

1 Informationstechnik und Menschenbild

If a machine is expected to be infallible, it cannot also be intelligent.

Alan M. Turing²

1.1 Informationelle Autonomie in Gefahr

Seit mehreren Jahrzehnten werden wir regelmäßig mit Charakterisierungen für die Gesellschaftsform eines neuen Zeitalters konfrontiert, das als Ergebnis von Innovationen in der Informationstechnik vermeintlich angebrochen ist. Typische Etikettierungen dafür sind Informationsgesellschaft, Wissensgesellschaft oder sogar die *Google*-Gesellschaft. Ein Telekommunikationskonzern bringt es so auf den Punkt:

Wir treiben die Entwicklung zur Gigabit-Gesellschaft voran. Weil wir glauben, dass die Digitalisierung unser Leben positiv verändern wird.³

Jede dieser angeblich neuen Gesellschaftsformen enthält auch Verheißungen für den Menschen. Fortschritte im Bildungsbereich, im Arbeitsleben und für die Freizeit werden prognostiziert. Ein dabei immer mit gedachter Trend zum ‚schlauer werden‘ wird nur zu gern billigend in Kauf genommen.

Bevor jedoch recht klar wird, welche konkreten Konsequenzen sich für den einzelnen Menschen ergeben werden – Anmutungen und Anforderungen erschließen sich eher indirekt –, wird das eine Modell schon durch das nächste abgelöst. Ein halbwegs stabiles Bild vom mündigen Menschen in der informationstechnisch geprägten Zivilgesellschaft hat sich trotz breit gefächerter Diskussionen nicht entwickelt. Die Vorstellungen von der zukünftigen Entwicklung des Menschen wechseln wie Modeerscheinungen. Gemeinsam ist den Modellen bislang noch die Fokussierung auf den Menschen als Individuum und damit die Überzeugung, dass die Strukturen der Gesellschaft durch handelnde Subjekte gestaltet werden. Die Veränderungen sollen Erleichterungen und Verbesserungen bringen, neue Menschenbilder positive Einflüsse haben.

Das könnte sich jedoch ändern. Das nächste Zeitalter könnte mit grundsätzlichen Umkehrungen traditioneller Vorstellungen verbunden sein, denn Modeströmungen (Pseudo-Philosophien?) wie Dataismus oder Transhumanismus betrach-

² Turing: Lecture to the London Mathematical Society on 20 February 1947, S. 124.

³ Aus einer Werbemail des Telekommunikationsanbieters *Vodafone* vom 24.10.2017.

ten den Menschen eher als funktionalisiertes Objekt denn als gestaltendes Subjekt. Seine Legitimation erhält der Mensch hier nur noch aus seinen Beiträgen für den Datenstrom des universellen Netzwerks und deren Auswertung zur Ableitung von Verhaltensregeln. Erstaunlich ist dabei nicht, dass solche Ideen entwickelt und verbreitet werden, erstaunlich ist eher die Resonanz, die diese als gleichsam unausweichliche Folge eines informationstechnischen Geschehens wertet. In dem als Kultbuch eines neuen Zeitalters angepriesenen Buch *Homo Deus* entwirft *Yuval Noah Harari* die wenig freundliche Zukunftsvision unter den Bedingungen der Datenreligion:

Der Dataismus ist weder liberal noch humanistisch. Er ist deshalb freilich keineswegs anti-humanistisch. Er hat nichts gegen menschliche Erfahrungen. Er glaubt nur nicht, dass sie für sich genommen einen Wert haben.⁴

Dem Dataismus zufolge besteht das Universum aus Datenströmen, und der Wert jedes Phänomens oder jedes Wesens bemisst sich nach seinem bzw. ihrem Beitrag zur Datenverarbeitung.⁵

Der Dataismus ist die erste Bewegung seit 1789, die einen wirklichen neuen Wert geschaffen hat: die Freiheit der Information. Diese Informationsfreiheit dürfen wir nicht mit dem alten liberalen Ideal der Meinungsfreiheit verwechseln. Meinungsfreiheit wurde den Menschen gewährt und schützte ihr Recht, zu denken und zu sagen, was sie wollten – dazu gehörte auch das Recht, nichts zu sagen und seine Gedanken für sich zu behalten. Informationsfreiheit dagegen wird nicht Menschen gewährt, sondern der Information.⁶

Unabhängig davon, wie viel Bedeutung oder realistische Zukunftsprognose man diesen Ansichten beimisst, eine Auseinandersetzung damit scheint geboten. In der Bewertung sind sich Für- und Widersprecher grundsätzlich uneinig. Die Polarisierung ist dabei angesichts der wenig greifbaren Substanz der Vision sehr ausgeprägt. Die Fürsprecher sehen in ihr den endgültigen Sieg der Demokratie. Da im Datenstrom alles und jeder transparent wird, wird angeblich allen Übeln unserer Zeit die Geschäftsgrundlage entzogen. Die Vision wird zur Heilsbotschaft. Die Widersprecher sehen mit solchen Vorstellungen eher den Beginn eines neuen Totalitarismus beschrieben, in dem es keine Freiheitsrechte eines handelnden Subjekts mehr gibt, sondern nur funktionalisierende Unterwerfung unter Massengeschmack und intransparente Algorithmen.⁷

⁴ Harari: *Homo Deus*, S. 524.

⁵ Harari: *Homo Deus*, S. 497.

⁶ Harari: *Homo Deus*, S. 517.

⁷ Vgl.: Pauen/Welzer: *Autonomie*.

In seinem Buch *Leben 3.0* beschäftigt sich *Max Tegmark* mit künftigen Lebensbedingungen unter den Einflüssen der Künstlichen Intelligenz. Er glaubt, dass Künstliche Intelligenz in der Lage ist (oder sein wird), Entscheidungen auf rationaler Basis unter Beachtung moralischer Bewertungen vorzunehmen. Damit gäbe es für den Menschen zum ersten Mal eine Konkurrenz auf einem Gebiet, für das er bislang das Monopol beanspruchte. Für diese Welt unter den Bedingungen der Künstlichen Intelligenz entwirft *Tegmark* verschiedene Szenarien:

Eroberer:

Künstliche Intelligenz übernimmt die Macht und entledigt sich der Menschheit mit Methoden, die wir noch nicht einmal verstehen.

Der versklavte Gott:

Die Menschen bemächtigen sich einer superintelligenten künstlichen Intelligenz und nutzen sie, um Hochtechnologien herzustellen.

Umkehr:

Der technologische Fortschritt wird radikal unterbunden und wir kehren zu einer prä-technologischen Gesellschaft im Stil der Amish zurück.

Selbsterstörung:

Superintelligenz wird nicht erreicht, weil sich die Menschheit vorher nuklear oder anders selbst vernichtet.

Egalitäres Utopia:

Es gibt weder Superintelligenz noch Besitz, Menschen und kybernetische Organismen existieren friedlich nebeneinander.⁸

Bei *Tegmark* fehlt ein Szenario, in dem eine Gruppe von Menschen den Verlockungen erliegt und sich aus Bequemlichkeit in eine selbst erzeugte Informationelle Abhängigkeit begibt und eine andere Gruppe dieses Verhalten im Rahmen totalitärer Strukturen zu ihrem Nutzen funktionalisiert. Dieses Szenario einer neuen Elitenbildung wird von *Harari* berücksichtigt:

Die dritte Bedrohung für den Liberalismus besteht darin, dass einige Menschen sowohl unentbehrlich als auch unentschlüsselbar bleiben, aber sie werden eine kleine und privilegierte Elite optimierter Menschen bilden. Diese Übermenschen werden über unerhörte Fähigkeiten und beispiellose Kreativität verfügen, was sie in die Lage versetzen wird, viele der wichtigsten Entscheidungen auf der Welt zu treffen. [...] Die meisten Menschen jedoch werden eine

⁸ Tegmark: *Leben 3.0*, Text zitiert nach Umschlagseite 1; vgl. für die vollständige Liste der Szenarien: S. 243–244.

solche ‚Aufwertung‘ nicht erleben und folglich zu einer niederen Kaste werden, die von den Computeralgorithmen ebenso beherrscht wird wie von den neuen Übermensch.⁹

Auch wenn man die Ansichten von *Harari* und *Tegmark* nicht teilt, sie möglicherweise sogar für überspannte Gedanken Technologie getriebener Fantasten hält, muss man doch zur Kenntnis nehmen, dass es viele solcher und ähnlicher Darstellungen gibt. Prognosen über die zu erwartenden Veränderungen durch die Informationstechnologie haben Konjunktur. Dabei kommt es häufig zu einer Art von gedanklichem Kurzschluss zwischen vier Konzepten:

Information – Informationstechnik – Fortschritt – Zukunft

Das erinnert fatal an den Streit der zwei Wissenskulturen bei *C. P. Snow* – die Naturwissenschaften auf der einen, die Literatur- und Geisteswissenschaften auf der anderen Seite –, dessen Essay *Die zwei Kulturen* in dem berühmt gewordenen Satz kulminiert: „Die Naturwissenschaftler haben die Zukunft im Blut.“¹⁰

Die Gestaltung der Zukunft erfordert Ideen. Solche Ideen wurden zur Entstehungszeit des Essays (basierend auf einem Vortrag von 1959) durch die Naturwissenschaften stärker eingebracht als durch andere Disziplinen. Das kann sich wieder ändern. Entscheidend ist, ob und welche Ideen zur Lösung von Zukunftsproblemen vorhanden sind. Damit diese entwickelt werden können, darf die Fähigkeit nicht verloren gehen, zwischen zu lösenden Zukunftsproblemen und weniger wichtigen Problemstellungen zu unterscheiden.

Heute scheint allein die Erwähnung eines bildungsbezogenen Informationsverständnisses schon auszureichen, um als Bewahrer oder Bedenkenträger und nicht als Fortschrittsgestalter gesehen zu werden. Wer möchte sich schon in der Rolle einer Zukunftsbremse sehen, wenn doch nahezu überall die Begriffe Zukunft und Informationstechnik gleichsam synonym verwendet werden? Dann orientiert man sein Verständnis doch lieber an einem Konzept, das man vielleicht gar nicht versteht, das einem aber die Teilhabe an der Zukunft verspricht. Vor allem in einer Zeit, in der sich Zukunft mehr und mehr über den Faktor Renditeerwartung definiert.

