



Georg Souvignier
Frank Vogelsang (Hg.)

Durch Digitalisierung zur Freiheit 4.0?

Georg Souvignier und Frank Vogelsang (Hg.)

Durch Digitalisierung zur Freiheit 4.0?

Georg Souvignier und Frank Vogelsang (Hg.)

Durch Digitalisierung zur Freiheit 4.0?

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

wbg Academic ist ein Imprint der wbg
© 2021 by wbg (Wissenschaftliche Buchgesellschaft), Darmstadt
Die Herausgabe des Werkes wurde durch die
Vereinsmitglieder der wbg ermöglicht.
Satz und eBook: Satzweiss.com Print, Web, Software GmbH
Gedruckt auf säurefreiem und
alterungsbeständigem Papier
Printed in Germany

Besuchen Sie uns im Internet: www.wbg-wissenverbindet.de

ISBN 978-3-534-40603-6

Elektronisch ist folgende Ausgabe erhältlich:
eBook (PDF): 978-3-534-40604-3

Dieses Werk ist mit Ausnahme der Einbandabbildung als Open-Access-Publikation im Sinne der Creative-Commons-Lizenz CC BY-NC-ND International 4.0 (»Attribution-NonCommercial-NoDerivatives International«) veröffentlicht. Um eine Kopie dieser Lizenz zu sehen, besuchen Sie <http://creativecommons.org/licenses/by-nc-nd/4.0/>. Jede Verwertung in anderen als den durch diese Lizenz zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages.

Inhalt

<i>Frank Vogelsang / Georg Souvignier:</i> Einführung – Durch Digitalisierung zur Freiheit 4.0?	6
<i>Martina Hessler:</i> Ambivalenz der Technik – Freiheit und Notwendigkeit	18
<i>Gregor Schöner:</i> Die neue künstliche Intelligenz und ihre möglichen Auswirkungen auf Mensch und Gesellschaft	35
<i>Gerhard Lakemeyer:</i> Kognitive Robotik – Chancen und Risiken	57
<i>Andreas Schuppert:</i> Big Data in der Medizin – Bedrohung oder Chance für die Freiheit?	78
<i>Philipp Schönthaler:</i> Das große Zeitalter einer neuen Dichtkunst – Automation, Autonomie, Schreiben	114
<i>Dan Verständig:</i> Die Berechnung der Bildung – Über das Verhältnis von Mündigkeit und Bildung in der digitalen Welt	133
<i>Axel Siegemund:</i> Leitplanken 4.0 für die Freiheit 4.0? Von der Rationalität des Freiheitsverlustes	154
<i>Markus Mühling:</i> Freiheit durch Digitalisierung? Theologische Bewertung einer fragwürdigen Hoffnung	171
Autoren	182

Einführung – Durch Digitalisierung zur Freiheit 4.0?

Die Moderne ging schon immer mit einem großen Freiheitsversprechen einher. Zunächst zielten die Hoffnungen auf eine Demokratisierung der Gesellschaften. Doch spätestens im 20. Jahrhundert stiegen auch die Erwartungen an eine Alltagstechnik, die die Freiheitsgrade vieler Menschen kontinuierlich vergrößern sollten. Das galt etwa für die Kommunikation: das persönliche Telefon, das galt auch für die Mobilität: das persönliche Automobil. Die digitalen Medien steigerten dieses Verfügbarkeitsversprechen nun dramatisch. Das Internet machte Daten weltweit verfügbar, die Geräte sind, kaum entwickelt, so preiswert und allgemein verfügbar, so dass Jugendliche schnell zu „digital natives“ wurden. Die Sprache verriet in den ersten Jahren eine neue Leichtigkeit im Umgang mit der Technik: Menschen „browsten im Netz“ oder „surften auf Internetseiten“. Freundschaften konnten quer über den Globus durch soziale Medien geschlossen oder gepflegt werden, „Connectedness“ und lose soziale Bindungen waren das Ziel.

In all dem zeigte sich scheinbar die Erfüllung eines großen Freiheitsversprechens. Doch seit einigen Jahren werden auch Schattenseiten der digitalen Technologien deutlicher. Nicht nur werden weltweit Daten verfügbar, auch die eigenen, persönlichen Daten können von Anderen zu fremden Zwecken missbraucht werden. Die Leichtigkeit, schnell antworten zu können, wandelt sich in den Zwang, schnell zu antworten. Digitale Algorithmen scheinen vorzugeben, was uns die digitalen Medien zur Verfügung stellen und was nicht. Es zeigt sich: Freiheit und digitale Medien stehen in einem komplexeren Verhältnis zueinander, als es auf den ersten Blick den Anschein hat. Dabei ist die Entwicklung noch lange nicht an ihr Ende gekommen. Inwiefern erhöht der Einsatz autonomer Systeme mit Künstlicher Intelligenz die per-

sönliche Freiheit? Wird sie nicht zugleich eingeschränkt, weil zunehmend (Vor-)Entscheidungen auf technische Systeme verlagert werden? Wie kann und wie wird das unser Leben, das zwischenmenschliche Miteinander, gesellschaftliche Werte und unser Selbstverständnis durch den immer umfassenderen Einsatz digitaler Technologien verändern? Worauf müssen wir bei der Einführung dieser Technologien achten, damit fundamentale Werte gewahrt bleiben und Freiheiten nicht verloren gehen?

Der vorliegende Band dokumentiert eine interdisziplinäre Tagung, die im Herbst 2019 in der Bischöflichen Akademie des Bistums Aachen in Kooperation mit der Stiftung „Theologie und Natur“ und der Evangelischen Akademie im Rheinland stattgefunden hat. Sie thematisierte aus sehr unterschiedlichen Perspektiven die Möglichkeiten, Grenzen und Konsequenzen digitaler Technologien für die Erweiterung und Gestaltung menschlicher Freiheit. Ist das Freiheitsversprechen, das die Technikentwicklung der Menschheit von Beginn an begleitet hat, durch die Potenziale der digitalen Technologien seiner Erfüllung nähergekommen? Die interdisziplinäre Tagung diskutierte diese Frage aus sehr unterschiedlichen Perspektiven. So kamen die Geschichtswissenschaft, technische Wissenschaften wie Robotik, Informatik und Biomedizin, aber auch die Literaturwissenschaft, Pädagogik, Technikethik und Theologie zu Wort. Die Vielzahl unterschiedlicher Disziplinen sollte eine Engführung auf einen bestimmten Begriff von „Freiheit“ oder auf einen bestimmten Begriff von „Technik“ und bestimmte technologische Probleme verhindern. Denn beide Aspekte des Themas, die digitalen Technologien wie auch die Freiheitsvorstellungen, sind sehr variantenreich. So wenig es die eine digitale Technologie gibt, so wenig gibt es den einen normativen Begriff von Freiheit.

Eine besondere Aktualität erhalten die Ergebnisse dieser interdisziplinären Tagung, die ein halbes Jahr vor dem Ausbruch der Corona-Krise in Europa stattgefunden hat, durch die neue Aufmerksamkeit, die die digitalen Medien in den Zeiten der Pandemie gewonnen haben. Die digitalen Technologien sind viel stärker in den Fokus der gesellschaftlichen Aufmerksamkeit gerückt. Die gesellschaftlichen Veränderungen im Zuge der Pandemie sind für unser Thema alles andere als eindeutig: Die gewonnenen Freiheiten des Homeoffice, die Möglichkeit, sich nicht durch den Berufsverkehr quälen zu

müssen, stehen der Einbindung in die häuslichen Verhältnisse und den damit einhergehenden Doppelbelastungen gegenüber. Die Freiheit, selbstbestimmter die Arbeit organisieren zu können, steht der Notwendigkeit gegenüber, auch außerhalb der üblichen Zeiten zu arbeiten. Schon an dieser ersten Gegenüberstellung wird deutlich: Das Verhältnis von Freiheit und digitalen Technologien gleicht eher einem Spannungsfeld als einer parallelen und sich wechselseitig verstärkenden Entwicklung. Die Gestaltung des Verhältnisses ist grundlegend für die künftige gesellschaftliche Entwicklung, sie ist eine zutiefst politische und zugleich eine kulturelle Herausforderung!

Martina Heßler zeigt in ihrem geschichtswissenschaftlichen Beitrag, dass die Vorstellungen von Freiheit in der europäischen Geschichte eng mit dem Projekt neuzeitlicher Humanität verbunden sind. Doch die Beziehung der Vorstellungen von Humanität und damit auch von Freiheit zur technischen Entwicklung war und ist von Beginn an ambivalent. Ihr Verhältnis schwankt zwischen einem Freiheitsversprechen durch neue Technologien, etwa indem sie die Autonomie der Menschen vergrößern und einer Freiheitsbedrohung durch die Technik, die sich in einer entfremdeten Arbeit, eingebunden in einer übermächtigen Maschinerie, zeigen kann. Technik verheißt auf der einen Seite die Freiheit von Übeln und von körperlichen Beschränkungen und die Freiheit von der Abhängigkeit gegenüber anderen Menschen. Auf der anderen Seite setzt die Technik eigene, technikimmanente Zwänge, die die Handlungsmöglichkeiten wiederum einschränken. Der Philosoph Ernst Cassirer hat vorgeschlagen, den Wert der Technik gerade an der Frage zu beurteilen, inwieweit sie die menschliche Freiheit in einem humanistischen Sinne vergrößert oder inwieweit sie sie einschränkt. Dieses ambivalente Verhältnis von humaner Freiheit und technischen Systemen lässt sich auch anhand der Einführung digitaler Technologien nachvollziehen. Zu Beginn der Entwicklung, in der zweiten Hälfte des 20. Jahrhunderts, stehen wiederum zunächst große Freiheitserwartungen. Alle Menschen sollen über einen Personal Computer verfügen, mit dessen Hilfe sie sich miteinander verknüpfen und selbstbestimmt kommunizieren können. Dies ist die Vision von Innovatoren und technischen Gründern wie Sherry Turkle oder Steve Jobs. Doch mit der zunehmenden Ausbreitung der Technologien im Alltag ist die anfängliche Euphorie schnell wieder verfliegen. In den ersten Jahrzehnten des

21. Jahrhunderts stehen eher die Verzerrungen und Gefahren im Mittelpunkt. Statt einer freien Information wird nun beklagt, dass in den digitalen sozialen Medien Bubbles entstehen, also Informationsblasen von Gleichgesinnten, die Außenstehende kaum noch wahrnehmen. Menschen kommunizieren de facto nicht frei, sondern eingebunden in ihren Kommunikationsblasen. Die Ambivalenzen, die die Entwicklung der Technik immer schon in ihrem Verhältnis zur Freiheit charakterisierte, haben sich nicht aufgelöst, sondern durch die Einführung digitaler Technologien eher noch verschärft.

Nach der historischen Einführung folgen Beiträge aus technischen Perspektiven. Ist auch das Verhältnis der Technik zur menschlichen Freiheit ambivalent, so ist zugleich unbestreitbar, dass die technischen Systeme selbst an Freiheitsgraden gewinnen, sie werden immer autonomer. Gregor Schöner ruft zu Beginn seines Beitrags eine allgemeine Definition für Systeme künstlicher Intelligenz in Erinnerung: Diese Systeme nehmen Signale der Umgebung wahr und verhalten sich dementsprechend. Seit dem Beginn der Forschung an künstlicher Intelligenz bildete der Vergleich mit der Intelligenz biologischer Wesen, natürlich insbesondere des Menschen, den Hintergrund. Um Systeme künstlicher Intelligenz schaffen zu können, ist es notwendig, ein Verständnis der inneren Zustände zu gewinnen, die reproduzierbar ein bestimmtes Verhalten herbeiführen. Allerdings erreichte auch intensive Forschung bislang keinen Zugang zu dem, was man allgemeine Intelligenz nennt, denn diese setzt ein Konzept von Welt voraus, in dem auch neue Dinge und Handlungen eingebunden werden können. Bis zum heutigen Tage sind Systeme künstlicher Intelligenz auf einen begrenzten und wohl definierten Satz von Daten angewiesen. Es gibt gewisse Fortschritte, seitdem Verfahren mit Wahrscheinlichkeitskalkülen verwendet werden, die einfache, sinnvolle Reaktionen auch in einer komplexeren Umgebung möglich machen. Einen großen Sprung nach vorne ermöglichte das neuronale Programmieren, das sich die Funktionsweise vernetzter Gehirnzellen zum Vorbild nahm. Allerdings: die Zahl der Parameter in diesen Netzen ist sehr groß und nur mit leistungsfähigen Rechnern zu bewältigen. Auch wird eine sehr große Menge strukturierter Daten benötigt, um die Systeme zu programmieren und zu optimieren. So werden die heute bekannten Systeme künstlicher Intelligenz ermöglicht, etwa als Fahrassistenzen oder Bilderkennung. Doch Schöner

bleibt kritisch und bezweifelt, dass es in absehbarer Zeit möglich sein könnte, auf diese Weise Modelle menschlicher Intelligenz zu schaffen. Das größte Problem: Viele relevante Anteile des Weltwissens gibt es nur in einer geringen Zahl von Datensätzen, so dass für diese Ausnahmen nicht genügend Lernmaterial vorliegt. Noch schwieriger ist es, wenn das Wissen nur implizit vorliegt, also nicht als klare Verknüpfung definierter Daten. Die menschenanaloge Metaphorik ist deshalb bei Künstlicher Intelligenz fragwürdig, nach Schöner sollte man eher von „Parameteranpassung“ als von „Lernen“ sprechen. Doch auch wenn das hochgesteckte Ziel der Menschenähnlichkeit nicht erreicht wird, sind neue KI-Systeme starke Instrumente für eine Vielzahl von Anwendungen. Diese können jedoch ethisch inakzeptable Folgen mit sich bringen, z. B., indem Korrelationen in den Ausgangsdatsätzen verallgemeinert und auf den Einzelfall in der Zukunft angewandt werden. Wenn eine ethnische Minderheit in diesen Datensätzen statistisch häufiger verschuldet ist als der Bevölkerungsdurchschnitt, wird eine KI, die mit diesen Daten optimiert wurde, einen Vertreter dieser Gruppe generell als weniger kreditwürdig klassifizieren. Eine Diskriminierung ist offenkundig. Aufgrund der Verfügbarkeit von wirtschaftlichen und politischen Daten ist auch eine bewusste Manipulation beim Kaufverhalten und in der politischen Meinungsbildung nicht ausgeschlossen. Wenn die Bedeutung virtueller Welten größer wird, wächst schließlich auch die Gefahr, dass KI-Systeme Nutzerinnen und Nutzern künstliche Welten präsentieren, die unter Berücksichtigung derer individuellen Profile manipuliert sind.

Gerhard Lakemeyer geht der Frage autonomer Systeme aus der Perspektive der kognitiven Robotik nach. Zunächst zeigt er, dass erstaunlicherweise während des gesamten 20. Jahrhunderts ein autonomer Roboter nur eine Zukunftsvision und Gegenstand von Science Fiction war. Erst gegen Ende des Jahrhunderts kam es zum Durchbruch. Der Roboter „Rhino“ konnte sich als einer der Ersten relativ autonom in wechselnden Umgebungen bewegen. Damit war die Grundlage für das autonome Fahren gelegt, das jetzt immer mehr zur Marktreife gelangt. Nach Lakemeyer ist jedoch festzuhalten, dass es sich dabei um eine isolierte, hochspezifische Realisierung autonomer Robotik handelt. Die Robotik macht zwar kontinuierlich Fortschritte und entwickelt sich rasch, dennoch liegen komplexe Objektmanipulationen wie

bei der Fertigkeit der menschlichen Hand auch heute noch außerhalb des technisch Möglichen. Auch über ein Weltwissen verfügen die technischen Systeme noch lange nicht. Eine entscheidende Hürde für eine Autonomie, die der menschlichen Freiheit ähnelt, ist das Handeln in einer sich wandelnden Umgebung. Menschen haben eine Urteilskraft, die ihnen ermöglicht, auch in ungewohnten Situationen angemessen zu reagieren. Das Handeln von Robotern dagegen basiert auf Erfahrungswissen, das immer Lücken aufweist. Trotz dieser aktuell bestehenden Grenzen warnt Lakemeyer jedoch zugleich vor dem Gefahrenpotenzial, das von militärischen Anwendungen der Robotik ausgeht, die bereits heute im Bereich des Möglichen liegen. Weiterhin muss das Handeln von Robotern auch im zivilen Bereich ethischen Regeln unterliegen: Sie dürfen Menschen nicht verletzen und müssen deren Befehle so gut wie möglich ausführen. Roboter sind technische Systeme, deren menschengefährdende Fehlleistungen auch in Ausnahmesituationen nicht akzeptiert werden. Eine wesentliche Voraussetzung für die Akzeptanz und Integration künstlich intelligenter Roboter in gesellschaftliche Vollzüge sieht Lakemeyer darin, Transparenz über die Fähigkeiten, Funktionsweisen und Grenzen dieser Systeme herzustellen. Neben diesen konkreten Grenzen sind fundamentale Fragen nach wie vor offen: Sollten Roboter zu Emotionen in der Lage sein? Haben sie Bewusstsein? Können Roboter künstlerisch kreativ sein? Schließlich bleibt die Frage, ob intelligente Roboter künftig die Freiheit der Menschen begrenzen können.

Die Verbindung von Freiheit und den Techniken der Medizin war schon immer eng. Krankheiten binden und reduzieren das Leben, Heilungen schaffen dagegen Befreiung. Andreas Schuppert diskutiert in seinem Beitrag zunächst die Frage, was überhaupt als Krankheit anzusehen ist und inwieweit diese die Freiheit des Menschen einschränkt. Es wird deutlich: Bei weitem nicht jede Einschränkung der Handlungsmöglichkeiten eines Menschen kann sogleich Krankheit genannt werden. Schuppert geht es in seinem Beitrag aber nicht so sehr um Fragen der klassischen Medizin, sondern vielmehr um eine Medizin mit intensiver Datenverarbeitung durch digitale Technologien. Wie eng schon heute die moderne Medizin mit digitaler Datenauswertung verknüpft wird, zeigt ja nicht zuletzt die aktuelle Bekämpfung der Corona-Pandemie. Auch die fortgeschrittene Gendiagnostik und die Moleku-

larbiologie bieten Möglichkeiten für eine datenbasierte Medizin. Hier werden sehr komplexe Datensätze erhoben, die für das Design von Arzneien und neueren Therapieansätzen genutzt werden können. Doch die Molekularmedizin stößt an ihre Grenzen, wenn komplexe Umweltbedingungen und ein variables Verhalten der Patienten einen großen Einfluss auf den Krankheitsverlauf haben. Es kommt dann darauf an, durch Erhebung umfangreicher Datensätze dem Ziel der „Real World Evidence“ nahezukommen. Eine große Hürde stellen in diesem Zusammenhang seltene Erkrankungen dar, über die es nur begrenzte Daten gibt. Vollends problematisch ist die Erhebung und sachgemäße Auswertung von Daten, die für die präventive Medizin notwendig wäre. Schon die Pandemie-Erfahrung zeigt: Hier können direkte und harte Zielkonflikte gerade mit den Werten von Freiheit und Selbstbestimmung der Patientinnen und Patienten entstehen. Präventive Maßnahmen können, wenn sie nicht von den Einzelnen gewollt werden, das Leben von Menschen massiv einschränken. Am Ende steht ein Blick in die Zukunft: Programme mit so genannter Künstlicher Intelligenz sind sehr gut darin, Muster in großen Datensätzen zu erkennen. Machen solche Systeme letztendlich die moderne Medizin effizienter? Schuppert ist da zurückhaltend. Er verweist auf viele praktische Grenzen, auf die eine digitalisierte Medizin auch heute noch stößt, bei komplexen Bedingungen einer Datenerhebung und vor allem beim Datenschutz für hochsensible persönliche Daten. Denn durch einen verbreiteten Missbrauch wären bei allem medizinischen Fortschritt grundlegende gesellschaftliche Freiheitsrechte in Gefahr. Dennoch vertritt er unmissverständlich die Position, dass der Verzicht auf die Nutzung von Big Data im medizinischen Kontext mit Verlust von Freiheit und Lebensqualität verbunden wäre. Die Ambivalenz der Technik in Bezug auf den Wert der Freiheit zeigt sich also auch hier in aller Deutlichkeit.

Der nun folgende, dritte Teil der Tagungsbeiträge umfasst einige geisteswissenschaftliche Perspektiven, die sich mit kulturellen Auswirkungen der digitalen Technologien auseinandersetzen. Den Anfang macht der Literaturwissenschaftler Philipp Schönthaler. Auch er setzt historisch ein und erinnert an die Diskussionen um die Einführung von Schreibmaschinen zu Beginn des 20. Jahrhunderts. Erstaunlich ist die Nähe der damaligen Technik-enthusiasten und -kritiker zu denen heutiger Positionen angesichts digitaler

Medien und Künstlicher Intelligenz. Diese Interpretation der Technik zielt auf eine Kritik an der bürgerlichen Vorstellung vom Subjekt, an dem autonom handelnden, kreativen Genie. Ist es möglich, so die strittige Frage, dass neuere Technologien nicht nur eine Hilfe beim Abfassen von Texten sind, sondern auch den geistigen Akt des Verfassens von Texten übernehmen können? In der zweiten Hälfte des 20. Jahrhunderts versuchten Akteure digitale Maschinen und menschliche Autoren immer mehr einander anzunähern. Was zunächst Theorie blieb, mag in der Gegenwart – im Jahr 2020 – Realität werden können, wie die aufsehenerregende Software GPT-3 der kalifornischen Firma OpenAI nahelegt. Hier wird einem technischen Artefakt zugesprochen, eigenständig und kreativ Texte erstellen zu können. Doch kann man die Leistung der digitalen, hoch vernetzten Maschinen auch anders deuten. So beurteilt Schönthaler auch die neueren Versuche, menschliche Akteure durch autonome Maschinen zu ersetzen, vor allem als ideologische Programme. Sie stehen für eine bestimmte Weltsicht, weniger für eine nüchterne technische Möglichkeit. Es geht, so Schönthaler, zentral darum, wie das Verhältnis des Menschen zu Schrift und Sprache gedacht wird. Und hier zeigt sich, dass Sprache und Schrift durch digitale Medien eher abgewertet werden. Man interpretiert sie als Datensätze, versucht sie zu ersetzen etwa durch biometrische Daten. Eine klassische Hermeneutik, die Suche und die Unterscheidung von Sinn und Bedeutung fallen damit aus. Wäre es aus dieser Weltsicht nicht besser, wenn Gehirne gleich direkt miteinander kommunizieren ohne den Umweg der Sprache? Alles soll in verfügbare und berechenbare Daten verwandelt werden. Diese Digitalökonomie ist eine Emanzipierung von Sprache und Schrift. Die Eigenständigkeit des menschlichen Schreibens und Sprechens ist damit aber nicht verloren, so die positive Bilanz Schönthalers zum Schluss, sie wird sich auch in einer technischen Zukunft als unverzichtbar erweisen.

Die problematische Rolle von Algorithmen, die die Vielzahl von Daten und Datenträgern in Netzwerken beherrschbar machen, steht auch am Anfang der Überlegungen von Dan Verständigs Beitrag aus der Perspektive der Pädagogik. Der Umgang mit den digitalen Medien ist heute vom Gedanken von „Big Data“ getrieben. Dieser Zugang zur Wirklichkeit gibt sich den Anschein von Objektivität, tatsächlich sind die Daten aber sehr interpreta-

tionsbedürftig, de facto sind sie politisch und ökonomisch beeinflusst. Die Einführung von digitalen Technologien geht oft mit der Vorstellung einher, Datensätze ließen sich objektiv und sehr effizient bearbeiten. Der Imperativ der Sichtbarkeit beherrscht etwa die sozialen Medien: Nur das ist wirklich und relevant, was eine größere Sichtbarkeit erlangen kann. Wie lassen sich aber Bildungsprozesse in einer solchermaßen strukturierten sozialen Welt gestalten? Bildung wie Verständig sie definiert, qualifiziert das Individuum für den Umgang mit dem Ungewissen, Unbekannten. Bildung kommt zum Zuge, wenn die Routinen des Alltags und die Routinen der Datenverarbeitung durchbrochen werden. Ziel der Bildung ist die Förderung eines selbständigen und mündigen Lebens auch in der durch digitale Technologien vermittelten Welt. Anders ausgerichtet sind allerdings die Aktionspläne, die politische Instanzen wie die Kommission der Europäischen Union auf den Weg bringt: Hier geht es um digitale Kompetenz, die direkt mit der Stärkung des Wirtschaftsstandortes verknüpft ist. Alles zielt auf die Kompetenz im Umgang mit digitalen Technologien, nicht ihre kritische Reflexion. Optimierung von Bildungsprozessen meint dann immer zugleich auch Standardisierung. Der Weg der Bildung ist aber mit einem Mündigkeits- und Freiheitsversprechen verbunden, das nicht technologisch eingelöst werden kann. Diese Mündigkeit wird vielmehr durch Algorithmen digitaler Technologien eher gefährdet. Verständig wertet damit nicht einfach digitale Technologien ab, er fordert aber eine Transparenz im Umgang mit den Algorithmen etwa über Open Source Systeme. Die reale Gefahr besteht, dass digitale Systeme in der Bildung auch als umfassende Überwachungssysteme wirken können. Bildung braucht aber genuin Freiheit, sie ist immer eine offene Suchbewegung.

Der Theologe und Ingenieur Axel Siegemund betrachtet das Verhältnis von Freiheit und digitalen Medien aus der Perspektive der Ethik. Auch er setzt historisch an: Die digitalen Medien reihen sich als Herausforderung für das menschliche Selbstbild in die Vielzahl der naturwissenschaftlichen Entdeckungen der Neuzeit ein. Kann ein freier Wille bei einem Aggregat mit künstlicher Intelligenz entstehen? Siegemund unterscheidet grundsätzlich zwischen Willen und Intelligenz – beide sind nicht notwendigerweise miteinander verknüpft. Die Freiheit muss aber dem Willen zugesprochen

werden. Auch Methoden der Stärkung des eigenständigen Willens, etwa in der Psychoanalyse, haben nicht bei der Intelligenz eingesetzt, sondern bei der Variierung der Willensäußerungen. Wenn also digitale Intelligenz nicht unmittelbar Willensfreiheit fördern kann, so kann sie aber Bedingungen schaffen, die den Willen stärker binden. Siegemund warnt davor, den Menschen nach der Art technischer Systeme zu verstehen. Denn letztere haben immer nur so viele Freiheitsgrade, wie ihre Konstrukteure ihnen zubilligen. Dagegen ist die Freiheit des Menschen von anderer Qualität. Allerdings ist diese Freiheit nicht erst durch die Technik in Frage gestellt worden, sie war und ist in vielen Debatten eine umstrittene Größe. Die Differenz zwischen Menschen und künstlichen technischen Systemen wird gerade hier deutlich: Menschen problematisieren ihre Freiheit. Freiheit ist nicht einfach eine Eigenschaft, sondern die Fähigkeit, eine gegebene Situation, die zum Handeln zwingt, anzuerkennen und ihr einen Sinn zu geben. Damit weist die Freiheit auf sinnhafte Bezüge innerhalb einer Kultur. Anhand des Verhältnisses von Wissenschaft und Ethik macht Siegemund deutlich, dass Ethik nicht einfach für eine absolute Freiheit steht, sondern für eine solche Freiheit, die sich immer auch auf – moralische – Ordnungen beziehen muss. Es geht nicht um Freiheit an sich, sondern um die Bedingungen, unter denen sich Freiheit entfalten kann. Die digitalen Technologien stellen deshalb eine fundamentale Anfrage an das klassische Menschenbild. Ist der Mensch überhaupt frei oder nicht eher eine komplexe Rechenmaschine? Hier wird der Geist mehr im Sinne einer digitalen Realisierung verstanden, was den Menschen zunehmend eindimensionaler erscheinen ließe. Der Mensch kann aber nicht auf bestimmte Qualitäten reduziert werden, er ist ein komplexes leibliches Geflecht aus Körper und Geist. Zum Ende bezieht Siegemund ein theologisches Argument ein, um menschliche Freiheit zu bestimmen: der Mensch und sein Freiheitsverhältnis zur Welt ist durch seine Gottesbeziehung konstituiert, die nicht durch digitale Technologien affiziert werden kann. Das ermöglicht aber einen freien Umgang mit den Möglichkeiten der Technologien, sie sind stets pragmatisch auf ihre Tauglichkeit zu prüfen. Schließlich stellt Siegemund hierzu Regeln auf: Die digitale Identität muss so immer nur eine Möglichkeit unter mehreren sein, digitale Technologien müssen revidierbar bleiben.

Markus Mühling unternimmt in seinem Beitrag eine dezidiert theologische Betrachtung des Verhältnisses von Freiheit und digitalen Technologien. Den christlichen Glauben versteht er als eine Weglinienperspektive, als eine Art des Wahrwertnehmens. Mit diesem Begriff markiert er, dass menschliche Wahrnehmung niemals wertfrei ist. Entscheidend ist in dieser Bestimmung der Bewegungsaspekt, christliche Existenz findet sich nicht einfach in einer vorgegebenen, objektiv beschreibbaren und statischen Welt wieder, sondern zeigt sich durch Beteiligung, durch Tätigkeit. Der Weltzugang entsteht durch Narrationen, die unmittelbare Lebensvollzüge zum Ausdruck bringen können. Dadurch erscheinen zum Beispiel einer Christin, einem Christen andere Menschen unmittelbar als Schwestern und Brüder – nicht erst im Zuge eines Nachdenkens. Welt erschließt sich Christinnen und Christen insbesondere auch durch die Narrationen biblischer Texte, die ihre Lebensvollzüge prägen. Dieses dynamische und vernetzte Bild menschlicher Existenz hat Folgen für das Verständnis von Freiheit: Diese kommt nicht isolierten Individuen zu, sondern Personen in Beziehung. Diese Personen erleben Freiheit nicht in einer abstrakten Wahlmöglichkeit, sondern dann, wenn sie ihre immer eingebundene und restriktive Lebensvorgaben vor allem als Gabe verstehen und bejahen können. Wie verhält sich nun diese Freiheit zu den digitalen Technologien? Letztere haben die Besonderheit einer fundamentalen Diskontinuität. Diese Bestimmung wird in ihrer Tragweite deutlich, wenn Mühling wieder auf die Bewegung als Existenzausdruck zurückkehrt und terminologisch zwischen „Wayfaring“ und „Transport“ unterscheidet. Letzteres ist eine Bewegung, die von den zu verbindenden statischen Punkten bestimmt ist. Diese Verbindungen können berechnet und technisch hergestellt werden. Der Weg als solcher hat demgegenüber keine besondere Relevanz. Dagegen ist bei dem Ersteren die Bewegung selbst primär, nicht die Punkte, die sie miteinander verbindet. Es geht um einen Lebensvollzug, der eben nicht so leicht technisiert werden kann. Mit dieser Unterscheidung ist deutlich: „Transport“ lässt sich gut durch digitale Technologien bewerkstelligen, nicht aber „Wayfaring“ als Lebensvollzug. Mit diesen beiden Begriffen ist eine Grundlage geschaffen, die Möglichkeiten und Grenzen digitaler Technologien kritisch zu beurteilen. Sofern sie für „Transport“ eingesetzt werden, haben sie ihren begrenzten Sinn und können sogar

Freiheit erweitern. Aber „Wayfaring“ als Lebensvollzug bedeutet auch, Wege zu verzögern, Umwege zu gehen, sich nicht auf ein Ziel zu fixieren. Deshalb kommt es in gesellschaftlichen Diskussionen darauf an, die digitalen Technologien auf ihren beschränkten Anwendungsbezug zu relativieren. Auch im Bereich des „Transport“ dürfen sie nicht absolut gesetzt werden, sie schaffen nur relative Verbesserungen, aber im Bereich des „Wayfaring“ haben sie keinen Platz, da sie die Welt der unmittelbaren Beteiligung nur verarmen und damit die Freiheit des Menschen einschränken würden.

Im Verlauf der Tagung wurde immer wieder deutlich, dass Technik und Freiheit in vielfältiger Weise miteinander verflochten sind. Dabei lassen sich drei Grundlinien unterscheiden. In der ersten ist Freiheit das Ziel menschlichen Handelns und Technik eine Hilfe, diese Freiheit zu erlangen. Die Freiheitsverheißung war immer schon mit der Technikentwicklung verbunden und war es im Besonderen zu Beginn der Entwicklung der digitalen Technologien. Inwieweit die Technologien dieses Versprechen einlösen, ist nach wie vor umstritten. Auf diesen Konflikt gehen mehrere Autoren ein. Eine zweite Linie fragt danach, in welchem Umfang durch die erweiterten technologischen Möglichkeiten die Freiheitsgrade der technischen Aggregate wachsen können. Schon früh war in der europäischen Kulturgeschichte von einer vom Menschen geschaffenen, selbsttätigen Kreatur die Rede. Die digitalen Technologien haben diese Erwartungen beflügelt. Jedoch: Können sie sie auch einlösen? Die dritte Linie wiederum setzt diese beiden Freiheitsvorstellungen in ein Verhältnis: Kann man anhand von autonomen Robotern und Computernetzen auch etwas über die Freiheit von Menschen lernen? Ist menschliche Freiheit mehr als die Summe der Freiheitsgrade hochkomplexer Systeme? Dies ist im Kern eine weltanschauliche Frage, die insbesondere in den Beiträgen aus den Geisteswissenschaften diskutiert wird. Trotz der Vielzahl unterschiedlicher disziplinärer Zugänge steht am Ende natürlich kein abschließendes Urteil in einer so grundlegenden Fragestellung. Es wurden aber interessante und weiterführende Argumentationslinien und Themenstellungen erkennbar, an denen künftige Projekte anknüpfen können.

Ambivalenzen der Technik – Freiheit und Notwendigkeit

Das Verhältnis von Freiheit und Technik ist im Kontext der Digitalisierung zu einem viel diskutierten Thema geworden. Vor allem das Internet ging mit begeisterten und teils utopischen Erwartungen einher: Emanzipation, Ermächtigung, Freiheit, die Unabhängigkeit von Autoritäten waren die Begriffe, hinter denen sich euphorische Hoffnungen verbargen.¹ Die 1996 vom Netzaktivisten John Perry Barlow formulierte „Unabhängigkeitserklärung des Cyberspace“ sprach im Namen der Freiheit den „Regierungen der industriellen Welt, (den) müden Giganten aus Fleisch und Stahl“ die Macht ab und erklärte den Cyberspace zum freien, regellosen und auch körperlosen Raum. „(M)it keiner größeren Autorität [...] als der, mit der die Freiheit spricht“,² wende er sich an die Regierungen.

Heute ist allerdings Ernüchterung eingekehrt. Wie Sascha Lobo nach dem NSA-Skandal frustriert schrieb, sei das Internet „kaputt“, die Erwartungen enttäuscht.³ Überwachung, Manipulation, Bedrohung, der Verlust der Privatsphäre, der Verlust von Freiheit werden inzwischen diskutiert. Vincent Cerf warnte kürzlich vor einer staatlichen Regulierung des Internets, die, so auch er, das Internet „kaputt“ mache.⁴

¹ Vgl. Turner, F., *From Counterculture to Cyberculture: Steward Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, Chicago/London 2006.

² Barlow, J. P., *Unabhängigkeitserklärung des Cyberspace*, in: Baumgärtel, T. (Hg.), *Texte zur Theorie des Internets*, Stuttgart 2017 (Text von 1996), 67.

³ Lobo, S., „Das Internet ist nicht das, wofür ich es gehalten habe“, in: FAZ-Net, 12. Januar 2014, <https://www.faz.net/aktuell/feuilleton/medien/sascha-lobo-das-internet-ist-nicht-das-wofuer-ich-es-gehalten-habe-12747989.html>, Zugriff 22. September 2020.

⁴ Cerf, V., *Macht das Internet nicht kaputt*, in: FAZ, 7. September 2020.

Historisch betrachtet, oszillierten die Visionen und Erwartungen an Technik seit der Moderne stets zwischen einem Freiheitsversprechen und der Warnung vor dem Verlust menschlicher Freiheiten. Tatsächlich aber war Technik nie nur Garant für Freiheit oder nur Sinnbild der Unterdrückung von Menschen. Vielmehr war sie stets ambivalent. Vielleicht wurde dies selten so offensichtlich wie im Kontext des Internets. Selten wurde wohl auch so intensiv und gesellschaftlich breit verankert über Freiheit und Technik diskutiert wie heute, vielleicht auch, weil das Internet und Digitalisierung so umfassend alle Lebensbereiche betrifft und weil die Erwartungen an Freiheit so überbordend waren und so massiv enttäuscht wurden.

Freiheitsversprechen und Ängste um Freiheitsverlust im Kontext neuer Technologien finden sich aber seit Jahrhunderten. Oft wurde dies allerdings auch in anderen Begriffen verhandelt: Technik als Schicksale, als Inbegriff von Unfreiheit, die Frage der Autonomie, aber auch Befreiung, also „frei sein von“.

Der Freiheitsbegriff ist zweifellos ein mit Pathos aufgeladener Begriff. Freiheit gilt als anthropologische Grundbestimmung, als „Projekt der neuzeitlichen Humanität“,⁵ als „zentrale Kategorie für die Beschreibung menschlicher Handlungen und Entscheidungen“,⁶ um nur einige emphatische Bestimmungen von Freiheit zu zitieren. Sartre betonte, dass Freiheit „nicht eine Eigenschaft der Menschen“ sei, „sondern aller menschlichen ‚Natur‘ vorausgehend (sic!), gleichbedeutend mit Existenz“.⁷

Es ist hier nicht der Ort die mannigfachen und komplexen Deutungen und Konnotationen des Freiheitsbegriffs zu rekonstruieren und zu diskutieren. Gleichwohl sind einige einführende Bemerkungen und Einordnungen notwendig, um das Verhältnis von Technik und Freiheit zu reflektieren.

⁵ Krings, H., Freiheit, in: Ders./ Baumgartner, H. M./ Wild, Chr. (Hg.), Neues Handbuch philosophischer Grundbegriffe, Freiburg 2011, 493–510, hier 493.

⁶ Pauen, M., Freiheit, in: Kolmer, P./ Wildfeuer, A. G. (Hg.), Neues Handbuch philosophischer Grundbegriffe, Band 1, Absicht – Gemeinwohl, Freiburg/München 2011, 801–817, hier 801.

⁷ Spaemann, R., Freiheit, in: Ritter, J., Historisch-Philosophisches Wörterbuch der Philosophie, Band 2, D-F, Darmstadt 1972, Spalte 1064–1098.

1. Freiheit – Begriffsgeschichtliches

Eine grundlegende Definition von Freiheit bezeichnet die Möglichkeit, so zu handeln, wie es dem eigenen Willen entspricht. Freiheit umfasst die Willensfreiheit, die Handlungsfreiheit, die politische Freiheit, die Meinungsfreiheit, die Entscheidungsfreiheit, die Wahlfreiheit. Bereits Aristoteles fasste Freiheit als konkretes Wählenkönnen.⁸ Thomas Hobbes oder John Locke stellten den Begriff der Handlungsfreiheit in das Zentrum: die Fähigkeit „to act or not to act according as we shall choose or will“.⁹

Unterschieden wird grundsätzlich, und das ist für die folgenden Ausführungen wichtig, zwischen einer positiven und einer negativen Bestimmung von Freiheit.

Die negative Bestimmung bezeichnet das Freisein von äußeren Zwängen, die Freiheit von Hindernissen (durch Zwang, Kausalität, Schicksal, Kontingenz). Die Freiheit *von* etwas, ist zentral für diesen negativen Freiheitsbegriff.¹⁰ Thematisiert wird dies seit der Antike: die Freiheit vom Schicksal, die Freiheit von politischem Despotismus und die Freiheit von den Übeln des Daseins.

Der positive Freiheitsbegriff meint Freiheit *zu* oder Freiheit *für* etwas. Er impliziert insbesondere die Möglichkeit der Selbstbestimmung, der freien Entscheidung und freien Wahl,¹¹ die Möglichkeit der Selbstgesetzgebung und ist damit eng gekoppelt an den Begriff der Autonomie. Die beiden Begriffe sind häufig nur analytisch zu trennen. Bedingt das eine doch das andere. So auch, wenn in der Antike Freiheit als „das, was sich ohne äußeren Zwang aus sich selbst entfaltet“¹² bezeichnet wird – die Freiheit *von* äußerem Zwang *für* die Selbstentfaltung.

Entscheidend für unser modernes Verständnis ist die hier schon anklingende Vorstellung von Freiheit, die seit der Renaissance und der Aufklärung

⁸ Precht, P., Freiheit, in: Ders./ Burghard, F.-P. (Hg.), Metzler Lexikon Philosophie, Begriffe und Definitionen, Stuttgart 2008, 187–189, hier 187.

⁹ Spaemann, Freiheit (Fußnote 7), Spalte 1096.

¹⁰ Precht, Freiheit (Fußnote 8), 187.

¹¹ Ebd.

¹² Spaemann, Freiheit (Fußnote 7), Spalte 1065.

die „unbehinderte allseitige Entfaltung der menschlichen Persönlichkeit“,¹³ also die Selbstbestimmung und Selbstverwirklichung, zu zentralen Begriffen des Menschseins machten und einen positiven Freiheitsbegriff bezeichnen.

Zur Bestimmung des Freiheitsbegriffs wurde des Weiteren auf das Urheberprinzip verwiesen, das bereits Aristoteles formulierte: Eine Handlung ist frei, wenn sie auf den Handelnden zurückzuführen ist.¹⁴ Das Urheberprinzip verweist zugleich auf Grenzen der Freiheit, indem es mit der Rückführung einer Handlung auf eine Person den Begriff der Verantwortung evoziert. Freiheit findet stets ihre Grenzen: in Notwendigkeiten, Kontingenzen, vor allem aber eben in der Verantwortung und moralisch-ethischen Dimensionen und in der Existenz eines Anderen.

2. Freiheit und Technik

Erstaunlicher Weise hat der Jahrtausende alte Freiheitsdiskurs Technik kaum thematisiert. Die historisch und philosophisch einschlägigen und umfassenden Überblicksartikel zu Freiheit, so zum Beispiel in den Geschichtlichen Grundbegriffen oder im Historischen Wörterbuch der Philosophie, auch im Metzler Handbuch Philosophie, schildern die feinsten Verästelungen und Semantiken und vor allem den Wandel des Begriffs der Freiheit seit der Antike. Auffälliger Weise wird Technik jedoch fast nie erwähnt, zu stark ist der Freiheitsbegriff offensichtlich mit Politik, Staat und dem Subjekt, der Person verbunden. Dies erstaunt einerseits, verweist aber andererseits darauf, dass die Herausforderungen und Hoffnungen, die Technik für menschliche Freiheit bedeutet, historisch neu sind und dass vor allem seit dem 20. Jahrhundert so intensiv über Technik und Freiheit diskutiert wird. Überhaupt ist Technik seit Ende des 19. Jahrhunderts ein Reflexionsbegriff.

Gleichwohl ist die Auslassung der Rolle der Technik aus den Freiheitsdiskursen erstaunlich, denn Technik ist zweifellos in einem ganz grundsätzlichen Sinne schon immer eine Möglichkeitsbedingung von Freiheit. Sie

¹³ Freiheit, in: Schischkoff, G. (Hg.), Philosophisches Wörterbuch, Stuttgart 1982, 201–203, hier 202.

¹⁴ Pauen, Freiheit (Fußnote 6), 802.

ermöglicht die Befreiung von Zwängen und Kontingenzen der Natur, von körperlichen Beschränkungen und Grenzen, von lästigen Tätigkeiten, sie schafft Freiheiten, indem sie menschliche Fähigkeiten erweitert und sie eröffnet neue Handlungsoptionen.

So haben einige Technikphilosophen, und damit diejenigen, die von der Technik herdachten (nicht diejenigen, die von der Freiheit herdachten), das Verhältnis von Freiheit und Technik im 20. Jahrhundert reflektiert. Ernst Cassirer schrieb beispielsweise 1930: In der „Frage über Wert und Unwert der Technik [...] geht es nicht um Lust oder Unlust, um Glück oder Leid, sondern um Freiheit oder Unfreiheit“.¹⁵ Er verbindet Technik mithin ganz grundsätzlich mit Freiheit und Unfreiheit.

Günther Anders verknüpfte Freiheit mit dem *homo faber* und betonte, wie andere auch, der Mensch sei das Wesen, das die Freiheit habe, die eigenen Lebensbedingungen zu gestalten. Menschen sind, anders als Tiere, nicht gleichsam „natürlicher Teil“ der Welt:

„Das Tier ist in seiner Welt zu Hause. Der Mensch aber – *homo faber* – muss sich in ihr erst ‚häuslich einrichten‘. D. h.: die Welt [...] wird nicht a priori vorgefunden. Damit er ‚in seiner Welt‘ lebe, schafft und vermittelt sich der Mensch eine andere; er *verändert* die Welt.“¹⁶

Der letzte Satz ist entscheidend: Menschen müssen und haben – anders als die Tiere – die Möglichkeit, die Welt zu verändern. Dies beschreibt eine fundamentale Freiheit, die Freiheit sich die Welt einzurichten, sie zu verändern und sie zu gestalten. Dabei spielt Technik eine zentrale Rolle.

Diese Freiheit der Gestaltung der Welt, die nur dem Menschen eignet, verweist aber zugleich auf Ambivalenzen oder stärker noch auf Aporien des Verhältnisses von Freiheit und Technik: Die Gestaltung der Welt ist eine menschliche Freiheit und sie ist zugleich eine Notwendigkeit, aus der wiederum neue Zwänge und Unfreiheiten erwachsen. Wurde Technik einerseits in diesem grundsätzlichen Sinne thematisiert, so findet sich in den Diskursen

¹⁵ Cassirer, E., Form und Technik (1930), in: Ders., Gesammelte Werke. Hamburger Ausgabe, hg. von B. Recki, Band 17: Aufsätze und kleine Schriften (1927–1931), Hamburg 2004, 139–183, hier 172f.

¹⁶ Anders, G., Die Weltfremdheit des Menschen. Schriften zur philosophischen Anthropologie, München 2018, 186.

um Technik des Weiteren die eingangs schon erwähnte dichotome Gegenüberstellung von Befreiung und Bedrohung, die für viele Technikdiskurse typisch ist. Beide Diskursfiguren sollen im Folgenden näher betrachtet werden. Die Ängste vor einem Freiheitsverlust durch Technik finden sich seit dem 18. Jahrhundert. Sie waren bis in die 1970er Jahre sehr stark und werden in jüngster Zeit erneut thematisiert. Blickt man auf die Freiheitsversprechen, so sind insbesondere drei Stränge zu betonen: die Freiheit von den Übeln des Daseins, die Freiheit vom Körper, die Freiheit vom Anderen.

3. Technik als Bedrohung der menschlichen Freiheit

Für die Vorstellung, dass Technik menschliche Freiheit gefährdet, spielten bis in das 20. Jahrhundert hinein vor allem der Begriff der Maschine und das Maschinenhafte eine zentrale Rolle. Blicken wir beispielsweise in die Enzyklopädien der Frühen Neuzeit, so findet sich, im Krünitz im 18. Jahrhundert, unter dem Stichwort „maschinenmäßig“ der Hinweis „nach Art einer Maschine, d. i. ohne eigene freye Bewegung, ohne eigene vernünftige Einsicht“.¹⁷ Vernunft, die eng mit Freiheit verknüpft war, und selbstbestimmte, freie Bewegung waren das, was der Maschine fehlte und das, was, wie im Krünitz weiter ausgeführt wurde, die Menschen im Unterschied zur Maschine auszeichnet. Menschen waren frei, Maschinen nicht.

Menschliche Selbstbeschreibung des 18. und 19. Jahrhundert betonten stets die menschliche Vernunft und Freiheit, im Vergleich und in Abgrenzung zur Maschine. Jean Paul betonte, den „freien Geistern“ stünden die „knechtischen Maschinen“ gegenüber.¹⁸ Aus dieser polaren Gegenüberstellung der knechtischen Maschinen und der freien Menschen resultierte aber auch die wohlbekannteste Angst, die Maschinen könnten die Menschen verknechten. Die Vorstellung, Maschinen machten die Menschen unfrei, findet sich seit dem 18. Jahrhundert, aber mit der Industrialisierung und vor allem im 19. und in

¹⁷ Krünitz, J. G., Oeconomische Encyclopädie oder allgemeines System der Staats- Stadt- Haus- und Landwirthschaft, 1773 bis 1858, <http://www.kruenitz1.uni-trier.de>, Zugriff 27. März 2021.

¹⁸ Zitiert nach Meyer-Drawe, K., Maschine, in: Wulf, Chr. (Hg.), Vom Menschen. Handbuch Historische Anthropologie, Weinheim/Basel 1997, 732.

der ersten Hälfte des 20. Jahrhunderts prägte diese Angst die Diskurse und wurde zu einem weit verbreiteten Topos der Maschinenkritik. Prominent ist hier Karl Marx, der polemisierte, Menschen würden die Knechte und Anhängsel der Maschine werden. Die mechanisierte Arbeitswelt sah Marx als das Reich der Notwendigkeit, dem er das Reich der Freiheit gegenüberstellte, in dem Menschen selbstbestimmt und frei seien. Ähnliche Debatten finden sich im frühen 20. Jahrhundert im Kontext der Taylorisierung, des Fordismus, schließlich der Automatisierung. Interessanter Weise ist der Diskurs um die Digitalisierung heute ein anderer. Zwar gibt es die Angst vor der Überlegenheit der Maschine, nun in Form Künstlicher Intelligenz, und dem Überflüssigwerden der Menschen, doch die kulturkritische Angst vor der Verknechtung der Menschen im Arbeitsprozess durch Maschinen wich einer Vorstellung oder einer Behauptung von Technik als Assistenz und Partner der Menschen, die die Menschen befreien, ihnen helfen würden.

Im 19. Jahrhundert bis in die 1970er Jahre bestanden jedoch Ängste, Menschen würden bei der Nutzung der Maschinen selbst maschinenmäßig, würden sich also der unfreien, vorgegebenen Arbeitsweise von Maschinen anpassen und verlören damit ihre Selbstbestimmung.

In der Kulturkritik des 20. Jahrhunderts spielte dabei der Begriff der Autonomie eine zentrale Rolle. Autonomie im Sinne der Selbstbestimmung ist, wie erwähnt, seit der Aufklärung zentral für den Freiheitsbegriff. Vor allem in den 1950er und 1960er Jahren entwickelte sich eine Debatte um die „Autonomie“ der Technik.¹⁹ Auch zuvor, in den 1920er und 1930er Jahren, finden sich in der deutschen philosophischen Diskussion ähnliche Gedanken, jedoch unter dem Begriff der „Technik als Schicksal“ oder als Macht. Schicksal und Macht stellen Gegenbegriffe zur Freiheit dar und genau in diesem Sinne wurde Technik interpretiert. Oswald Spengler betonte die Einflusslosigkeit, die Unfreiheit der Menschen gegenüber technischen Entwicklungen: „In der Tat vermögen weder die Köpfe noch die Hände etwas an dem Schicksal der Maschinenteknik zu ändern, die sich aus innerer, [...] Notwendigkeit

¹⁹ Vgl. ausführlicher: Heßler, M., Technik und Autonomie, in: Hirsch-Kreinsen, H./ Karacic, A. (Hg.), Autonome Systeme und Arbeit. Perspektiven, Herausforderungen und Grenzen der Künstlichen Intelligenz in der Arbeitswelt, Bielefeld 2019, 247–274.

entwickelt hat.²⁰ Des Weiteren bediente er sich, wie viele Zeitgenossen, der Herr und Knecht- oder Herr und Sklaven-Metapher. Der „Herr der Welt (also der Mensch) wird zum Sklaven der Maschine“,²¹ so Spengler, und damit unfrei. Ein weiterer typischer Topos im Kontext des Autonomie-Diskurses ist der der Totalität des Technischen. Technik sei zu einer Totalität geworden, der alles untergeordnet werde,²² es gibt demnach keine Freiheit in einer technisierten Welt.

Die Denkfigur ist selbst wiederum totalitär, anthropozentristisch und undifferenziert, denn sie suggeriert Technik mache Menschen stets unfrei. Technik wird hier als Kollektivsingular und in abstrakter Weise behandelt; häufig lag dieser Kritik tatsächlich das Modell der Maschine zugrunde. Des Weiteren fällt auf, dass Freiheitsverlust mit Unterordnung unter Technik verbunden wurde. Menschen, so die Ängste, müssten sich der Maschine unterordnen, verlören ihre Selbstbestimmung, seien ohne Einfluss, ohne Autonomie. Die Angst vor dem Freiheitsverlust verweist deutlich auf grundlegende Vorstellungen von menschlicher Freiheit, die von Menschen als selbstbestimmte, in der Hierarchie der Entitäten obenstehend, konzipiert. So sollte Technik die Menschen eigentlich von Kontingenzen, Zwängen und vom Schicksal befreien, doch schien die Maschine, so die Beschreibungen der ersten Hälfte des 20. Jahrhunderts, zum Schicksal zu werden und damit zum Verlust der Freiheit und der Selbstbestimmung und zur Unterordnung zu führen.

Einerseits wurde in der sozialwissenschaftlichen und historischen Technikforschung diese Denkweise seit den 1970er und 1980er Jahren als Technikdeterminismus kritisiert und die Gestaltbarkeit von Technik sowie die Freiheiten von Nutzer/innen, Technik zu formen, zu gestalten und sie eigenwillig zu nutzen und anzueignen, betont und damit das Narrativ der unfrei machenden Technik stark in Frage gestellt.

Andererseits finden wir ähnliche Ängste und Bedenken auch heute. Sie beziehen sich inzwischen allerdings weniger auf Maschinen der Arbeitswelt

²⁰ Spengler, O., *Der Mensch und die Technik. Beitrag zu einer Philosophie des Lebens*, München 1931, 74.

²¹ Ebd., 75.

