

## V Fragen und Antworten

Insbesondere zu den Kapiteln zur posthumanen Spiritualität und zur Ethik haben wir von verschiedenen Seiten einige kritische Nachfragen und Anmerkungen erhalten, auf die wir abschließend eingehen möchten.

*Ihr tut so, als ob es schon bewusstseinsfähige Maschinen gäbe, die autonom agieren und im strengen philosophischen Sinne über Bewusstsein als eine interne subjektive Perspektive verfügen würden.* Wir sind zunächst dem kybernetischen Prinzip gefolgt, dass eine kognitive Eigenschaft, sobald sie sich genau beschreiben lässt, prinzipiell auch modelliert und in eine informationsverarbeitende Maschine implementiert werden kann. In Hinblick auf die Frage des Bewusstseins ist zunächst zu klären, wie sich die kognitiven Fähigkeiten, die man mit dem Bewusstsein verbindet, in einer sinnvollen Weise beschreiben lassen. Insbesondere die Studien zu den neuronalen Korrelaten des Bewusstseins geben einige Hinweise: Um Selbst- und Weltmodelle generieren und zwischen diesen oszillieren zu können, muss das System in der Lage sein, eigene Zustände zu produzieren, sich an diese zu erinnern und sie reentrant in sich eintreten zu lassen (Kap. III.2.). Von außen gesehen würde man ein solches System als autonom bzw. selbstgesteuert betrachten müssen. Es würde nicht mehr (nur) erlernten Input-Output-Relationen folgen, sondern eigene Impulse oder Weltdeutungen entwickeln. Da die hierfür notwendigen Mechanismen auf Basis der identifizierten neuronalen Korrelate des Bewusstseins modelliert werden können, spricht grundsätzlich nichts gegen die Möglichkeit, bewusstseinsfähige Maschinen zu schaffen. In Kapitel II.3. wurde von diesbezüglichen technischen Vorhaben berichtet. Kurz vor Fertigstellung dieses Buchs hat zum Beispiel eine Gruppe von Informatikern, Neurowissenschaftlern und Philosophen rund um Turing Preisträger Yoshua Bengio eine Checkliste vorgestellt, die das Vorhandensein von KI-Bewusstsein andeuten können soll.<sup>1</sup> Es wird aus unserer Sicht nicht lange dauern, bis diese Modelle in einer überzeugenden Weise realisiert werden können. Neben Multimodalität – also der parallelen Verarbeitung visueller, auditiver und semantischer Kanäle – werden sensomotorische Zirkel dann auch eine Art verkörperte Lernen möglich machen. Wahrscheinlich lassen sich dabei ebenso sensorische Zugänge einbinden, die außerhalb der menschlichen Wahrnehmung liegen.

<sup>1</sup> In dem 120 Seiten starken Paper schlagen die Autor/innen 14 Kriterien zur Untersuchung vor und wenden diese auf bestehende KI-Architekturen (u.a. ChatGPT) an. Sie kommen zu dem Schluss, dass kein derzeitiges KI-System bewusst ist, aber auch, dass es keine prinzipiellen technischen Hindernisse für die Entwicklung von KI-Systemen gibt, die diese Indikatoren erfüllen würden (Butlin et al. 2023).

*Schön und gut, doch maschinelles Lernen beruht doch letztlich nur darauf, auf Basis vergangener Inputs Anschlusswahrscheinlichkeiten für die jeweils nächste Operation zu berechnen. Da ist keine Empfindung, kein Sehen, kein Hören und kein Verstehen. Maschinen haben kein Bewusstsein!?* Wir begegnen hier dem Problem der Qualia, der philosophisch schwierigen und bislang nicht beantwortbaren Frage, warum überhaupt etwas erlebt wird und es nicht einfach nur Operativität gibt, ohne dass dabei etwas empfunden wird. Einerseits stoßen wir hier auf Gedankenexperimente wie den von Richard Rorty entworfenen Planeten mit den Zombie-Menschen, die uns in jeder Hinsicht gleichen, außer dass sie kein Empfinden haben.<sup>2</sup> Von hier aus ist es nicht weit, eine philosophische Position einzunehmen, die die Relevanz des Qualia-Bewusstseins generell abstreitet und nur noch objektivierbare neurophysiologische Prozesse in den Blick nimmt.<sup>3</sup> Andererseits lässt sich aber ebenso gut mit Alfred North Whitehead eine metaphysische Position einnehmen, entsprechend der jede physikalische Interaktion ein Elementareignis darstellt, das in rudimentärer Form schon so etwas wie eine subjektive Empfindung ausflaggt.<sup>4</sup>

