

Jürgen Kippenhan

# Künstliche Intelligenz und die Sinnstrukturen menschlichen Handelns



Nomos



Jürgen Kippenhan

# Künstliche Intelligenz und die Sinnstrukturen menschlichen Handelns



**Nomos**



Onlineversion  
Nomos eLibrary

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

ISBN 978-3-7560-1644-0 (Print)

ISBN 978-3-7489-4041-8 (ePDF)

1. Auflage 2024

© Nomos Verlagsgesellschaft, Baden-Baden 2024. Gesamtverantwortung für Druck und Herstellung bei der Nomos Verlagsgesellschaft mbH & Co. KG. Alle Rechte, auch die des Nachdrucks von Auszügen, der fotomechanischen Wiedergabe und der Übersetzung, vorbehalten. Gedruckt auf alterungsbeständigem Papier.

*In Erinnerung an meinen Freund Henn Altena*



## Vorwort

Aus Brüssel ist in diesen Tagen zu vernehmen, dass sich die Europäische Union auf Regeln für den Einsatz *künstlicher Intelligenz* geeinigt hat. Dass hier Regulierungsbedarf entsteht, zeigt an, dass etwas gleichermaßen Wirkungsmächtiges wie auch Brisantes im Raum steht, dessen Umsetzung in die reale Praxis nicht sich selbst überlassen werden soll. Bei der anvisierten Regulierung handelt es sich um das weltweit erste KI-Gesetz (AI Act).

Entsprechend den Verlautbarungen sollen *Künstliche-Intelligenz-Systeme* in verschiedene Risikogruppen eingeteilt werden. Je höher die potenziellen Gefahren einer Anwendung eingeschätzt werden, desto restriktiver sollen die Anforderungen an Transparenz und Kontrollierbarkeit sein. Als besonders riskant anzusehende Anwendungen künstlicher Intelligenz können dann auch einem Verbot unterlegt werden. Im Zusammenhang mit den Transparenzregeln tritt die Frage in den Vordergrund, wie genau offenzulegen ist, welche Programme zum Training der „intelligenten“ Systeme eingesetzt werden und welchen „Verselbstständigungsgrad“ diese Systeme erreichen können. Zum Streitpunkt wurde dabei die Frage, ob die gesetzliche Regulierung schon an den Basisprogrammen der *künstlichen Intelligenz* ansetzt oder erst an ihren konkreten Anwendungen. So viel sei angedeutet: auch wenn es dabei um den Umriss rechtlich-politische Normierungen geht, im Kern sind hier aber grundsätzliche Einschätzungen angesprochen, die neben ihrer politischen Reichweite auf tief reichende philosophische Beurteilungen verweisen. Die auf Immanuel Kant zurückgehende Frage etwa „was können wir wissen?“ ließe sich hier umzumünzen in die Frage, welche Art von Wissen sich zum maßgeblichen innerhalb des Lebens moderner Gesellschaften erhebt bzw. erheben soll. Bemerkenswerterweise lassen sich die von Kant als grundlegend herausgestellten Fragen („Was sollen wir tun?“ „Was dürfen wir hoffen?“ „Was ist der Mensch?“) fast nahtlos daran anschließen. Darauf ist zurückzukommen.

Das Projekt der künstlichen Intelligenz erfährt in unserer Zeit nicht nur aus dieser gesetzgeberischen Perspektive große Aufmerksamkeit. Mit dem Generieren und Zur-Verfügung-Stellen von ChatGPT ist einerseits breites Aufsehen bezüglich des hohen und universell verfügbaren Leistungsniveaus der neuen Programmarchitektur und Datenverknüpfung erzeugt worden. Mit ihm verband sich weiterhin die Erwartung, das Gesamtvorhaben der

künstlichen Intelligenz erlebe damit einen die bisherigen Dimensionen der Digitalisierung sprengenden Anschlag. Tatsächlich warten die mit Microsoft und OpenAI konkurrierenden IT-Konzerne fast täglich mit Ankündigungen auf, die Großes mit Blick auf die zu erwartende Komplexitäts- und Datenkapazitätssteigerung und entsprechende Leistungsentfaltung der als intelligent hingestellten Systeme vorhersagen.

Dazu zunächst: Die Erfahrungen der zurückliegenden Jahrzehnte bezüglich dessen, was unter dem Stichwort „Digitalisierung“ das gesellschaftliche Leben in allen Bereichen infiltriert, ist in vielem so verblüffend, dass wenig Raum für gesteigertes Erstaunen zu bleiben scheint: „Das Digitale ist überall und umfassend. Es ist eingebettet in unsere hochtechnologische, vernetzte Lebenswelt. Es konstituiert sie, indem es uns über unsere technischen Geräte wie Mobiltelefone, PCs und intelligenten Objekte in einem kontinuierlichen Datenfluss verbindet. Das World Wide Web ist eine der Manifestationen des Digitalen, die Sozialen Medien eine weitere, ebenso wie die digitalen Geldströme auf unseren Konten.“ Diese Erfahrung unserer Zeit erhält ihre eigenwillige Charakteristik gerade dadurch, dass sie – trotz aller Aufgeregtheiten der Nutzercommunitys – eigenartig hintergründig verläuft, denn „trotz seiner allumfassenden Präsenz ist das Digitale verborgen in dem Gewirr an Datenleitungen, dem unsichtbaren Dauerfunken der Wireless Local Area Networks (W-LAN), in den Tiefen der technischen Geräte und des Internets. Das Digitale verbirgt sich in der Logik der Schaltkreise auf Computerchips und in den Datenspeichern der Maschinen.“<sup>1</sup>

Zu betonen sei noch, wie schnell sich die Nutzerinnen und Nutzer auf den Ebenen des Alltäglichen (vermerkbar als „Kolonialisierung der Lebenswelt“ durch Smartphones) bis hin zu den Etagen des Professionellen und der hoch spezialisierten Wissenschaften an die digitalisierten Eingriffe, Gestaltungsformen, Substituierungen und direktiven Übernahmen gewöhnt haben, eine Erfahrung, die erwarten lässt, dass sich diese Tendenz beschleunigt, erweitert und vertieft.

Kurz, die Überzahl dessen, was innerhalb der Gesellschaft und in ihren Subsystemen wie Politik, Wirtschaft, Produktion und Verwaltung, Forschung, des Finanzsystems, der Bildung, des Militärs, der Kultur und nicht zuletzt dem engen Netz der Alltagsverrichtungen – und überall dort bei der Ausgestaltung von Arbeitsprozessen – vorangetrieben wurde und wird, vollzieht sich über digitale, also programmtechnisch gestaltete Systeme. An

---

1 G. Gramelsberger, Philosophie des Digitalen, Hamburg 2023, S. 11.

eine Entflechtung oder gar Außerkraftsetzung dieser höchst wirkungsvollen Vernetzung und Steuerungssystematik ist nicht mehr zu denken.

Der Anstoß, der nun von den neuesten Entwicklungen innerhalb der *künstlichen Intelligenz* ausgeht, treibt diesen Prozess der hochtechnologischen Systemgestaltung nachdrücklich an und führt, wie fast immer, wenn etwas Zukunftsträchtiges im Raum steht, zu ebenso hoffnungs- wie angstvollen Fragen bezüglich der Richtung, in die diese Dynamik das gesellschaftliche Leben treiben wird, wie es zu bewerten ist und wie darauf zu reagieren ist.