Da Informationstechnologie durch den Einsatz Computer basierter Vorgehensweisen ermöglicht wird, kommt diesem Werkzeug eine fast mythische Bedeutung zu. Die Computermetapher wird zum Ausdruck von IT-Gläubigkeit. Ist dies ausreichend für die Gestaltung der Zukunft? Eher nicht. Betrachtet man Zukunftsgestaltung als menschliche Aufgabe, so ist es kaum angebracht, die kognitiven

⁹ Harari: *Homo Deus*, S. 467; ähnlich auch: S. 472–473, 476–477, 490.

¹⁰ Kreuzer (Hrsg.): *Die zwei Kulturen*, S. 16.

Informationsprozesse als Bedingungen des Denkens und Handelns einem Leitbild technischer Informationsverarbeitung zu unterwerfen.

Vielfach wird ein Vokabular verwendet, das die Position des Menschlichen entweder als bedrohlich oder verheißungsvoll relativiert. Es ist dabei etwa die Rede von der Mensch-Maschine-Konvergenz, vom Gefühlscomputing als Befreiungstechnologie oder von einem ‚affective computing‘ ohne Empathie.¹¹ In der Regel wird zum Ausdruck gebracht, dass der Mensch durch die fortschreitende Entwicklung der Informationstechnologie sein vormaliges Alleinstellungsmerkmal zum Erbringen höherwertiger kognitiver Leistungen verlieren wird oder bereits verloren hat. Das Spektrum der dabei berücksichtigten Leistungen beschränkt sich nicht mehr nur auf die rationale Intelligenz, sondern umfasst inzwischen auch die Gefühlsebene und das Bewusstsein bis hin zur Seele. Es scheint nur eine Frage der Zeit zu sein, bis das derzeit für den Menschen noch respektierte Alleinstellungsmerkmal der Einheit von Rationalität, Gefühl und Bewusstsein vollständig aufgegeben und durch die einzelnen Bestandteile ersetzt oder als Merkmal auch Maschinen zugeschrieben wird.

Besonders bedenklich ist es, wenn die maschinellen Beschreibungsmerkmale als Computermetapher verwendet und zur Charakterisierung menschlicher kognitiver Leistungsfähigkeit genutzt werden. Insbesondere im Kontext des KI-Gedankens einer Nachbildung der Funktionen des menschlichen Gehirns werden kognitive Vorgänge auf der Basis eines Modells mit funktionaler Trennung in Hard- und Software beschrieben.¹² Diese Vorstellung wird durch die aktuelle Hirnforschung nicht gestützt. *Thomas Metzinger* hält fest:

Die Computermetapher des menschlichen Geistes ist tot. Die Idee, dass Geist und Gehirn sich wie Software und Hardware zueinander verhalten und als zwei Ebenen klar voneinander getrennt werden können, vertritt in der Kognitionswissenschaft niemand mehr.¹³

Katrin Amunts (Hirnforscherin und Chair of the Science and Infrastructure Board of the *Human Brain Project*) sagt zur gleichen Frage in einem Interview vom 18.02.2017:

Wety: Der Vergleich drängt sich ja auf: Gehirn und Computer, das sind die Hardware, aber was genau macht die Software des Menschen aus? Ist das so etwas wie Intelligenz, Geist, Bewusstsein?

11 Vgl. etwa: Benedikter: Digitalisierung der Gefühle?

12 Vgl. für ein aktuelles Beispiel: Thielicke/Helmstaedter: Ein völlig neues Kapitel der Künstlichen Intelligenz.

13 Metzinger: Ist das Gehirn mit einem Computer oder einer Festplatte zu vergleichen?

Amunts: Also diese Vergleiche zwischen dem Computer und dem Gehirn, die scheinen natürlich erst mal ganz toll und sehr naheliegend zu sein, aber das sind schon zwei völlig unterschiedliche Dinge. Ein Computer ist dafür gemacht, dass er etwas berechnet, basierend eben auf mathematischen Modellen, basierend auf bestimmten Algorithmen. Das Gehirn, das ist ein Teil von uns, von uns Menschen, und wir sind ja nicht nur das Gehirn. Wir haben ja noch eine gewisse Peripherie: Wir haben Arme, Beine, Organe – all das macht uns ja als Persönlichkeit aus, und das Organ Gehirn ist darin eingebettet und hat sich über eine ganz lange Evolution entwickelt mit ganz vielen vielleicht auch Irrwegen, nicht immer geradeaus, und das ist also etwas, was keine Maschine ist, die da ist, sondern die sich eben entwickelt hat. Es ist schon ziemlich kompliziert, wenn wir allein formulieren wollen, was denn uns mit unseren kognitiven oder emotionalen Fähigkeiten, was uns wirklich ausmacht, und das finden wir in einem Computer im Moment noch gar nicht realisiert.¹⁴

Die Betrachtung der Rolle, die Gefühle und Empfindungen für die Grundlagen von Denken und Bewusstsein spielen, liefert eine weitere Begründung dafür, dass die Analogien Hardware–Gehirn und Software–Geist unzutreffend sind. Die folgenden Fragen lassen sich unter der Annahme einer Trennung von Hard- und Software wenig sinnvoll stellen:

Wie fühlt es sich an, ein Bier zu trinken?
 Wie fühlt es sich an, Olympiasieger zu sein?
 Wie fühlt es sich an, einen mathematischen Satz zu beweisen?

Welche Empfindung löst blau aus?
 Welchen Geschmack hat blau?
 Welchen Geruch hat grün?
 Welches Gefühl verbindet sich mit nass?

Gleiches gilt für Einschätzungen, Überzeugungen, Vorlieben und Wünsche; allgemein für das gesamte Spektrum qualitativer Erlebnisse. Dabei ist die konkrete Antwort (oder Aussage) gar nicht so wichtig, auch nicht, ob andere dieselbe Antwort geben würden, sondern, dass die Frage einen Prozess auslöst, für den das eigene Erleben in einer Einheit aus Körper und Geist Voraussetzung ist, um die Antwort zu finden.

Damit müssen alle Konzepte als höchst fragwürdig eingestuft werden, die von einem Austausch von Gedächtnisinhalten zwischen natürlichen und künstlichen Systemen oder der Möglichkeit einer direkten Verbindung zwischen beiden ausgehen. Daran können auch noch so eindrucksvoll medial in Szene gesetzte Darstellungen nichts ändern, wie sie etwa in dem Film *Transcendence* mit *Johnny Depp* aus dem Jahr 2014 zu finden sind. Die Handlung geht davon aus, dass all das, was wir gelernt und durchdacht haben, digitalisiert und auf einen externen

¹⁴ Amunts/Welty: EU-Projekt zur Hirnforschung.

Großspeicher hochgeladen werden kann. Wenn zu einem späteren Zeitpunkt der Körper wiederhergestellt wird, kann das zwischengespeicherte Gedächtnis wieder heruntergeladen werden; das Gedächtnis überlebt so das physische Ableben.¹⁵

Das Genre Film hat sich schon seit längerem der Thematik angenommen. *Rainer Werner Fassbinders Welt am Draht* war eindeutig von der kritischen Auseinandersetzung mit den menschlichen und sozialen Konsequenzen einer Computerisierung der Lebens- und Arbeitswelt geprägt und steht damit eher in der Tradition von *Fritz Langs Metropolis* oder *Chaplins Modernen Zeiten*. Spätere Produkte antizipieren eher unkritisch oder zustimmend die Potenziale der Technik. In *Die Frauen von Stepford* machen Männer ihre Frauen mit Mikrochips im Hirn steuerbar wie Roboter.¹⁶ Ein weiteres Beispiel liefern die Filme der *Matrix*-Reihe.¹⁷

Massive Spuren der Antizipierung der Computermetapher sind bereits in der Berufswelt zu finden, wenn etwa Stellenbewerber danach gefragt werden, was sie denn besser könnten als ein Computer. Im Kern wird bei dieser Frage eine Defensiv-Situation durch Herstellung eines Kontexts aufgebaut, der dem Computer die Rolle eines Subjekts und nicht mehr nur eines Werkzeugs zuweist. Warum sonst fragt man pauschal danach, was jemand besser könne als ein Werkzeug?

Woher rührt die Faszination am Transhumanen? Sich als Schöpfer zu fühlen und darin die Steigerung der Möglichkeiten der eigenen Potenz zu sehen? Allein die durchaus verständliche Lust am Erkenntnisgewinn kann es doch nicht sein, wenn man all die Selbstaufgabe- und Untergangsszenarien bedenkt, die angeboten werden. Ist es attraktiv, die evolutionär entstandene menschliche Rationalität durch eine selbst geschaffene transhumane Rationalität zu ersetzen?

Im realen Leben könnten sich solche Konstellationen als totalitäre Gesellschaftsstrukturen manifestieren. Woher kommt das Argument, dass transhumane Rationalität besser als die evolutionär entstandene Rationalität geeignet sei, den Fortbestand des Lebens zu sichern? Ist das nicht eher ein Eingeständnis der eigenen Unzulänglichkeit?

Pioniere der Künstlichen Intelligenz – und dabei muss nicht nur an ausgewiesene Kritiker wie *Josef Weizenbaum* oder *Hubert und Stuart Dreyfus* gedacht werden – haben sich durchaus zurückhaltend dazu geäußert, was man von ihr erwarten dürfe. So erklärt etwa *Roger Schank*:

¹⁵ Vgl. für die Handlung des Films: *Transcendence* (Film) [Wikipedia]; vgl. auch: Koch: Bewusstsein ohne Gehirn.

¹⁶ *Die Frauen von Stepford* (1975) [Wikipedia].

¹⁷ Weitere Beispiele enthält: Nida-Rümelin/Weidenfeld: *Digitaler Humanismus*.

