit der Schluss ziehen, dass es sich bei diesen Mythen um eine Verschwörung der Tech-Elite handelt? Viel wahrscheinlicher ist es, dass der Großteil der Vertreter der Singularitätsthese selbst von der Vorstellung eines stattfindenden, unerbittlichen Techno-Darwinismus überzeugt sind und sich in ihrem Wirken tatsächlich am Wendepunkt der Zeit begreifen. Diese Selbsttäuschung braucht in keinem Widerspruch dazu zu stehen, dass dieselben aus der Propagierung dieser Visionen ihre Vorteile ziehen.

Wie sehr die Mythen entgegen der Behauptung, einer menschenent hobenen Erkenntnis zu entspringen, ihrem Trägerkreis und deren Interessen verpflichtet bleiben, äußert sich darin, dass die Szenarien der technologischen Singularität von einer unhinterfragten Marktlogik durchsetzt sind. Es ist entlarvend, wie die Erzählung von der großen Disruption für alles bis auf das neoliberale Wirtschaften gilt. Das vorgeblich *aufgefundene* Prinzip aufsteigender Ordnung von Information stellt sich für die *Mind Children* oder die hochgeladenen Substrate stets auch als das wirtschaftliche Sinnvolle heraus. Die unterstellten Instinkte der hyperintelligenten Maschine, des kosmischen Eroberers, sind stets die eines übersteigerten *homo oeconomicus*. Mehr noch: Das Votum der schnelllebigen Digitalwirtschaft eines *Change or Perish* wird in der Post-Singularität selbst zur zentralen Wahrheit erhoben. Nirgendwo zeigt sich dies deutlicher als in dem Zustand der *Emulated Minds*, deren neues digitales Zuhause eine gesteigerte Darwin-Welt des gegenseitigen Fressens und Gefressen Werdens im Nanosekundenbereich bedeutet. Die Szenarien der Post-Singularität stimmen allesamt darin überein, dass sie den *Grenzen des Wachstums*, die der *Club of Rome* 1972 postulierte, die Prophetie von einem unendlichen, exponentiellen Wachstum entgegenstellen.⁶³⁵ Diese Heilsprophetie des unbegrenzten Wachstums ist jedoch nur zum Preis der Auflösung des Menschen zu haben; im Umkehrschluss zielen die Mythen auf eine Befreiung und Entfesselung des digitalen Kapitalismus von den Beschränkungen biologischer, endlicher Menschen.

634 Vgl. Ganascia, *Le Mythe de la Singularité*, 77–80.

635 Vgl. Krüger, *Virtualität und Unsterblichkeit*, 400.

„For now, it is clear that profit, and not cosmic evolution, is the driving spirit of planetization – its major metaphor, its omnipotent and universal truth.“⁶³⁶

6.3 Der bleibende Wert des Mythos für die technische Kultur

Was für Schlüsse lassen sich für den Umgang mit den Mythen der Künstlichen Intelligenz ziehen? Ganascia kommt in seiner Untersuchung zu dem Ergebnis, man müsse die Wissenschaft um Künstliche Intelligenz vom Mythos befreien. Er geht davon aus, dass es seit den 1960er Jahren eine heimliche Verselbstständigung der immer populäreren Mythen um Künstliche Intelligenz gegeben habe, wobei die Mythos-tradition sich seither als Wissenschaft ausbebe. Er greift dazu Oswald Spenglers Verwendung des Begriffes *Pseudomorphose* auf. Damit will er zum Ausdruck bringen, dass – auch wenn von außen betrachtet die wissenschaftlichen Namen und Begriffe erhalten blieben – innerhalb der Disziplin der Mythos die Wissenschaft in dem Maße unterwandert habe, dass die Rationalität der Irrationalität gewichen sei, das kritische Prüfen zugunsten von neuen Gewissheiten anerkannter Autoritäten aufgegeben wurde und so sich der Humanismus in einen technologischen Posthumanismus verwandelt habe.⁶³⁷ Nach Ganascia ist der Singularitätsmythos und alles, was auf ihm aufbaut, in dem Sinne *Fake-Science*, von dem die Forschung und der wissenschaftliche wie öffentliche Auseinandersetzung gereinigt gehört.

Ihm ist beizupflichten, dass die zu beobachtende Verkürzung der öffentlichen Debatte auf Künstliche Superintelligenzen und digitale Substrate deshalb wenig hilfreich ist, weil sie den Blick von den drängenden Fragen ablenkt, die mit der weiteren Verbreitung autonomer Systeme relevant werden.⁶³⁸ Nur ist das Ideal einer vollständigen Säuberung der Wissenschaft vom Mythos selbst ideologisch. Vielmehr scheint doch der Mythos innerhalb der Wissenschaft und Technik aus

⁶³⁶ Davis, *Techgnosis*, 312.

⁶³⁷ Vgl. Ganascia, *Le Mythe de la Singularité*, 125–126.

⁶³⁸ Vgl. Christoph Kehl, Christopher Coenen, *Technologien und Visionen der Mensch-Maschine-Entgrenzung*, 17.

demselben, notwendigen Überschuss an Freiheit zu entstehen, der für die sie selbst unerlässlich ist. Technologie und Technik tragen ihre transzendente Signatur schon immer mit sich; technische Innovation ist stets das Produkt von Kreativität und damit Ausdruck von Freiheit. Zur Wissenschaft und Technik gehört somit jenseits aller Fehlprognosen der Futurologen und Science-Fiction-Autoren die Fantasie der Umsetzbarkeit, der strategischen Planbarkeit und Konstruierbarkeit.⁶³⁹

Man kann die Konvergenzerzählungen und ihre Genugtuung, den archimedischen Punkt gegen den Menschen zu wenden, nicht einfach entmythologisieren und meinen, sie dadurch ethisch zu lösen. Stattdessen ist die über den kleinen Kreis der wirklichen *Singularitans* hinausgehende Popularität und Resonanz dieser Zukunftsmythen und das ihr zugrundeliegende Narrativ der Überwältigung und Passivierung des Menschen durch Technik ernst zu nehmen. Die Resonanz auf diese Erzählungen mag gerade deshalb so groß ausfallen, weil sie eine schon gegenwärtig empfundene Ohnmacht im Umgang mit den sich aufdrängenden Informationstechnologien trifft. Die postsinguläre Anthropologie spiegelt möglicherweise das gegenwärtige Misstrauen gegenüber einer fortschreitenden Technisierung der Umwelt wider, die die Lebenswelt eher droht zu verarmen, denn zu bereichern. Dieser Eindruck verstärkt sich darin, dass die politische Einhegung dieser supranationalen Tech-Unternehmen bisher nicht gelungen ist.

Damit stellt sich die Frage, wie mit der Mythosbedürftigkeit der technischen Kultur ethisch reflektiert umzugehen ist. Die Vereinseitigung der Zukunft in einer teleologischen Geschichtsphilosophie kann zumindest nur dort unwidersprochen bleiben, wo sie vollends als alternativlose Wirklichkeit akzeptiert wird. (Technische) Kultur hat jedoch niemals den Charakter des Selbstverständlichen. Die KI-Visionen scheinen auch deshalb so erfolgreich, weil sie die wenigen öffentlichkeitswirksamen Versuche sind, einen in die Zukunft tragenden Sinnzusammenhang von modernen Informationstechnologien und (Post-)Mensch herzustellen. Die liberale Idee müsste deshalb sein, durch kon-

639 Vgl. Macho, Technische Utopien und Katastrophenängste, 13.

kurrierende Entwürfe diese Macht über die Zukunft aufzuteilen und zu begrenzen.⁶⁴⁰

Das Entwerfen der Zukunft und damit der Glaube an die kreative Kraft der Utopie hat seit den Totalitarismen des 20. Jahrhunderts einen schweren Stand in der politischen Philosophie. Karl Poppers (1902–1994) Auseinandersetzung mit dem Faschismus des 20. Jahrhunderts und seine daraus hervorgehende Warnung vor Utopismus und Historizismus war die letzten sieben Jahren prägend. Mit seinem Ideal einer offenen Gesellschaft distanzierte sich Karl Popper deshalb bewusst von jedwedem Vorausgreifen der Zukunft.⁶⁴¹ Die proklamierte Visionslosigkeit hat seither nicht zuletzt in Bezug auf die technische Kultur ihre Probleme klar aufgezeigt. Denn eine radikale Offenheit der Zukunft im Zeitalter der Hochtechnologien schließt die Möglichkeit der Katastrophe immer schon mit ein. Hans Jonas hat die Gefahr erkannt, dass die selbstgewählte Zukunftsblindheit auch in die Verantwortungslosigkeit für zukünftige Generationen umschlagen kann. Darauf reagierte er mit seiner Heuristik der Furcht. Mit ihr lässt er den Blick in die Zukunft jedoch allein unter der Maßgabe der technologisch-induzierten Katastrophe zu, um aus der Furcht vor der drohenden Katastrophe und der Verzerrung des Menschen dieser politisch entgegenzusteuern.⁶⁴² Wie Trutz Rendtorff anmerkte, führt Jonas Prinzip des *in dubio pro malo* dazu, in Folgen der Ungewissheit grundsätzlich nichts zu tun.⁶⁴³ Nur zu wissen, was man nicht will, kann Menschen nicht aktivieren. Die Ermöglichungsbedingung für Politik, ihr Gestaltungswille, ihr gemeinsam Verbindendes und Verpflichtendes bedarf neben der beständigen ideologischen Kritik einer schemenhaften Vision gelingenden Zusammenlebens, die in Anbetracht des Stellenwerts der Informationstechnologien auch diese sinnhaft integriert. Es geht nicht darum, in jenen Fanatismus abzugleiten, vor dem Popper und Jonas gleichermaßen warnen, in welchem das Heute für (Alb-)Traumbilder eines unerreichbaren

640 Vgl. Sarah Spiekermann, „Der Mensch: ein Fehler“ (23.03.2019).

641 Vgl. Karl R. Popper, *Die offene Gesellschaft und ihre Feinde 1. Der Zauber Platons* (Uni-Taschenbücher 472), 8. Aufl., München 2003, XIX.

642 Vgl. Jonas, *Das Prinzip Verantwortung*, 8.

643 Vgl. Rendtorff, *Vielspältiges*, 143.

Übermorgens geopfert wird. Das „Auch-anderssein-Können“⁶⁴⁴ und damit die Kontingenz der Geschichte soll nicht verneint werden. Die Pluralisierung an Zukunftsvisionen soll Verantwortung nicht schmälern, sondern steigern. Denn der Vorteil dieser Fiktionen ist nicht, dass sie die Zukunft vorhersagen, die möglichst mit der Realität übereinstimmt, sondern dass sie eine transparente Perspektive anbieten, die man teilen kann, auf deren Basis man planen, diskutieren und prinzipiell Konsens herstellen kann. Die Pluralität an vorstellbarer Zukunft erweitert den Spielraum, auf deren Grundlage überhaupt diskutiert wird.⁶⁴⁵

Francis Bacon hat in seiner Utopie *Nova Atlantis* gewusst, was für Gefahren er seiner idealen Gesellschaft aufbürdet, wenn alle ihre Impulse allein von einer Technokaste ausgehen. Aus diesem Unbehagen heraus hat Bacon seine fiktive Bruderschaft des Haus Salomonis unter eine strenge Mönchsregel gestellt, die es den Brüdern verbietet, ihre Apparate einzusetzen, um Uneingeweihte hinter Licht zu führen.⁶⁴⁶ Die Geschichte jedoch hat gezeigt, dass in nicht-idealen Gesellschaften eine solche Selbstbindung, wie sie auch gegenwärtig Unternehmen wie Google oder Facebook über eigene Ethikrichtlinien sich auferlegen, ihren Zweck nicht erfüllt. Viel eher scheint das Votum des englischen Historikers Lord Acton zu gelten, nach dem absolute Macht auch absolut korrumpiert.⁶⁴⁷ Man sollte das Entwerfen der Zukunft daher nicht allein jenen überlassen, die auf ihre Erzählung absolut verpflichtet und dabei ihr Partikularinteresse als allgemeines Interesse missverstehen.