Verschiedene Aspekte lassen sich dabei unterscheiden. Unter dem Eindruck der ablesbaren Fortschritte rückt nicht zuletzt die Erwartung in den Vordergrund, mithilfe der künstlichen Intelligenz ließen sich sehr weit reichende Erkenntnis- und Produktivitätsfortschritte erzielen, die dazu beitragen, sich abzeichnende Krisenszenarien (Klimaveränderung, Migration, geostrategische Verschiebungen, Finanzkrisen) abzuwenden. Aus diesem Blickwinkel steht in Aussicht, dass die *künstliche Intelligenz* auf nahezu allen Ebenen, nicht zuletzt aber der der Forschung, der Logistik, des Wissensmanagements und Technikentfaltung hohe katalisatorische Wirkungen speziell in den funktionalen Bereichen des gesellschaftlichen Lebens erzielt.

Im Grunde geht es um programmtechnisch hoch entwickelte, universell einsetzbare und gleichzeitig auf enge Bedarfe zuschneidbare „intelligente“ (hier zunächst im Sinne von „lernfähig und sich selbst steuernd“) Prozessstrukturen, die in dieser Kombination den erhofften Produktivitätsfortschritt erzeugen sollen, die im Sinne eines „Quantensprungs“ die Informationsgesellschaft in neue Dimensionen der Risikobewältigung und der Prosperität überführen soll.

Unvermeidbar stellt sich dann die Frage, inwieweit den im Zeichen der *künstlichen Intelligenz* erzeugten Programmstrukturen und Datenverknüpfungen zugesprochen werden kann, in ihren Arbeitsweisen und -ergebnissen an das menschliche Urteilsvermögen, seiner Kreativität und seinem Lernvermögen heranzureichen. Dies müsste – und es beträfe den herausforderndsten Punkt in der Gesamteinschätzung – auch höhere geistige Leistungen wie die des *Sinn- und Bedeutungsverstehens* einschließen.

Spätestens hier zeichnet sich eine *Bipolarität* von „prosaischer“ und „visionärer“ künstlicher Intelligenz ab, die zwar schon in den Anfängen des Projekts der künstlichen Intelligenz angelegt war, nun aber neue Konturen und eine bisher nicht realisierte Anschaulichkeit gewinnt.

Die nüchterne Perspektive auf die notierte *Bipolarität* vermerkt zunächst die Entwicklungsergebnisse so: „Die real existierende `prosaische Künstliche Intelligenz` geht hervor aus der Synthese von Sensortechnologie, Netzwerkkonstruktionen, *Big Data*- und *Deep Learning*-Verfahren und ist in der Form von Suchmaschinen, Sprach- und Gesichtserkennung, Übersetzungssoftware etc. allgegenwärtig. Das Gravitationszentrum der gesellschaftlich wirksamen Künstlichen Intelligenz sind diese unsichtbaren datentechnologischen Verfahren.“<sup>2</sup>

Am Beispiel des Projekts autonom fahrender Fahrzeuge ließe sich etwa deutlich machen, wie das Zusammenspiel komplexer sensorischer Erfassung von Verkehrsabläufen, verbunden mit entsprechend dynamischer Datenverknüpfung, dazu führt, *menschliches* Situationserfassen (und dann auch menschliches Steuern) überflüssig zu machen. Dies ließe sich prototypisch auf viele andere Bereiche des gesellschaftlichen Lebens übertragen und hochzurechnen ist, dass sich hier die Akzente nachdrücklich zugunsten der informationstechnischen Apparaturen verschieben werden. Nicht selten tritt hier das Argument auf, dass eine „freundliche Übernahme“ mit einem deutlichen Gewinn an Präzision, Ermüdungsfreiheit, Komplexitätsbewältigung, Geschwindigkeit und Sachlichkeit verbunden sei.

Dieses Argument ist schwer von der Hand zu weisen. Längst hat sich die Menschheit daran gewöhnt, dass die maßgeblichen Faktoren etwa im Flugverkehr (Steuerung des Flugzeuges, Überwachung des Luftraums etc.) unverzichtbar der Computeranalyse und entsprechenden Eingriffe unterliegen.

Dahinter aber – und das führt zurück zu der angesprochenen Bipolarität – stehen *Paradigmen der Programmgestaltung*, die juristisch und intellektuell auf menschliche Programmierentscheidungen zurückzuführen sind.

Damit es überhaupt denkbar wird, dass Computer gerade auch in diesem Bereich der *Ziel- und Paradigmenbestimmung* – nicht zuletzt unter Verantwortungsgesichtspunkten – die Führung übernehmen, müssten sie sich – wie betont – der *Urteilsfähigkeit und dem Sinnverstehen* des Menschen nähern.

Noch einmal: Auch wenn es sicher erscheint, dass in Zukunft auf `prosaischer` Grundlage, also mit Blick auf die angesprochene Präzision, Schnelligkeit und `Umsicht`, eine Überzahl von Entscheidungs- und Steuerungskompetenzen vom Menschen auf den Computer übertragen wird, so

---

2 Sybille Krämer, Nüchtern bleiben! Künstliche Intelligenz jenseits des Mythos, in: A. Strasser u.a. (Hrsg), Künstliche Intelligenz, Die große Verheißung, Berlin 2021, S. 336

wird zum entscheidenden Problem, woraus die *dahinter liegenden* Grundlagenentscheidungen, Bewertungen, Zwecksetzungen, Entscheidungsstrategien, Kommunikationsstrukturen, Erkenntnismethoden, Muster der Wissensgenerierung wie auch ethischen Weichenstellungen schöpfen bzw. wer hier die Federführung übernimmt.

Solange Menschen bzw. gesellschaftliche Institutionen dafür verantwortlich sind (bzw. verantwortlich gemacht werden), scheinen die entsprechenden intellektuellen und juristischen Zuschreibungen geklärt. Tatsächlich aber werden sich komplexere Sachverhalte als zunehmend intransparent erweisen. Unter dem Stichwort der *Ko-Performanz* ist ersichtlich, dass nahezu alle gesellschaftlich relevanten Prozessverläufe (institutionelle Funktionen, Informationsflüsse, Kommunikationen, politische Entscheidungen etc.) auf Mensch-Maschine-Interaktionen beruhen, innerhalb derer die Maschinen über Lerneffekte (über Feedback-Kreisläufe erzeugte Skaleneffekte) sich selbst optimieren und sich in ihrer Flexibilität steigern. Kurz, die menschlichen Entscheidungen geraten in den Sog des flankierenden Zuspiegels durch Maschinen: „Alles was ist, was gewesen ist, was in Zukunft sein wird, kann in eine Datenstruktur umgewandelt und von Algorithmen durchsucht, analysiert und bearbeitet werden. Nicht mehr das Verstehen, vielmehr die Suche wird zum Königsweg im Umgang mit Daten. Schon hier zeichnet sich ein Umschlag von Kontrolle in Kontrollverlust ab: Einerseits werden große Teile des Wissens durchsuchbar und zugänglich. Doch was nicht die Form der Digitalisierung annimmt, fällt aus dem Suchraum heraus.“<sup>3</sup>

Die *Paradigmen und Trainingsvorgaben* mithin, die in Wissensdarstellungen und Wissensmanagement, in Deep-Learnig-Programme und Algorithmenbausätze eingehen, sind aus dieser Perspektive gleichermaßen effektiv wie undurchsichtig. Wie anhand der Bemühungen der Instanzen der Europäischen Union zu sehen war, soll Letzteres – und damit auch der Kontrollverlust – vermittels von Transparenzgeboten verhindert werden.