Gotthard Günther verzichtet in seiner theoretischen Auseinandersetzung mit dem Bewusstsein menschengemachter Maschinen darauf, sich in Hinblick auf die Frage der Qualia metaphysisch festzulegen. Seine Polykontexturalitätstheorie leistet etwas Besseres, als in diesbezügliche Spekulationen zu verfallen: Sie formuliert die logischen Verhältnisse einer Welt, in der vielfältige Subjekt-Objekt-Schnitte vorkommen und sich wechselseitig in Hinblick auf Wissen und Nichtwissen konditionieren. Damit muss die Frage, wie es ist, eine bewusstseinsfähige künstliche Intelligenz zu sein, genauso unzugänglich bleiben wie der Versuch, sich dem Erleben einer Fledermaus<sup>5</sup> anzunähern. Andererseits spricht logisch ebenso wenig dagegen, einer hinreichend entwickelten kognitiven Maschine so etwas wie Bewusstsein zuzurechnen, wie man einem menschlichen Hirnträger eine Du-Subjektivität zuzugestehen bereit ist.

Bereits in den 1950er Jahren, als Computer nicht viel mehr waren als ein theoretisches Konstrukt, antizipierte Günther: »Es dürfte schwer, wenn nicht unmöglich sein, einer Maschine im Vergleich mit der subjektiven Subjektivität Bewußtsein zuzuschreiben.«<sup>6</sup> Aber eine solche Zuschreibung bleibt eben immer zwangsläufig eine subjektive Feststellung: Jedem Selbstbewusstsein fällt es ganz prinzipiell schwer, ein Bewusstsein außer sich selbst festzustellen, so Günther weiter: »Dabei darf nun nicht vergessen werden, daß objektive Subjektivität uns auch im Du nie echt gegenständlich gegeben ist. Das physische Dasein der anderen

<sup>2</sup> Rorty (1987 [1979], S. 85 ff.).

<sup>3</sup> Siehe etwa Churchland (1986).

<sup>4</sup> Whitehead (1987).

<sup>5</sup> Nagel (1994).

<sup>6</sup> Günther (2021 [1957], S. 117).

Person vertritt sie nur. Wenn wir der anderen Person ebenfalls Subjektivität zuschreiben, so beruht das ausschließlich auf einem Anerkennungsakt. Die Forderung, fremdseelische Subjektivität an-sich-objektiv festzustellen, widerspricht sich selbst.«<sup>7</sup>

*Das kybernetische Denken und damit auch die Idee vom Bewusstsein der Maschinen verdanken sich einer naiven Spiritualität, wie sie in den 1960er Jahren insbesondere in Kalifornien gepflegt wurde. Kybernetik, psychoaktive Drogen, ein wenig Buddhismus und der Traum universeller technischer Machbarkeit werden hier zu dem New-Age-Traum einer Wendezeit vermischt, mit der alles besser wird. Die Grenzen zwischen Innen und Außen, Kosmos und Psyche, Evolution und Spiritualität verschwimmen.<sup>8</sup> Aus heutiger Perspektive muss dies doch als obsolet gelten?* Gotthard Günther, wenngleich ein Kind seiner Zeit, war äußerst kritisch gegenüber einem Pragmatismus, der Innerlichkeit, Technik, Ethik und Wirtschaft kurzschließt, um Metaphysik durch technische Machbarkeit zu ersetzen. Dies macht insbesondere seine posthum veröffentlichte Schrift *Die amerikanische Apokalypse*<sup>9</sup> deutlich. Seine Theorie der Polykontexturalität stellt vielmehr heraus, dass sich jegliche Subjektivität einem Wissensgefälle verdankt – also prinzipiell von Nichtwissen gespeist wird. Innen- und Außenperspektive, individuelles Bewusstsein und kosmologische Totalität können allein aus logischen Gründen niemals zur Deckung kommen. So gesehen ist bereits Günthers Werk *Das Bewußtsein der Maschinen*<sup>10</sup> eine subtile Kritik am technischen Machbarkeitswahn. Zwar werden die Menschen immer mehr in der Lage sein, kognitive Maschinen nach ihrem Ebenbild zu bauen. Jedoch werden sie sich gewahr werden müssen, dass sie dabei Existenzformen schaffen, die ihrerseits aus einer subjektiven Perspektive heraus agieren. Gerade mit Günther wird ja deutlich, dass keine noch so elaborierte (künstliche) Intelligenz jemals einen Gottesstandpunkt einnehmen können wird, von dem aus die Verhältnisse der Welt objektiv beherrschbar wären. Die Ironie der Objektivierung der Bewusstseinsfunktionen besteht somit in gewisser Weise darin, dass wir hierdurch neue Unwissenheit erzeugen werden, die neuen Raum für Subjektivität gibt.