²² Ebd., 79.

und das maschinenhaft Werden der Menschen, sondern eher auf neue Medien und sie beschreiben andere Formen der Unfreiheit, nämlich Abhängigkeiten: Abhängigkeiten vom Handy und von sozialen Medien und stellen damit die Frage, ob der Umgang mit ihnen noch selbstbestimmt sei, sowohl hinsichtlich der Zeit, die soziale Medien in Anspruch nehmen, als auch im Hinblick auf soziale Zwänge, beispielsweise zur Selbstdarstellung, die durch die Nutzung sozialer Medien entstehen, oder die jüngst diskutierte „Fear of Missing Out“, also der Angst, etwas zu verpassen, sobald man nicht online ist. Damit wird derzeit eine andere Unfreiheit gegenüber der Technik diskutiert: nämlich die Frage, ob Menschen ihre Selbstbestimmung verlieren, weil sie sich scheinbar freiwillig einer Fülle von digitalen Gadgets unterwerfen, von den Selbstvermessungsgeräten, die ihren Puls, ihre Aktivitäten, ihren Schlaf kontrollieren, aufzeichnen und in Clouds speichern, bis hin zu den sozialen Medien, die, wie erwähnt, Freiheiten der Selbstdarstellung und Kommunikation schaffen, die zugleich aber Zwänge darstellen und mit einem freiwilligen Verzicht auf Freiheit und Selbstbestimmung einhergehen.

In beiden Fällen, den Ängsten der ersten Hälfte des 20. Jahrhunderts als auch gegenwärtigen Bedenken, haben wir es mit der Behauptung einer Umkehrung eines Freiheitsversprechens zu tun. Maschinen sollten, so das Argument, die Menschen entlasten und frei machen, aber es bestand die Angst, dass genau das Gegenteil passiert, dass nämlich die Menschen unfrei und den Maschinen untergeordnet werden.

Historisch zurückgeblickt, zeigt sich, dass die Angst vor Freiheitsverlust sich vor allem auf den Verlust von Freiheit im Sinne von Selbstbestimmung und menschlicher Autonomie bezieht. Die Kernfrage, die also seit dem 19. Jahrhundert diskutiert wird, ist, ob Technik zu mehr oder weniger Selbstbestimmung führt – eine komplexe Frage, die für verschiedene Zeiten, für verschiedene Technologien, für verschiedene soziale Gruppen unterschiedlich zu beantworten ist und für deren Beantwortung auch genau zu fragen ist, was jeweils mit Selbstbestimmung gemeint ist. Wenn beispielsweise in den 1970er Jahren die „Freie Fahrt für freie Bürger“ verteidigt wurde, ist das eine andere Selbstbestimmung als die freie Entfaltung der menschlichen Persönlichkeit in demokratischen Staaten.

4. Die Freiheit *von* – Befreiung von „den Übeln des Daseins“

Bereits Aristoteles sah in Automaten eine Möglichkeit, Sklaven zu ersetzen. Waren Sklaven die Bedingung für die freie Existenz der Nicht-Sklaven, der Freien, so dachte Aristoteles darüber nach, dass die Freien der Sklaven nicht mehr bedürften, wenn Automaten deren Rolle übernehmen könnten.²³ Die Bedingung der Freiheit ist in dieser Figur ein Sklave oder eine Maschine, ein Automat. Hier findet sich die immer wiederkehrende Vorstellung der Technik als Sklave, Knecht oder Diener der Menschen, die ihnen zu freier Tätigkeit verhilft und vor deren Umkehrung, also der Mensch als unfreier Sklave der Maschine, wie gerade gezeigt, so große Ängste bestanden. Maschinen sollten frei *von* Notwendigkeit machen. Für diese Erwartungen an Technik gibt es unzählige Beispiele, häufig waren sie ein zentraler Bestandteil der Legitimation von Technik. Ein kurzer Blick in die Diskurse im Kontext der Arbeitswelt müssen an dieser Stelle zur Exemplifizierung der Denkfigur genügen.

So betonten in den 1950er Jahren Ingenieure, Unternehmen und auch Teile der Politik, dass die Automatisierung die Menschen von lästigen Tätigkeiten, von Monotonie und Eintönigkeit befreien würde, wie sie beispielsweise im Taylorismus herrschten. Menschen würden frei, so war das Argument, für höhere Tätigkeiten. Technik befreie von unangenehmer und anstrengender Arbeit und mache frei für sinnvolle und schöpferische Tätigkeiten, die zur Selbstbestimmung und dem „wahren Menschsein“ führen würden.²⁴ Im Hintergrund steht der bereits erwähnte emphatische Freiheitsbegriff der Aufklärung, also die Freiheit als Entfaltung der menschlichen Persönlichkeit, die durch Technik ermöglicht werden sollte, indem sie frei mache vom Reich der Notwendigkeit.

Weniger pathetisch als in den 1950er Jahren ist heute im Kontext der Digitalisierung die Rede von der Entlastung von Routinetätigkeiten und dem Freiwerden für anspruchsvolle und kreative Arbeiten, was auch mit einer in-

²³ Aristoteles, Politik, Hamburg ⁴1981, 4.

²⁴ Heßler, M., Die Ersetzung des Menschen? Die Debatte um das Mensch-Maschinen-Verhältnis im Automatisierungsdiskurs, in: Zeitschrift für Technikgeschichte, Bd. 82 (2015) 2, 109–136.

interessanten Verschiebung einer Freiheit nicht nur *von* Arbeit, sondern auch hin zu Freiheit für Kreativität und Selbstbestimmung *im* Arbeitsprozess darstellt.

Die Figur einer Befreiung von den Übeln des Daseins wird verknüpft mit dem Versprechen der Freiheit *für* oder *zu*, also einem positiven Freiheitsbegriff – dem Versprechen einer Freiheit zu schöpferischen Tätigkeiten, ob jenseits der Sphäre der Arbeit oder innerhalb des Arbeitsprozesses. Die *Befreiung* ermöglicht Freiheit im Sinne der Selbstbestimmung. Dies gilt auch für nächsten Punkt.

5. Die Freiheit *von* – Befreiung von körperlichen Beschränkungen und Grenzen

Der Philosoph Arnold Gehlen beschrieb den Menschen in Anschluss an Herder als Mängelwesen. „Der Mensch“ so Arnold Gehlen, sei „sinnesarm, waffenlos, nackt.“ Diese Mängel des Menschen machen Technik notwendig, die dann wiederum als „Verstärkertechnik“ genutzt werden kann.²⁵ Technik ist demnach dem von Natur aus schwachen und nicht überlebensfähigen Mängelwesen Mensch ein Instrument, um überhaupt zu überleben, aber – für die Frage nach der Freiheit entscheidend – auch, um über sich selbst hinauszuwachsen. Technik entlastet und überbietet die Fähigkeiten des menschlichen Körpers. Aus der Notwendigkeit, Technik zu nutzen, um zu überleben, wird die eingangs bereits erwähnte fundamentale Freiheit der Menschen, die Welt zu gestalten und die körperlichen Schranken nicht nur zu kompensieren, sondern sie zu sprengen. Gehlen selbst nannte das Flugzeug, aber wir können sämtliche Technologien der Kommunikation und der Mobilität bis zur Raumfahrt hinzufügen, die Freiheiten über Raum und Zeit gewähren und damit Freiheiten ermöglichen, die über die menschliche körperliche „Rohfassung“, wie Günther Anders formulierte, hinausgehen. Schon Ende des 19. Jahrhunderts wurde die Freiheit, schließlich v. a. die individuelle Freiheit, die durch das Fahrrad und vor allem das Automobil ermöglicht wurde,

²⁵ Gehlen, A., Die Seele im technischen Zeitalter. Sozialpsychologische Probleme in der industriellen Gesellschaft, Frankfurt a. M. 2006, 6.

betont. Freiheit meint hier die Be-Freiung vom Körper, die Unabhängigkeit von körperlichen Grenzen, der körperlichen „Reichweite“.

Das Versprechen der Be-Freiung von körperlichen Beschränkungen erreichte vor allem mit dem Computer eine neue Dimension. Der Computer wurde häufig, wie eingangs schon angedeutet, als „Befreiungstechnologie“ gefasst – ganz im Gegensatz übrigens zur Maschine, die mit der starken Konnotation des Mechanischen, und damit Unfreien viel stärker mit Ängsten vor dem Verlust der Freiheit einherging.

Frühe Internetnutzungen der 1990er Jahre betonten euphorisch die Freiheit, die mit der Immaterialität und Körperlosigkeit des Internets einherging. Sherry Turkle feierte, neben vielen anderen, die Möglichkeiten im virtuellen Raum, neue Identitäten, andere Rollen und Figuren auszuprobieren, wie es durch die Nichtsichtbarkeit des Körpers möglich wurde.²⁶ Befreit von den sichtbaren körperlichen Merkmalen, befreit von der eigenen Körperlichkeit entstand eine neue Freiheit, sich in und mittels einer neuen Technologie, einem second life oder einer virtuellen Welt, neu zu erfinden, sich in anderen Rollen zu verwirklichen, zu einem anderen Selbst zu finden. Dieser Freiheitsbegriff steht in der Tradition der Selbstbestimmung, der Emanzipation, der Freiheit von Zuschreibungen durch gesellschaftliche Erwartungen und Interpretationen, aber er meint auch eine Freiheit vom Körper.

Zugleich ist es ein Freiheitsbegriff, der die Vielfalt der Möglichkeiten, der Wahl- und der Handlungsoptionen betont. Wie Ernst Cassirer lange vor dem Internet formulierte: „Die Technik fragt nicht in erster Linie nach dem, was ist, sondern nach dem, was sein kann.“²⁷ Mit anderen Worten: „In der Technik werden Möglichkeiten erprobt und verwirklicht; es wird Neues geschaffen; es werden Grenzen in einem fort erweitert.“²⁸

Edward Snowden betonte in seinen kürzlich erschienen Memoiren das be rauschende Freiheitsgefühl, das in den 1990er Jahren mit dem Internet verbunden war: „die Freiheit, sich etwas völlig Neues auszudenken [...] die

²⁶ Turkle, Sh., *Life on Screen*, New York 1995.

²⁷ Cassirer, *Form und Technik* (Fußnote 15), 176.

²⁸ Recki, B., Ernst Cassirer (1874–1945), in: Hefßler, M./ Liggieri, K. (Hg.), *Technik-anthropologie. Handbuch für Wissenschaft und Studium*, Baden-Baden 2020, 145–151, hier 149.

Möglichkeit des Experimentierens, die neuen Ausdrucksformen“. Er bezeichnete die damalige Online-Welt als „angenehmste und erfolgreichste Anarchie“.29

Während im Kontext der Vorstellung, Technik befreie von lästigen Tätigkeiten stets betont wurde und wird, dass sie die Menschen frei mache für höhere, sinnvollere Tätigkeiten, für Kreativität und Freiheit und dies einem emphatischen Freiheitsbegriff entsprach, der auf die Persönlichkeitsentwicklung zielte und vor dem Hintergrund humanistischer Vorstellungen der Bildung von Persönlichkeit formuliert wurde, so war die Erhöhung der Handlungsoptionen im Kontext digitaler Technologien und die Befreiung von körperlichen Grenzen mit einer spielerischen Freiheit der Selbstfindung, des Ausprobierens, des Experimentierens und der Regellosigkeit verbunden.

Die Befreiung von körperlichen Grenzen geht aber noch viel weiter und verbindet sich auch insofern mit der Freiheit von den „Übeln des Daseins“, als der Körper optimiert und damit eine Befreiung von körperlichen Gebrechen, von Alter und vom Tod anvisiert wird. Eine Form der Selbstbestimmung, die sich gegen das Schicksal, das Unbeeinflussbare von Alter und Tod richtet. Die körperlichen Grenzen zu erweitern, den Körper zu gestalten und selbst über ihn zu bestimmen, ist ein Freiheitswunsch und Freiheitsversprechen, das heute im Kontext der Digitalisierung, des Internets sowie von Praktiken der Selbstvermessung und Selbstoptimierung gesucht und wohl teils gefunden wird. Gleichwohl reicht diese Vorstellung weit zurück bis in das 19. Jahrhundert, vor allem was die Überwindung des Todes mittels Technik angeht. Schon Ende des 19. Jahrhunderts und zu Beginn des 20. Jahrhunderts wurden biotechnologische Methoden erdacht, um den menschlichen Körper loszuwerden oder Embryonen frei von körperlichen Beschränkungen zu züchten, so von Haldane, Bernal, Reade oder H. G. Wells.³⁰

²⁹ Snowden, E., *Permanent Record: Meine Geschichte*, Frankfurt a. M. 2019, 62.

³⁰ Vgl. z. B. Hagemester, M., *Die Neue Menschheit. Biopolitische Utopien in Russland zu Beginn des 20. Jahrhunderts*, Frankfurt a. M. 2005; Heil, R., *Human Enhancement – Eine Motivsuche bei J. D. Bernal, J. B. S. Haldane and J. S. Huxley*, in: Coenen, Chr. u. a. (Hg.), *Die Debatte über „Human Enhancement“*. Historische, philosophische und ethische Aspekte der technologischen Verbesserung des Menschen, Bielefeld 2010, 41–62.

Gerade hier wird etwas sichtbar, was wichtig zu betonen ist. Die Freiheit von lästigen Tätigkeiten und von körperlichen Grenzen sind sehr alte Freiheitswünsche. Sie entstanden nicht erst mit einer neuen Technik, die Menschen gleichsam auf neue Ideen brachte. Technik ermöglicht vielmehr Befreiungen, über die schon lange nachgedacht wurde, die schon lange angestrebt wurden und die heute in ganz neuer, gesteigerter Weise möglich zu werden scheinen. Dies verweist auf die kulturelle Konstruktion von Technik. Technik verstärkt häufig vorhandene Freiheitskonzepte, sie verschiebt und verändert sie aber auch, wenn sich beispielsweise Selbstbestimmung auf die Verfügbarkeit über den Körper, das Alter und den eigenen Tod bezieht.

6. Befreiung *von* – Befreiung vom Anderen

Technik ermöglicht zunehmend eine Freiheit vom Anderen, die Unabhängigkeit oder das nicht Angewiesensein auf andere Menschen. Dies geht einher mit einer Verfügbarkeit, die als Freiheit empfunden wurde und wird, die allerdings viele soziale und ethische Implikationen hat, so die Befreiung vom anderen als einem lästigen Anderen, als einschränkend, als nicht verfügbar nach eigenen Wünschen, das nicht Rücksicht nehmen müssen

Dies verweist auf eine weitere Verschiebung des Freiheitsbegriffs, die wenig mit dem emphatischen Begriff der Freiheit zu tun hat, der an Verantwortung und, wie in der Antike, an das „Gute“ gebunden war. Vielmehr verweist die Befreiung vom Anderen auf eine Effizienz- und Verfügungskultur.

Eine Form der Unabhängigkeit vom Anderen begann bereits mit der Mechanisierung der Arbeit, indem sie Unternehmern die Freiheit gab, Menschen durch Maschinen zu ersetzen. Fahrkartenautomaten, Nahrungsmittelautomaten, das Online Shopping verschaffen die Freiheit, einzukaufen, wann wir möchten. Die Erfüllung von Dienstleistungen ist nicht mehr auf die Präsenz eines Anderen angewiesen.

Ein drastisches Beispiel stellen die derzeit diskutierten Sexroboter dar.³¹ Sexpuppen sind allerdings seit dem 19. Jahrhundert auf dem Markt. Sie wurden als Ersatz für nicht vorhandene Frauen angepriesen und schufen, so das

³¹ Vgl. Devlin, K., *Turned On. Science, Sex and Robots*, London u. a. 2018.

Argument bereits im 19. Jahrhundert, eine Unabhängigkeit von der Notwendigkeit eines lebendigen Gegenübers.³² Die seit den 1990er Jahren auf dem Markt existierende Real Dolls bieten die Möglichkeit, diese Dolls ganz frei zu gestalten, von der Augenfarbe, den Körpermaßen, dem Mund, der Form der Nase bis zur Haarfarbe.

Die eingangs zitierte Freiheit der Gestaltung, die ausschließlich den Menschen zukommt, wird hier zu einer Verfügbarkeit und zur Wunscherfüllung, die eine für menschliche, soziale Beziehungen problematische Freiheit vom Anderen bringt, indem sie zur Befreiung vom Anderen als eigenständigem Gegenüber mit einer eigenen Willensfreiheit wird.

7. Fazit und Ausblick – Konjunktoren und Wandel des Freiheitsbegriffs

7.1. Historischer Wandel des Verhältnisses von Freiheit und Technik

In der Frühen Neuzeit bis in das 20. Jahrhundert hinein wurde Technik, insbesondere die Maschine, häufig mit Unfreiheit assoziiert, mit der Gefahr, dass sie die Menschen verknechte und zu einem Verlust von Freiheit führe. Dies betraf vor allem die Arbeitswelt, während seit Mitte des 19. Jahrhunderts die Mobilitäts- und Kommunikationstechnologien völlig neue, vor allem individuelle Freiheiten versprachen.

Der Computer wiederum wurde vielfach als Befreiungs- und politische Emanzipationstechnologie interpretiert, beispielsweise bereits im Kontext des Automatisierungsdiskurses der 1950er Jahre (Befreiung für höhere Tätigkeiten), vor allem aber in den technikeuphorischen 1990er Jahren, nun als Instrument der Ermächtigung und der Teilhabe. Erst in jüngster Zeit wandelten sich die Debatten in Ängste vor Unfreiheit und vor dem Verlust der Selbstbestimmung.

³² Vgl. Smith, M., *The Erotic Doll: A Modern Fetish*, New Haven/London 2013.

Historisch können wir deutlich Konjunkturen beobachten: Eine weitaus stärkere Debatte um die Kehrseite der Technik als Gefahr des Freiheitsverlustes seit dem 19. Jahrhundert bis in die 1960er bis 1970er Jahre hinein, euphorische Erwartungen an Freiheit durch Technik im Kontext der Digitalisierung, vor allem in den 1990er Jahren und eine hochgradig ambivalente Interpretation der Technik heute.

7.2. Wandel von Freiheitsbegriffen

Einerseits rekurrieren die im Kontext der Technik gemachten Freiheitsversprechungen und -erwartungen auf seit der Antike bestehende Freiheitsbegriffe, die um die Befreiung von den „Übeln des Daseins“ (negativer Freiheitsbegriff) einerseits und um Autonomie, Selbstbestimmung, Selbstverwirklichung, persönliche Entfaltung (positiver Freiheitsbegriff) andererseits kreisen.

Gleichzeitig veränderte, wie gezeigt, Technik Konzepte von Freiheit. Freiheit als Steigerung der Handlungsoptionen oder Freiheit als persönliche Entfaltung wird zu einer Freiheit die eine Gestaltbarkeit, Verfügbarkeit über den Körper, Freiheit von körperlichen Grenzen, Freiheit vom Anderen meint. Freiheit meint dann, Grenzen zu überschreiten und unabhängig vom Anderen zu sein.

7.3. Technik als Freiheit und Notwendigkeit

Technik ist eine Möglichkeitsbedingung der Freiheit der Menschen. Darüber besteht kein Zweifel. Gleichzeitig führt sie zu neuen Unfreiheiten.

Historisch betrachtet erhöhte Technik zweifellos Handlungsoptionen und Gestaltungsmöglichkeiten. Zugleich wurde die Verschränkung von Menschen und Technik immer enger. Dieses Angewiesensein auf Technik, die uns täglich begleitet und mit uns handelt, tritt meist nur im Falle des Nicht-Funktionierens zu tage.

Der Schriftsteller Walter Kempowski beschrieb dies im Kontext seines in den 1980er Jahren erworbenen Computers, der ihm ganz neue Freiheiten des Schreibens brachte, wie er sich freute. Das Echolot-Projekt wäre nicht möglich gewesen, ohne den Computer, Mensch und Computer schrieben das Werk, d. h. nicht nur der Mensch allein. Doch war diese von Kempowski empfundene Freiheit des neuen Schreibens dahin, wenn, was anfangs häufig passierte, der Computer nicht funktionierte. Dann musste Kempowski auf den Olivetti-Mann warten und fühlte sich vorübergehend unfrei.³³

Ernst Cassirer hatte Technik grundsätzlich als Form der Freiheit begriffen. Er wies aber auch darauf hin, dass für jede Technik konkret zu prüfen sei, ob und in welchem Maße sie der Freiheit diene – und ich möchte hinzufügen: was für einer Freiheit.

³³ Vgl. Kempowski, W., Alkor. Tagebuch 1989, München 2003, 72.

Die neue künstliche Intelligenz und ihre möglichen Auswirkungen auf Mensch und Gesellschaft

In diesem Kapitel führe ich zunächst in die „neue“ künstliche Intelligenzforschung und ihre Ergebnisse ein. Dazu gebe ich einen kurzen Überblick über die Entwicklung der künstlichen Intelligenz als interdisziplinäres Forschungsgebiet. Dabei bemühe ich mich um eine Vermittlung auf begrifflicher Ebene, ohne in mathematisch-technische Details zu gehen. Auf dieser Basis spreche ich dann einige der aktuell diskutierten möglichen Auswirkungen auf den Menschen an. Dabei argumentiere ich, dass auch bei der neuen künstlichen Intelligenz die behauptete oder angenommene Nähe zur menschlichen Kognition eine Fehleinschätzung ist. Andererseits gehe ich auf wichtige gesellschaftliche Auswirkungen ein, die weniger überraschend sind, deren Eintreten aber wahrscheinlich ist und zum Teil bereits begonnen hat.

1. Was ist die neue künstliche Intelligenz? Eine kurze begriffliche Geschichte

1.1. Anfänge

Die Idee eines künstlich geschaffenen mensch- oder tierähnlichen Wesens, das eigenständig handeln kann, reicht in die Geschichte der Zivilisation weit

zurück.¹ So hat in der griechischen Mythologie Hephaistos einen metallenen Krieger, den Talos, geschaffen und die aus Lehm geformte Pandora zum Leben erweckt. Der jüdische Mythos des „Golem“ kreist um eine Figur dieser Art. Der Begriff des Roboters stammt aus einem tschechischen Theaterstück der 20er Jahre von Karel Čapek, das bereits den potenziellen Konflikt zwischen Menschen und künstlichen, menschenartigen Wesen zum Thema hatte.

Erste wissenschaftliche Arbeiten zum Thema autonomer Roboter hatten Organismen zum Vorbild.² Die Erzeugung von Verhalten in Organismen wurde dabei in der Begrifflichkeit der Kybernetik abstrahiert: Autonome Agenten oder Roboter erzeugen ihr Verhalten autonom auf der Basis der eigenen Sensorinformation. Sie erzeugen Verhalten, indem Motoren den Körper bewegen. Die Sensoren sind im Körper verankert, so dass die von den Motoren erzeugte Bewegung die Sensoren durch die Umwelt transportieren. Daher die Vorstellung des geschlossenen Regelkreises: Die von Sensorinformation abhängige Bewegung erzeugt wiederum neue Sensorinformation, die wieder neue Bewegung bestimmt. Die Theorie hierzu wurde im und nach dem zweiten Weltkrieg entwickelt, angeregt unter anderem aus militärischen Anwendungen, z. B. der automatischen Verfolgung von Flugzeugen mit Hilfe der ersten Radarsysteme.³

Die künstliche Intelligenz begann als Vision solcher autonomer „intelligenter“ Agenten, die die Welt wahrnehmen und in der Welt handeln würden. Die Intelligenz bestünde dabei in der Erzeugung von Verhalten, das geeignet wäre, bestimmte Ziele unter variablen Umweltbedingungen zu erreichen. Dies ist auch heute noch die grundlegende Zielstellung der künstlichen Intelligenz, so im Vorwort des Standardlehrbuchs der künstlichen Intelligenz: „We define AI as the study of agents that receive percepts from the environment and perform actions.“⁴

¹ Mayor, A., *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology*, Princeton/NJ 2018.

² Grey, C., *An imitation of life*, *Scientific American*, 182(2) (1950), 42–45.

³ Wiener, N., *Cybernetics: or Control and Communication in the Animal and the Machine*, Cambridge/MA 21947.

⁴ Russell, S. J./ Norvig, P., *Artificial intelligence: A modern approach*, Hoboken/NJ 42021.

1.2. Kognition als Informationsverarbeitung

Die Idee der künstlichen Intelligenz war also von Anfang an inspiriert von der Analogie mit biologischen Formen von Intelligenz und insbesondere mit der menschlichen Intelligenz. Der Begriff der Intelligenz selbst nährt sich aus dieser Analogie. Er kennzeichnet zunächst nur ein Verhalten als besonders erfolgreich und zweckmäßig. Bei biologischen Wesen wird der Erfolg dabei evolutionär am Überleben und Weitergeben des Erbgutes gemessen. Bei künstlichen Agenten ist viel weniger klar, was Erfolg ausmacht. Auf jeden Fall stellt man an ein intelligentes System die Erwartung, dass es flexibel auf zunächst unbekannte Umweltbedingungen reagiert und auch inhärent über eine gewisse Komplexität im Handeln verfügt, die etwa über die reine Optimierung fester Reiz-Reaktionsketten hinausgeht.

In der Tat entstand die moderne Auffassung von Intelligenz aus der „kognitiven Revolution“, einer Gegenbewegung zum Behaviorismus, der Verhalten nur aus beobachtbaren Reiz-Reaktions-Verkettungen verstehen wollte.⁵ Solche Verkettungen wurden angesehen als Ergebnis einer einfachen Maximierung der Belohnung, die das Tier durch ein Verhalten erlangt. Die Überwindung der begrifflichen Enge des Behaviorismus bestand in der wiedererstellten Erlaubnis, über innere Zustände eines Agenten und dessen Nervensystems zu sprechen, auch wenn diese nicht direkt der Beobachtung zugänglich sind. Solche kognitiven Zustände wurden gefordert, um mit mathematisch rigorosen Methoden die grammatische Struktur menschlicher Sprachen zu erklären.⁶ Sie waren zur Erklärung einfacher Verhaltensweisen dann nützlich, wenn die Verbindung zwischen Reiz und Verhalten nicht mehr sehr direkt war, etwa weil sie Gedächtnis oder das Erzeugen einer ganzen Abfolge von Handlungen erforderte.

In den 1950er Jahren kam es dann zu einer interessanten Parallelentwicklung. Es entstand gerade die Informatik, also die Lehre vom „Berechnen“ mit digitalen Computern. Diese Lehre stützte sich auf Begriffe wie Daten, die im „Gedächtnis“ des Computers gespeichert sind und Instruktionen, die solche

⁵ Wasserman, E. A., Behaviorism, in: Wilson, R. A./ Keil, F. C. (Hg.), MIT Encyclopedia of the cognitive sciences, Cambridge/MA 1999, 77–80.

⁶ Chomsky, N., Syntactic Structures, Den Haag 1957.

Daten als Eingangsinformation nutzen, um am Ausgang eine „Antwort“ zu erzeugen. Dabei können die Instruktionen, die zu Algorithmen zusammengestellt werden, selbst Daten anlegen, also etwa Zwischenergebnisse ableiten und so mit bisher unvorstellbar komplexen Abhängigkeiten der Antworten von den Eingangsdaten umgehen.

Diese Begrifflichkeit der Informationsverarbeitung wurde nun genutzt, um über menschliche Kognition zu sprechen. Typische kognitive Aufgaben, die man im Labor untersuchen kann, wurden modelliert als Berechnungen der Antwort aus der Eingangsinformation. Im Gegensatz zur Reiz-Reaktionsauffassung des Behaviorismus konnte diese Berechnung aber auf umfangreiches internes Wissen zurückgreifen, das im Gedächtnis gespeichert ist oder von den kognitiven Prozessen selbst erzeugt und verwaltet wird. Ihren Höhepunkt fand diese Entwicklung in mathematischen Modellen der kognitiven Architektur des menschlichen Geistes wie „ACT-R“⁷ oder „SOAR“⁸, die auch formal menschliche kognitive Leistungen als Algorithmen in einer Programmiersprache modellieren.

In dem neuen Gebiet der künstlichen Intelligenz hatten sich zur gleichen Zeit Informatiker das Ziel gesetzt, eine Theorie der „allgemeinen Intelligenz“ zu entwickeln, die sowohl die menschliche Intelligenz emulieren, diese aber auch verallgemeinern würde.⁹ Hier gab es also eine enge Verzahnung der Begrifflichkeit, mit der die menschliche Kognition beschrieben und erklärt werden sollte und die gleichzeitig der Entwicklung künstlicher intelligenter Systeme dienen sollte.

1.3. Das Kernproblem der klassischen KI: wie kommt das Wissen in das System?

In der Informatik führte dieses Forschungsprogramm zu viel Fortschritt, der heute die gesamte Informatik durchdringt und zum Teil die Softwaretechnik

⁷ Anderson, J. R., Rules of the mind, Hillsdale/NJ 1993; Ders., How can the human mind occur in the physical universe, New York 2007.

⁸ Newell, A., Unified Theories of Cognition, Cambridge/MA 1990.

⁹ Ebd.

prägt, die hinter vielen Alltagssystemen steckt. Ein Beispiel ist die objektorientierte Programmierung, die es durch Verkapselung von Information innerhalb von Modulen – oder Klassen – erleichtert, einmal geschriebene Algorithmen in neuem Kontext wiederzuverwenden. In der Theorie menschlicher Kognition wurde ein analoges Konzept diskutiert, die Vorstellung, dass einzelne Module der Informationsverarbeitung nicht kognitiv penetrierbar sind, so dass das Bewusstsein keinen Zugang zu den inneren Zuständen solcher Module hat und deren Inhalt auch nicht verändern,¹⁰ – ein weiteres Beispiel für die parallele begriffliche Entwicklung der kognitiven Psychologie und der künstlichen Intelligenzforschung.

Zu dem erhofften Durchbruch zu einer allgemeinen künstlichen Intelligenz kam es jedoch nicht. Ein Kernproblem war die Schwierigkeit, Weltwissen in ein solches System so einzubringen, dass dieses relevantes Denken oder Handeln entwickeln konnte. Jedes Stück Wissen musste als Regel abstrahiert und konkret in einen Algorithmus gegossen werden. Selbst einfachste Dinge – oder gerade einfachste Dinge – erforderten komplexe Analysen. Ein Glas Milch eingießen? Dazu muss man wissen, dass Milch im Kühlschrank ist, wo der Kühlschrank sich befindet, dass man ihn öffnen muss, um an die Milch zu kommen, wie man die Tür aufmacht, wie die Milchflasche aussieht, wie man sie anfasst damit sie nicht aus dem Greifer rutscht etc. – jeweils eine Doktorarbeit pro konkreter Handlung.

Dieses Problem des fehlenden Weltwissens war besonders schwierig anzugehen im Bereich der Wahrnehmung. In der visuellen Wahrnehmung z. B. gewinnen wir aus sensorischen Signalen des Auges Information über die Beschaffenheit der Umgebung, also über die Objekte, auf die wir unsere Handlungen richten werden. Da uns dies so leichtfällt, erscheint es selbstverständlich. Tatsächlich ist es ein unterbestimmtes Problem. Wir besitzen nur zweidimensionale Bildinformation und müssen dreidimensionale Information über Objekte und deren Abgrenzung gegenüber dem visuellen Hintergrund ableiten, die auch noch unabhängig sein soll davon, wie gerade die Objekte beleuchtet werden oder welche Seite sie uns zuwenden. Dieses Problem war jahrzehntelang praktisch unlösbar. Roboter, die eigenständig

¹⁰ Fodor, J., *The Modularity of Mind*, Cambridge/MA 1982.

handeln sollten, waren effektiv blind. Sie waren nicht nur blind, sie konnten auch nichts hören, das heißt, aus auditorischen Signalen keinen Aufschluss über die Umwelt oder die Absichten eines menschlichen Sprechers gewinnen. Und sie hatten keinen Tastsinn, konnten also auch durch Berührung nichts über die Folgen ihrer Bewegungen erschließen.

Der Weg zur künstlichen Intelligenz hat daher eine lange Durststrecke durchstanden. Es gibt immer wieder die Hoffnung, dass ein Durchbruch durch die schiere Akkumulation von Regelwissen erzielt werden kann – zuletzt das Watson System von IBM für abstrakte Domänen wie Medizinwissen –, aber im Moment ist noch immer nicht klar, ob solche Systeme echte Autonomie erlangen werden. Nützlich sind sie in wohl definierten Domänen, in denen sie auf bereits formatierten Daten aufbauen können.

1.4. Zurück zum Reflex, probabilistische Methoden

Die jetzt als „klassisch“ bezeichnete künstliche Intelligenzforschung ist gescheitert, wenn man sie an ihrem Ziel misst, autonome Agenten zu entwickeln, die sich in ihnen a priori unbekanntem Umwelten zielorientiert verhalten. Dieses Scheitern gab einer Reihe von alternativen Zugängen Auftrieb. Besonders sichtbar wurde der Ansatz der sogenannten „behavior based robotics“, den Rodney Brooks vom MIT prominent vertreten hat.¹¹ Grob zusammengefasst war die Idee, die Analogie zu Organismen wieder auf einer weniger abstrakten Ebene zu nutzen. Elementare Verhalten wurden in etwa so wie Reflexe oder feste Verhaltensmuster mit ihren eigenen sensorischen Eingängen und dedizierten motorischen Ausgängen entworfen. Solche Verhalten konnten sich gegenseitig ein- oder ausschalten, oder auch „subsumieren“, wo also ein Verhalten ein anderes nutzte. Relativ simple robotische Fahrzeuge konnten so mit recht einfachen Sensoren dazu gebracht werden, sich zumindest einmal vernünftig zu fortzubewegen. Das Verhalten war rein reaktiv, konnte aber in geeigneten Umwelten durchaus sinnvolle Aufgaben erfüllen. Ein sichtbares Produkt waren die ersten Staubsaugerroboter.

¹¹ Brooks, R. A., Intelligence without representation, *Artificial Intelligence* 47 (1991), 139–160.

Rodney Brooks selber war Mitgründer der Firma iRobots, deren Roomba Staubsauger diese Ideen teilweise nutzten.

Methoden aus der Wahrscheinlichkeitsrechnung machten es möglich, die Information von solchen einfachen Sensoren – die meist Distanzen schätzten – systematisch zu akkumulieren, um so Karten der Umwelt aufzubauen, die autonomen Vehikeln auch die Planung von Bewegungen erlauben. Diese sogenannten probabilistischen Ansätze¹² haben zu einem Innovationsschub in der autonomen Robotik geführt,¹³ der das autonome Fahren in den Bereich des Möglichen gebracht hat, wie der Leser sicherlich aus der Tagespresse weiß.

1.5. Neuronale Methoden

In der menschlichen Kognitionsforschung gab es ebenfalls eine Gegenbewegung zur „klassischen“ Theorie der Informationsverarbeitung. Der ursprüngliche Anstoß kam von dem Wunsch, die Erklärung kognitiver Leistungen mit dem, was wir über das Gehirn wissen, in Einklang zu bringen. Ganz offensichtlich gibt es im Gehirn keinen Computer mit zentraler Recheneinheit und frei zugänglichem Speicher. Es gibt auch keine zentrale Uhr, die in hohem Takt Berechnungen auslöst. Das Gehirn besteht aus Neuronen, die durch synaptische Verbindungen ein gigantisches Netzwerk bilden. Die Prozesse in den Neuronen sind eher langsam auf einer Zeitskala von zirka einer Millisekunde, während moderne Computer im Bereich von Gigahertz rechnen, also 10^9 Takte pro Sekunde, oder 0,000001 Millisekunden pro Takt, unglaublich viel schneller als das Nervensystem! Gängige Computer haben viele Gigabyte Arbeitsspeicher – meiner zum Beispiel 16 Gigabyte, also zirka 10^{10} Byte. Das kommt der geschätzten Anzahl von Neuronen im Gehirn von zirka 10^{12} schon nahe. Dabei ist das Gehirn aber unglaublich viel dichter und komplexer vernetzt. Während die meisten Bits nur mit einem Datenbus verbunden sind, hat das typische Neuron im Gehirn circa 10000 Verbindungen zu anderen Neuronen. Im Schnitt sind nur 5 Synapsen

¹² Thrun, S., Probabilistic algorithms in robotics, *AI Magazine*, 21(4) (2000), 93–109.

¹³ Thrun, S./ Burgard, W./ Fox, D., *Probabilistic Robotics*, Boston/MA 2005.

erforderlich, um eine Verbindungen von jedem beliebigen Neuron zu einem beliebigen anderen zu finden.¹⁴ Es ist also offensichtlich, dass Gehirne ganz anders aufgebaut sind und daher auch ganz anders funktionieren müssen als Computer.

Gehirne entwickeln sich, und zwar einerseits durch Wachstum im Embryo und Baby, in der Kindheit und der Jugend, andererseits durch das Verändern der synaptischen Verbindungen zwischen Neuronen beim Lernen. Wir lernen ein Leben lang. Das ist durchaus aufwändig. Die Muttersprache lernt man über Jahre, ebenso wie Lesen, Schreiben und Rechnen. Gehirne können Wissen also nicht „herunterladen“ wie Computer.

Die Gegenbewegung zur Computeranalogie setzt auf neuronale Netze als Basis einer Theorie der Intelligenz. Modelle solcher neuronalen Netze wurden schon seit den 1950er Jahren untersucht, mit verschiedenen Ansätzen, die teilweise auch wieder verworfen wurden.¹⁵ Ein Problem war, diese Netze mathematisch zu analysieren. Dies gelang ironischerweise erst, als durch digitale Computer numerische Simulationen in großem Umfang möglich wurden. Das zweite Problem waren die vielen Parameter dieser Modelle, die die Platzierung und Stärke der synaptischen Verbindungen im neuronalen Netz beschreiben. Das werden schnell gigantische Mengen von Parametern, wenn die Anzahl der Neurone wächst. Die möglichen Verbindungen zwischen zwei Schichten von nur je 100 Neuronen haben schon 10000 Parameter. Solche Parameter so einzustellen, dass das Netzwerk eine sinnvolle Funktion hat, ist von Hand nicht zu schaffen. Stattdessen wurden Algorithmen erfunden, die diese Parameter automatisch einstellen, in einer Form von Optimierung, die in der Theorie neuronaler Netze „Lernen“ genannt wird, obwohl die Analogie zu Lernmechanismen in Nervensystemen meist eher lose ist. Diese Optimierung wird auch „maschinelles Lernen“ genannt und beruht darauf, die Funktion eines neuronalen Netzes anhand von Beispieldaten so anzupassen,

¹⁴ Herculano-Houzel, S., The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost, *Proceedings of the National Academy of Sciences of the United States of America*, 109 (SUPPL.1) (2012), 10661–10668, <https://doi.org/10.1073/pnas.1201895109>.

¹⁵ Medler, D. A., A Brief History of Connectionism, *Neural Computing Systems* 1 (1989), 61–101, <https://doi.org/10.1002/9781118266885.ch2>.

dass die gewünschten Antworten gegeben werden. Das Netzwerk kann dann sinnvolle Antworten auch für Daten geben, die nicht beim Optimieren genutzt wurden. Zum Beispiel trainiert man ein Netzwerk mit vielen Bildern von handgeschriebenen Ziffern so, dass sie die entsprechende Ziffer als Antwort geben. Das Netzwerk kann dann auch neue Exemplare von handgeschriebenen Ziffern erkennen.

Die mathematischen Grundlagen dieser Methoden kristallisierten sich ab den 1980er Jahren.¹⁶ Eine Kette von Innovation hatte gegen Ende des letzten Jahrhunderts zu einer gewissen Reife dieser Modelle geführt. Diese wurden einerseits als „Konnektionismus“ zur Erklärung experimenteller Befunde in der Kognitionswissenschaft herangezogen, andererseits zunehmend auch als technische Modelle für Aufgaben der künstlichen Intelligenz genutzt, gerade für die künstliche Wahrnehmung, als dem Erkennen von Objekten auf der Basis von Bildern.

In den letzten 10 Jahren gab es einen Durchbruch in den technischen Anwendungen dieser Netzwerke.¹⁷ Auch hierbei war wieder ironischerweise die digitale Computertechnik entscheidend. Durch die zunehmende Rechenkraft der Rechner konnten immer größere neuronale Netzwerkmodelle simuliert werden, bis hin zu den sogenannten „deep networks“ mit vielen Schichten von Neuronen. Die explodierende Anzahl von Parametern dieser Netzwerke konnte anhand der unglaublichen Menge von Beispieldaten, die über das Internet zugänglich wurden, optimiert – oder „gelernt“ – werden. Kein Zufall, dass diese Forschung vor allem in Internetfirmen wie Google, Facebook und Microsoft oder bei Chipherstellern wie Nvidia oder Intel stattfindet. Solche Systeme erbringen seit einigen Jahren in speziellen Anwendungen übermenschlich gute Leistungen. So können solche Modelle

¹⁶ Aggarwal, C. C., *Neural Networks and Deep Learning – A Textbook*, Zürich 2018.

¹⁷ Schmidhuber, J., *Deep learning in neural networks: An overview*, *Neural Networks* 61 (2015), 85–117, <https://doi.org/10.1016/j.neunet.2014.09.003>; Lecun, Y./Bengio, Y./Hinton, G., *Deep learning*. *Nature*, 521 (2015), 436–444, <https://doi.org/10.1038/nature14539>.

Verkehrszeichen besser erkennen als Menschen¹⁸ und schlagen den Schach- oder Go-Weltmeister.¹⁹

1.6. Die neue künstliche Intelligenz und data science

Zunehmend werden solche neuronalen Modelle mit den anderen Methoden der künstlichen Intelligenz verknüpft, vor allem mit den probabilistischen Methoden und natürlich mit flexiblen, modularen Softwaresystemen. Die neuronalen Methoden lösen dabei zumindest zu einem Teil das Problem des Weltwissens. Sie ermöglichen den Systemen nämlich die Ankoppelung an die Welt durch Sensoren und das Entdecken von Struktur in diesen Sensordaten. Solche Struktur kann man als Wissen über die Umwelt auffassen.

Dies führt bereits jetzt zu zahlreichen Produkten, wie Systemen der Fahrerassistenz, automatischen Überwachungssystemen, Biometrik zur Identifikation von Individuen, und vielem mehr. Zusammen mit anderen Methoden des maschinellen Lernens und der „data science“, auf die ich hier nicht näher eingehen kann, gelingt heute die automatische Erzeugung kognitiver Leistungen überall dort, wo sehr große Mengen von Beispieldaten zur Optimierung dieser hoch-parametrischen Systeme zur Verfügung stehen. Dabei können Daten auch rein explorativ, wie im „data mining“ oder im unüberwachten Lernen durchforstet werden. Dies befähigt solche Systeme zu Rückschlüssen über Vorgänge oder Verhaltensweisen, aus denen die Daten entstanden sind. So ist es zum Beispiel sogar mit ganz klassischen statistischen Methoden möglich aus dem Einkaufsverhalten von Kunden – besonders wenn durch eine Kundenkarte verschiedene Einkäufe miteinander verbunden werden können – sehr spezifische Schlüsse zu ziehen. Berühmt ist die Fähigkeit, die Schwangerschaft und den ungefähren Zeitpunkt der Entbindung von Kundinnen vorherzusagen.

¹⁸ Stallkamp, J./ Schlipfing, M./ Salmen, J./ Igel, C., Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition, *Neural Networks* 32 (2012), 323–332, <https://doi.org/10.1016/j.neunet.2012.02.016>.

¹⁹ Botvinick, M./ Ritter, S./ Wang, J.X./ Kurth-Nelson, Z./ Blundell, C./ Hassabis, D., Reinforcement Learning, Fast and Slow, *Trends in Cognitive Sciences* 23(5) (2019), 408–422, <https://doi.org/10.1016/j.tics.2019.02.006>.

Nur die eigentliche Robotik leistet hartnäckig Widerstand. Hier gibt es noch fundamentale Schwierigkeiten auf dem Weg zum autonomen Roboter.

2. Wie nahe kommt die neue KI der menschlichen Kognition?

Die Wahrnehmung, dass künstliche intelligente Systeme dank der neuen Durchbrüche beim maschinellen Lernen dem Traum – oder Alptraum – echter autonomer Agenten nahekommen oder kommen können, möchte ich zuerst einmal kritisch hinterfragen. Dann erst möchte ich auf die absehbaren Stärken der neuen Verfahren eingehen und deren mögliche Auswirkungen auf wirtschaftliche und gesellschaftliche Prozesse diskutieren.

Die erste Frage, die man sich vielleicht stellen sollte, ist wahrscheinlich, ob die kritischen Probleme, die den ersten Anlauf zu künstlichen intelligenten Systemen gebremst haben, von der neuen KI gelöst werden. Dies war vor allem das Problem des Weltwissens, im Englischen auch als „background knowledge“ bezeichnet. Neuronale Netze haben hier durchaus eine gewisse neue Qualität gebracht, indem sie es ermöglichen, Begriffe mit sensorischem Datenmaterial zu verknüpfen. Ich kann zum Beispiel einem Computersehsystem sehr viele Filmausschnitte oder Fotos von Verkehrsschildern zeigen, jeweils mit der richtigen Bezeichnung des Schildes, und dem System so beibringen, Verkehrszeichen in Videodaten zu erkennen. Ich muss die Daten unter hinreichend vielen Bedingungen erheben, also bei Nebel und Schneefall, in der Dämmerung, nicht im Abblendlicht eines Fahrzeugs, teilweise verdeckt von Ästen, etc. Habe ich solche Daten, dann kann das System sogar zu übermenschlich gute Erkennungsleistungen fähig werden.²⁰

Viele der sehr sichtbaren Erfolge dieser Vorgehensweisen kommen aus Konstellationen, in denen solche Daten „billig“ sind, sowohl im Sinne geringer wirtschaftlicher Kosten als auch im Sinne von geringem theoretischem Aufwand in der Aufbereitung der Daten. Die besonders attraktiven Fallbeispiele sind immer Daten aus dem Internet. Hier gibt es unfassbar viele Bilder

²⁰ Vgl. Stallkamp et al., Man vs. computer (Fußnote 18).

und Videos, die oft mit Text gepaart sind, der die Bilder oder Filme beschreibt. In diesen kann man automatisch nach Bezeichnungen suchen und so Lernbeispiele für die Objektklassifikation und sogar die Szenenerkennung gewinnen.

Aber in Daten aus der realen Welt ist das schon schwieriger. Ich kann Millionen von Kilometern auf Autobahnen fahren und dabei Bilddaten sammeln, ohne dass ich relevante Lernbeispiele über Landstraßen ohne Markierung und mit scharfen Kurven gewinne. Seltene Beispiele sind unter Umständen dennoch besonders wichtig. Bilder oder Filmabschnitte, die kurz vor oder während eines Unfalls anfallen, sind sehr viel wichtiger als die Millionen von Bildern, die beim normalen Fahren anfallen. Das maschinelle Lernen muss aber speziell gesteuert werden, um solchen seltenen Daten Bedeutung und Gewicht zu geben. Und in vielen Fällen sind solche kritischen Daten auch gar nicht in der notwendigen Vielfalt zu erheben.

Von daher bieten die Techniken der neuen KI zwar in bestimmten Fällen neuen Zugang zu Weltwissen, in anderen Fällen ist dies jedoch nur sehr beschränkt der Fall. Diese Methoden verschieben das Problem der Formalisierung von Weltwissen in ein Problem des Kuratierens von Daten. Diese Verschiebung ist keineswegs immer erfolgversprechend.

Ein Beispiel, das gerade für gesellschaftliche Fragen interessant erscheinen mag: Kann ich mit Methoden des maschinellen Lernens aus Überwachungsvideos automatisch Information erlangen über Handlungsweisen von Menschengruppen? Kann ich zum Beispiel den Beginn einer Schlägerei vorhersagen oder den Einsatz einer Massenpanik bei einer Veranstaltung? Wir haben uns in einem Projekt einmal etwas Ähnliches vorgenommen und dabei entdeckt, dass es hier inhärente Begrenzungen gibt. Die Menge an relevanten Beispielen wäre selbst unter Ausräumung der rechtlichen und praktischen Randbedingungen extrem klein im Vergleich zur inhärenten Variabilität solcher Daten.²¹ Man kann also sagen, dass die neuen Methoden der KI in einzelnen Fällen interessante Speziallösungen liefern, es aber

²¹ Horn, D./ Houben, S./ Schöner, G., Erste Ansätze zur automatischen Erkennung von Gruppenverhalten mithilfe des Computersehens, in: Reichertz, J./Keysers, V. (Hg.), *Emotion-Eskalation-Gewalt: Wie kommt es zu Gewalttätigkeiten vor, während und nach Fußballspielen?*, Weinheim/Basel 130–147.

andererseits relevante Probleme der Gewinnung von Weltwissen gibt, in denen diese Methoden nicht greifen.

Noch begrenzter ist der Beitrag der neuen Methoden auf der motorischen Seite. Will ich ein künstlich intelligentes System nicht nur dazu befähigen, Ereignisse und Objekte in der Umwelt zu erkennen, sondern auch auf Objekte der Umwelt durch robotisches Handeln einzuwirken, dann muss ich letztendlich den Roboter in Kontakt mit Objekten bringen. Das ist aber weiterhin eines der schwierigsten Probleme der autonomen Robotik. Selbst die ausgeklügelte Regelungstechnik der modernen Robotik versagt, wenn Kontakt mit nicht gut bekannten oder modellierten Objekten auftritt, also etwa wenn deren Form und Oberflächeneigenschaften nicht gut bekannt sind oder deren Verformbarkeit bei mechanischem Kontakt nicht gut modelliert ist. In der Industrierobotik ist Kontakt immer sehr begrenzt. Materialbearbeitung wie Umformung, Bohren oder Polieren geschieht überwiegend mit Spezialmaschinen, nicht mit frei programmierbaren Robotern, und schon gar nicht mit autonomen Robotern, die Objekte selbst wahrnehmen müssen. Man muss sich nur einmal ansehen, wie gut – oder schlecht – Roboter autonom handeln – zum Beispiel in der berühmten „DARPA challenge“, die man im Internet nachsehen kann, wenn man nach „Darpa fails“ sucht. Dies wäre damit zu vergleichen, wie gut künstlich intelligente Systeme Schach oder sogar Go spielen – siehe oben. So wird deutlich, dass die physische Wechselwirkung mit Objekten in der Welt ein bisher kaum gelöstes Problem ist.

Ein Teil dieses Problems ist auch, wie wir mit solchen autonomen intelligenten Systemen als Menschen kommunizieren wollen. Wir stellen uns ja vor, dass wir einem solchen System Befehle geben auf einer hohen Ebene, unter Benutzung menschlicher Sprache, vielleicht noch von Gesten oder dem Vorzeigen von Objekten. Das gelingt aber weiterhin nicht. Anstatt einem Küchenroboter einfach nur einmal ein Omelett vorzukochen, braucht es eine ganze Doktorarbeit, um eine Teilhandlung zu realisieren.²² Das liegt dann immer noch am fehlenden Weltwissen, also den vielen impliziten und graduierten Randbedingungen für solches Handeln.

²² Beetz, M./ Klank, U./ Kresse, I./ Maldonado, A./ Mösenlechner, L./ Pangercic, D./ Rühr, Th./ Tenorth, M., Robotic Roommates Making Pancakes, in: 11th IEEE-RAS International Conference on Humanoid Robots, Bled, Slovenia 2011.

Ohne eine effiziente Schnittstelle zum menschlichen Nutzer kann man dann doch vielleicht eher seinen Hund auf neue Assistenzaufgaben trainieren als einen Assistenzroboter. Wir wissen, wie aufwendig und begrenzt die Leistungen sind, die man einem Tier antrainieren kann. Dies liegt gerade auch daran, dass solche Tiere kein Verständnis der Aufgabe und deren Lösung haben auf dem hohen Niveau von Abstraktion, das wir durch Sprache erreichen.

Hierdurch wird auch wieder deutlich, dass die Bezeichnung „lernende Systeme“ für solche KI-Systeme irreführend ist. Menschen und andere Tiere lernen aus Erfahrung, also aus ihrem eigenen Handeln in einer Umwelt, die sie aktiv explorieren. Das maschinelle Lernen kann man eher als Optimierung oder Parameteranpassung beschreiben, bei denen also das Lernziel entweder explizit als Fehlersignal geliefert wird, oder implizit durch die Auswahl der Datenbeispiele definiert wird.

In gewissem Sinne emuliert solches Lernen das, was im Laufe der Evolution als Optimierung der Funktion in Organismen geschehen ist. Oder es beschreibt das Ergebnis eines Entwicklungsprozesses. Wenn dem so ist, dann ist es nicht weiter verwunderlich, dass nur in engen Domänen mit eng umgrenzten Begrifflichkeiten solche „brute force“-Lernmethoden erfolgreich sind.

Es ist nicht das erste Mal in der Geschichte der KI, dass solche verführerischen Namensgebungen eine Analogie zum menschlichen Handeln andeuten, die eigentlich so nicht wirklich besteht. Das war auch der Tenor der Kritik an Begriffen der jetzt „klassischen“ KI, ja schon am Begriff der „intelligenten Systeme“ selbst.