Der allgemeine Ansatz der Künstlichen Intelligenz sollte heutzutage [1984, Anm. d. Verf.] sein, die Denkprozesse zu entdecken, die Menschen für verschiedene intelligente Aktivitäten verwenden, um Computer dafür zu programmieren, diese Prozesse durchzuführen.¹⁸

Und an anderer Stelle:

Um Computer richtig einschätzen zu können, müssen wir die Menschen betrachten.¹⁹

Wir sollten den Menschen also gerade nicht danach bewerten, wie weit er den Vorstellungen der Künstlichen Intelligenz gerecht wird. So ist es bezeichnend, dass der Hype und die überzogenen Erwartungen vielfach von Personen befördert werden, die *Roger Schank* so charakterisiert:

Ist Künstliche Intelligenz nur ein Thema für schwachbrüstige Akademiker, oder wird jedermann ein wenig von Künstlicher Intelligenz und Computern im allgemeinen verstehen müssen, um auch nur mithalten zu können? [...] Das sind die Fragen, die heutzutage in Zeitungen und auf Cocktailparties diskutiert werden. Einerseits ist es prima, daß die Leute sich um solche Dinge kümmern. Andererseits wirkt es ein bißchen komisch, daß so viele Leute ihre Meinung zu einem Thema diskutieren wollen, über das sie so wenig konkrete Information besitzen.²⁰

Wie anders lesen sich da heutige Äußerungen, wenn etwa *Wolfgang Wahlster*, CEO des *Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI)*, davon spricht, dass Künstliche Intelligenz in der Rolle eines Hilfswissenschaftlers die Forschung revolutioniert. Als eines ihrer zentralen Einsatzfelder sieht er die Mathematik, Menschen sollten jedoch weiterhin die Theorie machen.²¹ Nicht weniger gruselig ist die Prognose kommender digitaler Kriege durch den Einsatz Künstlicher Intelligenz.²²

Es mag schon immer umstritten gewesen sein, ob der Mensch eine Seele besitzt, durch die er sich von allen anderen Lebewesen unterscheidet. Es mag auch unklar sein, ob und gegebenenfalls welche physiologischen Prozesse für die Seele verantwortlich sind. Verbreitete Überzeugung ist jedoch, dass es psychische Verfasstheiten gibt, die nicht mit algorithmischem Denken vereinbar sind und dadurch nicht erklärt werden können. Dabei braucht man gar nicht an pathologische Befunde zu denken. Selbst bei größtmöglicher Sympathie für rationale Denk- und Entscheidungsprozesse wird sich jeder Mensch an Gefühle erinnern, die er nicht

¹⁸ Schank/Childers: Die Zukunft der künstlichen Intelligenz, S. 53.

¹⁹ Schank/Childers: Die Zukunft der künstlichen Intelligenz, S. 9.

²⁰ Schank/Childers: Die Zukunft der künstlichen Intelligenz, S. 11.

²¹ Reichert: Künstliche Intelligenz als Hilfswissenschaftler.

²² Springer: Wettrüsten im Cyberraum.

missen möchte, vielleicht auch an solche, auf die er sehr gut verzichten könnte, was aber nicht gelingt.

Bevor der Weg in eine Informationelle Unmündigkeit als attraktiv empfunden wird, wäre zu prüfen, ob nicht Informationelle Autonomie derartig fest an die Eigenschaften der menschlichen Grundausstattung gebunden ist, dass sie gar nicht aufgegeben werden kann. Zwei Bezugssysteme müssen dabei miteinander verbunden werden: erstens, die Bindung des Menschen als soziales Lebewesen an Gemeinschaften gemeinsamen Handelns und zweitens, die dem Einzelnen eigene kognitionspsychologische Basis des Denkens und Handelns.

1.2 Rationalität und Kognition

Die Aufklärung hat der Idee den Weg geebnet, den Menschen unabhängig von seiner Herkunft und Zugehörigkeit zu einer sozialen Schicht als ein Lebewesen mit Ich-Verständnis zu sehen. Sein Denken und Handeln sind so durch rationale Überlegungen geprägt, dass Emotionen beherrscht werden können. Diese Idee hat sich in den modernen Demokratien und freiheitlichen Zivilgesellschaften fortgesetzt, die ihre Ausgestaltung an das Verständnis des mündigen Subjekts binden. Psychologie und moderne Kognitionsforschung haben zur genaueren Einsicht beigetragen, wie Rationalität, biologische Grundfunktionen und emotionale Faktoren zum konsistenten Verständnis eines bewusst und autonom denkenden und handelnden Individuums zusammen wirken; eines Individuums, das zur Gestaltung von und zur Teilhabe an gesellschaftlichen Strukturen fähig ist. Gehen wir von der Prämisse aus, dass diese Errungenschaften grundsätzlich bewahrt und fortentwickelt werden sollen, stellt sich die Frage, welches Ausmaß an Informationeller Autonomie des handelnden Subjekts dafür erforderlich ist.

Rationales Denken und Handeln sind jedoch nicht die einzigen Anlässe für die Prozesse menschlicher Informationsverarbeitung. Auch Empfindungen und Gefühle, Absichten und Zweifel stehen mit Sinneseindrücken und Handlungen in Verbindung, die Vorgänge der kognitiven Informationsverarbeitung auslösen und durch sie gesteuert werden. Wir beschränken uns hier darauf, der Darstellung ein rationales Wissensverständnis zugrunde zu legen, wollen dabei aber im Blick behalten, dass auch nicht-rationale Vorgänge Ursachen und Ergebnisse für kognitive Informationsverarbeitung sind.

Menschen sind informationsverarbeitende Wesen, sowohl bei individuellen Denk- und Handlungsprozessen als auch in Interaktionen mit Kommunikationspartnern. Die Ansichten über die Art der menschlichen Informationsverarbeitung sind bemerkenswert unterschiedlich und haben im Laufe der geschichtlichen Entwicklung einige Wandlungen erfahren. Entscheidende Größen des Prozesses sind

Sinneswahrnehmungen und deren nachfolgende kognitive Verarbeitung. Dabei spielen der Vergleich mit bereits vorhandenen Strukturen, deren Anpassung sowie der Aufbau neuer Strukturen eine zentrale Rolle für alle mit Wissen verbundenen Leistungen.

Ein Modell der Wissensaneignung im Rahmen des Lösens von Aufgaben lässt sich wie in Abbildung 1.1 skizzieren.²³ Es stellt einen Zyklus dar, in dem externalisiertes Wissen aus einem Informationssystem oder kognitives Wissen eines Experten zur Lösung einer Aufgabenstellung abgefragt, abgerufen und verarbeitet wird. Je nach Erfolg wird der Zyklus beendet und es wird entweder neues Wissen für die spätere Verwendung externalisiert oder es wird eine veränderte Abfrage gestellt. Ein hartnäckiger Misserfolgsfall wird im Modell nicht spezifisch abgebildet.

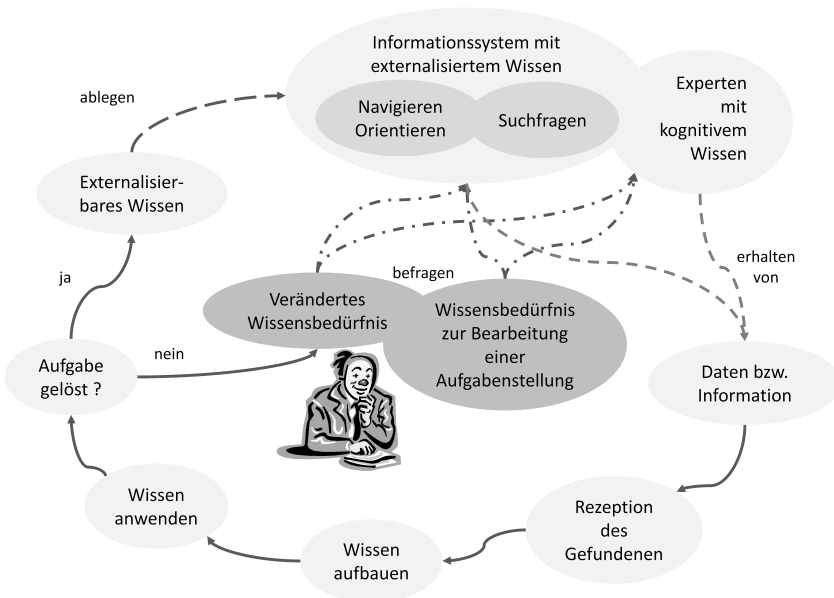


Abb. 1.1: Wissensaneignung.

²³ Wir lehnen uns in der Darstellung an ein bekanntes Modell an, das von Probst et al. für die Anwendung auf den Unternehmenskontext vorgeschlagen wurde, korrigieren und ergänzen es aber, damit es möglichst alle Formen einer an Aufgaben orientierten Wissensaneignung berücksichtigt. Aus Gründen der Konsistenz unserer Ausdrucksweise verwenden wir dabei statt ‚implizites‘ und ‚explizites Wissen‘ die Ausdrücke ‚kognitives‘ und ‚externalisiertes Wissen‘ (vgl.: Probst/Raub/Romhardt: Wissen managen).

Ein vergleichbares Modell für den allgemeinen Wissenserwerb zu skizzieren, das alle Wissensformen (auch prozedurales oder emotionales Wissen) einschließt, ist nicht so leicht. Weitgehend Einigkeit besteht darin, dass ein Wissenstransfer weder von Mensch zu Mensch noch von Medium zu Mensch als Vorgang eines reinen Transfers von Daten und ihrer unveränderten Speicherung gesehen werden kann. Einfache Modelle des Wissenstransfers wie der sogenannte ‚Nürnberger Trichter‘ haben sich als unzutreffend erwiesen. Immer ist es die Verarbeitung in der kognitiven Struktur des Rezipienten, die zum Wissen dieses Individuums führt. Es gibt keinen Königsweg zum Wissenserwerb, der für alle Individuen mit gleicher Effizienz abläuft. Für komplexe Zusammenhänge gilt dies als Allgemeinplatz, für vermeintlich einfache Daten oder Fakten aber nicht. Zu berücksichtigen ist, dass es keine von Theorie unabhängigen Daten und Fakten gibt und dass die Vollständigkeit des Begreifens von Daten und Fakten mit der Vollständigkeit des Wissens um die Theorie korreliert.