„Der Gewinn des Fortschritts ist ein Gewinn an Pluralität der Kultur, und die Signatur der Freiheit ist nicht Eindeutigkeit, der Preis der Freiheit ist eher der Mangel an Eindeutigkeit, oder positiv ausgedrückt: eine Erweiterung des Lebensraumes der Kultur.“⁶⁴⁸

⁶⁴⁴ Jonas, *Das Prinzip Verantwortung*, 339–341.

⁶⁴⁵ Vgl. Elena Esposito, *Die Fiktion der wahrscheinlichen Realität* (Edition Suhrkamp 2485), 3. Aufl., Frankfurt am Main 2014, 61.

⁶⁴⁶ Vgl. Bacon, *Neu-Atlantis*, 213.

⁶⁴⁷ „*Power tends to corrupt and absolute power corrupts absolutely.*“ Vgl. John Emerich Edward Dalber-Acton, „Appendix – Letter to Bishop Creighton. (5 April 1887)“, in: John Neville Figgis/Reginald Vere Laurence (Hgg.), *Historical Essays and Studies*, London 1907, 503–507. Hier 504.

⁶⁴⁸ Rendtorff, *Strukturen und Aufgaben technischer Kultur*, 14.

Die Frage, was für eine Kultur förderlich ist, kann nicht aus Wissenschaft und Technik allein beantwortet werden. Eine tatsächliche Pluralisierung von Entwürfen kann allein in einer Demokratisierung der Deutungshoheit innerhalb der technischen Kultur bestehen. Bei einer Technologie wie dem maschinellen Lernen, die das Potential in sich birgt, ähnlich wie die Elektrizität sämtliche Lebensbereiche zu durchziehen, muss die Art und das Ausmaß ihrer Nutzung von möglichst verschiedenen Blickwinkeln aus betrachtet werden. Dazu bedarf es der Aufbereitung des notwendigen Grundlagenwissens, um die disziplinen-eigene Hermetik aufzubrechen. Öffentlichkeitsarbeit muss hier einen anderen Sinn haben als Werbung und Interessenvertretung.⁶⁴⁹ Damit steht gleichsam die Forderung, dass die vielfältigen Institutionen der offenen Gesellschaft wie Parteien, Kunst, Universitäten, Gewerkschaften und Kirchen Formate finden, in denen sich Menschen anders als bloß passive Konsumenten fertiger digitaler *Services* verstehen lernen und die Aura des Fatalen aufgebrochen wird, um vielfältige Stellungnahmen zu ermöglichen. Bei allem Idealismus ist zugleich festzuhalten, dass es bisher noch eine weitgehend offene Frage bleibt, wie eine *antizipatorische Governance* mit möglichst unterschiedlichen Interessenvertretern konkret organisiert werden könnte. Möglicherweise ist die gesellschaftliche Auseinandersetzung sinnvollerweise mit einem differenzierten Blick auf die jeweiligen Anwendungsfelder und ihre moralischen Implikationen leichter zu führen, aus denen sich allein implizit bzw. erst im Abgleich ethische Muster und größere Sinnbilder erschließen.⁶⁵⁰

Die Freiheit des Menschen in der Geschichte aufrecht zu erhalten und somit die Warnungen Poppers und Jonas ernst zu nehmen – und gleichzeitig doch Visionen eines guten Lebens in Antizipation technologischer Innovationen zu ermöglichen – bedarf neben einer Demokratisierung der technischen Kultur weiterhin der hermeneutischen Vorsicht. Blumenberg beispielsweise warnt vor der Ärmlichkeit der Phantasie. Er sagt, die literarische Gattung der Utopie beweise, wie schwierig es für Menschen ist, aus dem Wissen, dass die Wirklichkeit auch anders sein könnte, tatsächlich konkrete, positive Alternativen zu

649 Vgl. ebd., 21.

650 Vgl. Kehl, Möglichkeiten und Grenzen ethischer Technikgestaltung, 158.

denken. Oftmals werde der Negation des Bestehenden derart begegnet, dass die Phantasie sich ihr Gelände in die Zukunft wieder nur anhand bekannter Muster baue.⁶⁵¹ Wie an den Konvergenzerzählungen der Künstlichen Intelligenz auffällt, zieht auch die rein technizistische Zukunft ihre religiösen Schatten nach sich. Das Christentum bleibt ein Fundus, bei dem sich trotz oder gerade wegen der fortschreitenden Säkularisierung bedient wird:

„Perhaps the West has written itself into a narrative trap and cannot escape its old grandiose fairy tale of fulfillment and annihilation, a story that, like all good stories, both demands and staves off its own end.“⁶⁵²

Das Geschichtsbewusstsein der Konvergenzerzählungen, in denen die Gewissheit um die Selbstläufigkeit des technologischen Fortschritts apokalyptisch wird, hält an den christlichen Voraussetzungen und Konsequenzen fest, die Vergangenheit stets als Vorbereitung für die Erfüllung in der Zukunft zu verstehen.⁶⁵³ Gegenüber einer solchen Sakralisierung wie Dämonisierung des technologischen Fortschritts liegt die besondere Aufgabe der Theologie darin, Absolutionstendenzen und Handlungszwängen innerhalb der technischen Kultur zu wehren, die den Ort der Lebensführung schon jetzt in einen Un-Ort verlegen wollen. Quasi-religiöse Übersteigerungen der Erwartungen an Digitalisierung sind offenzulegen und auf das Maß menschlicher Entscheidungen zurückzuführen.⁶⁵⁴ Die Gegenwart ist keine Zeit der apokalyptischen Bedrängnis, der nichts bliebe, außer auf die Künstliche Superintelligenz zu warten. Ernst Troeltsch wies bereits darauf hin, dass allein in der Entdringlichung aller Naherwartungen das eigene Handeln wieder in das Licht der Dauer treten könne und sich der Mensch erst so den positiven Aufgaben der Heiligung und Bewältigung in der Welt zu stellen vermag.⁶⁵⁵

651 Vgl. Blumenberg, *Arbeit am Mythos*, 179.

652 Davis, *Technosis*, 255.

653 Vgl. Karl Löwith, *Weltgeschichte und Heilsgeschehen. Die theologischen Voraussetzungen der Geschichtsphilosophie* (Urban-Taschenbücher 2), 8. Aufl., Stuttgart 1990, 170.

654 Vgl. Charbonnier, *Digitalisierung: Theologische Selbstklärungen und Gegenwartsinterpretationen. Eine Skizze*, 249. Ebenso: Vgl. *Evangelische Kirche in Deutschland, Freiheit digital. Die Zehn Gebote in Zeiten des digitalen Wandels.*, 50.

655 Troeltsch, *Grundprobleme der Ethik*, 154.

„Die kritische Relevanz der Ethik für die Zukunftserwartung liegt darin, dass die ethisch reflektierte Zukunftserwartung der Tendenz widersteht, die gegebene Welt um des künftigen Guten willen zu negieren, und statt dessen die Zukünftigkeit des Guten im Vertrauen der Hoffnung als Moment der Gegenwart des gegebenen und aufgegebenen Lebens selbst bestimmt.“⁶⁵⁶

Neben der Kritik bedarf die Kultur als gemeinsamer Bewältigungsraum des Lebens zugleich einer Vorstellung des gemeinsam geteilten Guten, auf die sich Fortschritt als ein Annähern hinbewegen kann.⁶⁵⁷ Der christliche Glaube findet diesen Fixpunkt, der das eigene Handeln zur Hoffnung hin ausrichtet und in den Dienst des Guten stellt, in der Verheißung vom Reich Gottes.⁶⁵⁸ Die christliche Hoffnung, die die christliche Freiheit aus sich heraussetzt, geht über die Unzulänglichkeiten der Wirklichkeit hinaus und stellt das Unvollständige nicht als antiquiert und lebensunwürdig dar, sondern sieht es in Gott aufgehoben und in dem Sinne gerechtfertigt, dass der Mensch sich und seine Welt nicht am Maßstab des Absoluten zu messen hat. Das Vertrauen auf dieses gemeinsame Gute lässt menschliche Weltgestaltung nicht vergebens werden, sondern nimmt den Menschen in die Pflicht und betont die nicht abzutretende Verantwortung in der je eigenen Lebensführung für die gemeinsame Lebenswelt. Die Ermutigung zum Guten geht darüber hinaus, dass trotz allem Unverständnis, Streit und Gewalt menschliches Miteinander und echte Gemeinschaft möglich ist, weil Menschen mehr sind als die Summe ihrer Vergangenheit. Schließlich reicht die Hoffnung sogar über das eigene, partikuläre Leben und seine Geschichte hinaus, wodurch die eigene Endlichkeit nicht ihren Schatten bereits auf das Leben wirft, sondern die eigene Endlichkeit sich bereits im Leben in der Ewigkeit Gottes heil wissen darf.

„Es liegt auf der Hand, dass diese beiden Zielbestimmungen – die Betonung der Vorläufigkeit und Partikularität menschlich-geschichtlicher

⁶⁵⁶ Rendtorff, *Ethik*, 128–129.

⁶⁵⁷ Vgl. Anselm, *Warum Wissenschaft Religion braucht*, 369.

⁶⁵⁸ Vgl. Pannenberg, *Grundlagen der Ethik*, 71.

Ordnungen sowie die gemeinwohlverpflichtete Formulierung von Vorstellungen des gemeinsam geteilten Guten als eines Korridors für die politische Auseinandersetzung – in einer spannungsvollen wechselseitigen Ergänzung und Korrektur stehen.⁶⁵⁹

Ein besonnener Umgang mit dem Mythos innerhalb der technischen Kultur, wie er in diesem Buch unter dem Vorzeichen christlicher Selbst- und Weltdeutung entfaltet wurde, entlastet die Technik und befreit sie zu ihrem Technisch-Sein. Freiheit drückt sich in der Technik dann gerade in ihrem instrumentellen Charakter aus; an ihr hat sich kein Richtspruch von Heil oder Unheil über den Menschen zu vollziehen. Stattdessen orientiert sich ihre Funktionalität an individuellen Bedürfnissen wie gemeinsam getragenen Vorstellungen des Guten von Menschen in dieser Welt. Die neuen Informationstechnologien um Künstliche Intelligenz sind ein mächtiges Werkzeug, dessen Nutzen in positivem Sinne entfesselt wird, wenn man es aus dem festgefahrenen Vergleich mit dem Menschen löst. Denn es ist gerade ihre andere Form der Intelligenz, mit der maschinelles Lernen für den Menschen einen tatsächlichen Mehrwert darstellt. Sie bietet noch unbedachte Möglichkeiten, den Menschen in seiner individuellen Lebensführung als auch Gesellschaften insgesamt zu unterstützen, zu entlasten und zu inspirieren. Anstatt die Maschine zu humanisieren und den Menschen zu mechanisieren, kann sich in der besonnenen Implementierung von KI die technische Kultur entgegen aller posthumanistischen Träume als Ausdruck einer humanen Kultur erweisen.

659 Anselm, Albrecht, *Öffentlicher Protestantismus*, 49.

Literaturverzeichnis

- Alfonseca, Manuel/Cebrian, Manuel/Fernandez Anta, Antonio/Coviello, Lorenzo/Abeliuk, Andrés/Rahwan, Iyad, „Superintelligence Cannot be Contained: Lessons from Computability Theory“, *Journal of Artificial Intelligence Research (JAIR)* 70 (2021), 65–76 (= Alfonseca et al.).
- Alix, When hype is harmful: Why what we think is possible matters, <https://medium.com/swlh/when-hype-is-harmful-why-what-we-think-is-possible-matters-e7988db6f643> (27.9.2019) (= Alix).
- Anders, Günther, *Die Antiquiertheit des Menschen*. Bd. I: Über die Seele im Zeitalter der zweiten industriellen Revolution, München 1956 (= Anders).
- , *Die Antiquiertheit des Menschen*. Bd. II: Über die Zerstörung des Lebens im Zeitalter der dritten industriellen Revolution, München 1980 (= Anders).
- , „Apokalypse ohne Reich. Radikale Überlegungen zum atomaren Zeitalter“, in: *Die atomare Drohung*. Radikale Überlegungen zum atomaren Zeitalter, München 1993, 207–210 (= Anders).
- Anselm, Reiner, „Warum Wissenschaft Religion braucht“, in: Werner Matt/Birgit Brida/Wolfgang Ortner (Hgg.), *Das Montagsforum*. Versuche, die Welt zu verstehen; Festschrift anlässlich des 75. Geburtstages von Dkfm. Dr. Heinz Bertolini, Dornbirn 2010, 364–372 (= Anselm).
- , „Die Ethik im Kreis der theologischen Fächer“, in: Markus Buntfuß/Martin Fritz (Hgg.), *Fremde unter einem Dach?* Die theologischen Fächerkulturen in enzyklopädischer Perspektive, Berlin/Boston 2014, 115–125 (= Anselm).
- , „Politische Ethik“, in: Wolfgang Huber/Torsten Meireis (Hgg.), *Handbuch der Evangelischen Ethik.*, München 2015, 195–263 (= Anselm).
- Anselm, Reiner/Albrecht, Christian, *Öffentlicher Protestantismus*. Zur aktuellen Debatte um gesellschaftliche Präsenz und politische Aufgaben des evangelischen Christentums, Zürich 2017 (= Anselm/Albrecht).
- Anselm, Reiner/Schleissing, Stephan, „Einführung. Zum Ort der »Ethik« im Werk Trutz Rendtorffs“, in: Reiner Anselm/Stephan Schleissing (Hgg.), *Ethik*. Grundelemente, Methodologie und Konkretionen einer ethischen Theologie, Tübingen 2011, XI–XXII (= Anselm/Schleissing).