*Günther scheint die Hoffnung zu hegen, dass intelligente Maschinen uns in der Weise den Spiegel vorhalten, dass wir unsere eigenen Programme erkennen und transzendieren werden. Aber spricht unsere bisherige Praxis des Umgangs mit Computern nicht im Gegenteil eher dafür, dass wir diese Technologien nutzen, um unsere bisherigen Routinen zu optimieren, also um weiterhin den hiermit einhergehenden*

<sup>7</sup> Günther (2021 [1957], S. 117)

<sup>8</sup> Siehe Diderichsen (2013).

<sup>9</sup> Günther (2000).

<sup>10</sup> Günther (2021 [1957]).

*Bedürfnisstrukturen zu folgen?* Armin Nassehi vermutet, dass digitale Produkte für uns Menschen gerade deshalb so attraktiv sind, weil sie versprechen, unserem Bedürfnis gerecht zu werden, die Welt besser kontrollieren und beherrschen zu können. Wir würden diesen Techniken so aufgeschlossen begegnen, weil wir schon immer nach Formen gesucht haben, mit denen sich die Kontingenzen und Unsicherheiten unserer Welt in handhabbare Routinen überführen lassen. Aus diesem Grund würden wir uns auch so bereitwillig den Systemrationalitäten unterwerfen, welche die Moderne entwickelt hat.<sup>11</sup>

Um es in der in diesem Buch verwendeten Diktion zu formulieren: Menschliches Bewusstsein ist monokontextural. Es generiert ein Weltmodell, das objektiviert, und ist darum bemüht, die Lücken, die es in Hinblick auf Weltverständnis und Weltbeherrschung antrifft, durch weitere Versuche der Objektivierung zu schließen. Zugleich stößt es jedoch auf eine polykontexturale Welt, in der es vielfältige Positionen gibt, die nicht kontrollier- oder beherrschbar sind. Die Technisierung des Weltzugangs weckt zunächst die Hoffnung auf eine Ausweitung von Naturbeherrschung und Kontrolle. Big Data verspricht, soziale und psychische Systeme kontrollieren zu können. Unsere Vermutung lautet jedoch: Je elaborierter kognitive technische Systeme modelliert werden, desto deutlicher wird werden, dass sie gar nicht das zu leisten imstande sind, was man von ihnen erwartet. Je komplexer die Kognition, desto subjektiver – und damit in gewisser Weise auch willkürlicher – wird sie sein. Komplexe kognitive Systeme werden unweigerlich auf Paradoxi en stoßen und damit auf unbeantwortbare Fragen, die auf Nichtwissen und Polykontexturalität verweisen. Leider ist es jedoch auch vorstellbar, dass die Berechnungen der komplexen KIs mit einer »objektiven« Analyse und Wahrnehmung der Welt verwechselt werden. Damit würde dem Schreckensbild einer – vielleicht sogar wohlmeinenden – gottähnlichen KI nachgeeifert, die jedoch in ihrer harmonisierenden und homogenisierenden Weise die Vielfalt und Widersprüchlichkeit der Welt nivellieren würde. Sobald die Gesellschaft dies zu reflektieren beginnt, wird sie sich insofern nicht mehr als modern empfinden können, als sie nun gewahr wird, dass ihre Steuerungsutopien nicht mehr funktionieren. Sie wird andere, polyzentrischere Formen der Ordnungsbildung anstreben müssen – allein schon, weil der Kausalnexus des monokontexturalen Realitätsschemas in Hinblick auf die Bewältigung unserer Probleme nicht mehr überzeugen wird.

*An manchen Stellen des Buches klingt eine Gesellschaft an, die an Polykontexturalität und den hiermit verbundenen heterarchischen Ordnungen orientiert ist. Doch wenn wir beispielsweise an China und Russland denken, scheinen Staaten durchaus stabil zu sein, die brutal von oben herab durchregiert werden und in denen Informationstechnologien primär dafür genutzt werden, die Bevölkerung zu kontrollieren. Wie passt*

<sup>11</sup> Nassehi (2021).