Besonders drastisch aber weichen die erreichbaren kognitiven Leistungen der Systeme künstlicher Intelligenz ab von den Erwartungen gerade auch derer, die hier mögliche katastrophale Szenarien bis hin zur „Machtergreifung“ durch solche Systeme sehen – wie etwa in reißerischen Aussagen des Technologieunternehmers Elon Musk, von denen später noch die Rede sein wird. Schon rein begrifflich haben KI-Systeme bisher keinerlei Form von Motivation, Antrieb, Grundbedürfnissen und mithin auch keinen „Willen zur Macht“. Dies ist zunächst auch deshalb der Fall, weil wir die Antriebe von Menschen und anderen Tieren theoretisch nur sehr begrenzt verstehen und

deshalb auch nur begrenzt mathematisch formalisieren können. Das sogenannte Verstärkungslernen – auf Englisch „reinforcement learning“ – kratzt hier nur an der Oberfläche, indem es bestimmte Handlungen mit einem Verstärkungssignal belohnt. Vom „instrumental conditioning“ von Tieren, woher diese Ideen stammen, wissen wir, wie beschränkt das Repertoire solcher belohnungsorientierten Handlungen ist. Das ist ja ein Grund dafür, warum es so schwer ist, seinem Hund Kunststückchen beizubringen. Menschen werden angetrieben durch viel allgemeinere, aber auch tiefere Anreize.²³ Dazu gehört zum Beispiel auch der Wunsch, uns den anderen gegenüber mitzuteilen, also freiwillig und ohne Erwartung einer Belohnung Information auszutauschen.²⁴ Wir beginnen gerade erst, die zugrundeliegenden Fragen in den Neurowissenschaften anzugehen.²⁵

3. Gesellschaftliche Herausforderungen aus der neuen KI

Diese Kritik an den Begrifflichkeiten der neuen KI hat eine gewisse Ähnlichkeit zur früheren Kritik an der klassischen KI: Zu viel wurde versprochen, die Analogie zum Menschen wurde überbetont und Erwartungen geweckt, die nicht annähernd erfüllt werden konnten. Ähnlich ist aber auch der mögliche Fehler im Umkehrschluss. Die klassische KI hat ja durchaus sehr viel erreicht und enorme Auswirkungen auf die Informatik gehabt. Sie hat nur nicht das übersteigerte Ehrgeizige erreicht.

Auch die neue KI erreicht sehr viel, nur nicht die Emulierung menschlichen, motivierten, autonomen Handelns in realen Umwelten. Was die neue KI besonders gut kann, ist, überraschend starke Schlüsse aus scheinbar unstrukturierten Daten zu ziehen. Dabei ist es besonders wichtig zu erkennen, dass diese Daten eben nicht in Datenbanken formalisiert sein müssen. Datenbanken gibt es seit über 50 Jahren und sie haben die erste Phase der automatischen Datenverarbeitung angetrieben. Von daher ist der Begriff der

²³ Tomasello, M., *A natural history of human morality*, Cambridge/MA 2016.

²⁴ Tomasello, M., *A natural history of human thinking*, Cambridge/MA 2014.

²⁵ Feldman Barrett, L., *How emotions are made: The secret life of the brain*, Boston/MA 2017.

„Digitalisierung“ irreführend. Denn dass Daten „digitalisiert“, also in computer-gerechte Formate gebracht werden, ist gerade nicht was die neuen Methoden auszeichnet.

Vielmehr ist deren Stärke, dass sie Schlüsse ziehen aus inhärent analogen Daten, die also noch nicht begrifflich aufbereitet sind, nicht strukturell modelliert wurden, vielleicht sogar multi-medial sind, aus ganz unterschiedlichen Prozessen stammen und einfach zusammengebracht wurden. Natürlich werden die Verfahren auf Computern umgesetzt und dazu die Daten auch in binärer Form gespeichert. Aber diese triviale Form der Digitalisierung ist nicht das neue Element. Digitalisierte Fotos, Filme oder Musik sollen sich ja gerade nicht – oder nur vernachlässigbar – von den analogen Originalen unterscheiden.

Den Methoden des maschinellen Lernens gelingt es dann, wenn auch in eng umgrenzten Domänen, echte Klassifikationsleistungen auf solchen Daten zu erbringen. Sie können beispielsweise Personen als kategorisch verschieden voneinander erkennen, ein Foto als Hans Müller und nicht als Fritz Schulze klassifizieren, und zwar bei vielen verschiedenen Fotos, die sich als Bilder deutlich unterscheiden, aber alle Hans Müller abbilden. Oder sie erkennen aus einer Zeitreihe von Bildern einer Kamera im Auto, ob man in eine Kreuzung hineinfahren darf oder nicht. In beiden Fällen ist das diskrete an der Datenverarbeitung eine makroskopische Entscheidung: Hans oder Fritz, fahren oder nicht. Dass die Daten mikroskopisch digital vorliegen, ist zwar technisch notwendig, aber nicht das Wesentliche.

Diese Form der digitalen Verarbeitung analoger Daten durch Methoden des maschinellen Lernens erschließt neue Funktionen und Leistungen. Diese können eine neue Welle auslösen, in der menschliche Leistungen durch maschinelle ersetzt werden: automatisch Texte erstellen, automatisch Prognosen erarbeiten, automatisch relevante Urteile zu einem Sachverhalt finden, automatisch entscheiden, ob ein Kreditantrag ein Risiko impliziert und vieles andere mehr. Diese Stärken treten immer auf, wenn es um relativ klar umrissene Fragen geht, es dazu viele Beispieldaten gibt, und die entsprechende menschliche Arbeit umfangreich und relativ eintönig ist.

Die neuen technischen Möglichkeiten werden enorme wirtschaftliche Auswirkungen haben. Hierin liegt wahrscheinlich die größte und unmittelbarste

Wirkungskraft der neuen KI, wie mehrfach sehr deutlich von Jack Ma, dem Gründer des chinesischen Internetriesen Alibaba argumentiert wurde. Die öffentliche Debatte zwischen Jack Ma und Elon Musk im August 2019 thematisierte genau den Kontrast zwischen der futuristischen Vision einer neuen KI, die „die Macht ergreift“ und diesen konkreten wirtschaftlichen Fragen. Gerade auch gehobene Tätigkeiten sind von den neuen Möglichkeiten bedroht, denn die neue KI geht gut mit Daten um, aber nicht mit echter Motorik, echtem Handwerk. Die Effizienz von Rechtsanwälten kann durch automatische Textsuche und auch Textgenerierung so erhöht werden, dass die Nachfrage nach Rechtsanwälten sinkt, was sich aktuell bereits abzeichnet – dies vor allem im angelsächsischen Rechtssystem, weil die dort übliche umfangreiche Suche nach einschlägigen Urteilen automatisiert werden kann. Bankkaufleute und Versicherungssachbearbeiter können ihre Effizienz ähnlich steigern – mit entsprechenden Auswirkungen auf die Anzahl der benötigten Arbeitskräfte. Dieses Muster zieht sich durch viele Branchen. Die möglichen Auswirkungen auf den Arbeitsmarkt treten erst verzögert auf, da der Aufbau der neuen Techniken zunächst mehr Personal erfordert. Erst wenn die Techniken gut etabliert sind, werden Kostenreduktionen durch entsprechenden Abbau konkretisierbar. Dieser Prozess fügt sich kontinuierlich an die bereits stark umgesetzten Wellen von Digitalisierung an. So haben Reisebüros einen wichtigen Teil ihres Geschäfts durch die erleichterte Online-Buchung des Endverbrauchers verloren. Nun kann sogar die individuelle Beratung automatisch durch ein KI getriebenes Dialogsystem geleistet werden.

Andererseits entstehen durch die neuen technischen Möglichkeiten auch neue Produkte und Services, die wiederum Wachstum erzeugen mögen. Wie die Bilanz der gegenläufigen Trends am Ende aussieht, ist nicht einfach vorherzusagen. Es kann sogar sein, dass es im Endeffekt zu einer Erhöhung der Nachfrage nach qualifizierten Mitarbeitern kommen wird. Die Herausforderung ist aber vor allem der sehr schnelle Wandel, noch stärker beschleunigt als in früheren Wellen der Digitalisierung. Diese hohe Änderungsrate erzeugt Verwerfungen, sowohl in der Anpassung von Unternehmen und Märkten als auch in der Ausbildung und Mobilität von Arbeitskräften. Dies wird auch durch die Globalisierung kompliziert, die die rasche Verschiebung wirtschaftlicher Aktivität zwischen Nationen und Handelsblöcken ermöglicht.

Jenseits der wirtschaftlichen Auswirkungen gibt es weitere konkret fassbare gesellschaftliche Herausforderungen. Prominent diskutiert wird das Problem des „Bias“, den Algorithmen des maschinellen Lernens institutionalisieren können. Es ist eher die Ausnahme, dass die Daten, die für das maschinelle Lernen genutzt werden, gleichmäßig verschiedene Untergruppen der Bevölkerung oder Untergruppen von Situationen/Fallbeispielen repräsentieren. Gibt es beispielsweise deutlich mehr Männer in einem bestimmten Beruf, dann lernt ein automatisches Übersetzungsprogramm, Begriffe mit uneindeutigem grammatikalischem Geschlecht der Quellsprache in das grammatikalisch männliche Geschlecht der Zielsprache zu übersetzen, da es hierfür viel mehr Beispiele gibt. Lernt ein Algorithmus, die Kreditwürdigkeit von Kunden vorherzusagen, dann wird die in den Daten abgebildete Korrelation zwischen ethnischer Zugehörigkeit und finanzielle Leistungsfähigkeit dazu führen, dass der Algorithmus Vorhersagen macht, die vorurteilsbehaftet erscheinen, was als Diskrimination gegen bestimmte ethnische Gruppen interpretiert werden könnte. Es gibt sehr viele Beispiele und Möglichkeiten für solche Formen von unerwartetem „Bias“ in der neuen KI.

Nun sind die Tendenzen, die solche Algorithmen entdecken, nur eine Abbildung realer Korrelationen, keine Manipulationen des Algorithmus, der also kein eigenes „Vorurteil“ hat. Es ist eine politische, ethische, und auch wirtschaftliche Entscheidung, inwieweit solche Tendenzen in der Dienstleistung, die der Algorithmus erbringt, zulässig sind. Das Problem ist aber, solche Tendenzen überhaupt zu entdecken, und dann erst, mit ihnen umzugehen. Auch bei den klassischen Methoden der computergestützten Analyse von Daten – etwa mit Hilfe konventioneller Datenbasen – traten Probleme dieser Art bereits auf. Sie waren jedoch leichter zu erkennen. Die mathematische Natur der Methoden der neuen KI macht die Algorithmen weniger transparent und erschwert so die Aufdeckung solcher Tendenzen. Die Möglichkeit, die Methoden relativ blind auf neue, unformatierte Datensätze anzuwenden, die ja gerade einer der Vorteile der Verfahren ist, macht auch die Aufdeckung solcher Tendenzen schwierig.

Aktuell wird dieses Problem in der Forschung umfangreich untersucht²⁶ und ist von Anwendern erkannt worden. Ein subtiler Aspekt ist, dass man Hypothesen über mögliche Ungleichheiten in den Daten bzw. Diskriminierungen in den Entscheidungen eines KI-Systems haben muss, um solche Tendenzen aufzudecken. Diese Hypothesen sind nicht frei von Ideologie.

Ein zweiter Bereich gesellschaftlicher Probleme, die nicht rein wirtschaftlicher Natur sind, hängt mit der Wechselwirkung zwischen KI-Systemen und dem menschlichen Benutzer zusammen. Das Verhalten des Nutzers liefert aus Sicht des KI-Systems Daten. Diese Daten kann das KI-System nutzen, um sein eigenes Verhalten zu anzupassen. Es entsteht eine Wechselwirkung zwischen Nutzer und KI-System, das den Nutzer auf mehr oder weniger subtile Weise „steuern“ kann. Ein einfacher Fall ist die „gamification“ solcher Wechselwirkung. Dabei werden Anreize erzeugt, um den Nutzer „bei der Stange zu halten“, also weiter an der Computeranwendung zu interessieren, mit der der Nutzer wechselwirkt. Die meisten Computerspiele nutzen bereits solche Elemente. Spieler bekommen Punkte, Preise, visuelle und auditorische Reize, können sich neue Spielmöglichkeiten verdienen, und so weiter. All dies wird fein justiert, um den Spieler motiviert zu halten.

Im nächsten Schritt kann ein KI-System aus den Daten auch lernen, also Vorlieben des Nutzers schätzen. Das tun bereits alle großen Internetplattformen wie YouTube, Facebook etc. Die Algorithmen können die Texte, Bilder und Videos, die Nutzer ansehen, klassifizieren und dadurch Analogien zum Verhalten anderer Nutzer erstellen. Sie können so dem Nutzer Angebote machen, die gemäß dieser Analogien zu den Präferenzen des Nutzers „passen“. Wie der Nutzer auf solche Angebote reagiert, liefert dabei neue, ideale Lerndaten, denn das System kann seine Vorhersage mit dem eingetretenen Verhalten vergleichen, also zum Beispiel, welcher der angebotenen Links angeklickt wurde.

²⁶ Raji, I. D./ Smart, A./ White, R. N./ Mitchell, M./ Gebru, T./ Hutchinson, B./ Smith-Loud, J./ Theron, D./ Barnes, P., Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing, in: FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33–44, <https://doi.org/https://doi.org/10.1145/3351095.3372873>.

Dieses Lernen vom Nutzer ist die Grundlage des Geschäftsmodells von Google und anderen Anbietern von Online-Werbung. Damit wird Werbung personalisiert. In der einfachsten Form kommt nach dem Betrachten einer Webseite über Lederwaren eine Werbung für Lederwaren im nächsten Google Fenster. Das hat bestimmt mancher Leser schon erlebt. Viele Dimensionen der Persönlichkeit, der wirtschaftlichen Verhältnisse, der Interessen, etc. eines Nutzers können aus dem Verhalten bei der Nutzung des Internets oder des Mobiltelefons geschätzt werden. Auf der Basis solcher Daten kann das Interneterlebnis eines Nutzers rundum gestaltet werden. Solche Daten werden auch gehandelt, so dass andere Dienstleister Nutzer ansprechen können.

Dies mag für manche schon an sich problematisch erscheinen. Ganz offensichtlich problematisch ist dies, wenn es zu politischen oder verbrecherischen Zwecken genutzt wird. Totalitäre Regime wie China und Russland nutzen solche Methoden bereits jetzt zur Unterdrückung oppositioneller Stimmen – auch mit Zersetzungsmaßnahmen. Im Nachgang zu den Enthüllungen Edward Snowdens haben wir erfahren, dass auch westliche Länder Daten so nutzen, offenbar weit über reine Internetdaten hinaus.

Problematisch mag aber auch der kybernetische Effekt sein, der entsteht, wenn Nachrichten oder sonstige Inhalte auf die Entscheidungen von Nutzern reagieren. Ein viel beachtetes Problem ist die Entstehung von Kommunikationsblasen: Nutzern wird Information angeboten, die in einem für die Algorithmen messbaren Sinne den Informationen ähneln, die Nutzer bisher schon angesehen haben. So wird ein Nutzer nach und nach in eine Blase gezogen, in denen Information ständig zuvor Studiertes zu bestätigen scheinen. Dies tritt in Resonanz mit dem als „confirmation bias“ bekannten psychologischen Phänomen:²⁷ Menschen nehmen Information, die ihre Vormeinung bestätigt, stärker wahr als Information, die ihre Vormeinung widerlegt. Insgesamt entsteht ein Effekt der Abtrennung einzelner Nutzer vom allgemeinen Informationsfluss, und deren Einschluss in eine Meinungsgruppe.

²⁷ Nickerson, R. S., Confirmation bias: a ubiquitous phenomenon in many guises, Review of General Psychology, 2(2) (1998), 175–220.

Auch dies ist ein Effekt, den es klassisch durchaus bereits gab. Ein Bürger mit politisch linker Einstellung las traditionell eine eher linke Tageszeitung, die seine politische Einstellung weiter festigte. Die neuen Methoden skalieren nur diese klassischen Effekte, allerdings drastisch. Denn Meinungsblasen können jetzt sehr lokal auf sehr kleine Gruppen konvergieren, ohne redaktionelle Prinzipien, die gegen Abdriften in Extreme absichern. Gerade Facebook oder Twitter, wo die meisten Beiträge von anderen Nutzern stammen, erlaubt hier das Entstehen von Kommunikationsblasen, in deren Kern es regelmäßig keine einzige faktisch richtige Information gibt. Ein in seiner Radikalität geradezu verstörendes Beispiel war das als „Pizzagate“ bekannte Gerücht, dass demokratische Politiker Teil eines pädophilen Verbrecherringes seien, der von einer Pizzeria in Washington aus organisiert würde.

Insoweit junge Generationen ihre Meinungsbildung, sogar ihre eigene Identitätsbildung, zunehmend auf Medien und Online-Kommunikationsformen stützen, in denen Methoden des maschinellen Lernens solche manipulativen Einflüsse ermöglichen, gehen von den angedeuteten gesellschaftlichen Problemen der neuen KI möglicherweise fundamentale Bedrohungen des freiheitlich-demokratischen Gemeinwesens aus.²⁸

Zum Schluss möchte ich noch hinweisen auf etwas abstraktere Dimensionen dieser gleichen Trends, die unter dem Schlagwort „Hyperrealität“ diskutiert werden. Hier ist die Vorstellung, dass Menschen zunehmend die Welt durch technische Medien erleben. Was bisher Zeichentrickfilme waren, sind jetzt durch „CGI“ – Computer Generated Imagery – erweiterte Filme, in denen hoch-realistische visuelle Darstellungen reiner Computersimulation mit realen Bildern nahtlos fusioniert werden. Heutige Kinofilme nutzen solche Techniken routinemäßig in ständig wachsendem Umfang. Auch hier hilft die neue KI, vor allem, indem sie die Kosten für solche Techniken reduziert und die Erzeugung neuer Inhalte teilweise automatisiert. Gleichzeitig findet eine radikale Verschiebung der alltäglichen Erfahrung weg von der realen „3D“-Welt zu den vielen Formen virtueller „2D“-Welten statt, also vom Smartphone, Tablet, Computer, TV-Bildschirm oder Computerspiel bis hin zur virtuellen Realität, die gerade einen Eintritt in den Endverbrauchermarkt zu

²⁸ Snyder, T., *The road to unfreedom: Russia, Europe, America*, New York 2018.

schaffen scheint.²⁹ Die Frage ist, was mit Menschen geschieht, deren Erleben überwiegend auf diesem Wege stattfindet.³⁰ Einzelne Elemente kann man schon erahnen. Teenager laden ja umfangreich Fotos und andere Elemente auf Internet-Plattformen hoch, über die ein Teil ihrer Sozialisierung stattfindet. Dabei spielen „Selfies“ eine besondere Rolle. Bereits jetzt wird dabei häufig Software genutzt, die diese Selfies automatisch retuschiert, also zum Beispiel Pickel entfernt oder die Hauttönung ändert. Jugendliche, deren Erleben der Gleichaltrigen vorwiegend auf solchen idealisierten Bildern beruht, entwickeln vielleicht völlig verzerrte Vorstellung vom menschlichen Aussehen, und dies vor allem unbewusst.

4. Zusammenfassung

Die neue künstliche Intelligenz, deren Methoden des maschinellen Lernens es klassischen Algorithmen erlaubt, Muster in Daten zu erkennen und so eine Anbindung an Sensorik und Motorik zu ermöglichen, ist weit davon entfernt, menschliches Verhalten in der realen Welt zu emulieren. Für eng umschriebene Aufgaben können diese neuen KI-Systeme unter Umständen aber sogar übermenschliche Leistungen erbringen. Dennoch fehlt ihnen grundsätzlich die Orientierung auf Ziele, der Wille solche Ziele zu erreichen, und die grundlegenden biologischen Antriebe, wie das Vermeiden von Gefahr, die Exploration von Neuem, und das Interesse an der Kommunikation mit Anderen. Von den neuen KI-Systemen gehen daher Gefahren für unsere Freiheit als Menschen eher nicht dadurch aus, dass diese KI-Systeme sich wie Menschen verhalten und die Macht ergreifen wollen. Gefahren gehen von den neuen Systemen auf weniger exotische Weise aus: Die KI-Systeme beschleunigen Trends, die durch die Digitalisierung aller Prozesse der Wirtschaft, der Kommunikation und der Unterhaltung schon seit Jahrzehnten stattfinden. Dies stellt Gesellschaften vor wirtschaftliche, aber zunehmend auch ethische, mentale, und kulturelle Probleme. Hier kann quantitative Skalierung

²⁹ Desmurget, M., *La fabrique du crétin digital— les dangers des écrans pour nos enfants*, Paris 2019.

³⁰ Vgl. auch hierzu Snyder, *The road to unfreedom* (Fußnote 28).

durchaus zu neuen Qualitäten führen, die uns als Menschen vor grundlegende Fragen stellt, auch zu unserer Identität und zu unserer Verankerung in der Wirklichkeit.

Kognitive Robotik – Chancen und Risiken

1. Einführung

Roboter faszinieren die Menschen bereits seit mindesten einhundert Jahren. Der Begriff selbst stammt von dem tschechischen Wort *robota*, was in etwa mit Frondienst oder Knechtschaft übersetzt werden kann, und wurde 1920 von dem tschechischen Künstler und Schriftsteller Josef Čapek geprägt, um damit maschinenähnliche Wesen in einem Theaterstück seines Bruders Karl zu bezeichnen. Auch wenn es noch lange keine echten Roboter geben sollte, fanden sie schnell ihren Weg in Science-Fiction Filme, von *Metropolis* aus dem Jahr 1927, über *Der Tag, an dem die Erde stillstand* von 1951, bis zu der *Terminator*-Filmreihe der letzten Jahrzehnte. Es ist wahrscheinlich den Science-Fiction Filmen und der Literatur geschuldet, dass die Menschen sich Roboter meist als menschenähnliche (humanoide) Maschinen vorstellen, auch wenn diese in der Realität eher die Ausnahme sind, aber dazu später mehr.

Lassen Sie uns mit der Frage beginnen, wann man einem Roboter das Prädikat „kognitiv“ zuerkennen kann oder sollte. In gewisser Weise hängt dies direkt von der Art und Weise ab, wie Roboter eingesetzt werden, insbesondere, ob die Umgebung, in der sie agieren, vorhersehbar ist oder nicht. Vorhersehbar heißt hier, dass sich der Roboter selbst keine „Gedanken“ machen muss, wie seine Umgebung aussieht und alle Bewegungen und Aktionen bis ins kleinste Detail vorausberechnet bzw. von Menschen programmiert werden können. Ein typisches Beispiel sind etwa Schweißroboter in der Automobilindustrie, die unermüdlich, präzise und schnell Karosserieteile zusammenschweißen. Diese Industrieroboter sind zwar in der Lage, höchst komplexe Handlungsabläufe durchzuführen, es sind aber immer wieder die-

selben Abläufe, und alles funktioniert nur unter der Annahme, dass die Teile sich immer an exakt denselben Stellen befinden. In der Informatik spricht man daher von einer *deterministischen* Welt, in der man sich auch ohne kognitive Fähigkeiten, sozusagen blind zurechtfinden kann.

Ganz anders sieht die Sache aus, wenn Roboter in unstrukturierten oder für sie gar völlig unbekanntem Umgebungen agieren müssen. Ein einfaches Beispiel sind die bekannten Staubsaugerroboter. Bevor ein solcher Roboter das erste Mal in einer Wohnung seinen Dienst antreten kann, muss er zunächst einmal seine Umgebung erkunden. Dazu wird er die Zimmer mehr oder weniger systematisch abfahren und sich eine Karte erstellen. Dabei wird er wahrscheinlich auch gefährliche Stellen wie Treppenabgänge markieren und sich merken, wo sich seine Ladestation befindet. All das funktioniert nur, wenn man den Roboter mit Sensoren ausstattet, die es ihm erlauben, die Umgebung wahrzunehmen. Dazu gehören zum Beispiel Sonarsensoren, die es ermöglichen, mit Hilfe von Ultraschallsignalen den Abstand zu den nächstgelegenen Hindernissen zu vermessen. Sobald eine Karte erstellt ist, kann der Roboter sich recht sicher in der Wohnung bewegen und seine Arbeit verrichten, muss dabei aber immer damit rechnen, dass plötzlich unerwartete „Hindernisse“ wie Katzen, Hunde oder Menschen vor ihm auftauchen oder Möbelstücke verrückt worden sind und geeignet darauf reagieren. Bei einer Begegnung mit einer Katze reicht es in der Regel einfach stehen zu bleiben oder auszuweichen, und bei verändertem Mobiliar wäre es vielleicht sinnvoll die Karte anzupassen. Kognitive Roboter sind daher solche, die ihre Umwelt aktiv wahrnehmen und flexibel auf Änderungen reagieren können. Man beachte, dass mit „kognitiv“ nicht unbedingt eine außergewöhnliche Intelligenz verbunden sein muss. Staubsaugen ist nun mal nicht, wie man so schön im Englischen sagt, *Rocket Science*. Und die heutigen Staubsaugerroboter lassen auch noch viel zu wünschen übrig, da sie etwa Schmutz nicht von Legosteinen oder Schmuckstücken unterscheiden können. Trotzdem benötigen sie kognitive Fähigkeiten zur Wahrnehmung und Interpretation ihrer Umwelt.

Im Folgenden werden wir uns zunächst kurz mit der Entwicklung der kognitiven Robotik seit ihren Anfängen Mitte der sechziger Jahre befassen, gefolgt von einem Überblick über den Stand der Technik und Bereichen, wo die

Entwicklungen noch ziemlich am Anfang stehen. Ich werde auf die Gefahren eingehen, wenn Roboter als autonome Waffensysteme eingesetzt werden, und mit einigen persönlichen Gedanken zum Begriff der Freiheit in Bezug auf Roboter enden.

2. Eine kleine Geschichte der kognitiven Robotik

Die Geschichte der kognitiven Robotik beginnt 1966 mit der Entwicklung des mobilen Roboters *Shakey*, der bis 1972 am Stanford Research Institute, heute SRI, entwickelt wurde. Shakey war ein mannshohes Ungetüm auf Rädern, ausgestattet mit Kameras, Sonar- und Stoßsensoren, und konnte in einer künstlichen, wohnungsähnlichen Umgebung Objekte von einem Ort zum anderen schieben. Sein Name war der Tatsache geschuldet, dass er beim Fahren ordentlich hin und her wackelte. Die meisten Berechnungen führte Shakey auf einem externen Computer aus, mit dem er über eine Funkverbindung kommunizierte und der selbst einen ganzen Raum einnahm. Um die Bildverarbeitung, die damals noch in den Kinderschuhen steckte, zu vereinfachen, bestanden die Objekte aus großen geometrischen Figuren wie Würfeln oder Quadern. Dies alles klingt heute nicht besonders aufregend, war aber damals ein Meilenstein der Roboterforschung und machte Furore in den Medien. Was Shakey auszeichnete war, dass er nicht nur die zwar künstliche, aber doch für ihn unbekannte Umgebung erkunden konnte, sondern auch selbständig Handlungspläne berechnen konnte, um ein Objekt von A nach B zu transportieren. Die eigens dafür entwickelte Sprache STRIPS zur Beschreibung von Planungsproblemen ist sogar noch heute in Gebrauch.

Nach diesen ersten Erfolgen brachten die folgenden Jahre allerdings relativ wenige Fortschritte, was unter anderem an den Kosten lag, Roboter zu entwickeln, aber auch an der sehr begrenzten Rechenleistung der damaligen Zeit, vor allem, wenn man nicht auf externe Rechner angewiesen sein wollte. Erst Anfang der 90er Jahre kam wieder Bewegung in die Entwicklung kognitiver Roboter, und es gab auch die ersten Firmen wie *Real World Interface* (RWI) oder *Nomadic Technologies*, die mobile Roboterplattformen für Forschungszwecke produzierten. Die wirklichen Durchbrüche wurden aller-

dings auf dem Gebiet der Navigationssoftware erzielt. Während Shakey noch ein weitgehend deterministisches Weltbild hatte, was zu allerlei Problemen führte, wurden nun Algorithmen für die Roboternavigation entwickelt, die mit Unsicherheiten aufgrund verrauschter Sensordaten oder unpräziser Bewegungen des Roboters umgehen konnten, vor allem mit Methoden der Wahrscheinlichkeitstheorie. Dies führte dazu, dass Roboter plötzlich in der Lage waren, sich sicher in menschlichen Umgebungen zu bewegen. Ein eindrucksvolles Beispiel war der an der Universität Bonn entwickelte Roboter *Rhino*, einem RWI B21, der 1997 erstmals demonstrierte, dass Roboter auch außerhalb eines Forschungslabors einsetzbar sind, im Falle von Rhino als Museumsführer im Deutschen Museum Bonn. Der geistige Vater von Rhino war damals Sebastian Thrun, unterstützt von Wolfram Burgard und Dieter Fox. Alle drei zählen heute zu den renommiertesten Robotikforschern weltweit.

Ein weiterer Durchbruch erfolgte 2005 im Rahmen der *DARPA¹ Grand Challenge*. Bei diesem Wettbewerb ging es darum, mit einem Fahrzeug ca. 200 km autonom durch die Mojave Wüste in Kalifornien zu fahren. Von ursprünglich 195 Teams schafften es 23 in die Endrunde, wovon schließlich fünf das Ziel erreichten, allerdings nur vier innerhalb des vorgegebenen Zeitlimits von 10 Stunden. Sieger war das *Stanford Racing Team* mit seinem Fahrzeug *Stanley*, einem modifizierten VW Touareg. Geleitet wurde das Team von Sebastian Thrun, den es nach seiner Promotion in die USA gezogen hatte. Viele der verwendeten Methoden zur Umgebungserkennung und Navigation gehen auf diejenigen zurück, die vorher im Rhino-Projekt entwickelt wurden, und bilden auch heute noch die Grundlage des autonomen Fahrens. Sebastian Thrun wechselte später zu Google, um dort die Sparte für autonomes Fahren zu gründen, und heute leitet er ein Startup, das ein autonomes, elektrisch getriebenes Flugtaxi entwickelt. Neben vielen anderen Ehrungen wurde ihm für seine Leistungen 2020 der Ingenieurpreis der RWTH und der Stadt Aachen verliehen.

Die Methoden zur sicheren Navigation, die mit Rhino ihren Anfang nahmen, haben nicht nur das autonome Fahren geprägt, sondern werden heute über-

¹ DARPA steht für Defense Advanced Research Project Agency, einer Forschungsförderungsabteilung des amerikanischen Militärs.

all standardmäßig bei mobilen Robotern eingesetzt, einschließlich der bereits erwähnten Staubsaugerroboter, die wegen ihres günstigen Preises auch einen der wenigen kommerziellen Erfolge in der Robotik darstellen. Ansonsten werden mobile Roboter bisher meist für Botengänge in Hotels oder Krankenhäusern eingesetzt. Im industriellen Kontext spielen sie auch noch keine große Rolle, mit Ausnahme vielleicht der Intralogistik, wo Roboterflotten in Warenlagern mehr und mehr für Transportaufgaben eingesetzt werden. Ansonsten ist es immer noch eine Vision von Industrie 4.0, Roboter flexibel in der Produktion im Zusammenspiel mit Menschen und Maschinen einzusetzen, aber bis zur Realisierung ist es noch ein weiter Weg. In meiner eigenen Forschungsgruppe arbeiten wir seit Jahren daran, im Rahmen der *RoboCup Logistikliga* Roboter für die Produktion in einer simulierten Fabrik zu entwickeln. Die Aufgabe der Roboter besteht dabei im Wesentlichen darin, Güter mit Hilfe verschiedener Maschinen zu produzieren und innerhalb vorgegebener Zeitfenster abzuliefern. Es sei hier am Rande erwähnt, dass der RoboCup ursprünglich im Jahr 1997 ins Leben gerufen wurde, um Fußball spielende Roboter zu entwickeln, die bis 2050 in der Lage sein sollen, die menschlichen Fußballweltmeister zu besiegen. Ob dieses Ziel erreicht wird, steht noch in den Sternen, aber diese Vision hat dazu geführt, dass sich jedes Jahr tausende Forscher, Studierende und auch Schüler bei den RoboCup-Weltmeisterschaften treffen, um mit ihren Robotern in verschiedenen Ligen gegeneinander anzutreten und sich zu messen. Mittlerweile sind auch Ligen jenseits des Fußballs entstanden, darunter die RoboCupRescue-Liga, bei der es darum geht, Menschen mit Hilfe von Robotern aus Trümmern zu retten, und die bereits genannte Logistikliga. Der RoboCup hat sich als sehr motivierend erwiesen, insbesondere um Schüler und Studierende für das Thema Robotik zu begeistern. Das gilt nicht zuletzt auch für meine Arbeitsgruppe, zumal wir schon mehrmals den Weltmeistertitel in der Logistikliga erringen konnten.

Während das Problem der Navigation dank der in den 1990er Jahren entwickelten Ideen weitgehend als gelöst angesehen werden kann, ist dies im Fall der Objektmanipulation noch lange nicht absehbar. Zwar wird intensiv an der Entwicklung leistungsfähiger Greifer geforscht, aber diese sind in ihren Fähigkeiten noch meilenweit von dem entfernt, was menschliche Hände

leisten können. Punktuell gibt es sicher Erfolge wie die kürzlich von *OpenAI* entwickelte Hand, die innerhalb von Sekunden Rubik's Cube – einhändig! – lösen kann, aber die Flexibilität, mit der Menschen ihre Hände fast beliebig einsetzen können, bleibt bislang unerreicht.

Im Bereich der humanoiden Roboter, also Robotern in Menschengestalt, steht man noch vor großen Herausforderungen, auch wenn sie in Science-Fiction Filmen quasi zum Standardrepertoire gehören. Ein Problem haben wir bereits kennengelernt, nämlich funktionstüchtige Hände zu entwickeln. Darüber hinaus stellt das Gehen auf zwei Beinen eine besondere Herausforderung dar. Auch wenn der gesunde Mensch sich kaum Gedanken darüber macht, aufrechtes Gehen ist ein hochkomplexer Vorgang, und es gibt nur wenige Roboter, die dies zufriedenstellend leisten. Eine Ausnahme ist der von der Firma *Boston Dynamics* entwickelte Roboter *Atlas*, der sich auch in unwegsamem Gelände sicher auf zwei Beinen bewegt. Was ihm allerdings noch fehlt, ist die Intelligenz dabei auch noch sinnvolle Aufgaben zu erledigen. Warum das so ist, werden wir später noch diskutieren. Aber vielleicht sollte man sich zunächst einmal fragen, warum es überhaupt sinnvoll ist Roboter mit menschenähnlicher Form zu bauen. Abgesehen von ästhetischen und anderen nicht-funktionalen Erwägungen liegt der Hauptgrund schlicht und einfach in der Tatsache, dass Umgebungen, die von uns Menschen erschaffen wurden, speziell auf unsere Körper abgestimmt sind. Denken Sie nur an ihre Wohnung mit Treppen, Regalen auf Augenhöhe und vielem mehr. Wenn uns dort ein Roboter einmal zur Hand gehen soll, dann macht es durchaus Sinn, wenn er einen ähnlichen Körper besitzt wie wir.

Trotzdem gibt es auch Anwendungen, die nicht unbedingt eine humanoide Form des Roboters verlangen. So hat die eben schon genannte Firma *Boston Dynamics* den hundeähnlichen Roboter *Spot* entwickelt, den man etwa zur Objektüberwachung einsetzen kann und der mit seinen vier Beinen recht agil Treppen steigen oder über Hindernisse hinweggehen kann. Außerdem gibt es schwimmende Unterwasserroboter, fliegende Drohnen und sogar schlangenähnliche Roboter, um die Kanalisation zu erkunden. Drohnen werden uns später noch beschäftigen.

3. Wo steht die kognitive Robotik heute?

Wir haben gesehen, dass sich die kognitive Robotik seit Shakey recht eindrucksvoll weiterentwickelt hat, vor allem seit den 1990er Jahren. So können sich die Roboter mittlerweile auch in unbekanntem Umgebungen gut zu rechtfinden und sicher navigieren. Dabei handelt es sich allerdings in der Mehrzahl um fahrende Maschinen. Selbstfahrende Autos gehören auch in diese Kategorie und sind geradezu ein Paradebeispiel, wozu kognitive Roboter heute fähig sind. Wie schon erwähnt, begann die Erfolgsgeschichte des autonomen Fahrens mit der DARPA Grand Challenge im Jahr 2005. Zwei Jahre später fand dann bereits die *Urban Challenge* statt, wo es bereits darum ging mit selbstfahrenden Autos in einer simulierten Stadt in Kalifornien vorgegebene Ziele anzufahren. Um ein möglichst realistisches Szenario zu erreichen, fuhren dort auch von Menschen gesteuerte Autos umher. Aus Sicherheitsgründen waren dies aber eigens für diesen Zweck engagierte Stuntmen und -women, da man noch nicht wusste, wie sich die autonomen Fahrzeuge verhalten würden. In der Tat kam es auch zu kleineren Unfällen, wobei zum Glück Menschen nicht zu Schaden kamen. Am Ende gewann ein Team der Carnegie Mellon University und das Team von Sebastian Thrun wurde Zweiter. Seit dieser Zeit hat es rasante Fortschritte im autonomen Fahren gegeben. Das liegt zum einen an den schon erwähnten Navigationsalgorithmen, aber insbesondere auch an den Fortschritten im maschinellen Lernen. Die dazu nötigen Daten liefern die autonomen Autos während der vielen Testfahrten selbst mit Hilfe der zahlreichen Sensoren. Mittlerweile sind diese Fahrzeuge bereits viele Millionen Kilometer gefahren. Für das Finden einer Route von A nach B verwenden sie auch die uns bekannten GPS-basierten Navigationssysteme. Was gelernt werden muss, ist dabei hauptsächlich das richtige Verhalten während der Fahrt. Dabei helfen auch Informationen zur aktuellen Verkehrslage und Bildverarbeitungsprogramme zur Erkennung von Verkehrsschildern, Personen und anderen Fahrzeugen, die auch jetzt schon in vielen Assistenzsystemen herkömmlicher Autos eingesetzt werden. Mittlerweile meistern autonome Fahrzeuge fast jede Verkehrssituation, wobei ich allerdings eine Betonung auf „fast“ legen möchte. Es kommt zwar selten zu Fehlverhalten, aber leider manchmal auch zu kata-

strophalen Fehlern mit tödlichem Ausgang für die Insassen. Warum dieses letzte Quäntchen Fahrsicherheit immer noch nicht erreicht worden ist, wird uns später noch beschäftigen.

Die meisten humanoiden Roboter wirken beim Gehen immer noch recht unbeholfen, auch wenn es, wie erwähnt, erste Exemplare wie Atlas gibt, dessen aufrechter Gang dem des Menschen schon verblüffend ähnlich ist. Viel zu tun gibt es noch im Bereich der Manipulation. Es gibt zwar bereits ausgeklügelte mechanische Arme und Hände, auch solche, die mit taktilen Sensoren ausgestattet sind, aber die Steuerung dieser komplexen Gebilde erreicht noch lange nicht die Fähigkeiten, die wir von Menschen gewohnt sind.

Eine wesentliche Schwäche heutiger Roboter sind die immer noch sehr eingeschränkten kognitiven Fähigkeiten. Mein Versuch einer Definition kognitiver Roboter zu Anfang dieses Beitrags war ja recht anspruchslos und erlaubt es, selbst Staubsaugerroboter als kognitiv zu bezeichnen. Wenn man die Messlatte etwas höher anlegen möchte und etwa erwartet, dass ein kognitiver Roboter vorausschauend und klug agiert, dann gibt es durchaus Fortschritte, aber man stößt immer noch schnell an Grenzen.

So sind Roboter bereits jetzt in der Lage komplexe Aktionsfolgen zu planen, um ihre Ziele zu erreichen. Daran arbeiten wir zum Beispiel in meiner Arbeitsgruppe, um unseren Robotern in der Produktionslogistik beizubringen, möglichst effizient und in Teamarbeit Produkte herzustellen. Zwar ist für jedes Produkt im Vorhinein bekannt, aus welchen Teilen es besteht und welche Maschinen zur Produktion benötigt werden, für die Herstellung eines bestimmten Produkts müssen aber die jeweiligen Maschinen rechtzeitig reserviert werden und vor allem müssen die Aufgaben auf die Roboter verteilt werden, um möglichst parallel fertigen zu können. Hinzu kommt, dass es immer mehrere Aufträge gleichzeitig gibt, wobei bestimmte Deadlines eingehalten werden müssen, die auch Einfluss auf die Reihenfolge der Abarbeitung haben. Es handelt sich also um eine recht komplexe Planungsaufgabe, die von den Robotern in Eigenregie erledigt werden muss.

Auch wenn ein Plan gefunden ist, besteht ein wesentliches Problem darin, dass die Planausführung immer wieder zu Fehlern führt. Aktionen können fehlschlagen, etwa durch Fallenlassen eines Werkstücks, unvorhergesehene Dinge können passieren wie das Versperren des Wegs durch einen anderen

Roboter, von katastrophalen Abstürzen von Maschinen und Robotern ganz zu schweigen. Es ist immer noch Forschungsthema Wege zu finden, damit Roboter auf solche Unwägbarkeiten flexibel und vernünftig reagieren können.

An dieser Stelle mag der Einwurf kommen, wieso der Fortschritt hier anscheinend so mühsam und langsam vonstatten geht, wo doch die künstliche Intelligenz in den letzten Jahren so bahnbrechende Erfolge vorzuweisen hat. Man denke nur an *AlphaGo*,² ein von der Firma Deepmind entwickeltes Spieleprogramm, das vor einigen Jahren Lee Sedol, einen der weltbesten Go-Spieler, 4:1 geschlagen hat; oder *Pluribus*,³ ein an der Carnegie-Mellon-Universität in Zusammenarbeit mit Facebook entwickelter Poker-Bot, der 2019 gegen fünf Weltklasse-Pokerspieler überraschend deutlich gewonnen hat. Der Unterschied zu kognitiven Robotern liegt schlicht und einfach in der Tatsache, dass diese Spiele-Bots in einer wohldefinierten künstlichen Welt agieren. Wie bei Schach oder Dame hat Go klare Spielregeln, und zu jedem Zeitpunkt können beide Spieler die gerade vorliegende Spielsituation vollständig erfassen. Dies erlaubt es Spieleprogramme zu entwickeln, die Züge millionenfach ausprobieren und dabei aus ihren Fehlern lernen, um schließlich übermenschlich gut Go zu spielen. Bei Poker ist die Lage noch um einiges komplizierter, da ein Spieler, egal ob Mensch oder Maschine, ja nicht in die Karten der anderen schauen kann und trotzdem eine Strategie entwickeln muss, abhängig vom eigenen Blatt und dem Verhalten der anderen geschickt zu bluffen. Aber im Grunde ist die Welt des Pokers immer noch recht überschaubar im Vergleich zur realen Welt, sodass es auch hier möglich ist, einen Poker-Bot wie Pluribus millionenfach gegen sich selbst spielen zu lassen und aus Fehlern zu lernen, gekoppelt mit Ideen aus der mathematischen Spieltheorie und vielem mehr.

Die Art und Weise, wie es AlphaGo und Pluribus gelungen ist, Go bzw. Poker zu meistern, ist sicher faszinierend, und den Entwicklern gebührt dafür

² Silver, D./ Huang, A./ Maddison, C. et al., Mastering the game of Go with deep neural networks and tree search, *Nature* 529 (2016), 484–489, DOI:10.1038/nature16961.

³ Brown, N./ Sandholm, T., Superhuman AI for multiplayer poker, *Science* 365 (2019), 885–890, DOI: 10.1126/science.aay2400.

höchste Anerkennung. Aber leider sind diese Methoden nicht oder nur sehr bedingt auf Roboter anwendbar. Das liegt an mindestens zwei grundsätzlichen Problemen. Zum einen gibt es keine fest definierten Spielregeln, sondern die Art und Weise, wie Roboter mit ihrer Umwelt interagieren können, umfasst schier unendlich viele Möglichkeiten. Dies gilt selbst für scheinbar einfache Aufgaben wie den Tisch zu decken. Während jeder von uns weiß, wie man das macht, bin ich sicher, dass es noch niemandem gelungen ist, einen Tisch mehrmals auf die exakt gleiche Art und Weise zu decken, und damit meine ich nicht unbedingt das Endergebnis, sondern die Folge der Aktionen, bis ins kleinste Detail der Bewegungen von Armen, Beinen und Händen, die zu einem gedeckten Tisch führen. Zum anderen ist das Agieren in der realen Welt sehr zeitraubend. Mit anderen Worten, es ist einfach nicht möglich Dinge millionenfach auszuprobieren, ganz zu schweigen von den Kosten, die damit verbunden wären. Außerdem kann Fehlverhalten schnell katastrophale Folgen haben. Man denke nur an einen Roboter, der aus Versehen beim Versuch den Tisch zu decken die Treppe herunterfällt.

Man mag einwenden, dass man stattdessen Dinge vielleicht besser in einer Computersimulation ausprobieren sollte. In der Tat arbeiten wir in der Robotik viel mit Simulatoren, insbesondere um neue Software zu entwickeln und zu testen, bevor man sie auf den echten Robotern einsetzt. Damit die Simulation möglichst realitätsnah ist, muss sie allerdings physikalische Gegebenheiten wie Reibung oder Schwerkraft möglichst genau nachbilden. Das ist wiederum recht zeit- und rechenintensiv, sodass man selbst in Simulationen nur vergleichsweise wenige Handlungsfolgen ausprobieren kann, also nur einen winzigen Bruchteil dessen, was in den künstlichen Spielwelten von Go oder Poker möglich ist, wo Physik überhaupt keine Rolle spielt.⁴

⁴ Das wäre anders, wenn die Maschinen tatsächlich spielen müssten, also Karten in der Hand halten oder Steine auf dem Brett platzieren. Nicht ohne Grund wird den Spiele-Bots bei einem Turnier Mensch gegen Maschine immer noch ein Mensch zur Seite gestellt, um diese scheinbar trivialen Tätigkeiten zu übernehmen.

4. Wo steht die kognitive Robotik noch am Anfang?

Lassen Sie mich nun auf Bereiche eingehen, wo die kognitive Robotik noch ziemlich am Anfang steht im Vergleich zu dem, was menschliche Intelligenz auszeichnet. Um dies zu veranschaulichen, schauen wir uns noch einmal das autonome Fahren an und insbesondere die Frage, warum autonomes Fahren zwar in der Regel gut funktioniert, aber es immer wieder Situationen gibt, wo diese Autos Fehlentscheidungen treffen. Dazu muss man wissen, dass diese Systeme ihre Entscheidungen weitgehend aufgrund von Erfahrungen treffen. Im einfachsten Fall werden dazu zunächst große Datenmengen mit Hilfe menschlicher Fahrer gesammelt. Dabei werden zu den jeweils vorliegenden Sensordaten die Aktionen des Fahrers wie bremsen, beschleunigen oder lenken notiert. Später werden damit maschinelle Lernverfahren wie neuronale Netze trainiert. Nach erfolgreichem Training sind diese Lernverfahren dann in der Lage, in Verkehrssituationen, die denen während der Lernphase ähneln, das Richtige zu tun. So kann man mit genügend großen Datenmengen viele, wenn nicht die meisten Szenarien meistern. Ein Haken bei der Sache ist allerdings, dass es ab und zu dennoch zu Situationen kommt, die keiner der bisher gesehenen entsprechen. Man denke etwa an einen plötzlich auftretenden Schneesturm oder die Begegnung mit einem Tier. In den Daten wird man dafür kaum Beispiele finden, und das gleiche gilt für Szenarien, die mit einiger Wahrscheinlichkeit oder zwangsläufig zu einem Unfall führen würden. Das autonome Fahrzeug wird dann zwar immer noch eine Entscheidung treffen, aber eben nicht unbedingt die richtige, da die Trainingsdaten nichts Vergleichbares enthalten haben.

Wir Menschen funktionieren da zum Glück ganz anders. Auch wenn unser Verhalten in der Regel auf Erfahrungen beruht, sind wir ohne weiteres in der Lage, uns auf völlig neue Situationen einzustellen. Jeder, der zum ersten Mal mit seinem Fahrzeug in einen Schneesturm gerät, wird instinktiv die Geschwindigkeit reduzieren und eventuell nach einer Möglichkeit suchen am Straßenrand zu halten. Wir werden dabei von unserem „gesunden Menschenverstand“ geleitet. So wissen wir einfach, dass es keine gute Idee ist, bei schlechter Sicht einfach stehen zu bleiben, da die Gefahr groß ist, dass uns ein Hintermann übersieht und auffährt. Das heißt sicher nicht, dass wir Men-

schen immer das Richtige tun, aber was uns auszeichnet ist eine ungeheure Flexibilität uns auf Unvorhergesehenes einzustellen und selbst aus brenzligen Situationen in der Regel heil herauszukommen. Diese Fähigkeit fehlt heutigen Robotern noch völlig und stellt eine große wissenschaftliche Herausforderung dar. Es ist aus meiner Sicht auch unwahrscheinlich, dass man einer Lösung allein mit den heutigen, datengetriebenen Lernverfahren näherkommen wird, schon allein deshalb, weil es die nötigen Daten oft nicht in ausreichender Zahl oder gar nicht gibt. Während die Erfolge der KI der letzten Jahre weitgehend auf eine verbesserte Mustererkennung zurückzuführen sind, liegt ein Grundproblem der derzeitigen Verfahren in der Tatsache, dass sie überhaupt nicht verstehen, was sie da eigentlich erkennen. Ein neuronales Netz, das auf einem Bild einen Hund erkennt, trifft diese Entscheidung allein basierend auf Ähnlichkeiten zu den vielen anderen Hundebildern, mit denen es vorher trainiert wurde.⁵ Es hat überhaupt keine Vorstellung davon, was ein Hund wirklich ist, geschweige denn welche Rolle Hunde für Menschen spielen. Diese Art des tiefen Verstehens, wie unsere Welt aufgebaut ist und funktioniert, scheint aber eine Grundvoraussetzung für gesunden Menschenverstand zu sein. In dieser Hinsicht sind selbst Kleinkinder Robotern heutzutage noch weit überlegen.⁶

Ein weiteres und in gewisser Weise noch drängenderes Problem als gesunder Menschenverstand verbirgt sich hinter der Frage, inwieweit kognitive Roboter in der Lage sind oder sein sollten, nach ethischen Prinzipien zu handeln. Diese Frage beschäftigte bereits den Science Fiction Autor Isaac Asimov, der 1950 in seiner Kurzgeschichte „*I, Robot*“ die folgenden drei Robotergesetze formulierte:⁷

1. Ein Roboter darf kein menschliches Wesen verletzen oder durch Untätigkeit zulassen, dass einem menschlichen Wesen Schaden zugefügt wird.

⁵ Dies führt auch dazu, dass man solche neuronalen Netze leicht aufs Glatteis führen kann, indem man einzelne Bildpunkte leicht verändert, sodass für die Maschine aus einem Hund eine Katze wird, obwohl das Bild für uns Menschen immer noch genauso aussieht wie vorher.

⁶ Eine Empfehlung zur Vertiefung dieses Themas ist das kürzlich erschienene Buch Marcus, G./ Davis, E., *Rebooting AI: Building Artificial Intelligence We Can Trust*, New York 2019.

⁷ Asimov, I, *I, Robot*, New York 1950.

2. Ein Roboter muss den ihm von einem Menschen gegebenen Befehlen gehorchen – es sei denn, ein solcher Befehl würde mit Regel eins kollidieren.
3. Ein Roboter muss seine Existenz beschützen, solange dieser Schutz nicht mit Regel eins oder zwei kollidiert.

Diese Regeln wurden in der Zwischenzeit ergänzt und verfeinert, wobei die wesentliche Aussage aber erhalten bleibt, dass Roboter Helfer der Menschen sein sollen und ihnen keinen Schaden zufügen dürfen. Allerdings steckt wie immer der Teufel im Detail, wenn man genauer hinsieht. Dies lässt sich sehr schön am Beispiel des autonomen Fahrens illustrieren: Stellen Sie sich vor, Sie sitzen allein in einem solchen Fahrzeug und plötzlich läuft ein Kind auf die Straße. Stellen Sie sich weiterhin vor, dass es für das Fahrzeug nur die Möglichkeit gibt, entweder das Kind zu überfahren oder gegen eine Mauer zu fahren und dabei den Insassen, also Sie, zu töten. Eine weltweit geführten Umfrage⁸ ergab, dass viele Menschen in einer solchen Situation durchaus bereit sind sich selbst zu opfern. Wenn man Ihnen allerdings die Frage stellt, ob sie bereit wären in ein autonomes Fahrzeug einzusteigen, das diese Entscheidung für sie trifft, dann sah die Antwort schon anders aus.⁹ Dahinter steckt vielleicht auch, dass wir zwar bereit sind Menschen derartige Entscheidungen zu überlassen und durchaus nachsichtig sein können, wenn es in Extremsituationen zu Fehlentscheidungen kommt, wir bei Robotern diese Nachsicht aber eher vermissen lassen.

Ich möchte das eben skizzierte moralische Dilemma, das zum Glück nur sehr selten in der Realität vorkommt, gar nicht bewerten, es zeigt aber deutlich, vor welchen großen Herausforderungen Entwickler von kognitiven Robotern stehen, die diesen Maschinen ethisches Verhalten beibringen möchten. Und selbst wenn man sich auf gewisse Prinzipien einigen könnte, ist es überhaupt noch nicht klar, wie man das technisch umsetzen könnte. Forschung zu diesem Thema hat zwar begonnen, steht aber noch weitgehend am Anfang. Für autonome Fahrzeuge wie für fast alle mir bekannten Roboter gilt, dass sie derzeit noch keinerlei ethisches Verhalten an den Tag legen.

⁸ Awad, E. et al., The Moral Machine Experiment, Nature 518 (2018), 59–64.

⁹ Bonnefon, J.-F. et al., The Social Dilemma of Autonomous Vehicles, Science 352(6293) (2016), 1573–1576.

Ein weiterer Aspekt, der gerade in Bezug auf humanoide Roboter oft zu Fehleinschätzungen und Missverständnissen führt, sind Emotionen. Viele kennen wahrscheinlich den freundlich winkenden Roboter *Pepper* der Firma Softbank, der oft, vielleicht zu oft in den Medien verwendet wird, wenn es um das Thema Robotik oder allgemein um Digitalisierung geht. Das Gesicht von Pepper wurde sehr wahrscheinlich mit der Intention entwickelt, den Roboter als sympathisches Gegenüber erscheinen zu lassen. Dabei ist Pepper immer noch von recht einfacher Gestalt. Es wurden bereits viel komplexere Roboterköpfe entwickelt, die fast menschenähnliche Gesichtszüge besitzen und damit Emotionen wie Freude, Ärger oder Trauer ausdrücken können. Dies sollte aber nicht darüber hinwegtäuschen, dass diese Maschinen absolut keine Emotionen empfinden, sondern nur wir als Betrachter Emotionen in die Maschinen hineinprojizieren. Daran ist zunächst auch gar nichts auszusetzen und hat durchaus seine Berechtigung, etwa beim Einsatz von Robotern als Begleiter von alten, vereinsamten Menschen, ein Bereich, der gerade in Japan bereits in der Praxis erprobt wird. Bedenklich wird die Sache erst, wenn Menschen dazu neigen, einem Roboter Fähigkeiten zuzutrauen, die er gar nicht besitzt, nur weil der uns freundlich anzulächeln scheint. In diesem Zusammenhang wird gerne von Anthropomorphisierung gesprochen, also der Zuschreibung menschlicher Eigenschaften, und wir sollten uns gerade bei Robotern immer wieder vor Augen führen, dass dies bislang weitgehend Illusion ist.