- Arendt, Hannah, „The Conquest of Space and the Stature of Man“, *The New Atlantis. A Journal of Technology and Society* (2007), 43–55 (= Arendt).
- , *Vita activa oder Vom tätigen Leben*, 19. Aufl., München/Berlin/Zürich 2018 (= Arendt).
- Armstrong, Stuart/OhEigartaigh, Sean, „Who Knows Anything about Anything about AI?“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 46–60 (= Armstrong/OhEigartaigh).
- Asprem, Egil, „The Magus of Silicon Valley. Immortality, Apocalypse, and God Making in Ray Kurzweil’s Transhumanism“, in: Ehler Voss (Hg.), *Mediality on Trial.*, Oldenburg 2020, 397–412 (= Asprem).
- Augustinus, Aurelius, *De Civitate Dei*. (Comm. Thomas Waleys et Nicolaus Trivet), Basel 1489 (= Augustinus).
- Bächle, Thomas Christian, *Mythos Algorithmus. Die Fabrikation des computerisierbaren Menschen*, Wiesbaden 2014 (= Bächle).
- Bacon, Francis, *Neues Organon*. Teilband 1: Lateinisch – Deutsch, 2. Aufl., Hamburg 1999 (= Bacon).
- , „Neu-Atlantis“, in: Klaus J. Heinisch (Hg.), *Der utopische Staat.*, Reinbek bei Hamburg Februar 2017, 171–219 (= Bacon).
- Bacon, Roger, „Epistola de secretis operibus artis et naturae, et de nullitate magiae“, in: J. S. Brewer (Hg.), *Rogeri Bacon Opera quaedam hactenus inedita.*, London 1859, 521–552 (= Bacon).
- Bartneck, Christoph/Lütge, Christoph/Wagner, Alan R., *Ethik in KI und Robotik*, München 2019 (= Bartneck/Lütge/Wagner).
- Bastian, Hans-Dieter, „Problemanzeigen einer kybernetischen Theologie“, *Evangelische Theologie* 28 (1968), 334–344 (= Bastian).
- Beck, Ulrich, *Risikogesellschaft*. Auf d. Weg in e. andere Moderne, 1. Aufl., Frankfurt am Main 1986 (= Beck).
- Benjamin, Walter, *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. Drei Studien zur Kunstsoziologie, Berlin, Frankfurt am Main 2018 (= Benjamin).
- Bense, Max, „Kybernetik oder Die Metaphysik einer Maschine (1951)“, in: Elisabeth Walther (Hg.), *Ausgewählte Schriften: in vier Bänden.*, Stuttgart/Weimar 1998, 429–448 (= Bense).

- Bernal, John Desmond, *The World, The Flesh And The Devil*. An Enquiry into the Future of the Three Enemies of the Rational Soul, London 1929 (= Bernal).
- Bertaux, Pierre, *Mutation der Menschheit*. Zukunft und Lebenssinn, Frankfurt am Main 1963 (= Bertaux).
- Bertaux, Pierre/Gehlen, Arnold, *Maschine. Denkmaschine, Staatsmaschine*. Entwicklungstendenzen der modernen Industriegesellschaft, Hamburg/Berlin 1963 (= Bertaux/Gehlen).
- Bloch, Ernst, *Das Prinzip Hoffnung*. In fünf Teilen. Kapitel 1–37, Frankfurt am Main 1959 (= Bloch).
- Blumenberg, Hans, *Arbeit am Mythos*, 6. Aufl., Frankfurt am Main 2001 (= Blumenberg).
- Bostrom, Nick, *Superintelligence*. Paths, dangers, strategies, Oxford 2014 (= Bostrom).
- , „Ethical Issues in Advanced Artificial Intelligence“, in: Wendell Wallach/Peter Asaro (Hgg.), *Machine Ethics and Robot Ethics*, London 2020, 69–75 (= Bostrom).
- Brand, Lukas, *Künstliche Tugend*. Roboter als moralische Akteure, 1. Aufl., Regensburg 2018 (= Brand).
- Brin, David/Broderick, Damien/Bostrom, Nick/Chislenko, Alexander “Sasha”/Hanson, Robinson/More, Max/Nielsen, Michael/Sandberg, Anders, „A Critical Discussion of Vinge’s Singularity Concept“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader*. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future, Chichester 2013, 395–418 (= Brin et al.).
- Brittnacher, Hans Richard, „Zu Grunde gehen in Sumpf, Schlamm und Fluten. Zur longue duree apokalyptischer Metaphorik“, in: Olaf Brieese/Richard Faber/et al. (Hgg.), *Aktualität des Apokalyptischen: zwischen Kulturkritik und Kulturversprechen*, Würzburg 2015, 99–116 (= Brittnacher).
- Bultmann, Rudolf, „Neues Testament und Mythologie. Das Problem der Entmythologisierung der neutestamentlichen Verkündigung“, in: H.W. Bartsch (Hg.), *Kerygma und Mythos*. Ein theologisches Gespräch, Hamburg 1960, 15–48 (= Bultmann).

- Butler, Samuel, „Darwin among the machines“, *The Press (Christchurch in New Zealand)* (13.06.1863) (= Butler).
- , *Erewhon*. or, Over the range, New York 1917 (1872) (= Butler).
- Chalmers, David J., „The Singularity: A Philosophical Analysis“, *Journal of Consciousness Studies* (2010), 7–65 (= Chalmers).
- Charbonnier, Ralph, *Technik und Theologie*. Ein theologischer Beitrag zum interdisziplinären Technikdiskurs unter besonderer Berücksichtigung der Theologie F.D.E. Schleiermachers, Marburg 2003 (= Charbonnier).
- , „Digitalisierung: Theologische Selbstklärungen und Gegenwartsinterpretationen. Eine Skizze“, *Zeitschrift für Pädagogik und Theologie* 70 (2018), 238–250 (= Charbonnier).
- Christoph Kehl/Christopher Coenen, *Technologien und Visionen der Mensch-Maschine-Entgrenzung*. Sachstandsbericht zum TA-Projekt »Mensch-Maschine-Entgrenzungen: zwischen künstlicher Intelligenz und Human Enhancement«, Berlin 2016 (= Christoph Kehl/Christopher Coenen).
- Clarke, Arthur Charles (Hg.), *Die sieben Sonnen*. Ein utopisch-techn. Abenteuerroman (Goldmanns Zukunftsromane Bd. Z. 13), München 1960.
- Crutzen, Paul J., „Die Geologie der Menschheit“, in: Paul J. Crutzen/Michael D. Mastrandrea/Stephen H. Schneider/Mike Davis/Peter Sloterdijk (Hgg.), *Das Raumschiff Erde hat keinen Notausgang*. Energie und Politik im Anthropozän, Berlin 2011, 7–10 (= Crutzen).
- Dabrock, Peter, Geheimnis, Freiheit, Verzeihen: Warum Big Data an die Lehre von der Vorsehung erinnert, https://zeitzeichen.net/archiv/2014_November_theologie-der-big-data (2.8.2019) (= Dabrock).
- Dalber-Acton, John Emerich Edward, „Appendix – Letter to Bishop Creighton. (5 April 1887)“, in: John Neville Figgis/Reginald Vere Laurence (Hgg.), *Historical Essays and Studies*., London 1907, 503–507 (= Dalber-Acton).
- Davis, Erik, *Techgnosis*. Myth, magic, mysticism in the age of information, 1. Aufl., New York 1998 (= Davis).
- Deutsch, Karl W., *The Nerves of Government*. Models of political Communication and Control, 2. Aufl., London 1963 (= Deutsch).

- Deutscher Pavillon der Architekturbiennale 2021, 2038 | Home, 21.5.2021, <https://2038.xyz/> (21.5.2021) (= Deutscher Pavillon der Architekturbiennale 2021).
- Deutschlandfunk Kultur, Künstliche Intelligenz – Der Mensch wird das alles nicht mehr verstehen, 17.2.2021, https://www.deutschlandfunkkultur.de/kuenstliche-intelligenz-der-mensch-wird-das-alles-nicht.1008.de.html?dram:article_id=399264 (17.2.2021) (= Deutschlandfunk Kultur).
- Devanthro GmbH, *Robodies: General Purpose Remote Bodies*, 16.11.2020, <https://devanthro.com/> (16.11.2020) (= Devanthro GmbH).
- Dierken, Jörg, „Technik als Kultur“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 21–37 (= Dierken).
- Dittmann, Franz, „Mensch und Roboter – ein ungleiches Paar.“, in: Fabian Karsch/Arne Manzeschke (Hgg.), *Roboter, Computer und Hybride. Was ereignet sich zwischen Menschen und Maschinen?*, Baden-Baden 2016, 17–46 (= Dittmann).
- Dreyfus, Hubert L., *What Computers can't do: a Critique of Artificial Reason*, New York 1972 (= Dreyfus).
- Dubbels, Elke, „Apokalypse der „Massen“ in wissenschaftlichem Diskurs, Drama und Film (1892–1927)“, in: Olaf Briese/Richard Faber/et al. (Hgg.), *Aktualität des Apokalyptischen: zwischen Kulturkritik und Kulturversprechen.*, Würzburg 2015, 39–60 (= Dubbels).
- Dumouchel, Paul/Damiano, Luisa, *Living with robots. Vivre avec les robots*, Cambridge, Massachusetts/London 2017 (= Dumouchel/Damiano).
- Eckert, Michael, *Transzendieren und immanente Transzendenz. Die Transformation der traditionellen Zweiweltentheorie von Transzendenz und Immanenz in Ernst Blochs Zweiseitentheorie*, Wien 1981 (= Eckert).
- Eichener, Elis/Peuckmann, Niklas, „Mind Uploading als Seele 4.0. Trauerkultur in Zeiten der Digitalisierung“, *Zeitschrift für Evangelische Ethik* 65 (2021), 114–127 (= Eichener/Peuckmann).
- Ellison, Harlan, *Ich muss schreien und habe keinen Mund. Erzählung*, München 2014 (= Ellison).
- Engineered Arts, Ameca: The Future Face of Robotics, <https://www.engineeredarts.co.uk/robot/ameca/> (20.1.2022) (= Engineered Arts).