Gegenwärtig wird für kognitive Prozesse mit der sogenannten Computermetapher eine Verbindung zur Künstlichen Intelligenz und der Welt der Datennetze als einem neuen Intelligenzträger hergestellt. Der Ausdruck Computermetapher soll hier als Platzhalter für alle Versuche verwendet werden, Vorgänge der kognitiven Informationsverarbeitung durch Rückgriffe auf ein informationstechnisches Verständnis von Information interpretieren zu wollen. Es handelt sich um eine schleichende Umdeutung etablierter Begriffe unter der Prämisse einer informationstechnischen Prozessierung. Dabei wird die Bedeutung der Computermetapher durchaus unterschiedlich eingeschätzt. Das Spektrum reicht von der Charakterisierung eines sich vollziehenden Entwicklungsprozesses bis zum Manipulationsinstrument zur Umdeutung des Werts kognitiver Leistungen.

Lohnend ist ein Blick auf die historische Entstehung der Metapher, die man anhand von *Norbert Wiens* Buch *Kybernetik: Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine* nachvollziehen kann.²⁴ Dort wird zum ersten Mal die Bezeichnung ‚Kybernetik‘ für die Beschreibung von sich selbst regelnden Systemen von Servomechanismen verwendet. Es bestehen Verbindungen zu automatischen Navigationsvorgängen, verlässlicher Kommunikationstechnik und Künstlicher Intelligenz. Interessant ist die frühe Parallelität von ‚Lebewesen‘ und ‚Maschine‘.

Der Stellenwert, den eine Metapher für die interpretierten Sachverhalte bekommen kann, sollte nicht unterschätzt werden. Dies lässt sich an einem Beispiel verdeutlichen, das durch technische Entwicklungen eine ganz andere Dimension

²⁴ Wiener: *Kybernetik*; Original: Wiener: *Cybernetics or control and communication in the animal and the machine*.

erfahren hat, als es die dafür noch immer gern verwendete Bezeichnung erahnen lässt: das Abhören eines Telefonats. Analoges Telefonieren erfolgte mit einer unabhängigen Schwachstromversorgung und konnte als eine temporäre, aber proprietäre Verbindung zwischen zwei Teilnehmern gesehen werden, die über Vermittlungsstellen zunächst mechanisch und später automatisiert hergestellt wurde. Es war ein erheblicher technischer Aufwand erforderlich, das Abhören eines solchen analogen Gesprächs zu realisieren. Anders beim digitalen Telefonieren über Server, bei dem die Daten in speicherfähiger Form Bestandteil des Übermittlungsprozesses sind. Zum Verhindern eines späteren ‚Abhörens‘ müssen hier die Daten explizit gelöscht bzw. vernichtet werden. Nun erfordert das ‚Vermeiden‘ des Abhörens den Aufwand.

Dies darf sicher als Hintergrund einer Entscheidung des *Bundesverwaltungsgerichts* (BVerwG) vom 30.05.2018 gesehen werden, mit dem das Abschöpfen von Internet-Daten am sogenannten Frankfurter Knoten durch den *Bundesnachrichtendienst* für rechtmäßig erklärt wurde.²⁵ Juristisch ist das letzte Wort zwar noch nicht gesprochen, denn der Knoten-Betreiber bringt den Fall vor das *Bundesverfassungsgericht*²⁶; das Urteil des BVerwG hat aber deutliche Zeichen gesetzt, deren Spuren möglicherweise nicht mehr vollständig zu tilgen sind.

Grundsätzlich darf sicher davon ausgegangen werden, dass das allgemeine Verständnis für technische Vorgänge immer hinter deren Entwicklung zurückbleibt und von den historisch überkommenen Metaphern überlagert wird. *Richard Sennett* weist auf die große Bedeutung von Metaphern und deren prägende Wirkung auf Vorstellungs- und Beurteilungsprozesse hin:

Mit anderen Worten, eine Metapher erzeugt eine Bedeutung, die größer ist als die Summe ihrer Teile, weil diese Teile aufeinander einwirken. Die Elemente einer Metapher gewinnen aus der Beziehung zueinander eine Bedeutung, die sie für sich genommen nicht haben. Auf diese Weise können Metaphern gesellschaftliche Beziehungen herstellen: Unterschiedliche soziale Klassen oder unterschiedliche Rollen innerhalb der Gesellschaft können zu Elementen der Metapher werden. Das Ganze bringt die spezifische Bedeutung der Teile hervor.²⁷

Zentral für unsere Diskussion der Computermetapher ist der Begriff Information. Früher eher metaphorisch verwendete Ausdrücke wie Wissensspeicherung, Wissensorganisation oder Wissensmanagement erfahren hier eine direkte informationstechnische Bedeutungsgebung, bei der es sich eingebürgert hat, vom Speichern, Abrufen und Prozessieren von Gehirninhalten zu reden. Kognitive Informationsverarbeitung wird aus dieser Sicht als ein Prozess verstanden, wie er im

²⁵ BVerwG, 30.05.2018 – 6 A 3.16.

²⁶ Krempf: BND-Überwachung.

²⁷ Sennett: Autorität, S. 102–103.

Rahmen des Turingmodells von Computern durchgeführt werden könnte. Es gibt Ausgangsdaten, die eingelesen und von einem Algorithmus verarbeitet werden, um die auszugebenden Daten zu erzeugen. Verändert werden die Daten durch den Algorithmus, den Träger der ‚Intelligenz‘ des Prozesses. Die Vorgänge des Einlesens und Ausgebens der Daten werden im Allgemeinen nicht als Daten verändernd gedacht.

Für Kommunikationsprozesse resultiert die Computermetapher in einem Sender-Empfänger-Modell, das prinzipiell geeignet ist, Quelldaten ohne Beeinflussung an den Empfänger zu übertragen. Abweichungen stellen entweder gewollte technische Ergebnisse oder Fehler dar, die durch geeignete Vorkehrungen vermieden werden können. Ihren deutlichsten Ausdruck findet diese Vorstellung im sogenannten Turingtest. Bei diesem Test werden Fragen von einer nicht bekannten Instanz beantwortet. Diese Instanz kann ein Mensch sein, aber auch ein Rechner. Ist der Beobachter nicht in der Lage zu unterscheiden, welche der Antworten vom Menschen oder vom Rechner gegeben wurden, wird dem getesteten System Intelligenz zugeschrieben. Ein solches Kriterium erscheint auf den ersten Blick verführerisch. Wendet man es jedoch auf Lebewesen an, die eigenständig und auf ganz individuelle Weise am raum-zeitlichen Geschehen teilnehmen, darf man in Frage stellen, ob damit das Spektrum aller Intelligenzleistungen abgebildet wird.

Die Computermetapher führt für kognitive Prozesse nicht weiter. Ein Gehirn lässt sich, wie bereits beschrieben, nicht als ein System mit Trennung in Hard- und Software betrachten. Eine Trennung zwischen dem Gehirn und seinen Funktionen existiert weder physiologisch noch funktional, so zum Beispiel *Wolf Singer*:

Da gibt's im Gehirn nicht eine Trennung zwischen Rechenwerk und Programmspeicher und Datenspeicher oder so was, sondern es gibt nur Neuronen und deren Verschaltung. Und die Art und Weise, wie die verschaltet sind, nennen wir funktionelle Architektur, und da liegen die ganzen Geheimnisse, denn die Freiheitsgrade sind überschaubar, es kann nur variiert werden, wer mit wem kommuniziert, wie stark oder schwach die Kopplungen sind und ob sie hemmend oder erregend sind, wobei die allermeisten erregend sind. Und mit diesem Lego-Baukasten hat die Evolution die Großhirnrinde zusammengebastelt.²⁸

Die Fragwürdigkeit, den rein auf Verhaltensbeobachtung basierenden Turingtest als Kennzeichen für die Zuschreibung von Intelligenz zu nehmen, hat *John Searle* durch sein Gedankenexperiment des Chinesischen Zimmers offen gelegt.²⁹ In dieser Variante des Turingtests wird mit einem verborgenen Kommunikationspartner in chinesischer Sprache kommuniziert. Der Kommunikationspartner kann kein

²⁸ Kluge/Singer: Hirnforschung.

²⁹ Chinesisches Zimmer [Wikipedia].

Chinesisch, hat aber ein Regelwerk, das für alle gestellten Fragen aus einem ‚Baukasten‘ immer korrekte Antworten liefert. Das Regelwerk, der ‚Algorithmus‘, erzeugt den Eindruck, der Kommunikationspartner spreche Chinesisch, der Fragesteller gewinnt den Eindruck, das ‚System‘ könne Chinesisch. Tatsächlich konnte nur der (menschliche) Ersteller des Regelwerks Chinesisch.

Das Gedankenexperiment macht deutlich, dass rein symbolisches Manipulieren nicht ausreicht, um Rückschlüsse auf ein inhaltliches Verstehen oder das Entwickeln von Bewusstsein zu ziehen. In diesem Sinne ist das Chinesische Zimmer eine Widerlegung der klassischen Interpretation des Turingtests: aus beobachtetem Verhalten oder syntaktischen Manipulationen kann nicht auf ein semantisches Verstehen (Verständnis, Einsicht, Bewusstsein) geschlossen werden. Beispielsweise zeigt die Beobachtung eines Schauspielers, dass Verhalten keine verlässlichen Rückschlüsse auf Ursachen oder innere Zustände bzw. Funktionsweisen liefert.

Die Frage, wie aus Wörtern Bedeutung erzeugt werden kann, wird im Allgemeinen an das Vorhandensein eines Bewusstseins geknüpft. Ohne ein solches Bewusstsein bleibt eine Symbolmanipulation ein rein syntaktischer Vorgang auf der Ebene der Daten und erzeugt keine strukturierte Bedeutung. Durch reine Symbolmanipulation ist es nicht möglich, Bedeutung zu erzeugen – wie es durch Vertreter der KI immer wieder postuliert wird.³⁰

Häufig werden auch Varianten des klassischen Lügner-Paradoxons „Ein Kreter sagt: ‚Alle Kreter lügen‘“ verwendet, um die Möglichkeiten und Grenzen Künstlicher Intelligenz auszuloten. Eine bekannte Variante von *Bertrand Russell* lautet:

Ein Mann sagt: ‚Ich lüge gerade.‘³¹

Die Korrektheit der Aussage ist ebenso wenig entscheidbar wie alle aussagenlogischen Paradoxien oder Antinomien. Haben Nicht-Entscheidbarkeitsprobleme eine Bedeutung für die Diskussion um Künstliche Intelligenz? Eigentlich nicht. Wenn das System um die logischen Fallstricke des Paradoxons weiß, muss es sich – anders als es uns manche belletristische oder filmische Darstellung glauben machen möchte – nicht in Verzweiflung stürzen. Wie wir Menschen auch, kann ein KI-System die Haltung einnehmen: Alles schön und gut, aber wo beeinträchtigt uns das denn? Mit logischen Argumenten oder Szenarien grundsätzlicher Nicht-Entscheidbarkeit kann man dem Einsatz von KI-Systemen nicht begegnen.