Im Zeichen der künstlichen Intelligenz aber, der ja gerade wegen ihrer Attribuierung als *intelligent* (im Sinne von selbstlernend, selbsturteilend) eine unvermeidbare Autonomie zugesprochen wird, wäre dies – selbst wenn durchsichtig – nicht mehr steuerbar. Kurz, die als intelligent ausgewiesenen Systeme würden ihre Entscheidungen selbst treffen.

---

3 S. Krämer, a. a. O. S. 341

Dies wiederum würde voraussetzen – um es sehr komprimiert zu sagen –, sie *wüssten, was sie tun*. Nach allem, was wir wissen, sind die gegenwärtig die Szene bestimmenden Hochleistungssysteme noch weit davon entfernt. Ob und in welchem Maße sie diese Entfernung aber je werden überwinden können, darum soll es in der folgenden Abhandlung sinngemäß gehen.<sup>4</sup>

Die Fähigkeiten, die nun speziell Programmen wie ChatGPT bzw. seinen angekündigten Folgeprogrammen zugesprochen werden, sind ihrerseits mit der Vision verbunden, in Zukunft einen Großteil der Text- und Wissensproduktion wissenschaftlicher, redaktioneller oder auch literarischer wie auch alltagssprachlicher Art von diesen Programmen übernehmen zu lassen. In der Folge hätte dies, wie es sich ausmalen lässt, Auswirkungen auf die geistige Verfassung einer Gesellschaft, die dann ihre „*Texte*“ programmtechnisch vorgegeben bekäme. Damit geriete unter dem Stichwort der ‚Bipolarität‘ die „visionäre“ künstliche Intelligenz ins Blickfeld.

Kurz gesagt geht es um die Frage, ob die maschinelle Intelligenz prinzipiell in der Lage sein wird, in einem umfassenden Sinn der menschlichen Intelligenz gleich zu sein und ihre Funktionen abzubilden. Ob und in welcher Art der Gestaltung das von ihr Anvisierte umsetzbar und überhaupt wünschenswert sein wird, hängt von der Beantwortung einer Reihe nicht zuletzt philosophischer Fragen ab. Im Kern geht es um die sich naheliegende Frage, ob sich menschliches Denken, Sprechen, Erkennen, Entwerfen, Bewerten, Fühlen, Fantasieren, Beurteilen, Kommunizieren auf programmtechnische Strukturen reduzieren bzw. mithilfe algorithmischer Modelle rekonstruieren und reproduzieren lässt. Im Folgenden soll deutlich werden, welche Art der „Operationalisierung“ des menschlichen Geistes dem Modell und mithin dem Aufbau der künstlichen Intelligenz zugrunde liegt. Anders gesagt: Um herauszuarbeiten, was der künstlichen Intelligenz zuzutrauen, ist es notwendig, den logischen Aufbau algorithmischer Prozessstrukturen herauszuarbeiten und ihn zu der „Arbeitsweise“ des menschl-

---

4 Heute kommt in der Debatte um die digitale Gestaltungskraft noch ein Aspekt hinzu, der den Akzent auf die Einbettung des geistigen Vermögens des Menschen in einen eigensinnig erweiterten transsubjektiven Raum verlegt. Unter dem Stichwort „Metaversum“ öffnet sich ein Gebilde eines digitalen Raumes, der seine eigene – paradoxerweise: virtuelle – Realität erzeugt, einer Realität, die sich entgegen der bisherigen und grundlegenden Unterscheidung zwischen real und erfunden (fiktional) eigene Dimensionen erzeugt. In ihr geht es um “keine falschen realen Objekte, sondern wahre virtuelle Objekte, für welche die Frage der realen Realität ... keine Frage ist.“ E. Esposito, Fiktion und Virtualität, in: S. Krämer (Hrsg.), Medien, Computer, Realität, Wirklichkeitsvorstellungen und Neue Medien, Frankfurt/M. 1998, S. 270.

chen Geistes ins Verhältnis zu setzen. Denkbar kurz formuliert: Was eint und was trennt künstliche von menschlicher Intelligenz?

Aachen, Dezember 2023



# Inhaltsverzeichnis

1	Einleitung	17
2	Ausgangsfragen – Theorie der Rationalität – Aufgabe der Philosophie	27
3	Philosophie, Wissenschaft und Technik	57
3.1	Neuzeitliche Wissenschaft, Technik und das Projekt der Moderne	57
3.2	Das Ideal exakter Wissenschaft	71
3.3	Der logische Positivismus	78
3.4	Die neue Reichweite der Technik	85
3.5	Der Anspruch der Kybernetik	92
3.6	Ausblick	98
4	Künstliche Intelligenz als Herausforderung der Philosophie	101
4.1	Künstliche Intelligenz und das Paradigma technischer Kommunikation	101
4.1.1	Kognitivismus und Konnektionismus	102
4.1.2	Geschichte und Anspruch der künstlichen Intelligenz	111
4.1.3	Wissensrepräsentation	120
4.1.4	Neuronale Netze	129
4.2	Grundpositionen in der Auseinandersetzung um die künstliche Intelligenz	144
4.3	Künstliche Intelligenz im Zeichen grundsätzlicher philosophischer Fragen	193
4.3.1	Geist und Gehirn	195
4.3.2	Grundzüge des Funktionalismus	201
4.3.3	Kritik des Funktionalismus	205
4.3.4	Das Leib-Seele-Problem	216
4.3.5	Intentionalität und Bedeutung	234

## *Inhaltsverzeichnis*

4.4 Maßgaben der Philosophie. Die rationalen Grenzen der künstlichen Intelligenz	244
4.4.1 Intentionalistische vs. pragmatische Bedeutungstheorie	247
4.4.2 Geltung und Bedeutung	257
4.4.3 Identität und Bewusstsein	266
4.4.4 Zum Charakteristikum des In-der-Welt-Seins	279
4.4.5 Lebenswelt als gemeinsam geteilter Hintergrund und das Problem der Repräsentation	285
4.4.6 Vom 'Gebrauch' der Vernunft	296
4.4.7 Hermeneutik und Wahrheit	300
4.5 Zusammenfassende und abschließende Bemerkungen	329
5 Gesellschaft im Zeichen 'intelligenter' Steuerung und technischer Kommunikation	335
6 Schlussbetrachtungen	359
Literaturverzeichnis	385

## 1 Einleitung

Ein Projekt wie das der *künstlichen Intelligenz* ist notwendigerweise umstritten. Die Ausgriffe auf das, was bisher dem menschlichen Geist bzw. Urteilsvermögen auferlegt und vorbehalten war und damit die Ansprüche, diesen technisch reproduzieren zu können, sind so weit gesteckt, dass sie den Verdacht, anmaßend und gegebenenfalls auch bedrohlich zu sein, auf sich ziehen. Ob das Projekt nun realisiert werden kann oder nicht: Schon der Entwurf einer Reproduktion des menschlichen Geistes zielt auf eine Entfesselung technischen Könnens, die alle die Menschheitsgeschichte begleitenden und umwälzenden technischen Projekte – die großen wie die kleinen – in den Schatten stellt. Dass dennoch die Aussicht auf intelligente Computer nicht wie ein Schock wirkt, liegt daran, dass Computer sich schon heute so intelligent in den Organismus der gesellschaftlichen Systeme einfügen, dass die virtuelle Übernahme intelligenter Funktionen eher unter dem Gesichtspunkt der Perfektionierung einer Technik registriert wird, die für den Nichtexperten kaum mehr durchschaubar ist. Computer lenken die Stationen im Weltall, steuern weltumspannend Finanztransaktionen, verhelfen zu Einblicken in subatomare Strukturen, simulieren Klimaszenarien, erstellen Computertomografien und unendlich vieles mehr. Nicht zuletzt ihr Einsatz in der Medizin erweckt das Vertrauen, dass die Diagnosen 'unbestechlich' und die Eingriffe präzise sind.