- Esposito, Elena, *Die Fiktion der wahrscheinlichen Realität*, 3. Aufl., Frankfurt am Main 2014 (= Esposito).
- Eurich, Claus, „Mythos und Maschine“, in: Bernd Flessner (Hg.), *Nach dem Menschen*. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur, Freiburg im Breisgau 2000, 19–41 (= Eurich).
- European Council on Ethics in Science and Technology, *A Definition of AI: Main Capabilities and Disciplines* 08.04.2019 (= European Council on Ethics in Science and Technology).
- , *Ethics Guidelines for Trustworthy AI*, Brüssel 08.04.2019 (= European Council on Ethics in Science and Technology).
- , *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*, Brüssel 21.04.2021 (= European Council on Ethics in Science and Technology).
- Evangelische Kirche in Deutschland, *Freiheit digital. Die Zehn Gebote in Zeiten des digitalen Wandels. Eine Denkschrift der Evangelischen Kirche in Deutschland, Leipzig 2021* (= Evangelische Kirche in Deutschland (EKD)).
- Evers, Dirk, „Der Mensch als Turing-Maschine?“, *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie* 47 (2005), 101–118 (= Evers).
- Fischer, Johannes, *Theologische Ethik*. Grundwissen und Orientierung, Stuttgart 2002 (= Fischer).
- , „Vier Ebenen der Narrativität. Die Bedeutung der Erzählung in theologisch-ethischer Perspektive“, in: Karen Joisten (Hg.), *Narrative Ethik*. Das Gute und das Böse erzählen, Berlin 2007, 235–252 (= Fischer).
- , *Verstehen statt Begründen*. Warum es in der Ethik um mehr als nur um Handlungen geht, 1. Aufl., Stuttgart 2012 (= Fischer).
- Flessner, Bernd, „Antizipative Diffusion. Science Fiction als Akzeptanzbeschleuniger und Wegbereiter einer multitechnokulturellen Gesellschaft“, in: ders. (Hg.), *Nach dem Menschen*. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur, Freiburg im Breisgau 2000, 245–264 (= Flessner).

- , (Hg.), *Nach dem Menschen*. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur (Rombach Wissenschaften Reihe Litterae 72), 1. Aufl., Freiburg im Breisgau 2000.
- , „Die Rückkehr der Magier. Die KI als lapis philosophorum des 21. Jahrhunderts“, in: Dierk Spreen/Bernd Flessner (Hgg.), *Kritik des Transhumanismus*. Über eine Ideologie der Optimierungsgesellschaft, Bielefeld 106–63, 2018 (= Flessner).
- Floridi, Luciano, „What the Near Future of Artificial Intelligence Could Be“, in: Christopher Burr/Silvia Milano (Hgg.), *The 2019 Yearbook of the Digital Ethics Lab*, Cham 2020, 127–142 (= Floridi).
- Foerst, Anne/Kather, Regine, *Von Robotern, Mensch und Gott*. Künstliche Intelligenz und die existentielle Dimension des Lebens, Göttingen 2008 (= Foerst/Kather).
- Freud, Sigmund, „Über einige Übereinstimmungen im Seelenleben der Wilden und der Neurotiker. III. Animismus, Magie und Allmacht der Gedanken.“, *Imago. Zeitschrift für Anwendung der Psychoanalyse auf die Geisteswissenschaften* (1913), 1–21 (= Freud).
- Frisch, Max, *Homo Faber*. Ein Bericht, Frankfurt a.M. 1977 (= Frisch).
- Ganascia, Jean-Gabriel, *Le Mythe de la Singularité*. Faut-il craindre l'intelligence artificielle?, Paris 2017 (= Ganascia).
- Gehlen, Arnold, *Die Seele im technischen Zeitalter*. Sozialpsychologische Probleme in der industriellen Gesellschaft, 104. Aufl., Reinbek bei Hamburg 1976 (= Gehlen).
- Geraci, Robert, „Apocalyptic AI: Religion and the Promise of Artificial Intelligence“, *Journal of the American Academy of Religion* (2008), 138–166 (= Geraci).
- Geraci, Robert M., *Apocalyptic AI*. Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality, Oxford 2010 (= Geraci).
- Gödel, Kurt, „Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I“, *Monatshefte für Mathematik und Physik* 38–38 (1931), 173–198 (= Gödel).
- Goertzel, Ben, „Artificial Intelligence and the Future of Humanity“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader*. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future, Chichester 2013, 128–137 (= Goertzel).

- Good, Irving John, „Speculations Concerning the First Ultraintelligent Machine“, *Advances in Computers* 6 (1966), 31–88 (= Good).
- Gräb-Schmidt, Elisabeth, „Der Homo Faber als Homo Religiosus. Zur anthropologischen Dimension der Technik“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 39–55 (= Gräb-Schmidt).
- Graf, Friedrich Wilhelm, „Von der creatio ex nihilo zur »Bewahrung der Schöpfung«: Dogmatische Erwägungen zur Frage nach einer möglichen ethischen Relevanz der Schöpfungslehre“, *Zeitschrift für Theologie und Kirche* 87 (1990), 206–223 (= Graf).
- Grunwald, Armin, „Technische Zukunft des Menschen? Eschatologische Erzählungen zur Digitalisierung und ihre Kritik“, in: Mitscherlich-Schönherr/Reiner Anselm/Martin Heinze (Hgg.), *Das Gelingen der künstlichen Natürlichkeit. Mensch-Sein an den Grenzen des Lebens mit disruptiven Biotechnologien*, Berlin/Boston 2021, 314–331 (= Grunwald).
- Gumbrecht, Hans Ulrich, *Weltgeist im Silicon Valley. Leben und Denken im Zukunftsmodus*, Basel 2018 (= Gumbrecht).
- Günther, Gotthard, *Das Bewußtsein der Maschinen. Eine Metaphysik der Kybernetik*, 3. Aufl., Baden-Baden 2002 (= Günther).
- Guthmüller, Bodo, „»Trasumanar significar per verba / non si poria«. Zu Dantes erstem Gesang des Paradiso“, *Deutsches Dante-Jahrbuch* 82 (2007), 67–86 (= Guthmüller).
- Gutmann, Mathias/Knifka, Julia, „Biomorphic and Technomorphic Metaphors – Some Arguments Why Robots do not Evolve, Why Computing is not Organic, and Why Adaptive Technologies Are not Intelligent“, in: Michael Decker/Mathias Gutmann/Julia Knifka (Hgg.), *Evolutionary Robotics, Organic Computing and Adaptive Ambience. Epistemological and Ethical Implications of Technomorphic Descriptions of Technologies*, Zürich 2015, 53–78 (= Gutmann/Knifka).
- Haberer, Johanna, *Digitale Theologie. Gott und die Medienrevolution der Gegenwart*, München 2015 (= Haberer).
- Habermas, Jürgen, „Technik und Wissenschaft als ›Ideologie‹“, in: Jürgen Habermas (Hg.), *Technik und Wissenschaft als ›Ideologie‹*, Frankfurt am Main 1968, 48–119 (= Habermas).

- Haeckel, Ernst, *Gemeinverständliche Vorträge und Abhandlungen*, Bonn 1902 (= Haeckel).
- Hagner, Michael, „Vom Aufstieg und Fall der Kybernetik als Universalwissenschaft“, in: Michael Hagner/Erich Hörl (Hgg.), *Transformation des Humanen. Beiträge zur Kulturgeschichte der Kybernetik*, Frankfurt am Main 2008, 38–71 (= Hagner).
- Hanson, Robinson, „What Will It Be Like To Be an Emulation?“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 298–309 (= Hanson).
- Harari, Yuval Noah, *Homo Deus. Eine Geschichte von Morgen*, 1. Aufl., München 2017 (= Harari).
- Haraway, Donna, „A Cyborg Manifesto. Science, Technology, and Socialist-Feminism in the Late Twentieth Century“, in: Donna Jeanne Haraway (Hg.), *Simians, Cyborgs, and Women. The Reinvention of Nature*, New York 2015, 149–181 (= Haraway).
- Heidegger, Martin, „Die Herkunft der Kunst und die Bestimmung des Denkens. Vortrag gehalten am 4. April 1967 in der Akademie der Wissenschaften und Künste in Athen“, in: Petra Jaeger/Rudolf Lütke (Hgg.), *Distanz und Nähe. Reflexionen u. Analysen zur Kunst d. Gegenwart*. Hrsg. von Petra Jaeger u. Rudolf Lütke. [Walter Biemel zum 65. Geburtstag gewidmet, Würzburg 1983, 11–22 (= Heidegger).
- , *Die Technik und Die Kehre*, 9. Aufl., Stuttgart 1996 (= Heidegger).
- Heisenberg, Werner, *Das Naturbild der heutigen Physik*, 1. Aufl., Hamburg 1955 (= Heisenberg).
- Herbert, Frank, *Dune*, New York 2016 (= Herbert).
- Herrgen, Matthias, „‘Leben’ bei Darwin und im Darwinismus“, in: Stephan Schaede (Hg.), *Das Leben. Historisch-systematische Studien zur Geschichte eines Begriffs*, Tübingen 2012, 45–60 (= Herrgen).
- Hoffmann, E. T. A., „Der Sandmann“, in: Carl Georg von Maassen (Hg.), *E. T. A. Hoffmanns Sämtliche Werke. Historisch-kritische Ausgabe mit Einleitungen, Anmerkungen und Lesarten von Carl Georg von Maassen*, München, Leipzig 1909, 3–42 (= Hoffmann).
- Höhne, Florian, „Darf ich vorstellen: Digitalisierung. Anmerkungen zu Narrativen und Imaginationen digitaler Kulturpraktiken in theologisch-ethischer Perspektive“, in: Jonas Bedford-Strohm/Florian Höhne/Julian Zeyher-Quattlander (Hgg.), *Digitaler Strukturwandel*

- der Öffentlichkeit*. Interdisziplinäre Perspektiven auf politische Partizipation im Wandel, Baden-Baden 2019, 25–46 (= Höhne).
- Horkheimer, Max/Adorno, Theodor W., *Dialektik der Aufklärung*. Philosophische Fragmente, 23. Aufl., Frankfurt am Main August 2017 (= Horkheimer/Adorno).
- Hörl, Erich/Hagner, Michael, „Überlegungen zur kybernetischen Transformation des Humanen“, in: dies. (Hgg.), *Transformation des Humanen. Beiträge zur Kulturgeschichte der Kybernetik.*, Frankfurt am Main 2008, 7–37 (= Hörl/Hagner).
- Huber, Wolfgang, „Fortschritt durch Freiheit. Zur Erwartung politischer Emanzipation“, in: Wolfgang Lienemann (Hg.), *Fortschrittsglaube und Wirklichkeit*. Arbeiten zu einer Frage unserer Zeit, München 1983, 47–65 (= Huber).
- , (Hg.), *Von der Freiheit*. Perspektiven für eine solidarische Welt (Beck'sche Reihe 6065), München 2012.
- Huxley, Julian, „Transhumanism“, *Journal of Humanistic Psychology* 8 (1968), 73–76 (= Huxley).
- Innerhofer, Roland, „Die technische Modernisierung des künstlichen Menschen in der Literatur zwischen 1800 und 1900“, in: Bernd Flessner (Hg.), *Nach dem Menschen*. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur, Freiburg im Breisgau 2000, 69–99 (= Innerhofer).
- Itskov, Dmitry, 2045 Initiative, 4.8.2020, <http://2045.com/> (4.8.2020) (= Itskov).
- Jalobeanu, Dana, „Bacon's Brotherhood and its classical Sources: Producing and Communicating Knowledge in the Project of the Great Instauration“, in: Claus Zittel (Hg.), *Philosophies of Technology*. Francis Bacon and his Contemporaries, Leiden/Boston 2010, 197–230 (= Jalobeanu).
- Jasanoff, Sheila/Kim, Sang-Hyun (Hgg.), *Dreamscapes of Modernity*. Sociotechnical Imaginaries and the Fabrication of Power, Chicago/London 2015.
- John McCarthy/Marvin L. Minsky/Nathaniel Rochester/and Claude E. Shannon, „A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955“, *AI Magazin* 27 (2016), 12–14 (= John McCarthy et al.).

