

INTELLIGENZ IM ZEITALTER IHRER TECHNISCHEN REPRODUZIERBARKEIT. ZUR DIALEKTIK ARTIFIZIELLER VERNUNFT

INTELIGÊNCIA NA ERA DE SUA REPRODUTIBILIDADE TÉCNICA:
SOBRE A DIALÉTICA DA RAZÃO ARTIFICIAL

Alexander Schubert*

ABSTRACT: The steady rise of generative AI raises the question of what impact the technical reproducibility of intelligence has on our thinking in general and our understanding of reason in particular. There are different opinions on the question whether AI can develop something like a humanoid consciousness and whether it represents an opportunity or a threat for the future. In this essay, I discuss this question with reference to the dialectic of (self-)consciousness as laid out by Hegel in his philosophy of mind and subjective spirit. I will examine the thesis that artificial intelligence represents the consequent further development of what Hegel called “verständiges Denken”. Building on Max Horkheimer’s critique of instrumental reason and Walter Benjamin’s concept of technological reproducibility of art, I will outline the commonalities and differences between natural and artificial intelligence, taking into account contemporary debates about the potential consciousness of AI or AGI. I will show that the dialectical ambiguity of the former continues in the latter and that emancipatory and totalitarian tendencies go hand in hand in both. In contrast to current approaches, which usually view AI from a subject-centered perspective of cognitive science or action theory, a systemic-dialectical concept of AI based on social interactivity can enable a critical position that does not play natural and artificial intelligence off against each other, but recognizes the emancipatory potential in both and helps to protect both from the danger of totalitarian thinking.

KEYWORDS: AI / AGI; Human versus Artificial Intelligence; (Self-)Consciousness; Understanding Thinking / Instrumental Reason; Philosophy of Mind and Subjective Spirit; Hegel; Benjamin; Horkheimer

RESUMO: O crescimento constante da IA generativa levanta a questão de que impacto a reprodutibilidade técnica da inteligência tem sobre nosso pensamento em geral e nossa compreensão da razão em particular. Existem diferentes opiniões sobre a questão de saber se a IA pode desenvolver algo como uma consciência humanoide e se ela representa uma oportunidade ou uma ameaça para o futuro. Neste ensaio, discuto essa questão com referência à dialética da (auto)consciência conforme exposta por Hegel em sua filosofia do espírito e do espírito subjetivo. Examinarei a tese de que a inteligência artificial representa o desenvolvimento consequente daquilo que Hegel chamou de “verständiges Denken” (pensamento do entendimento). Baseando-me na crítica de Max Horkheimer à razão instrumental e no conceito de Walter Benjamin sobre a reprodutibilidade tecnológica da arte, delinearei as semelhanças e diferenças entre inteligência natural e artificial, levando em conta os debates contemporâneos sobre a consciência potencial da IA ou IAG. Mostrarei que a ambiguidade dialética da primeira continua na segunda e que tendências emancipatórias e totalitárias caminham lado a lado em ambas. Em contraste com as abordagens atuais, que geralmente veem a IA de uma perspectiva centrada no sujeito da ciência cognitiva ou teoria da ação, um conceito sistêmico-dialético de IA baseado na interatividade social pode possibilitar uma posição crítica que não coloca inteligência natural e artificial uma contra a outra, mas reconhece o potencial emancipatório em ambas e ajuda a proteger ambas do perigo do pensamento totalitário.

PALAVRAS-CHAVE: IA / IAG; inteligência humana *versus* inteligência artificial; (auto)consciência; pensamento do entendimento / razão instrumental; filosofia do espírito e espírito subjetivo; Hegel; Benjamin; Horkheimer

* Dr. Alexander Schubert earned a master’s degree in sociology and a doctorate (PhD) in philosophy from Goethe University in Frankfurt am Main, Germany. He is a private scholar and author of philosophical publications, and a member of the Internationale Hegelvereinigung and the Internationale Hegel-Gesellschaft. E-mail: alex.schubert.1@icloud.com. ORCID: <https://orcid.org/0000-0002-1834-3787>