Es bleibt allerdings die Frage, ob es in Zukunft möglich oder gar erstrebenswert sein wird, Roboter mit echten Emotionen auszustatten. Vieles spricht schließlich dafür, dass Intelligenz und Emotionen bei uns Menschen eng miteinander verknüpft sind. Das Problem ist nur, dass wir noch weit davon entfernt sind zu verstehen, was Emotionen sind und welche Rolle sie genau spielen im Zusammenhang mit Intelligenz, vielleicht abgesehen von Basisinstinkten wie Angst vor Schlangen, die bereits unseren Vorfahren halfen zu überleben. Solange wir also nicht in der Lage sind, das Phänomen Emotion wirklich wissenschaftlich zu durchdringen, wird es kaum möglich sein, Artefakte wie Roboter damit auszustatten, selbst wenn wir es wollten.

Ganz ähnlich sieht es beim Thema Bewusstsein aus. Jeder gesunde Mensch hat es, aber wir wissen immer noch nicht, wie es zustande kommt und auf

welchen Prinzipien es beruht. Daher ist es auch nicht verwunderlich, dass Roboter bisher keinerlei Bewusstsein besitzen. Nichtsdestotrotz ist Bewusstsein Gegenstand intensiver Forschung in den Neurowissenschaften. Christof Koch, einer der führenden Köpfe auf diesem Gebiet,¹⁰ ist sogar recht optimistisch, dass es in Zukunft gelingen wird, den Geheimnissen dieses Phänomens auf die Spur zu kommen, auch wenn der Gedanke an Roboter mit bewusstem Verstand noch sehr utopisch erscheint, ganz abgesehen von der Frage, ob wir solche Geschöpfe wirklich unter uns haben möchten.

Bei Kreativität und Robotern sieht die Lage vielleicht etwas hoffnungsvoller aus, zumindest auf den ersten Blick. Es gibt zum Beispiel Maschinen, die in der Lage sind, gefällige Musikstücke zu komponieren oder Bilder zu malen, die aussehen, als stammten sie von van Gogh oder Rembrandt. Nur wenige Menschen wären dazu in der Lage, und wenn, dann würden wir Ihnen sicher ein gehöriges Maß an Kreativität zubilligen. Aber ob diese Maschinen wirklich kreativ sind, mag durchaus bezweifelt werden. Es kann gut sein, dass sie doch nur auf bestimmte, vortrainierte Muster reagieren und uns damit Kreativität nur vorgaukeln. Ein besonderer Aspekt der menschlichen Kreativität liegt darin, dass sie sich oft auf mehr als einen Bereich erstreckt. So war Einstein nicht nur ein genialer Physiker, sondern auch ein recht guter Musiker mit einer Vorliebe für Mozart. Ob es einmal Maschinen geben wird mit vergleichbaren Eigenschaften, steht noch in den Sternen.

5. Müssen wir uns keine Sorgen machen?

Ein Fazit aus den beiden vorherigen Abschnitten sollte vielleicht sein, dass die kognitive Robotik sich seit Shakey über die Jahrzehnte zwar eindrucksvoll weiterentwickelt hat, diese Maschinen aber im Vergleich zum Menschen immer noch recht schlichte Gesellen sind, die schnell an ihre kognitiven Grenzen stoßen. Mit anderen Worten, wir sind von Dystopien, bei denen Roboter die Weltherrschaft übernehmen und uns Menschen ersetzen, zum

¹⁰ Faszinierende Einblicke in die Gedankenwelt von Christof Koch liefert sein autobiographisches Werk: Koch, Ch., *Bewusstsein: Bekenntnisse eines Hirnforschers*, Wiesbaden 2013.

Glück noch weit entfernt. Es ist aus heutiger Sicht noch nicht einmal klar, ob dies je möglich sein wird. Müssen wir uns also keine Sorgen machen, dass Roboter in absehbarer Zeit Unheil anrichten werden? Leider ist dem nicht so.

Das Problem besteht nämlich darin, dass es keiner besonderen Intelligenz und schon gar keiner Superintelligenz bedarf, um mit Robotern großen Schaden anzurichten. Denken wir zum Beispiel an unbemannte Flugkörper, besser bekannt als Drohnen. Man kann sie leicht mit Kameras und anderen Sensoren bestücken, und die heutigen Methoden der künstlichen Intelligenz reichen aus, damit sie sich autonom und sicher in der Luft bewegen können. Mehr noch, sie können ohne weiteres Objekte am Boden erkennen, und dazu gehören eben auch Menschen. Von hier ist es nicht mehr weit sich vorzustellen, wie man aus einer Drohne eine Waffe macht, die selbständig Jagd auf Menschen macht. Es bedarf auch keiner großen Fantasie sich vorzustellen, dass die militärische Forschung nicht nur der Großmächte sich intensiv mit diesem Thema beschäftigt. Von Menschen ferngesteuerte Drohnen gibt es schon seit langem, aber die neue Qualität liegt in der Möglichkeit, kostengünstig Waffensysteme zu bauen, die völlig autonom in der Lage sind zu entscheiden Menschen zu töten.

Diese imminente Gefahr bewegt die KI-Forschergemeinde schon seit einigen Jahren und hat 2015 zu einem offenen Brief geführt, der zu einem Verbot solcher Waffen aufruft und der bereits von mehr als 30 000 Menschen unterschrieben wurde, darunter 4 500 KI-Forscher, aber auch bedeutenden Persönlichkeiten wie Elon Musk oder dem leider verstorbenen Stephen Hawking.¹¹ Um der Öffentlichkeit die Bedrohung zu veranschaulichen, haben Mitglieder der Bewegung, darunter Professor Stuart Russell aus Berkeley, den Videoclip *Slaughterbots* drehen lassen, in dem mit drastischen Szenen die Jagd auf Menschen durch Drohnen simuliert wird und der bereits millionenfach auf YouTube abgerufen wurde. Darüber hinaus gibt es auch Bestrebungen der UNO, ein Verbot dieser todbringenden Waffensysteme zu erwirken. Allerdings sind die Erfolgsaussichten dafür zurzeit leider eher gering, aber aufgeben sollte man deshalb sicher nicht, denn die Konsequenzen eines Scheiterns könnten verheerend sein.

¹¹ <https://futureoflife.org/open-letter-autonomous-weapons>, Zugriff 7. April 2021.

6. Roboter und Freiheit

Auch wenn Roboter derzeit noch in vielerlei Hinsicht unzulänglich sind, ist zu erwarten, dass sich die Technik und insbesondere die kognitiven Fähigkeiten stetig weiterentwickeln werden, und sie mehr und mehr Einzug in unser tägliches Leben halten werden. Daher lassen Sie mich zum Schluss noch auf die Frage eingehen, inwieweit Roboter auch Einfluss auf unsere Freiheit nehmen können.¹² Ich möchte mich dabei auf ganz praktische Erwägungen beschränken.

Ein vielleicht offensichtlicher und durchaus positiver Aspekt dürfte sein, dass uns Roboter Tätigkeiten abnehmen, von denen wir gerne befreit werden. Schon heute gilt das, wenn auch mit Abstrichen, für das Staubsaugen in der Wohnung und später vielleicht für anspruchsvollere Tätigkeiten im Haushalt wie Abwaschen oder Hemden bügeln. Die meisten von uns dürften sich darüber freuen und die gewonnene freie Zeit für angenehmere Dinge nutzen oder auch dazu einfach einmal nichts zu tun. Schon jetzt hat man natürlich die Freiheit nicht zu staubsaugen oder die Wohnung nicht aufzuräumen, aber die Konsequenzen dürften auf Dauer für die meisten unerträglich sein. Also halten wir fest, dass uns eine Roboterhaushaltshilfe gelegen käme, uns Freiheiten für angenehmere Dinge im Leben zu schaffen.

Darüber hinaus kann es aber auch dazu kommen, dass uns Roboter Dinge abnehmen, die uns selbst Spaß machen, wir jedoch keine Wahl haben uns dagegen zu entscheiden.¹³ Ein Szenario wäre wiederum das autonome Fahren. Es ist durchaus denkbar, dass es irgendwann in der Zukunft keine von Menschen gesteuerten Fahrzeuge mehr geben wird. Dafür mag es viele gute Gründe geben: Sicherheit, Klimaschutz, Ökonomie und vieles mehr. Aber es gibt nun einmal Menschen, zu denen auch ich gehöre, denen Autofahren Spaß macht. Wenn einem die Möglichkeit genommen wird selbst ein Fahr-

¹² Man könnte sich auch Gedanken über den Freiheitsbegriff in Bezug auf Roboter machen. Wegen der auf absehbare Zeit doch sehr begrenzten kognitiven Fähigkeiten von Robotern halte ich dieses Thema allerdings für verfrüht.

¹³ Wem Hausarbeit Spaß macht, der hätte wahrscheinlich immer die Möglichkeit sich keinen Haushaltsroboter anzuschaffen.

zeug zu steuern, käme das für viele mit einem Gefühl des Freiheitsverlusts einher.

Davon abgesehen sollte man sich vielleicht auch über einen gewissen Kontrollverlust Gedanken machen, wenn man von einem autonomen Fahrzeug von A nach B befördert wird. Schließlich wird so ein Auto in der Regel selbst eine Route wählen, ob sie uns nun passt oder nicht. Die meisten von uns haben dies sicher schon beim Taxifahren in einer uns bekannten Umgebung erlebt, wenn wir uns wundern, warum der Fahrer gerade diese Strecke wählt. Ich denke Kontrollverlust wird immer eine gewisse Rolle spielen, sobald wir Aufgaben auf Roboter übertragen. Falls wir die Wahl haben, etwa im eigenen Haushalt, sollte man immer abwägen, ob die Vorteile wie die gewonnene Freiheit sich anderen Dingen zu widmen es wirklich wert sind.

Neben Kontrollverlust gibt es noch einen weiteren Aspekt, den es beim Übertragen von Aufgaben auf Roboter zu beachten gilt. Es könnte sehr wohl dazu führen, dass wir Menschen dadurch grundlegende Fähigkeiten verlieren. Man denke zum Beispiel an Navigationssysteme, sei es im Auto oder auf dem Smartphone. Wenn wir uns nur noch auf solche Technik verlassen, vernachlässigen wir unseren eigenen Orientierungssinn, der bisher ohne Übertreibung eine zentrale Rolle im Überleben unserer Spezies gespielt hat. Oder denken wir an die großen Fortschritte der letzten Jahre im Bereich der Übersetzung von einer Sprache in eine andere, wodurch der dolmetschende Roboter C-3PO aus der *Star Wars* Serie ein Stück weit realistischer geworden ist. Und dann stellt sich irgendwann die Frage, warum wir noch Fremdsprachen lernen sollen, wenn Maschinen das vielleicht viel besser können als wir. Aber würde uns damit nicht auch ein Stück unserer eigenen Intelligenz verloren gehen?

Gesellschaftlich brisanter ist allerdings die Frage, ob Roboter nicht eines Tages Arbeitsplätze vernichten werden. Das Problem ist heute schon akut im Zuge der Automatisierung, wo insbesondere im Bereich der Produktion bereits viele Menschen ihren Job verloren haben, und es ist absehbar, dass sich dieser Trend, gerade im Hinblick auf die rasante Entwicklung der künstlichen Intelligenz, auch im Dienstleistungsbereich fortsetzen wird. Um noch einmal das Beispiel des autonomen Fahrens zu bemühen: falls dies eines Tages tatsächlich Realität wird, werden allein in Deutschland hunderttausende

Berufsfahrer und Kuriere ihre Arbeit verlieren. Bis uns humanoide Roboter Arbeitsplätze wegnehmen, wird es dagegen sicher noch eine ganze Weile dauern, schon allein, weil deren Fähigkeiten, wie wir gesehen haben, noch sehr beschränkt sind. Mit anderen Worten, die Berufe des Gärtners oder Tischlers, in denen sowohl Geschicklichkeit als auch Kreativität gefordert sind, wird es voraussichtlich noch lange geben. Dennoch werden sich Politik und Gesellschaft Gedanken machen müssen, wie Arbeit in Zukunft gestaltet und auf die Menschen verteilt werden soll. Optimisten werden vielleicht sagen, dass es bisher immer gelungen ist, neue Arbeitsplätze zu schaffen. Natürlich wird es auch neue Berufe geben. Jemand muss schließlich die Hard- und Software für die Roboter der Zukunft entwerfen, aber nicht jeder hat das Talent Ingenieur zu werden. Und so wird es mit einiger Wahrscheinlichkeit Verlierer geben, um die wir uns kümmern müssen. Ohne sinnvolle Arbeit dürfte es vielen schwer fallen ein erfülltes Leben zu führen, und auch das hat in meinen Augen mit Freiheit zu tun.

7. Abschließende Bemerkungen

Wir haben gesehen, dass kognitive Roboter heute noch weit entfernt sind von den Exemplaren, die wir aus Science-Fiction Filmen kennen, und schon gar von den dunklen Gesellen, die mit übermenschlicher Intelligenz die Herrschaft über die Welt erlangen wollen. Und das ist auch gut so. Vielleicht kann dieser Beitrag ein wenig helfen, etwaige Befürchtungen in diese Richtung zu zerstreuen. Aber auch wenn Roboter heute oft noch unbeholfen sind, wird die Entwicklung voranschreiten, und sie werden irgendwann in unseren Alltag Einzug nehmen und unser Leben beeinflussen. Aufgabe der Wissenschaft ist es dafür zu sorgen, dass diese Maschinen sich uns gegenüber als vertrauenswürdig erweisen, und dazu gehören sowohl Aspekte der Sicherheit als auch ethisch korrektes Verhalten. Um Missverständnissen vorzubeugen, wird es auch wichtig sein, dass die Hersteller von Robotern transparent kommunizieren, was ihre Maschinen können und was nicht. Nur so, glaube ich, wird es gelingen, dass Menschen bereit sind Roboter in ihrer Umgebung nicht nur zu dulden, sondern im Bewusstsein ihrer Stärken und Schwächen

bestmöglich zu nutzen. Mit etwas Glück sollte es dann auch gelingen, dass Roboter uns helfen neue Freiheiten zu gewinnen.

Andreas Schuppert

Big Data in der Medizin – Bedrohung oder Chance für die Freiheit?

1. Medizin und Freiheit – ein Widerspruch?

Um Big Data und Künstliche Intelligenz (KI) in der Medizin im Kontext von Freiheit zu diskutieren, lohnt es sich, ein gemeinsames Verständnis von medizinischem Handeln überhaupt im Kontext von Freiheit zu erarbeiten.

Medizin und Freiheit scheinen im besten Fall zwei völlig verschiedene Bereiche zu sein. Wenn Medizin im Kontext Freiheit diskutiert wird, dann wird häufig die Medizin als Verlust von Freiheit diskutiert. Und sind nicht in der Tat medizinische Maßnahmen, seien es Operationen, Dialysen oder die regelmäßige Einnahme von Medikamenten, Situationen, in denen dem einzelnen von einer Krankheit betroffenen Patienten die Verfügungsmacht über sein Handeln genommen ist, eine klare Einschränkung von Freiheit? Noch mehr, im Kontext des Umgangs mit der Pandemie SARS-CoV2, werden von Regierungen jeder Couleur auf Anraten von Medizinern freiheitsreduzierende Maßnahmen nicht nur einzelnen Patienten, sondern allen Menschen, vorgeschrieben. Ist dies eine internationale Diktatur der Virologen, ein von der Medizin beinahe im Sinne einer Verschwörung betriebener Entzug der Freiheit nicht nur von Patienten, sondern allen Menschen, wie manche behaupten?

Diese Haltung übersieht jedoch, dass Medizin selbst kein unabhängiges Handeln darstellt, sondern stets eine Reaktion auf eine unerwünschte Veränderung des Gesundheitszustands, in diesem Sinne einer Krankheit, eines oder mehrerer Menschen ist. Es stellt sich natürlich die Frage, wie ein uner-

wünschter Gesundheitszustand definiert ist. Hier soll davon ausgegangen werden, dass der betroffene Mensch selbst seinen Gesundheitszustand subjektiv als unerwünscht empfindet, darunter leidet und damit zum Patienten wird. Diese subjektive Empfindung kann dabei durchaus von einer sogenannten objektiven Empfindung im Sinne der Einschätzung durch Dritte abweichen.

In der Medizin muss dabei unterschieden werden zwischen der kurativen Medizin, die Heilung im Sinne eines Zurückversetzens des Gesundheitszustandes eines Menschen in seinen Ausgangszustand zum Ziel hat, der palliativen Medizin, die nicht den Gesundheitszustand selbst verändern will, sondern bei gegebenem Gesundheitsverlust ein würdevolles, schmerzfreies und möglichst selbstbestimmtes Leben bis zum Tode anstrebt, sowie der präventiven Medizin, die das Auftreten der Krankheit selbst zu verhindern versucht und deren Handlungsfeld daher nicht der Patient, sondern der „gesunde“ Mensch ist. Letztere wird daher häufig nicht der Medizin, sondern dem Themenfeld „Public Health“ zugeordnet. In allen drei Feldern der Medizin finden wir Daten und deren Nutzung im Sinne der jeweiligen medizinischen Zielsetzung. Besonders im Bereich der präventiven Medizin spielt die Auswertung großer Datenmengen, „Big Data“, eine essenzielle Rolle. Besonders wenn kurative Maßnahmen nicht vorhanden sind und der präventive Ansatz die einzige Möglichkeit ist Menschenleben zu retten, können aus der Analyse der Daten auch freiheitsbeschränkende Maßnahmen hergeleitet werden, wie sich in der SARS-CoV2-Pandemie gezeigt hat. Allerdings muss gerade am Beispiel der SARS-CoV2-Pandemie klar betont werden, dass die eigentliche Bedrohung der Freiheit das Virus ist, das sich nicht an irgendwelche zivilisatorischen Regeln hält. Die Analyse der Daten ist daher nur in der undankbaren Rolle des Überbringers der schlechten Nachricht, keinesfalls die Ursache selbst.

In Abgrenzung zu den oben genannten medizinischen Handlungsfeldern werden Big Data und Künstliche Intelligenz heute schon intensiv im Kontext Lifestyle und Selbstoptimierung eingesetzt. Hier ist die Zielsetzung jedoch nicht die Heilung oder Vermeidung einer Krankheit, sondern eine wie auch immer definierte Verbesserung der Leistungsfähigkeit des Körpers. Die Diskussion dieser Anwendungen von Big Data soll hier nicht geführt werden.

Eine Diskussion der freiheitsreduzierenden Maßnahmen der präventiven, kurativen oder palliativen Medizin muss daher stets im Kontext der Krankheit erfolgen. Dabei ist eine genaue Definition von Krankheit nicht einfach. Viele Krankheiten wie Influenza, Krebs, Alzheimer oder die Masern sind offensichtlich negativ empfundene, jedoch natürliche, nicht durch menschliches Handeln hervorgerufene Veränderungen unseres Gesundheitszustands, die mit einer Einschränkung der individuellen Freiheit im Sinne einer Einschränkung von Handlungsoptionen der betroffenen Patienten einhergehen. Die Krankheit ist dabei, wie wir heute wissen, primär nicht eine Folge des eigenen Handelns, sondern ein natürlicher Prozess, der weitgehend zufallsgetrieben abläuft. Beeinflussbar ist die Erkrankung in diesem Sinne nicht primär, sondern nur durch Erhöhung der eigenen Widerstandsfähigkeit, der Resilienz, gegenüber den zufällig auftretenden krankmachenden Faktoren, einem der Felder der Präventivmedizin.

Diese unerwünschten Veränderungen des Gesundheitszustands werden ohne weiteres als Krankheiten verstanden, kuratives medizinisches Handeln soll „wieder gesund“ machen, d. h. die alten Handlungsoptionen wiederherstellen und damit die durch die Krankheit verlorene Freiheit wieder rekonstituieren. Medizin ist also Wiedererschafferin der Freiheit, die uns durch den natürlichen Prozess der Krankheit genommen wurde.

Problematischer sind hingegen Gesundheitszustände, die als unerwünscht empfunden werden, die jedoch nicht auf einem Veränderungsprozess beruhen, sondern genetisch bedingt sind, d. h. in der Regel bei der Verschmelzung von Eizelle und Samenzelle beim Befruchtungsakt zustande gekommen sind, erweitert um Zustände, die sich während der embryonalen Entwicklung eines Menschen vor der Geburt entwickeln. Viele dieser Zustände schränken die Handlungsoptionen, damit die Freiheit, der Betroffenen teilweise massiv ein, sie sollen daher hier ebenfalls als Krankheiten im Sinne eines medizinischen Handlungsfelds gesehen werden.

Deutlich problematischer müssen jedoch Gesundheitszustände gesehen werden, die nur spezielle Handlungsoptionen der Betroffenen einschränken. Ist eine genetische Mutation, die zur Laktoseunverträglichkeit führt, eine Krankheit? Sie beschränkt ohne Zweifel unsere Freiheit in Bezug auf den Konsum von Milchprodukten. Dasselbe gilt für Mutationen, die mit einer Al-

koholunverträglichkeit einhergehen. Ist der Entzug der Freiheit, Alkohol zu trinken, eine Krankheit? Noch problematischer muss die Einschränkung der Freiheit zu Spitzenleistungen in spezifischen Handlungsfeldern durch genetische oder andere gesundheitliche Dispositionen gesehen werden. Ist die gesundheitsbedingte Unfähigkeit, 100 Meter in 10 Sekunden zu laufen, eine Krankheit? Eher nicht, obwohl die Freiheit, die meisten Sportdisziplinen auf höchstem Niveau zu betreiben, dadurch eingeschränkt ist. Allerdings sind gerade diese Freiheiten Ziele der Selbstoptimierungsbestrebungen, durchaus auch unter Einsatz von Big Data und künstlicher Intelligenz.

In diesem Kontext muss das Thema Altern ebenfalls diskutiert werden, da es einer der Haupttreiber von Erkrankungen ist. Alterung ist ein objektiv natürlicher Prozess, der alle Individuen betrifft und subjektiv meist als unerwünscht empfunden wird, ohne Zweifel nicht zuletzt durch einen damit einhergehenden Verlust von Handlungsoptionen und daher Freiheit. Ist Alterung daher eine Krankheit? Die Tatsache, dass Alterung mit einer Reduktion der Resilienz einhergeht und damit einem zunehmenden Erkrankungsrisiko, macht Alterung zunehmend zu einem medizinischen Handlungsfeld. Seine Universalität macht es zu einem Forschungs- und Investitionsfeld für Big Data und Künstliche Intelligenz. Trotzdem, oder vielleicht gerade deswegen, soll hier auf die speziellen Herausforderungen von Big Data für die Bekämpfung der Alterung nicht eingegangen werden.

In diesem Aufsatz werden Big Data und Künstliche Intelligenz nur in Bezug auf medizinische Handlungsfelder diskutiert, die eine Prävention oder Reversion von schwerwiegenden, individuellen Einschränkungen der Gesundheit, begleitet von Einschränkungen der individuellen Freiheit, anstreben.

2. Medizin und Big Data – bringt Big Data die Medizin weiter in Richtung zu mehr Freiheit?

2.1. Handlungsfeld kurative Medizin – Big Data und die effiziente Entwicklung neuer Therapien

Fortschritte der kurativen Medizin erfordern die Entwicklung neuer Therapien, entweder im Sinne einer Verbesserung des Standes der Kunst, durch schnellere Wirkung oder weniger Nebenwirkungen, oder die Entwicklung von Therapien gegen Erkrankungen, die heute noch nicht therapiert werden können. Am Beispiel der Entwicklung neuer Medikamente soll diskutiert werden, warum die Analyse großer Patientendatensätze durch Big Data-Analysetechnologien für die Entwicklung von Therapien gegen eine Vielzahl bisher nur unzureichend therapierbaren Erkrankungen ein notwendiges Hilfsmittel sein kann.

Kurative Medizin mit dem Ziel der Heilung von Krankheiten und Verletzungen, folglich der naturbedingten Reduktion von Einschränkungen der Handlungsfreiheit des Menschen, ist neben der Produktion von Nahrung zweifellos eine der ältesten Tätigkeiten der Menschen, die schon in der Steinzeit betrieben wurde. Über Jahrtausende hinweg wurde sie, soweit bekannt, auf Basis von Beobachtungen sowie Versuch und Irrtum betrieben, assoziiert mit Versuchen zu einer kausalen, mechanistischen Erklärung der Gesundheitsstörungen. Gesundheitsstörungen wurden dabei häufig als Störungen eines Gleichgewichts interpretiert, sei es als Störung der Harmonie zwischen einem Gott und dem Menschen, dem Befall des Menschen durch einen Dämon, wie zum Beispiel in der Bibel berichtet, oder als Ungleichgewicht von Säften bei Hippokrates und Galen¹ oder Ungleichgewicht von Yin und Yang in der chinesischen Medizin.² Daraus abgeleitet entwickelten sich Diagnose- und Therapieansätze, wobei empirische Erfahrungen und Beobachtungen gesammelt und tradiert wurden. Dabei stammten die Daten wie heute aus

¹ https://de.wikipedia.org/wiki/Corpus_Hippocraticum, Zugriff 27. März 2021.

² https://de.wikipedia.org/wiki/Traditionelle_chinesische_Medizin#Das_Qi_in_der_Spannung_von_Yin_und_Yang, Zugriff 27. März 2021.

Beobachtungen, die jedoch wegen mangelnder technologischer Möglichkeiten nicht weltweit vernetzt gespeichert und ausgewertet wurden, sondern in erfahrungsbasierten Regelwerken individueller Ärzte gespeichert und tradiert wurden. Insofern kann die Medizin des Altertums bis ins Mittelalter als eine Anwendung von Big Data interpretiert werden, wobei durch die erfahrungsbasierte Erstellung von Regelwerken eine implizite Anonymisierung der Daten erfolgte. Dass dieser beobachtungsgetriebene Ansatz auch bei nach unserem heutigen Verständnis völlig falschem mechanistischem Verständnis zu verblüffenden praktischen Erfolgen führte, ist durch Funde dokumentiert.

Erst in der Renaissance wurden Leichen seziiert,³ um ein tiefergehendes mechanistisches Verständnis über die Krankheiten und ihnen zu Grunde liegenden Mechanismen zu gewinnen und damit die Beschränkungen der rein beobachtungsgetriebenen Medizin der Antike und des Mittelalters zu überwinden.

Empirisch gesehen waren diese neuen Ansätze jedoch lange Zeit eher ein Misserfolg, zumindest im Sinne des Patientenwohls. Entscheidende Fortschritte der Medizin, zum Beispiel im Bereich der Wundversorgung, wurden von Praktikern, oft Feldärzten bei Truppen, durch Beobachtung und Zufall erreicht. Ebenso die Entwicklung der Impfungen, sei es durch Inokulation oder später durch Vakzination, wurde durch beobachtende Landärzte wie Edward Jenner entdeckt und entwickelt.⁴ Auch die Entwicklung der Hygiene brachte entscheidende Verbesserungen im Gesundheitszustand der Bevölkerung, wurde jedoch ebenfalls von statistisch arbeitenden Praktikern auf Basis vergleichender, sorgfältiger Beobachtungen vorangetrieben. Erst durch die Entdeckung der Bakterien und ihrer Identifizierung als kausalem Auslöser von Krankheiten, unter anderen durch Louis Pasteur und Robert Koch, die Pathologie sowie die Entwicklung der Narkose und der aseptischen Operationsmethoden konnte die wissenschaftlich-kausal orientierte Medizin ihr Potential entfalten. Dennoch wurden auch später wesentliche Erfolge

³ Vesalius, A., *De humani corporis fabrica*, 1543.

⁴ <https://www.aerzteblatt.de/archiv/3914/Edward-Jenner-200-Jahre-Pockenschutz>, Zugriff 27. März 2021.

durch Beobachtung und Empirie erzielt. So dauerte die Aufklärung des Wirkungsmechanismus von Aspirin – zugelassen 1895⁵ – mehr als 90 Jahre.

Gerade bei der medikamentösen Behandlung von Krankheiten und ihrer Folgen dominierte trotz der – bis heute im Wesentlichen akzeptierten – Erklärung der Chemotherapie durch Paul Ehrlich⁶ lange ein eher zufallsgetriebener Prozess: Neue Moleküle wurden in den Laboren der chemischen Industrie synthetisiert und dann auf ihre Anwendbarkeit in unterschiedlichsten Einsatzfeldern, unter anderem auch als Pharmakon, getestet. Alternativ wurde versucht, Naturstoffe mit bekannter Wirkung chemisch so zu optimieren, dass ihr Potential erhalten blieb, ohne die Nebenwirkungen akzeptieren zu müssen. Im Bereich der Pharmaka bildet gerade die Acetylsalicylsäure ein gutes Beispiel: die schmerzlindernde Wirkung der Salicylsäure als Komponente der Weidenrinde *Salix* war lange bekannt, allerdings auch ihre erheblichen Nebenwirkungen und schlechte Dosierbarkeit. Erst durch die gezielte Modifikation durch Acetylierung konnte daraus ein verträgliches, gut dosierbares Medikament im großen Stil produziert werden. Allerdings zeigte es sich seit den 1980er Jahren, dass die Effizienz dieses eher zufallsgetriebenen Ansatzes deutlich nachließ. Zum einen lag das daran, dass für die medikamentös „einfachen“ Erkrankungen, zum Beispiel bakterielle Infektionen, bei denen es ausreicht, ein Bakterium abzutöten, ohne dem Patienten zu schaden, durch die Entdeckung der Antibiotika eine prinzipielle Lösung vorlag. Hierfür maßgeblich ist die Tatsache, dass Bakterien und Menschen sich so fundamental unterscheiden, dass es ausreichend viele für das Bakterium essenzielle Mechanismen gibt, die beim Menschen nicht vorkommen. Daher ist eine Blockade dieser Mechanismen für das Bakterium tödlich, für den Menschen jedoch harmlos. Leider ist dieser biologische Unterschied zwischen dem „krankmachenden Faktor“ und der Biologie des Patienten

⁵ Kuhnert, N., Hundert Jahre Aspirin, *Chemie in unserer Zeit* 33 (4) (1999), 213–220, doi:10.1002/ciuz.19990330406.

⁶ Valent, P./ Groner, B./ Schumacher, U./ Superti-Furga, G./ Busslinger, M./ Kralovics, R./ Zielinski, C./ Penninger, J. M./ Kerjaschki, D./ Stingl, G./ Smolen, J. S./ Valenta, R./ Lassmann, H./ Kovar, H./ Jäger, U./ Kornek, G./ Müller, M./ Sörgel, F., Paul Ehrlich (1854–1915) and His Contributions to the Foundation and Birth of Translational Medicine, *Journal of innate immunity* 8 (2) (2016), 111–120, doi:10.1159/000443526, PMID 26845587.

eher die Ausnahme: Krebszellen sind Zellen des Patienten, die „nur“ nicht mehr geregelt funktionieren. Sie enthalten dieselben Funktionalitäten wie die gesunden Zellen des Patienten, nutzen diese jedoch in unregelmäßiger Intensität. Eine molekulare Intervention in Krebszellen ohne Schaden für gesunde Zellen gestaltet sich daher naturgemäß deutlich schwerer. Die mit der „klassischen“ Methode gefundenen Chemotherapeutika hatten daher starke Nebenwirkungen. Außerdem blieb die Wirkstoffsuche für weite Klassen von Erkrankungen, von Autoimmunerkrankungen über neurodegenerative Erkrankungen bis hin zu viralen Infekten unbefriedigend.

Erst die Entwicklung der Biochemie, die zelluläre Prozesse als eine Abfolge chemischer Reaktionen analysiert, erlaubte den Beginn einer systematischen Erforschung und Entwicklung von Pharmaka. Einen weiteren Schub brachten die Entschlüsselung der Proteinstruktur durch Proteinkristallographie, die Entschlüsselung des menschlichen Genoms sowie der Verfügbarkeit von Hochdurchsatzscreeningverfahren. Gerade durch letztere zeigt sich, wie ein im Prinzip alter Ansatz, nämlich die zufallsgetriebene Suche nach neuen, therapeutisch wirksamen Substanzen, durch neue Technologien optimiert werden konnte: es wurde möglich, mehrere 100 000 Substanzen pro Tag auf ihre grundsätzliche Wirksamkeit an zellulären Modellen der Krankheit, die durch molekularbiologische Verfahren entwickelt werden, automatisiert zu testen. Aufbauend auf den – viel zu vielen – Treffern in diesem ersten Prozess, wurde in der Folge der pharmakologische Forschungsprozess systematisiert, um sukzessive die aussichtsreichsten Moleküle zu identifizieren, die dann in einem stark regulierten klinischen Testprozess in drei Phasen auf ihr therapeutisches Potential sowie akzeptables Risikoprofil am Patienten getestet werden. Nur wenn alle Phasen erfolgreich durchlaufen wurden, kann ein neues Molekül als Medikament zugelassen werden. De facto wurde durch Kombination von unterschiedlichsten Technologien ein weitgehend standardisierter Workflow etabliert, der jedoch letztlich auf den Ergebnissen des immer noch weitgehend zufallsgetriebenen Screenings beruht.

Trotz einiger Erfolge dieses Ansatzes für einzelne Erkrankungen blieb der erhoffte große Erfolg auf breiter Front aus. Gerade bei der Entwicklung von Medikamenten gegen komplexe Erkrankungen wie Krebs, Alzheimer oder

Autoimmunerkrankungen versagen mehr als 80 % der im Labor erfolgreichen Medikamentenkandidaten in der klinischen Prüfung. Während noch vor 20 Jahren viele Substanzen in einer frühen Phase der klinischen Prüfung ausschieden, fallen heute die hoffnungsvollen Medikamentenkandidaten oft erst in der letzten klinischen Prüfung, den sogenannten Phase-3-Tests mit mehreren tausend Patienten, wegen mangelnder Wirksamkeit oder einem inakzeptablen Risikoprofil aus. Nicht zuletzt hierdurch haben sich die Kosten und die Entwicklungsdauer neuer Medikamente trotz aller Investitionen in Technologie und Forschung in den letzten zwanzig Jahren deutlich erhöht. Schlimmer noch, für viele Erkrankungen besonders der alternden Gesellschaften, wie Alzheimer, fehlen therapeutische Optionen.

Warum blieb der wissenschaftlich-mechanistische Ansatz hinter den Erwartungen zurück? Ein wesentlicher Grund hierfür liegt in der Grundannahme, dass die zellulären Modelle, an denen im ersten Forschungsschritt die Moleküle auf potenzielle Wirksamkeit getestet werden, tatsächlich die kausalen Ursachen der Erkrankung repräsentieren. Dies ist in der Regel durchaus der Fall. Bedeutet dies jedoch auch, dass eine Intervention am fehlerhaften molekularen Mechanismus in der Zelle auch zu einer Heilung der Krankheit führt? Auch diese Annahme trifft teilweise durchaus zu, wie das – frühe – Beispiel der Chronischen Myeloischen Leukämie zeigte, der durch genau diesen Ansatz ihr tödlicher Schrecken genommen wurde. Weitere Erfolge waren gezielt ansetzende Krebsmedikamente, die deutlich weniger schwere Nebenwirkungen aufweisen als die alten Chemotherapeutika und zumindest das Fortschreiten der Erkrankung verzögern. Leider blieben sie eher die Ausnahme als die Regel. Auch konnte die erhoffte kurative Wirkung meist nicht erreicht werden.

Es zeigte sich zum einen, dass eine Intervention auf zellulärer Ebene durch ein Medikament, die in der Regel die Blockade der Wirkung eines funktionalen Moleküls in den Zellen bedeutet, als Antwort der Zelle eine Ausweichreaktion hervorruft, entweder durch Selektion von Zellen, bei denen die Blockade zufällig erfolglos blieb, Umorganisation der innerzellulären Mechanismen⁷ oder gar durch Umorganisation der extrazellulären Umgebung.

⁷ Lee, M. J. et al., Cell, 149(4) (2012), 780–794.

Diese zur Resistenz führenden Mechanismen können mehrere Stunden bis Wochen benötigen und lassen sich daher im Laborversuch, besonders im auf hohen Durchsatz optimierten Screeningverfahren, nur sehr schwer analysieren.

Ein weiteres Problem liegt in der hohen Komplexität der Erkrankungsmechanismen sowie der Wirkungsmechanismen von Medikamenten im Menschen, die durch eine Vielzahl von Faktoren, sogenannten Covariaten, beeinflusst werden: angefangen von der genetischen Disposition des einzelnen Individuums, dem Lebensstil, anderen Co-Erkrankungen und dem Alter beeinflussen viele anderen Faktoren den Verlauf einer Erkrankung sowie den Erfolg einer Therapie. Es ist offensichtlich, dass diese Faktoren im Laborversuch nicht nachgestellt werden können.

Diese Aporie begleitet die wissenschaftliche Entwicklung der Medizin seit ihrem Ursprung. Da bis heute keine ethisch vertretbaren Alternativen verfügbar sind, wurde das Design der klinischen Studien daraufhin ausgerichtet, dass (nur) der direkte Nutzen der Medikamente für die Zielerkrankung und (nur) die direkten Nebenwirkungen getestet werden. Um die Zahl der Probanden in der Studie sowohl aus ethischen als auch aus Kostengründen möglichst klein zu halten und trotzdem ein klares statistisches Signal zu bekommen, werden daher die Tests in erster Linie an Probanden durchgeführt, die nur die Zielerkrankung haben dürfen und keine weiteren Medikamente einnehmen. Dieses Vorgehen ist bis heute alternativlos, allerdings repräsentieren dadurch die Studienergebnisse nicht notwendigerweise die Wirkungen und Nebenwirkungen in der realen Gesellschaft, bei der besonders ältere Patienten häufig unter mehreren Erkrankungen leiden und auch mehrere Medikamente parallel einnehmen müssen, was zu Recht kritisiert wurde.

Daher wurde seit der Jahrtausendwende zunehmend die Analyse der Daten von „wirklichen“ Patientenkohorten als komplementärer Ansatz eingeführt, die sogenannte „Real World Evidence“(RWE)-Studie. Hierbei werden sogenannte Beobachtungsdaten ausgewertet, Daten, die beim normalen Arzt- oder Klinikaufenthalt erhoben werden. Zusätzliche Daten, die in einer gezielten klinischen Studie erhoben werden, fallen bei RWE-Studien daher nicht an. Das Ziel ist dabei, zum einen die Wirkung von Therapien in der realen Welt vor dem Hintergrund der Vielzahl von individuellen Prädispositi-

onen der Patienten kennenzulernen, zum anderen tiefere Einblicke in Details der Erkrankungen, insbesondere im Kontext ihrer Wechselwirkungen, zu erhalten. Hieraus werden neue Erkenntnisse erwartet, die helfen können, die oben diskutierten Schwachstellen im Forschungs- und Entwicklungsprozess neuer Therapien zu beseitigen. De facto kann man den RWE-Prozess interpretieren als „Back to the roots“: man versucht, wie im Altertum, aus der Beobachtung zu lernen, nur eben auf einer völlig neuen technischen Basis und in Kooperation mit dem mechanistischen Forschungsansatz.

So stringent – und alternativlos – RWE-Analysen als Komplement zum mechanistischen Forschungsansatz erscheinen, so bringen sie auch hohe wissenschaftliche und gesellschaftliche Herausforderungen mit sich:

- Beobachtungsdaten umfassen konzeptuell nur die Daten, die der behandelnde Arzt auf Basis des momentanen Stands der Medizin für diagnostisch notwendig erachtet. Es wird daher niemals zu einzelnen Patienten ein vollständiges Datenbild vorliegen, wie es in Studien erhoben wird. Die Daten haben daher in etwa die Struktur eines Schweizer Käse.
- Nicht erfasste Parameter, die einen Einfluss auf Erkrankung und Therapie haben können, sogenannte „Hidden Parameter“, stellen die Auswertung vor besondere Herausforderungen. Sie erzeugen spezielle Datenverteilungen, sogenannte Korrelationen, die mit klassischen statistischen Methoden zu fehlerhaften Ergebnissen führen können, besonders wenn sehr viele Parameter analysiert werden. Ein bekannter dokumentierter Fall ist die Häufigkeit des Arztbesuchs. Bei Patienten, die, aus welchen Gründen auch immer, häufig einen Arzt aufsuchen, werden Krankheiten tendenziell in einem frühen Stadium entdeckt und verlaufen daher weniger schwer. In den Daten kann das zu dem Paradoxon führen, dass Patienten, die an einer durchaus schweren chronischen Erkrankung leiden, eine akute Erkrankung daher deutlich besser überstehen als Patienten, die ansonsten gesund sind.
- Beobachtungsdaten repräsentieren stets die Interaktion zwischen Arzt und Patient inklusive der Therapieentscheidung des Arztes. Im Gegensatz zu klinischen Studien fehlt die Kontrollgruppe sowie die doppelte „Verblindung“, durch die weder Arzt noch Patient wissen, ob sie eine neue Therapie oder den State of the Art bekommen. Das bedeutet, dass

die Interpretation der Daten stets die Arzt-Patient-Interaktion berücksichtigen muss. Hierzu ist das Methodenarsenal leider noch unterentwickelt.

- Konzeptuell sollen alle Patienten mit ihren unterschiedlichsten individuellen Kovariaten erfasst werden, um gerade die konzeptuellen Schwächen von gezielten Studien zu kompensieren. Wegen der Vielzahl von potenziell relevanten Faktoren müssen daher per definitionem sehr viele Patienten erfasst werden, sonst kann prinzipiell das Ziel nicht erreicht werden, was zu einem konzeptuellen Konflikt mit dem aus juristischer Sicht gebotenen Prinzip der Datensparsamkeit führt.
- Die Analyse der Daten mit dem Ziel der Aufdeckung möglicher, bisher unbekannter Interaktionen verschiedenster Einflussfaktoren auf Erkrankung und Therapie erfordert neue Datenanalyseverfahren aus der Künstlichen Intelligenz, insbesondere im Kontext der oben diskutierten speziellen Datenstruktur von RWE-Beobachtungsdaten.

2.2 Seltene Erkrankungen – ohne Big Data chancenlos?

Ein Teilbereich der kurativen Medizin, der eines der klassischen Anwendungsbereiche von Big Data darstellt, ist das Feld der Seltenen Erkrankungen. Eine Seltene Erkrankung ist definiert als eine Erkrankung, die bei weniger als 5 von 10 000 Menschen auftritt,⁸ unabhängig vom auslösenden Faktor. Durch ihre Seltenheit ist es sehr unwahrscheinlich, dass ein Arzt der Primärversorgung in seinem Berufsleben mehr als einen Patienten mit derselben Seltenen Erkrankung zu Gesicht bekommt. Der durchaus bewährte Faktor der Ärztlichen Erfahrung als essenzieller Bestandteil von Diagnose und Therapie ist in diesem Fall definitionsgemäß irrelevant. Bestenfalls kann ein Arzt der Primärversorgung erkennen, dass ein vorliegendes Krankheitsbild nicht in die ihm bekannten Schemata passt. Die Ursachen Seltener Erkrankungen sind sehr vielfältig und nicht im Detail bekannt, oft sind sie genetischen Ursprungs oder die Folge einer Infektion mit einem seltenen Er-

⁸ <https://www.bundesgesundheitsministerium.de/themen/praevention/gesundheitsgefahren/seltene-erkrankungen.html>, Zugriff 27. März 2021.

reger, Parameter, die im Rahmen einer ungezielten Verdachtsdiagnose nicht erhoben werden. Daher muss die Kategorisierung eines Krankheitsbildes als Seltene Erkrankung auf Basis einer Vielzahl unspezifischer Merkmale getroffen werden, die einzeln genommen meist mit einer „normalen“ Erkrankung kompatibel sind. In der Regel weist nur das gemeinsame Auftreten vieler solcher Merkmale auf eine Seltene Erkrankung hin, wenn diese Kombination von diagnostischen Merkmalen mit keiner „normalen“ Erkrankung erklärt werden kann. Daher ist schon eine korrekte Klassifizierung als „Seltene Erkrankung“ nicht trivial. Meist werden die entsprechenden Leiden falsch klassifiziert, indem Teile des komplexen Krankheitsbildes vernachlässigt werden und damit unzureichende Therapieentscheidungen getroffen. Die berichteten, oft langjährigen Leidensgeschichten von Patienten bis zur richtigen Diagnose, legen für das Dilemma einer auf mangelnden Daten beruhenden zwangsläufig mangelnden Erfahrung der behandelnden Ärzte ein beredtes Zeugnis ab.

Ähnlich gelagert stellt sich die Situation bei komplexen Erkrankungen dar, die in einer Vielzahl therapeutisch relevanter Untergruppen vorliegen können. Dies ist zum Beispiel typisch für viele Tumorerkrankungen, die per se keine Seltenen Erkrankungen darstellen, jedoch in so vielen genetisch unterschiedlichen Variationen vorliegen können, dass jede einzelne Variante wieder recht selten ist. Auch hier ist eine Abstimmung der therapeutischen Optionen vor dem Hintergrund der kombinatorischen Details der Erkrankung geboten, ein Arbeitsfeld, das heute unter dem Stichwort Personalisierte Medizin betrieben wird. Der Unterschied zur Problematik der Diagnose Seltener Erkrankungen besteht in der Regel darin, dass die kausal relevanten genetischen Daten im Fall einer Krebsdiagnose typischerweise erhoben werden und die entsprechenden Datenbanken zur Verfügung stehen.

In beiden Fällen ist es jedoch offensichtlich, dass eine grundlegende Verbesserung der Situation durch eine systematische Analyse sehr großer Patientendatenbestände nach denselben Merkmalskombinationen gelingen kann. Dabei müssen diese Datenbestände eine automatisierte Suche auf nationaler, besser internationaler Ebene ermöglichen, was definitionsgemäß dem Einsatz von Big Data entspricht.

2.3 Präventive Medizin – ohne Big Data zum Scheitern verurteilt?

Komplementär zur kurativen Medizin, deren Objekt der schon erkrankte Mensch ist, zielt die präventive Medizin auf die Verhinderung von Krankheiten durch geeignete Maßnahmen. Hierzu gehört sowohl die Entwicklung von Impfstoffen, aber auch die Entwicklung von Maßnahmen zur Reduktion von Ansteckungsrisiken bei Epidemien, wie sie zum Beispiel im Kontext der SARS-CoV2-Pandemie dringend erforderlich sind. Darüber hinaus stellt sich aber auch die Frage, ob Nahrungsergänzungsmittel, Diäten oder andere spezielle Präventionsmaßnahmen – auf Zielgruppen mit erhöhtem Krankheitsrisiko hin zugeschnitten – einen sinnvollen und messbaren Beitrag zur Gesunderhaltung leisten können. Ganz allgemein bleibt die Frage zu beantworten, ob ein wie auch immer gearteter gesunder Lebensstil einen messbaren Beitrag zur Gesunderhaltung leistet oder nur bei einer speziellen Gruppe einen Effekt hat.

Hier stoßen wir sehr schnell auf ethische und praktische Dilemmata:

- das Entstehen von komplexen Erkrankungen aus dem gesunden Zustand heraus auf molekularer Ebene ist wesentlich weniger gut erforscht als die Krankheit selbst und deren mögliche Therapie. Ein einfacher Grund liegt darin, dass gesunde Menschen nicht zum Arzt gehen und damit auch keine diagnostischen Daten erfasst werden. Ein weiterer Grund liegt darin, dass die Erhebung der molekularen Daten, wie sie z. B. bei Krebserkrankungen durchgeführt wird, in der Regel mit einem invasiven, nicht risikofreien, Eingriff, wie einer Biopsie, verbunden ist. Dies kann mit kurativer Absicht bei Patienten ethisch geboten erscheinen, invasive Eingriffe bei gesunden Menschen nur aus Forschungszwecken sind es in der Regel nicht, insbesondere da mehrere Eingriffe notwendig wären, um die Entstehung und Entwicklung der Erkrankung zu beobachten. Außerdem ist in der Regel bei Auftreten einer Erkrankung in der frühen Phase, in der die kurativen Optionen oft noch vielversprechend sind, der direkte Beginn einer Therapie geboten, wodurch umgekehrt jedoch die Beobachtung der Entwicklung der Erkrankung in frühen Stadien unmöglich wird. Daher ist die Datenlage und in Folge das

Verständnis der Entstehung und Entwicklung von Erkrankungen bis auf wenige Ausnahmen, wie bei langsamen, noch gutartigen Tumorstufen, nicht ausreichend. Wenn überhaupt, dann liegen Daten der Erstdiagnose vor, sozusagen Schnappschüsse des Erkrankungsgeschehens. Die dynamische Entwicklung, sogenannte longitudinale Daten, sind in aller Regel nicht vorhanden.

- Ein weiteres Dilemma, das insbesondere die Überprüfung der Wirksamkeit von Präventionsmaßnahmen betrifft, ist die Zeitspanne bis zu einem eventuell sichtbaren Erfolg einer Maßnahme. Bei Fragen nach optimaler Ernährung oder anderen Maßnahmen zur generellen Erhaltung der Gesundheit wird dies besonders deutlich: um einen Erfolg messen zu können, muss das nicht-Eintreten der Krankheit beobachtet werden. Dies erfordert jedoch in der Regel eine Jahrzehnte dauernde, eine sehr große Kohorte umfassende und daher kostspielige Studie, die jegliche kommerzielle Verwertung der entsprechenden Maßnahme allein schon wegen der limitierten Patentlaufzeiten unmöglich macht.

Ein Ausweg aus den beschriebenen Dilemmata könnte wieder in der Analyse von Beobachtungsdatensätzen liegen. Dabei ist es jedoch essenziell, die Vielfalt der individuellen Heterogenität zu erfassen und in all ihren Kombinationen auf den jeweiligen Einfluss auf den Gesundheitszustand quantitativ zu erfassen. Wie noch diskutiert werden wird, nimmt jedoch der Datenbedarf, der für eine valide quantitative Aussage benötigt wird, extrem schnell mit der Zahl der potentiellen Einflussparameter zu. Die – im schlimmsten Fall – notwendige Datenmenge folgt dabei dem aus dem bekannten indischen Schachbretträtsel bekannten exponentiellen Verhalten: wenn es für jeden Einflussparameter nur zwei mögliche Zustände gibt, etwa positiver oder negativer Einfluss, dann müssen bei 2 potenziellen Parametern 4 Kombinationen, bei 10 Parametern 1024 Kombinationen, bei 20 Parametern aber schon über eine Million Kombinationen geprüft werden. Im – nie realisierbaren – Idealfall benötigt man für die Überprüfung der Wirkung einer Kombination mindestens einen Datensatz, also einen Patienten. Bei 20 Parametern wären also Beobachtungsdaten von mehr als einer Million Personen nötig, von denen jeweils der volle Datensatz erfasst sein muss.

Ein weiterer wichtiger Einsatzbereich von Big Data für die Präventivmedizin trat im Kontext der SARS-CoV2 Pandemie ins Rampenlicht: Offensichtlich verlaufen die meisten Erkrankungen sehr milde und benötigen keine besonderen therapeutischen Maßnahmen. In einigen Fällen jedoch verläuft die Erkrankung sehr schwer, mit sehr langem Bedarf an Intensivtherapie und – zum Zeitpunkt des Verfassens des Beitrags – noch nicht ausreichend erforschten Nebenwirkungen, und kann tödlich enden. Es ist offensichtlich, dass solche schweren Erkrankungsverläufe einen massiven Entzug von Freiheit der Betroffenen bedeuten. Andererseits bedeuten auch die epidemiologisch gebotenen Maßnahmen zur Verhinderung einer unkontrollierten Verbreitung der Epidemie einen – nach Meinung des Verfassers ungleich kleineren, zeitlich beschränkten – Eingriff in individuelle Freiheitsrechte, allerdings verbunden mit durchaus hohem wirtschaftlichem Schaden und für Viele. Der ethisch optimale Weg wäre daher ein gezielter Schutz der Personen, die ein Risiko für einen schweren Erkrankungsverlauf haben. Wie können diese Personen identifiziert werden? Wiederum ist eine wissenschaftlich solide experimentelle Studie ethisch indiskutabel. Es bleibt also nur die sorgfältige Analyse des Erkrankungsverlaufs der Patienten vor dem Hintergrund ihrer individuellen Konstitution. An dieser Stelle begegnet uns jedoch das oben schon diskutierte Dilemma: Aus der bisherigen, kurzen Erfahrung der Medizin kann eine Vielzahl von Faktoren einen schweren Verlauf bewirken. Die statistische Analyse der einzelnen, isolierten Faktoren ergibt jedoch Risikogruppen, die so groß sind, dass ein effektiver Schutz illusorisch erscheint. So gehören alle Personen älter als 60 Jahre, zusätzlich Menschen mit Übergewicht und chronischen Vorerkrankungen, zur Risikogruppe. Ist eine so grobe Charakterisierung noch hilfreich für eine Lehre zum Handeln? Dies ist eher fraglich.