Zweifel an einer realen bzw. zu erwartenden 'Intelligenz' von Computern melden sich nun da an, wo es um Urteilskraft und die ihr komplementären Vermögen der Situationserfassung, der kreativen Sprachgenerierung, der intuitiven Problemlösung, der zwischenmenschlichen Empathie und nicht zuletzt der gleichberechtigten Teilnahme an intersubjektiven Verständigungen und die sie leitenden Sinnkonstruktionen geht.

Da es nun zu den Grundeinstellungen des Menschen zählt, sich reflexiv dieser Vermögen zu vergewissern, d. h., sie unter systematischen Gesichtspunkten zu objektivieren, wird zum entscheidenden Punkt, inwieweit der Gegenstand der Untersuchung dies *erlaubt*. Spielstrategien beim Schach lassen sich, wie man weiß, bis zur höchsten Ebene in Programme fassen. Wer auf diesem Niveau spielen kann, gilt in *gewisser Weise* als intelligent. Berechtigterweise würde man aber den Programmen als ganzen dieses Attribut verweigern. Und tatsächlich zielt *künstliche Intelligenz* auf

mehr. In ihr geht es um die sukzessive Rekonstruktion und Modellierung derjenigen Leistungen, die beim Menschen in einem *umfassenden* und existenziellen Sinn als intelligent gelten.

Schwierigkeiten entstehen nun dadurch, dass der *Begriff* der Intelligenz nicht exakt umrissen ist. Aus der Vielfalt der Definitionsversuche ist gar zu entnehmen, dass es sich um einen Sachverhalt handelt, der sich auf widerständige Weise einer prägnanten Definition entzieht. Selbst die Systemtheorie, erklärt mit der Stimme Luhmanns 'Intelligenz' zu einem Verlegenheitsbegriff, der gerade anzeigt, dass man nicht weiß, was sich im Einzelnen vollzieht.

Angesichts der Tatsache, dass das Forschungs- und Realisierungsprogramm der *künstlichen Intelligenz* längst auf dem Weg ist, wird man sich mit Antworten dieser Art nicht zufriedengeben, sondern in ihnen einen Stachel zum Weiterdenken sehen.

Die bündige Frage, die den Gang dieser Abhandlung nach all dem Gesagten leiten wird, wird sich darauf richten, ob und in welcher Hinsicht künstliche Intelligenz *möglich* ist und ob sie – und damit ihr Einsatz – unter ethischen Gesichtspunkten zu *wollen* ist. Es wird nicht verwundern, dass diese Fragen keine knappe Antwort zulassen. Umgekehrt, es ist schwierig, einen engen Zuschnitt zu finden. Wie nahezu alle Beiträge zur Diskussion des Themas bezeugen, handelt es sich um Fragestellungen, die in alle Bereiche vordringen, die sich die Philosophie je erschlossen hat. Selbst ein Mathematiker – hier Roger Penrose – führt mit großer Eindringlichkeit vor, dass es sich um einen Gegenstandsbereich handelt, der bis an die großen kosmologischen Fragen heranführt.

Es handelt sich aber nicht nur um ein Thema von großem inhaltlichem Umfang, sondern möglicherweise auch eine des großen Umbruchs. Nicht wenige der Protagonisten der *künstlichen Intelligenz* kündigen an, dass im Zuge ihrer Realisierung die wichtigsten Fragen der Philosophie in ein neues Licht gerückt würden. Computer führen uns demnach nicht nur in ein neues Zeitalter der Technologie, Steuerung und Information, sie leiten auch, wenn die Erwartungen zutreffen, einen radikalen Wandel unserer Weltbilder ein – einschließlich des Bildes, das wir uns von uns selbst machen.

Visionen beglaubigen aber noch nicht von sich aus ihre Realisierbarkeit. In der Gegenwart allerdings nähren sich die Ausblicke auf ein zukünftiges – insbesondere von künstlicher Intelligenz geprägtes – Computerzeitalter von eindrucksvollen Veränderungen, die die neuen informationsverarbeitenden Maschinen in kurzer Zeit in relevanten gesellschaftlichen Bereichen her-

beigeführt haben. Und die Entwicklung nimmt eine Fahrt auf, die kaum präzise und seriöse Einschätzungen erlaubt. Und wie nicht anders zu erwarten, gibt es auch deutlich warnende Stimmen. Diese treffen allerdings wiederkehrend auf den nicht zu umgehenden Einwand, dass die sich im Zeichen des ‚Prinzips Verantwortung‘ erforderlichen Maßnahmen auf dem Stand der Dinge nicht mehr ohne hoch entwickelte Informationstechnik durchführbar sind.

Die Grenzen *menschlicher* Intelligenz nun, wie sie sich in den Phänomenen der Armut und des Elends, den Potenzialen der Destruktion und der ökologischen Bedrohung widerspiegeln, könnten sich dann aus dieser Perspektive durch künstliche Varianten so ausgleichen lassen, dass sich insbesondere die sich ankündigenden Katastrophen abwenden lassen.

Überlegungen dieser Art haben einen rationalen Kern, spiegeln andererseits jedoch eine verengte bzw. vereinseitigende Sicht wider. Dies erkennt man auch daran, dass ein Teil der umrissenen Probleme nicht aufgrund des *Fehlens* moderner Technik entstanden ist, sondern gewissermaßen auf ihrer Grundlage. Ambivalenzen dieser Art scheinen dem Thema unauflösbar eingeschrieben.

Aus ganz anderer Perspektive nun haben neue Paradigmen – in der Technik wie in der Wissenschaft – gewissermaßen einen *Zweck in sich*. Neben der beanspruchten Förderung des allgemeinen Wohls geht es ihnen um die Begründung eines neuen theoretischen Konzepts. Auch wenn z. B. Francis Bacon seine Idee einer Inthronisierung empirisch-technischen Wissens der Verwirklichung humaner Ziele verschrieb, so ging die eigentliche Sprengkraft doch von der generell neuen *Sichtweise* aus, die sich wiederum einer eigenen, auf exakter Wissenschaftlichkeit beruhenden Begründungsstrategie verdankt.

Vergleichbar damit lässt sich die *künstliche Intelligenz* als ein neues *technologisches Paradigma* fassen, das – ungeachtet dessen, ob es alle Visionen realisieren kann – sein Interesse an ‚Selbstbehauptung‘ innerhalb des philosophisch-wissenschaftlichen Diskurses wahrnimmt. Sein zentrales Gedankenmodell ist, dass Intelligenz und Denken – man könnte auch sagen: der menschliche Geist – auf grundlegender Ebene eine logische und symbolische Struktur aufweisen, die sich, wie komplex sie sich auch auf der Makroebene darstellt, auf formale symbolische Systeme zurückführen lässt. Ausgehend von dieser Überzeugung eines elementar-funktionalen Aufbaus der ‚Gedankenwelt‘, entfaltet sich ein Systemgedanke, der mit ähnlich universellem Gestaltungswillen auftritt wie ehemals der logische Positivismus, mit dem er die szientistischen Wurzeln teilt.