Einleitung

Als Walter Benjamin im Jahr 1936 seinen Aufsatz *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*¹ veröffentlichte, beschrieb er damit den Prototyp einer technologisch bedingten Disruption in der Produktion kultureller Erzeugnisse, welche gravierende Auswirkungen auf die gesellschaftliche Rezeption von Kunst und letztlich auch auf Verhaltensweisen innerhalb der sozialen und politischen Verhältnisse hat. Im Gegensatz zur damals verbreiteten kulturpessimistischen Technikfeindlichkeit auf der einen und einer unkritisch futuristischen Technikeuphorie auf der anderen Seite untersuchte er die eigentümliche Dialektik des technischen Fortschritts, wie sie sich am Beispiel der Entwicklung von Film- und Radioproduktion darstellte. Diese Entwicklung beinhaltete auf der einen Seite den Verlust der Autonomie des klassischen Kunstwerks und damit verbunden seiner „Aura“ sowie die Überwindung der Kultur des individuellen Sich-Versenkens in das Kunstwerk zugunsten der Zerstreuung und potenziellen Mobilisierung der Massen. Auch wenn damit die Gefahr einer totalitären Ästhetisierung von Massenbewegungen besteht, wie es der Faschismus und auch der Stalinismus letztlich auch bestätigte, sieht Benjamin in dieser Disruption auch ein kritisches, emanzipatorisches Potenzial, auf das wir später (in Kapitel 5) noch einmal zurückkommen werden.

Eine ähnliche Dialektik wie Benjamin seinerzeit bei der Kunstproduktion beobachten wir heute im Bereich der Denkproduktion, nämlich der technischen Reproduzierbarkeit von Intelligenz. Auch wenn „Disruption“ heute ein Marketing-Buzzword der Technologie-Apologeten des Silicon Valley ist, muss man zugestehen, dass es sich bei künstlicher Intelligenz (KI) tatsächlich um eine technologische Disruption handelt, und zwar um eine, die sich noch im Anfangs- und Entwicklungsstadium befindet und deren Chancen und Risiken derzeit noch nicht vollumfänglich abzuschätzen sind.

Selbst ausgewiesene Experten aus Theorie und Praxis sind sich nicht einig, ob die technologische Entwicklung eher ein Fluch oder ein Segen ist. Die einen sehen eine Gefahr darin, dass die auf künstlicher Intelligenz basierenden Technologien sich so dynamisch entwickeln, dass sie potenziell der wissenschaftlichen und mehr noch der sozialen Kontrolle entgleiten, und dass KI potenziell in der Lage sein könnte, eine totalitäre Weltherrschaft zu erringen. Andere sehen eher das Potenzial darin, durch hyperintelligentes Datenmanagement und Selbstperfekti-

¹ In: BENJAMIN, W., **Gesammelte Schriften (GS)**, Bd. I.2, Frankfurt am Main, Suhrkamp 1974, pp. 431-508 (zwei Fassungen).

onierung mögliche Fehler auszumerzen, die in der Natur des Menschen liegen. Es scheint unstrittig, dass wie bei jeder neuen Technologie Chancen und Risiken Hand in Hand gehen, uneinig ist man sich aber darüber, was von beidem überwiegt. Auch gibt es eine kontroverse Einschätzung, ob mögliche Risiken in der Natur des Konzepts künstlicher Intelligenz an sich liegen, oder lediglich in der Anwendung durch den Menschen, was sozusagen die Möglichkeit des *dual use* beinhalten würde: „An sich“ sei die Technologie neutral, aber je nachdem ob sie in die richtigen oder falschen Hände geriete ...

In gewisser Weise haben sowohl die Position der Apologeten als auch die der Kritiker ihre Berechtigung. Denn ist es nicht gerade die Dialektik des Fortschritts generell, dass darin emanzipatorische und totalitäre Tendenzen Hand in Hand gehen? In seiner *Kritik der instrumentellen Vernunft* spricht Horkheimer vom „widerspruchsvollen Charakter eines jeden Fortschritts“, welcher immer wieder in Barbarei umzuschlagen droht.² Und umgekehrt: Könnte diese Dialektik etwa auch bedeuten, dass sich gemäß Hölderlins Bonmot in den Gefahren der technologischen Entwicklung auch etwas „Rettendes“ ankündigt?³