- Jonas, Hans, *Das Prinzip Verantwortung*. Versuch einer Ethik für die technologische Zivilisation, 7. Aufl., Frankfurt am Main 1987 (= Jonas).
- , „Warum die Technik ein Gegenstand für die Ethik ist: Fünf Gründe“, in: Hans Lenk (Hg.), *Technik und Ethik*, Stuttgart 1993, 81–91 (= Jonas).
- , *Das Prinzip Leben*. Ansätze zu einer philosophischen Biologie, 1. Aufl., Göttingen 1994 (= Jonas).
- , *Gnosis*. Die Botschaft des fremden Gottes, 1. Aufl., Frankfurt am Main/Leipzig 2008 (= Jonas).
- Joy, Bill, Why the Future Doesn't Need Us, 1.4.2000, <https://www.wired.com/2000/04/joy-2/> (12.8.2020) (= Joy).
- Julien Offray de La Mettrie, *L'homme machine. Der Mensch eine Maschine, Leipzig 1875* (= Julien Offray de La Mettrie).
- Jung, Volker, *Digital Mensch bleiben*, München 2018 (= Jung).
- Kaltenborn, Olaf, *Das künstliche Leben. Die Grundlagen der Dritten Kultur, München 2001* (= Kaltenborn).
- Käsemann, Ernst, „Die Anfänge der christlichen Theologie“, in: Ernst Käsemann (Hg.), *Exegetische Versuche und Besinnungen II.*, Göttingen 1964, 82–104 (= Käsemann).
- Katzenbach, Christian, I miss Imaginaries that spell out Ideas of AI as a Public Good, 21.5.2019, <https://www.hiig.de/en/the-global-ai-race-analysing-different-strategies-and-imaginaries/> (14.12.2020) (= Katzenbach).
- Kehl, Christoph, „Möglichkeiten und Grenzen ethischer Technikgestaltung. Das Beispiel der Mensch-Maschine-Entgrenzung“, in: Mitscherlich-Schönherr/Reiner Anselm/Martin Heinze (Hgg.), *Das Gelingen der künstlichen Natürlichkeit*. Mensch-Sein an den Grenzen des Lebens mit disruptiven Biotechnologien, Berlin/Boston 2021, 151–170 (= Kehl).
- Kelly, Kevin, The Three Breakthroughs That Have Finally Unleashed AI on the World, <https://www.wired.com/2014/10/future-of-artificial-intelligence/> (9.6.2020) (= Kelly).
- Kirste, Moritz/Schürholz, Markus, „Entwicklungswege zur KI“, in: Volker Wittpahl (Hg.), *Künstliche Intelligenz*. Technologie, Berlin, Heidelberg 2019, 21–35 (= Kirste/Schürholz).
- Koslowski, Peter, „Mensch-Maschine-Hybride: Dinge, die sprechen, und Maschinen, die unsere Brüder werden“, in: Christoph Hubig/

- Peter Koslowski (Hgg.), *Maschinen, die unsere Brüder werden. Mensch-Maschine-Interaktion in hybriden Systemen*, Paderborn 2008, 191–202 (= Koslowski).
- Krüger, Oliver, *Virtualität und Unsterblichkeit*. Gott, Evolution und die Singularität im Post- und Transhumanismus, 2. Aufl., Freiburg i. Br./Berlin/Wien 2019 (= Krüger).
- Kurzweil, Raymond, *The Singularity Is Near. When Humans Transcend Biology*, New York/Cambridge/et al. 2005 (= Kurzweil).
- , *The Age of Spiritual Machines. Die Intelligenz der Evolution*. Wenn Mensch und Computer verschmelzen, 1. Aufl., Köln 2016 (1999) (= Kurzweil).
- Lang, Jürgen P., „Karl W. Deutsch, The Nerves of Government. Models of Political Communication and Control, New York 1963“, in: Steffen Kailitz (Hg.), *Schlüsselwerke der Politikwissenschaft.*, Wiesbaden 2007, 92–95 (= Lang).
- LesswrongWiki, Roko's basilisk, 2.1.2020, https://wiki.lesswrong.com/wiki/Roko%27s_basilisk (4.8.2020) (= LesswrongWiki).
- Levy, David, *Love and Sex with Robots. The Evolution of Human-Robot Relations*, New York/London/Toronto 2008 (= Levy).
- Lierfeld, Karl Johannes, *Künstliche Superintelligenz und/oder Ethik. Utopien, Dystopien, Disruptionen*, Bochum 2019 (= Lierfeld).
- Löwith, Karl, *Weltgeschichte und Heilsgeschehen. Die theologischen Voraussetzungen der Geschichtsphilosophie*, 8. Aufl., Stuttgart 1990 (= Löwith).
- Luther, Martin, „Bekenntnis (1528). Übertragung: Christopher Spehr“, in: Dietrich Korsch (Hg.), *Deutsch-deutsche Studienausgabe / Luther, Martin*. Band 1 Glaube und Leben, Leipzig 2012, 551–570 (= Luther).
- , „Von der Freiheit eines Christenmenschen (1520). Übertragung: Dietrich Korsch“, in: Dietrich Korsch (Hg.), *Deutsch-deutsche Studienausgabe / Luther, Martin*. Band 1 Glaube und Leben, Leipzig 2012, 277–317 (= Luther).
- , „Von der weltlichen Obrigkeit: Wie weit man ihr Gehorsam schuldet (1523)“, in: Hellmut Zschoch/Johannes Schilling/Albrecht Beutel/Dietrich Korsch/Notger Slenczka (Hgg.), *Deutsch-deutsche Studienausgabe / Luther, Martin*. Band 3 Christ und Welt, Leipzig 2016, 217–290 (= Luther).

- Macho, Thomas, „Technische Utopien und Katastrophenängste“, *Gegenworte* (2002), 11–13 (= Macho).
- Mack, Eric, Elon Musk: ‘We are summoning the Demon’ with Artificial Intelligence, 27.8.2020, <https://www.cnet.com/news/elon-musk-we-are-summoning-the-demon-with-artificial-intelligence/> (28.8.2020) (= Mack).
- Mager, Astrid/Katzenbach, Christian, „Future Imaginaries in the Making and Governing of Digital Technology: Multiple, Contested, Commodified“, *New Media & Society* (2020), 1–12 (= Mager/Katzenbach).
- Marcuse, Herbert, „Einige gesellschaftliche Folgen moderner Technologie“, in: *Schriften*. Band III: Aufsätze aus der Zeitschrift für Sozialforschung 1934–1941, Frankfurt am Main, 286–319 (= Marcuse).
- Marquard/Odo, „Zeitalter der Weltfremdheit?: Ein Beitrag zur Analyse der Gegenwart“, *Giessener Universitätsblätter* (1985), 9–20 (= Marquard/Odo).
- Marquard, Odo, „Entlastungen. Theodizeemotive in der neuzeitlichen Philosophie“, in: *Zukunft braucht Herkunft*. Philosophische Essays, Stuttgart 2015, 124–145 (= Marquard).
- McArthur, Neil/Twist, Markie L.C., „The Rise of Digisexuality: Therapeutic Challenges and Possibilities“, *Sexual and Relationship Therapy* 32 (2017), 334–344 (= McArthur/Twist).
- McLuhan, Marshall, *Understanding Media*. The Extensions of Man, 1. Aufl., Cambridge, Mass 1994 (= McLuhan).
- Meireis, Torsten, „O daß ich tausend Zungen hätte. Chancen und Gefahren der digitalen Transformation politischer Öffentlichkeit – die Perspektive evangelischer Theologie“, in: Jonas Bedford-Strohm/Florian Höhne/Julian Zeyher-Quattlender (Hgg.), *Digitaler Strukturwandel der Öffentlichkeit*. Interdisziplinäre Perspektiven auf politische Partizipation im Wandel, Baden-Baden 2019, 47–62 (= Meireis).
- Metzinger, Thomas, *Being no One*. The Self-Model Theory of Subjectivity, 1. Aufl., Cambridge, Mass. 2003 (= Metzinger).
- Misselhorn, Catrin, *Maschinen mit Moral?* Grundfragen der Maschinenethik, Ditzingen 2018 (= Misselhorn).
- Moltmann, Jürgen, „Ist die Welt unfertig?“, *Evangelische Theologie* 72 (2012), 326–336 (= Moltmann).

- Moravec, Hans P., *Mind Children. The Future of Robot and Human Intelligence*, Cambridge 1988 (= Moravec).
- , *Mind Children. Der Wettlauf zwischen menschlicher und künstlicher Intelligenz*, Hoffman und Campe 1990 (= Moravec).
- , *Computer übernehmen die Macht. Vom Siegeszug der künstlichen Intelligenz*, 1. Aufl., Hamburg 1999 (= Moravec).
- , „Pigs in Cyberspace“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 177–181 (= Moravec).
- More, Max, „Digital Self-Transformation: Expanding Personal Extropy“, *Extropy* #10 4 (1993) (= More).
- , „A Letter to Mother Nature. Original 1992“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 449–450 (= More).
- , „The Enhanced Carnality of Post-Biological Life“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 222–230 (= More).
- More, Max/Vita-More, Natasha (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013.
- Müller, Oliver, „Von der Selbstüberschreitung zur Selbstersetzung. Zu einigen anthropologischen Tiefenstrukturen des Transhumanismus“, in: Mitscherlich-Schönherr/Reiner Anselm/Martin Heinze (Hgg.), *Das Gelingen der künstlichen Natürlichkeit. Mensch-Sein an den Grenzen des Lebens mit disruptiven Biotechnologien*, Berlin/Boston 2021, 333–350 (= Müller).
- Nachtwey, Oliver/Seidl, Timl, *Die Ethik der Solution und der Geist des digitalen Kapitalismus*, Frankfurt am Main (= Nachtwey/Seidl).
- Nassehi, Armin, *Muster. Theorie der digitalen Gesellschaft*, München 2019 (= Nassehi).
- Nature, „Let’s talk about Sex Robots“, *Nature* 547 (2017), 138 (= Nature).

- Naumann, Friedrich, „Im Eisenwerk“, in: Wolfgang Philipp (Hg.), *Der Protestantismus des 19. und 20. Jahrhunderts.*, Wuppertal 1988, 213–218 (= Naumann).
- , „Jesus, der Sozialist“, in: Wolfgang Philipp (Hg.), *Der Protestantismus des 19. und 20. Jahrhunderts.*, Wuppertal 1988, 222–224 (= Naumann).
- Nemitz, Paul, „Constitutional Democracy and Technology in the Age of Artificial Intelligence“, *Philosophical transactions. Series A, Mathematical, Physical, and Engineering Sciences* 376 (2018), 1–14 (= Nemitz).
- Newell, Allen/Simon, Herbert A., „GPS, a program that simulates human thought (1961)“, in: Edward Feigenbaum (Hg.), *Computers and Thought.*, Menlo Park, Calif. 1995, 279–293 (= Newell/Simon).
- Nida-Rümelin, Julian/Weidenfeld, Nathalie, *Digitaler Humanismus. Eine Ethik für das Zeitalter der künstlichen Intelligenz*, München 2018 (= Nida-Rümelin/Weidenfeld).
- Nietzsche, Friedrich, *Jenseits von Gut und Böse. Zur Genealogie der Moral* 1988 (= Nietzsche).
- Nietzsche, Friedrich/Simon, Josef, *Also sprach Zarathustra. Ein Buch für Alle und Keinen*, Stuttgart 1994 (= Nietzsche/Simon).
- Ohly, Lukas, *Ethik der Robotik und der Künstlichen Intelligenz. Theologisch-Philosophische Beiträge zu Gegenwartsfragen Band 22*, Bristol 2019 (= Ohly).
- Pannenberg, Wolfhart, *Das Glaubensbekenntnis. ausgelegt und verantwortet vor den Fragen der Gegenwart*, 6. Aufl., Gütersloh 1995 (= Pannenberg).
- , *Grundlagen der Ethik. Philosophisch-theologische Perspektiven*, Göttingen 1996 (= Pannenberg).
- , „Fluch und Segen der Arbeit. Das Phänomen menschlicher Arbeit aus theologischer Sicht (1986)“, in: Wolfhart Pannenberg (Hg.), *Beiträge zur Systematischen Theologie*. Bd. 2 Natur und Mensch – und die Zukunft der Schöpfung, Göttingen 2000, 202–214 (= Pannenberg).
- , „Geist und Energie. Zur Phänomenologie Teilhards de Chardin (1971)“, in: Wolfhart Pannenberg (Hg.), *Beiträge zur Systematischen Theologie*. Bd. 2 Natur und Mensch – und die Zukunft der Schöpfung, Göttingen 2000, 55–65 (= Pannenberg).