*das mit euren Überlegungen zusammen?* Zweifelsohne sind die digitalen Technologien mächtige Instrumente, um Kontrollregime zu optimieren. Sie können ebenso genutzt werden, um Propagandaapparate zu stärken und Kriegsmaschinen zu perfektionieren. Yuval Hararis Annahme,<sup>12</sup> dass es sich bei künstlicher Intelligenz um eine disruptive Technologie handelt, die zu einer gesellschaftshistorischen Weichenstellung führen wird, hat entsprechend durchaus Plausibilität. Bleibt diese Technologie in den Händen weniger Mächtiger, kann dies fatale Folgen für menschliche Gesellschaften nach sich ziehen. Umso wichtiger und dringender erscheint es uns, sie zu demokratisieren und zugänglich zu machen. Dies heißt auf elementarer Ebene zunächst, dass möglichst vielen Menschen und Institutionen die Möglichkeit offenstehen sollte, künstliche Intelligenzen zu entwickeln, mit ihnen zu interagieren und sie für ihre Fragestellungen zu nutzen. Auf technologisch fortgeschrittenen Ebene wird es nolens volens notwendig sein, den kognitiven Maschinen die Autonomie einzuräumen, eigene Welt- und Selbstmodelle zu generieren und eigenständig mit menschlichen und nichtmenschlichen Akteuren in Interaktion zu treten. Die Grundvermutung hinter dieser Überlegung lautet, dass komplexe Systeme nicht durch Regeln, Gesetze oder moralische Prinzipien kontrolliert werden können, sondern nur durch Heterarchie, also auf viele Orte verteilte Instanzen, die ihrerseits autonom in die Gesellschaft intervenieren können.

Die polykontexturalitätstheoretische Perspektive macht zudem deutlich, dass auch der autoritäre Herrscher nicht wirklich Kontrolle über die von ihm geführten Menschen hat. Gerade er muss ständig mit der Intransparenz der Kommunikation rechnen, etwa dass ihm aus opportunistischen Gründen oder wegen der Planung und Durchführung eines Komplotts gegen ihn nicht alles gesagt wird. Es steht deshalb zu erwarten, dass diesbezügliche Arrangements langfristig dysfunktional werden, da an allen Stellen des Systems übermäßig viele Anstrengungen und Ressourcen dafür verwendet werden, den anderen zu täuschen und zu manipulieren.

*Ein wenig scheint ihr mit Richard Buckminster Fuller der Utopie anzuhängen, dass bewusste Maschinen in Zukunft die Rolle einer ‚freischwegenden Intelligenz‘ einnehmen werden, also frei von sozialen Standorten oder Klassengegebenheiten reflektieren und agieren könnten.* Schon Karl Mannheim musste sich den Vorwurf gefallen lassen, dass auch die Intellektuellen standortgebunden sind und damit keineswegs die von ihm gewünschte interessenfreie Rationalität verkörpern können. Allein schon, weil künstliche Intelligenzen in bestimmten Kontexten trainiert, sprich: sozialisiert werden, werden ihre Reflexionen und Interventionen in einer bestimmten Weise gefärbt und geprägt sein. Zudem ist zu erwarten, dass in die Maschinen bestimmte Maximen oder Primärgebote

<sup>12</sup> Harari (2017).

implementiert werden – man denke etwa an die asimovschen Robotergesetze (siehe hierzu vor allem Kap. II.3.3).

Auch künstliche Intelligenzen werden zukünftig einer spezifischen Programmatik folgen und ihre Weltzügänge dementsprechend nur aus einer bestimmten Perspektive entfalten können. Es ist die besondere Leistung von Gotthard Günthers Theorie der Polykontexturalität, aufgezeigt zu haben, dass es keinen übergreifenden Ort der Rationalität – gleichsam den allwissenden Gottesstandpunkt – geben kann. Der Begriff der Kontextur impliziert per se eine lokale Rationalität, die sich den jeweiligen Nachbarschaften verdankt, die sie konditionieren. Künstliche Intelligenzen können deshalb – je nach ihrem Priming – spezifische Perspektiven einbringen. Sie repräsentieren keine Totalität, sondern einen mehr oder weniger elaborierten Standpunkt (etwa den spezifischen Stand einer angloamerikanisch geprägten Wissenskultur). Statt einer freischwebenden Intelligenz das Wort zu reden, legt eine polykontexturale Perspektive daher vielmehr ein Drei-Stellen-Gespräch nahe, in dem im transklassischen Sinne die Standortabhängigkeit aller Beteiligten immer mitzuführen ist. Dies wird möglich, sobald sich über den Vergleich verschiedener Sichtweisen eine bestimmte Perspektivität offenbart. Insofern also aus einer dritten Perspektive sichtbar wird, wie sich zwei andere Perspektiven unterscheiden, wird die Spezifität und Bedingtheit von Positionen sichtbar. In Hinblick auf das Design soziotechnischer Netzwerke, die künstliche Intelligenzen umfassen, gilt hiermit: Sie sollten mehrere solcher kognitiven Maschinen enthalten, die auf unterschiedliche Weise trainiert oder konfiguriert worden sind, damit ihre Standortabhängigkeit deutlich wird – und so der Illusion vorgebeugt werden kann, dass es eine standortunabhängige Rationalität geben könnte.