Eine genauere Charakterisierung der wahren Risikogruppe erfordert jedoch offensichtlich Informationen über die in den jeweiligen Patienten vorliegenden Kombinationen von Risikofaktoren, die jede allein eventuell noch harmlos, im Zusammenwirken jedoch fatal sein können. Hier schlägt jedoch wieder das oben diskutierte kombinatorische Dilemma zu: ein Datensatz, der wirklich die dringend erforderlichen Informationen liefern könnte, müsste die Daten von hunderttausenden bis Millionen von Patienten umfas-

sen. Weltweit sind diese Daten sicher verfügbar. Die Durchführung eines solchen Programms, obwohl es ohne Zweifel dringend geboten ist, erfordert jedoch offensichtlich internationale Kooperation und gemeinsame Auswertung sehr großer Patientendatensätze – beides Herausforderungen, die heute diametral den individuellen Freiheitsrechten und der dadurch implizierten Frage nach der Data Ownership und Datensparsamkeit gegenüberstehen.

Wir stehen folglich vor einer Aporie: entweder einer Beschränkung der Freiheiten Aller oder der Akzeptanz einer schweren gesundheitlichen Schädigung oder gar des Todes vieler Mitmenschen oder aber eine Einschränkung des Freiheitsrechts auf Eigenbestimmung über die Daten. Es muss hier nochmals klar konstatiert werden, dass diese Aporie nicht etwa eine menschliche Erfindung ist, sondern aus der Natur des Virus herrührt.

Sowohl für die kurative als auch die präventive Medizin ist es folglich offensichtlich, dass die systematische Analyse und Verfügbarkeit sehr großer Patientendatensätze aus Beobachtungsdaten, die nicht im Rahmen gezielter, mit einem klar definierten Fokus erstellten Studien erhoben werden, essenziell für die Auflösung konzeptueller Dilemmata sind. Sie könnten die effiziente und zielgerichtete Entwicklung von gesundheitserhaltenden und gesundheits-restituierenden Maßnahmen in medizinischen Kontexten ermöglichen, die durch entweder komplexe Erkrankungsbilder, seltene Erkrankungen oder neu auftretenden Erkrankungen geprägt sind.

Im nächsten Abschnitt soll daher genauer betrachtet werden, was Big Data und Künstliche Intelligenz im oben diskutierten Kontext tatsächlich sind, was heute möglich ist und welcher zukünftige Bedarf sich hieraus ableiten lässt.

3. Big Data und Künstliche Intelligenz – just more of the same oder ein Quantensprung der Technologie?

3.1. Big Data – nur ein Haufen Daten?

Big Data wird gemeinhin als ein Sammelbegriff für eine große Menge an Daten und deren Auswertung gebraucht. Dabei muss vorausgeschickt werden, dass Daten allein keinen Wert an sich darstellen – Daten sind keine Information! Als Beispiel sei ein Telefonbuch angeführt: zumindest in den Zeiten vor dem Aufkommen des Internet enthielt es sehr viele Daten, zu jeder Person – im Klarnamen! – die Adresse sowie die Telefonnummer. Es gab nur in Ausnahmefällen die Möglichkeit, eine Streichung einzelner Informationen zu erreichen. Wenn der Name bekannt war, konnte daraus jeder ohne Schwierigkeiten die Adresse sowie die Telefonnummer identifizieren. Umgekehrt war dies jedoch nicht möglich: die Ordnungsstruktur war durch die alphabetische Ordnung des Namens festgelegt, für die Adresse sowie die Telefonnummer gab es keinerlei Ordnungsstruktur. Somit bedeutete eine Suche nach einem Namen zu einer gegebenen Telefonnummer das Durchsuchen des gesamten Telefonbuchs, einer Aufgabe, die von Hand extrem mühsam war. Ähnliche lexikalische Ordnungen waren in Lexika zu finden: es war leicht herauszufinden, dass ein Tiger ein Säugetier ist, jedoch extrem mühsam, alle Säugetiere aufzulisten. Diese Beispiele sollen zeigen, dass die Verfügbarkeit von Daten ohne Weiteres nur beschränkt die Extraktion von Information ermöglicht, die nicht a priori in der Ordnungsstruktur der Daten implementiert ist.

In der medizinischen Analogie würde ein „Patientenbuch“, in dem alle Patienten des Landes nach Anfangsbuchstaben des Nachnamens geordnet sind, und das alle diagnostischen Daten enthält, zwar alle Daten enthalten, die entsprechend der im letzten Kapitel diskutierten Anwendungsfelder von Big Data in der Medizin notwendig wären, ohne die Technologie zu einer automatisierten elektronischen Umsortierung im Sinne einer „Musterbildung“ wäre das Patientenbuch jedoch für die oben diskutierten Zwecke weitge-

hend wertlos: für jeden Fall müsste ein Arzt oder Wissenschaftler das Gesamt-Buch mit Millionen von Patienten durchlesen.

Viele Daten allein sind daher sicher noch nicht das, was im Sprachgebrauch unter Big Data subsummiert wird.

Hierzu passt ein Vergleich: Daten werden auch als das Rohöl des 21. Jahrhunderts bezeichnet. Die zunehmende Verfügbarkeit von Öl im 20. Jahrhundert hat nicht nur die Mobilität, sondern mit Hilfe der Chemie letztlich fast alle Bereiche unseres Lebens nachhaltig beeinflusst. Zweifellos hat diese Entwicklung vielfältige Optionen auf die Realisierung von Freiheiten eröffnet, die vorher nicht denkbar erschienen. Dies betrifft nicht nur die Mobilität, sondern auch die Gesundheit hat letztlich profitiert: die meisten Medikamente wurden letztlich aus Komponenten des Rohöls synthetisiert, bis in den letzten 30 Jahren zunehmend biotechnologische Produktionsverfahren entwickelt wurden und auch die für die Medizintechnik unverzichtbaren Kunststoffe dürfen nicht vergessen werden.

Können Daten auch die Rolle des Rohöls als Enabler von Freiheiten übernehmen? Sicher ist, dass das Potential der auf Rohöl basierenden Chemie zunehmend ausgereizt erscheint und komplexe Syntheseverfahren zunehmend biotechnologisch realisiert werden, während das Potential der Nutzung von Daten noch nicht vergleichbar der Nutzung des Rohöls Anfang des 20. Jahrhunderts entwickelt ist. Der Vergleich trifft jedoch noch aus einer ganz anderen Blickrichtung: Rohöl allein ist eine klebrige, giftige Flüssigkeit, für die es zu Beginn des 20. Jahrhunderts außer dem Betrieb von Petroleumlampen keine Verwendung gab. In etwa auf diesem Niveau können wir die Nutzung der Daten heute sehen. Erst die Entwicklung einer zunehmend ausgefeilten Technologie zur Verteilung, Aufreinigung und weiteren Modifikation hin zu immer komplexeren Molekülen mit zunehmend komplexeren Funktionalitäten, getrieben durch kontinuierliche Fortschritte der chemischen Forschung und Technologie, ermöglichte die Entwicklung des Rohöls zum Fortschrittsmotor des 20. Jahrhunderts. Genauso ist die Rolle der Daten zu sehen: für sich allein genommen sind sie zwar durchaus hilfreich – etwa als Telefonbuch oder Lexikon –, ihr volles Potential können sie jedoch erst im Kontext geeigneter Technologie entfalten: Die Daten müssen zugänglich, recherchierbar in Bezug auf alle Komponenten, nicht nur in Bezug auf eine

fest vorgegebene Ordnung, interpretierbar und reproduzierbar sein.⁹ Hierzu wird Technologie benötigt: Speichertechnologie, die eine effiziente, schnell abrufbare Speicherung sehr großer Datenmengen erlaubt und Verifizierungstechnologie, die automatisch Fehler in sehr großen Datenmengen erkennt, die von Menschen nicht mehr durchsucht werden können. Nicht zuletzt wird Technologie benötigt, die aus den Myriaden von Daten Information in Form von Regeln, Mustern oder Modellen extrahiert, die im Kontext ihrer Anwendungen interpretierbar sind und im Sinne einer Lehre zum Handeln in – medizinischen – Anwendungsfeldern neue Optionen eröffnen.

Im Kontext der medizinischen Handlungsfelder kann man die unter „Big Data“ subsumierten Daten in vier Gruppen klassifizieren, mit jeweils speziellen Datenstrukturen, Anwendungsschwerpunkten, und Herausforderungen an die benötigte Technologie:

- Daten, bei denen von jeder Probe – in der Regel von jedem Patienten – sehr viele (> 10 000) Parameter desselben Typs erfasst werden, werden als „Fat Data“ bezeichnet. Sie werden typischerweise mit Hilfe der sogenannten „-omics“-Technologien generiert: sie bestimmen zum Beispiel von jeder Zellprobe alle Genmutationen – auf mehr als 20 000 Genen –, oder die Expression – in etwa die Aktivität – aller Gene, oder aller Proteine (mehr als 100 000), oder aller Metabolite und vieles andere mehr. In der Regel werden daher für jede Probe alle Parameter erfasst, die Datenstruktur enthält daher wenige Lücken. Für jeden Datentyp (Gene, Proteine etc.) wird dabei eine spezifische Technologie benötigt, so dass in der Regel – auch aus Kostengründen – von einer Probe jeweils nur ein Datentyp in seiner ganzen Breite erfasst wird. Diese Technologien sind heute schon im klinischen Betrieb für spezielle Anwendungen etabliert, besonders in der Onkologie. Trotz der durch die rasante Entwicklung der Messtechnologien dramatisch gesunkenen Kosten sind diese jedoch immer noch so hoch, dass die meisten Daten in gezielten Studien anfallen, die deutlich weniger Probanden umfassen als jeweils Parameter gemessen werden. Eine entsprechende Tabelle sieht daher sehr breit, aber vergleichsweise gedrungen aus, woher der Name „Fat Data“ herrührt. In-

⁹ <https://www.go-fair.org/fair-principles>, Zugriff 27. März 2021.

zwischen sind zwar sehr viele Daten (> 1 Million Genexpressionsdatensätze) insgesamt frei verfügbar, diese stammen jedoch aus sehr heterogenen Studien, die jeweils spezifische, systematische Abweichungen enthalten, und müssen daher mühsam zusammengeführt werden.

- Im Gegensatz dazu werden „Real World Evidence“-Beobachtungsdaten als „Lean Data“ bezeichnet: zwar werden oft einige hundert unterschiedliche Parameter erfasst, die sehr unterschiedliche medizinische Prozesse abbilden. Die Zahl der Patienten ist jedoch deutlich größer. Auf spezielle Krankheiten hin gesammelte Datenkohorten schließen einige tausend Patienten ein, intensivmedizinische Datensätze einige zehntausend Patienten, Datensätze von Krankenversicherern oder staatlichen Institutionen wie dem britischen National Health Service umfassen oft die Daten von mehreren Millionen Patienten, so dass die Daten in Tabellenform eine eher schlanke Struktur hätten. Im Gegensatz zu typischen -omics-Datensätzen, in denen alle entsprechenden Parameter des erfassten Typs gespeichert werden, werden in „Real World Evidence“-Daten nur die Parameter erfasst, die der behandelnde Arzt beim jeweiligen Patienten erhebt, alle anderen Daten bilden die sogenannten „missing Values“. In sehr großen Datensätzen, die unterschiedlichste Patienten erfassen, z. B. Daten von Krankenversicherern, sind daher zwar sehr viele Patienten eingeschlossen, allerdings bestehen für jeden einzelnen Patienten die Datensätze zu einem Großteil aus Lücken. Daher erfordert die Analyse solcher Datensätze spezielle Verfahren der Künstlichen Intelligenz.
- Zunehmende Bedeutung erhalten Daten, die mit einer engen Abtastrate kontinuierlich in sogenannten Monitoring-Systemen als sogenannte Zeitreihen erfasst werden. Als Beispiel seien EKG-Daten, Pulsrate, oder Blutzucker genannt, die heute nicht nur in einer kurzen Untersuchung beim Arzt, sondern über leichte, ständig am Körper tragbare Geräte, sogenannten Wearables, über Tage und Wochen hinweg kontinuierlich einen Datenstrom erzeugen. Hierbei werden zwar jeweils vergleichsweise wenige Parameter erhoben, dafür jedoch mit einer zeitlich beinahe unbegrenzten Dauer und Abtastrate, so dass wiederum „Big Data“ zur Analyse anstehen. Hierzu gehören auch Spracherkennungssysteme, die aus

Sprachmelodien oder Sprechweisen medizinisch relevante Informationen extrahieren können. Insbesondere für Patienten mit einem erhöhten Risiko für das Auftreten von plötzlichen kritischen Syndromen, wie Herzrhythmusstörungen oder für Diabetespatienten können solche kontinuierlichen Monitoringsysteme einen echten Vorteil gegenüber der etablierten Praxis bieten: seltene Probleme, die bei auch regelmäßigen Besuchen einer Arztpraxis nicht auftreten müssen, werden sofort erkannt, so dass direkt eine zielgerichtete ärztliche Intervention ausgelöst wird, die kritische Gesundheitszustände vermeiden kann. Daher soll dieser Datentyp hier „Monitoring-Daten“ genannt werden.

- In der medizinischen Spezialdisziplin Radiologie werden Bildaufnahmen des Inneren von Patienten, die mit Röntgen-, Ultraschall-, MRT- oder anderen bildgebenden Verfahren gewonnen werden, zur Diagnose von Erkrankungen interpretiert. Gerade in der Interpretation von Bildern geschah in den letzten 10 Jahren einer der größten Technologiesprünge der Künstlichen Intelligenz, ausgelöst durch die Entwicklung effizienter Verfahren für das sogenannte „Deep Learning“. Daher wird die Radiologie generell als ein kanonisches Anwendungsfeld der Künstlichen Intelligenz und Big Data in der Medizin gesehen, in der die Entwicklung schon weit fortgeschritten ist, bis hin zur Zulassung entsprechender Verfahren durch die zuständigen Regulierungsbehörden. Der Einsatz von KI-basierten Bildanalysen wird jedoch zunehmend auch für andere Felder der Medizin, z. B. die Augendiagnostik und Dermatologie, entwickelt. Jedes Bild wird hierfür in seine Bestandteile, die Pixel, zerlegt, deren Zusammenwirken zu einem Bild von entsprechenden Algorithmen analysiert wird. Da ein Bild aus Millionen von Pixeln bestehen kann, müssen Bilddaten als besonderer Typ von Big Data in der Medizin betrachtet werden, auch wenn die Zahl der Patienten, von denen Bilddaten zur Verfügung stehen eher klein ist. Insofern ist die Problematik der Verfügbarkeit und Analyse von medizinischen Bilddaten ähnlich zur Problematik der „Fat Data“, insbesondere der Genomsequenzierungsdaten von Patienten. Allerdings sind die optimalen Analyseverfahren für Bilddaten und -Omicdaten deutlich unterschiedlich, so dass mit einem

gewissen Recht Bilddaten als eine eigene Klasse von medizinisch relevanten „Big Data“ gesehen werden können.

3.2 Künstliche Intelligenz – die Technologie zur Transformation von Big Data zur Information

Künstliche Intelligenz ist heute zweifellos ein Hype mit positiver oder negativer Annotation, je nach Sichtweise. Für die Diskussion der Relevanz von Big Data in der Medizin für die Freiheit ist es daher essenziell, die hierfür relevanten KI-Themenfelder klar zu definieren und den Kern ihrer Bedeutung für die Entwicklung der Medizin darzustellen.

In diesem Kontext verstehen wir unter KI-Algorithmen auf Computern lauffähige Programme, die

- auf Basis von individuellen Patientendaten medizinisch relevante Aussagen treffen können, die von Menschen nur schwer in derselben Qualität und Geschwindigkeit getroffen werden können
- die ihre Fähigkeit, medizinisch relevante Aussagen zu treffen, auf Basis von Daten lernen. Sie benötigen für dieses Training keine expliziten Informationen oder Regeln. Daher kann erwartet werden, dass entsprechende KI-Algorithmen
 - „objektiv“ trainiert werden: sie lernen nur von Daten und sind nicht von Meinungen beeinflusst.
 - Aussagen zu neuen Zusammenhängen treffen, die von Menschen übersehen worden sind.

Wieso wird erwartet, dass KI- Algorithmen Menschen zumindest in einigen medizinisch relevanten Aufgaben überlegen sein sollen? Ein wesentlicher Unterschied zwischen der Informationsverarbeitung von Menschen und Computern liegt in der Architektur des Gehirns und der Architektur des Computers. Das menschliche Gehirn kann abstrakt als ein riesiges Netzwerk von miteinander verknüpften Nervenzellen, den Neuronen, betrachtet werden. Jedes Neuron kann dabei aktiv oder inaktiv sein. Je nach Aktivitätsgrad stimuliert es dabei die Neuronen, mit denen es verknüpft ist, ebenfalls zu Aktivität, die jedoch abklingt, wenn keine Stimulation vorliegt. Dabei sind ei-

nige Teile des Netzwerks mit der Außenwelt verbunden und können dadurch von außen stimuliert werden, z. B. durch den Sehnerv oder den Gehörnerv. Durch das Zusammenwirken all dieser Knoten im Netzwerk des Gehirns entstehen unsere gesamte Gehirnaktivität, unsere mentalen Fähigkeiten sowie, medizinisch relevant, die unbewusste Steuerung der Organfunktionen. Dieses Zusammenwirken hängt dabei wesentlich von der Verknüpfungsstruktur der Knoten ab, die im Laufe des Lernprozesses eines Menschen erst gebildet – bzw. umgebildet – wird. Sie ist nicht per se vorgegeben. Die Kommunikation der Nervenzellen untereinander erfolgt dabei durch elektrische Stimulation, die allerdings im Vergleich zur elektrischen Stimulation elektronischer Komponenten sehr langsam ist.

Im Gegensatz dazu besteht ein Computer aus elektronischen Komponenten, die – wie im Gehirn die Neuronen – nur einfache Grundoperationen durchführen können, allerdings sehr viel schneller. Ein weiterer wesentlicher Unterschied ist, dass die Verschaltung dieser elektronischen Komponenten eine per Konstruktion aufgeprägte hierarchische Ordnung aufweist, so dass durch Kopplung der elektronischen Komponenten komplexere funktionale Module, wie Rechenwerke, Grafikchips etc. entstehen, die wiederum zu einem Computer zusammengeschaltet werden.

In der Konsequenz haben dadurch beide, das Gehirn und der Computer, jeweils spezifische Stärken-Schwächen-Profile: während Computer bei mathematischen Rechenoperationen in der Geschwindigkeit dem Gehirn meilenweit überlegen sind, ist das Gehirn dem Computer bei komplexen Aufgaben, die sich nicht einfach auf mathematische Grundoperationen reduzieren lassen, überlegen, zum Beispiel der Mustererkennung in Bildern und der Sprache.

Über die Ursache dieses Unterschieds kann man spekulieren. Es ist jedoch naheliegend, dass die Fähigkeit, schnell und zuverlässig komplexe Muster in Bildern und Geräuschen zu interpretieren einen Evolutionsvorteil darstellt: je früher die Bild- und Tonsignale eines sich anschleichenden Raubtiers – oder einer Beute – erkannt werden konnten, desto eher konnte sich der in seinen sonstigen körperlichen Fähigkeiten unterlegene Urmensch darauf einstellen und geeignet reagieren. Im Gegensatz dazu war die Fähigkeit gut und schnell rechnen zu können offensichtlich kein Evolutionsvorteil: erst mit

Aufkommen der Zivilisation wurden Rechenfähigkeiten überhaupt relevant, dann jedoch konnten wenige ausgebildete Spezialisten die notwendigen Berechnungen vornehmen. Für einen klassischen Evolutionsvorteil waren mathematische Fähigkeiten eher nicht notwendig.

Durch die Optimierung der kognitiven Fähigkeiten auf Mustererkennung in Bildern und Geräuschen ist unser Gehirn sehr gut an unsere dreidimensionale Welt + Zeit angepasst. Das Erkennen des Zusammenspiels von bis zu drei oder vier unterschiedlichen Faktoren ist für ein menschliches Gehirn intuitiv auch ohne rationale Analyse möglich. Darüber hinaus nehmen die intuitiven Fähigkeiten jedoch rasch ab. Das Zusammenspiel von mehr als etwa 5 Faktoren intuitiv zu erfassen dürfte selbst trainierten Menschen schwerfallen. Für die Medizin ergibt sich hieraus die Konsequenz, dass Erkrankungen, für deren Diagnose das Zusammenwirken von vielen diagnostischen Parametern notwendig ist, nicht mehr intuitiv richtig diagnostiziert werden können. Hierzu werden Regelsysteme benötigt, die dieses Zusammenwirken in Form expliziter logischer Regeln beschreiben. Diese Regelsysteme sind seit alters her eine Grundlage der Medizin. Sie wurden durch Beobachtung an Patienten und darauf basierendes Erfassen der gefundenen Muster in logischen Systemen im Konsens entwickelt. Dieser deduktive, auf menschlicher Erfahrung basierende, Prozess wird jedoch wiederum durch die begrenzte Fähigkeit des Gehirns zum Erfassen von komplexen Interaktionen vieler Faktoren limitiert. Ein weiteres Problem entsteht durch die limitierte Kapazität des Gehirns, viele kleine, wenig komplexe, Regelsysteme zu einem großen, komplexen Regelsystem zu integrieren.

Es stellte sich daher schon sehr früh in der Entwicklung der Künstlichen Intelligenz die Frage, ob Algorithmen mit dem Schwerpunkt der quantitativen Verknüpfung vieler Faktoren die menschlichen Grenzen überwinden könnten. Ein erster Schritt hierzu war die Entwicklung von diagnostischen Expertensystemen seit den 1970er Jahren.¹⁰ Hierbei wurde versucht, das Wissen über diagnostische Regeln von medizinischen Experten zu erfragen und dann in ein komplexes formales und widerspruchsfreies Regelsystem zu integrieren, das dann mit Hilfe von Computern überall, sozusagen auf Knopf-

¹⁰ Shortliffe, E. H., Computer-Based Medical Consultations: MYCIN, New York 1976.

druck, auch von Nicht-Experten ihres jeweiligen Fachs eingesetzt werden kann. Es zeigte sich jedoch sehr bald, dass die Entwicklung von Expertensystemen deutlich hinter den Erwartungen zurückblieb. Ein Grund war, dass die Extraktion von formal widerspruchsfreien Regeln von Experten deutlich schwieriger war als erwartet. Häufig waren die Regeln nicht eindeutig in Form von Wenn-Dann-Strukturen formalisierbar, und das Zusammenführen mehrerer Regeln konnte in logischen Inkonsistenzen resultieren. Dies erforderte eine Abschwächung der Logik in Richtung einer Fuzzy-Logik, die „Graubereiche“ behandeln konnte. Als ein gravierendes praktisches Problem erwies sich darüber hinaus, dass die Extraktion der Regeln einen unerwartet hohen Aufwand für die Experten darstellte.

Nicht zuletzt hierdurch wurde das Konzept selbstlernender Systeme, das Maschinelle Lernen (ML) befördert. Selbstlernende Systeme sind Algorithmen, die ihr Verhalten nur mit Hilfe eines Trainings aus Daten selbst lernen. Sie benötigen folglich kein Expertenwissen, das wie bei den Expertensystemen explizit in ein Computerprogramm integriert wird.

Maschinelles Lernen subsummiert generische Algorithmen, die konzeptuell jeden funktionalen Zusammenhang zwischen vielen verschiedenen Größen darstellen können, ohne irgendeine Information über den realen Zusammenhang zu kennen. Grundsätzlich sind in der Mathematik solche Algorithmen nicht neu. Seit dem 19. Jahrhundert wurde eine Vielzahl von ihnen für unterschiedlichste Anwendungen in den Naturwissenschaften und im Ingenieurwesen entwickelt. Alle diese „klassischen“ Algorithmen werden jedoch mit wachsender Zahl der darzustellenden Größen schnell ineffizient, so dass komplexere Zusammenhänge nicht zuverlässig dargestellt werden können. Diese Lücke füllt das Maschinelle Lernen.

Die Architektur der prominenteste Klasse der ML-Algorithmen, die Neuronale Netze, insbesondere die tiefen neuronalen Netze (DNN), ist dabei dem Aufbau des Gehirns nachempfunden. DNN bestehen aus einer großen Zahl von sogenannten Knoten, deren Verhalten den Neuronen des Gehirns nachempfunden ist. Die Knoten sind in einem Netzwerk miteinander verbunden und tauschen über diese Verbindungen Informationen über ihren jeweiligen Aktivitätszustand aus. Die Aktivität eines Knotens wird dabei über die Summe der Aktivitäten der mit ihm verbundenen Knoten abgebildet, so dass

sich durch den Austausch im Netzwerk der Knoten komplexe Aktivitätsmuster einstellen können. Da spezielle Knoten des Netzwerks die darzustellenden Größen als Eingänge und Ausgänge haben, kann bei geeignetem Training des Netzwerks der Zusammenhang zwischen den darzustellenden Größen dargestellt werden. Mit Hilfe des trainierten Netzwerks können dann die Zusammenhänge zwischen neuen, bisher unbekanntem Daten rekonstruiert werden.

In der Praxis ist die Architektur der DNN in der Regel in mehreren Schichten strukturiert, wobei nur Knoten in jeweils benachbarten Schichten miteinander verbunden sind. Je nach Einsatzfeld können die Schichten noch spezielle Strukturen aufweisen. Es kann mathematisch gezeigt werden, dass die Architektur der DNN für eine große Klasse von funktionalen Zusammenhängen, wie sie prominent zum Beispiel in der Bild- und Sprachverarbeitung, der Signalanalyse und der Textverarbeitung auftreten, gegenüber allen anderen Algorithmen des maschinellen Lernens, und den „klassischen“ Algorithmen sowieso, deutlich überlegen ist. Daher konnten in den letzten Jahren echte Quantensprünge in Anwendungen erreicht werden, bei denen Bild- oder Signalverarbeitung die Schlüsselstellen zum Erfolg sind.

In der Medizin reichen potenzielle Anwendungen mit einer konzeptuellen Überlegenheit des Maschinellen Lernens von der Grundlagenforschung über den klinischen Einsatz bis hin zu Home-Monitoring Systemen. In der Grundlagenforschung kann zum Beispiel mit Hilfe von DNN die Faltung von Proteinen mit deutlich besserer Genauigkeit und Geschwindigkeit berechnet werden als mit allen gängigen Methoden. Da die Struktur von Proteinen der Schlüssel für die gezielte Suche nach neuen Wirkstoffen ist, hat diese Entwicklung das Potential zu einer deutlich effizienteren Medikamentenforschung.

Auch in der klinischen Forschung wird ein großes Potential gesehen. Hier gilt es, besonders bei komplexen Erkrankungen, durch Integration sehr vieler verschiedener diagnostischer Daten eine gesamtheitliche Prognose des Erkrankungs- und Therapieverlaufs zu erreichen. Einsatzgebiete reichen dabei von der Onkologie bis zur Intensivmedizin.¹¹

¹¹ https://www.researchgate.net/publication/347804780_Physicians'_requirements_and_expectations_of_future_Artificial_Intelligence_applications_for

Ein weiteres Einsatzgebiet der ML-Technologie von heute ist die Radiologie. Hier gilt es, sehr große Mengen an Röntgenbildern nach krankheitsspezifischen Mustern zu durchsuchen. Studien haben gezeigt, dass eine gut trainierte künstliche Intelligenz den Experten-Radiologen nicht ersetzen, wohl aber durch das automatisierte Aussortieren kritischer Fälle maßgeblich unterstützen kann. Ähnliche Unterstützungsfunktionen werden aus vielen weiteren Anwendungsfeldern, von der Dermatologie bis zur Pathologie, beschrieben. Ein weiterer Vorteil einer künstlichen Intelligenz ist, dass durch die Unterstützungsfunktion auch weniger erfahrene Ärzte vom Erfahrungsschatz der Maschine profitieren können und damit die Qualität der medizinischen Versorgung in der Breite verbessert werden kann.

Im Bereich der Präventivmedizin eröffnen sich dadurch ebenfalls neue Möglichkeiten, da eine automatisierte kontinuierliche Überwachung der Vitalfunktionen von potenziell gefährdeten Patienten die Einweisung in eine stationäre Behandlung eventuell ohne Risiko zu vermeiden vermag.

Es darf dabei aber nicht verschwiegen werden, dass vor der Realisierung des Potentials von Big Data in der Medizin mit Hilfe des Maschinellen Lernens bzw. Künstlicher Intelligenz noch erhebliche Hürden zu nehmen sind.

4. Herausforderungen im Einsatz von Big Data in der Medizin

4.1. Technologische Herausforderungen

Obwohl in der breiten Diskussion der Eindruck entstehen kann, dass technologieeitig mit DNN alle notwendigen Technologien zur Verfügung stehen, muss dies klar verneint werden. Es bleiben für einen breiten Einsatz auch technologieeitig erhebliche Hürden zu nehmen:

- Sind DNN die Lösung für alle Fragen? Dies muss klar verneint werden. DNN haben überragende Stärken in – mathematisch – recht genau definierbaren Anwendungsfeldern, die oben diskutiert wurden, sie zeigen jedoch eine eher unterdurchschnittliche Performance in vielen anderen Feldern. Insbesondere sind DNN unterdurchschnittlich performant, wenn sehr unterschiedliche Datentypen in ein gemeinsames Modell integriert werden müssen, insbesondere wenn die Datentypen nicht die Struktur von Bild- oder Zeitreihendaten haben. Außerdem ist das Trainingsverhalten auf stark unvollständigen Daten, wie sie in RWE-Datensätzen vorliegen, ebenfalls noch suboptimal. Leider sind alternative Konzepte des maschinellen Lernens ebenfalls noch nicht in der Lage, die notwendige Genauigkeit zu gewährleisten.
- Der Energiebedarf zum Training sehr großer DNN, zum Beispiel für die Bildanalyse, stellt ein wenig diskutiertes Problem dar. Spracherkennungssysteme benötigen für das Training einer einzigen Anwendung ca. 1 000 kWh, der Energiebedarf für das Training sehr großer Bilderkennungssysteme kann bis zum Energiebedarf von 5 Autos über den kompletten Lebenszyklus reichen.¹² Auch wenn dieser Energiebedarf nur einmalig pro Anwendung anfällt – der spätere Einsatz hat einen eher vernachlässigbaren Energiebedarf – zeigt sich trotzdem sehr deutlich, dass nur große Firmen überhaupt den notwendigen Aufwand leisten können. Dieser Aufwand lohnt sich nur für Anwendungen mit einem

¹² <https://www.spektrum.de/news/kuenstliche-intelligenz-verbraucht-fuer-den-lernprozess-unvorstellbar-viel-energie/1660246>, Zugriff 27. März 2021.

sehr breiten Nutzungspotential, viele spezialisierte Anwendungen, die gerade in der Medizin eine große Rolle spielen könnten, kommen daher nicht in Betracht. Hier werden dringend effizientere Hardwaretechnologien benötigt, wie neuromorphe Computer oder Quantencomputer, die jedoch selbst noch in der Entwicklungsphase sind.

- Selbst in dem Fall, dass das Training eines KI-Systems erfolgreich durchgeführt wurde, stellt sich die Frage nach dessen Zuverlässigkeit. Zweifellos können die KI-System zuverlässige Prognosen innerhalb des Datenbereichs durchführen, mit dem sie trainiert wurden. Wenn sie sich aus diesem Datenbereich entfernen, verlieren sie jedoch konzeptuell sehr rasch ihre Zuverlässigkeit. Für einen Einsatz in der Medizin müssen solche Situationen jedoch mit sehr hoher Zuverlässigkeit erkannt werden, wofür ebenfalls noch Entwicklungsarbeit geleistet werden muss.
- Generell stellt sich die Frage nach der Erklärbarkeit einer auf Basis eines KI-Systems berechneten Prognose. Sowohl für die Zulassung solcher Systeme als auch für das Vertrauen von Ärzten und Patienten ist es essenziell, dass die Entscheidungsfindung nachvollziehbar ist. Gerade hier haben die heute etablierten KI-Systeme eine erhebliche Schwäche, die gerade durch ihre komplexe, generische Architektur hervorgerufen wird. Explizit interpretierbare Entscheidungsregeln lassen sich heute noch sehr schwer aus einer trainierten KI extrahieren. Dies ist ein aktives Feld der Forschung.

4.2 Herausforderungen an Datenverfügbarkeit und Datennutzung

Für das Training einer künstlichen Intelligenz sind in jeder Anwendung sehr große Datenmengen erforderlich, die mit der Komplexität des Anwendungsproblems stark zunehmen können („Fluch der Dimensionalität“). Für Anwendungen in der Medizin bedeutet dies, dass Daten von sehr vielen Menschen benötigt werden, in der Regel mindestens mehrere 1000 bis hin zu mehr als 100 000 Patienten. Es ist offensichtlich, dass diese technisch nötigen Datenmengen nicht für jeden Anwendungsfall in einer eigens entwor-

fenen Studie erhoben werden können, hierfür wäre der Aufwand viel zu hoch. Es muss daher auf schon existierende Daten zurückgegriffen werden, die, zumindest verstreut, auch vorliegen. Hierbei ergeben sich sowohl technische als auch rechtliche und ethische Herausforderungen:

- Diese Daten wurden nicht für den Zweck einer nachträglichen Auswertung erhoben, sondern entweder im Rahmen von Studien, zur Dokumentation oder zur Abrechnung von ärztlichen Leistungen. Jede einzelne Datenbank umfasst daher weder Patienten in der nötigen Zahl noch alle nötigen Parameter pro Patient. Technisch gesehen müssen daher Daten aus unterschiedlichen Quellen zusammengeführt werden, was erhebliche datentechnische Probleme aufwirft, da es momentan keine einheitlichen Standards gibt.
- Die Datenqualität muss sichergestellt werden. Hierzu werden heute noch alle Daten von Ärzten überprüft und nach Befunden klassifiziert, bevor mit Ihnen ein KI-System trainiert wird. Um Fehler zu vermeiden, werden die Daten zusätzlich meist von mehreren Ärzten unabhängig begutachtet. Dieser Schritt stellt einen erheblichen Aufwand dar, der nur für sehr wenige Anwendungen, und auch dort nur von großen Firmen, realisierbar ist. Allein dieser Aufwand könnte dazu führen, dass die Entwicklung medizinisch relevanter KI-Systeme in der Hand weniger großer, international operierender Konzerne liegen wird, die dadurch implizit Einfluss auf die Gesundheitsversorgung nehmen können. Alternative Technologien, die zumindest ohne den aufwändigsten Schritt, nämlich die manuelle Befundung der Daten auskommen, sogenannte unüberwachte Verfahren, erreichen heute noch nicht die notwendige Performance und benötigen nochmals deutlich mehr Daten für ihr Training.
- In Bezug auf Datenverfügbarkeit sind jedoch nicht nur technische Probleme zu lösen. Wesentlich größere Hürden stellen die Herausforderungen an Datenschutz und Datenzugang dar: Während in einer Medikamentenstudie alle teilnehmenden Patienten aufgeklärt werden und in die Teilnahme an der konkreten Studie einwilligen, ist es bei den erforderlichen retrospektiven Datenanalysen zum Training einer künstlichen Intelligenz nicht praktikabel, nachträglich alle Patienten aufzuklären und eine Einwilligung einzuholen. Aktuelle Initiativen zu einem soge-

nannten „Broad Consent“, bei dem Patienten ihre Daten für die Forschung freigeben können, werden erst langfristig zum Tragen kommen. Es ist daher zu erwarten, dass die Entwicklung von KI-Systemen in der Medizin primär in den Ländern stattfinden muss, in denen die Datenverfügbarkeit gegeben ist. Länder, in denen dies nicht gegeben ist, werden die trainierten Systeme einkaufen müssen. Wie oben diskutiert, ist dies jedoch mit einem erhöhten Risiko für die Patienten verbunden, da die jeweiligen Gesundheitsversorgungssysteme und Diagnose- und Therapiestrategien im internationalen Kontext sich deutlich unterscheiden. Die Übertragbarkeit von KI-basierten Prognosesystemen zwischen Ländern ist keinesfalls trivial und bedarf noch ausgiebiger Forschung. Eine schnelle Umsetzung ist daher nicht zu erwarten.

- Es muss dabei ausdrücklich betont werden, dass es beim Training eines medizinischen KI-Systems nicht um eine Re-Identifizierung eines Patienten geht, sondern im Gegenteil sind gemeinsame Muster zu finden, die für viele Patienten gelten. Hier treffen international unterschiedliche Rechtsauffassungen aufeinander.
- Des Weiteren sind KI-Systeme für den klinischen Einsatz als medizintechnische Produkte zu bewerten, besonders wenn die von ihnen aus Big Data extrahierten Prognosen die ärztliche Therapieentscheidung beeinflussen können. Dies gilt besonders, wenn sie, etwa als Teil eines medizintechnischen Gerätes, direkt in den Therapieprozess eingreifen. In der Konsequenz müssen solche KI-Systeme die entsprechenden regulatorischen Zulassungsvoraussetzungen erfüllen, insbesondere muss die Wirkung und Sicherheit in klinischen Studien geprüft werden, ähnlich zu neuen Medikamenten. Diese Prüfung erfordert einen erheblichen Aufwand, der in aller Regel nur von größeren Unternehmen geleistet werden kann. Da viele Entwicklungen der KI in der Medizin jedoch entweder im akademischen Kontext, in kleineren Start-Up Firmen oder in IT-Unternehmen stattfinden, die wenig Erfahrung mit den regulatorischen Prozessen haben, gibt es auch hier eine organisatorische Hürde für die Nutzung des Potentials in der Praxis. Daher ist bis heute ein starker Trend zu KI-Lösungen für den Forschungsbereich und für den Lifestyle

Bereich zu beobachten, der naturgemäß weniger regulatorische Anforderungen stellt.

- Last but not least ist das Know How des Arztes in der Anwendung der KI-Tools bei der Analyse von Big Data ein nicht zu vernachlässigender Erfolgsfaktor. Alle Erfahrungen in der Anwendung von Big Data und KI in der Vergangenheit haben stets gezeigt, dass die Erfahrung des Anwenders zu den Top 3 der wichtigsten Erfolgskriterien gehört. Es ist sehr fraglich, ob sich diese Situation in den nächsten Jahren durch die weitere Entwicklung der Technologie merklich ändern wird. Hier stoßen wir auf ein weiteres Problemfeld, das einer Lösung bedarf: in der Ausbildung der Ärzte heute spielen moderne Datenanalysetechnologien sowie deren Grundlagen keine Rolle, geschweige denn in der Vergangenheit. Die Technologien werden daher von de facto Laien angewendet. Dies muss kein unüberwindlicher Hinderungsgrund sein, schließlich fahren die meisten Menschen Auto, ohne zu verstehen, wie ein Auto genau funktioniert. Allerdings gibt es einen Unterschied: beim Autofahren gefährden wir nur Wenige, und wir reizen das Potential der Autotechnologie nicht aus (hierfür ist erhebliches Training erforderlich). Ein Arzt gefährdet jedoch stets Menschen, wenn er seine Werkzeuge nicht optimal einsetzt. Auch erscheint ein optimaler Einsatz von Technologien in der Medizin relevanter als im Alltagsgebrauch. Es ist daher zu erwägen, ob für den Einsatz von Big Data in der Medizin zusätzliche Ausbildungen für Mediziner verbindlich eingeführt werden sollten. Erste Ansätze hierzu sind durchaus schon vorhanden, z. B. als berufsbegleitende Aufbaustudiengänge.

5. Big Data in der Medizin – ein notwendiger gesellschaftlicher Diskurs

Für eine endgültige Bewertung von Big Data in der Medizin ist es heute sicher noch viel zu früh. Obwohl erste Anwendungen zugelassen sind, kann von einer schon existierenden Welle von KI-Tools in der Medizin sicher nicht gesprochen werden, so dass die Bewertung auf Basis der sich noch im For-

schungs- und Entwicklungsstadiums befindlichen Technologien erfolgen muss.

Wie oben diskutiert wurde, basiert der momentane Hype um KI in der Medizin nicht nur auf Technologiebegeisterung, sondern durchaus auch auf einem zunehmenden Druck auf die Gesundheitsversorgung in einer alternden Gesellschaft, die von zunehmend komplexen Erkrankungsbildern geprägt ist und offensichtlich auch nicht vor plötzlich auftretenden katastrophalen Ereignissen wie der SARS-CoV2 Pandemie gefeit ist. Von diesen Herausforderungen an die Gesundheitsversorgung geht ein erheblicher Druck auf die sozialen Systeme aus, der durchaus Konsequenzen für die individuelle Freiheit haben kann, wie wir gegenwärtig in der Pandemie erleben.

Es gibt gute Gründe für die Annahme, dass Big Data und Künstliche Intelligenz Technologien sind, die zumindest einen erheblichen Beitrag zur Bewältigung dieser Herausforderungen leisten können und damit eine durch beschränkte medizinische Ressourcen bedingte Einschränkung an persönlicher Freiheit verhindern können.

Umgekehrt darf nicht übersehen werden, dass Big Data in der Medizin auch inhärente gesellschaftliche Probleme aufwerfen kann, die einer Diskussion und einer politischen Lösung bedürfen:

- Auch wenn für die Entwicklung von medizinischen KI-Systemen eine Identifizierung der hinter den jeweiligen Daten stehenden Person weder nötig noch irgendwie angestrebt ist, so dass theoretisch anonymisierte /pseudonymisierte Daten ausreichen, stellt sich die Frage eines potenziellen Risikos einer Deanonymisierung/Depseudonymisierung. Dieses Risiko ist nach heutigem Stand der Technik nicht völlig auszuschließen, so dass die Frage gestellt werden muss, ob das Risiko einer mit krimineller Absicht möglichen Deanonymisierung/Depseudonymisierung die Entwicklung von Big Data Technologien per se inakzeptabel macht oder ob nicht nur die Handlung selbst unter Strafandrohung gestellt werden sollte. Eine zukünftige ideale Lösung ist sicherlich in einem Broad Consent zu sehen, in dem ein Patient freiwillig seine pseudonymisierten/anonymisierten Daten für die Forschung ohne Bezug auf eine konkrete Studie zur Verfügung stellen kann. Da diese Lösung jedoch erst zukünftig zu erhebende Daten betreffen kann, würde die Einschränkung

auf eine Broad-Consent-Lösung für die Datennutzung für die Forschung ein effektives Moratorium für einen nicht zu unterschätzenden Zeitraum bedeuten, insbesondere für longitudinale Datenanalysen.

- Ein weiteres Problem besteht durch die im internationalen Kontext extrem unterschiedlichen rechtlichen und regulatorischen sowie wirtschaftlichen Rahmenbedingungen für Big Data und KI in der Medizin. Eine auf den ersten Blick gesellschaftlich eventuell akzeptable Lösung zum ersten Punkt, nämlich eine Verlagerung der Entwicklungen in Länder wie USA oder China, die jeweils sehr unterschiedliche, jedoch grundsätzlich weniger restriktive Rahmenbedingungen bieten, beinhaltet bei genauerem Hinschauen erhebliche Risiken für die Freiheit sowie die Sicherheit: ein für die Gegebenheiten dort konzipiertes und dort sicher zuverlässiges System lässt sich, wie oben diskutiert wurde, nicht ohne Weiteres in einen völlig anderen medizinischen Kontext implementieren. In der Konsequenz müsste das medizinische Versorgungssystem an die Gegebenheiten des entwickelnden Landes angepasst werden. Alternativ müssten erhebliche Tests mit Bedarf an großen Datenmengen durchgeführt werden bzw. eine erhebliche Unsicherheit toleriert werden. All diese Alternativen sind ebenfalls mit einem Verlust an Handlungsmöglichkeiten verbunden, so dass auch die Vermeidungsstrategie sicher nicht ohne einen Verlust an Freiheit realisiert werden kann.
- Grundsätzlich besteht die heute zu beobachtende Tendenz, dass die Entwicklung von Big Data und KI in der Medizin sich weniger auf die „harten“ Probleme mit sehr hohem Potential für die Patientenversorgung konzentriert, sondern sich auf weniger regulierte Themenfelder, wie Monitoring von Vitalfunktionen für Life-Style und Selbstoptimierung, fokussiert. Diese Anwendungen scheinen auf eine hohe Akzeptanz in einer ausreichend großen Community zu stoßen, so dass eigentlich dringend benötigte Ressourcen für Fortschritte im medizinischen Kerngebiet verlorengelangen.
- Ein politischer Diskurs ist zweifellos nötig im Bereich von versicherungsrelevanten Nutzungen von Big Data in der Medizin. Die Prognosegenauigkeit der entsprechenden KI-Tools ist zwar auf den ersten Blick oft verblüffend hoch, reicht gerade bei komplexen Erkrankungen jedoch

selten aus, um konkrete Handlungsempfehlungen abzugeben. Als Beispiel seien Risikoanalysen für spätere kardiovaskuläre Probleme genannt. Das Wissen, auf Grund eines komplexen Musters in Daten ein um 5 % erhöhtes Risiko für einen Herzinfarkt zu haben, ist für das Individuum wenig relevant. Es reicht nicht aus, um spezielle Therapien zu beginnen, in Konsequenz empfiehlt sich da die bekannte gesunde Lebensweise. Für diesen Ratschlag ist jedoch eine Big-Data Analyse nicht nötig, eine gesunde Lebensweise ist immer zu empfehlen. Daher ist der individuelle Nutzen mit heutigem Stand der Technik außer in speziellen Einsatzbereichen, wie Home-Monitoring bei chronisch erkrankten Patienten, noch beschränkt. Eine solche Risikoprognose ist jedoch für einen Krankenversicherer oder Lebensversicherer extrem relevant: es besteht hier bei Einsatz von Big Data Analysen dieselbe ethische Grundproblematik wie bei Genomanalysen. In diesen Fällen ist eine Einschränkung der persönlichen Freiheit durchaus möglich, ohne dass umgekehrt ein entsprechend hoher Nutzen für das Individuum realisiert werden kann.

Zusammenfassend besteht im Bereich Big Data in der Medizin ganz sicher ein hoher Bedarf für einen gesellschaftlichen Diskurs und politische Entscheidungen, um die notwendige Balance zwischen individuellen Rechten heute und einer nachhaltig hochwertigen Gesundheitsversorgung in der Zukunft zu halten. Eines ist jedoch sicher: die Unterlassungsalternative wird uns sicher Freiheit nehmen. Simple Unterlassung wird sehr wahrscheinlich mit einem höheren Verlust an Freiheit und Lebensqualität verbunden sein als eine ausbalancierte Strategie für die zügige Entwicklung einer Technologie, die uns helfen kann, die Herausforderungen einer alternden Gesellschaft zu bewältigen.

Das große Zeitalter einer neuen Dichtkunst – Automation, Autonomie, Schreiben

Im *Prager Tagblatt* erscheint 1922 ein Artikel mit dem lakonischen Titel „Die Schreibmaschine“. Zu diesem Zeitpunkt ist sie alles andere als neu. Seit Anfang des Jahrhunderts werden Schreibmaschinen mit stetig wachsenden Absatzzahlen seriell produziert, in den USA etabliert sich neben Remington die Firma Underwood 1901, deren Modell No. 5 zum Inbegriff der modernen Schreibmaschine wird. In Deutschland übernimmt die Firma Adler die Marktführung, als Fahrradwerk steigt sie 1896 in die Schreibmaschinenproduktion ein, 1922 verkauft sie 32 000 Exemplare. Unter Autorinnen und Autoren verfängt die Schreibmaschine dennoch nicht. Allenfalls Journalisten verwenden sie, schließlich sind es Verlage, die zunehmend Typoskripte statt unleserlicher Manuskripte verlangen – mit dem Effekt, dass nun Ehefrauen oder Sekretärinnen tippen. Autorinnen bleiben bei der Handschrift. Ohne Einbindung in eine etablierte literarische Praxis kann die Schreibmaschine im Feuilleton daher zum Ausgangspunkt eines Phantasmas werden: „Der Tag, an dem es gelungen sein wird, den Schriftsteller ganz auszuschalten, und die Schreibmaschine unmittelbar in Tätigkeit zu setzen, wird das große Zeitalter neuer Dichtkunst einleiten.“¹

Der Beitrag stammt vom österreichischen Journalisten und Autor Alfred Polgar, seine Pointe ist satirisch, nicht visionär. Nimmt man ihn ernst, fußt er auf einer Missdeutung: Die Schreibmaschine zeichnet sich nicht durch ihre Selbsttätigkeit, sondern ihre ökonomische und geschlechtliche Kodierung aus. Um 1900 ging der Begriff des Typewriters zeitweise von der Maschine

¹ Polgar, A., „Die Schreibmaschine“, in: Ders., *Kleine Schriften*, Band 4, 246–248, 246.

auf maschinenschreibende Frauen über. Die Maschinen eröffnen ihnen die bis dahin von Männern dominierte Bürowelt.

Polgar knüpft zwar an ein Phantasma an, das schon von den Automaten des Barocks her vertraut ist. Berühmt sind die Ente oder der Flötenspieler von Jacques Vaucanson, 1774 stellt Pierre Jaquet-Droz seinen mechanischen „Schreiber“ vor, der einen beliebigen Text mit bis zu 40 Zeichen schreiben kann. Polgar spinnt seine Maschinenfiktion dagegen während der Hochphase der Mechanisierung, die der Taylorismus und Fordismus zu Beginn des 20. Jahrhunderts einläuten. Ab den Zehnerjahren wandern die tayloristischen Effizienzobsessionen von der Fabrik über Büros bis in den häuslichen Alltag hinein. Das Schreiben bleibt davon nicht unberührt. Frank Bunker Gilbreth, neben Taylor einer der Begründer des Scientific Managements, führt Ende der Zehnerjahre eine Reihe von Schreibmaschinenexperimenten durch, die die Leistungsfähigkeit des Tippens steigern sollen: „Gemessen werden die Geschwindigkeit und die Fehlerhäufigkeit, aber auch die exakten Bewegungen der Finger, der Hände und des Körpers.“² Im Zuge der neuen Arbeitswissenschaften und Psychotechnik wird der Mensch auf der Grundlage physiologischer Leistungen neu entworfen und über Begriffe und Funktionen bestimmt, die empirisch und apparativ vermessen werden können.

Aus der Psychologie kommend ist der Begriff des Automatismus bereits zuvor Thema. Dort geht es mit willkürlichen motorischen oder automatischen Phänomenen um physische, psychische und geistige Bewegungsabläufe, die sich jenseits der mentalen Kontrolle ereignen. Besonders im Anschluss an die Philosophie von Henri Bergson in Frankreich und William James in den USA taucht gleichzeitig die Frage auf, ob mechanistische Abläufe als Grundlage der Kreativitätserzeugung dienen können. In Polgars Maschinenvisionen fließen die beiden Stränge des Automatismus als gleichförmiger, unbewusster Handlungsablauf einerseits, der Basis einer neuen materialistischen Kreativität andererseits, zusammen. Anfangs schränkt die Maschine die „menschliche Mitarbeit immer mehr und mehr“ ein, schließlich übernimmt sie das Schreiben aber nicht nur als physischen, sondern als

² Stiegler, B., Der montierte Mensch, Paderborn 2016, 118.

geistigen Akt. Dies wirft die Frage auf, inwiefern es sich hier bloß um eine fixe Idee handelt, oder fügt sich Polgars Maschinenfiktion in die Signatur seiner Epoche? Aber falls dem so wäre: Sollte man sie dann nicht eher in einem avantgardistischen Manifest, weniger im spielerischen Format eines Feuilletonartikels vermuten?

1. Die Schreibmaschine als blinder Fleck

In seinem *Technischen Manifest der futuristischen Literatur* von 1912 hatte der italienische Futurist Tommaso Filippo Marinetti zu einer Abschaffung des Ichs und der bürgerlichen Literatur aufgerufen; die neue Kunst solle dagegen „direkt ins Universum eintreten und eins mit ihm werden“.³ In ihrer Absage an die klassische Autorschaft, die die Genese und Einheit des Textes an das individuelle Dichtersubjekt und dessen schöpferischen Geist bindet, kommen Futurismus, Dadaismus und Surrealismus trotz aller politischen und ästhetischen Divergenzen überein. Der gemeinsame Nenner ihrer ästhetischen Verfahren, mit denen sie das Schreiben aus der bürgerlichen Tradition lösen und revolutionieren wollen, lässt sich darin identifizieren, dass sie es von einer geistigen in eine händische Praxis umwandeln. „Die schreibende Hand scheint sich vom Körper zu lösen“, heißt es in der Erläuterung zum *Technischen Manifest der futuristischen Literatur*, „und sich ins Freie zu verlängern, weit weit weg vom Gehirn, das ebenfalls losgelöst vom Körper ätherisch geworden ist und von oben mit erschreckender Klarheit die unerwarteten Sätze betrachtet, die der Feder entfließen.“⁴ Die Avantgarden bedienen sich einer ganzen Reihe ästhetischer Verfahren: automatisches Schreiben, der Einsatz der Schere, um die reproduzierte Schrift des Massenmediums Zeitung zu bearbeiten, die Zufallsproduktion oder Montage. Den Verfahren ist gemein, dass sie das Schreiben als veräußerlichten und materialistischen Prozess reformulieren. Obwohl die Schreibmaschine ihren Anliegen entgegengekommen wäre, fehlt sie im Katalog der avantgardistischen Instrumente.

³ Marinetti, F. T., *Manifeste des Futurismus*, Berlin 2018, 71.

⁴ Ebd., 52.

1907, zwei Jahre vor Erscheinen von Marinettis erstem Manifest im Pariser *Le Figaro*, ist in einer französischen Fachzeitschrift zu lesen: „Die Schreibmaschine verhält sich zur Feder wie das Fahrrad zum Gehen, das Automobil zum Fuhrwerk. Sie ist ein Instrument der Beschleunigung, das heißt des Fortschritts.“⁵ Marinetti, der den Fortschritt fetischisiert, träumt zwar von diversen Mensch-Maschine-Symbiosen und erklärt die Geschwindigkeit zur „neuen ethischen Religion“ des Zeitalters; er selbst bleibt aber bei der Handschrift. In seinen Listen von Maschinen, die er wie Flugzeug oder Bombe als existierende vergöttert oder wie das schnurlose Telefon als zukünftige imaginiert, taucht sie ebenfalls nicht auf.