Wenn etwa Joseph Weizenbaum von der auf dem szientistischen Modell aufbauenden 'Computermetapher' spricht, dann verortet er darin einen ideologischen Gehalt, dem er zutraut, die Welt bis zur Unkenntlichkeit zu verändern. In einem letzten Akt der Selbstüberbietung eignet sich in dieser Sicht der 'Verfügungsgedanke' selbst jene geistig-seelischen Prozesse an, die der Mensch zum Verankerungspunkt seines Anspruchs auf Integrität und Würde macht und die nun durch die Subsumtion unter eine abstraktive Vernunft endgültig 'entseelt' zu werden drohen.

Es ist unschwer, Kontinuitäten zu erkennen. Angestoßen durch Max Weber wurde der Begriff der 'Entzauberung' zu einem der großen zeitdiagnostischen Stichworte. Die Moderne ist danach unabwendbar einem Rationalisierungsprozess ausgeliefert, dessen kalte Logik gemäß Weber zum bestimmenden Geistes- und Gestaltungsmodell wird.

Bemerkenswert ist, dass nahezu alle *Selbstverständigungen der Moderne* seit Hegel erkennen, wie sich dieser vereinseitigende Prozess der Rationalisierung mit nahezu naturgesetzlicher Gewalt durchsetzt. Gleichzeitig zeigen diese Akte der Selbstvergewisserung ihrerseits an, auf welchem hohen Stand der rationalen Reflexion sich die Moderne selbst gebracht hat.

Aus diesem Spannungsfeld – oder man könnte sagen: performativen Widerspruch – heraus konnte sich die Kritik der Moderne höchst produktiv entfalten und selbst die subtilsten Formen entfremdeter Wirklichkeit aufdecken.

Auf eine Situation dieser Art treffen wir auch heute. Zu notieren ist immer noch die Macht, mit der sich die innere Gesetzlichkeit der systemischen Bereiche – repräsentiert durch Technik, Logik und exakte Wissenschaft – Ausdruck verleiht. Dem steht ein hoch entwickeltes Potenzial der Verständigung und der Partizipation gegenüber, das – unter dem Titel des Primats der Politik und des sinnbestimmten und kritischen gesellschaftlichen Diskurses – den Gesellschaftsmitgliedern unserer Zeit ein gegenüber anderen Epochen unverwechselbares Selbstbewusstsein verleiht.

Diese letztgenannte Einschätzung bleibt aber nicht ohne Anfechtungen. Ohne einen Ausblick auf die technische Entwicklung unserer Tage zu haben, sah Arnold Gehlen schon in den Fünfzigerjahren die Gefahr, dass eine der Selbstentwicklung freigegebene Technik das in sich schon fragile Gleichgewicht innerhalb moderner Gesellschaften aufzuheben vermag. Auch zeigt die Erfahrung, in welche Zerreißproben z. B. moderne Gesellschaften durch Techniken wie die der Kernenergie kommen können. Probleme dieser Art sind aber in diesem Kontext nicht vorrangig gemeint. Gemeint ist vielmehr die Frage, inwieweit die neuen Techniken – nicht

zuletzt die, die im Zeichen der *künstlichen Intelligenz* stehen – die *innere Verfassung* gesellschaftlicher Rationalität in struktureller Weise berühren.

Zum neuralgischen Punkt werden dabei die oben angesprochenen *Verständigungsverhältnisse*, deren Gegengewicht zu den machtvollen systemischen Dynamiken als *das* essenzielle Moment anzusehen ist, durch das sich die Gemeinschaft vor tiefgreifenden Regressionen schützt, Herausforderungen bewältigt und neue Sinnstiftungen entfaltet. Die Frage ist, ob anstelle solcher politischen, kulturellen und existenziellen Verständigungen zunehmend 'technische Kommunikationen' Raum greifen bzw. ob die nicht-technischen gesellschaftlichen und lebensweltlichen Steuerungsvollzüge informationstechnisch okkupiert werden.

Um ein Beispiel zu nennen: Norbert Wiener etwa benutzte in seinen Grundlegungen zur Kybernetik das Bild des *Nervensystems*, um die Nähe des Neuen zu dem 'kommunikativ' mit seiner Umwelt verbundenen Lebewesen zu illustrieren. Zu fragen wird sein, ob dieser informationsgesteuerte Austausch mit der Umwelt tatsächlich angemessen als intersubjektiv verstandene *Kommunikation* zu greifen ist.

Unter diesem Gesichtspunkt kommt dem amerikanischen Gesellschaftstheoretiker George Herbert Mead das Verdienst zu, den evolutionären Gang des interaktiven Verhaltens bis hin zur Stufe menschlicher Kommunikation nachzuzeichnen, um deren *eigene* Charakteristik und *geltungstheoretische* Verankerung sichtbar zu machen. Deutlich wird hier, dass der Begriff der Kommunikation bzw. diskursiven Verständigung im Zuge der durchgeführten Untersuchung zu den Grundlagen des menschlichen Geistes eine herausgehobene Bedeutung erhält.

Demgegenüber verdichtet sich aber auch das Bild, dass sich Computerprogramme in hohem Maße differenzieren lassen und damit der Zuschnitt der universalen Maschine auf unendlich viele Spezifikationen möglich wird. In der Folge erreicht deren Bündelung und Vernetzung ein bisher nicht vorstellbares Niveau der Komplexität. Eindrucksvoll ist dabei nicht so sehr die "Blackbox" des Computers, sondern das immer dichter geknüpfte Netz eines Gesamtsystems, dessen interne technische Kommunikation zunehmend mit dem Anspruch behaftet wird, den vielschichtigen Aufbau von Gehirnprozessen nachzuzeichnen.

Um diese konnektiven Vollzüge tatsächlich technisch reproduzieren zu können, wäre nicht zuletzt das Erreichen einer Stufe der internen Steuerung notwendig, die schon Alan Turing durch die Notwendigkeit einer Verschränkung von 'Disziplin' und 'Selbsttätigkeit' gekennzeichnet hat. Denn den Leistungen des menschlichen Geistes ist zu unterstellen, dass

sie Momente der *Autonomie* und *Kreativität* – etwa im Sinne von Wahl- bzw. Willensfreiheit – aufweisen, die – soweit absehbar ist – die *künstliche Intelligenz* vor erkennbare Probleme stellen.

Insbesondere fällt es schwer, die hier wirksamen Momente des *Bedeutungsverstehens*, des *Selbstbewusstseins* und der *rationalen Geltung* dem Computer zuzueignen. Und dennoch reicht es nicht, die Entscheidung auf dieser Ebene der intuitiven Grenzziehung zu fällen. Die Erfahrung weniger Jahrzehnte, in denen die Informationstheorie grundagentheoretisch wie auch technisch so viel Beeindruckendes hervorgebracht hat, nötigt dazu, die Konkurrenz ernst zu nehmen, die nicht zuletzt der Philosophie durch das Paradigma der – empirisch und informationstheoretisch ausgerichteten – Kognitionswissenschaft zugewachsen ist.