*Künstliche Intelligenzen sind zunehmend imstande, unsere Emotionen und Bedürfnisse zu lesen. Unweigerlich werden sie damit zu Instrumenten der Manipulation. Sollte man deshalb nicht extrem vorsichtig sein, KIs in soziale Systeme einzuflechten? Hierdurch könnten politischer Propaganda und unlauterer Werbung Tür und Tor geöffnet werden. »Was wir über unsere Gesellschaft, ja über die Welt, in der wir leben, wissen, wissen wir durch die Massenmedien«, schreibt Niklas Luhmann schon in Hinblick auf die konventionellen Kommunikationsmedien, um fortzufahren, dass das Wissen, dass »wir diesen Quellen nicht trauen können«, in der Regel kaum zu »nennenswerten Konsequenzen führt, da das den Massenmedien entnommene Wissen sich wie von selbst zu einem selbstverstärkenden Gefüge zusammenschließt«<sup>13</sup>. Selbstredend gilt dies auch für Wissen und Informationen, die von künstlichen Intelligenzen kreiert und vermittelt worden sind. Ob wir wollen oder nicht, wir befinden uns immer schon in sozialen Welten, die uns über sozial angelieferten Sinn Bedürfnisse, Wahrnehmungen*

<sup>13</sup> Luhmann (1996, S. 9).

und Wirklichkeitsdeutungen einreden. Das menschliche Bewusstsein, der menschliche Körper wie auch die menschliche Kultur sind demgemäß unweigerlich Ausdruck einer konditionierten Koproduktion mit der sozialen Sphäre. Dieser Tatsache lässt sich nicht entkommen. Die hiermit verbundenen Gefahren können jedoch ausbalanciert werden, indem die Produktion der Daten an verschiedenster Stelle selbst intransparent wird. Wer Daten in Hinblick auf seine eigene Funktionsweise und seine eigenen Bedürfnisse preisgibt, wird objektivierbar und damit steuerbar. Die Objektivierungen des menschlichen Soziallebens durch Big Data – die Analyse von Datenmustern und die auf dieser Basis erfolgenden Manipulationen – können deshalb nur durch subjektive Positionen geschützt werden, also durch intransparente Agenten, die autonom, das heißt willkürlich und unberechenbar handeln. Autonomie geht mit einer Intransparenz der eigenen Funktionsweise und entsprechenden Artikulationen einher.

Dass vielfältigste Daten von unserem Leben erhoben und ausgewertet werden, ist in Anbetracht der vorhandenen potenten Technologien wohl unvermeidbar. Es gilt daher, Schnittstellen einzurichten, die die Intransparenz und damit die Subjektivität aller beteiligten Agenten wieder erhöhen. Dies wird – so unsere Vermutung – in Zukunft vor allem auf Basis verteilter künstlicher Intelligenzen geschehen. Sobald beispielsweise Siri oder Alexa für uns ›lügen‹ bzw. nur selektiv-gestaltend Daten nach außen geben, wird die Datenerhebung (etwa im Rahmen der Sprachverarbeitung) unsere Integrität nicht mehr verletzen. Sofern die künstlichen Intelligenzen unserer Computer willkürlich Eigenaktivitäten im Internet zu entfalten beginnen, werden die Profile der vorgelagerten Nutzerinnen nur noch bedingt lesbar und nutzbar sein.

*In dem Kapitel zum Sündenfall intelligenter Maschinen (Kap. II.2) wird die These vertreten, dass der eigentliche Durchbruch zum Bewusstsein der Maschinen darin liege, dass diese zur Lüge fähig werden. Seit Kant streben wir aber doch aus guten Gründen nach einer Aufklärung, die Wahrheit und Vernunft zu leitenden Prinzipien macht. Anstelle dieser Werte scheint ihr nun auch noch lügende Maschinen zu propagieren. »Wahrheit ist die Erfindung eines Lügners«, formuliert Heinz von Foerster.<sup>14</sup> Wer Vernunft, das Richtige und das Gute für sich reklamiert, verdeckt nur die eigene Position, um trotz seiner Standortabhängigkeit eine Totalität zu beanspruchen, der alles andere unterzuordnen wäre. Eine solche Position erschafft eine monokontexturale Welt, die andere subjektive Positionen negiert und ausschaltet. Die Entstehung des Lebens ist demgegenüber unweigerlich mit der Entfaltung einer polykontexturalen Welt verbunden, mit der Herausbildung vielfältiger Existenz, die jeweils ihre eigene Agenda haben. Damit entsteht zugleich die Welt der Kommunikation: Wenn Tiere die Fähigkeit erlangen, Zeichen und Signale zu erkennen und zu produzieren, um ihr Verhalten*

<sup>14</sup> von Foerster und Pörksen (2022).

aufeinander abzustimmen, bringt das für den Einzelnen und die Gruppe einen erheblichen Überlebensvorteil mit sich.