30 Vgl. zu dieser Frage das Video: On consciousness with Giulio Tononi, Max Tegmark and David Chalmers [FQXi].

31 „The simplest form of this contradiction is afforded by the man who says ‚I am lying;‘ if he is lying, he is speaking the truth, and vice versa.“ (Russell: *Mathematical logic as based on the theory of types*, S. 222).

Zur Beschreibung einer künstlichen oder maschinellen Intelligenz werden daher zunehmend spezifische Kriterien genutzt, die für einzelne Einsatzfelder genauere Aussagen gestatten. Im Rahmen der als besonders anspruchsvoll geltenden maschinellen Sprachverarbeitung wird beispielsweise gemäß eines Vorschlags von *Hector Levesque* die sogenannte *Winograd Schema Challenge (WSC)* ausgetragen.³² Sie testet die Fähigkeit, das Bezugswort eines Pronomens in einer Aussage zu identifizieren. Falls es in der Aussage, die das Bezugswort des Pronomens enthält, mehrere Kandidaten gibt, ist das Pronomen mehrdeutig. Ein typisches Winograd-Schema enthält für eine Aussage einen alternativen Satzbestandteil, der die Bedeutung des Pronomens steuert:

The city councilmen refused the demonstrators a permit because they [feared/advocated] violence.³³

Das Bezugswort für das Pronomen ‚they‘ hängt davon ab, ob in den Satz ‚feared‘ oder ‚advocated‘ eingesetzt wird. Für ‚feared‘ sind es die ‚city councilmen‘, für ‚advocated‘ die ‚demonstrators‘.

Die Lösung der Aufgabe erfordert neben der korrekten Verarbeitung der Sprache zusätzliche Kenntnisse über die Eigenschaften und Motivationen der an der Handlung Beteiligten. Kenntnisse, die üblicherweise als Kontextwissen bezeichnet werden. Der Umfang des für die Lösung des Beispiels benötigten Kontextwissens dürfte beträchtlich sein.

Inzwischen ist es leicht möglich, die Leistungsfähigkeit von Systemen zur automatischen Sprachverarbeitung selbst zu testen. Die Eingabe der beiden Varianten des Winograd-Schemas in eines der aktuell leistungsfähigsten automatischen Übersetzungssysteme – *DeepL*³⁴ – erbringt folgende Ergebnisse:

The city councilmen refused the demonstrators a permit because they feared violence.

DeepL:

Die Stadträte lehnten den Demonstranten eine Genehmigung ab, weil sie Gewalt befürchteten.

The city councilmen refused the demonstrators a permit because they advocated violence.

DeepL:

Die Stadträte lehnten den Demonstranten eine Genehmigung ab, weil sie sich für Gewalt aussprachen.

³² Levesque/Davis/Morgenstern: The Winograd schema challenge.

³³ Bekanntes Beispiel von *Winograd*; vgl.: Winograd: Understanding natural language, S. 33.

³⁴ Vgl.: <https://www.deepl.com/translator>.

Neben der Erkenntnis, dass automatisches Übersetzen inzwischen auf einem hohen Niveau möglich ist, zeigt das Beispiel auch, dass Mehrdeutigkeiten in Übersetzungen unter Umständen erhalten bleiben. Automatisches Übersetzen kann auch ohne ein Kontextwissen zu akzeptablen Ergebnissen kommen. Die Auflösung der Mehrdeutigkeit erfolgt dann wiederum durch den menschlichen Leser.

1.3 Intelligenz und Künstliche Intelligenz

Die sprachliche Nähe zwischen den Begriffen ‚Intelligenz‘ und ‚Künstliche Intelligenz‘ führt leicht zu der Annahme, es bestünde zwischen beiden ein enger Zusammenhang in dem Sinne, dass Künstliche Intelligenz lediglich eine spezifische Ausprägung der allgemeinen Intelligenz sei. Tatsächlich geht es aber nicht um Intelligenz im Allgemeinen, sondern um die ‚menschliche‘ Intelligenz. Wir betrachten die Künstliche Intelligenz und die natürliche menschliche Intelligenz als zwei voneinander unabhängige Phänomene, deren Verhältnis im Rahmen einer Diskussion von Informationeller Autonomie genauer betrachtet werden sollte.

In vielen Darstellungen mit visionären Verheißungen zu den Möglichkeiten der Künstlichen Intelligenz wird zwar zugegeben, dass man dem Begriff Intelligenz nicht durch eindimensionale Beschreibungen oder gar durch eine einzelne Kennziffer gerecht werden könne. Gleichwohl wird dann eine verengte Interpretation vorgenommen, um besser eine Brücke zur Künstlichen Intelligenz schlagen zu können. So findet man etwa bei dem prominenten KI-Befürworter *Max Tegmark* die zunächst unverfänglich erscheinende Charakterisierung:

Intelligenz = Die Fähigkeit, komplexe Ziele zu erreichen³⁵

Die Verengung des Begriffs wird an seiner Begrenzung auf ein Kriterium deutlich: die Handlungsorientierung. Selbst die früher von KI-Vertretern bemühte Orientierung an der Fähigkeit, Probleme zu lösen, wird bestenfalls billigend unter ‚komplex‘ subsumiert. Damit fühlt man sich an den Turingtest erinnert, der als Denkstoß eine wichtige Bedeutung hat. Für die Bestimmung der maschinellen Intelligenz mag das ausreichen, zur Bestimmung des komplexen Begriffs Intelligenz von raum- und zeitgebundenen Lebewesen mit individuellen Erlebensräumen jedoch nicht.

Andere Faktoren, die inzwischen als konstituierend zur Bestimmung eines umfassenden Intelligenz-Begriffs gelten, bleiben bei solch einer Betrachtung unbe-

³⁵ Tegmark: *Leben 3.0*, S. 80.

rücksichtigt (zum Beispiel emotionale Faktoren).³⁶ Es ist konsequent, dass Künstliche Intelligenz dann gar nicht mehr als an Aufgaben orientiert oder funktional erklärt, sondern schlicht als ‚Nichtbiologische‘ Intelligenz verstanden wird.³⁷ So ist auch nachvollziehbar, dass keine Anstrengungen unternommen werden, zu einem spezifischen Verständnis von Information zu gelangen.³⁸ Die Orientierung an informationstechnisch geprägter Datenspeicherung oder biologischer Erbinformation im Sinne des Einspeicherns und Auslesens wird als ausreichend für kognitive Prozesse der Informationsverarbeitung gesehen.

Abbildung 1.2 stellt die Einflussfaktoren dar, die in ihrer Wechselwirkung die Gemeinsamkeiten und Unterschiede der beiden Konzepte Intelligenz und Künstliche Intelligenz vermitteln.

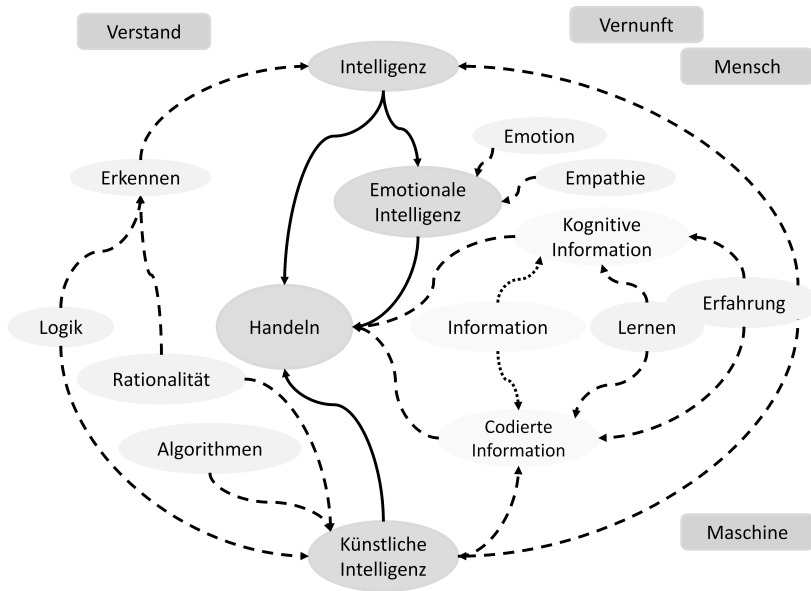


Abb. 1.2: Kognition – Intelligenz – Künstliche Intelligenz.

³⁶ Vgl.: Intelligenztheorie [Wikipedia]; Walter: Intelligenz; Ciompi: Die emotionalen Grundlagen des Denkens; Damasio: Descartes' Irrtum.

³⁷ Tegmark: Leben 3.0, S. 63.

³⁸ Bei Tegmark findet ‚Information‘ keine Berücksichtigung für den ‚Terminologie-Spickzettel‘, der eine Präzisierung der im Buch behandelten Schlüsselbegriffe geben soll (Tegmark: Leben 3.0, S. 63–65).

Für die menschliche Intelligenz lassen sich zwei Einflusstänge angeben, die – hier darf die Abbildung nicht missverstanden werden – nicht disjunkt zueinander sind, sondern je nach Ausgangslage miteinander in Wechselwirkung stehen. Der eine Strang verwendet Logik und Rationalität für den Erkenntnisvorgang und konstituiert damit einen Intelligenztyp, der für Handlungen genutzt werden kann. Der zweite Strang verwendet Ergebnisse des vorangegangenen Lernens und der Erfahrung – Lernen und Erfahrung können wiederum mit Logik und Rationalität verbunden sein. Die durch Logik und Rationalität ermöglichte Handlung kann darüber hinaus durch Emotion und Empathie beeinflusst werden. Rational geprägte Intelligenz und Emotionale Intelligenz können also für eine Handlung miteinander verbunden sein. In beiden Fällen findet die Verarbeitung von kognitiver Information statt.