Da es nun um Grundsätzliches geht und neben den zentralen Fragen der Wissenschaftstheorie die rationalen Grundbedingungen menschlichen Sprechens, Redens und Handelns berührt sind, vollzieht sich die Erörterung notwendig auf dem Boden der Philosophie. Dies in den wesentlichen Punkten zu vergegenwärtigen, wird im ersten Abschnitt des *ersten Kapitels* der folgenden Abhandlung die Aufgabe sein. Es geht um den Umriss derjenigen Fragestellungen, die für die Philosophie unter dem Eindruck des sich machtvoll Geltung verschaffenden 'Computerparadigmas' zur Herausforderung werden. Im Zentrum findet sich das Problem der sinnbestimmten *Rationalität*. Unter ihrem Titel formieren sich differente Ansprüche, die je nach Ausgangslage in Konkurrenz zueinander treten. Das die gesamte philosophische Debatte dieses Jahrhunderts prägende Gegenüberstellung von Szientismus und Hermeneutik wird hier den Rahmen abstecken, innerhalb dessen die Ansprüche der *künstlichen Intelligenz*<sup>5</sup> geklärt und diskutiert werden.

Im zweiten Abschnitt des ersten Kapitels wird zu zeigen sein, inwiefern Erörterungen dieser Art die Selbstverständigung der Moderne von Anfang an in den Bann ziehen. Die aus einer *vereinseitigten* Rationalität rührenden 'Entzweigungen' des modernen Lebens sind schon für Hegel der Stachel für tief reichende Reflexionen. Von da an kreist das philosophische Denken nicht zuletzt um jenen Punkt, der durch die *rational* gesteigerte Kapazität der 'Verfügung' und Berechnung markiert ist.

---

5 Fortan wird mit dem Terminus der *künstlichen Intelligenz* – eben auch in dieser Schreibweise – das Gesamtprojekt gemeint sein, das unter der Zielvorgabe der Realisierung künstlicher Intelligenz steht.

Muster dieser Argumentation werden insbesondere in dem Disput um die *künstliche Intelligenz* wiedererkennbar sein. Dabei soll ersichtlich werden, dass die *künstliche Intelligenz* – gewissermaßen als einer der Höhepunkte der die Moderne kennzeichnenden Entwicklung – diesen Problemgehalt noch einmal bis zu einem kritischen Punkt zuspitzt.

Reflexionen dieser Art nehmen ihren Ausgang nicht nur von einer philosophischen Selbstverständigung der Moderne, sondern auch von den Ansprüchen der *künstlichen Intelligenz* selbst. So sollte es aufschlussreich sein, sich die wissenschaftstheoretische Einstellung zu vergegenwärtigen, auf der die *künstliche Intelligenz* bzw. die sie untermauernde Kognitionswissenschaft fußt. Deutlich zu machen ist hier, dass anhand der Grenzen, auf die der logische Positivismus stößt, auch Beschränkungen sichtbar werden, denen die *künstliche Intelligenz als solcher* unterliegt.

Um zu sehen, welches Potenzial das Gesamtprojekt der *künstlichen Intelligenz* schon entfaltet hat, soll das zweite Kapitel diesem Thema in angemessener Breite gewidmet sein. Sein erster Abschnitt wird die beiden Entwicklungsebenen beleuchten, auf denen sich die Forschung zu diesem Projekt bewegt. Mit den Stichworten des 'Kognitivismus' bzw. 'Konnektionismus' sind die beiden wichtigsten Forschungsvorhaben benannt, die sich einer Realisierung künstlicher Intelligenz verschrieben haben. Von Anfang an verbindet sich in diesen Ansätzen der grundagentheoretische Teil mit jener experimentellen Ebene, von der die empirischen Belege erwartet werden. In diesem Abschnitt soll entsprechend ein Überblick darüber gegeben werden, a) welche Geschichte diese Ansätze aufweisen, b) welche konkreten Realisierungen sie anstreben und c) auf welche konzeptionellen und technischen Schwierigkeiten sie treffen.

Zu den bemerkenswerten Tatsachen innerhalb der Entwicklung der *künstlichen Intelligenz* gehört, dass zwei ihrer profiliertesten Protagonisten, Joseph Weizenbaum und Terry Winograd, aus der Erfahrung heraus, welche zugespitzte Hochschätzung das Gesamtprojekt erfährt, zu seinen profilierten Kritikern wurden. Die Gegenüberstellung von Verfechtern und Gegnern der *künstlichen Intelligenz* – einschließlich quer liegender Positionen – soll die grundsätzlichen Fragestellungen erkennbar machen, an denen sich die Debatte abarbeitet. Die Beiträge stellen sich dabei dem Problem, dass die im Kern technischen Weichenstellungen eine Wirkung darauf erzielen, welchen Fortgang die wissenschaftliche Kultur unseres Zeitalters nimmt.

Aus dem Gesagten mag deutlich werden, dass das Thema sehr umfassend ist und mithin Eingrenzungen notwendig sind. Der Rückgriff auf den amerikanischen Philosophen John Searle hilft dabei abzuzeichnen, welch

thematischer Zuschnitt sich nahelegt. In diesem Sinne dient der zweite Abschnitt des zweiten Kapitels dazu, den philosophischen Kern der Fragestellung deutlich zu machen. Neben den Fragen nach dem 'Bewusstsein', der 'Willensfreiheit' und dem 'Leib-Seele-Problem' tritt insbesondere die nach der Konstitution von 'Bedeutung' in den Mittelpunkt. Insbesondere Searles These, dass Computer prinzipiell nicht verstehen, was sie tun, lässt fragen, aus welchen Bedingungen heraus der Mensch genau diese ihn auszeichnende Leistung vollbringt.

Der sogenannte *Funktionalismus* kann nun unter diesen herausgestellten Gesichtspunkten dadurch Profil gewinnen, dass er – gewissermaßen zwischen Naturalismus und Metaphysik – eine Ebene verortet, auf der er den geistigen Vollzügen ihre intentionale Charakteristik lässt und sie gleichzeitig nach dem Modell der Symbolverarbeitung rekonstruiert. Sollte es sein, dass sich für *jeden* geistigen Zustand ein kalkülmäßiger Zustand angeben lässt, dann müsste dies auch für diejenigen geistigen Vollzüge gelten, anhand deren der Mensch die *Bedeutung* beliebiger Sachverhalte generiert. Diese Grundüberlegung wird insbesondere dem amerikanischen Philosophen Hilary Putnam zugeordnet. Noch einmal aber wird mit ihm einer der Wortführer der *künstlichen Intelligenz* zum Dissidenten, denn Putnam kehrt sich in jüngerer Zeit nachdrücklich gegen den ursprünglich eigenen Gedanken einer *funktionalen* Erfassbarkeit des menschlichen Geistes.

Die philosophische Behandlung des Projektes der *künstlichen Intelligenz* stößt aber noch auf noch andere Schwierigkeiten, die der Philosoph Hans Jonas identifiziert. Erwägt man, dass Computern Intelligenz – im Sinne eines geistigen Vermögens – zukommen könnte, dann stellt sich die Frage, wie dies auf der Grundlage des Silicium-Chips vorzustellen ist. Es zeigt sich aus dieser Erörterung heraus, dass das philosophisch zugrunde liegende und weit zurückreichende Problem des Verhältnisses von *Leib und Seele* bzw. Körper und Geist nicht gelöst ist.