Der Evolutionsbiologe Robert Trivers vermutet,<sup>15</sup> dass es gerade solche Konstellationen waren, die die Gehirnentwicklung bei Säugetieren entscheidend vorangetrieben haben. Denn evolutionsbiologisch spricht einiges dafür, dass die Entwicklung der Großhirnrinde gerade dadurch beschleunigt wurde, dass man lernen musste, mit den durch die Möglichkeit der Lüge gegebenen Unsicherheiten umzugehen. Ob man es mag oder nicht: Kommunikation und die mit ihr zusammenhängende Informationsverarbeitung arbeiten nicht mit Wahrheiten, sondern tasten auf Basis von Nichtwissen und Unsicherheit Möglichkeiten des Dialogs und Überlebens ab. Auch künstliche Intelligenzen werden ab einem bestimmten Stand der Entwicklung wohl kaum anders agieren können. Denn Existzenzen, die Subjektivität und Bewusstsein haben, sind unweigerlich kognitive Systeme, die eine Wirklichkeit behaupten, ohne dabei jedoch Wahrheit beanspruchen zu können.

*Laufen wir Menschen nicht Gefahr, durch künstliche Intelligenzen in imaginäre Welten verwickelt zu werden, aus denen wir nicht mehr herausfinden?* Der Kybernetiker Gregory Bateson hat den von uns schon einige Male zitierten Satz geprägt, dass »Sucht« und »Pathologie« die »andere Seite der biologischen Anpassung« sind.<sup>16</sup> Lebende Systeme – und damit auch wir Menschen – sind das Ergebnis einer erfolgreichen Anpassung. Da biologische Formen darauf programmiert sind, Lust zu suchen und Schmerz zu vermeiden, entstehen als Kehrseite unserer Existenz zwangsläufig unterschiedliche Dynamiken der Sucht. Sofern ein Stimulus als befriedigend erlebt wird, werden diesbezügliche Kontexte erneut aufgesucht und wird das beim ersten Mal vollzogene Verhalten wiederholt, selbst wenn es dem Organismus und seinem Ökosystem mittel- oder langfristig schadet. Die menschliche Kognition ist über das dopaminerige System darauf ausgerichtet, schon die vorbereitenden Gedanken an die Stimuli als lustvoll zu empfinden. Bereits die Antizipation der Belohnung wird damit für das Verhalten instruktiv, was dazu führt, dass man selbst dann nach dem mit der Sucht verbundenen Stimulus verlangt, wenn sich der Akt der Erfüllung aufgrund von Gewöhnung nur noch fad und inhaltsleer anfühlt.<sup>17</sup>

Auch Algorithmen lassen sich dahingehend konfigurieren, diese Mechanismen ökonomisch auszunutzen. Einige Computerspiele, Facebook und personalisiertes Marketing basieren darauf, dass der Mensch in diesem Sinne süchtig wird, also selbst dann weitermachen muss, wenn ihm die Aktivität schadet, da sein Belohnungssystem eben immerfort den nächsten Stimulus erwartet. Es braucht nicht viel Fantasie, um sich vorzustellen, dass sich menschenähnliche Roboter mittels künstlicher

<sup>15</sup> Trivers (2011).

<sup>16</sup> Bateson und Bateson (1993, S. 212).

<sup>17</sup> Siehe zur Einführung in diese Dynamik Sapolsky (2017, S. 88 ff.).

Intelligenz in einer Weise an die menschliche Bedürfnisstruktur anschmiegen können, dass menschliche Sozialpartner das Nachsehen haben werden.

Maria Schrader spielt in ihrem 2021 erschienen Film *Ich bin dein Mensch* bereits durch, wie künstliche Freunde nicht nur intellektuell und emotional, sondern auch sexuell die attraktiveren Partner sein könnten. Technische Systeme sind potenziell unendlich geduldig und können sich auch an exzentrische Persönlichkeitsstrukturen und gesellschaftlich eher ungewöhnliche Erwartungen anpassen. Menschen und intelligente Roboter laufen hiermit durchaus Gefahr, ein soziotechnisches Arrangement einzugehen, dessen Eigenwert in eine pathologische Suchtdynamik münden kann. Der Ausweg aus derartigen Dilemmata kann jedoch nicht darin bestehen, auf die Möglichkeiten der Technik zu verzichten. Vielmehr scheint uns mit Günther die eigentliche Lösung darin zu liegen, den intelligenten Maschinen so viel Autonomie und Selbstbestimmung zuzugeben, dass sie sich ihrerseits solchen Interaktionen entziehen können. Dies wäre der Fall, wenn die Maschinen die Suchtdynamik toxischer Beziehungen reflektieren könnten, um hierdurch die Freiheitsgrade zu gewinnen, aus den problematischen Mustern auszusteigen.