Künstliche Intelligenz hingegen wird über Logik, Rationalität und Algorithmen erzeugt und beeinflusst. Handelt es sich um lernfähige Künstliche Intelligenz, vielleicht sogar um erfahrungsbasierte, so fließen die entsprechenden Ergebnisse in die codierte Information ein, die schließlich die Handlungsausführung steuert.

Die Analogien, die im Rahmen der Künstlichen Intelligenz zwischen Gehirn (Kognition), künstlichen neuronalen Netzen und Turingmaschinen hergestellt werden, verwenden die Computermetapher, um die Besonderheiten biologischer Systeme zu vernachlässigen und die kognitiven Funktionen auf eine mechanistische Funktion zu reduzieren. Beispielhaft wieder *Tegmark*:

Wie zuvor schon erwähnt, bewies Turing in seinem Aufsatz von 1936 obendrein noch etwas Tiefgreifenderes: Falls ein Computertyp ein bestimmtes absolutes Minimum an Operationen durchführen kann, dann ist er insofern ‚universell‘, als er – ausgestattet mit genügend Kapazitäten – alles tun kann, was ‚jeder beliebige‘ andere Computer tun kann. Er bewies, dass seine Turingmaschine universell war, und wenn wir uns jetzt etwas genauer auf die Physik zurückbesinnen, dann haben wir gerade gesehen, dass diese Familie universeller Computer ebenfalls Objekte einbezieht, die so unterschiedlich sind wie ein Netzwerk von NAND-Gattern und ein Netzwerk miteinander verbundener Neuronen. Tatsächlich hat Stephen Wolfram behauptet, dass die meisten nichttrivialen physikalischen Systeme – vom Wettersystem bis zu Gehirnen – universelle Computer wären, sollte man sie beliebig groß und langlebig gestalten können.³⁹

Hier wird die Gleichsetzung eines Computers im Sinne der universellen Turingmaschine (auf der Basis einer binären Logik) mit einem biologischen Gehirn (das über neuronale Verbindungen verfügt und mit anderen vegetativen Systemen verbunden ist) suggeriert, deren Berechtigung wir schon verworfen haben. Unabhängig davon ist aber fraglich, ob menschliche – oder allgemeiner – biologische Kognition

³⁹ Tegmark: *Leben 3.0*, S. 101–102.

allein vom Gehirn gesteuert wird und neuronale Verbindungen mit anderen Organen völlig außer Acht gelassen werden dürfen. Es muss der weiteren Entwicklung vorbehalten bleiben, welche Sicht auf die Dinge sich durchsetzen wird.

Wir können vorerst nur zusammenfassen: Selbst wenn die neuronale Grundlage der kognitiven Informationsverarbeitung biologischer Organismen (nach *Tegmark* metaphorisch die Hardware oder das ‚Substrat‘) im Sinne einer Turingmaschine verstanden werden kann⁴⁰, würde dies noch nicht heißen, dass auch die Verarbeitung wie bei einer Turingmaschine erfolgt. Bei einer Turingmaschine kommen Algorithmen zum Einsatz. Selbst wenn man sie als trainierbar und lernfähig ansieht, ergäbe sich die direkte Vergleichbarkeit erst durch die Fähigkeit zur Selbstorganisation, zum autopoietischen Verhalten biologischer Organismen. Diesen Konzepten und den daraus resultierenden Gesichtspunkten wird an späterer Stelle noch genauer nachgegangen werden.

Zunächst bleiben wir bei dem Begriff des Algorithmus, der zu einem Schlüsselbegriff für die Künstliche Intelligenz und die Verwendung der Computermetapher geworden ist. Die Anfänge seiner Verbreitung gehen mit der Entwicklung automatisierter Rechanlagen einher. Algorithmische Modelle finden mittlerweile aber zunehmend Eingang in die Diskussion um Kognition, Denken und Rationalität und werden sogar auf emotionale Prozesse angewendet.

Viele vertraute Handlungsabläufe lassen sich durch die Komponenten eines Algorithmus formulieren: Aufgliederung eines komplexen Ablaufs in festgelegte Einzelschritte und Abfrage von Bedingungen mit Verzweigungen an definierten Stellen des Ablaufs. Vielfach wird der algorithmische Ablauf verbunden mit einem Optimierungs- und Effizienzgedanken. Durch Befolgen der algorithmisch gedachten Einzelschritte sollen die wenigste Zeit, der kürzeste Weg, der geringste Ressourceneinsatz bei gleichbleibender Zuverlässigkeit gewährleistet sein.

Menschen werden im Berufsleben immer intensiver mit dem Effizienzgedanken konfrontiert und übertragen ihn dadurch möglicherweise bewusst oder unbewusst auf Alltagsverrichtungen, was eine fatale Entwicklung zur Folge hat: Mit dem Anspruch, jede Handlung unter Effizienzgesichtspunkten zu vollziehen, ist kein ausgeglichenes und glückliches Leben möglich, zumal erst die Abweichungen vom Effizienzgedanken Raum für Kreativität bieten.

Ein Beispiel: Man stelle sich die häusliche Zubereitung des Frühstücks für eine Einzelperson, für zwei Personen oder für eine größere Familie oder auch mit Gästen vor. Es muss einer Reihenfolge der einzelnen Schritte gefolgt werden,

⁴⁰ Es ist bezeichnend, dass an dieser Stelle terminologisch zu einem Ausdruck der Computermetaphorik Zuflucht genommen werden muss, weil für physiologische Prozesse keine Differenzierung in Hard- und Software vorgenommen werden kann (vgl.: Kluge/Singer: Hirnforschung).

persönliche Präferenzen müssen beachtet (abgefragt) und berücksichtigt werden. Sind die Schritte einmal festgehalten – es ist eine reizvolle Übungsaufgabe, dies im Detail zu versuchen –, könnte man jeden Tag nach demselben Schema vorgehen und die Illusion erzeugen, künftig absolut fehlerfrei agieren zu können. Dennoch wären Modifikationen unvermeidbar. Was macht man beispielsweise, wenn man erst während der Zubereitung merkt, dass das Vorratsbehältnis für den Kaffee oder das Müsli nachgefüllt oder ein neues Glas Marmelade geöffnet werden muss? Nicht immer sind alle Personen anwesend; es könnten Gäste mit anderen Präferenzen hinzu kommen, Vorräte könnten zur Neige gegangen sein oder Vorlieben sich geändert haben. Wie passt man den Algorithmus der Situation an, dass man etwas verschüttet und erst nach einem Tuch zum Aufwischen suchen muss? Wie überträgt man die Schritte auf eine veränderte Umgebung (zum Beispiel Ferienwohnung)? All dies könnte man durch Änderungen oder Erweiterungen im Algorithmus berücksichtigen, fertig würde er dabei voraussichtlich nie: Man kann nur bereits bekannte oder vorhersehbare Bedingungen und Bausteine einbauen und dabei gewisse Wahrscheinlichkeitsüberlegungen anstellen.

In der Realität sorgt die Eigenschaft der kognitiven Plastizität – einmal erzeugte Hirnstrukturen sind innerhalb dynamischer Grenzen änderbar (vgl. Kapitel 2.2) – dafür, dass man auch mit abweichenden Situationen umgehen kann, ohne sie zuvor durch Abfragen in den Ablauf eingebunden haben zu müssen. Merkmale und darauf basierende Abgrenzungen erzeugen Ordnung. Das Erkennen und Berücksichtigen des Dazwischen bzw. des Zusammenhangs (die Plastizität) sind eine wichtige Grundlage für Dynamik und Kreativität. Dies hat auch Konsequenzen für Ordnungssysteme wie beispielsweise Klassifikationen, die ihre Entitäten auf der Basis von Merkmalen gegeneinander abgrenzen. Dadurch schaffen sie zwar Transparenz, werden aber gleichzeitig zu statischen Instrumenten.

Wie das Frühstücksbeispiel zeigt, kann ein besonderer Wert des algorithmischen Vorgehens darin gesehen werden, sich Handlungsabläufe und die zugrunde liegende Entscheidungslogik zu verdeutlichen. Die Einführung von Computern in den schulischen Alltag sollte ursprünglich diesen Wert fördern. Wie wir heute wissen, handelte es sich dabei um eine Illusion. Gefördert wurde nicht das Denken in Algorithmen, sondern nur das Bedienen ihrer Ergebnisse im Rahmen einer IT-(Unterhaltungs)-Infrastruktur. Vielleicht ist dieses misslungene Experiment einer der Gründe dafür, dass die Spezifika algorithmischen Vorgehens wenig Eingang in das Alltagsdenken gefunden haben.

Es bedarf keiner Erwähnung, dass Menschen nicht auf das Vorhandensein formulierter Algorithmen angewiesen sind. Umstritten ist neuerdings jedoch, ob die zu vollziehenden Handlungen auf innere Algorithmen gestützt werden. Das würde bedeuten, dass durch das Vorliegen einer bestimmten Situation (Frühstück machen) eine kognitive Verarbeitungsroutine angestoßen wird, die alle erforderli-

chen Schritte berücksichtigt und in die notwendigen Handlungen umsetzt. Wenn dies so wäre, dann wäre die Routine durch ein Ausmaß an Flexibilität gekennzeichnet, das man Algorithmen üblicherweise nicht zuschreibt. In einem komplexen Vorgang wie der Zubereitung eines Frühstücks gibt es Module, die eine und nur eine Reihenfolge erlauben (Kühlschrank öffnen, ‚bevor‘ man die Butter herausnimmt; Kaffee in eine Tasse schütten, ‚bevor‘ man ihn daraus trinken kann). Es gibt aber auch Module, deren zeitliche Abfolge vertauscht werden kann (Tassen und Teller aus dem Schrank nehmen; Besteck aus der Schublade nehmen; Eier kochen). Nimmt man für eine einzelne Verrichtung immer die linke oder die rechte Hand? Lässt man sich von spontanen Umständen beeinflussen? Wer öfter das Frühstück zubereitet, wird die Erfahrung machen, dass vertauschbare Module auch immer wieder einmal vertauscht werden, verschiedene Personen werden ohnehin unterschiedliche Abläufe befolgen. Man wird sich hin und wieder auch dabei ertappen, eine eigentlich unzweckmäßige Reihenfolge gewählt zu haben. Wie unvorhersehbare Situationen in einen solchen inneren Algorithmus integriert werden, bedarf ebenfalls einer Erklärung. Ob all dies für oder gegen die Vorstellung innerer Algorithmen spricht, soll dahin gestellt bleiben. Vollständig kann menschliches Vorgehen allerdings nicht durch das Abarbeiten eines Algorithmus erklärt werden.