- Pias, Claus, „Unruhe und Steuerung. Zum utopischen Potential der Kybernetik“, in: Jörn Rüsen (Hg.), *Die Unruhe der Kultur. Potentiale des Utopischen*, Weilerswist 2004, 301–325 (= Pias).
- , „Zeit der Kybernetik – Eine Einstimmung“, in: ders. (Hg.), *Cybernetics / Kybernetik. The Macy-Conferences 1946–1953*, Berlin 2004, 9–41 (= Pias).
- Popper, Karl R., *Die offene Gesellschaft und ihre Feinde 1. Der Zauber Platons*, 8. Aufl., München 2003 (= Popper).
- Pratt, Gill A., „Is a Cambrian Explosion Coming for Robotics?“, *Journal of Economic Perspectives* 29 (2015), 51–60 (= Pratt).
- Precht, Richard David, *Jäger, Hirten, Kritiker. Eine Utopie für die digitale Gesellschaft*, München 2018 (= Precht).
- Quantic Dream, *Detroit: Become Human* 2018 (= Quantic Dream).
- Ramge, Thomas, *Mensch und Maschine. Wie künstliche Intelligenz und Roboter unser Leben verändern*, Ditzingen 2018 (= Ramge).
- Rapp, Hans Reinhard, *Mensch, Gott und Zahl. Kybernetik im Horizont d. Theologie*, 2. Aufl., Hamburg 1971 (= Rapp).
- RealDoll, Product Site HarmonyX, <https://www.realdoll.com/product/harmony-x/> (22.1.2022) (= RealDoll).
- Rendtorff, Trutz, „Strukturen und Aufgaben technischer Kultur“, in: Dietrich Rössler (Hg.), *Möglichkeiten und Grenzen der technischen Kultur. Symposium Hotel Schloß Fuschl, Österreich* 1981, 9–21 (= Rendtorff).
- , *Vielspältiges. Protestantische Beiträge zur ethischen Kultur*, Stuttgart/Berlin 1991 (= Rendtorff).
- , *Ethik. Grundelemente, Methodologie und Konkretionen einer ethischen Theologie*, 3. Aufl., Tübingen 2011 (= Rendtorff).
- Riedl, Matthias, „Die Welt als Kloster. Joachim von Fiore und sein Verfassungsentwurf für die zukünftige Menschheit“, in: Katharina Martin/Christian Sieg (Hgg.), *Zukunftsvisionen zwischen Apokalypse und Utopie.*, Würzburg 2016, 127–147 (= Riedl).
- Rothblatt, Martine, „Mind is Deeper than Matter: Transgenderism, Transhumanism, and the Freedom of Form“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 317–326 (= Rothblatt).

- Rothe, Richard, *Stille Stunden*, Wittenberg 1872 (= Rothe).
- Russel, Stuart, AI Open Letter – Future of Life Institute, <https://futureoflife.org/ai-open-letter/?cn-reloaded=1> (19.9.2019) (= Russel).
- Sandberg, Anders, „An Overview of Models of Technological Singularity“, in: Max More/Natasha Vita-More (Hgg.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester 2013, 376–394 (= Sandberg).
- Sartre, Jean-Paul, *Geschlossene Gesellschaft*. Stück in e. Akt (= Huis clos), Reinbek bei Hamburg 1975 (= Sartre).
- Scheler, Max, *Das Ressentiment im Aufbau der Moralen*, Frankfurt am Main 1978 (= Scheler).
- , *Die Stellung des Menschen im Kosmos*, 12. Aufl., Bonn 1991 (1928) (= Scheler).
- Schipper, Bernd U., „Zwischen apokalyptischen Ängsten und chilias-tischen Hoffnungen. Die religiöse Dimension moderner Utopien“, in: Reto Sorg/Stefan Bodo Würffel (Hgg.), *Utopie und Apokalypse in der Moderne.*, Paderborn 2010, 47–61 (= Schipper).
- Schleiermacher, Friedrich, *Ethik*. mit späteren Fassungen der Einleitung, Güterlehre und Pflichtenlehre, Hamburg 1981 (= Schleiermacher).
- , *Über die Religion*. Reden an die Gebildeten unter ihren Verächtern (1799), 2. Aufl., Berlin 2001 (= Schleiermacher).
- Schleissing, Stephan, „Laien, Experten, Propheten: Zur Rolle von Theologen in Technikdiskursen“, in: Anne-Maren Richter/Christian Schwarke (Hgg.), *Technik und Lebenswirklichkeit. philosophische und theologische Deutungen der Technik im Zeitalter der Moderne*, Stuttgart 2014, 203–216 (= Schleissing).
- Schmidt, Hermann, „Regelungstechnik. Die technische Aufgabe und ihre wirtschaftliche, sozialpolitische und kulturpolitische Auswirkung“, *Zeitschrift des Vereins deutscher Ingenieure* (1941), 81–88 (= Schmidt).
- Schnabel, Ulrich, Hirnforschers Mondfahrt. Human Brain Project, 27.11.2019, <https://www.zeit.de/2019/49/human-brain-project-hirnforschung-ausstellung-finanzierung-neurowissenschaften> (27.1.2022) (= Schnabel).
- Schütt, Hans-Werner, *Auf der Suche nach dem Stein der Weisen*. Die Geschichte der Alchemie, München 2000 (= Schütt).

- Schwarke, Christian, „Technik und Theologie. Was ist der Gegenstand einer theologischen Technikethik?“, *Zeitschrift für Evangelische Ethik* (2005), 88–104 (= Schwarke).
- , „Einleitung“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 9–17 (= Schwarke).
- Schwarz, Elke, *Death Machines. The Ethics of Violent Technologies*, Manchester 2018 (= Schwarz).
- Searle, John R., „Minds, Brains, and Programs“, *The Behavioral and Brain Sciences* (1980), 417–457 (= Searle).
- Shannon, Claude Elwood/Weaver, Warren, *The Mathematical Theory of Communication*, Illinois 1964 (= Shannon/Weaver).
- Shelley, Mary, *Frankenstein. Or, the modern Prometheus*, London 2012 (= Shelley).
- Singler, Beth, „Roko’s Basilisk or Pascal’s? Thinking of Singularity Thought Experiments as Implicit Religion“, *Implicit Religion* 20 (2018), 279–297 (= Singler).
- Singularity University, About SU – Singularity University, 17.2.2021, <https://su.org/about/> (17.2.2021) (= Singularity University).
- Sölle, Dorothee, *Atheistisch an Gott glauben. Beitr. z. Theologie*, Olten/Freiburg i. Br. 1968 (= Sölle).
- , *Stellvertretung. Ein Kapitel Theologie nach dem Tode Gottes*, Stuttgart 1982 (= Sölle).
- Spaemann, Robert/Löw, Reinhard, *Natürliche Ziele. Geschichte und Wiederentdeckung teleologischen Denkens*, Stuttgart 2005 (= Spaemann/Löw).
- Spengler, Oswald, *Der Untergang des Abendlandes. Umriss einer Morphologie der Weltgeschichte*, 230. Aufl., München 1998 (= Spengler).
- Spiekermann, Sarah, *Digitale Ethik. Ein Wertesystem für das 21. Jahrhundert*, München 2019 (= Spiekermann).
- , „Der Mensch: ein Fehler“ (23.03.2019) (= Spiekermann).
- Suarez-Nani, Tizina, „Individualität und Subjektivität der Engel im 13. Jahrhundert. Thomas von Aquin, Heinrich von Gent und Petrus Johannis Olivi“, *Das Mittelalter* 11 (2006), 29–48 (= Suarez-Nani).
- Taylor, Charles, *Modern Social Imaginaries*, 4. Aufl., Durham 2004 (= Taylor).

- Tegmark, Max, *Life 3.0. Being Human in the Age of Artificial Intelligence*, New York 2017 (= Tegmark).
- Teilhard de Chardin, Pierre, *Der Mensch im Kosmos*, 7. Aufl., München 1964 (= Teilhard de Chardin).
- The Human Brain Project, Brain Models and Simulation, <https://www.humanbrainproject.eu/en/brain-simulation/> (25.1.2022) (= The Human Brain Project).
- The Singularitarian Principles, Eliezer S. Yudkowsky: The Singularitarian Principles Version 1.0.2, 23.10.2020, <https://museum.netstalking.ru/cyberlib/lib/critica/sing/singprinc.html> (21.8.2021) (= The Singularitarian Principles).
- Theunissen, Michael, „Freiheit und Schuld – Freiheit und Sünde“, in: Hans-Richard Reuter/Heinrich Bedford-Strohm/Helga Kuhlmann/Karl-Heinrich Lütcke (Hgg.), *Freiheit verantworten*. Festschrift für Wolfgang Huber zum 60. Geburtstag, Gütersloh 2002, 343–356 (= Theunissen).
- Thiede, Werner, „Digitaler Turmbau zu Babel“, in: Severin Lederhilger (Hg.), *Gott und die digitale Revolution.*, Regensburg 2019, 47–72 (= Thiede).
- Tillich, Paul, *Die religiöse Substanz der Kultur*. Schriften zur Theologie der Kultur, 2. Aufl., Stuttgart 1975 (= Tillich).
- Tipler, Frank J., „The Omega Point as Eschaton: Answers to Pannenberg's Questions for Scientists“, in: Carol Rausch Albright/Joel Haugen (Hgg.), *Beginning with the End: God, Science, and Wolfhart Pannenberg.*, Chicago 1997, 156–194 (= Tipler).
- , „If You Can't Beat 'em, Join 'em — Frank Tipler“, in: John Brockman (Hg.), *What to think about Machines that think*. Today's leading Thinkers on the Age of Machine Intelligence, New York/London/Toronto/Sydney/New Delhi/Auckland 2015, 30–31 (= Tipler).
- Tödt, Heinz Eduard, „Ambivalenz des Fortschritts. Zur Urteilsfindung über das Wesen der Macht“, in: Wolfgang Lienemann (Hg.), *Fortschrittsglaube und Wirklichkeit*. Arbeiten zu einer Frage unserer Zeit, München 1983, 141–156 (= Tödt).
- Troeltsch, Ernst, „Grundprobleme der Ethik. Erörtert aus Anlass von Hermanns Ethik“, in: Hartmut Spenner (Hg.), *Ethik / Grundprobleme der Ethik.*, Waltrop 2002, 44–94.125–178 (= Troeltsch).

- Turing, Alan Mathison, „Computing Machinery and Intelligence“, *Mind* LIX (1950), 433–460 (= Turing).
- Ulam, Stanislaw, „Tribute to John von Neumann 1903–1957“, *Bulletin of the American Mathematical Society* 64 (1958), 1–49 (= Ulam).
- Verne, Jules, *Das Karpatenschloß*, Berlin/Herrsching 1984 (= Verne).
- Vinge, Vernor, „The Coming of Technological Singularity: How to Survive the Post-Human Era“, in: NASA. Lewis Research Center (Hg.), *Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*. 1993, 11–22 (= Vinge).
- Vondung, Klaus, „Utopische Entwürfe – apokalyptische Visionen. Träume vom besseren Leben“, in: Katharina Martin/Christian Sieg (Hgg.), *Zukunftsvisionen zwischen Apokalypse und Utopie.*, Würzburg 2016, 45–59 (= Vondung).
- Walker, Mark, „Uploading and Personal Identity“, in: Damien Broderick/Russell Blackford (Hgg.), *Intelligence Unbound. The Future of Uploaded and Machine Minds*, Chichester, West Sussex 2014, 161–177 (= Walker).
- Watson, David, „The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence“, in: Christopher Burr/Silvia Milano (Hgg.), *The 2019 Yearbook of the Digital Ethics Lab.*, Cham 2020, 45–65 (= Watson).
- Way of the Future Church, *Way of the Future*, 15.3.2018, <http://www.wayofthefuture.church/> (15.1.2020) (= Way of the Future Church).
- Weber, Max, „Wissenschaft als Beruf“, in: Dirk Kaesler (Hg.), *Schriften 1894 – 1922.*, Stuttgart 2002, 474–511 (= Weber).
- Weber, Tomas P., *Science Fiction*, 1. Aufl., Frankfurt am Main 2005 (= Weber).
- Weizenbaum, Joseph, „ELIZA – A computer Program for the Study of Natural Language Communication between Man and Machine“, *Communications of the ACM* 9 (1965), 36–45 (= Weizenbaum).
- , *Die Macht der Computer und die Ohnmacht der Vernunft*, 1. Aufl., Frankfurt am Main 1977 (= Weizenbaum).
- Weizsäcker, Carl Friedrich von, *Die Verantwortung der Wissenschaft im Atomzeitalter*, 7. Aufl., Göttingen 1986 (= Weizsäcker).
- Weizsäcker, Carl Friedrich von/et al., „Erklärung der 18 Atomwissenschaftler vom 12. April 1957“, in: *Die Verantwortung der Wissenschaft im Atomzeitalter.*, Göttingen 1986, 50–52 (= Weizsäcker/et al.).