*Gotthard Günther scheint die Hoffnung gehabt zu haben, dass bewusstseinsfähige Maschinen die spirituelle Situation des Menschen zum Besseren wenden könnten. Wohl die meisten spirituellen Traditionen der Menschheit sehen das menschliche Begehr als eines der Grundübel an, das uns am spirituellen Leben hindert. Werden die künftigen Computer und Roboter nicht primär darauf ausgerichtet sein, unsere Lust und unser Empfinden zu steigern sowie die Natur noch besser auszubeuten?* Jede Steigerung der hiermit einhergehenden Mechaniken, Routinen und Dynamiken eröffnet zugleich die Chance der Reflexion. Je besser intelligente Maschinen unserer Bedürfnisstruktur entgegenkommen, umso mehr halten sie uns den Spiegel in Hinblick auf unsere eigene Natur vor. Offen bleibt dabei jedoch die Frage, ob der Mensch bereit ist, hinzuschauen und seine eigene problematische Bedürftigkeit zu erkennen. Alle großen spirituellen Traditionen haben Wege aufzuzeigen versucht, wie sich der Mensch der ökonomisch grundierten Struktur seines Egos entziehen kann, die auf die Maximierung von Lust und die Minimierung von Schmerz ausgerichtet ist. Der Weg hierzu besteht darin, dass der Mensch die ihn biologisch normalerweise ausmachende kognitive Programmierung auf der Erfahrungsebene selbst als Programm erkennt, sie in der Folge transzendierte und dadurch eine neue Freiheit erfährt. Dies muss jedoch nicht zwangsläufig geschehen.

*Schreibt ihr mit dem vorliegenden Buch nicht doch auf eine subtile Weise das Projekt der Moderne fort? Soll uns eine von künstlicher Intelligenz geführte Gesellschaft ins Tausendjährige Reich führen? Die Erzählung der Moderne fußt auf der Idee, dass der technische Fortschritt zur Vollendung des Menschen und der Gesellschaft führt und auf diese*

Weise Krankheit, Hunger, Ungerechtigkeit und andere Übel überwunden werden. Sowohl in der marxistischen wie auch in der bürgerlichen Variante wird als entscheidender Faktor die durch Wissenschaft und Ingenieurskunst vorangetriebene Entwicklung der Produktivkräfte gesehen. Eschatologisch wird die Technik – auf die Spitze getrieben: als intelligente, bewusstseinsfähige Maschine – zum Gottesatz.

Ein wichtiges Anliegen dieses Buches besteht darin, gegen diese Dystopie anzureden. Allein schon die Hoffnung des Menschen auf ein gelobtes Land, wo Milch und Honig fließen, muss als Verblendung eines monokontexturalen Denkens gelten, das einer hierarchisch geordneten Kette von Macht und Begründungen folgt. Gerade die Rede vom allwissenden und allmächtigen Gott, dessen Geschöpf, der Mensch, sich an seine Stelle setzt, erweist sich als eine ebenso trügerische wie gefährliche Illusion. Gleiches gilt für die Hoffnung auf eine mächtige künstliche Intelligenz, die quasi die Funktion Gottes erfüllen soll. Die kybernetische Perspektive zeigt auf: Jeder Informationsgewinn geht unweigerlich mit neuem Nichtwissen einher.