In diesem Essay möchte ich diese Dialektik am Beispiel von künstlicher Intelligenz und ihrem Verhältnis zur natürlichen Intelligenz näher beleuchten. Dabei werde ich auch der Frage nachgehen, was die technische Reproduktion von Intelligenz für das Verständnis von „Intelligenz“ generell bedeutet. Und ob deren technische Reproduzierbarkeit nicht sogar darauf verweist, dass „künstliche“ Intelligenz mehr Gemeinsamkeiten mit der „natürlichen“ Intelligenz hat, als man gemeinhin vermuten würde. Das beinhaltet auch die Frage, ob das Problematische an der künstlichen in Wahrheit nicht auch auf eine Problematik der natürlichen Intelligenz verweist. Ich werde mich dabei unter anderem auf Hegels Philosophie des Geistes beziehen, da diese ein Instrumentarium bereit hält, die Dialektik der — menschlichen wie künstlichen — Intelligenz in ihrer Dynamik zu begreifen.

1. Die Streitfrage: Kann künstliche Intelligenz ein Bewusstsein entwickeln?

Bewusstsein [...] beinhaltet die Fähigkeit, sich selbst und die Umgebung zu erkennen, zu verstehen und zu erleben. Menschen haben ein Bewusstsein, das mit subjektiven Erfahrungen, Selbstrefle-

² KIV, p. 141.

³ „Wo aber Gefahr ist, wächst das Rettende auch.“ (HÖLDERLIN, F., Patmos, In: **Sämtliche Werke**, hg. v. F. Beissner and A. Beck. Band 6, Stuttgart, Kohlhammer, 1975, p. 328.)

xion und der Fähigkeit zur Intentionalität verbunden ist, also der Fähigkeit, Absichten zu haben und Ziele zu verfolgen.

(*ChatGPT*)⁴

Nachdem das Thema künstliche Intelligenz in den letzten Jahren vor allem in wissenschaftlich-technisch, politisch oder ökonomisch orientierten Publikationen debattiert wurde, hat es jetzt auch in der breiteren philosophischen Diskussion Fahrt aufgenommen. Das heißt nicht, dass in der Vergangenheit nicht auch einzelne Wissenschaftler das Phänomen der KI unter philosophischen Gesichtspunkten diskutiert hätten. Diese, wie z.B. Jehoschua Bar-Hillel in den 1950er oder Hubert L. Dreyfus in den 1970er Jahren legten den Schwerpunkt zumeist auf die Abgrenzung der KI zur menschlichen Intelligenz, um aufzuzeigen, zu was Computer oder Maschinen (immer noch) nicht in der Lage sind.⁵

Da die Technologie sich jedoch seither signifikant weiterentwickelt hat und mit Large Language Modellen (LLMs) wie *ChatGPT*, *xAI* oder *Deepseek* in der breiten Bevölkerung angekommen ist, und posthumanistische KI-Apologeten wie Ray Kurzweil der KI in nicht allzu ferner Zukunft die Überlegenheit gegenüber dem Menschen prophezeien,⁶ stellt sich die Frage neu, wie intelligent künstliche Intelligenz eigentlich ist oder, im besten Fall, wie intelligent sie in Zukunft sein könnte. Insbesondere, ob KI in Zukunft auch so etwas wie ein Bewusstsein, Selbstbewusstsein oder gar Machtbewusstsein entwickeln könnte.

Dabei sind viele Experten auch heute noch der Ansicht, künstliche Intelligenz sei „nowhere near developing systems that are genuinely intelligent“, wie Brian Cantwell Smith sagt.⁷ Dabei wird in dieser Debatte oft auf Kognitions- und Neurowissenschaften Bezug genommen. Anton Friedrich Koch zitiert etwa den Neurowissenschaftler Johann Grolle, der zu dem Schluss kommt, dass selbst Insekten oder Oktopusse mehr Bewusstsein besitzen als Computer, die dieses lediglich simulieren oder vortäuschen.⁸ Emily Bender spricht in ähnlicher

⁴ Antwort von ChatGPT auf die Frage: „Was verstehst du unter Bewusstsein?“ – Aufruf 20.02.2024. – Zur Problematik der wissenschaftlichen Zitierfähigkeit von Chatbots wie ChatGPT vgl. meine Anmerkung 58 weiter unten, p. 16.