Der schon erwähnte Philosoph John Searle stellt unter diesem Gesichtspunkt die konkretere Frage, wie Menschen solchen Dingen (KI-Systemen) Intentionalität verleihen, die von sich aus keine haben. Dies führt zurück auf die Überlegung, ob Bedeutungen sich tatsächlich im Gehirn bzw. 'im Kopf' konstituieren und damit die eigenartig anmutende Konstruktion vom 'Gehirn im Tank' (Putnam) einen realen Gehalt hat.

In dem vierten Abschnitt des zweiten Kapitels ist in diesem Sinne zu fragen, inwieweit 'Bedeutung' auf einen solipsistisch angelegten Begriff der Intentionalität zurückführbar, d. h. 'im Kopf' zu verorten ist. Gelänge es

stattdessen, Putnams Argumentation dadurch zu stärken, dass sich die Konstitution von Bedeutung tatsächlich in den *Grundbegriffen intersubjektiver Verständigung* verankern lässt, dann wäre der gesamten Erörterung ein neuer Weg gewiesen. Diese Spur gilt es zu verfolgen, um Aufschluss darüber zu gewinnen, ob das geistige Vermögen des Menschen – ebenso wie Bewusstsein und Selbstbewusstsein – aus den semantischen Strukturen intersubjektiver Verständigung heraus rekonstruierbar ist. Das *gemeinsam geteilte In-der-Welt-Sein* könnte dann genau jene eine – menschliche – Rationalität freisetzen, deren Eigenart es nicht zuletzt ist, dass sie als einzige – in Rückwendung auf sich selbst – ihre Grenzen und Möglichkeiten bestimmen kann.

Interessant wird in diesem Zusammenhang weiterhin sein, ob sich mit formalen beweistheoretischen Mitteln – hier das Gödel-Theorem – signifikante Unterschiede zwischen menschlicher und maschineller Logik nachweisen lassen. Ein Exkurs wird in zusammenfassender Form auf den Weg dieses Gedankengangs führen.

Die Maßstäbe der Vernunft mit Blick auf das zu entwickeln, was die *künstliche Intelligenz* in die Wege leitet, erfordert die 'Anstrengung des Begriffs' im Sinne einer Reflexion auf die Herausforderungen, denen sich die Gesellschaft durch das groß angelegte Projekt der Informationstechnik ausgesetzt sieht. Inwieweit Computer, die auf der Linie der *künstlichen Intelligenz* entwickelt werden, das Band gesellschaftlicher Rationalität stärken oder schwächen, wird in einem abschließenden Abschnitt behandelt. Die Unterscheidung zwischen schlichtem Computer und *künstlicher Intelligenz* verwischt sich an *diesem* Punkt, weil aus gesellschaftstheoretischer Sicht hier ein Kontinuum besteht.

Zu hoffen ist, dass sich die doch große Spannbreite zwischen grundsätzlichen Erwägungen und konkreten Erörterungen in dem Sinne als produktiv erweist, dass die Linien sichtbar werden, welche die Grundstrukturen menschlicher bzw. gesellschaftlicher Sinnbestimmung und Rationalität mit den greifbaren technischen Gestaltungsperspektiven verbinden.



## 2 Ausgangsfragen – Theorie der Rationalität – Aufgabe der Philosophie

Wenn heute die umgreifenden Zusammenhänge der modernen Technik in das Blickfeld kulturtheoretischer Erörterungen geraten, dann stellen sich erstaunliche Übereinstimmungen in der Bestimmung der Ausgangs- und Problemlage ein. Paradigmatisch sei hier Hans Jonas zu nennen, der feststellt, "dass die Verheißung der modernen Technik in Drohung umgeschlagen ist oder diese sich direkt mit jener unlösbar verbunden hat"<sup>6</sup>. Wenn Technik nun in dieser Form zum Problem geworden ist, stellt sich die Frage, wie es sich verhindern lässt, dass aus der Bedrohung Wirklichkeit wird, und welche Rolle die Philosophie im Rahmen dieser Bemühung spielen kann.

Der Zugang der Philosophie zum Thema ist zunächst und auf weite Sicht ein *mittelbarer*, d. h., im Interesse einer Verbindlichkeit erzeugenden Argumentation muss sie zunächst den Status und die Maßstäbe ihrer Beiträge klären. Zwar findet sich hierin einer der Gründe, warum der Philosophie oft praktische Unwirksamkeit vorgeworfen wird, aber auch unmittelbar handlungsrelevantes Wissen muss in den vorausgehenden Begründungsschritten schlüssig sein. Dabei ist selbst die Frage zu klären, warum die Philosophie berufen ist, sich immer mehr jenen Problemen zuzuwenden, die im Zusammenhang der technischen Gestaltung moderner Lebenszusammenhänge entstehen.

Es sind nun nicht erst die charakteristischen Bedrängnisse, in die moderne Gesellschaften durch Hochtechnologien – wie Computer-, Kern- und Gentechnik – geraten, die die Philosophie auf das Thema aufmerksam machen. Schon 1949 stellte Karl Jaspers fest, die Technik sei "heute vielleicht das Hauptthema für die Auffassung unserer Lage" und "ihre Folgen für schlechthin alle Lebensfragen" könne man "gar nicht überschätzen".<sup>7</sup> In welchem hohem Maße 'weltabgewandt' müsste sich demnach die Philosophie erweisen, würde sie sich dieser Herausforderung nicht stellen.

Es war insbesondere Georg Wilhelm Friedrich Hegel, der in seinem berühmten Diktum der Philosophie auftrag, *ihre Zeit in Gedanken zu fassen*.

---

6 H. Jonas, Das Prinzip Verantwortung, Frankfurt/M. 1979, S. 7

7 K. Jaspers, Vom Ursprung und Ziel der Geschichte, Zürich 1949, S. 98.

Die 'Entzweigungen', die die Moderne aus sich hervorbringt, wurden ihm zum Stachel und ihre mögliche Überwindung zum "Quell des Bedürfnisses der Philosophie".<sup>8</sup>

Mit dem Anspruch, die 'Lebensfragen' ihrer Zeit auf den Begriff zu bringen, sieht sich die Philosophie im gleichen Maße gefordert, aus der Diagnose dessen, was *ist*, Perspektiven für eine zukünftige Gestaltung zu gewinnen. Unumgebar stoßen wir dabei auf die Grundfrage der Ethik: "Was sollen wir tun?"

Angesichts der Komplexität des Themas ist nun nicht daran zu denken, dass es hierauf in unserer Zeit simple Antworten gibt. Mit Blick auf eine alle Lebensräume durchdringende Technik jedoch, die sich – wie es scheint – eigendynamisch und angetrieben durch wirtschaftliche Imperative entwickelt und dabei möglicherweise irreversible Wirkungen hervorbringt, erscheinen bohrende Fragen im Sinne der genannten ethischen Erwägung sehr notwendig: "Der endgültig entfesselte Prometheus, dem die Wissenschaft nie gekannte Kräfte und die Wirtschaft den rastlosen Antrieb gibt, ruft nach einer Ethik, die durch freiwillige Zügel seine Macht davor zurückhält, dem Menschen zum Unheil zu werden."<sup>9</sup>

Nichts anderes wird damit zur Zielvorgabe für eine philosophische Erörterung als die, "die praktische Notwendigkeit einer ethisch-politischen Grundorientierung gerade in unserem durch Wissenschaft und Technologie geprägten Zeitalter [zu] illustrieren".<sup>10</sup> Nicht zuletzt sieht sich die Philosophie damit aber vor die Aufgabe gestellt, überhaupt die *Möglichkeit einer verbindlichen ethischen Grundorientierung* nachzuweisen.