Die Maschine hätte es aber auch erlaubt, sich der Feder als klassischem Schreibutensil autoritärer, männlicher Autorschaft bequem zu entledigen. 1889 stellt *Wilhelm Spemann's Illustrierte Zeitschrift für das Deutsche Haus* verschiedene Schreibmaschinentypen vor und klärt darüber auf, was „Das Schreiben mit der Maschine“ bedeutet: „Maschinen überall, wohin das Auge blickt“, so eröffnet der Beitrag, um zu erläutern, dass es angesichts dieser Entwicklungen „nur natürlich“ sei, dass sich Ingenieure auch der Mechanisierung des Schreibens gewidmet haben, und zwar mit der Konsequenz, „die Feder, das Symbol männlichen geistigen Schaffens, durch eine Maschine zu ersetzen.“⁶ Noch in den 20er Jahren greifen Tristan Tzara oder Hans Arp hingegen zur Schere, um auf der Basis der gesetzten und technisch reproduzierten Schrift der Zeitung mit dem linearen Schriftregime der Feder zu brechen und eine intentionslose Textproduktion herbeizuführen.

Polgars Schreibmaschinenreflexion verweist rückblickend auf eine Leerstelle, die die Schreibmaschine in der Praxis, aber mehr noch in der Imagination der europäischen Avantgarden markiert. Dies irritiert nicht zuletzt deshalb, weil die Fiktion Polgars, der von der Wiener Moderne herkommt und darin einem Ästhetizismus zugerechnet wird, den die Avantgarden attackieren, weitaus radikaler ausfällt als deren Proklamationen. So vehement

⁵ *Revue dactylographique et mécanique* 1 (April 1907), 4, zit. n. Gardey, D., *Schreiben, Rechnen, Ablegen. Wie eine Revolution des Bürolebens unsere Gesellschaft verändert hat*, Konstanz 2019, 97f.

⁶ *Das Schreiben mit der Maschine*, in: *Wilhelm Spemann's Illustrierter Zeitschrift für das Deutsche Haus*, Band 2, 1889, 863.

die Avantgarden gegen das humanistische und hermeneutische Subjekt der Aufklärung vorgehen, und es negieren wollen, es bleibt auf die eine oder andere Weise in die Textproduktion – als Intuition, Körper oder Unbewusstes – involviert. Polgar entkoppelt das Dichtersubjekt dagegen vollständig von der Textgenese und dreht die klassischen Kategorien auf den Kopf.

2. Die Fiktion einer autonomen Schreibmaschine

Um 1800 bildet sich in der Literatur ein modernes Verständnis der Autonomie aus. Einerseits wird sie parallel zur Ökonomisierung des Buchmarkts als eigenständige Sphäre entworfen. Andererseits wird sie direkt an die Textgenese gekoppelt. Der Begriff des Genies wird auch deshalb wichtig, weil es die Eigengesetzlichkeit und Einheit des Textes ans Individuum bindet. Prominent wird der Autonomiegedanke erneut im Ästhetizismus um 1900, der ihn auf die Formel des *l'art pour l'art* bringt. Die Avantgarden polemisieren sowohl gegen diesen Ästhetizismus, dem sie den Rückzug in ein „inhaltloses, bequemes und unbewegtes Leben“ (Dadaistisches Manifest von 1918) vorwerfen, als auch gegen die bürgerliche Literatur generell.

Die Kritik der Dadaisten an der Autonomie speist sich vor allem aus den Erfahrungen des Ersten Weltkriegs. Sie machen die bürgerliche Literatur für die Mobilisierung nationaler Ideologien und Idealisierung des Kriegs mitverantwortlich. Mit deutschen Soldaten, die mit Goethe im Tornister an die Front marschierten, hat der literarische Humanismus seine Komplizenschaft mit der Barbarei erwiesen. Weil der Anspruch auf Autonomie die Sphären getrennt hält, verschleiert sie die Produktionsbedingungen der bürgerlichen Literatur, konkret ihre Beziehung zu Ökonomie und Politik. Literatur, Kultur und Geist, so Raoul Hausmann, werden als „neues Mittel zur Stabilisierung der Bourgeoisie missbraucht“. Die Zufalls-, automatischen und regelgeleiteten Schreibverfahren der Avantgarden zielen nicht nur auf den Bruch mit der Tradition. Sie sind gleichzeitig dem Versuch geschuldet, das Schreiben neu zu begründen, und das heißt jenseits der Werte der bürgerlichen Literatur, die um Begriffe wie Autonomie, Genie oder Geist zentriert sind.

Die Schreibmaschine hätte die Reformulierung des Schreibens fördern können, um die Schreibpraxis von einer inneren und geistigen in eine veräußerlichte, materialistische und physiologische Aktivität zu überführen. Zwar haben die Avantgarden kein Interesse an einer großen Dichtkunst, ihre Arbeiten zielen eher auf Desillusionierung und Anti-Kunst. Dennoch hätte sich Polgars Fiktion mit ihrem Vorhaben vertragen, denn zuallererst erfüllt die Aufwertung der Maschine eine simple Funktion: Nur als autonome kann sie zum Urheber ästhetisch eigengesetzlicher Werke werden. Polgars Verschaltung von Automation und Autonomie basiert darin auf einem ebenso einfachen wie effektiven Kniff: In dem Moment, in dem das romantische Genie seine Glaubwürdigkeit einbüßt, lässt er es die Plätze wechseln und vom Menschen in die Maschine wandern. Wenn er den Autor mit der autonomen Maschine darüber hinaus von der Arbeit des händischen Schreibens freistellt, versieht er den Wandel sogar mit utopischen Vorzeichen.

3. Computergeneriertes Schreiben

Ihren satirischen Gehalt verliert die Vorstellung, dass die Maschine einerseits ein Schreiben jenseits menschlicher Agency verwirklichen und andererseits das Zeitalter einer neuen Dichtkunst einleiten könne, in der zweiten Hälfte des 20. Jahrhunderts. „Der herkömmliche Schriftsteller, so wie er bis jetzt gewesen ist“, schreibt Italo Calvino in *Kybernetik und Gespenster*, „ist bereits eine schreibende Maschine“ gewesen, der nun potenziell von Computern ersetzt werden könne; was in „romantischer Terminologie Genius“ hieß, kann mit dem Computer auf eine mathematische Basis zurückgeführt werden.⁷ Ab den Vierziger- und Fünfzigerjahren vollzieht sich mit dem Aufkommen des Digitalcomputers und der Kybernetik eine Analogisierung des menschlichen Denkens mit Maschinen. Dies wird einerseits zur Basis, um die Fähigkeiten des Menschen auf die Maschine zu übertragen, andererseits soll die Funktionsweise des Computers nun die des menschlichen Denkens er-

⁷ Calvino, I., *Kybernetik und Gespenster. Überlegungen zu Literatur und Gesellschaft*, München/Wien 1984, 20f.

klären. Nicht dass Maschinen schreiben können, steht für Calvino zur Disposition, sondern ob Computer auch ebenso „die eifersüchtig gehüteten Attribute der psychologischen Intimität, der erlebten Erfahrung, der Unvorhersehbarkeit sprunghafter Launen“ von Menschen wiedergeben können.

Weder in den Sechzigerjahren, noch danach lösen sich die an die Rechenanlagen gerichteten Erwartungen ein. Dennoch bleiben sie unterschwellig präsent, in den letzten Jahren sind sie sogar mit neuer Wucht zurückgekehrt. Die jüngste Veröffentlichung der GPT-3 Software des kalifornischen Unternehmens OpenAI im Sommer 2020, scheint hier nur einen vorläufigen Höhepunkt innerhalb eines Hypes zu bilden, den die maschinellen Lernverfahren seit den Zehnerjahren des 21. Jahrhunderts ausgelöst haben. „GPT steht für ‚Generative Pretrained Transformer‘ und bedeutet, dass das Modell ohne menschliche Hilfe arbeitet“, erklärt die *Süddeutsche Zeitung* nach der Bekanntgabe einer Testversion: „Es schreibt Kurzgeschichten, Songtexte, Betriebsanleitungen, Bilanzanalysen, juristisch korrekte Abhandlungen oder eine Harry-Potter-Fortsetzung im Stil von Raymond Chandler.“⁸ Der Mediensendienst *Bloomberg* spekuliert sogar darüber, ob man sich im Rückblick auf 2020 nicht zuallererst an GPT-3, weniger an Covid-19 oder die amerikanische Präsidentschaftswahl erinnern wird.⁹ Steht uns nun also doch noch das Zeitalter einer neuen Dichtkunst bevor, und das frei von der Ironie Polgars?

Zwar lässt sich eine Software wie GPT-3 in eine Verbindung mit Polgar, oder wenn man noch weiter zurückgehen will, dem Automaten von Jaquet-Droz stellen. Ein solcher Vergleich verfehlt aber die entscheidenden Aspekte der technologischen Entwicklungen. Wenn die gegenwärtigen Erfolge einerseits wie die Einlösung einer Fiktion anmuten, die sich lange angekündigt hat, haben sich die technologischen Bedingungen andererseits fundamental gewandelt. Daher stellt sich auch die Frage, was eine Software wie GPT-3 überhaupt bedeutet, mittlerweile anders. Wie die Maschinenfiktionen Polgars oder Calvinos sind die gegenwärtigen Techniken in soziale, ökonomische und kulturelle Zusammenhänge eingebettet. Dass sich die Technik nicht im Technischen erschöpft, gilt heute mehr denn je.

⁸ Moorstedt, M., Federhalter, in: *Süddeutsche Zeitung*, 6. August 2020, 9.

⁹ Cowen, T., *Artificial Intelligence Is the Hope 2020 Needs*, in: *Bloomberg*, 21. Juli 2020.

Gegenüber Polgar und Calvino haben sich die technologischen Grundlagen heute grundsätzlich geändert. Für sie lag die Maschine noch als vergleichsweise simples und in sich geschlossenes Gerät vor. Heute sind wir mit einem globalen Gefüge verteilter und vernetzter Computer konfrontiert. Schon deshalb greift eine Lesart, die eine Anwendung wie GPT-3 als Erfüllung einer historischen literarischen Imagination versteht, zu kurz. Gleichzeitig stellt dies vor die Herausforderung, wie sich das technologische Gefüge überhaupt umreißen und in eine Beziehung zum Schreiben setzen lässt.

Die Spekulation, ob wir in Zukunft tatsächlich autonom generierte Erzählungen lesen werden, hängt nicht nur von den kulturellen Kontexten ab, die darüber entscheiden, mit welchen Zuschreibungen und Bedeutungen die digitalen Objekte ausgestattet werden. Vielmehr bleiben computergenerierte Romane angesichts des Wandels, dem die Sprache und Schrift in der digitalen Sphäre unterworfen sind, letztlich nur ein Epiphänomen. Wie lassen sich aber die Entwicklungen fassen, aus denen eine Technik wie GPT-3 hervorgeht? Um dieser Frage nachzugehen, greife ich auf eine weitere Fiktion zurück. Wie bei Polgar und Calvino kommt es hier weniger auf die Richtigkeit, sondern ihren heuristischen Wert an: Der phantasmatische Überschuss erlaubt es, besser nach der Bedeutung der Technologie zu fragen.

4. Universalmaschinen

2015 veröffentlicht der britische Autor Tom McCarthy im *Guardian* eine Maschinenfiktion unter dem Titel „The death of writing – if James Joyce were alive today he’d be working for Google“.¹⁰ Der Titel weist schon darauf hin, dass McCarthy sich weniger um eine computergenerierte Dichtkunst im allgemeinen kümmert, vielmehr orientiert sich seine Fiktion an einem kanonischen Roman der Moderne. James Joyce *Ulysses* (1922) schildert einen Tag im Leben seines Protagonisten Leopold Bloom, er folgt ihm vom morgendlichen Aufstehen über seine Botengänge durch die Stadt und bis zur nächtlichen Rückkehr nach Hause, die letzte Szene spielt im ehelichen Schlafzim-

¹⁰ McCarthy, T., The death of writing – if James Joyce were alive today he’d be working for Google, in: *Guardian*, 7. März 2015.

mer. Berühmt ist der Roman unter anderem für seine scheinbar hierarchiefreie Schilderung sämtlicher Nebensächlichkeiten und Details, die er in einer chronologischen Abfolge notiert. McCarthy verbindet dies mit einem literarischen Projekt, das auf Totalität aus ist, möglichst alles aufzuzeichnen.

Was Joyce noch händisch begrenzt und selektiv ausführte, kann die Technologie heute aber viel umfänglicher leisten. Ein neuer Ulysses wird daher maschinell generiert sein. McCarthy geht aber noch weiter, wenn er schreibt, dass er im Grunde schon da sei: „Weit davon entfernt, gar nicht geschrieben werden zu können, wird der alles umfassende Große Report in einem Fort um uns her verfasst – aber dahinter steht kein menschlicher Romancier, sondern ein neutrales und indifferentes binäres System, dessen einziges Ziel es ist, sich selbst fortzuspinnen, ein auto-alphaaisierendes und auto-omegaisierendes Skript.“ Auch in diesem Szenario ist der Autor ausgeschaltet und das Schreiben an selbsttätige und vom Menschen autonome Maschinen übergegangen.

Wie Polgars Schreibmaschine ist auch McCarthys binäres System, das alles autonom und automatisch erfasst und speichert, und daraus so etwas wie einen kohärenten Text erzeugt, ein Phantasma. Schon die periodisch auftauchenden Leaks aus der Digitalökonomie legen nahe, dass es mit der Autonomie der Automatisierung nicht so weit her ist. Zuletzt geriet Apple im Juli 2019 in die Schlagzeilen, weil das Unternehmen Leute beschäftigt, die heimlich Transkriptionen von Gesprächen anfertigen, die sie über ihre Sprachassistentin Siri abhören. Was bedeutet das aber?

Die Automatisierung und autonome Funktionsweise der Sprachsysteme bildet ein vorrangiges Ziel der Technologiebranche. 2016 berichtete der Mediendienst *Bloomberg* bereits: „Amazon, Apple, Microsoft und Chinas Baidu haben sich auf eine weltweite Jagd nach Terabytes menschlicher Sprache gemacht. In Metropolen rund um den Globus hat Microsoft Modell-Apartments eingerichtet, um aufzuzeichnen, wie sich Freiwillige in einem heimischen Setting unterhalten.“¹¹ Die Sprachsysteme mögen in vielerlei Hinsicht auf

¹¹ Cao, J./ Bass, D., Why Google, Microsoft and Amazon love the sound of your voice, in: *Bloomberg*, 13. Dezember 2016, <https://www.bloomberg.com/news/articles/2016-12-13/why-google-microsoft-and-amazon-love-the-sound-of-your-voice>, letzter Zugriff 26. Oktober 2020.

menschliche Eingriffe angewiesen bleiben, kennzeichnend für die Entwicklungen ist jedoch, dass es sich bei all dem um ein ökonomisches und ideologisches Projekt handelt. Rita Railey spricht von einer „neuen technolinguistischen Doxa“, die darauf fußt, dass „alles, jede Einschreibung und jeder Sprechakt jederzeit und auf Abruf verfügbar gemacht werden soll, egal wo wir sind.“¹²

Was auf die Sprache zutrifft, lässt sich in ähnlicher Weise auf sämtliche Umweltdaten übertragen, die ebenfalls umfassend aufgezeichnet werden sollen. Zwar fällt die Aufzeichnung unstrukturierter Umweltdaten auf der offenbar auch McCarthys Maschine basiert, noch komplexer als die Sprachverarbeitung aus. Die Umwelt lässt sich nicht einfach erschließen, sie muss erst algorithmisch urbar gemacht werden. Wie im Fall der Sprachsysteme steht ein Heer an global verstreuten und unterbezahlten Crowd- oder Clickworkern bereit, die die Daten prüfen und kategorisieren. „Es gab Sachverständige für den Darm wie Herr Pradhan“, schreibt die *New York Times* 2019 über die Firma iMerit, „und Spezialisten, die einen guten von einem schlechten Husten unterscheiden können. Es gab Sprachexperten und Personen, die Straßensituationen identifizierten. Wer ist ein Fußgänger? Handelt es sich dort um eine doppelte gelbe oder eine gestrichelte weiße Linie? Ein Roboterauto muss den Unterschied eines Tages wissen müssen.“¹³ Unter der Hand wirft die Tätigkeit eines global verstreuten, digitalen Proletariats einen Schatten auf Polgars Utopie und die neuerlichen Versprechen, dass die Automatisierung und algorithmische Autonomie zu einer Entlastung oder gar Emanzipation von der Arbeit führen werde.

McCarthys Fiktion hebt dagegen auf den geschilderten Umstand ab, dass sich mit der gegenwärtigen Technologie auch eine an den Roman der Moderne gebundene Vorstellung einlösen könnte. Im Unterschied zu GPT-3, dessen Stärke in der Reproduktion von bereits vorhandenen Mustern liegt, und das sich deshalb vornehmlich am Konventionellen orientiert, lädt

¹² Railey, R., Algorithmic Tanslations, in: *The New Centennial Review* 16/1 (2016), 115–138, 122.

¹³ Metz, C., AI Is Learning From Humans. Many Humans, in: *New York Times*, 16. August 2019.

McCarthy seine algorithmische Maschine mit der Programmatik des modernistischen Romans auf, dessen Anspruch ein ästhetischer und epistemologischer war, die Gesellschaft in ihrer schon damals nicht zu bewältigenden Fülle zu durchdringen und darzustellen. Ich hebe dagegen auf einen anderen Aspekt ab. Im Unterschied zu McCarthy's Fiktion, der die algorithmische Maschine nochmals an eine bestimmte Vorstellung der Literatur, und was sie leisten soll, zurückbindet, lautet mein Argument, dass das entscheidende Moment des technologischen Gefüges zuallererst jenseits der Literatur zu suchen ist. Seine Bedeutung liegt in der generellen Transformation, wie der Mensch in der Beziehung zur Sprache und Schrift gedacht wird.

5. Die algorithmische Maschine der Digitalökonomie

Polgars Fiktion mag sich in gewisser Weise realisiert haben, die Pointe liegt jedoch darin, dass sich gleichzeitig die Vision der Maschine gewandelt hat. Lapidar formuliert liegt diese nicht in einer perfekten Simulation, sondern in einer Emanzipation von der Sprache und Schrift. Dies vollzieht sich auf mindestens zwei Ebenen. Zum einen stellt die natürliche Sprache innerhalb der digitalen Sphäre einen Mangel dar, den es zu überwinden gilt. Zum anderen wird sie aufgrund ihrer maschinensprachlichen Kodierung zu einem Rohstoff, der sie in einem ungekannten Maß zu kommodifizieren erlaubt. Beides lässt sich als direkter Angriff auf die Sprache verstehen, wie sie im Verständnis des Menschen als sprachbegabten Tier artikuliert ist. Im 18. Jahrhundert inthronisieren Autoren wie Johann Gottfried Herder die Sprache sogar als Königsweg zum menschlichen Bewusstsein. Erst die Sprache bringt das Bewusstsein und die spezifische Wahrnehmung der Welt überhaupt erst mit hervor. In der digitalen Kommunikationstechnologie verlieren sprachliche Zeichensysteme, aber selbst die Verbalisierung von Gedanken dagegen an Bedeutung. Exemplarisch lässt sich dies an drei Momenten nachvollziehen: Biometrie, Soziometrie und Telepathie.

Die Biometrie hat die statistische Erfassung des Menschen über biometrische Daten zum Ziel. Ihre moderne Vorgeschichte hat sie in der Anthropometrie des 19. Jahrhunderts. Allgemein kommt sie der Notwendigkeit

moderner Gesellschaften nach, die Identität des Einzelnen – ursprünglich handelt es sich um rückfällige Straftäter – möglichst effizient und verlässlich zu identifizieren. In der Digitalökonomie verschmelzen kapitalistische mit staatlichen und sicherheitstechnischen Interessen. Im vorliegenden Kontext interessiert lediglich, dass die biometrische Identifikation ohne Fragen des Bewusstseins, der Sprache oder sinnlichen Wahrnehmung auskommt. Der Nachweis der Identität wird allein durch den statistisch-biologisch identifizierbaren Körper erbracht, Geist, Bewusstsein oder eine sprachliche Selbstauskunft sind ohne Belang.

Die Biometrie gründet auf der Natur. Aus technologischer Perspektive ist sie daher fehleranfällig, weil Natur evolutionsbiologisch Abweichung bedeutet. Dennoch erlaubt sie eine stabile statistisch-mathematische Identifizierung des Einzelnen, und ist deshalb von der Entsperrung des Smartphones bis zur Regulierung des Grenzverkehrs für die Digitalökonomie als vergleichsweise einfache Lösung nach wie vor attraktiv.

Weitrechender ist die soziometrische und verhaltensbasierte Erfassung der Menschen. Die Hybris der Techelite, dass sie und ihre Dienste uns besser kennen als wir uns selbst, basiert auf dem Wechsel der Wissensregister von sprachlichen auf außersprachliche Parameter. Die Prämisse lautet, dass unser Verhalten von Microexpressions in der Gesichtsmimik und Körpersprache, wie sie das Affective Computing dechiffriert, über das Klick- und Kaufverhalten in Onlinetransaktionen, bis hin zu Bewegungsprofilen oder unserem individuellen Fahrverhalten, besser und zuverlässiger Auskunft über uns und unsere zukünftigen Aktivitäten gibt, als wir es selbst je könnten. Calvinos Frage, ob die Maschine unsere eifersüchtig gehüteten Attribute der psychologischen Intimität, der erlebten Erfahrung, der Unvorhersehbarkeit sprunghafter Launen aufzeichnen können, wird von Leuten wie Eric Schmidt, ehemaliger Executive Chairman von Alphabet, nicht nur bejaht – seine Pointe wäre vielmehr, dass sie dies viel besser können als wir.

Die Sprache bildet in all dem nur einen Datenträger neben anderen, entscheidend ist weniger was wir sagen, sondern dass die Sprache in dem was und wie wir es sagen, zu einem weiteren Datenspender werden kann, dessen Verwendung aber nichts mit unseren Intentionen zu tun hat und sich zudem unserer Kenntnis entzieht. „Nicht was in ihren Sätzen steht, zählt, sondern

deren Länge und Komplexität, nicht was Sie auflisten, sondern die Tatsache, dass sie eine Liste aufstellen, nicht das Bild selbst, sondern die Wahl des Filters und der Grad der Sättigung, nicht was Sie enthüllen, sondern wie Sie Privates mit andere teilen oder eben nicht“,¹⁴ so erklärt Shoshanna Zuboff diesen instrumentellen Zugriff auf die Sprache. In *Das Zeitalter des Überwachungskapitalismus* spricht sie von zwei Texten, die unsere Interaktion in der digitalen Sphäre produziert, neben der Kommunikation auf der Oberfläche, der den von uns generierten und zugänglichen Text bildet, entsteht im Hintergrund ein „Schattentext“, der letztlich der mächtigere ist und unsere digitale Identität auswertet und festschreibt, ohne dass wir Zugang zu diesem Text hätten. Die Emanzipation der digitalen Kommunikation von der Sprache und Hermeneutik bedeutet demnach kein Verstummen. Vielmehr sollen die digitalen Subjekte so viel wie möglich kommunizieren, aber die quantitative Steigerung ist darin rein kapitalistisch motiviert und zielt auf jenen Verhaltensüberschuss, den Zuboff als Merkmal des Überwachungskapitalismus identifiziert. In der Produktion von Verhaltensdaten, die sich für kommerzielle Zwecke verwenden lassen, wird die Sprache zu einem Rohstoff, der sich profitabel ausbeuten lässt. Frederic Kaplan hat die Digitalökonomie auf den Begriff eines linguistischen Kapitalismus gebracht, dem die Sprache als maschinenlesbarer Kode unweigerlich unterworfen ist.

Am radikalsten vollzieht sich die Abkehr von der Sprache schließlich in der Telepathie. In einer Q & A-Session 2015 identifiziert sie Mark Zuckerberg als „ultimative Kommunikationstechnologie“: „Ich glaube, die Technologie macht es möglich, dass wir uns eines Tages vollständige, gehaltvolle Gedanken direkt zusenden. Du kannst an etwas denken und deine Freunde werden sofort in der Lage sein, es auch zu erleben, sofern du es teilen willst. Dies wäre dann die ultimative Kommunikationstechnologie.“¹⁵ Die Optimierung der Kommunikation bedeutet schneller, direkter und mehr zu kommunizieren. Eine natürliche Sprache, die auf Semantik und hermeneutischen Sinn

¹⁴ Zuboff, Sh., *Das Zeitalter des Überwachungskapitalismus*, Frankfurt a. M./New York 2018, 316.

¹⁵ Zuckerberg, M., *Building Global Community*, Facebook post, 16. Februar 2017, <https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634>, letzter Zugriff 26. Oktober 2020.

gründet, ist für die Anforderungen eines reibungslosen Kapitalismus zu langsam und in ihrer Mehrdeutigkeit zu fehleranfällig. Außerdem unterliegt sie als Medium, dass einerseits der Interpretation bedarf und darin andererseits manipulationsanfällig ist, einem grundsätzlichen Verdacht.

Dass die Technologie der hermeneutischen Sprache unverträglich gegenübersteht, zeigt sich bereits in den Anfängen der modernen Informationstheorie. Als deren Begründer erklärt Claude Shannon Semantik und Sinn gleich auf der ersten Seite der Vorrede zu seiner mathematischen Kommunikationstheorie für irrelevant. Die Nachricht spielt nur insofern eine Rolle, als dass sie aus einem möglichen Vorrat von Nachrichten ausgewählt werden muss. Shannon definiert sie als Zufallsfolge in der Zeit, die wahrscheinlichkeits theoretisch berechnet werden kann. Die Telepathie, an deren technologischen Umsetzung neben Zuckerberg das Militär, Universitäten und weitere Techentrepreneure wie Elon Musk mit seiner 2017 gegründeten Firma Neuralink forschen, lässt sich in der Negation von Semantik und Hermeneutik in eine direkte Nachfolge der modernen Informationstechnologie stellen.

Obwohl bereits erste Erfolge im Bereich der Gehirn-zu-Gehirn-Kommunikation gemeldet worden sind, lassen sich Fakt und Fiktion des Vorhabens kaum trennen, zumal die Forschung stark von PR-Maßnahmen überformt ist. Musk hat im Juli 2019 angekündigt, dass Neuralink die ersten Interfaces zwischen Gehirn und Maschine 2020 an Erkrankten austesten werde. Im August 2019 publizierten von Facebook engagierte Wissenschaftlerinnen in *Nature Communications*, dass es ihnen erstmals gelungen sei, Wörter in Echtzeit aus der Beobachtung von Gehirnaktivitäten abzuleiten. Jenseits von philanthropischen Pilotprojekten ist das Interesse der Telepathie für Unternehmen wie Facebook klar: Mit ihr fallen die Kommodifizierung und Kontrolle der Kommunikation potenziell in eins. Darin fügt es sich konsequent in den „Vorhersageimperativ“ (Zuboff), der die Digitalökonomie antreibt, und der darauf fußt, dass verlässliche Vorhersagen letztendlich erschöpfende Daten erfordern. Dass mit der Telepathie gleichfalls die Verhaltensmodifikation näher rückt, liegt ebenso nahe – und illustriert nebenbei, dass die Digitalökonomie die Subjekte, die sie zu erfassen und erforschen vorgibt, als solche überhaupt erst mit hervorbringt.

Ein konkretes Szenario, in dem der Traum der Telepathie greifbarer wird, lässt sich anhand einer Projektbeschreibung von Palantir entnehmen. Mittlerweile wird der Börsengang des 2003 gegründeten Start-up diskutiert, dass durch geheime Predictive Policing-Systeme, die Straftaten vorhersagen sollen, auf sich aufmerksam gemacht hat. Zwar fällt der Begriff der Telepathie nicht namentlich, sehr wohl gehört die fortschreitende „Mensch-Maschine-Symbiose“ aber zu den erklärten Unternehmenszielen. Der Name Palantir stammt aus J.R.R. Tolkiens Romantrilogie *Herr der Ringe* und beschreibt dort eine magische Kristallkugel, die es erlaubt, Ereignisse der Vergangenheit oder Zukunft über große Distanzen hinweg zu sehen.

Ein System für die Streitkräfte konkretisiert nun Zuckerbergs Traum eines ultimativen Kommunikationstools, dreht die Kommunikationssituation aber in gewisser Weise um: Das Subjekt ist nicht Sender, sondern Empfänger der Information. „Palantir Defense“, so lautet die Projektbeschreibung, „erlaubt es Frontkämpfern auf beispiellose Weise mit allen ihren Daten aus allen ihren Systemen von einem einzigen Zugriffspunkt aus zu interagieren. Unstrukturierter Nachrichtenverkehr, strukturierte Identitätsdaten, Beziehungsgrafiken, Tabellenkalkulationen, Telefonie, Dokumente, Netzwerkdaten, Sensordaten – sogar bewegte Videobilder können simultan und intuitiv durchsucht werden, ohne dass eine spezialisierte Abfragesprache nötig wäre.“¹⁶

Soweit lässt sich der Service auch als Version von McCarthys selbstschreibendem System verstehen, das darin kaum weniger fantastisch anmutet. Gleichzeitig nimmt es jedoch zwei Präzisierungen vor. Erstens macht ein solches System nur Sinn, wenn die Mitschrift der Maschine gebündelt und an eine Adressatin gerichtet werden kann: Hier ist der Empfänger ein in Kampfhandlungen involvierter Frontkämpfer, keine Leserin eines großen Reports. Zweitens macht das Szenario klar, was es heißt, dass die Daten einem „vollständigen Universum aus Daten“ entstammen. Sie müssen nicht nur verfügbar und in Echtzeit erfasst und ausgewertet, sondern darüber hinaus homogenisiert und in eine einheitliche Sprache übersetzt werden. Genau darin

¹⁶ Palantir Homepage, <https://www.palantir.com/solutions/defense/>, letzter Zugriff 26. Oktober 2020.

besteht aber die angebliche Leistung: „Palantir Defense versetzt Frontkämpfer in die Lage, Berge aus Daten schnell in Aktionspläne zu verwandeln, die jene Fragen aufwerfen, auf die sie eine Antwort brauchen und das in einer Sprache, die sie verstehen.“ In dem Szenario ist der Traum einer Universalsprache aufgerufen, der auch McCarthys Maschine zugrunde liegt. Diese umfasst nicht nur die Übersetzung sämtlicher gesprochener und geschriebener Fremdsprachen, sondern gleichzeitig alle erdenklichen Umweltdaten in eine verständliche Sprache. An dieser Stelle interessiert jedoch weniger diese Übersetzungs- und Homogenisierungsleistung, sondern dass es sich bei dieser Universalsprache um eine Befehlssprache handelt. Was am „einzelnen Zugangspunkt“ als „Sprache, die sie verstehen“ ankommt, ist ein „Aktionsplan“. Der Sinn dieser Sprache liegt in der Handlungsanweisung, keiner sinnorientierten Erschließung des Subjekts oder der Welt. Wie die digitale Schrift kommt sie ohne Repräsentation aus, und lässt sich direkt operationalisieren. In nichts anderem liegt ihre gewünschte Effizienz. Die Sprache muss nicht mehr verbalisiert oder in Zeichen überführt werden, sondern kann direkt wirksam werden. Eingeschmolzen auf maximale Effizienz bedeutet Kommunikation in diesem Modell, dass Information und Verstehen in der Handlungsanweisung oder Handlung zusammenfallen.

6. Schreiben im Horizont der algorithmischen Maschine

Kehrt man abschließend zu Polgars Schreibmaschinenfiktion zurück, birgt sie eine doppelte Ironie. Aus der Satire scheint Wirklichkeit zu werden und unbeabsichtigt hat sie sich darin hellstichtiger als die Programme der historischen Avantgarden erwiesen. Das Schreiben ohne Agency wird heute von der algorithmischen Maschine der Digitalökonomie realisiert, und das heißt an jenem blinden Fleck, den die Automatisierung des Schreibens durch die Maschine in ihrer Imagination hinterlassen hatte. Zwar ist die selbsttätige Maschine als autonome nach wie vor Fiktion. Sowohl im Scheitern der Implementierung, als auch dort, wo sie gelingt, darin aber Effekte erzeugt, die selbst nicht Teil des Systems sind, bleibt sie auf menschliche Interventionen und die Sinnorientierung der natürlichen Sprache angewiesen. Dennoch

strebt die Digitalökonomie auf eine Emanzipation von der Sprache und Schrift.

Das heißt wohlgermerkt nicht, dass die Sprache einfach verschwindet. Beispielsweise schwören die Technoavantgarde und Unternehmen wie Facebook auf das Storytelling als äußerst wirksame Form der Kommunikation. Eine Paradoxie des Phantasmas einer non-verbalen Kommunikation liegt darin, dass Zuckerberg und Musk, um Visionen wie die der Gehirn-zu-Gehirn-Übertragung verständlich zu machen und zu verbreiten, auf die natürliche Sprache und Stories zurückgreifen müssen. Das ist nur ein Beispiel, warum die Hegemonie des Codes nicht als einfache Ablösung von der natürlichen Sprache zu fassen ist. Zwar sind die Machtverhältnisse klar verteilt. Nicht zuletzt gewinnt die Digitalökonomie ihre Hegemonie aufgrund des Codes, der den referentiellen Zeichengebrauch der Sprache durch einen formal-operativen ersetzt. Dennoch verlaufen die Entwicklungen stets parallel, vielschichtig und verschlungen. Die algorithmische Maschine bleibt in all dem ein heterogenes und poröses Gefüge, das zahlreiche Einfallstore, Anknüpfungspunkte und Fluchtwege für die Sprache aufweist. Weil die Implementierung sich darüber hinaus nur stückweise, uneinheitlich und unvorhergesehen vollzieht, liegt es nahe, dass sie im selben Takt neue Erzählungen hervorbringen wird, wie sie mit alten Vorstellungen Schluss zu machen versucht. Zweifellos ließe sich hier ein Einsatzpunkt für eine Literatur suchen, die sich auf eine Auseinandersetzung mit den technologischen Bedingungen einlässt.

Einfache Schlüsse lassen sich für Literatur und Schreiben daraus aber nicht ziehen. Selbst die Anknüpfung an literarische Traditionen, sei es der Mainstream oder die Avantgarden, die nach über hundert Jahren mittlerweile selbst klassisch und eine Konvention wie jede andere geworden sind, stellen keine einfachen Blaupausen bereit. Die Digitalökonomie beherrscht das Spiel der Zerstörung und Neubegründung, De- und Reterritorialisierung inzwischen weitaus besser und wirkmächtiger als künstlerische Avantgarden. Außerdem unterwirft sie die Gesellschaft darin als Ganze einem Experiment, das sämtliche Ordnungen der Repräsentation in offene Prozesse überführt. „Datengesteuert zu sein bedeutet ohne Zielvorstellung oder Hypothese zu beginnen, keine stabile Basis anzunehmen, und seine Resultate

ständig zu ändern.“¹⁷ Nicht nur angesichts des Kontrollregimes der algorithmischen Maschine fällt es daher schwer, die referentielle Auflösung von Zeichenordnungen länger emanzipatorisch ausdeuten.

Gleichzeitig setzt die algorithmische Maschine die herkömmliche Sprache aber nicht einfach wieder in ihr Recht, auch nicht als bewahrender Gegenentwurf zur algorithmischen Maschine. Die Sprachkritik der Avantgarden, sei es, dass sie propagandistisch missbraucht werden kann oder epistemologisch unzuverlässig ist, ist nach wie vor unabgeholten. Ähnliches gilt für ihre Revolte gegen soziale Strukturen und Konventionen, oder die Kritik des Humanismus, der sich nicht nur durch die Kriege und seinen Eurozentrismus diskreditiert hat. Alles spricht dafür, dass der Mensch mittlerweile ebenso in seiner Beziehung zu anderen Arten, der Umwelt und Natur gedacht werden muss.

Vor diesem Hintergrund erweist sich die Preisgabe der Autonomie, wie sie die Avantgarden vornahmen, heute als fragwürdiges Unterfangen, und das nicht nur, weil die Autonomievermutungen heute eher in algorithmische Systeme als in Menschen gelegt werden. Bereits Theodor W. Adorno argumentiert, dass an der ästhetischen Autonomie in dem Moment festzuhalten ist, in dem revolutionäre Hoffnungen erloschen sind. Nur wo es der Literatur gelingt, eine Differenz zu ziehen, kann sie das Bewusstsein für eine andere Ordnung als die Bestehende, die auch dem vergangenen Leid Rechnung trägt, aufrechterhalten. Dies erreicht sie nicht durch Abschottung. Verfehlt wird die Autonomie nach Adorno dort, wo sie sich gar nicht erst auf das ihr Heteronome einlässt.

Dieser Zwiespalt, einerseits eine Autonomie zu behaupten, und sich andererseits auf das Heteronome der Literatur einzulassen, vollzieht sich heute, so meine These, vor dem Hintergrund der algorithmischen Maschine und ihrer Tendenz, sich von der Sprache und Schrift insgesamt zu lösen. Folgt man diesem Argument, lautet die Herausforderung, wie sich eine Autonomie des Schreibens wiedergewinnen lässt. Die Autonomie ist hierbei niemals einfach gegeben, weder lässt sie sich aus der Instanz der Autorin selbst noch der Literatur als Institution ableiten. Stattdessen kann sie nur immer wieder neu

¹⁷ Halpern, O./ LeCavalier, J./ Calvillo, N./ Pietsch, W., Test-bed Urbanism, in: *Public Culture* 25/2 (2013), 273–306, 294.

errungen werden. Ein wesentlicher Ansatzpunkt könnte hier darin liegen, das Schreiben gegenüber der algorithmischen Organisation des Individuums und der Gesellschaft stark zu machen. Das Schreiben und die algorithmische Verwaltung der Individuen lassen sich als Techniken der Vernetzung, Verknüpfung und Verkettung beschreiben. Ausgehend davon, ließe sich das Schreiben nicht nur als Technik und Praxis entwerfen, die konkrete Satzverkettungen stiftet, die quer zu algorithmischen Prozessen und Verknüpfungsregeln verlaufen, es vermag darin die Einbettung des Subjekts in die Satz- und Textgefüge zu reflektieren. Darüber hinaus kann die Literatur sich dem öffnen, was die algorithmischen und automatisierten Verkettungen ausschließen oder gar nicht erst mit einbeziehen, weil es den Formalisierungsprozessen der Programmierung entgeht. Seine Bedeutung behält das Schreiben in diesem Entwurf nicht zuletzt, weil die Satzverkettungen als eigengesetzliche in ihren Setzungen nachvollziehbar bleiben und ein Beziehungsgeflecht knüpfen, das sich in die Wahrnehmung heben und als Alternative zu algorithmischen Verknüpfungstechniken greifbar wird.

Die Berechnung der Bildung – Über das Verhältnis von Mündigkeit und Bildung in der digitalen Welt

1. Einleitung

Algorithmen, Big Data, Machine Learning und Künstliche Intelligenz zählen zu den bemerkenswerten technischen Entwicklungen der letzten Jahre. Alle Konzepte basieren auf dem Grundgedanken der Berechnung von Prozessen auf Basis der jeweils gewonnenen und generierten Daten. Letztere sind gleichzeitig die Grundlage für die Neumodellierung und Re-Konfiguration der algorithmischen Systeme, die uns inzwischen ganz natürlich umgeben. Sei es das Empfehlungssystem beim Online-Shopping oder dem Musik- oder Videostreaming, das Bewertungssystem in Tausch- oder gar Partnerbörsen oder die algorithmisch kuratierten Nachrichten, wie sie uns personalisiert im Social Stream oder auch in den Suchergebnissen einzelner Suchmaschinen allgemein sichtbar gemacht werden. Wir befinden uns in einer von digitalen Technologien durchzogenen Welt, die sich in ihrer besonderen Komplexität auszeichnet und die Herstellung von individueller Orientierung damit nicht nur erleichtert. Es sind Algorithmen, die uns die Suchergebnisse aufzeigen und es uns im sprichwörtlichen Meer von Daten und Informationen überhaupt erst ermöglichen zu sehen und Neues zu erkunden. In dieser Hinsicht ist der Position des Kulturwissenschaftlers Stalder zuzustimmen, wenn er konstatiert, dass wir angesichts „der von Menschen und Maschinen

generierten riesigen Datenmengen“ ohne Algorithmen und deren Selektionsleistung blind wären.¹ Dementsprechend drängt sich die Frage nach der Abhängigkeit von Algorithmen hier ganz unweigerlich auf, denn sie steht im direkten Zusammenhang zu Fragen der Freiheit und Autonomie. Für die Fragen der Bildung ergeben sich hier zweierlei Konsequenzen: Erstens erlauben es diese Entwicklungen, den Bildungsbegriff selbst kritisch und unter den aktuellen Rahmenbedingungen auf ein Neues zu befragen. Zweitens halten algorithmische Systeme natürlich auch Einzug ins Bildungssystem und prägen nicht nur die Informationsbeschaffung, sondern auch die Wissensarbeit² in einer digitalen Welt und damit die Institutionen selbst³ ganz maßgeblich. Dies geschieht vorrangig über kommerzielle Angebote und die Aushandlungen über den spezifischen Einsatz digitaler Medien und digitaler Technologien im Bildungskontext. Hardware und Software lassen sich insbesondere im Kontext der Bildung nur schwer trennen, wie langjährige Debatten um die Ausstattung an deutschen Schulen oder Hochschulen verdeutlichen. Für die Fragen der Bildung ergibt sich hieraus eine komplizierte Verflechtung von teils privaten Anbietern und öffentlichen Bildungsinstitutionen einerseits und grundlegenden Fragestellungen darüber, was die jeweiligen Technologien Neues ermöglichen oder wie sie bisherige Probleme und Herausforderungen reproduzieren.

Der Beitrag nimmt sich dieser Problematik an, indem er das hier skizzierte Verhältnis sowie die Implikationen digitaler Technologien für die Fragen der Bildung diskutiert. Dies geschieht in drei Schritten. Zunächst wird grundlegend auf Big Data und die Rhetorik der Zahlen eingegangen, wie sie in unseren gesellschaftlichen Zusammenhängen beobachtet werden kann, um auf die Implikationen von datengetriebenen Verfahren hinzudeuten und so schließlich Bildung unter den Bedingungen des Digitalen befragen zu können. Die hieraus abgeleiteten Erkenntnisse, Konsequenzen und Schlüsse

¹ Stalder, F., *Kultur der Digitalität*, Berlin 2016, 13.

² Holze, J., *Digitales Wissen – Bildungsrelevante Relationen zwischen Strukturen digitaler Medien und Konzepten von Wissen*, <https://doi.org/10.25673/4666>.

³ Williamson, B., *Big data in education: The digital future of learning, policy and practice*, London 2017; Williamson, B., *Who owns educational theory? Big data, algorithms and the expert power of education data science*, *E-Learning and Digital Media*, 14(3) (2017), 105–122, <https://doi.org/10.1177/2042753017731238>.

werden dann zusammenfassend entlang des transformativen Potenzials von Selbst- und Weltverhältnissen vor dem Hintergrund von Bildungstechnologien (EdTech) kritisch diskutiert.

2. Big Data und die Rhetorik der Zahlen

Große Datenmengen, Big Data und die dazugehörigen mathematischen Modelle, wie sie über Algorithmen abgebildet werden, verdeutlichen, wie der starke Glaube an mathematische Grundlagen und naturwissenschaftliche Gesetzmäßigkeiten mehr und mehr Einzug in die Organisation des Sozialen hält. Big Data ist in diesem Zusammenhang zu einem populären Schlagwort geworden, welches hinter sich ebenfalls eine nicht unbeträchtliche Menge von Perspektiven, Interessenlagen und Sinnzuweisungen versammelt.⁴

Zahlen sind in Deutungshorizonte eingeflochten, die ganz grundlegend eine arithmetische Ordnung der Relationen in Raum und Zeit evozieren.⁵ Sie erlauben damit eine Komplexitätsreduktion und Abbildbarkeit von Wirklichkeit, die auch die Herstellung von Orientierung beeinflussen.

Die Repräsentation der Zahlen wird dabei von einer zumindest zweigeteilten Plausibilisierungslogik getragen, wie sie Duttweiler und Passoth entlang

⁴ Vgl. Manovich, L., Trending: The Promises and the Challenges of Big Social Data, in: Gold, M. K. (Hg.), *Debates in the Digital Humanities*, Minneapolis/MN 2011, 460–475; boyd, d./ Crawford, K., *Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon*. *Information, Communication & Society* 15(5) (2012), 662–679, <https://doi.org/10.1080/1369118X.2012.678878>; Geiselberger, H./ Moorstedt, T. (Hg.), *Big Data: Das neue Versprechen der Allwissenheit*, Berlin 2013; Gitelman, L. (Hg.), „Raw data“ is an oxymoron, Cambridge/MA 2013; Verständig, D., *Die Ordnung der Daten – Zum Verhältnis von Big Data und Bildung*, in: Iske, S./ Fromme, J./ Verständig, D./ Wilde, K. (Hg.), *Big Data, Datafizierung und digitale Artefakte* (Bd. 42), Wiesbaden 2020, 115–139, https://doi.org/10.1007/978-3-658-28398-8_7; Reichert, R. (Hg.), *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*, Bielefeld 2014.

⁵ Vgl. Rebiger, B., *Zahlensymbolik*, in: *Religion in Geschichte und Gegenwart* (Bd. 8), Tübingen, 1776–1777.

der Selbstvermessungsthematik pointiert auf den Punkt bringen:⁶ Zum ersten verweisen „Kurven, Statistiken, Tabellen oder Kuchendiagramme dezidiert auf Wissenschaftlichkeit“ und zum zweiten suggerieren „(stilisierte) Bilder und Grafiken [...] die vermeintlich unmittelbare Repräsentation der Wirklichkeit“, was eine kaum hinterfragte Evidenz hervorbringt.⁷ Das Vertrauen in die Daten ist schließlich eng mit dem Vertrauen in Zahlen verbunden und bringt zugleich auch alle Probleme mit, die seit jeher an die Logik der Zahlen gebunden sind. Es ist nicht nur die Blindheit gegenüber der Menge an Daten, sondern vielmehr die steigende Unsichtbarkeit der Genese der Zahlen mit ihrer Präsenz und ihren Repräsentationsmodi, die sich als eine fortwährende Herausforderung darstellt. Die Herausforderung liegt darin begründet, zu erkennen, dass allein die Entstehung oder Produktion von Daten keineswegs frei von Interpretationsleistungen sind. Ebenso wenig sprechen die Zahlen und Daten für sich selbst, denn sie sind letztlich ein Resultat verschiedener Prozesse, die sich selbst der Sichtbarkeit entziehen.

In der Konsequenz werden die technologischen Entwicklungen zum einen von Verschiebungen sozialer Strukturen und der Reproduktion sozialer Ungleichheiten begleitet.⁸ Zum anderen weisen etwaige Lösungsstrategien hierzu nur eine bestimmte und oftmals sehr geringe Halbwertszeit auf, da

⁶ Duttweiler, S./ Gugutzer, R./ Passoth, J.-H./ Strübing, J. (Hg.), Self-Tracking als Optimierungsprojekt?, in: *Leben nach Zahlen*, Bielefeld 2016, 9–42, <https://doi.org/10.14361/9783839431368-001>.

⁷ Ebd., 13.

⁸ Vgl. Iske, S./ Klein, A./ Kutscher, N., *Digitale Ungleichheit und formaler Bildungshintergrund: Ergebnisse einer empirischen Untersuchung über Nutzungsdifferenzen von Jugendlichen im Internet*. Kompetenzzentrum Informelle Bildung 2004, <http://www.kib-bielefeld.de/externelinks2005/digitaleungleichheit.pdf>; Zillien, N., *Digitale Ungleichheit: Neue Technologien und alte Ungleichheiten in der Informations- und Wissensgesellschaft*, Wiesbaden ²2009, <https://doi.org/10.1007/978-3-531-91493-0>; Hargittai, E., *Second-Level Digital Divide: Differences in People's Online Skills*, *First Monday*, 7(4) (2002), <https://doi.org/10.5210/fm.v7i4.942>; Hargittai, E., *Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites*, *The ANNALS of the American Academy of Political and Social Science* 659(1) (2015), 63–76, <https://doi.org/10.1177/0002716215570866>; O'Neil, C., *Weapons of math destruction: How big data increases inequality and threatens democracy*, New York 2016; Eubanks, V., *Automating inequality: How high-tech tools profile, police, and punish the poor*, New York 2017.

durch digitale Technologien eben nicht nur die Spielregeln geändert werden, es wird ein komplett neues Spielfeld eröffnet. Die hieraus entstehende Brisanz für die Fragen der Bildung ergibt sich aus einer gewissen Doppelläufigkeit, da bislang bekannte Probleme, wie Zugangsfragen zu digitalen Technologien und Informationen längst nicht ad acta gelegt werden können und gleichzeitig neue Probleme, wie die Auswirkungen der Datafizierung und die ökonomischen Tendenzen des Netzes weitere Problemfelder, wie jene der *ökonomischen*, aber auch *politischen* Überwachung deutlich machen. In ökonomischer Hinsicht sind es Unternehmen, die über personenbezogene Daten und Datenprofile ganze Geschäftsmodelle ausrichten und damit die Datenproduktion noch undurchsichtiger wird. Die politische Überwachung greift tief in gesellschaftliche Prozesse ein und verweist auf eine komplizierte Auseinandersetzung,⁹ die als Grundlage für das Verständnis um eine Überbewertung der Zahlen dienen kann. Es sollen damit die Konsequenzen herausgestellt werden, wenn jene Prinzipien der Datenproduktion, -evaluation und -repräsentation durch Verknüpfungen in den Bildungsbereich einziehen.

Um die Rhetorik der Zahlen genauer in den Blick nehmen zu können, scheint es sinnvoll, zumindest die jüngere Entwicklungsgeschichte des Computers zu berücksichtigen. Natürlich könnte eine historische Betrachtung der Zahlen und mathematischen Auseinandersetzungen wesentlich weitergehen, jedoch verdeutlicht die Entwicklung der universalen Rechenmaschinen einen entscheidenden Punkt. Es geht weniger um die Technologie selbst, sondern vielmehr um die sie begleitenden Argumentationsmuster, denn Befürworter von Computern haben weitreichende Behauptungen über ihre inhärente transformative Kraft aufgestellt: neu und anders als frühere Technologien, waren sie sicher, dass viele unserer bestehenden sozialen Probleme berechnet und durch eine Risikominimierung gar gelöst werden können. Damit wäre dann nicht nur die Technologie von den positiven Einwirkungen

⁹ Vgl. hierzu Gaycken, S. (Hg.), *Jenseits von 1984: Datenschutz und Überwachung in der fortgeschrittenen Informationsgesellschaft: eine Versachlichung*, Bielefeld 2013; Gaycken, S./ Kurz, C., 1984.exe. *Gesellschaftliche, politische und juristische Aspekte moderner Überwachungstechnologien*, Bielefeld 2015, <https://nbn-resolving.org/urn:nbn:de:101:1-201512039542>; Bauman, Z./ Lyon, D., *Daten, Drohnen, Disziplin: Ein Gespräch über flüchtige Überwachung*, Berlin 2013.

des Rechners betroffen, sondern auch die Politik und schließlich die Gesellschaft als Ganzes. Diese Argumentationen waren begleitet von einem strengen Glauben an die Zahlen und Rechner. In „The Cultural Logic of Computation“¹⁰ stellt David Golumbia diese Orthodoxie in Frage und argumentiert stattdessen, dass Computer „durch und durch“ kulturell sind – dass es keinen Teil der offensichtlichen technologischen Transformation gibt, der nicht von historischen und kulturellen Prozessen geprägt sei oder der sich der bestehenden Kulturpolitik entziehe. Aus der Perspektive transnationaler Konzerne und Regierungen dienen Computer und rechenbasierte Architekturen der Stabilisierung von bestehenden Machtverhältnissen viel mehr, als dass sie Mittel bereitstellen, diese zu verschieben oder gar anzufechten. Trotzdem hat sich unser Denken über Computer zu einer fast unsichtbaren Ideologie entwickelt, die Golumbia als „Computationalism“ bezeichnet – eine Ideologie, die unser Denken nicht nur über Computer, sondern auch über so weitreichende ökonomische und soziale Trends wie die Globalisierung prägt. Eine Bestätigung erfährt die Diagnose durch die zuvor angerissene Problematik um die ökonomische, aber auch politische Überwachung des Einzelnen und gesellschaftlicher Strukturen.

Golumbias Kernargument ist, dass die zeitgenössische westliche Kultur zu tiefst durch Formen der Hierarchie und Kontrolle strukturiert ist, die ihren Ursprung in der Entwicklung und Nutzung von Computern in den letzten 50 Jahren haben. Die leitende These ist also nicht, dass unsere Gesellschaft und die Menschen dabei sind, sich durch Computer grundlegend zu verändern. Vielmehr ist Veränderung ein dem Menschen inhärentes Phänomen, zu dem nun rechenbasierte Technologien zählen. Es geht dann also nicht darum, dass der Mensch sich durch technologische Entwicklungen und „Enhancements“ zu einem Wesen entwickelt, welches eine kategorische Differenz zu sich selbst aufweist, wie er in Anlehnung an Hayles Begriff des „Posthumanen“¹¹ argumentiert. Veränderungen, die so grundlegend sind, dass sie das Menschsein und die Teilhabe an einer Gesellschaft völlig neu definieren, müssen daher entweder dieselbe Art von Veränderung sein, der wir immer

¹⁰ Golumbia, D., *The cultural logic of computation*, Cambridge/MA 2009.

¹¹ Hayles, K., *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*, Chicago 1999.

unterworfen sind, oder eine Art von Transzendenz der menschlichen Lebenswelt, die wir wohl nicht begreifen können, weil sie per Definition jenseits unseres menschlichen Verständnisses liegt.