⁵ Vgl. DREYFUS, H. L., **What Computers Still Can't Do: A Critique of Artificial Reason**, Cambridge/Ms., MIT Press, 1992 [1972].

⁶ KURZWEIL, R., **Menschheit 2.0. Die Singularität naht**, Berlin, Lola Books, 2014.

⁷ CANTWELL SMITH, B., **The Promise of Artificial Intelligence: Reckoning and Judgement**, Cambridge/Ms., MIT Press, 2019.

⁸ GROLLE, J., Interview in DER SPIEGEL 37/2023. – Vgl. KOCH, A.F., **Artificial Intelligence and the Problem of Consciousness, Introduction**, in www.academia.edu, 2024, p. 1. https://www.academia.edu/113325577/Artificial_Intelligence_and_the_Problem_of_Consciousness_New_Version.

Weise über LLMs als „stochastische Papageien“.⁹ Und Koch argumentiert ebenfalls in diese Richtung: „The inner architecture of programmable automata is too different from that of neuronal systems. Therefore, artificial intelligence is essentially parasitic on human intelligence, which means that there can only be what is called weak AI, not strong, i.e. autonomous, human-like AI.”¹⁰ Dabei versucht er, die Debatte über künstliche Intelligenz über den neuronalen Aspekt hinaus um eine philosophische Perspektive anzureichern, indem er sich vor allem auf die Qualia-Theorien, insbesondere auf Quine, Sellars, Chalmers et al. bezieht. Danach sei die Un erreichbarkeit eines menschenähnlichen Bewusstseins für die KI nicht nur neurowissenschaftlich, das heißt empirisch, sondern auch genuin philosophisch beziehungsweise logisch zu begründen.

Kochs Hauptargument liegt, wenn ich es richtig verstehe, darin, dass künstliche Intelligenz prinzipiell repräsentational sei, während diskursives, intentionales Denken grundsätzlich präsentational sei,¹¹ das heißt auf externe, seiende Dinge, phänomenale Qualitäten, kurz: Qualia gerichtet. Das entspräche dem logos apophantikos bei Aristoteles — eine Prädikation, welche auf die Dinge selbst abzielt und diese erfasst. Ohne qualitatives Bewusstsein, so Koch, gibt es kein diskursives Denken, und ohne eine gewisse Leiblichkeit gibt es keine qualitative Erfahrung externer Dinge. Deshalb, so Koch, kann es keine „starke“ künstliche Intelligenz geben, da diese keinen Zugang zu qualitativer Erfahrung haben kann.¹² Er kommt mithin zu dem Schluss, dass eine starke künstliche Intelligenz logisch unmöglich sei:

[A]rtificial intelligence is inherently representational, whereas strong, i.e. human intelligence is presentational at its basis. In order to create AI, one would have to find a way to connect the infinite spatiotemporal field of conscience with a physical automation in such a way that this automation would literally overlap with us and other sentient beings in this field. It seems impossible to even guess how this could be accomplished or what exactly would have to be accomplished. We may therefore confidently conclude that strong AI is *logically impossible*.¹³

⁹ BENDER, E.M., **On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?**, in: FAccT'21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, New York 2021, pp. 610-623.

¹⁰ KOCH, A.F., ebd.

¹¹ „[A]rtificial intelligence is essentially *re*-presentational, while *presentational* thinking is the mark of intentionality. AI will therefore always be essentially parasitic, inauthentic intelligence, i.e. weak AI.“ Ebd., p. 2.

¹² „[D]iscursive thinking cannot skip or bypass qualitative consciousness and [...] for this reason no strong AI would be possible without qualia.“ Ebd., p. 5.

¹³ Ebd., p. 8 – Hervorhebung von mir.

So plausibel diese These klingen mag, so sehr steht sie jedoch in zweifacher Hinsicht auf schwachen Füßen. Erstens verdankt sich die Behauptung der logischen Unmöglichkeit einer starken KI einer Herleitung, die ausschließlich empirischer Natur ist, nämlich dass wir uns nicht vorstellen können, wie eine Überlappung eines Automaten mit uns und anderen fühlenden Wesen aussehen könnte. Dabei ist diese Herleitung heute bereits empirisch leicht zu widerlegen: gibt es doch im Bereich der KI-Forschung genau das Bestreben, den Konnex zwischen Automation und Empfindung herzustellen, wie es etwa das Neuralink-Projekt von Elon Musk zum Programm erhoben hat.¹⁴ Und zweitens ist sie in theoretischer Hinsicht zirkulär, weil sie als Ergebnis konstatiert, was sie bereits als Definition voraussetzt, nämlich dass KI per se nur repräsentational und nicht präsentational sein kann. Hinter dieser fehlerhaften Zirkularität verbirgt sich jedoch das Problem einer dialektischen Zirkularität, wie sie Hegel als *bestimmende Reflexion* beschrieben hat.