- Wiener, Norbert, *God and Golem, Inc. A Comment on Certain Points where Cybernetics impinges on Religion*, Cambridge, Mass. 1964 (= Wiener).
- , *Kybernetik. Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine*, Düsseldorf 1992 (= Wiener).
- Wirth, Mathias, „Doketisch, pelagianisch, sarkisch? Transhumanismus und technologische Modifikationen des Körpers in einer theologischen Perspektive“, *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie* 60 (2018), 142–167 (= Wirth).
- , „Trans-Körper. Theologie im Gespräch mit Transhumanismus und Transsexualität“, *Zeitschrift für Evangelische Ethik* 62 (2018), 10–30 (= Wirth).
- Wittig, Frank, *Maschinenmenschen. Zur Geschichte eines literarischen Motivs im Kontext von Philosophie, Naturwissenschaft und Technik*, Würzburg 1997 (= Wittig).
- Woschech, Anke, „Zwischen Luftschloss und Prognose. Der Terminus der technischen Utopie im Fokus der Technikgeschichte“, in: Katharina Neumeister/Peggy Renger-Berka (Hgg.), *Technik und Transzendenz. Zum Verhältnis von Technik, Religion und Gesellschaft*, Stuttgart 2012, 147–164 (= Woschech).
- Wustmans, Clemens, „»Homo Deus« oder Menschsein in Grenzen und Relationalität?“, *Zeitschrift für Evangelische Ethik* 65 (2021), 46–51 (= Wustmans).
- Yudkowsky, Eliezer, Coherent Extrapolated Volition, <https://intelligence.org/files/CEV.pdf> (28.3.2022) (= Yudkowsky).
- , „Artificial Intelligence as a positive and negative Factor in Global Risk“, in: Nick Bostrom (Hg.), *Global Catastrophic Risks*. Oxford 2008 (= Yudkowsky).
- Zhou, Yuefeng/Fischer, H. Martin, „Intimate Relationships with Humanoid Robots: Exploring Human Sexuality in the Twenty-First Century“, in: Yuefeng Zhou (Hg.), *AI love you.*, Cham 2019, 177–184 (= Zhou/Fischer).
- Zittrain, Jonathan, „The Hidden Costs of Automated Thinking“, *The New York Times* (23.07.2019) (= Zittrain).
- Zizek, Slavoj, *Hegel im verdrahteten Gehirn*, Frankfurt am Main 2020 (= Zizek).

Register

Namensregister

A

- Adorno, Theodor 9, 31, 223
Alighieri, Dante 241f., 255
Allende, Salvador 138
Anders, Günther 5, 58f., 70, 104, 149,
154, 160–171, 196, 212, 240, 243, 245,
251, 263, 265
Arendt, Hannah 6, 215f., 221f., 243, 260,
266
Aristoteles 135,
Augustinus 108, 161f., 220

B

- Bächle, Thomas 12f., 87, 143
Bacon, Francis 104, 110–113, 157, 260, 273
Bacon, Roger 109
Barth, Karl 19, 175
Bastian, Hans-Dieter 143
Beck, Ulrich 44, 190
Beer, Stafford 138
Benjamin, Walter 32, 264
Bense, Max 149
Bernal, John Desmond 236
Bertaux, Pierre 136f.
Bloch, Ernst 110, 130, 244, 253
Blumenberg, Hans 6, 9, 13, 18, 166, 265,
268, 274f.
Brin, David 57
Bonhoeffer, Friedrich 19
Bostrom, Nick 14, 41–43, 53, 57, 62,
179–188, 191, 194, 234, 237f.
Butler, Samuel 113–118, 167

C

- Cameron, James 47
Cassirer, Ernst 32
Castoriadis, Cornelius 38
Chalmers, David 227, 231
Clarke, Arthur 129, 228
Crutzen, Paul 194

D

- Darwin, Charles Robert 6, 114, 116f.,
127–129, 144 167f., 198, 255, 261f., 269
Descartes, René 140f., 150
Deutsch, Karl Wolfgang 136
Diderot, Denis 193
Dreyfus, Hubert 88f.

E

- Ellison, Harlan 185
Ettinger, Robert 242
Eurich, Claus 11f., 268

F

- Fiore, Joachim von 104, 108, 166f., 191
Fischer, Johannes 19, 46, 50
Freud, Sigmund 250
Frisch, Max 139
Fukuyama, Francis 52

G

- Ganascia, Jean-Gabriel 53, 79, 95, 190,
246f., 268–270
Gehlen, Arnold 32, 137, 149, 196

- Geraci, Robert 55–61, 249
- Gödel, Kurt 92–94
- Goethe, Johann Wolfgang von 107, 182
- Good, John Irving 14f., 85, 89, 96f., 168, 170, 174
- Günther, Gotthard 144, 148
- H**
- Haberer, Johanna 75
- Haeckel, Ernst 128
- Hagner, Michael 151
- Hanson, Robin 234
- Harari, Noah Yuval 42, 55, 177–180, 204–208, 212, 215, 239
- Haraway, Donna 189, 232, 256
- Hawking, Stephen 42
- Heidegger, Martin 31, 34, 83, 148
- Herbert, Frank 117
- Herder, Johann Gottfried 110
- Hobbes, Thomas 140
- Hoffmann, Ernst Theodor Wilhelm 120–124
- Horkheimer, Max 9f., 31
- Huxley, Julian 241f.
- I**
- Itskov, Dmitry 229
- J**
- Jonas, Hans 27, 45, 54, 109, 127, 149f., 219, 247f., 272–274
- Jonze, Spike 48
- Jung, Volker 26, 177
- K**
- Käsemann, Ernst 54
- Kepler, Johannes 119
- Krüger, Oliver 59f., 128, 193f., 242, 264, 268f.
- Kubrick, Stanley 47
- Kurzweil, Raymond 14, 40, 53f., 91, 93, 97, 152, 191, 227, 230–232, 237f., 243f., 250, 254
- L**
- La Mettries, Offray de 140
- Lacan, Jacques 38
- Lang, Fritz 130
- Leibniz, Gottfried Wilhelm 139f., 164, 166
- Lierfeld, Karl Johannes 180, 195f., 206, 209, 211–214
- Lilje, Hanns 31
- Lisberger, Steven 48
- Luther, Martin 175f., 199, 217–221, 224, 252
- M**
- Macho, Thomas 51f., 109f., 127, 271
- Marcuse, Herbert 31
- Marquard, Odo 58, 266
- McCarthy, John 68
- Meireis, Thorsten 21, 50
- Metzinger, Thomas 142, 225
- Miessner, Benjamin 132
- Minsky, Marvin 68
- Moltmann, Jürgen 199–201
- Moravec, Hans 40f., 52f., 69, 97, 130, 132, 153, 169, 186–188, 190, 192–195, 234–240, 246, 249

- More, Max 16, 43, 57, 62, 231–235, 239f.,
244
- Morus, Thomas 108f., 112
- Mostow, Jonathan 48
- Musk, Elon 42, 54
- N**
- Naumann, Friedrich 28f.
- Neumann, John von 15, 237
- Nida-Rümelin, Julian 47, 75f., 92, 223
- Nietzsche, Friedrich Wilhelm 261f., 266
- O**
- Otto, Rudolph 10
- P**
- Pannenberg, Wolfhart 166, 175, 199, 202,
218f., 222f., 236, 238, 245, 251–253,
276
- Paracelsus 106, 111
- Paulus 53, 135, 156, 218, 242
- Pfister, Wally 48
- Pias, Claus 132, 135, 138f., 150f., 261
- Pindar 135
- Popper, Karl Raimund 272, 274
- Precht, Richard David 58, 207
- du Prel, Carl 129
- Proyas, Alex 48
- R**
- Rapp, Hans Reinhardt 146
- Rendtorff, Trutz 17, 20, 34f., 37, 175f., 198f.,
203, 219, 221, 225f., 253, 272f., 276
- Ritschl, Albrecht 19
- Roger, Carl 5
- Rom, Hippolyt von 135
- Rothblatt, Martine 232f., 239
- Rothe, Richard 28f.
- S**
- Samuel, Arthur 68
- Sartre, Jean-Paul 213
- Scheler, Max 6, 94, 263
- Schleiermacher, Friedrich Wilhelm
26–30, 215
- Schmidhuber, Jürgen 41, 97, 130, 169
- Schmidt, Hermann 135
- Schrader, Maria 48
- Searle, John 13, 89–92, 97
- Shannon, Claude Elwood 131, 151
- Shaw, George Bernard 2
- Shelley, Mary 119f., 190
- Sölle, Dorothee 25f., 37
- Spallanzani, Lazzaro 123
- Spengler, Oswald 160, 170, 261, 270
- Spiekermann, Sarah 67, 80f., 94, 197, 268,
272
- Spielberg, Steven 48
- Swift, Jonathan 119
- T**
- Tegmark, Max 41f., 56f., 153, 179, 181f.,
186f., 191, 193, 196, 201, 206f., 249,
- Teilhard de Chardin, Pierre 104, 136, 149,
154–160, 166–171, 179, 234, 249, 253
- Tillich, Paul 33f., 266
- Tipler, Frank Jennings 238f., 245
- Troeltsch, Ernst 30, 275
- Turing, Alan 4, 14, 68, 86–89, 91, 96, 141,
178f., 189, 231
- U**
- Ulam, Stanislaw 15

V

Verhoeven, Paul 48

Verne, Jules-Gabriel 124

Vinge, Vernor 15f., 52, 57, 60, 62, 97, 178,
191, 234, 262

W

Walker, Mark 230

Walter, William 132

Watson, John Broadus 141

Weaver, Warren 131

Weber, Max 36

Weidenfeld, Nathalie 47, 76, 92, 223

Weizenbaum, Joseph 1–9, 53, 70, 79–81,
98, 129, 213, 267

Weizsäcker, Karl Friedrich von 32, 42

Wells, Herbert George 112

Wenders, Wim 257

Wiener, Norbert 131f., 135, 139f., 144–147,
151f., 167, 182, 228

Wirth, Mathias 242, 251

O

Otto, Rudolf 10

Ovid 6

Y

Yudkowsky, Eliezer 41, 60f., 180–185,
192f., 239

Sachregister

A

- Alchemie, alchemistisch 62, 106f., 111,
119, 121, 147, 245–250
- Allbetroffenheit 44f., 54, 190
- Android 4, 48f., 105, 119, 123, 129f.,
208–211, 233
- Angst 4, 6, 18, 51f., 57f., 74, 106, 109f.,
122, 127, 137, 201, 207, 211, 214, 251,
263, 271
- Anthropozän 57, 173, 194
- Antiquiertheit, antiquiert 154, 192, 196,
262, 276
- Alpha-Prädator 180, 264
- Apokalypse, Apokalyptiker, apokalyptisch
38, 44, 52–63, 103, 130, 154, 160, 166,
168f., 198, 201, 259, 261f., 275

B

- Bewusstsein 12, 14, 19, 40, 43, 46–49, 56,
59, 72, 74, 77, 85, 89–92., 97–100, 103,
115, 127, 144, 147, 149, 151–159,
166–168, 174–176, 197, 201, 217,
227–233, 246, 276, 275
- Black Boxes 79–83

C

- Christ, Christentum, Christenheit 9, 20,
29, 37, 54f., 108, 135, 146, 159f., 175,
179, 217, 223, 243, 251f., 275
- Christus 108, 135, 146, 159, 169, 175,
202, 217, 219, 222f., 246, 252–254, 256
- Cyborg 17, 21, 43, 151, 189, 227f., 232,
255f.