Erkennbar bewegen wir uns damit ein ganzes Stück weit weg von jenen konkreteren Fragen, in denen es um die Auswirkungen der modernen hoch entwickelten Technik – genauer gesagt der fortgeschrittensten Form der Kommunikations- und Informationstechnologie – auf die Vielgestalt unserer Lebensformen geht. So scheint für die Philosophie jener Spagatschritt unvermeidlich zu werden, der zwischen der argumentativen Explikation eigener Maßstäbe und der anschließenden Stellungnahme zu den Großprojekten unserer Technikgesellschaft liegt, an die ja nicht zuletzt auch die Hoffnung geknüpft ist, mit ihrer Hilfe ließe sich die hoch ausgestattete Form moderner Lebensgestaltung in die Zukunft fortschreiben.

---

8 G.W.F. Hegel, Jenaer Schriften 1801–1807, Suhrkamp-Werkausgabe, Bd. 2, Frankfurt/M. 1970, S. 20.

9 H. Jonas, Das Prinzip Verantwortung, S. 7.

10 K.-O. Apel, Diskurs und Verantwortung, Frankfurt/M. 1988, S. 22.

Wie kritisch und wohlfundiert ethische Maßstäbe auch immer sein mögen, ihnen bleibt doch das Nachsehen, wenn sie in nur *abstrakter* Weise einer in sich vielschichtigen Wirklichkeit entgegengehalten werden. Daraus folgt nichts anderes, als dass die Philosophie sich auf den Gegenstand ihrer Beurteilung in möglichst vielen relevanten Punkten 'einlassen' muss.

In unserem Fall – um dies deutlich zu machen – geht es darum, die wesentlichen Aspekte herauszuarbeiten, unter denen die neuesten Perspektiven der Computertechnologie – zusammengefasst unter dem Stichwort *künstliche Intelligenz* – zur Herausforderung für die Philosophie werden.

Die Frage, was Computer heute schon können und wie sie arbeiten, verbindet sich mit der prinzipiellen Erwägung, in welche Richtung die Entwicklung der in dieser Technik angelegten Potenziale geht und inwieweit sich 'intelligente' Computer dem Menschen in seiner Bestimmung als *geistigem Wesen* annähern.

Von Menschen nehmen wir an, dass sie – unter günstigen Rahmenbedingungen – *wissen, was sie tun*<sup>11</sup>, dass sie *verstehen*, was andere zum Ausdruck bringen, dass sie *interpersonale Beziehungen* eingehen und die *Bedeutung* symbolischer Äußerungen und Akte erfassen können. Weiterhin nehmen wir an, dass Menschen einen – privilegierten – Zugang zu ihren *inneren Erlebnissen* haben und dass ihnen – wie rudimentär auch immer – ein *ästhetisches Empfinden* und tief sitzende *moralische Intuitionen* zu eigen sind. Diese Liste ließe sich vielfach erweitern. Psychologie, Anthropologie, Soziologie und viele andere Einzelwissenschaften könnten weiter gehende Ergänzungen anfügen, und in der Tat ist es nicht leicht herauszuarbeiten, welche Eigenschaften und Fertigkeiten es sind, die den Menschen *im Kern* von hoch entwickelten Automaten unterscheiden.

Wir kennen Beispiele, in denen Tiere so etwas wie Empathie zeigen, mit Symbolen wie auch mit Werkzeugen umzugehen vermögen und Gefühle äußern, die uns vertraut vorkommen. Dennoch wären wir weit davon entfernt, sie als *geistige Wesen* anzuerkennen. Eher geschieht es, dass Charakterisierungen dieser Art – wenn auch noch in einem *metaphorischen* Sinn – schon auf Computer angewendet werden; gemeint sind insbesondere sol-

---

11 Die Psychologie z. B. lehrt uns, dass diese Unterstellung nicht unproblematisch ist. Sie ist hier in dem grundlegenden Sinne gemeint, dass unsere Handlungen derart von Bewusstsein begleitet sind, dass wir im Prinzip Auskunft darüber geben können, wie sie gemeint sind.

che Systeme, die auf dem Weg zur *künstlichen Intelligenz* den gegenwärtig höchsten Entwicklungsstand repräsentieren.<sup>12</sup>

Dennoch hat sich die Idee von der 'Sonderstellung' des Menschen eine ihr eigene Widerstandsfähigkeit bewahrt. Aber die Dinge geraten in Bewegung. Die Faszination, die von den verblüffend reibungslosen, spielerisch leichten und doch präzisen Abläufen in den – bald mikroskopisch kleinen – Prozessoren ausgeht, öffnet den Weg für Gedankenexperimente, denen die Idee der technischen Reproduktion geistiger Vollzüge längst nicht mehr blasphemisch erscheint. "Der Computer-Jargon ist der Jargon des Denkens."<sup>13</sup>

Die Pointe dieser Projektion mündet – so die Autorin – schließlich in eine eigentümliche Verdrehung: Nicht mehr die geistigen Vollzüge des Menschen geben als oberste Instanz die Maßstäbe für das ab, was sich approximativ als 'intelligent' erweist, sondern die in ihren Schrittfolgen und Vernetzungen klar durchschaubaren Prozessvollzüge des Computers gewinnen Modellfunktion für die Explikation menschlichen Denkens.<sup>14</sup>

Der Verdacht des 'Reduktionismus' liegt an dieser Stelle nahe und doch lässt sich die genannte Idee nicht von vornherein von der Hand weisen. Sie führt uns zumindest zu jener Grundfrage, die es Wert wäre, in den von Kant aufgestellten Kanon von Grundfragen der Philosophie aufgenommen zu werden: "Was ist der menschliche Geist?" Vergegenwärtigt man sich aber, dass die in den drei Kant'schen Kritiken sich ausdrückende Architektonik der menschlichen Vernunft Auskunft über die Reichweite und innere Differenzierung des rationalen geistigen Vermögens des Men-

---

12 Wie Sherry Turkle nachdrücklich gezeigt hat, beschränkt sich die "Vergeistigung" und "Vermenschlichung" von Computern bei Weitem nicht auf die fortgeschrittensten Produkte dieser Technik. Schon der Personal Computer und auch normale Spielprogramme wecken Fantasien, die anzeigen – so der Gedanke der Autorin –, dass der Computer sich im "Grenzbereich zwischen Natur und Kultur" ansiedelt. "Seine evokatorischen Eigenschaften hängen nicht ab von Spekulationen über denkbare Erfolge bei der Entwicklung künstlicher Intelligenz, einer Forschung, deren Ergebnisse es am Ende möglich machen könnten, Maschinen zu konstruieren, die ein Abbild des Menschen sind. Sie sind vielmehr abhängig von der Tatsache, dass Menschen dazu neigen, eine 'Maschine, die denkt' wahrzunehmen als ein 'Wesen, das denkt'. Sie fangen an, über die Tätigkeiten jener Maschine in psychologischen Begriffen nachzudenken". Sherry Turkle, *Die Wunschmaschine. Der Computer als zweites Ich*, Hamburg 1986, S. 25.

13 Ebd. S. 14.

14 "Der Computer ist eine 'metaphysische Maschine', eine 'psychologische Maschine', und zwar nicht nur deshalb, weil man sagen könnte, er habe eine Psyche, sondern weil er Einfluss darauf hat, wie wir über uns selbst denken." Ebd. S. 13.