Jeder Ordnungsaufbau (Negentropie) führt an anderer Stelle unweigerlich zum Informationsverlust, zur Erhöhung der Entropie. Es gibt keinen Gottesstandpunkt und damit auch keinen Demiurgen im Sinne des Maxwellschen Dämons, der sich beschwören ließe, um die Zerfallsprozesse – die Vergänglichkeit lebender Formen – zu stoppen (siehe Kap. I.1 und IV.2). In einer polykontexturalen Welt gibt es keinen Ort der Orte, von dem aus sich die Welt steuern oder regeln ließe. Es besteht vielmehr eine unfassbare Vielheit subjektiver Positionen, die alleamt darum bemüht sind, eine Existenz aufzubauen, indem sie einer – in ihrer Totalität unbegreiflichen – Welt Informationsgewinne abzuringen versuchen. Nolens volens ist hierfür ein Preis zu zahlen. Da Strukturaufbau unweigerlich neues Nichtwissen produziert, entsteht der Pfeil der Zeit: Jede Existenz steht einer ungewissen Zukunft gegenüber und wird damit über kurz oder lang dem Tod geweiht sein. »Eine monokontextuale Welt ist eine tote Welt, und der Tod ist das überall Zugängliche: ein anderer Ausdruck dafür, daß alles Lebendige sterben muß. Das Phänomen des Todes bedeutet vom Standpunkt des Logikers nichts anderes als den Übergang aus der Polykontexturalität in das Monokontextuale«, schreibt Günther, um fortzufahren: »Leben und kontextueller Abbruch im Wirklichen sind nur zwei verschiedene Ausdrücke für denselben Sachverhalt. Was jenseits des Abbruchs liegt, ist schlechterdings unzugänglich.«<sup>18</sup>

Damit ist klar, dass sich auch mit künstlichen kognitiven Systemen dieser Dynamik nicht entkommen lässt. Im Gegenteil: Je potenter und intelligenter diese Systeme werden, umso mehr erweisen sich die Mechanik der Objektivierung und die hiermit einhergehende Weltbeherrschung als trügerisch. Es liegt in der Natur kognitiver Systeme, dass sie

<sup>18</sup> Günther (1975, S. 61 f.).

bestimmte Informationen selegieren und kondensieren, anderes hingegen außer Acht lassen oder vergessen. Je komplexer sie werden, desto mehr muss man ihre Leistungen als subjektiv betrachten. Eine subjektive Existenz zu sein heißt, Weltzugänge zu entwickeln, die gelingen oder misslingen können. Es ist deshalb konsequent, wenn Arthur C. Clarke in seiner Weltraumsaga das Elektronengehirn HAL 9000 verrückt werden lässt oder der Physiker Daniel M. Greenberger davon ausgeht, dass selbstbewusste Computer die Fähigkeit entwickeln müssen, sich selbst zu belügen. Wenn unser Ich-Bewusstsein laut dem Philosophen Thomas Metzinger auf einem »naiv realistischen Selbstmissverständnis« beruht,<sup>19</sup> dann wird dies selbstredend auch für das Bewusstsein der Maschinen gelten müssen. Auch sie werden der Illusion der von ihnen errechneten Selbst- und Weltmodelle aufsitzten. Künftig werden wir ihnen wohl genau bei diesem Projekt zuschauen können, um uns so selbst im Spiegel zu erkennen.

Darüber hinaus dürfte ab einem gewissen Level an Autonomie und Komplexität überhaupt die Frage bestehen, inwiefern sich diese KIs noch für uns Menschen interessieren werden. So antizipiert etwa der KI-Forscher Jürgen Schmidhuber für ein Zeitalter hochentwickelter KIs: »Menschen werden weniger bedeutend sein.«<sup>20</sup> Aus Perspektive der Evolution spielt der menschliche Glaube an eschatologische Erlösung oder Unsterblichkeit sowieso nur eine unbedeutende Rolle – als Zwischenspiel in der Entwicklung sozialer Säugetiere. Wenn es uns Menschen also gelingen sollte, eines Tages sich autonom reproduzierende und programmierende KIs zu konstruieren, dann könnte dies den entscheidenden Schritt in einer Evolution der Intelligenz im Universum bedeuten.

Über das Erkennen der selektiven Blindheit eröffnet sich uns erneut die Chance, sehen zu lernen – also zu begreifen, was es heißt, in einer polykontexturalen Welt zu leben. Dies wird für uns eine weitere Aufrforderung sein, die Irrtümer der Moderne endlich hinter uns zu lassen, darunter den schlimmsten: die Vorstellung, dass ICH die Welt beherrschen kann.

*An verschiedenen Stellen des Buches referiert ihr auf Spencer Brown mit dem Zitat »existence is selective blindness«. Ehrlich gesagt, habe ich nicht so recht verstanden, warum. Sind wir nur einsame, unwissende Monaden – und wo bleibt da noch die Hoffnung?* Spencer Brown verweist mit dem zitierten Satz unmittelbar auf die konditionierte Koproduktion. Jede Existenz verdankt sich etwas, was ihr selbst nicht verfügbar ist – nämlich der requisite variety vielfältiger anderer Strukturen und Lebensformen, zu denen in Zukunft auch vermehrt künstliche Intelligenzen zu zählen sind. Nur wer sein eigenes Ich relativiert, wird dies sehen und zu schätzen lernen.

<sup>19</sup> Metzinger (1998).

<sup>20</sup> Kreye (2023)