Damit wird das kulturelle Begriffsverständnis dieser Veränderungsdynamiken betont, welches sich auch bei Stalders Kultur der Digitalität¹² wiederfindet. Auch Stalder argumentiert entgegen einer radikalen und abrupten Veränderung, die nun in Form der digitalen Technologien und Medien in Erscheinung tritt. Vielmehr sei das Leben unter den Bedingungen des Digitalen von einer langen Entwicklung und damit verbundenen Ausdifferenzierung und Hybridität geprägt, wie er unter anderem an der Transformation der LGBT-Bewegung verdeutlicht:

„Wenn die Schwulenbewegung exemplarisch für die gesellschaftliche Liberalisierung der siebziger und achtziger Jahre steht, dann kann man ihre Transformation in die LGBT-Bewegung seit den neunziger Jahren – mit ihrer enormen Vielfältigkeit und Verflüssigung von Identitätsmodellen, der Betonung von Wandelbarkeit und Hybridität – als Zeichen für die Neuerfindung dieses Projekts im Kontext der dominant werdenden Kultur der Digitalität sehen.“¹³

Der hier beschriebene Wandel lässt sich abermals modernisierungstheoretisch aufladen, indem es hierbei nicht nur um eine „Diversifizierung und Verflüssigung kultureller Praktiken“ und sozialer Rollen geht, die damit längst nicht am Ende sind, sondern hier ein Bruch mit traditionellen Strukturen gewissermaßen als Aufforderung zu Selbsterprobung auferlegt und zur alltäglichen Aufgabe oder Normalität wird.

„Die zunächst subkulturell und unter existentiellstem Druck erprobten Lebensweisen werden allmählich zum Mainstream. Sie erweitern die Spannweite der relativ leicht verfügbaren Identitätsmodelle für alle, die irgendwie daran interessiert sind, sei es in Bezug auf Familienformen (etwa Patchworkfamilien, Adoption durch gleichgeschlechtliche Paare), Ernährungsweisen (etwa Vegetarismus und Veganismus), Gesundheits-

¹² Stalder, Kultur der Digitalität (Fußnote 1).

¹³ Ebd., 48f.

praktiken (etwa Ablehnung von Impfungen) oder andere Lebens- und Glaubensgrundsätze.“¹⁴

All diese verschiedenen Interessengruppen und Akteure treten mit dem Anspruch auf, ihre Positionen in der Öffentlichkeit zu artikulieren und sich somit eine Legitimation sozialer Bedeutung und Sinnzuweisung zu erwirken. In der Konsequenz treffen hier verschiedene nicht selten konträre Positionen aufeinander, was zu komplexen Aushandlungsprozessen führt, die – wenn sie in digitalen Öffentlichkeiten stattfinden – ganz grundlegend von den Metriken zur Sichtbarkeitssteigerung und Affirmation geprägt sind. Es zählt dann, wie vielen Menschen die eigene Position gefällt und wie viele Menschen die jeweiligen Positionen gesehen, geklickt und kommentiert haben. Das Argument als solches wird dadurch nicht zwingend, aber potenziell in den Hintergrund gerückt, da die Metriken analog zur Selbstvermessungsproblematik keineswegs zwingend eine qualitative Aussagekraft haben müssen. Zahlen werden dabei durch ihre Repräsentation und Einbettung in komplexe rechenbasierte Systeme zu einer eigenen Sprach- und Diskursform mit spezifischen Grammatiken und Semantiken des Digitalen. Hieraus ergibt sich eine komplexe Verflechtung sozialer und kultureller Praktiken mit digitalen Medien und den dahinterliegenden Technologien.¹⁵ Die Komplexität ergibt sich aus vielfältigen Datenverknüpfungen und den unterschiedlichen Geschäftsmodellen, die in ihrer Wirkweise oftmals wenig transparent gemacht und vor der Konkurrenz am Markt gut geschützt werden. Die Positionierung in digitalen Öffentlichkeiten und die Generierung von Sichtbarkeit wird somit mehr und mehr zu einem unbestimmten Spiel mit den algorithmischen Strukturen, die diese Räume erst aufspannen.

Dieses Spannungsverhältnis von Komplexität und Unsichtbarkeit der algorithmischen Strukturen sowie der digitalen Daten, die bewusst und unbewusst produziert werden, stellt eines der Grundprobleme für die Bildung

¹⁴ Ebd., 49.

¹⁵ Vgl. Verständig, D., Soziale Medien zwischen Disruption und Synthese – Eine bildungstheoretische Perspektive auf Praktiken des Codings zur Herstellung von digitalen Öffentlichkeiten, in: Klimczak, P./ Petersen, C./ Breidenbach, S. (Hg.), Die Gesellschaft im Spiegellabyrinth sozialer Medien. Kultur-, Sozial- und Computerwissenschaftliche Zugänge zur Onlinekommunikation, Wiesbaden 2020, 26–43, https://doi.org/10.1007/978-3-658-30702-8_2.

dar, wie es sich ähnlicher Form schon lange zeigt. Es handelt sich um die Problematik zur Herstellung von Orientierung und damit der individuellen Entwicklung des Menschen hinsichtlich seiner Persönlichkeit zu einem Menschsein, geprägt von kulturellen, geistigen und sozialen Merkmalen. Ausgehend von einschlägigen Diagnosen der Modernisierungstheorie¹⁶ zeigt sich heute, dass sich diese Problematik, um die Herstellung von Orientierung und dem Menschsein vor dem Hintergrund digitaler Technologien in mehrfacher Hinsicht verschärft hat und damit danach fragen lässt, wie Bildung in der digitalen Welt überhaupt gedacht werden kann.

3. Welche Bildung in der digitalen Welt?

Bildung wird als eine der zentralen Ressourcen gesehen, um in einer komplexen Welt zu bestehen, sich persönlich zu verwirklichen und in gesellschaftliche Prozesse einbringen zu können. Wenn Bildung der Schlüssel zum Erfolg für zukünftige Handlungs- und Gestaltungsfähigkeit in einer von digitalen Technologien geprägten Welt ist, dann stellt sich zunächst die Frage, was die Anforderungen an Bildung unter diesen Bedingungen überhaupt sein können. Über welche Bildung wird dann überhaupt gesprochen? In einem bildungstheoretischen Verständnis wird Bildung immer dann relevant, wenn sich bestehende Muster der Weltaufordnung im Sinne von Selbstverständlichkeiten auflösen und Menschen mit Unsicherheiten konfrontiert werden. Anders formuliert:

„Bildung bringt Unbestimmtheitsdimensionen zur Geltung; das macht gerade den offenen, experimentellen und suchenden Charakter aus. Intakte Routinen der Selbst- und Weltauslegung werden gerade in Bildungsprozessen außer Kraft gesetzt; sie werden würdig, befragt zu werden, also fragwürdig.“¹⁷

¹⁶ Vgl. Beck, U., Risikogesellschaft: Auf dem Weg in eine andere Moderne, Berlin 1986; Giddens, A., Konsequenzen der Moderne, Berlin 1996; Sennett, R., Der flexible Mensch: Die Kultur des neuen Kapitalismus, München ³2000.

¹⁷ Marotzki, W., Entwurf einer strukturalen Bildungstheorie: Biographietheoretische Auslegung von Bildungsprozessen in hochkomplexen Gesellschaften, Weinheim 1990, 153.

Alltägliche Handlungsrountinen sind notwendig, um ohne große Reflexions-schleifen agieren zu können und handlungsfähig zu sein. Digitale Technolo-gien bestärken uns in diesen Handlungsrountinen, indem sie durch Assistenz-, Anreiz- oder Empfehlungssysteme Gewohnheiten erfassen und uns in diesen Gewohnheiten bestärken oder durch Nudging in andere Richtungen lenken. Bildung hat so gesehen einen Bezug zur Autonomie des Individuums, da die Infragestellung und erneute Befähigung einen Moment des Kontrollverlusts sowie der Wiederherstellung von Orientierungsrahmen umfasst. Die Auseinandersetzung mit dem Bildungsbegriff ist unter dem Vorzeichen der digita-len Technologien insofern von gesteigerter Bedeutung, als sich die Diskussion über die Digitalisierung in der Bildung ganz zentral mit der Frage beschäftigt, was für die Zukunft gelehrt und gelernt werden soll. Damit ist die Auseinandersetzung um die Fragen der Bildung einerseits mit der Unsicherheit von Zukunftsentwürfen konfrontiert, andererseits eine Auffor-derung zur Mitgestaltung des digitalen Wandels. Auf der Basis gesellschaft-licher Verständigungsprozesse weisen Kompetenzlisten aus, was für die Zukunft wichtig ist und gelernt werden sollte. Sie sind vielfach so angelegt, dass sie ihre Zukunftsentwürfe verstecken, und so der weiteren Verständi-gung entziehen.

Bildung in der digitalen Welt, so lautet die Strategie der Kultusminister-konferenz, wie sie 2016 verabschiedet wurde. Das vorgestellte Handlungs-konzept bietet einen ersten und vieldiskutierten Ansatzpunkt,¹⁸ um Bildung

¹⁸ Vgl. Asmussen, M./ Schröder, C./ Hardell, S., Bildung in politischen Programmen – Eine pädagogische Revision der KMK-Strategie zur Bildung in der digitalen Welt, in: Leineweber, Ch./de Witt, C. (Hg.), Digitale Transformation im Diskurs. Kritische Perspektiven auf Entwicklungen und Tendenzen im Zeitalter des Digitalen, deposit_hagen, Publikationsserver der Universitätsbibliothek 2017, <https://medien-im-diskurs.de>; Dander, V., Ideologische Aspekte von „Digitalisierung“, in: Leineweber, Ch./de Witt, C. (Hg.), Digitale Transformation im Diskurs, deposit_hagen, Publikationsserver der Universitätsbibliothek 2018, 15, <https://medien-im-diskurs.de>; Schulmeister, R./ Loviscach, J., Mythen der Digitalisierung mit Blick auf Studium und Lernen, in: Leineweber, Ch./de Witt, C. (Hg.), Digitale Transformation im Diskurs, deposit_hagen, Publikationsserver der Universitätsbibliothek 2017, 12, <https://medien-im-diskurs.de>; Kerres, M., Bildung in der digitalen Welt: Wir haben die Wahl, denk-doch-mal.de Online-Magazin für Arbeit-Bildung-Gesellschaft 02–18 (2018), <http://denk-doch-mal.de/wp/michael-kerres-bildung-in-der-digitalen-welt-wir-haben-die-wahl/>;

im Hinblick auf den digitalen Wandel zu konturieren. Das Konzept sah vor, dass jede Schülerin und jeder Schüler bis 2021 eine digitale Lernumgebung und einen Zugang zum Internet nutzen können soll. Die übergeordnete Zielstellung dabei war es, die Rahmenbedingungen dafür zu schaffen, dass die Heranwachsenden zu einem „selbstständigen und mündigen Leben in der digitalen Welt“ befähigt werden. Diese Zielstellung ist in vielerlei Hinsicht sehr voraussetzungsvoll, denn sie ist nicht nur darauf beschränkt, eine gute digitale Infrastruktur und damit Zugang zu Informationen und Wissensbeständen zu ermöglichen. Vielmehr gilt es bei der Gestaltung der Rahmenbedingungen zur gesellschaftlichen Teilhabe für Heranwachsende auch die Lehrkräfte systematisch mit einzubeziehen. Schließlich sollten die Lehrkräfte zur Umsetzung der Strategie als grundlegende Voraussetzung selbst über einen sicheren Umgang mit den Medien verfügen, um diese professionell und didaktisch sinnvoll in ihrem jeweiligen Fachunterricht nutzen zu können. Hier muss der Spagat zwischen dem innovativen Neuen, also neuen Einsatzmöglichkeiten, die durch den Einsatz digitaler Medien hervorgerufen werden und dem Bewährten also den grundständigen Aufgaben und Anforderungen geschafft werden. Für Lehrkräfte heißt das in aller Regel, dass dem digitalen Wandel über Lehreraus-, Fort- und Weiterbildung Rechnung getragen werden muss. Damit wird eine Diskrepanz der letzten Jahre deutlich, die man nur schwer übersehen kann, denn die digitalen Medien und digitalen Technologien, wie wir sie heute vorfinden, sind eben kein Phänomen der letzten drei Jahre, sondern das sich stetig wandelnde Produkt unserer sozialen, kulturellen und politischen Aushandlungsprozesse. Bislang bleibt offen, wie sich ein digitaler Wandel nachhaltig mitgestalten lässt und

Kerres, M., Bildung in der digitalen Welt: Über Wirkungsannahmen und die soziale Konstruktion des Digitalen, *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 17 (2020), 1–32, <https://doi.org/10.21240/mpaed/jb17/2020.04.24.X>; Dander, V./ Bettinger, P./ Ferraro, E./ Leineweber, C./ Rummler, K./ Budrich, B., *Digitalisierung – Subjekt – Bildung: Kritische Betrachtungen der digitalen Transformation*, Leverkusen 2020; Herzig, B./ Klar, T.-M./ Martin, A./ Meister, D. M., Editorial: Orientierungen in der digitalen Welt, *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 39 (2020), i–vii, <https://doi.org/10.21240/mpaed/39/2020.12.12.X>.

wie Bildung im Horizont einer Kultur der Digitalität¹⁹ nicht nur konzipiert, sondern auch gelebt werden kann. Entscheidend dafür scheint die Reflexion über die bestehenden wirtschaftlichen Verwertungslogiken sowie die administrativ angelegten Optimierungs- und Standardisierungstendenzen²⁰ und deren Aktualisierung.

Der digitale Wandel ist auch eine der Prioritäten der Europäischen Kommission, die sie in ihrer Leitinitiative „Ein Europa für das digitale Zeitalter“ zum Ausdruck gebracht hat. Im aktualisierten Aktionsplan für digitale Bildung (2021–2027) greift die Kommission die aus der Pandemie um COVID-19 gewonnenen Erkenntnisse auf und entwirft eine Langzeitvision für den digitalen Wandel in der allgemeinen und beruflichen Bildung in der EU. Der Aktionsplan sieht zwei entscheidende strategische Ausrichtungen vor. Einerseits soll die Implementation und damit die infrastrukturelle Förderung von digitalen Technologien im Bildungssektor vorangetrieben werden. Es geht um „die Förderung der Entwicklung eines leistungsfähigen digitalen Bildungsökosystems“. Andererseits sieht der Aktionsplan den Ausbau digitaler Kompetenzen und Fertigkeiten für den digitalen Wandel vor. Die Ziele dieser langfristig angelegten Priorisierung sind unter anderem die Verbesserung der digitalen Infrastruktur im Bildungsbereich, eine höhere Qualität der Lernressourcen beim digitalen Lernen, eine Steigerung der Effizienz der Lernprozesse sowie die bessere Nachweisbarkeit von Lernergebnissen.

Auffallend dabei ist zunächst die Differenzierung von Lernen und Bildung. Es geht hierbei um das Erlernen von gewissen qualifikatorischen Faktoren und weniger um Bildung im engeren Sinne. Diese Ausrichtung ist getragen von einer doch sehr ergebnisorientierten Perspektive auf den Lernerfolg bzw. die Lernergebnisse, also den Output und die produktive Wendung institutioneller Bildung. Ein „leistungsfähiges digitales Bildungsökosystem“ sieht damit die effiziente und höchstmögliche freie Erarbeitung von Zertifikaten und Bildungsabschlüssen vor, die dann auch den Mobilitätsaspekt – zumindest innerhalb der EU – adressieren. Durch Standards werden diese qualifikatorischen Ziele erreichbar, beispielsweise über das European Credit

¹⁹ Stalder, Kultur der Digitalität (Fußnote 1).

²⁰ Dander, Ideologische Aspekte von „Digitalisierung“ (Fußnote 18).

Transfer System (ECTS), ein Instrument, das der Gliederung des Hochschulstudiums dient und die Gewichtung seiner Bestandteile transparent macht und somit nicht nur der Qualitätssicherung der Hochschullehre dient, sondern gleichzeitig dazu befähigen kann an unterschiedlichen Standorten zu studieren. Die Studierenden sammeln so Leistungspunkte, bis die geforderte Gesamtpunktzahl ihres Studiengangs erreicht ist und das Studium dann als erfolgreich gelten kann. Auch hier liegt die Rhetorik der Zahlen fest in den Prozess der Standardisierung eingeschrieben, da die bearbeiteten Themenfelder im Studienverlauf in Zahlen ausgedrückt werden, was zugleich eine gewisse Vergleichbarkeit unterschiedlicher Studienprogramme nahelegt, in der Tat jedoch nicht viel über die kompetitiven Zusammenhänge hinsichtlich des Erlangens von Studienleistungen aussagt.

Interessant ist zudem die noch immer nicht aufgelöste Zugangsproblematik zu digitalen Technologien. So wird der Aktionsplan unter anderem mit einer Studie des statischen Amtes der Europäischen Union (Eurostat) begründet, die zu dem Ergebnis kommt, dass viele Haushalte mit niedrigem Einkommen keinen Zugang zu Computern und Breitbandinternet haben, bzw. dass der Zugang in der EU abhängig vom Haushaltseinkommen sehr unterschiedlich ist.²¹ Zudem wird der Aktionsplan damit begründet, dass mehr als ein Fünftel der jungen Menschen in Europa nicht über grundlegende digitale Kompetenzen verfügen. Die Kommission hat sich dahingehend unter anderem auferlegt, die Teilnahme an der internationalen Studie zur Messung der Computer- und Informationskompetenzen (ICILS)²² zu fördern, in der länderübergreifende Daten über digitale Kompetenzen von Lehrkräften gesammelt werden. Daneben soll die Einführung eines EU-Ziels für die digitale Kompetenz von Schülerinnen und Schülern, um den Anteil der 13- und 14-jährigen Lernenden, die in den Bereichen Computer- und Informationskom-

²¹ Eurostat, Survey on ICT usage in households and by individuals (2019), https://ec.europa.eu/eurostat/statistics-explained/index.php/Digital_economy_and_society_statistics_-_households_and_individuals.

²² https://ec.europa.eu/education/resources-and-tools/document-library/the-2018-international-computer-and-information-literacy-study-icils-main-findings-and-implications-for-education-policies-in-europe_en.

petenz unterdurchschnittlich abschneiden, bis 2030 auf unter 15 % zu senken. Auch hier liegt die Rhetorik der Zahlen bereits in den Handlungsvollzügen tief eingeschrieben.

Angesichts der hohen Komplexität, der Ungewissheit zukünftiger Entwicklungen rund um digitale Technologien, also der hohen Veränderungsdynamik, muten derartige langfristig angelegt Pläne, die eine klare Zielvorstellung im Umgang mit Computern und Informationstechnologien haben, fernab einer realisierbaren Umsetzung. Dennoch wird die Notwendigkeit für die Kompetenz im Umgang mit digitalen Technologien nicht nur als wünschenswerte Zielvorstellung, sondern ebenso als notwendige Aufgabe beschrieben:

„Today more than ever, **being digitally competent is both a necessity and a right**. Participating actively, continuously and responsibly in society at all levels (political, economic, social, cultural and intercultural) means being able to harness the benefits and opportunities of the online world, while building resilience to potential risks.”²³

Die dahinterliegende Idee ist auch hier klar. Es geht darum, Europa zukünftig wettbewerbsfähig zu gestalten, denn einschlägige technologische Entwicklungen wie Künstliche Intelligenz, Machine Learning oder Big Data verändern die Qualifikationsanforderungen des Arbeitsmarktes und wirken sich wiederum auf die Kompetenzentwicklung für die Wirtschaft aus. Daher gilt es diese Zusammenhänge kritisch zu betrachten und das Verständnis über diese Strukturen zu schärfen. Jede/r Einzelne sollte verstehen, wie digitale Technologien Kommunikation, Kreativität und Innovation unterstützen können, und sich ihrer Chancen, Grenzen, Auswirkungen und Risiken bewusst sein. Das ist die Maßgabe des Aktionsplans für digitale Bildung (2021–2027).

Die Aushandlungen darüber, welche Kompetenzen und Fertigkeiten tatsächlich benötigt werden finden allerdings gerade noch statt, wie bei einem

²³ Europäische Kommission, Digital Education Action Plan (2021–2027). Resetting education and training for the digital age – Commission Staff Working Document 2020, https://ec.europa.eu/education/sites/default/files/document-library-docs/deap-swd-sept2020_en.pdf, 39, Hervorhebung im Original.

genaueren Blick in das Digital Competence Framework 2.0 (DigComp) deutlich wird. Dort werden ebenfalls grundlegende Kompetenzen für die digitale Welt entlang fünf verschiedener Felder identifiziert, die sich von Informations- und Datenkompetenz über Kommunikation und Zusammenarbeit, sowie der Produktion digitaler Inhalte über Fragen der Sicherheit hin zu Problemlösungskompetenzen erstrecken. Es ist kein Zufall, dass die Problemlösung als finaler Punkt in diesem Modell angemerkt wird. Ziel dieses Kompetenzrahmens ist auch hier die kritische Befähigung im Umgang mit digitalen Technologien, die sich im Kern auf die Urteilskraft und ein Differenzbewusstsein bezieht. Beispielsweise wird in Punkt 5.4 die Identifikation von Kompetenzlücken sowie eine offene, unterstützende Haltung und die Bereitschaft zur Vermittlung von Kenntnissen adressiert:

„To understand where one’s own digital competence needs to be improved or updated. To be able to support others with their digital competence development. To seek opportunities for self-development and to keep up-to-date with the digital evolution.”²⁴

Allen Positionspapieren und Konzepten ist die Notwendigkeit zum kreativen und kritischen Umgang mit digitalen Technologien gemein. Hier werden alle Papiere doch sehr konkret. Gleichzeitig ist der Weg dorthin mal mehr mal weniger diffus dargestellt. Dadurch werden zwei grundlegende Probleme des aktuellen Bildungsverständnisses offengelegt. Erstens werden Bildungsprozesse im Sinne der Menschwerdung hinter den jeweiligen Zielvorstellungen zumindest terminologisch verortet. Zweitens zielen die Strategien im Kern auf genau jene Problemstellung ab.

In dem Zusammenhang ist Bildung immer in zweierlei Richtungen zu denken: Einerseits auf das Individuum und die dem Menschen zu Grunde liegenden Potenziale und andererseits auf gesellschaftliche Rahmenbedingungen, also überindividuelle Aushandlungsprozesse. Während Letzteres auf überindividuelle machttheoretische Implikationen deutet, wie sie beispielsweise

²⁴ <https://ec.europa.eu/jrc/en/digcomp/digital-competence-framework>, Zugriff 10. April 2021.

durch Wimmer sowie Benner adressiert²⁵ und später durch Ricken diskutiert wurden,²⁶ bezieht sich zumindest die erste Denkrichtung des Menschseins in der Tradition des humboldtschen Bildungsgedankens und dem Streben nach der allgemeinen Bildung auf die individuelle Verortung des Menschen in der Welt. Damit lässt sich gerade in den hier genannten jüngeren Arbeiten eine Diskursformation ablesen, die einen bildungstheoretischen Komplexitätszuwachs verdeutlicht und zugleich die Befragung von Bildung unter zeitgenössischen gesellschaftlichen Rahmenbedingungen ermöglicht. Allgemeine Bildung ist damit zur Grundlage für das Menschsein und die Menschwerdung geworden. Die vielzitierte Aussage von Humboldt, die gleichzeitig Anstoß zur Kontroversen Auseinandersetzung gegeben hat und vor dem Hintergrund einer Subjektconstitution immer noch gibt, ist folgende:

„Der wahre Zweck des Menschen – nicht der, welchen die wechselnde Neigung, sondern welche die ewig unveränderliche Vernunft ihm vorschreibt – ist die höchste und proportionirlichste Bildung seiner Kräfte zu einem Ganzen.“²⁷

Dörpinghaus hält hierzu fest, dass Humboldt sich mit der Freiheit als erster Voraussetzung einer allgemeinen Bildung, die nicht revolutionär erzwungen oder staatlich befohlen werden kann, zur Idee einer *Humanität* als Allgemeinbildung bekenne.²⁸ Sie ist nicht auf eine spezifische Zielsetzung oder Tätigkeit hin ausgerichtet und löst sich von äußeren Zwecksetzungen sowie gesellschaftlichen Anforderungen. Was bleibt, ist ganz im Sinne Kants, der

²⁵ Wimmer, M., Die Gabe der Bildung. Überlegungen zum Verhältnis von Singularität und Gerechtigkeit im Bildungsgedanken, in: Masschelein, J./ Wimmer, M. (Hg.), *Alterität, Pluralität, Gerechtigkeit. Randgänge der Pädagogik*, Leuven 1996, 127–162; Benner, D., „Der Andere“ und „Das Andere“ als Problem und Aufgabe von Erziehung und Bildung, *Zeitschrift für Pädagogik* 45 (1999), 315–327.

²⁶ Ricken, N., *Die Ordnung der Bildung: Beiträge zu einer Genealogie der Bildung*, Wiesbaden 2006, 26.

²⁷ Humboldt, W. von, *Ideen zu einem Versuch, die Grenzen der Wirksamkeit des Staats zu bestimmen*, Breslau 1851, 9, <https://books.google.de/books?id=eavLesjp5xYC>.

²⁸ Dörpinghaus, A., *Die Grammatik der Allgemeinbildung*, in: Böttcher, W./ Heinemann, U./ Priebe, B. (Hg.), *Allgemeinbildung im Diskurs. Plädoyer für eine Kernaufgabe der Schule*, Hannover 2019, 35–53.

Verweis auf den Menschen selbst. Bildung wird hierbei nicht als die spezifische Ausbildung oder die Spezialisierung in Form einer differenzierten Expertenhaltung gesehen, sondern vielmehr in der Ausgewogenheit und damit der allgemeinen Form als erstrebenswert verstanden. Erst in der Wechselwirkung mit der Vielfalt der Welt, die in ihrem Reichtum mit der Pluralität der menschlichen Möglichkeiten korrespondiert, finden die im Menschen liegenden Möglichkeiten ihre Aktualisierung.

Pluralität und Vielfalt sind zwar Aspekte, die den hier dargestellten Konzepten zur Bildung im Horizont der Digitalität zumindest implizit innewohnen, jedoch ist die Potenzialentfaltung kritisch zu hinterfragen, wenn auf Basis von algorithmischen Empfehlungen, Nutzertracking und Überwachung das Lernverhalten analysiert und dann optimiert wird. Denn Optimierung steht der Irritation entgegen. Optimierung heißt zwangsläufig Standardisierung. Diese Standardisierung wird bereits über Kompetenzorientierung abgebildet und umgesetzt. Eine über Kompetenzen legitimierte Logik durchzieht sämtliche Debatten über Bildung, beginnend von der europäischen Perspektive auf digitale Bildung, hin zur Perspektive des Bundes und der Länder über regionale Bildungseinrichtungen, deren Aktivitäten und Handlungsspielräume darüber definiert und nicht zuletzt minimiert werden. Nun kann man unterschiedlicher Auffassung hinsichtlich der Intentionalität dieser Entwicklungen sein, deutlich wird jedoch die Forderung nach dem Streben von Autonomie und Freiheit, und zwar insofern, als man als mündiges Mitglied der Gesellschaft zu begründeten Entscheidungen auf Basis validierbarer Informationen beziehungsweise Daten kommen kann.

4. Auf dem Weg zur Mündigkeit

Der Weg zur Mündigkeit verweist insofern auf die Grundverfasstheit von Bildung, als hier der geschichtsträchtige Kampf des Menschen auf dem Weg zu sich selbst thematisiert wird. Nach Heydorn erfolgt der Weg zur Mündigkeit über Erziehung:

„Erziehung ist das uralte Geschäft des Menschen, Vorbereitung auf das, was die Gesellschaft für ihn bestimmt hat, fensterloser Gang [...] Die Geschichte der Erziehung hält an, die Gesellschaft hat ihren vorrationalen Charakter nicht überwunden, sondern ihn nur den veränderten Bedingungen gemäß modifiziert. Mit der Erziehung geht der Mensch seinen Weg durch das Zuchthaus der Geschichte.“²⁹

Erziehung geht der Bildung damit voraus und ist demnach die Bedingung für Bildung. Gleichwohl entfaltet Bildung dann ihre eigene Qualität, wenn sie ihrer selbst habhaft wird. Sie erlangt dann ihren eigenen gesellschaftlichen Antrieb, in der Umkehrung der gesellschaftlichen Verhältnisse.³⁰ Hierin liegt das Freiheitspotenzial des Bildungsversprechens, wie es Heydorn formuliert³¹ und von Michel-Schertges diskutiert wird.³² Der Weg zur Mündigkeit kann durch Technologie nicht verkürzt werden. Dennoch sind es aktuell verschiedene Bildungstechnologien (EdTech), die sich genau dieses Versprechen der Bildung zu eigen machen und damit zunächst pragmatische Lösungsvorschläge für die komplizierte Verfasstheit von Bildungsinfrastrukturen anbieten.

Es ist recht deutlich, dass ein solches Verständnis von Mündigkeit, welches schließlich Bildung ermöglicht, entgegen einer strategischen Ausrichtung und globalen Wettbewerbsfähigkeit steht, stattdessen bestehende soziale und technologische Differenzen tendenziell verstärkt. Die Pandemie um COVID-19 hat sowohl im Schul- als auch Hochschulkontext beispielhaft verdeutlicht wie agil sich Unternehmen entgegen der bisher staatlichen Bestrebungen positionieren können. Der Weg zur Mündigkeit ist angesichts dieser Tendenzen über den Diskurs zu Datenpraktiken und den Konsequenzen für

²⁹ Heydorn, H.-J., *Über den Widerspruch von Bildung und Herrschaft*, Bildungstheoretische Schriften Band 2, Frankfurt a. M. 1979, 9.

³⁰ Vgl. Michel-Schertges, D., *Sprache und Halbbildung: Zur gegenwärtigen Lage wissenschaftlicher Produktion*, in: Stederoth, D./ Novkovic, D./ Thole, W. (Hg.), *Die Befähigung des Menschen zum Menschen*, Wiesbaden 2020, 219–231, hier 220, https://doi.org/10.1007/978-3-658-24215-2_13.

³¹ Heydorn, H.-J., *Zu einer Neufassung des Bildungsbegriffs*, Berlin 1972.

³² Michel-Schertges, *Sprache und Halbbildung* (Fußnote 30).

jedes einzelne Gesellschaftsmitglied zu führen. Digitale Mündigkeit kann jedoch nicht bedeuten, dass man sich den digitalen Technologien entzieht. Gleichwohl wird digitale Mündigkeit durch die undurchsichtigen Strukturen der algorithmischen Systeme zweifelsohne ganz grundlegend zur Disposition gestellt. Wenn Mündigkeit über den Prozess der Erziehung und damit auch über die Bewusstwerdung der Umwelt verortet ist, dann sind Praktiken der Datensammlungen und Auswertung sowie die ökonomischen Analysen von großen Datensätzen, die den individuellen, aber auch überindividuellen lebensweltlichen Alltag beeinflussen – sich dennoch der Sichtbarkeit entziehen – ein Schlüsselproblem unserer Zeit, das es als Querschnittsproblematik der Digitalisierung zu bearbeiten gilt. Dabei handelt es sich um einen fortwährenden Prozess, für den ein Aushandlungsraum geschaffen und offengehalten werden muss. Sei es über die Förderung von Kompetenzen im Umgang mit digitalen Technologien, die stetige Reflexion von privatwirtschaftlichen Einflüssen in staatliche Bildungseinrichtungen³³ oder eben das Wecken der Neugier einzelner Menschen. Entgegen der Verwertungslogik unserer Zeit erscheint das folgende Postulat zur Aufgabe von Schule von Heydorn aktueller denn je:

„Unter der gegenwärtigen Bedingung wird es darauf ankommen, den Menschen für seine Universalität auszubilden, die in der Tendenz bereits angelegt ist. Das heißt Ausbildung aller seiner Organe als Ausbildung seiner Selbsterfassung.“³⁴

Nun könnte man argumentieren, dass es doch keineswegs problematisch sei, würden digitale Technologien diesen Prozess unterstützen, jedoch ergibt sich aus der Beschaffenheit der Systeme selbst eine nicht auflösbare Problemlage, dass diese Systeme oftmals einer einfachen behavioristischen Logik folgen und damit zeitgemäße Formen des Lernens gar nicht erst ermögli-

³³ Hartong, S., Algorithmisierung von Bildung. Über schrumpfende Spielräume für demokratisches (Ver-)Handeln und warum die EdTech-Industrie nicht das einzige Problem ist, denk-doch-mal.de Online-Magazin für Arbeit-Bildung-Gesellschaft (2020), 2–20, <http://denk-doch-mal.de/wp/sigrid-hartong-algorithmisierung-von-bildung/>.

³⁴ Heydorn, Neufassung des Bildungsbegriffs (Fußnote 31), 150.

chen. Im Grunde ist auch diese Logik geprägt von einer Form der instrumentellen Komplexitätsreduktion, die sich bis hin zu hochentwickelten rechenbasierten Systemen beobachten lässt. Freiheit würde im Hinblick auf EdTech auch bedeuten, dass man nicht nur Zugang zur Software – respektive Hardware – bekäme, sondern auch die Einsicht in die algorithmischen Strukturen ermöglicht wird. Dies würde den Einsatz von freier bzw. Open Source Software bedeuten. Damit wären allerdings noch nicht die Probleme der Überwachung und Datensammlung gelöst.

Eine in diesem Zusammenhang durchaus kontrovers diskutierte Ausprägung ist Learning Analytics.³⁵ Im Kern geht es darum, den individuellen Lernprozess auf Basis von Datenerfassung und Auswertung zu optimieren. Dabei werden nicht nur Gewohnheitsdaten über das jeweilige Nutzungsverhalten erhoben und somit eine Art digitale Beobachtung oder Überwachung der Lernenden auf Basis vergangener Handlungen einbezogen. Es werden zugleich Erkenntnisse und – wahrscheinlichkeitsbasierte – Prognosen für die Zukunft abgeleitet. Legitimiert wird dies über das Versprechen, dass so beispielsweise unerkannte Lernschwächen frühzeitig und gezielt adressiert werden können. Die Konsequenzen sind jedoch auch hier vielschichtig und nicht nur auf die Lernenden bezogen, die damit zumindest einen Teil der Verantwortung an ein System übergeben, sondern auch auf die Lehrenden, die darüber ein Abbild des Lernverhaltens bekommen, welches die Beobachtungsperspektive geprägt durch Daten und die Repräsentation der Daten beeinflusst. Nimmt man Heydorn³⁶ und die Hinwendung zu geschichtlichen Zusammenhängen ernst, dann erscheint eine globale Überwachung und die Erhebung digitaler Daten, eine Verknüpfung von Datenpunkten und Kontextualisierung mindestens genauso problematisch, wie die Wahl der Technologie. Mündigkeit hieße dann sich entgegen einer Schicksalsgläubigkeit für

³⁵ Shum, S. B./ Ferguson, R., Social Learning Analytics 15(3) (2012), 3–26; Williamson, B., Who owns educational theory? Big data, algorithms and the expert power of education data science, *E-Learning and Digital Media* 14(3) (2017), 105–122, <https://doi.org/10.1177/2042753017731238>; Ifenthaler, D./ Drachler, H., Learning Analytics: Spezielle Forschungsmethoden in der Bildungstechnologie, in: Niegemann, H./ Weinberger, A. (Hg.), *Handbuch Bildungstechnologie*, Berlin/Heidelberg 2020, 515–534, https://doi.org/10.1007/978-3-662-54368-9_42.

³⁶ Heydorn, Neufassung des Bildungsbegriffs (Fußnote 31).

eine *gemeinsame* Gestaltung von Lern- und Bildungsräumen in der digitalen Welt einzusetzen, denn ein humanistisch geprägter Bildungsbegriff ist nicht technisch induziert. Das hieße aber auch, dass eine Kritik an der EdTech-Industrie nur eine Seite der Medaille darstellt.

5. Fazit

Die Berechnung der Bildung hat eine lange Tradition. Der Beitrag hat verschiedene Phänomene und Probleme ins Verhältnis zu bildungs- und kulturtheoretischen Überlegungen gestellt, um auf die Komplexität hinter der Konzeption von Bildung in einer digitalen Welt hinzudeuten. Dabei wurde aufgezeigt, dass Bildung schon auf begrifflicher Ebene von einer diffusen und von unterschiedlichen Positionen begleiteten Prämisse geprägt ist, die nicht selten in Technikdeterminismus mündet. Inwiefern es sich dann um eine verkürzte Sicht auf digitale Bildung handelt, wurde entlang der Differenz zum Allgemeinen der Bildung diskutiert und anhand der Herausforderungen zur Umsetzung von Kompetenzrahmen sichtbar gemacht. Dabei wurden zwei grundlegende Herausforderungen identifiziert: Einerseits ergeben sich neue Freiheiten im Umgang mit den digitalen Technologien, die es vorher in dem Maße nicht gegeben hat. Andererseits führt die Erinnerung an die unhintergehbare Bedingtheit der Zahlen zur Ernüchterung, der man nicht selten mit dem Wunsch nach mehr Technologieverständnis entgegentritt. Es ist jedoch genau diese Bewegung, der sich Bildung auf institutioneller und administrativer Ebene aktuell ausgesetzt sieht: Die beständige Schaffung von Räumen zur kreativen Auseinandersetzung mit dem Digitalen, bei gleichzeitiger kritischer Reflexion über die adäquate Repräsentation der Welt durch Daten.

Leitplanken 4.0 für die Freiheit 4.0? Von der Rationalität des Freiheitsverlustes

1. Verteidigt die Ethik die Freiheit?

Man kann die kurze Geschichte der ethischen Beschäftigung mit der Digitalisierung, die in ihren Ursprüngen vor mehr als 50 Jahren als Auseinandersetzung mit der Kybernetik begann,¹ als eine Geschichte der Desillusionierung lesen. Sie stellt dann die Fortsetzung jener Entwicklung dar, die auf Kepler und Galilei zurück geht und bei Freud und Libet noch längst nicht endet. Diese Entwicklung war im Kern keine Geschichte aneinander gereihter Kränkungen, wie es häufig heißt, sondern es handelte sich um einen schrittweisen Perspektivwechsel, bei dem jedoch der Mensch immer im Visier blieb.² Hatten Kepler und Galilei die Erde aus dem astronomischen Zentrum heraus katapultiert, so hat die darauf aufbauende Raumfahrt den „blauen Planeten“ zu einer Ikone unseres technischen Zeitalters gemacht.³ Wir wissen nun zwar, dass wir selbst nicht das Zentrum der Galaxie bilden, sind aber nach wie vor vor allem an den Bedingungen interessiert, die uns unsere Umwelt zur Heimat werden lassen. Wir denken nach wie vor „geozentrisch“ und

¹ Reblin, K., *Kybernetik und Menschenbild*, in: *Information der Ev. Zentralstelle für Weltanschauungsfragen* 32/V, Stuttgart 1968.

² Siegemund, A., *Beyond the hurts: Warum fühlt sich der gekränkte Mensch so gesund?* in: *Vogelsang, F./ Meisinger, H. (Hg.), Beyond Darwin – die unabgeschlossene Geschichte des wissenschaftlichen Fortschritts*, Bonn 2010, 237–254.

³ Sachs, W., *Satellitenblick. Die Ikone vom blauen Planeten und ihre Folgen für die Wissenschaft*, in: *Braun, I./ Joerges, B. (Hg.), Technik ohne Grenzen*, Frankfurt a. M. 1994, 305–346.

unsere Urteile bleiben selbst dann „anthropozentrisch“, wenn wir dies dezidiert ausschließen wollen. Haben Freud und die – gegen Libet gewendeten – Interpretationen der Experimente rund um den Zusammenhang zwischen Reiz-Reaktionen und Handlungsfreiheit zu einer Infragestellung des Subjekts Mensch und seines freien Willens geführt, so hat doch gerade die rationale Nachvollziehbarkeit von Willensentscheidungen zur Stärkung individueller Freiheitsrechte und Verantwortungsfähigkeit geführt. Wer reflektiert, dass die eigenen Handlungen durch soziale Prägungen, natürliche Instinkte und neuronal nachweisbare Aktivitätspotentiale beeinflusst werden, kennt eben die Bedingungen, unter denen er frei sein kann.

Nun stecken wir mitten in einem neuen Perspektivwechsel; mit der Digitalisierung verändert sich wiederum nicht die Freiheit an sich, wohl aber werden ihre Bedingungen neu justiert. Willens- und Handlungsfreiheit stehen zur Debatte, weil sie aus neuen Voraussetzungen resultieren. Aber die Annahme, ein KI-System bräuchte beispielsweise nur ein Mindestmaß an – künstlicher – Intelligenz, um einen eigenen Willen entwickeln zu können, ist ja nicht aus technischen, sondern aus anthropologischen Gründen fragwürdig. Tatsächlich hat sich natürliche Intelligenz, insbesondere das menschliche Bewusstsein, aus willensbesetzten Vorformen heraus entwickelt. Das intentionale Handeln geht dem bewussten Handeln voraus, so dass wir bislang kein Beispiel dafür haben, dass Willen aus Intelligenz heraus entsteht. Vielmehr lehrt die Alltagserfahrung geradezu das Gegenteil. Da mit dem Maß an Intelligenz auch die Komplexität für Problemlösungen zunimmt, fällt es weniger intelligenten Menschen leichter, einen konkreten Willen auszubilden. Willensentscheidungen gehen nämlich immer mit Komplexitätsreduktion und Kontingenz einher. Der Willen ist daher schon von Duns Scotus als eine vom Intellekt unabhängige Instanz beschrieben worden. Um einen Willen ausbilden zu können, bedarf es also nicht in erster Linie einer besonderen Form von Intelligenz, wohl aber der reflexiven Befähigung zur Distanznahme vom eigenen Intellekt. Wenn sich durch KI die Möglichkeiten, intelligent zu sein, verändern, dann sind dies wiederum – nur – Variationen der Bedingungen, unter denen eine Willensbildung möglich wird, es ist nicht die freie Willensbildung selbst.

Die hierbei anstehenden Veränderungen sind denen früherer Entwicklungen durchaus vergleichbar. So haben Psychoanalyse und Verhaltenstherapie besondere Formen von Abhängigkeit wie Zwangsstörungen oder Suizidgefährdungen nicht nur erklärt, sondern auch therapierbar gemacht. Die Erkenntnis konkreter Unfreiheit ermöglicht hier im Erfolgsfall eine neue Wahrnehmung von Freiheit. Die Möglichkeit, den eigenen Willen zu stärken, bildet sich nicht durch höhere Intelligenz aus, sondern durch die Variation ihrer Einsatzbedingungen. Dieses Wissen kann auch gegen die Entfaltung persönlicher Freiheitsrechte gewendet werden. Ähnlich gelagert ist der digitale Zugriff auf menschliche Verhaltensweisen. Wer diese technisch rekonstruieren kann, der macht damit zuvorderst eine Aussage über die Manipulierbarkeit von Menschen, nämlich die grundsätzliche Möglichkeit, das kulturschaffende Handeln durch Prozesse und Funktionen beschreiben zu können. Damit geht selbstverständlich die Beeinflussbarkeit der Handlungsbedingungen einher. Handlungen selbst sind aber immer erst mittelbar betroffen.

Wird die – aus technischer Sicht notwendige – Reduktion von Handlungen auf Funktionen von Seiten der Ethik beklagt, dann meistens mit dem Hinweis, dass die Empiriker eben nach Gesetzmäßigkeiten fragen und die Techniker somit gar nichts anderes rekonstruieren können als determinierte Vorgänge. Was Biologie bzw. Physik ergründen und Bioinformatik bzw. Robotik konstruieren, sind Kausalbeziehungen, also unfreie Zustände und Prozesse. Dies gilt prinzipiell trotz Relativitäts- und Quantentheorie, denn die Schwierigkeiten, die sich mit diesen revolutionären Veränderungen im naturwissenschaftlichen Weltbild verbinden, haben ja das technische Weltbild nicht abgelöst, sondern weiter differenziert und damit verstärkt. Kausale Beziehungen beinhalten immer Determination, das heißt, wenn die Ausgangsbedingungen bekannt sind, kennen wir potenziell auch schon das Ergebnis. Darauf beruht nicht zuletzt die prognostische Qualität künstlicher Intelligenz und daran ändert auch die Möglichkeit offener Systeme nichts Grundsätzliches. Ein System, das sich in verschiedenen Weisen entwickeln kann, besitzt Offenheit nur innerhalb der Bedingungen, die ihm im Entwicklungsprozess zugestanden werden. Weil der Grad der Offenheit hier also von den Technikern bestimmt wird, scheint es oberflächlich so zu sein, als könne man den Menschen ebenfalls als ein solches offenes System ansehen. Im Ergebnis

wäre er dann aber unfrei, weil das Maß der Offenheit ja von außen bestimmt wird. Dieser Gedanke ist einerseits erhellend, andererseits springen wir hier aber doch zu kurz. Erhellend ist die experimentelle Gleichsetzung des Menschen mit einem technischen System, weil wir so erkennen können, dass unsere Fragen auch schon die Fragen vergangener Generationen gewesen sind. So ist die theologiegeschichtlich interessante Diskussion um die „Vorsehung Gottes“ eben eine Auseinandersetzung um das Maß unserer Freiheit angesichts ihrer Bedingtheit. Aber bereits hier sind die Ergebnisse differenziert zu betrachten. Während Erasmus von Rotterdam die Mitwirkungsmöglichkeiten des Menschen am – göttlichen – Heil betonte, war für Martin Luther ausschließlich Gott das Subjekt des Heils, der Mensch demnach in den letzten Dingen unfrei. Anders als es konfessionelle Differenzen vermuten lassen, besteht hier jedoch kein grundsätzlicher Widerspruch. Vielmehr geht es um die anthropologische Perspektive. Luther fühlte sich nämlich keineswegs als ein unfreier Mensch, sondern als freier Herr und dienstbarer Knecht zugleich.⁴ Dies ist keine dualistische Anthropologie, wie mitunter unterstellt wird – innere Freiheit unter äußerem Zwang –, sondern ein Verweis auf die Relativität von Welterfahrungen – der Mensch *coram mundo* – angesichts der Möglichkeit des Absoluten – der Mensch *coram deo*. Entscheidend ist nämlich hier nicht die Erkenntnis, dass Menschen in Bezug auf fixe göttliche Entscheidungen keine Einflussmöglichkeiten haben; entscheidend ist, wie diese Erkenntnis wahrgenommen wird, nämlich als Entlastung – so Luther – oder als Beschränkung – so Erasmus.

An dieser Stelle zeigt sich die Differenz zwischen dem Menschen und offenen Systemen. Während wir unseren (un-)freien Zustand wahrnehmen, reflektieren und interpretieren können, bleiben digitale Systeme dem Grad ihrer Offenheit gegenüber taub, unabhängig davon, wie hoch er ist. Ob ein Roboter „selbsttätig“ tätig sein kann, interessiert ihn überhaupt nicht, sondern die Ingenieure, die mit ihm zusammenarbeiten. Ob ein Rechner rechnen kann, weiß er nicht. Genau genommen rechnet er auch nicht, denn „Rechnen“ beinhaltet in unserem Sprachgebrauch immer schon die Fähigkeit, die

⁴ Luther, M., Von der Freiheit eines Christenmenschen, WA 1883ff., Bd. VII, 1897, 21, 1–4.

Plausibilität des Ergebnisses prüfen zu können. Ob richtig oder falsch gerechnet wird, erschließt sich ausschließlich für uns Menschen im Zusammenhang der kulturellen Bedingungen, unter denen der Rechenvorgang stattfindet. So ist die Gleichung $a+a=2a$ keinesfalls immer richtig, sondern nur dann, wenn es einen Menschen gibt, der die Richtigkeit erkennt. Ohne dieses Erkenntnisvermögen ist sowohl die Gleichung als auch die Behauptung, sie sei richtig, sinnlos. Selbst der falsche Rechenweg setzt einen Sinn für Sinn voraus.

Verfolgen wir diesen Gedanken weiter, dann sind digitale Prozesse nur im Rahmen ihrer kulturellen Einbettungsfaktoren überhaupt sinnvoll. Gleichzeitig fußen alle technischen Vorgänge auf der Annahme, dass naturgesetzliche Vorgänge prinzipiell auch ohne das Zutun des Menschen so ablaufen, wie es Physik und Mathematik beschreiben. Unter der Bedingung, dass Richtigkeit Sinnhaftigkeit voraussetzt, werden auch digital determinierte Vorgänge sinnfrei, sobald es keine Menschen mehr gibt, die das Digitale als Bedingung ihrer Existenz akzeptieren. Im Bereich des moralischen Handelns gibt es keine Gesetzmäßigkeiten, sondern hier haben wir es mit singulären, historisch gewachsenen und kulturell vermittelten Situationen zu tun. Genau darin aber besteht unsere Freiheit; ihre Bedingungen als unsere Aufgabe anzuerkennen, ist somit bereits ein ethisches Urteil, auf dem wir aufbauen können.

2. Ethik im digitalen Kulturzusammenhang

Nun könnte man dabei stehen bleiben und hätte dann im Prinzip jene Arbeitsteilung vorliegen, die bereits Max Weber formuliert hat.⁵ Bei ihm ist die Wissenschaft für Erklärungen zuständig und die Ethik, resp. die Religion, für den Sinn. Dies stimmt aber schon deshalb nicht, weil wir technische Systeme gar nicht konzipieren können, ohne ihnen Sinn zu unterstellen. Auch die Digitalisierung ist undenkbar ohne ihre lebensweltlichen Schnittstellen zu menschlichen Sinngebungen. Während wir bei der Entwicklung digitaler

⁵ Weber, M., Gesammelte Aufsätze zur Religionssoziologie, Bd. 1, Tübingen 1988 (1920), 564ff.

Optionen unsere Freiheit gebrauchen, indem wir ihre Bedingungen manipulieren, kann uns ihre Verwendung – ebenso wie die Abhängigkeit von einer göttlichen Vorsehung – belasten oder entlasten. Belastend wirkt die digital erzeugte Abhängigkeit, wenn wir sie als Unfreiheit erleben, entlastend, wenn sie uns Sicherheit gibt. Wenn uns der Algorithmus eines Onlinehändlers mit den passenden Kaufvorschlägen versorgt, erleichtert er die Suche. Wird uns aber permanent uninteressanter und überteuert Tand angeboten, fühlen wir uns manipuliert.

Auch ein kurzer Rückblick auf die Geschichte des Verhältnisses zwischen Ethik und Wissenschaft zeigt, dass es hier keine prinzipielle Aufgabenteilung, sondern eine kulturell geprägte Zuordnung gibt.⁶

1. Als Galileo Galilei den Lauf der Welt auf die heute allseits bekannten, aber damals neuen Gesetzmäßigkeiten zurückführte, traf er vor allem deshalb auf Widerstand, weil der damalige *common sense*, also die herrschende Moral bewahrt werden sollte. Galileis Fernrohr aber hatte die Perspektive grundsätzlich verändert. Er machte die berechnende, empirische Untersuchung zur vorherrschenden Erkenntnismethode und bedrohte damit – aus Sicht seiner Kritiker – die Überzeugungskraft des kulturell fest verankerten Weltentwurfs.
2. Als sich James Watson zur letzten Jahrtausendwende für die Selektion von Embryonen aussprach, schrieb der Theologe Dietmar Mieth in der FAZ, dass wir statt dieser „Ethik des Grauens“ eine Ethik bräuchten, die „eine Antwort darstellt, auf etwas, das vorgegeben ist.“ Auch hier stehen sich *common sense* und Wissenschaft gegenüber. Während die herrschende Moral durch den Würdegedanken geprägt ist, folgt der Biologe dem Machbarkeitspostulat.
3. Der Deutsche Ethikrat hat 2018 eine Stellungnahme „Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung“⁷

⁶ Vgl. zum Folgenden Schwarke, Ch., Watson, Galilei und die „zwei Kulturen“, in: Grenzen überschreiten, Festschrift zum 70. Geburtstag von Trutz Rendtorff, München 2001, 151–163.

⁷ Deutscher Ethikrat, Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung, Stellungnahme, Vorabfassung vom 30. November 2017, Berlin 2018.

erarbeitet. Darin wird festgehalten: „Souveränität galt als jene Eigenschaft Gottes oder eines absolutistischen Herrschers, kraft derer er absolut und unbedingt von anderen Mächten alles zu tun oder zu lassen imstande sei.“⁸ „Informationelle Freiheitsgestaltung [...] meint interaktive Persönlichkeitsentfaltung unter Wahrung von Privatheit in einer vernetzten Welt und ist gekennzeichnet durch die Möglichkeit, auf Basis persönlicher Präferenzen effektiv in den Strom persönlich relevanter Daten eingreifen zu können.“⁹

Die Digitalisierung steht demnach in einer längeren Folge von Auseinandersetzung mit uns selbst. Denn von der Souveränität Gottes oder einer unbedingten Würde sprechen wir nur, um unsere eigene Freiheit zu kultivieren. Ebenso ist „informationelle Freiheitsgestaltung“ die bewusste Gestaltung unserer Freiheitsbedingungen. Die Person bleibt Kriterium der Freiheit. Dazu gehört einerseits ihre Begrenztheit, die theologisch in Form der Mensch-Gott-Beziehung, ethisch als Würdepostulat beschrieben wird. Andererseits gehört die Befähigung zum Freiheitsgebrauch dazu.

Die Ethik ist also nicht nur Verteidigerin der Freiheit, sondern auch der gerade bestehenden Ordnung. Sie konfrontiert die Wissenschaft, die durch ihren Erkenntnisdrang ganz selbstverständlich zur Infragestellung der Ordnung beiträgt, mit dem herrschenden *common sense*. Die Beispiele sind durchaus paradigmatisch für die vergangenen 400 Jahre. Die Ethik hat immer die gefestigten Ordnungen ins Bewusstsein gerufen. Geozentrik, Würdepostulat, Persönlichkeitsentfaltung – das sind die hier die Stichworte. Tatsächlich geht es der Ethik aber nicht um die Begrenzung des wissenschaftlichen Erkenntnisdrangs, dies war schon im Fall Galilei keinesfalls der Beweggrund für die Auseinandersetzungen.¹⁰ Die Ethik erinnert vielmehr daran, dass die Perspektive von Wissenschaft und Technik zu je anderen Perspektiven im Rahmen des Kulturzusammenhangs in Beziehung gesetzt werden muss.