D

- Daten, Datenhunger, Datenschutz 9, 43,
51, 55, 65, 69–74, 77f., 82, 91f., 96–100,
153, 168, 174, 192, 203–208, 214, 216,
224f., 231f., 254
- Disruption, disruptiv 16, 40, 74–80, 177,
204, 269
- Digitalisierung 10, 19, 22, 26, 47, 55f.,
65–67, 249, 275

E

- Engel 110, 255–257, 267
- Eroberung, Eroberer 164, 180, 235, 267,
269
- Epistemologie, epistemologisch 21, 85,
99, 104, 110, 128, 131–153, 171, 173,
197, 248, 260
- Erlösung 21, 23, 57, 60, 62, 68, 238–251
- Ethik, Technikethik, Ethikrichtlinien
17–22, 25–37, 45f., 54, 77, 79, 83, 94f.,
146f., 217, 221, 226, 273, 276
- Evolution, evolutionär 12, 19, 40f., 56, 115,
118, 128f., 142, 152–156, 159, 166–169,
192, 197, 240f., 244, 246, 259, 270
- Emulation, Emulated Minds 98, 227–230,
234, 238, 245, 249f., 255f., 269
- Eschaton, Eschatologie, eschatologisch
54f., 59, 62, 159, 171, 173, 202, 218,
222, 241, 251, 256
- ### F
- Fatum, Fatalismus 51, 162, 170, 173, 190,
192, 198, 202, 232, 263, 268
- Fake Science 81, 270
- Frightful Five 73f., 77

- Fortschritt 4f., 8, 12, 15, 27–31, 39, 41, 51, 58f., 69, 82, 104, 110, 115f., 128, 152, 154, 160, 166–171, 173, 192–197, 205, 206, 209, 227, 239, 243, 245, 253, 262, 273, 275f.
- G**
- Gott, göttlich 6f, 19, 21, 25–31, 33, 35, 37, 56, 61, 66, 75, 105–107, 109, 112f., 119, 135, 143–148, 158–160, 162f., 167–169, 175–177, 189–191, 195, 197–202, 207, 212, 217–226, 239, 242, 246–248, 252–256, 262, 264–266, 276
- Geist, Geisteswissenschaften, Geistesgeschichte 10, 26, 46, 69, 74, 104, 108f., 125, 128, 132, 140, 143, 150, 154–159, 166–169, 176, 189–192, 212, 219, 228–233, 245–250, 253–256, 259
- Gehirn, Hirn 14, 85, 98f., 141–143, 182, 227, 229–231, 264
- Gleichsetzung 10, 12, 17, 22, 91, 103, 194, 227, 259
- Gnosis, gnostisch 62, 147, 245–247, 255
- Das Gute, Güter, Güterlehre 30f., 36f., 57, 59, 160, 164, 195, 199, 202, 215, 217, 221, 225f., 276f.
- H**
- Das Heilige, (Ent-)Heiligung, heilig 10, 46, 108, 118, 146, 176, 198, 205f., 253, 266, 275
- Humanum, Humanität, Humanismus 26–35, 47, 76, 137, 177, 205, 246, 249, 267, 270, 277
- I**
- Identität, Identitätsfrage, Identitätsverlust, Identifikation 3, 23, 51, 174f., 203, 209, 211, 227, 231, 237, 241f., 247, 250, 254
- Imagination, Imaginäre, imaginiert 6, 7, 14, 35, 38f., 45, 60, 63, 108, 174, 178, 190, 195, 203, 206, 208, 227, 232, 235, 260
- Institution, Institutionalisiere 38–40, 73, 77, 134, 173, 274
- K**
- Katastrophe, katastrophal 7, 22, 42, 50–59, 77, 104, 114, 126f., 136, 154, 158, 160, 182, 201, 272
- Kinder, Mind Children 40f., 52, 121, 127, 130, 170, 186–189, 194f., 202, 227, 235, 249–253, 269
- Konfrontation 23, 47, 77, 134, 174, 178, 186, 188, 190, 195, 203, 214, 221, 227, 252
- Konvergenz, konvergieren 13, 16–23, 52, 61f., 103f. 128, 133, 151, 153, 158, 161, 164–171, 175, 177, 189f., 194, 196–198, 201, 217, 227, 239, 245, 249, 259f., 263, 265, 267, 268, 271, 275
- Körper, körperlich 12, 48, 56, 57, 60, 94, 116, 123, 129, 140–143, 149, 151, 155, 183, 189, 207, 228f., 231–233, 236, 245, 255
- Kultur, technische Kultur, Kulturtheorie, Kulturkritik, Popkultur 19f., 25, 29, 30–37, 39, 45, 75, 103, 136, 137, 148, 151, 153, 160, 187, 192f., 197, 199, 202, 217, 219, 223, 228, 232, 243, 259f., 269–277

- Kontrolle, Selbstkontrolle, Kontrollverlust 198, 201, 220, 228, 236, 240–250, 257, 260–262
- 4, 48, 65f., 78, 100, 104, 111f., 132, 178f., 182, 189f., 192, 201, 215, 239, 248
- Kybernetik, kybernetisch 8, 22, 68, 104, 131–154, 168, 173, 182, 203, 205, 216f., 222, 225, 228, 230, 238, 248, 260
- L**
- Lebensführung 46, 100, 175f., 204, 221, 226, 253, 275–277
- Liebe 6, 21, 46, 48, 121f., 124, 160, 210f., 217f., 222–226, 262, 276
- Logos 6, 9, 33f., 44, 241, 244, 246, 265
- M**
- Machtübergabe, Machtwechsel 4, 186
- Magie, Magier, magisch 6, 8, 54, 62, 105, 120f., 123f., 129, 162, 247f., 250f.
- Maschinelles Lernen 54, 65, 70, 72, 82, 95f., 174, 266, 277
- Mind Upload 14, 21, 227f., 232–254
- Morphological Freedom 231, 239
- Muster 7, 23, 45–47, 63, 65, 67, 70, 72, 84, 89, 98–100, 115, 152, 166, 173, 244, 264, 274f.
- Mythos, Mythologisierung 2, 6–21, 35, 38, 40, 44f., 52f., 75, 87, 95, 119, 143, 145, 152, 166, 168, 190, 197f., 202, 242, 245–252, 259, 260–277
- N**
- Natur, Naturkräfte, Naturbeherrschung, Naturwissenschaften 7, 10f., 22, 26–30, 34, 36f., 50, 58f., 67, 96, 98, 103, 106, 108–113, 119, 123, 127, 129, 134, 137, 140–157, 166–169, 173, 189, 190, 195, 198, 201, 220, 228, 236, 240–250, 257, 260–262
- Narrative, narrativ 13, 45–52, 115, 142, 271, 275
- Neuronale Netze 11, 41, 69–71, 81–83, 98, 227
- O**
- Omega 154–169, 179, 182, 253
- P**
- Performanz, Performance 3f., 88, 91
- Prophetie, prophetisch 42, 58, 62, 69, 155, 269
- Plausibilität 83, 94f., 129, 185
- Politik 26f., 38–40, 72–76, 118, 130, 133, 135–138, 185, 187, 203, 205, 272
- Polis 76, 101, 130, 203–224
- Puritanismus 75
- R**
- Regierung, Regierungsapparate 49, 76, 135–138, 151
- Regulierung 76–78, 115, 268
- Risiko, Risikobewertung 44f., 78, 184, 190, 201
- S**
- Scham 5f., 161–163, 240, 243, 263, Science-Fiction 32, 47, 50, 57, 103, 105, 117, 119, 124, 129, 174, 228, 271
- Schöpfer, Schöpfung, Geschöpf 4, 7, 20, 23, 29f., 37, 50, 79, 101, 103, 105–107, 113, 119, 123, 128, 145f., 158, 176, 181, 188–202, 239, 248, 253, 259
-

- Schuld, Verschuldung 81–83, 98, 143,
162, 195–200, 220–222, 226, 251, 265f.
- Singularität, Singularitans 15f., 40f.,
52–54, 59–61, 95, 103, 152, 170,
173f., 178, 184, 191f., 194, 205–208,
228–231, 238–249, 259, 262–271
- Silicon-Valley 74–77, 196, 205, 207, 229
- Simulation, Simulation Game 3f., 7, 72,
88, 91f., 96, 111, 179, 238
- T**
- Täuschung, Selbsttäuschung, Vertauschung
3, 8, 103, 125, 161, 210, 224, 269
- Technikfolgenabschätzung 17f., 38, 54,
62, 174
- Theologie 17, 20, 25–33, 54, 141, 143–146,
175, 177, 197, 202, 214, 217, 220, 251,
267, 275
- Transhumanismus, Transhumanist 43, 57,
227, 229, 232f., 240–248, 251f., 260
- Transzendenz, transzendieren 35, 56,
83f., 154, 167, 177, 188, 193, 238,
243–249, 253–256, 271
- U**
- Universum 41, 49, 69, 118, 153f., 158f.,
164, 169, 171, 191–193, 201, 214, 228,
235–239, 244f., 249f., 255f.
- Überwältigung, Überforderung 18, 37,
50–52, 56, 113, 120, 128, 175, 194, 199,
214, 259–262, 271
- Unternehmen 4, 21, 54, 73–83, 95, 205,
209, 229, 268, 271, 273
- Unterwerfung 108, 113, 122, 127, 153,
255, 264
- Utopie, Utopismus, Utopia, utopisch 22f.,
41, 56, 58, 74, 76, 103f., 10–114, 116,
119, 127–130, 151, 153, 157, 166, 199,
206f., 219, 232, 234, 238, 253, 262,
272–274
- V**
- Verantwortung, Verantwortungsabgabe,
Verantwortungslosigkeit, verantwortungsvoll 20, 27, 32, 45, 49, 52, 77,
100, 135, 143, 146f., 177, 188, 194–202,
219, 223, 245, 259, 267f., 272, 273, 276
- Vernunft, vernünftig 10, 29–31, 46, 60,
83, 113, 126, 213
- Vertretung 23, 174, 203f., 208, 214–222,
227, 274
- W**
- Wahrheit 10, 28, 34, 4, 46, 105, 143, 171,
269
- Weltmodell 92f., 142
- Weltflucht 30, 217, 245
- Widerstand, das Widerständige 10, 51, 60,
84, 186, 205, 262
- Wirklichkeit 6, 10, 15, 18, 39, 46f., 92–94,
104, 107f., 116, 138, 142, 150–156, 166,
168, 171, 173, 182, 203, 213, 215f., 225f.,
244, 248–250, 254f., 259–262, 267, 271,
274, 276
- Wirtschaft, Ökonomie, wirtschaftlich,
ökonomisch 11, 22, 38, 40, 50, 63, 67,
72–83, 94, 138, 173, 179, 181f., 186,
196, 203, 205, 228, 234f., 253, 259, 269

Wissenschaft, Wissenschaftler 3f., 10, 12,
18, 28, 32, 36–44, 48, 67f., 81, 86, 104,
106–108, 113, 120–124, 129–133, 141,
143f., 146, 147, 149, 151f., 157, 163,
166, 177, 179, 189–192, 202, 228, 232,
234, 239, 245, 254, 259, 264, 267f.,
270f., 274

Künstlicher Intelligenz wird immer mehr zugetraut. Unter dem Deckmantel nüchterner Technikfolgenabschätzung sind in den letzten Jahrzehnten populäre Mythen entstanden, die die Menschheit an einem historischen Wendepunkt im Umgang mit ihrer Technik vermuten. Viele der Erzählungen schauen gebannt in die nahe Zukunft und befürchten den drohenden Machtwechsel zwischen Mensch und intelligenter Maschine. Visionen einer von ihrer Technik überforderten Menschheit mischen sich dabei mit einer apokalyptischen Rhetorik, nach der die Menschheit gar nicht anders könne, als ihre effizienteren Nachfolger zu konstruieren. Technik, die ehemals Ausdruck menschlicher Souveränität war, verkehrt sich zum neuen Subjekt der Geschichte. Diese Untersuchung macht die Auslieferung des Menschen an die Zukunft der Künstlichen Intelligenz als kontingenten Mythos durchsichtig. Im Abgleich der Zerrbilder menschlicher Freiheit, wie sie die Mythen der KI hervorbringen, wirbt der Autor für einen pragmatischen Umgang mit Technik und Innovation auf der Grundlage eines christlichen Freiheitsverständnisses, das auf Bewältigung anstatt Überwältigung setzt.

Die Dissertation wurde mit dem Wissenschaftspreis 2023 der Hanss-Lilje-Stiftung ausgezeichnet.

Yannick Schlote ist evangelischer Theologe und seit 2019 wissenschaftlicher Mitarbeiter am Lehrstuhl für Systematische Theologie und Ethik von Prof. Dr. Reiner Anselm (LMU München) und dem Zentrum für Technik-Theologie-Naturwissenschaften (TTN). Dort betreut er die ethische Begleitforschung des Diagnostikprojekts Bavarian Genomes. Seine Forschungsschwerpunkte bilden die Technik- und Medizinethik.

48,00 €
ISBN 978-3-487-16350-5



www.olms.de