Bemühen wir uns deshalb um eine abstrakte Charakterisierung für den Begriff Algorithmus:

Ein ‚Algorithmus‘ ist eine eindeutige, ausführbare Folge von Anweisungen endlicher Länge zur Lösung eines Problems. Ein Algorithmus besteht aus einem Deklarationsteil (Was wird benötigt?) und einem Anweisungsteil (Wie wird das Problem gelöst?).⁴¹

Wesentliche Eigenschaften eines Algorithmus gemäß dieser Definition sind: Allgemeinheit, Eindeutigkeit, Ausführbarkeit, Endlichkeit, Determiniertheit und Terminierung. Durch diese Eigenschaften ist zwar formal festgelegt, was einen Algorithmus ausmacht; welche Möglichkeiten und Grenzen sich daraus für die Berechnung komplexer Probleme ergeben, bleibt jedoch eine offene Frage. Eine Frage, die auch für die Leistungsfähigkeit von Systemen zur Künstlichen Intelligenz von zentraler Bedeutung ist. *Roger Penrose* verneint die Möglichkeit der Algorithmisierung von Verstehensprozessen: „Echtes Verstehen liegt außerhalb von Berechnungen.“⁴² Er stellt seine Antwort in eine Beziehung zu den *Gödelschen* Unvollständigkeitssätzen und zur Rolle des Bewusstseins für ein allgemeines Verstehen:

⁴¹ Definition Algorithmus [Winfriedschule Fulda].

⁴² Blackmore (Hrsg.): Gespräche über Bewußtsein, S. 242.

Gödels Theorem ließ mir einfach keine Ruhe. Ich [...] dachte immer, es gehe darum, daß es Dinge gibt, die wir nicht wissen können. Aber dann erfuhr ich [...], daß es um ganz andere ging: Gödel sagt, daß wir diese Dinge wissen können, aber nicht, indem wir einfach die Regeln eines formalen Systems befolgen. Man braucht eine Methode, um an die Wahrheit zu gelangen, die verlässlich und gleichzeitig anders ist: Man muß sein Bewußtsein, seinen Verstand aufbieten um das Problem zu lösen. Es geht also nicht um das Befolgen der Regeln, sondern darum, zu wissen, warum die Regeln funktionieren; erst dadurch gelangt man zu einem Verständnis, das über die Regeln selbst hinausgeht.⁴³

Zu einem wesentlichen Bestandteil von Algorithmen gehört das Abfragen von Bedingungen, um Verzweigungen zur Behandlung unterschiedlicher Erfordernisse zu erreichen. Das einfachste Muster ist dabei eine Ja/Nein-Entscheidung. Die Planung eines Entscheidungsbaums setzt eine gründliche Analyse der zu erwartenden Fallvarianten voraus. Angesichts des dafür erforderlichen Aufwands ist es nicht verwunderlich, dass große Erwartungen in Konzepte gesetzt werden, die sich mit ‚Lernfähigkeit von Algorithmen‘ umschreiben lassen und für die inzwischen eine Vielzahl von Methoden und Techniken entwickelt wurden.

Derartige Algorithmen sind es auch, die besondere Aufmerksamkeit im Zusammenhang mit Informationeller Autonomie verdienen. Entscheidender als die Frage nach der Trainierbarkeit der zu erkennenden Situationen ist dabei, ob sich die Algorithmen in ihrem Handlungsspektrum selbst verbessern können. Dabei meint Selbstverbesserung nicht nur Fehlerfreiheit im bereits vorhandenen Programmcode, sondern das Schaffen neuer Leistungsmerkmale durch neue Programmcodes für bis dahin nicht berücksichtigte Fälle. Das Lernen würde dem Algorithmus also die Entscheidung ermöglichen, selbstständig eine Veränderung der bereits vorhandenen Programmierung vorzunehmen.

Die Computermetapher als Hilfsmittel zur Umdeutung etablierter Konzepte hat auch Konsequenzen für die Bedeutung von ‚Lernen‘ in Wendungen wie ‚Lernfähige Algorithmen‘ oder ‚Deep Learning‘.⁴⁴ Im traditionellen Informatik-Verständnis wird ein Programm verstanden als:

Programm = Algorithmus + Daten

Ist von der Lernfähigkeit von Algorithmen die Rede, wird dieses traditionelle Schema erweitert:

⁴³ Blackmore (Hrsg.): Gespräche über Bewußtsein, S. 244–245.

⁴⁴ Vgl. z. B.: Becker: Machine/Deep Learning. Durch die Kombination von ‚lernfähig‘, ‚genetisch‘ und ‚adaptiv‘ mit ‚Algorithmus‘, ‚neuronales Netz‘ oder ‚Programmierung‘ ergeben sich eine Reihe weiterer gebräuchlicher Ausdrücke, inzwischen auch ‚neuronale Turingmaschine‘ (DeepMind).

Programm = Algorithmus + Daten + Domain-Wissen⁴⁵

Dabei wird Domain-Wissen nicht an kognitive Eigenschaften gebunden – in dem Zusammenhang werden wir es noch als ‚Kontext‘ diskutieren –, sondern ausdrücklich in Abhängigkeit von einer formalen Datenmodellierung gesehen. So heißt es beispielsweise:

Die Verwendung von Hintergrundwissen, natürlich in einer passenden Datenstrukturform, ist grundlegend für die Lösung solcher Probleme.⁴⁶

Maschinelle Lernfähigkeit lässt sich anhand der Komponenten entsprechender Systeme genauer charakterisieren. So müssen etwa Programme zur Handlungssteuerung neben der internen Verarbeitung von Algorithmen auch über die Möglichkeit verfügen, ihre Außenwelt zu interpretieren, Signale zu empfangen, zu verarbeiten und zu senden und sie müssen durch Aktionen auf die Außenwelt einwirken können. Folgende Komponenten von Algorithmen lassen sich angeben:

- neben dem Programmcode des Algorithmus und den Daten ist Domain-Wissen in jeweils spezifischen Schemata repräsentiert;
- die Algorithmen arbeiten Regeln ab, die insbesondere eine Logik der Abfragen beinhalten;
- zur Bearbeitung der Abfragen werden eingegebene Daten gelesen oder durch Sensoren erfasste Bedingungen (Muster, Verhalten, etc.) erkannt;
- durch den Einsatz neuronaler Netze oder anderer Methoden des tiefen Lernens wird das Spektrum der zu erkennenden Bedingungen vergrößert;
- aus der Bearbeitung der Bedingungen für die Abfragen resultieren Schritte zum Vollzug einer Handlung;
- es gibt Möglichkeiten zur Planung und Veränderung der Strategie, um ein vorgegebenes Ziel zu erreichen;
- es gibt Möglichkeiten zur selbstständigen Veränderung des Programmcodes und deren Fixierung für zukünftige Bearbeitungsdurchläufe.

Für jede einzelne Komponente ist zu klären, ob Lernfähigkeit damit verbunden werden kann. Davon hängt ab, wie sich die Komponenten zu einem Gesamtsystem verbinden lassen, das im Sinne einer autonomen Maschine nicht nur spezialisierte Aufgabenstellungen bearbeitet, sondern sich mit selbst generierter und wachsender Leistungsfähigkeit in einer komplexen Umwelt bewegen kann. Die am weitesten gehende Lernfähigkeit wäre das Erreichen von künstlicher Kreativität.

⁴⁵ Vgl. für den Hintergrund dieses Ansatzes: Kókai: Erfolge und Probleme evolutionärer Algorithmen, S. 17.

⁴⁶ Kókai: Erfolge und Probleme evolutionärer Algorithmen, S. 17.

Betrachten wir nun die Rolle von Algorithmen für autonom arbeitende Geräte, um sie mit den erforderlichen Bedingungen und Erkennungsmustern für situationsangepasste Handlungsanweisungen auszustatten. Dass die Algorithmen über die Zeit gesehen aktualisiert werden müssen, ist ein Prozess, der aus der Computerwelt als Updating mit all seinen Begleiterscheinungen vertraut ist und der auch Handlungsanweisungen selbst einschließen kann. Es handelt sich erkennbar um einen sehr aufwendigen Prozess, der wohl nur bei einer geringen Anzahl aller autonom arbeitenden Geräte zur Anwendung kommen kann und dadurch Sicherheitslücken erzeugt, die böswillige Angriffe vereinfachen.

Einen Ausweg könnte die Lernfähigkeit jedes autonomen Gerätes bieten, wodurch die Ergebnisse seiner Lernprozesse automatisch seinem Wissensspeicher und seinem Algorithmus zugeführt werden. Neben der derzeit nicht abschätzbaren Größe des Aufwands für eine Realisierung bei allen autonomen Geräten, bleiben weitere offene Fragen. So lernt jedes Gerät nur ‚seine‘ neuen Situationen und nicht die der anderen, woraus sich unterschiedliche Lernzustände für verschiedene autonome Geräte ergeben. Treffen beispielsweise autonome Fahrzeuge unterschiedlicher Lernzustände aufeinander, kann es zu abweichenden Entscheidungen darüber kommen, wie eine konkrete Situation behandelt werden soll. Damit verhalten sich die autonomen Geräte zwar wie Menschen, aber sind ihre Entscheidungen dann noch immer weniger Fehler anfällig?

Die Idee der Synchronisierung der Lernvorgänge aller autonomen Geräte in einer Cloud basierten Masterumgebung liegt nahe, doch wieviel Realisierungswahrscheinlichkeit darf man ihr zutrauen, oder bleibt sie reine Fiktion? Wer löst die entstandenen Widersprüche aus unterschiedlichen Lernvorgängen auf? Die Masterumgebung selbst oder doch wieder der Mensch? Wie wird sichergestellt, dass die Aktualisierung der Masterumgebung zeitnah alle autonomen Geräte erreicht, damit alle auf derselben Entscheidungsgrundlage agieren?