⁸ Ebd., 21.

⁹ Ebd., 40.

¹⁰ Leich, P., Die schwierige Beziehung von Ratio und Religio: Der Inquisitionsprozess gegen Galileo Galilei, in: Losch, A./ Vogelsang, F. (Hg.), Wissenschaft und die Frage nach Gott. Theologie und Naturwissenschaft im Dialog, Bonn 2015, 22–29.

In dieser Tradition steht auch der heutige Ruf nach einer Ethik für das digitale Zeitalter. Wer nach Ethik fragt, will den Fortschritt einhegen, weil er das Gewohnte infrage stellt und der Übergang human gestaltet werden muss. Im Fall der künstlichen Intelligenz geht es demnach nicht um die Verteidigung der Freiheit als Selbstzweck, sondern als Ausdruck einer religiös und politisch gewachsenen Ordnung, die dieser Freiheit nicht nur einen bedeutenden Wert zu zuzusagt, sondern in der alle, auch die digitalen Räume, an dieser Freiheit partizipieren.

Wenn die Wissenschaften die Freiheit infrage stellen, dann machen sie auf die Relativität unserer auf Freiheit basierenden Ordnung aufmerksam. Prognostisch wirkende Algorithmen zeigen uns, dass Freiheit nicht selbstverständlich ist. Wir würden die Ethik aber völlig missverstehen, wenn wir sie als Grenzposten der Gesellschaft angesichts der Angriffe aus Wissenschaft und Technik ansähen. Die herrschende Ethik verteidigt nicht den Heiligen Geist, der bekanntlich weht, wo er will, sondern den Zeitgeist, der durchaus gebunden ist. Aus diesem Grund ist das, was wir heute als menschenwürdig ansehen, ja gerade gegen die Zeitgeister der Vergangenheit durchgesetzt worden. Die Menschenrechte, die Demokratie, die Freiheit der Wissenschaften und der Religion – immer wurde im Namen einer moralischen Einsicht dagegen opponiert und immer hat die Ethik, und zwar sowohl in ihrem religiösen als auch in ihrem säkularen Gewand, den Zeitgeist gegen das Neue verteidigt. Der Beitrag der Ethik besteht nicht in einem restaurativen Kampf gegen die Digitalisierung, sondern in der Suche nach Wegen ihrer Humanisierung.

Dass Ethik dem Fortschritt dient, indem sie die Grenzen des Machbaren behutsam neu zeichnet, gilt bereits für die Indexkongregationen im 17. Jahrhundert, und es gilt erst recht für die heutigen Ethikräte, deren Aufgabe ja gerade darin besteht, die Kulturverträglichkeit von Innovationen sicher zu stellen. Der allgemeine Ruf nach Ethik, der heute aller Orten erschallt und dem sich Wissenschaft und Technik ausgesetzt sehen, resultiert aus der Suche nach Orientierung angesichts einer Transformation. Orientierung bekommt man aber niemals von der unsicheren Zukunft her, sondern immer aus der vermeintlichen Sicherheit des Vergangenen.

Insofern stellt Ethik angesichts der Digitalisierung ein doppelt paradoxes Geschäft dar. Die Quantifizierung unseres Daseins und die Beschreibung fundamentaler Lebenszusammenhänge im binären System stellen einen Angriff auf unsere etablierte Ordnung dar, die von der Unverfügbarkeit freier Menschen ausgeht. Die Einhegung der technischen Möglichkeiten ist demnach eine Beschränkung der Freiheit im Namen der Freiheit. Die relevante Frage ist also nicht, wie die Freiheit zu wahren ist, sondern wie ihre Bedingungen gestaltet sein sollen und was wir dafür bereit sind, zu zahlen. Denn zahlen müssen wir, und zwar entweder mit dem Verlust an Freiheitsgraden, weil uns unsichtbare Algorithmen steuern oder mit dem Verlust an Sicherheit, weil wir uns letzte Räume des Unberechenbaren erhalten. Wofür wir uns im konkreten Fall entscheiden, wird sich aus dem bestimmen, was wir uns vom Menschsein im Verhältnis zu unserer Vernunft erhoffen. Daher ist nun ein Blick auf das Menschsein im Digitalen zu werfen.

3. Die Grenzen der Vernunft

Die Wurzeln des Protests gegen eine durch das Digitale präjudizierte Ethik liegen in der Sorge begründet, die Digitalisierung könne dem ohnehin schon geschwächten abendländischen Menschenbild den Todesstoß versetzen. Wenn diese Sorge bislang auch noch nicht scharf artikuliert ist, so wird man doch zugestehen müssen, dass in ihr ein gewisses Gespür für die harten Konsequenzen zum Ausdruck kommt, die sich bereits aus den anthropologischen Aussagen der Kybernetik ergeben. Die Kybernetik ist seit den 1950er Jahren auch davon geprägt, die Funktionen von Mensch und Maschine einander gleich zu stellen.¹¹ Dies liefe tatsächlich auf die Zerstörung des klassischen Menschenbildes hinaus, wie es die griechische Philosophie und in ihrer Nachfolge die Scholastik und René Descartes ausgeprägt haben. Nach diesem Menschenbild ist die Besonderheit des Menschen seine Vernunft, die zugleich etwas Göttliches ist und sich konsequent von allem Materiellen abhebt. Gemäß den obigen Ausführungen ist dieses Menschenbild aber nicht nur

¹¹ Vgl. Steinbuch, K., *Automat und Mensch – kybernetische Tatsachen und Hypothesen*, Berlin/Heidelberg/New York 31965.

kulturell geprägt, sondern auch sachgemäß, insofern Vernunft im Unterschied zu bloßer Intelligenz als Befähigung zur Reflexion der Lebenssituation und damit als Differenzkriterium zwischen Mensch und Maschine gelten kann.

Die Vernunft verbindet in der Tradition Gott und Mensch, dies ist bereits eine Grundaussage der antiken Anthropologie. Aus dieser Voraussetzung der Identität von menschlicher und göttlicher *ratio* folgt sogleich die erkenntnistheoretische Möglichkeit der rationalen Gotteserkenntnis. Aber trotz der Idee, dass die Vernunft göttlichen Ursprungs ist, kennt auch die Antike bereits ihre Begrenztheit. Für die Stoa ist die Vernunft nur ein Bruchteil der göttlichen Vernunft, dennoch besteht auch hier Wesensgleichheit. Die Apotheose der geistigen Fähigkeiten des Menschen hat in der europäischen Geistesgeschichte eine Abwertung des Materiellen zur Folge, die bis in einen strengen Dualismus von Leib und Seele hinein mündete. In dieser Tradition stehen Hoffnungen, die dem Menschen eine Ewigkeit *in silicio* angedeihen lassen wollen. Dem steht vor allem die biblische Tradition entgegen, nach der der Mensch eine leiblich-seelische Einheit darstellt, so dass Geist und Materie in Bezug auf den Menschen zwar gedanklich voneinander unterschieden, aber nicht substantiell getrennt werden können.

Das nichtbiblische antike Denken ist heute noch wirksam, wenn von einer Trennung von Körper und Seele oder von einem Weiterleben der autarken Seele nach dem Tode gesprochen wird. Wird aber das Geistige grundsätzlich vom Körperlichen abgesondert, dann werden die heutigen Entwicklungen im Bereich der Digitalisierung und KI zu einer grundsätzlichen Anfrage an die Existenz des freien Geistes überhaupt. Wenn man den freien Geist des Menschen vom Bereich naturgesetzlicher Kausalität trennt und ihn metaphysisch überhöht, dann muss die digitale Rekonstruktion unserer Denk- und Gefühlsfunktionen als Angriff auf die Freiheit des Menschen verstanden werden.

Wird der Versuch, den Menschen in Analogie zu technischen Systemen zu verstehen, als Angriff auf das Selbstverständnis angesehen, dann steht dahinter die idealistische Annahme, dass der Geist separiert werden kann und sich die den Menschen definierenden Eigenschaften sämtlich in seiner geistigen Natur zeigen. Alles Körperliche ist demnach verzichtbar. Wenn nun menschliche Existenz mehr und mehr durch das körperlich-materielle bestimmt

wird, so bleibt dem Idealisten nur die Weltflucht, die sich in einer schwächeren Variante als kulturpessimistische Technikkritik äußert. Eine andere Sichtweise ergibt sich allerdings, wenn man auf eine derartige Metaphysik verzichtet. Zwar spricht die Philosophie erst seit dem Ende des 20. Jahrhunderts von einem sog. nachmetaphysischen Zeitalter, aber eine Entzauberung der Vernunft hat es doch schon vorher gegeben. So ist die jüdische Anthropologie vom Ineinander von Körper und Geist geprägt. Auch in den Thesen über den Menschen, die Martin Luther 1536 formuliert hat, heißt es über die Vernunft:¹²

„Und in der Tat ist es wahr, daß die Vernunft die Hauptsache von allem ist, das Beste im Vergleich mit den übrigen Dingen dieses Lebens und geradezu etwas Göttliches. [...] Gleichwohl, daß sie solche Majestät sei, weiß eben diese Vernunft nicht auf Grund von deren Ursache, sondern nur durch Rückschluß aus den Wirkungen. [...] Und was jämmerlich ist: Nicht einmal über ihren Entschluß oder ihre Gedanken hat die Vernunft volle und zuverlässige Gewalt, sondern ist darin dem Zufall und der Nichtigkeit unterworfen.“

Die Vernunft wird hier als eine höchst trügerische Angelegenheit beschrieben. Der Mensch, der sie sich zunutze macht, ist nicht davor gefeit, von ihr auf eine falsche Fährte geführt zu werden. Luther hält es also für möglich, dass uns die *ratio* etwas vorspiegelt, das nicht der Realität entspricht. Diese Sicht auf den vernunftbegabten Menschen kennen wir, wir hören sie heute u. a. aus den Neurowissenschaften. So meint Michael Pauen beispielsweise, dass das Ich eine Illusion und die menschliche Freiheit eine Größe ist, die keine kausale Wirksamkeit in der Realität hat.¹³ Thomas Metzinger pflichtet ihm bei, er nutzt die Metapher „Egotunnel“, um im Prinzip dasselbe zu sagen.¹⁴ Subjektbewusstsein und Ich-Bezug sind demnach Vorspiegelungen des Gehirns. Unsere Rationalität spiegelt uns folglich eine Freiheit vor, die sich der Erkenntnis durch die Rationalität entzieht. Darin stimmen Luthers

¹² Luther, M., *Disputatio de homine* (1536), WA 39, I, 175–177.

¹³ Pauen, M., *Illusion Freiheit? Mögliche und unmögliche Konsequenzen der Hirnforschung*, Frankfurt a. M. 2004.

¹⁴ Metzinger, Th., *Der Ego-Tunnel. Eine neue Philosophie des Selbst: Von der Hirnforschung zur Bewusstseinsethik*, Berlin 2009.

Thesen und die heutige Anthropologie überein. Offen bleibt die Frage, wie Pauen und Metzinger darauf kommen, dass sich jeder Mensch etwas vorspiegeln lässt, das nicht real ist und warum gerade sie die Möglichkeit haben, aus diesem Verwirrspiel der Vernunft auszubrechen. Dieser Kritikpunkt trifft aber genauso Martin Luther. Woher will er denn wissen, dass die Vernunft „blinde Kuh“ spielt?¹⁵ Er ist ja, ebenso wie die heutigen Neurowissenschaften, auch in dem Moment, in dem er die Vernunft kritisiert, von ihr gefangen.

4. Das Bild der KI

Der Ausweg, der hier zu suchen ist, muss nicht nur die Möglichkeit von Rationalität bewahren, durch ihn müssen wir auch erklären können, wie sich diese Rationalität zu unserer Freiheit verhält. Ich schlage vor, den oben skizzierten Weg weiter zu gehen und zwischen zwei Perspektiven zu unterscheiden. Die Möglichkeiten künstlicher Intelligenz stellen dann jene Freiheit infrage, über die wir verfügen, während sie selbst ja ein durchaus optionaler Weg bleibt, den *wir* gehen. Das Problem besteht darin, dass wir die Freiheit, prinzipiell auch anders leben zu können, immer weniger sehen. Dies zeichnete sich bereits in der Auseinandersetzung mit konventionellen Techniken ab, die sich ja ebenfalls in unseren Alltag einnisteten und die Frage aufwarfen, ob es ein Leben außerhalb überhaupt noch geben könne. Werden wir im digitalen Zeitalter endgültig das, was Herbert Marcuse 1967 als den eindimensionalen Menschen bezeichnet hat?¹⁶ Marcuse beschreibt, dass es sich bei der technologischen Rationalität um eine grundsätzlich neue Denkweise handelt. Diese Denkweise unterscheidet er von der aufgeklärten Denkform des freien Individuums. Vernunft und Autonomie seien in der technologischen Rationalität aufgegangen und hätten uns alle zu Anpassungsleistungen an die Technik gezwungen. Das Schicksal des eindimensionalen Menschen besteht darin, dass jede Form des Widerstands nicht etwa politisch problematisch wäre,

¹⁵ Luther würde seine Einsicht allerdings auf sein Offenbarungserlebnis, mithin auf die Mensch-Gott-Beziehung zurückführen. Dieser Weg ist den Neurowissenschaftlern allerdings versperrt.

¹⁶ Marcuse, H., Der eindimensionale Mensch. Studien zur Ideologie der fortgeschrittenen Industriegesellschaft, hg. v. P.-E. Jansen, Springe 2014 (1967).

sondern schlicht irrational ist. Die Anpassung an die Zweckrationalität der Maschine ist demnach sinnvoll und angenehm. Wer wollte da ausscheren?

So fällt es bislang kaum auf, dass die Anpassung durch die Aufgabe von Autonomie und die Beschränkung von Freiheit erkaufte wird. Für denjenigen, der sich von der technischen Rationalität leiten lässt, ist die Beschränkung der eigenen Freiheit rational. Jeder, der sich seinen Weg in den Urlaub digital vorausberechnen lässt und diesen Berechnungen folgt, hat einen Teil seiner Freiheit um der Freiheit willen delegiert. Wir unterwerfen uns der digitalen Rationalität nicht, weil sie uns zwingt, sondern weil sie förderlich ist. Folgen wir diesem Gedanken, dann ist mit KI jedoch keine grundsätzliche Infragestellung der Freiheit verbunden, sondern eine immerwährende Aufgabe. Dass wir die Zweckrationalität des Digitalen als angenehm erfahren, eben dies ist ja gerade die Art und Weise, in der wir das Digitale als Bedingung unserer Freiheit wahrnehmen – und gestalten müssen.

Anders gesagt: Der digitale Zeitgeist macht es sich bequem. Folgen wir der Bequemlichkeit, dann ist die Freiheit damit nicht abgeschafft, sondern in einer bestimmten Weise genutzt. Kommt es im Weiteren dazu, dass uns die bequemen Vorgaben der Algorithmen stören, so ist dies ein Zeichen für unsere doch immer noch vorhandene Freiheit. Problematisch ist, dass jeder Protest gegen intelligente Dienste, die das Leben erleichtern, ebenso wie der stille Verzicht unter Ideologieverdacht gerät. Es mag vielleicht lächerlich sein, bewusst auf ein Navigationsgerät zu verzichten, der Verzicht auf ein mobiles Telefon gilt auf einsamen Reisen heute bereits als Selbstgefährdung. Da die digitale Welt unseren Autonomieverlust als sinnvolle Datenkorrelationen ausweist, gehört die Kritik an ihr inzwischen ins Reich des Unsagbaren. Ja mehr noch, der Verlust der Freiheit wird nicht als das „Werk feindlicher und fremder Kräfte“ (Marcuse) erlebt, sondern als Ergebnis des Vernunftgebrauchs. Protest und Verzicht werden hingegen als irrational gebrandmarkt, weil der Technikgebrauch der Rationalität entspricht. Hier muss aber zwischen der Vernunft, die jedem Menschen grundsätzlich zukommt, und der heutigen Form ihres Gebrauchs unterschieden werden. Wer sich den Dia-

gnostikmethoden der modernen Medizin nicht aussetzen will, handelt in Bezug auf den aktuellen Vernunftgebrauch irrational; ein derartiger Verzicht ist aber nicht an sich unvernünftig.¹⁷

In all dem löst sich die Freiheit im digitalen Raum nicht auf, sondern ihre Bedingungen werden zu quasi herstellbaren Produkten. Dabei erlebt die Freiheit sogar eine Renaissance, denn wir betrachten sie nicht mehr als Ziel *unseres* individuellen und gesellschaftlichen Handelns, sondern wir unterwerfen uns gezielt unter Prozesse und Verfahren, die so zweckrational sind, dass wir *von ihnen* eine Steigerung unserer Autonomie erwarten. Das Freiheitsversprechen der Technik – siehe den Beitrag von Martina Hessler in diesem Band – ist auch im Bereich der KI so intakt wie eh und je, vorausgesetzt wir erleben konkrete KI als freiheitsfördernd. Solange uns die Algorithmen den Lebensvollzug versüßen, schränkt derjenige, der die Eindringtiefe des Digitalen reduziert, den Freiheitsraum ein, ohne dies als Beschränkung zu erfahren. Die Umkehrung, also die Wahrnehmung von Künstlicher Intelligenz als einer potenziell freiheitsraubenden Macht, setzt voraus, dass konkrete KI den Lebensvollzug hindert. Die Art der Wahrnehmung ist dabei wichtiger als die KI selbst.

Es sei hinzugefügt, dass sich auch der Streit zwischen Erasmus und Luther nicht an unterschiedlichen Göttern, sondern an verschiedenartigen Gottesbildern entzündet hat. Ganz parallel dazu werden wir heute sagen können, dass derjenige, der einen Roboterhund als Bereicherung erlebt, diesen genauso närrisch lieblosen kann wie ein konventioneller Hundehalter. Tiere bzw. Roboter dienen der Steigerung der Lebensqualität, weshalb die mit ihnen verbundene Unfreiheit gar nicht thematisiert wird. Desselbigengleichen wird es aber auch – weiterhin – Menschen geben, die sich von Hunden bedroht fühlen. Die positive Gestaltung der individuellen Freiheit oder ihre Bedrohung – eine konkrete KI macht offensichtlich beides zugleich möglich.

¹⁷ Es ist zu konstatieren, dass ein andersartiger Gebrauch der Vernunft in unserer Zeit häufig diskreditiert wird, indem fremde Moralvorstellungen als Unvernunft etikettiert werden. So wird der religiös begründete Verzicht auf medizinische Behandlungen – beispielsweise jener von Bluttransfusionen bei den Zeugen Jehovas – häufig als unvernünftig angesehen, obwohl es sich eigentlich um ein Werturteil handelt, bei dem Werthaltungen nur anders gewichtet werden als in der Mehrheitsgesellschaft.

Damit ist auch gesagt, was unsere Freiheit wirklich herausfordert: es ist nicht die KI, sondern das Bild, das wir von ihr haben.

5. Transzendente Freiheit und die Leitplanken 4.0

Menschliche Freiheit ist aus theologischer Perspektive immer nur relativ, denn sie gehört in den Kreis des Endlichen hinein. Alle unsere reproduzierbaren Denkprozesse sind endlich; darin unterscheidet sich eine christliche Anthropologie von der vorchristlich-antiken. Deshalb kann zumindest die theologische Ethik heute sehr offen sein gegenüber den Versuchen, unsere geistigen Prozesse technisch zu reproduzieren. Biblische Anthropologie und technischer Realismus stimmen in einem entscheidenden Punkt überein: Sie verstehen den Menschen ohne metaphysische Arbeitshypothese und verorten ihn innerhalb der Grenzen alles Endlichen und damit innerhalb der physikalischen Gesetzmäßigkeiten.¹⁸

Was den Menschen im theologischen Sinne ausmacht, findet seine Begründung nicht in einer Eigenschaft des Menschen, die Forschungsgegenstand irgendeiner Wissenschaft sein könnte. Auch die Tatsache, dass der Mensch Geistwesen ist, entscheidet nicht über sein Menschsein. Konstitutiv für den Menschen ist vielmehr die Gottesbeziehung, also etwas, das außerhalb des Menschen und auch außerhalb der Natur liegt, zu der wir gehören. KI ist im Gegensatz dazu keine Größe, die das Humanum konstituiert. Absolute Freiheit gehört ebenfalls zu den äußeren Dingen; sie ist nichts Immanentes, über das wir verfügen könnten. Damit sind die Möglichkeiten, den endlichen

¹⁸ Wenn der erste Schöpfungsbericht (Gen 1) sagt, Gott habe Pflanzen, Tiere und Menschen allein durch sein Wort geschaffen, dann stellt die jüdische Tradition diese Aussage der Metaphysik des babylonischen Gilgamesch-Epos entgegen. Bei den Babyloniern war das Verhältnis zwischen Schöpfer und Geschöpf stark kosmologisch-metaphysisch geprägt. Hier hatte Marduk die Göttin Tiamat mit dem Schwert in zwei Stücke geschlagen und aus den Teilen des erschlagenen Körpers den Himmel und die Erde erschaffen. Himmel und Erde sind also aus göttlicher Substanz gemacht. Im biblischen Schöpfungsdenken bleibt die Grenze zwischen Gott und Mensch hingegen gewahrt.

Menschen mit all seinen Fähigkeiten zu erforschen und diese Fähigkeiten technisch nachzuahmen, nur durch den Faktor Endlichkeit begrenzt.

Die Leitplanken, über die wir ethisch diskutieren müssen, folgen demnach keinen prinzipiellen religiös-weltanschaulichen, sondern ganz praktischen Prämissen. Unser Freiheitsbewusstsein speist sich unter anderem auch aus einer reformatorischen Tradition, in der es um die Freiheit geht, sich in der Gottesbeziehung selbst finden oder verlieren zu können. Versteht man Freiheit so, wie sie Naturwissenschaft und Technik verstehen müssen, nämlich immanent, dann ist sie eine Verzerrung, die auf das subjektive Bewusstsein zurück zu führen ist. Das haben, wie oben beschrieben, u. a. die Neurowissenschaften durchaus richtig erkannt. Freiheit, wie sie die Bibel versteht, kommt dem Menschen von außen zu, sie ist transzendent.

Welche Aufgabe kommt dann der Ethik zu? Es geht darum, die transzendente Freiheit erfahrbar zu machen, indem wir die immanente Freiheitsgestaltung als Aufgabe ansehen – und als Teil des Verfügungszusammenhangs. Es geht nicht darum, Freiheit gegen unser Verfügen ins Feld zu führen. Wer behauptet, KI sei ein Angriff auf die Freiheit, der folgt einer spezifischen Wahrnehmung. Dem gegenüber ist zu sagen, dass die Unfreiheit, die wir angesichts des digitalen Zugriffs auf unsere Existenz erleben, genau so konstruiert ist wie die Freiheit, die wir erleben könnten. Die Ethik wird die Entwicklung von KI-Systemen demnach als Gestaltung unserer relativen Freiheitsräume begreifen – und sich in unserer Kultur dafür einsetzen, dass die Technik diese Räume erweitert. Ich schlage zum Schluss einige normative Bedingungen vor, die einen humanen KI-Gebrauch ermöglichen:

1. Erhalten bleiben sollte die Freiheit, sich weltanschaulich positionieren zu können und dabei die Zwänge, die durch immanente Prozesse zustande kommen, möglichst klein zu halten. Konkret ist damit die Religionsfreiheit angesprochen, die vor allem die Freiheit beinhaltet, unsere Transzendenzfähigkeit zu gestalten. Religionsfreiheit ist hierbei zweierlei, eine Freiheit von äußeren Zwängen und eine Freiheit zum freien und öffentlichen Bekenntnis. Diese doppelte Freiheit in digitalen Räumen zu schaffen, ist eine schwierige, aber lohnenswerte Aufgabe, vor der wir als Erben unserer analogen Religionskulturen stehen.

2. Zweitens geht es um die Aufgabe, die eigene Identität gestalten zu können. Wir sind dabei, uns digitale Artefakte wie Smartphones und Implantate über Brain-Computer-Interfaces einzuverleiben. Biofakte verwischen die Grenze zwischen Mensch und Maschine, so dass zu unserer bisherigen leiblich-seelischen Identität eine digitale Identität hinzutritt. Die digitale Identität sollte nur eine unter vielen bleiben und kein Mensch sollte auf diese eine Identität festgelegt sein. Bedeutsam wird dies etwa bei der Frage nach der Vergesslichkeit digitaler Systeme.
3. Ein wesentlicher Punkt ist die soziale Anerkennung in und durch digitale Medien. Hier entwickeln sich neue Formen der Selbstperformanz und neuartige Kriterien für Inklusion und Exklusion. So ergibt sich einerseits die Notwendigkeit, diese Formen der Anerkennung relativ zur unbedingten Anerkennung des Menschseins, also seiner Würde, anzusehen. Verhaltenskontrollen anderer sind ebenso kritisch wie der Hang zur Selbstdarstellung.
4. Daraus erwächst viertens die Aufgabe, auch digitale Lebensäußerungen revidierbar zu gestalten. So muss es beispielsweise im Bildungssektor auch unter digitalen Bedingungen Lernorte geben, an denen Fehler gelöscht werden, so dass später keiner darauf zugreifen kann. Freiheit setzt Erprobungsräume voraus.

Glaubens- und Gewissensfreiheit, Selbstgestaltung, Würde und Revidierbarkeit, das sind vier ethisch-moralische Leitplanken für die digitale Welt, anhand derer die Ethik das Terrain absteckt, welches unserer heutigen Kultur am Übergang zum digitalen Zeitalter wichtig ist. Die theologischen Begriffe hierfür sind Bekenntnis, Buße, Gnade und Vergebung. In all dem geht es nicht um die Markierung fixer Grenzen, sondern um die Kulturverträglichkeit der digitalen Transformation. Diese Transformation wird wesentlich von den Bildern bestimmt, die wir mit KI verbinden.

Freiheit durch Digitalisierung? Theologische Bewertung einer fragwürdigen Hoffnung

Die Aufgabenstellung ist hinsichtlich all ihrer relationalen Bestandteile vieldeutig, also erst recht hinsichtlich der gesamten Relation. Weder „Digitalisierung“, noch „Freiheit“ sind eindeutige Begriffe, die auf bestimmte Phänomene referieren. Aber auch, was eine theologische Perspektive ist, ist nicht eindeutig zu beschreiben. Da insbesondere die phänomenalen Veränderungen des mit „Digitalisierung“ beschriebenen Sachverhalts so vielfältig sind, dass im Prinzip jedes der unterschiedlichen Phänomene einer theologisch sozialetischen Evaluation bedürfte, werde ich im Folgenden keines dieser recht unterschiedlich zu beurteilenden Phänomene in den Blick nehmen, sondern werde versuchen, im Folgenden etwas Grundlegendes zu beschreiben.

1. Die theologische Perspektive

Theologie als methodische Selbstreflexion des christlichen Glaubens ist weder auf ein System von Dogmen noch auf ein Wirklichkeitsverständnis gerichtet, noch bringt sie so etwas hervor. Der christliche Glaube ist vielmehr eine Weglinienperspektive oder eine Weise des Wahrwertnehmens, vermittelt im Medium der Narration des Evangeliums.¹ Unmittelbares Wahrnehmen ist nie wertfrei, sondern primär stets Wahrwertnehmen in Einheit. Wer-

¹ Vgl. Mühling, M., Post-Systematische Theologie I. Denkwege – Eine theologische Philosophie, Leiden/Paderborn 2020, 579–581.

te werden nicht an wertlose Fakten herangetragen, sondern aus der Einheit des Wahrwertnehmens ergeben sich erst sekundär-reflexive Objektivierungsmöglichkeiten.² Das unmittelbare Wahrwertnehmen ist allerdings nur hinsichtlich der kognitiven Reflexion unvermittelt. Es findet selbst in einem narrativen Medium statt, und dies ist die Verschränkung der primären Lebensgeschichten der Wahrnehmenden mit zahlreichen sekundärnarrativen Erzählungen, in denen die Wahrnehmenden selbst sich bewegen, aufwachsen, gebildet werden.

Im Unterschied zu rein literaturwissenschaftlichen Narratologien arbeite ich hier mit einer narrativen Ontologie des Werdens. Sie geht davon aus, dass „stories are lived before they are told,“³ so dass Erzählungen Bestandteile der Wirklichkeit selbst sind. *Primäre Narrationen* bilden das Geflecht der Wirklichkeit. Sie werden nicht von Autoren erzählt, sondern bringen Personen, Gegenstände und Ereignisse hervor. Davon unterschieden sind *sekundäre Erzählungen*, die alle zeichenhaften Verwendungen von Personen und Gemeinschaften darstellen. Sie können sich in mannigfacher Weise aufeinander beziehen. Das Geflecht primärer und sekundärer Erzählungen bildet die Welt. Davon unterschieden ist die transzendente Narrativität – so die Bedingung der Möglichkeit von primären und sekundären Narrationen –, die im Werden des dreieinigen Gottes besteht.⁴

Sind diejenigen sekundärnarrativen Erzählungen, die die Lebensgeschichte von Menschen bilden, vom Evangelium bestimmt, handelt es sich um christliches Wahrwertnehmen. Der christliche Glaube ist also stets in der durch das Evangelium vermittelten Unmittelbarkeit ein unmittelbares Wahrwertnehmen.

Ein Beispiel: Im christlichen Glauben werden andere Menschen nicht zuerst als andere Lebewesen wahrgenommen und dann durch kognitive Interpretationen als Schwestern oder Brüder interpretiert, sondern sie werden unmittelbar als Schwestern und Brüder in konkreten Situationen wahrge-

² Vgl. Mühling, M., *Perceiving Values in the Story of the Gospel. A Sketch in 11 Theses*, in: Mühling, M./Förster, Y./Gilland, D. (Hg.), *Perceiving Truth and Value*, Göttingen/Bristol 2020, 117–132, hier 120–121.

³ MacIntyre, A., *After Virtue. A Study in Moral Theory*, London³2007, 212.

⁴ Vgl. Mühling, *Post-Systematische Theologie* (Fußnote 1), 550.

nommen. Wer auf der Autobahn unterwegs ist und einen Notleidenden am Rande erblickt, muss in der Lage sein, unmittelbar ohne allzu ausführliche kognitive Reflexion auf das Wahrgenommene zu reagieren, indem er eine notleidende Schwester und einen notleidenden Bruder oder aber einen Betrüger wahrwertnimmt – ansonsten wäre er schon kilometerweit entfernt.

Neben der vermittelten Unmittelbarkeit des christlichen Wahrwertnehmens ist wichtig, dass eine Perspektive nie als Standpunkt, auch nicht als Addition von Standpunkten verstanden werden darf, sondern als leibliche Perspektive, die stets eine Weglinienperspektive ist, d.h. eine Perspektive in Bewegung. Perspektiven bestehen nicht im Einnehmen von Standpunkten oder deren Wechsel, sondern im kontinuierlichen Verlassen eines jeden Standpunktes.⁵ Auch hier verhält es sich umgekehrt, wie meist angenommen: Nicht die Addition von Standpunkten ergibt bewegte Perspektiven, sondern Standpunkte selbst sind Abstraktionen von Weglinienperspektiven.

Auch dies lässt sich mit einem Beispiel erläutern: Die visuelle Wahrnehmung des Menschen ist immer bewegt. Bei einer Bergwanderung erscheint der wahrgenommene Berg nie in Ruhe. Lediglich die künstlich hergestellte Fotografie generiert den Standpunkt. Allerdings wird die Abbildung – selbst dann, wenn es sich um ein Panoramafoto handelt – kaum so eindrücklich sein wie der tatsächlich wahrwertgenommene Berg auf der eigentlichen Weglinienperspektive.

Während das Gesagte eigentlich für alle nicht-präparierten Perspektiven gilt, gilt für die Besonderheit des christlichen Wahrwertnehmens, dass alles im Lichte der trinitarischen Struktur des Evangeliums wahrwertgenommen wird.

Das bedeutet, verkürzt gesagt,

- dass alle Phänomene nicht als Gegebenheiten wahrwertgenommen werden, sondern als Gaben des Schöpfers, die in einem (guten) intern relationalen Zusammenhang werden. Sie werden also in einem Liebesgefüge bzw. sind zur Liebe bestimmt.

⁵ Vgl. Masschelein, J., E-ducating the gaze. The Idea of a Poor Pedagogy, Ethics and Education 5 (2010), 43–53.

- dass alle Phänomene als gefallen und zurechtbringungsbedürftig, d. h. als in ihrem Beziehungsgefüge gestört, wahrgenommen werden, aber als zurechtbringungsbedürftig. Dabei gilt, dass diese Zurechtrückung wie das Gegebensein selbst immer als Gabe verstanden wird, also nicht durch die Summe aller weltlichen Aktanten hinreichend erklärbar ist.
- dass alle Phänomene durch die Versöhnung Christi als des Sohnes tatsächlich zurechtrückbar sind und in der Hoffnung des Geistes auf eine eschatische Vollendung angelegt sind, die sich ebenso nicht aus innerweltlichen Prozessen erklären lässt.⁶

2. Freiheit in der theologischen Perspektive

Freiheit ist ein Begriff zweiter Ordnung. Er bezeichnet keinen einzelnen Sachverhalt, sondern es ist immer zu fragen, um welche Freiheit von wem oder zu was in Bezug auf was es sich handelt. Während einige Begriffe, wie der der Handlungsfreiheit, theologisch in der Gegenwart unserer nordwestlichen Welt weniger Probleme bereiten, sind andere, wie der der Willensfreiheit, hoch umstritten und werden auch theologisch unterschiedlich verstanden. Für theologische Freiheitsbegriffe ist im Allgemeinen gesprochen Folgendes wichtig:

- Da der Mensch nicht ein Individuum ist, das sich wie in den klassischen Vertragstheorien erst sekundär zu einer Gesellschaft assoziiert, sondern als Bild des dreieinigen Gottes selbst als Person in Beziehung zu verstehen ist, so dass partikulare Besonderheit und Bezogenheit stets gleichursprünglich sind,⁷ kann Freiheit theologisch nicht als „Freiheit von etwas“, sondern nur als „Freiheit zu etwas“ verstanden werden.
- Da der Mensch als Geschöpf immer gefallen, zurechtgebracht und auf Erlösung angelegt erscheint, kann Freiheit auch nicht als Bindungslosigkeit verstanden werden, sondern ist strenggenommen immer als Bin-

⁶ Vgl. Mühling, Post-Systematische Theologie (Fußnote 1), 549–578.

⁷ Vgl. Schwöbel, C., *Imago Libertatis. Freiheit des Menschen und Freiheit Gottes*, in: Schwöbel, C. (Hg.), *Gott in Beziehung*, Tübingen 2002, 227–256.

dungswechsel zu verstehen.⁸ Recht verstandene Freiheit ist keine konstruktivistische, sondern sie bezieht sich auf einen Wechsel der Bindungen: von der Wahrwertnehmung in anderen Medien als dem Evangelium (z. B. in naturalistischen oder ökonomistischen Geschichten) hin zur Wahrwertnehmung im Medium des Evangeliums, die als der Wirklichkeit resonierender als andere Wahrwertnehmungen verstanden wird.

- Aufgrund des Sachverhalts von Freiheit als Bindungswechsel kann ein freier Zustand auch nicht abstrakt als ein Zustand von höheren Wahlmöglichkeiten gegenüber weniger Wahlmöglichkeiten verstanden werden. Theologisch gesehen ist man nicht freier, wenn man mehr Wahlmöglichkeiten für die Zukunft hat (etwa im Studium) als wenn man weniger Wahlmöglichkeiten für die Zukunft hat (etwa im Altenpflegeheim), sondern wenn man sein Werden nicht als *Gegebenheit*, sondern als *Gabe* verstehen kann.⁹ Oder anders ausgedrückt: Wenn man im Lichte des Evangeliums sein *fate* als *destiny* (Hauerwas) begreifen kann.¹⁰ Oder noch anders ausgedrückt: Wenn man sich als Werdend im Abenteuer des Reiches Gottes wahrwertnimmt.

3. Digitalisierung in theologischer Perspektive

Mit Digitalisierung werden, wie schon bemerkt, zahlreiche divergente Phänomene bezeichnet. Ich selbst möchte hier ein sehr basales Verständnis von „Digitalisierung“ verwenden: Mir geht es um die Idee und Durchführung, *phänomenales Wahrwertnehmen bzw. phänomenales Werden, das stets kontinuierlich erfolgt, in digitalen, also diskontinuierlichen und diskreten, binär dargestellten Repräsentationen darzustellen*. Die sogenannte digitale Revolution, das Ausgreifen von digitaler Informationstechnologie auf immer mehr und unterschiedlichere Bereiche der Lebenswelt hat diesen Digitalisierungs-

⁸ Vgl. dazu Mühling, M., *Liebesgeschichte Gott. Systematische Theologie im Konzept*, Göttingen 2013, 258–301.

⁹ Vgl. ebd., 177–193.

¹⁰ Vgl. Hauerwas, S., *A Community of Character. Toward a Constructive Christian Social Ethic*, Notre Dame 1981, 10.

begriff zumindest als Ursprung und als informationstechnische Basis stets in sich. Gibt es eine Möglichkeit, dieses basale Digitalisierungsverständnis theologisch zu bewerten?

Um dies zu tun, muss zunächst eine andere Unterscheidung eingeführt werden, die ursprünglich gar keine theologische Unterscheidung war, sondern von dem Aberdeener Sozialanthropologen Tim Ingold stammt, die Unterscheidung von *wayfaring* und *transport*.¹¹ Was hat es damit auf sich?

Wenn Perspektiven immer Weglinienperspektiven sind, ist primär nicht vom Sein, sondern vom Werden zu sprechen. Allerdings gibt es verschiedene Arten des Werdens bzw. der Bewegung:

Transport meint eine Bewegung zwischen zwei oder mehreren Punkten a, b, ... n. Im *transport* sind die Punkte vorgängig und die Bewegung dazwischen nachgängig. In einer idealen Welt des *transport* würde das transportierte Gut instantan, ohne Geschwindigkeit von a nach b gelangen. Die eigentliche Bewegung ist im *transport* also mediatisiert, bzw. es handelt sich um eine Scheinbewegung. Entsprechend ist *transport* immer *intentional* in folgendem Sinne: Die einzelnen Punkte, bzw. zumindest der Zielpunkt, müssen *vor* der eigentlichen Wegabschreitung in der *intentio* feststehen. Was während des Weges erscheint, kann dann entweder nur als hilfreich oder als hinderlich für die Zielerreichung angesehen werden. Die Bewegung des *transport* in seiner Gesamtheit bildet ein *Netzwerk*, das etwa mit Hilfe von Graphen dargestellt werden kann. Das Wissen, das nötig ist, um *transport* beschreiben zu können, ist *klassifikatorisches* Wissen.¹²

Beim *wayfaring* hingegen verhält es sich umgekehrt: Hier ist die Bewegung vorgängig. Punkte erscheinen nur als Verknotungen von geschwungenen Linien. Sie sind auch nicht wirklich mathematische Punkte, sondern nur kleine verkrümmte Bewegungslinien. Der Bergwandernde bleibt auch im Sitzen auf einer Alm nicht wirklich erstarrt. Vor allem aber dienen diese im *wayfaring* aus dem Bewegen sich ergebenden Verknotungen nie als Endziel, sondern höchstens als vorläufiges Zwischenziel, das dem Weitergehen dient: Im *wayfaring* ist nicht nur die Bewegung Selbstzweck und nicht mediatisiert, sie ist

¹¹ Ingold, T., *Lines. A Brief History*, London/New York 2007, 77–84, 91–92 und dazu Mühling, *Post-Systematische Theologie* (Fußnote 1), 133–154.

¹² Vgl. Mühling, *Post-Systematische Theologie* (Fußnote 1), 147f.

auch nicht eigentlich intentional, sondern *attentional*. Mit Attentionalität ist dabei gemeint, dass der Verlauf der weiteren Bewegung sich aus dem Wahrwertnehmen der Umwelt ergibt und zusammen mit dieser geformt wird. *Wayfaring* bildet auch kein Netzwerk, sondern ein *Gewebe*. Das nötige Wissen, um *wayfaring* praktizieren zu können, ist auch kein klassifikatorisches Wissen, sondern *narratives* Wissen.¹³

Natürlich sind dieses beides Idealtypen der Bewegung. Sie können auch ineinander umgewandelt, *invertiert* werden. Allerdings in der Regel nicht auf sinnvolle Weise: Wenn der Postbote *wayfaring* statt *transport* betreiben würde, würde ich endlos auf mein Paket warten. Wenn umgekehrt ein Studierender sein Studium im Modus des *transport* versteht als Mittel zur Erlangung von „Abschlüssen“, wird er zu einem schlechten Wissenschaftler.

Theologisch interessant ist nun, dass das, was klassisch Sünde genannt wird – also eine nicht mit der Wirklichkeit resonierende Wahrwertnehmung im Lichte des Evangeliums – als eine solche Inversion dargestellt werden kann allerdings weniger als eine Inversion von *transport* in *wayfaring*, als vielmehr grundlegend als eine Inversion von *wayfaring* in *transport*. Denn wenn man davon ausgeht, dass der Mensch zur *imago dei* bestimmt ist, und diese als *imago trinitatis* immer eine *imago dilectionis* ist – wenn also das Werden des Menschen wesentlich durch Liebe bestimmt ist, dann ist ersichtlich, dass Liebe und *wayfaring* wesentliche Gehalte gemeinsam haben: In der Liebe wie im *wayfaring* ist der Andere, das Begegnende, nicht zu instrumentalisieren, sondern Selbstzweck. Die Liebe kennt kein Ende, wie das *wayfaring*. Liebe erfordert Attentionalität und nicht Intentionalität. Unsere sehr zu unterscheidenden Bewegungen der Liebe können nicht erschöpfend klassifiziert werden, aber sie können leicht narrativ erzählt werden. Kurz, eine wichtige Erscheinungsform der Sünde, des Falles, kann als *Inversion* von *wayfaring* in *transport* dargestellt werden.

Was hat dies nun mit Digitalisierung zu tun? Wenn Digitalisierung sich primär auf diskontinuierliche Repräsentationen von kontinuierlichem Werden bezieht, dann liegt auf der Hand, dass sich zwar Bewegungen des *transport*s vorzüglich digital darstellen lassen, Bewegungen des *wayfaring* allerdings

¹³ Vgl. Mühling, Post-Systematische Theologie (Fußnote 1), 151.

nicht. Versucht man es dennoch, wird man unweigerlich in einer *Inversion* von *wayfaring* in *transport* landen.

Vielleicht kann die Lebenswelt der Moderne und Postmoderne zu einem großen Teil als eine solche gesteigerte Inversion von *wayfaring* in *transport* dargestellt werden. Deutlich ist, dass die sog. digitale Revolution diese Bewegung in allen Lebensbereichen forciert.

Aus theologischer Perspektive und im Wahrwertnehmen des christlichen Glaubens sind diese Veränderungen unserer Lebenswelt aber keineswegs neutral. Da die Inversion selbst die Verrückung im geschaffenen Beziehungsgefüge bezeichnet, gilt dies auch für die entsprechenden Digitalisierungsbemühungen.

Die entscheidende Frage für die theologische Beurteilung von Digitalisierungsbemühungen lautet also: Wo bezieht sich die Digitalisierung auf lebensweltliche Phänomene, die an sich dem *transport* zuzurechnen sind, und wo bezieht sie sich auf Phänomene des *wayfaring*? Bezieht sie sich auf letztere, ist Digitalisierung wie die Inversion tatsächlich als problematisch zu bezeichnen.

4. Digitalisierung durch Freiheit?

Was bedeutet nun unser Konzept für das Verhältnis der Digitalisierung zur Freiheit? Nehmen wir einmal an, wir könnten bei allen Phänomenen *transport* und *wayfaring* eindeutig unterscheiden.

Beim *transport* würde Freiheit zweierlei bedeuten:

Zum einen würde Freiheit bedeuten, die Wege zwischen den Zielen zu verkürzen oder ganz zu eliminieren, da dann keine Abweichungen und keine Hemmnisse unserer intentionalen Zielwahl mehr entgegen stehen würden. Viele Formen der Digitalisierung ermöglichen genau dies. Ein Beispiel wäre, wenn digitale Kommunikation an die Stelle von klassischer Kommunikation oder gar an die Stelle von realen Reisen zu Konferenzen tritt, wie wir es in der Coronakrise alle erleben. Auch hier ist die auf *transport* beruhende Digitalisierung nicht unproblematisch. Reisen als Erholungszeit fällt weg, leibliche Kommunikation ebenfalls.

Zum anderen würde im *transport* Freiheit eine Steigerung der Wahlmöglichkeiten, also der Auswahl zu erreichender Punkte, bedeuten. Auch hier ist vorstellbar, dass Digitalisierung die Wahrnehmung mehrerer Zielwahlen in kürzerer Zeit ermöglicht. Allerdings ist auch in einer idealen Welt des *transport* beides zweischneidig. Martin Wendte hat in Anlehnung an Heideggers Technikkritik darauf hingewiesen,¹⁴ dass Modernisierungserscheinungen mit einer Vertauschung von räumlichen und zeitlichen Ressourcen einhergehen. Während in vormodernen Gesellschaften räumliche Distanzen diejenige Schwierigkeit darstellen, die freiheitshemmend war, wurde dieses Freiheitshemmnis durch die zur Verfügung stehende Zeit ausgeglichen. Luthers Romreise 1510/11 oder 1511/12 dauerte beispielsweise aufgrund der Entfernung Erfurts von Rom über ein Jahr, wobei die eigentlichen Geschäfte nur wenige Tage in Anspruch nahmen.¹⁵ Unter Modernitätsbedingungen verhält es sich genau umgekehrt: Die Informations- und Kommunikationstechnologie lässt räumliche Distanzen als Freiheitshemmung nahezu verschwinden. Demgegenüber wird aber nun die Handlungszeit für Projekte zum Problem: Sie wird knapp und hemmt dadurch die möglichen Verwirklichungen von Projekten wieder.

In Phänomenen des *wayfaring* hingegen hat Freiheit eine andere Bedeutung. Sie hat einerseits die Bedeutung, *attentional* adäquat reagieren zu können und das eigene Werden diesen attentionalen Erfordernissen anpassen zu können. Es ist gewissermaßen *die Freiheit, Umwege gehen zu können* – ohne dass man den Umweg noch als Umweg empfinden würde, denn Umwege gibt es im *wayfaring* strenggenommen gar nicht. Es ist die Freiheit, sich binden und verpflichten zu lassen. Das aber erfordert Zeit. Selbst wenn alle Phänomene des *wayfaring* nicht invertiert wären, wäre klar, dass der Freiheitsgewinn im Bereich des *transport* einem Freiheitsgewinn im Bereich des *wayfaring* entgegenstände. Beachtet man aber, dass zunehmend mehr Phä-

¹⁴ Vgl. Wendte, M., Raumzeitimplosion und Schöpfungsfülle. Soziologisches, Philosophisches und Theologisches zur gesellschaftlichen Gestaltung von Raum und Zeit in der Spätmoderne, in: Gräb-Schmidt, E. (Hg.), Was heißt Natur? Philosophischer Ort und Begründungsfunktion des Naturbegriffs, Leipzig 2015, 55–81.

¹⁵ Vgl. zu Luthers Romreise u. a. Schneider, H., Martin Luthers Reise nach Rom – neu datiert und neu gedeutet, Berlin u. a. 2011.

nomene des *wayfaring* technologisch in *transport* invertiert werden, wird deutlich, dass hiermit ein immenser *Freiheitsverlust* einhergeht.

Andererseits bedeutet Freiheit im Bereich des *wayfaring* aber auch, diejenigen narrativen Mittel zu haben, dank derer man die attentionale Zielwahl nicht als *fate*, sondern als *destiny* verstehen kann. Faktisch heißt das, dass man sein Werden nicht als ohne Absicht hin- und hergetrieben, sondern als gnadenhafte Bestimmung in einem größeren Rahmen – dem Abenteuer des Reiches Gottes – erfährt. Wichtig ist nun, dass das Geschenk der Gnade dabei stets mehr Möglichkeiten gibt, als man sich ohne die Gabe der Gnade im Medium des Evangeliums jemals intentional erträumt hätte. Wer etwas geschenkt bekommt, etwa ein überraschendes Erbe, der hat mannigfache neue Möglichkeiten: die Möglichkeit, das Erbe anzunehmen oder abzulehnen, sowie zahlreiche weitere Wege, die sich ohne dieses Geschenk nicht ergäben. Das ist, worauf Ingolf Dalferth richtig hinweist, die eigentliche Pointe der Gratuität der Gnade.¹⁶ Um diese Gnade – und damit den Freiheitsgewinn – aber wirklich in der Verschränkung der Lebensgeschichten, der eigenen Lebensgeschichte mit der des Evangeliums, erfahren zu können, ist es wichtig, Zeit und Muße für die attentionale Werdensbestimmung des eigenen Lebens zu haben. Geschenke zu erhalten erfordert, nicht davonzurennen. Es erfordert, seine Narration der eigenen Bildung mit der des Schenkenden verschränken zu lassen – und zwar jenseits von allen intentionalen Absichten und Zweckrationalitäten. Gnade ist immer nur im Modus des *wayfaring* denkbar – nie im Modus des *transport*. Sie ist daher prinzipiell nicht digitalisierbar.

Wenn Digitalisierung eine Freiheitssteigerung im Sinne des *transport* erreichen kann (aber nicht muss), dann ist dieses Phänomen immer noch ambivalent. Selbst im positiven Sinne ist dieser Freiheitsgewinn aber nie ein Selbstzweck. Er muss nach der Logik des *transport* einem Ziel dienen. Ist dieses Ziel aber selbst wieder im Netzwerk des *transport* aufgehoben, wird letztlich Freiheit drastisch reduziert. Nur wenn dieser digitale Freiheitsgewinn sein Ziel in Gelegenheiten für mehr nicht-intentionale und nicht-digitalisierbare Bewegungen des *wayfaring* findet und so der zunehmenden

¹⁶ Vgl. Dalferth, I.U., Umsonst. Eine Erinnerung an die kreative Passivität des Menschen, Tübingen 2011, 110.

Inversion aller Lebensbereiche ins Digitale des *transport* gerade entgegen wirken könnte, wäre es tatsächlich ein Freiheitsgewinn.

Autoren

Prof. Dr. Martina Hessler

Technische Universität Darmstadt, Institut für Geschichte,
Abteilung für Technikgeschichte

Prof. Dr. Gerhard Lakemeyer

Rheinisch-Westfälische Technische Hochschule Aachen,
Lehr- und Forschungsgebiet Wissensbasierte Systeme/kognitive Robotik

Prof. Dr. Gregor Schöner

Ruhr-Universität Bochum, Institut für Neuroinformatik

Prof. Dr. Andreas Schuppert

Rheinisch-Westfälische Technische Hochschule Aachen,
Aachen Institute for Advanced Study in Computational Engineering Science

Dr. Philipp Schönthaler

Literaturwissenschaftler und Schriftsteller, Konstanz

Prof. Dr. Dan Verständig

Otto-von-Guericke-Universität Magdeburg,
Fakultät für Humanwissenschaften, Institut I: Bildung, Beruf und Medien,

Prof. Dr. Axel Siegemund

Rheinisch-Westfälische Technische Hochschule Aachen,
Institut für Katholische Theologie, Fachbereich Grenzfragen von Theologie,
Naturwissenschaft und Technik

Prof. Dr. Markus Mühling

Kirchliche Hochschule Wuppertal/Bethel, Systematische Theologie

Dr. Frank Vogelsang

Evangelische Akademie im Rheinland

Dr. Georg Souvignier

Bischöfliche Akademie des Bistums Aachen

Wissen verbindet uns

Die wbg ist ein Verein zur Förderung von Wissenschaft und Bildung. Mit 85.000 Mitgliedern sind wir die größte geisteswissenschaftliche Gemeinschaft in Deutschland. Wir bieten Entdeckungsreisen in die Welt des Wissens und ein Forum für Diskussionen. Unser Fokus ist nicht kommerziell, Gewinne werden reinvestiert.

Wir wollen Themen sichtbar machen, die Wissenschaft und Gesellschaft bereichern. In unseren Verlags-Labels erscheinen jährlich rund 120 Publikationen, darunter viele Werke, die ansonsten auf dem Buchmarkt nicht möglich wären. Wir bieten außerdem Zeitschriften, Podcasts und die wbg-KulturCard. Seit 2019 vergeben wir den höchstdotierten deutschsprachigen WISSEN!-Sachbuchpreis.

Vereinsmitglieder fördern unsere Arbeit und genießen gleichzeitig viele Preis- und Kulturvorteile.

**Werden auch Sie wbg-Mitglied.
Zur Begrüßung schenken wir Ihnen ein
wbg-Buch Ihrer Wahl bis € 25,-**

Mehr Infos unter wbg-wissenverbindet.de
oder rufen Sie uns an unter 06151 3308 330



wbg Wissen
Bildung
Gemeinschaft

Digitale technische Systeme waren von Beginn an mit Freiheitsversprechen verbunden. Doch lösen sie ein, was sie versprechen? In welcher Hinsicht erhöhen sie Freiheiten der Nutzer und der Gesellschaft – und was ist der Preis dafür? Gewinnen die digitalen Artefakte, etwa Systeme künstlicher Intelligenz, eine Freiheit, die der des Menschen analog ist? Dieser Band beleuchtet das Verhältnis von Technik und Freiheit aus technischen wie auch geisteswissenschaftlichen Perspektiven.

Dr. Frank Vogelsang, Dipl.-Ing und promovierter Theologe ist Direktor der Evangelischen Akademie im Rheinland. Er hat zur ethischen Bewertung der technischen Entwicklung veröffentlicht.

Dr. Georg Souvignier ist promovierter Physiker und Dozent für Grenzfragen von Theologie und Naturwissenschaft in der Bischöflichen Akademie des Bistums Aachen.

www.wbg-wissenverbindet.de
ISBN 978-3-534-40603-6



wbg Academic