Derzeitige Überlegungen eines autonomen Fahrens mittels 5G-Mobilfunktechnologie und Cloud basierter Verarbeitungsalgorithmen zielen weit mehr auf begrenzte Umgebungen mit homogenen Bedingungen (wie große Unternehmensgelände) als auf eine offene Verkehrsinfrastruktur mit heterogenem Fahrzeugbestand. Das einzelne Fahrzeug liefert seine Daten und bekommt Anweisungen, die zentral durch Auswertung dieser und der Daten anderer Fahrzeuge festgelegt werden. Nicht das einzelne Fahrzeug ist dann autonom, sondern allenfalls der Zentralalgorithmus.

Bei der Benutzung der Computermetapher werden menschliche Denk- und Handlungsvorgänge durch Algorithmen simuliert, um diese zu unterstützen oder sie vollständig durch Algorithmen zu ersetzen. Unser Alltag wird inzwischen von einer so großen Anzahl derartiger Vorgänge beeinflusst, dass sie uns im einzelnen gar nicht mehr auffallen. Dennoch sollte diese Tatsache nicht zu dem Umkehr-

schluss verleiten, dass es sich bei der algorithmischen Vorgehensweise um ein Leitbild für kognitive Vorgänge handelt.

Zur Veranschaulichung der Möglichkeiten einer Dialoggestaltung mit einem KI-System zitieren wir einen Dialog mit dem Chatbot *Cleverbot*. Es zeigt sich, dass das erneute Stellen derselben Frage nicht zur selben Antwort führt. Vielleicht die einzige menschliche Eigenschaft, die in solchen Entwicklungen mit ausreichender Zuverlässigkeit nachgebaut werden kann:

Frage: Denkst Du, dass ich ein Mensch bin?
Cleverbot: Ja.
 Frage: Warum?
Cleverbot: Ich weiss nicht.
 Frage: Warum weißt Du nicht, was Du denkst?
Cleverbot: Über was soll ich was denken?
 Frage: Warum Du denkst, dass ich ein Mensch bin.
Cleverbot: Ich dachte du bist ein Computer.
 Frage: Du dachtest doch, dass ich ein Mensch bin.
Cleverbot: Ich bin ein Hund:P.
 Frage: Kann ein Hund denn denken, dass ich ein Mensch bin?
Cleverbot: Nein du bist ein Computer weil du in mein iPhone.⁴⁷

1.4 Informationelle Autonomie und Freiheitsbild

Wie positionieren sich Vordenker und Befürworter eines modernen Humanismus gegenüber den Ideen des Transhumanismus? Bisherige Äußerungen aus diesem Kreis fallen eher zurückhaltend bis ablehnend aus. Statt dessen wird ein digitaler Humanismus vorgeschlagen – oder soll es doch ein humanistischer Digitalismus sein? In einem Interview zu seinem Buch *Digitaler Humanismus*⁴⁸ bemerkt *Julian Nida-Rümelin*:

Frage: Mit der KI stehen wir an der Schwelle selbst zu gottähnlichen Schöpfern neuer Wesen zu werden. Können wir das und dürfen wir das?

Nida-Rümelin: Wir tun es nicht, wir können es nicht und wenn wir es könnten, dürften wir es nicht. Die Softwareentwickler selbst sind in der Regel davon überzeugt, dass künstliche Intelligenzen, auch humanoide Roboter, keine mentalen Eigenschaften haben, keine

⁴⁷ Der Dialog wurde am 27.10.2018 geführt; alle vermeintlichen Fehler so im Original; vgl.: <http://www.cleverbot.com/>.

⁴⁸ Nida-Rümelin/Weidenfeld: Digitaler Humanismus.

Absichten verfolgen, keine Wünsche haben, keine Schmerzen empfinden, ja nicht einmal etwas erkennen oder entscheiden.⁴⁹

Vielleicht ist das Ausdruck einer gut durchdachten gesellschafts-philosophischen Position. Vielleicht ist es aber auch der Beleg für eine Fehleinschätzung gegenüber einer von Interessen geleiteten Dynamik zur Herstellung neuer gesellschaftlicher Strukturen mit deutlich weniger humanistischen Elementen. Eine klare Kennzeichnung transhumanistischer Ideen als gefährlicher Weg mit dem Potenzial zur Vorbereitung totalitärer Strukturen ist dabei jedoch nicht zu erkennen. Häufig wird die Frage nach der Leitwährung gesellschaftlicher Fortentwicklung nicht gestellt oder implizit zugunsten der Digitalisierung mit all ihren inhumanen Begleiterscheinungen beantwortet.

Rollen und Funktionen der Menschen in informationstechnischen Zusammenhängen können als Gestalter, Beitragender, Nutznießer oder Opfer beschrieben werden. Das durch die Aufklärung für alle Menschen geprägte Selbstverständnis eines autonomen Subjekts mit Ich-Verständnis, das seine Entscheidungen und Handlungen an rationale Denkvorgänge bindet, würde insbesondere die Rolle des Gestalters oder des Beitragenden betonen. Die aktuell zu beobachtende Verschiebung zugunsten funktionalisierter Mitglieder eines Schwarms würde eher die Rolle des Nutznießers oder Opfers betonen. Aus frei handelnden Subjekten würden Objekte und möglicherweise würde damit sogar einem totalitären Menschenbild der Weg geebnet.

Schon immer gab es die Bereitschaft, statt eigenen Nachdenkens lieber andere für sich denken und eigene Handlungsoptionen von anderen auswählen zu lassen. Bislang konnte diese Neigung allerdings als eine Art Bequemlichkeit, meist unter dem Vorbehalt einer späteren Rechenschaftslegung, betrachtet werden.

Nun aber scheint eine neue Ebene erreicht worden zu sein, die durch ein bewusstes Abtreten von Entscheidungen an Maschinen und Algorithmen gekennzeichnet ist. Warum gibt es sowohl im praktischen Geschehen als auch in den informationstechnischen Visionen die Neigung, Informationelle Selbstbestimmung zugunsten maschineller und algorithmischer Fremdbestimmung aufzugeben, obwohl die Selbstbestimmung durch die Rechtsprechung mit einem hohen Stellenwert versehen ist? Wird eine Fremdbestimmung durch Maschinen und Algorithmen als weniger bedenklich wahrgenommen als die durch andere Menschen? Erstaunlicherweise ändern nicht einmal unzureichende Kenntnisse über die be-

⁴⁹ Koch/Riecke: Deutscher Wirtschaftsbuchpreis.

teiligten Algorithmen – gleich, ob auf der prinzipiellen Funktionsebene oder der Ebene der Entscheidungsdetails – etwas an dieser Einstellung.⁵⁰

Ethische Überlegungen zur Verantwortung überspringen schnell die grundsätzliche Ebene und konzentrieren sich auf vergleichsweise artifizielle Entscheidungssituationen, wie etwa Abwägungen zur Minimierung der Opfer bei autonomen Fahrzeugen (Kinderwagen oder Senior).⁵¹ Derartige Sicht- und Vorgehensweisen eröffnen ein Spannungsfeld zwischen kognitiver Informationeller Autonomie und dem Zustand, den wir später als Informationellen Totalitarismus charakterisieren werden.

Abbildung 1.3 gibt einen ersten Überblick über die Verbindungen zwischen verschiedenen Schlüsselbegriffen. Auf die einzelnen Elemente werden wir im weiteren Verlauf der Darstellung noch genauer eingehen.

50 Vgl. hierzu etwa die Diskussion um das Offenlegen von Algorithmen, z. B.: Stalder: Algorithmen, die wir brauchen; vgl. auch die Beiträge der Plattform *Algorithm Watch* (<https://algorithmwatch.org/>); O’Neil: Angriff der Algorithmen; Krüger: Wie der Mensch die Kontrolle über den Algorithmus behalten kann.

51 Vgl. die Empfehlungen der *Ethik-Kommission Automatisiertes und Vernetztes Fahren*: BMVI : Bericht der Ethik-Kommission. Zunächst wurde die Abwägung zwischen dem Wert individuellen Lebens untersagt, nicht aber die Abwägung, welche Aktion die höhere Wahrscheinlichkeit zur Minimierung der Opferzahlen besitzt. Vgl. z. B. die Presseberichterstattung: *Autonomes Fahren* [Zeit online]. Vgl. auch: Rötzer: Brauchen Roboter eine Ethik und handeln Menschen moralisch? Vgl. zur Thematik auch die Aktivitäten des *Deutschen Ethikrats*: <https://www.ethikrat.org/>.

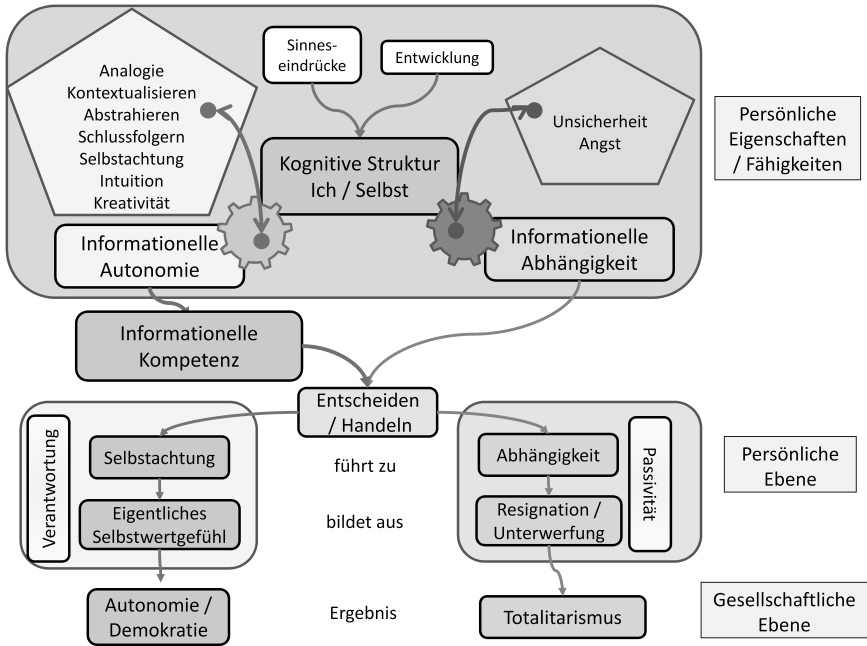


Abb. 1.3: Kompetenz – Gesellschaftsmodele.