Die Problematik der Intelligenz, wie sie in dieser Argumentation aufgegriffen wird, verweist auf das bereits von Descartes bis Leibniz thematisierte Leib-Seele-Problem, auf welches auch der von Koch zitierte Chalmers Bezug nimmt.¹⁵ Während sich Quine et al. durchaus noch im Horizont der rationalistischen Metaphysik oder des klassischen Nominalismusstreits bewegen, ist Hegels Reflexionslogik darüber jedoch längst hinaus. Denn logisch entspricht das Verhältnis von präsentationalem und repräsentationalem Denken der Dialektik von Setzen und Voraussetzen. Ersteres entspricht der setzenden Reflexion, letzteres der voraussetzenden. In der *Wissenschaft der Logik* weist Hegel nach, dass jede dieser beiden Reflexionsfunktionen ihr Gegenteil an ihr selbst enthält, dass also die Setzung auf Voraus-Setzungen beruht, und dass die Voraussetzung selbst ein Voraus-Setzen ist.¹⁶ Und diese Grundstruktur jeder Dialektik ist nicht nur genuin logischer Natur, sie affiziert auch in hohem Maße alle Elemente der sogenannten Realphilosophie. Das zeigt sich etwa beim Verhältnis von Geist und Natur im dritten Teil der *Enzyklopädie*, wo Hegel die Manifestation des Geistes als das „Setzen der Natur als seiner Welt;

¹⁴ Vgl. hierzu ŽIŽEK, S., *Hegel im verdrahteten Gehirn*, Frankfurt am Main, S. Fischer, 2020, p. 22 ff. – Aktuellen Presseveröffentlichungen zufolge sollen bei diesem Konnex auch schon erste Erfolge erzielt worden sein. Vgl. <https://www.zeit.de/wissen/2024-01/elon-musk-neuralink-implantat-chip-gehirn>.

¹⁵ „[W]e can say that a mental state is conscious if it has a *qualitative feel* – an associated quality of experience. These qualitative feels are also known as *phenomenal qualities*, or qualia for short. The problem of explaining these phenomenal qualities is just the problem of explaining consciousness. This is the really hard part of the mind-body-problem.” CHALMERS, D.J., *The Conscious Mind*, zit. bei KOCH, a.a.O., p. 2. Dieses Statement belegt ebenfalls die Zirkularität der Argumentation, die es zu überwinden trachtet.

¹⁶ Vgl. SCHUBERT, A., *Der Strukturgedanke in Hegels ‚Wissenschaft der Logik‘*, Berlin, Eule der Minerva, 2021, p. 117 ff., insbes. 125 ff.

ein Setzen, das als Reflexion zugleich Voraussetzen der Welt als selbständiger Natur ist“, beschreibt.¹⁷ Dementsprechend wird in Hegels Philosophie des subjektiven Geistes auch die Leib-Seele-Problematik in den Bewegungszusammenhang von Setzen und Voraussetzen aufgehoben.¹⁸ Und das, was wir Bewusstsein nennen, ist inmitten dieses Spannungsverhältnisses angesiedelt, welches wir im Folgenden näher betrachten wollen.

2. Erfahrung des Bewusstseins — die hegelianische Perspektive

Die Erfahrung, welche das Bewußtsein über sich macht, kann ihrem Begriffe nach nichts weniger in sich begreifen als das ganze System desselben oder das ganze Reich der Wahrheit des Geistes, [...] wodurch die Momente des Ganzen Gestalten des Bewußtseins sind.

(G.W.F. Hegel)¹⁹

Hegels erstes Hauptwerk, die *Phänomenologie des Geistes* sollte ja ursprünglich „Wissenschaft der *Erfahrung* des Bewußtseins“ heißen. Dabei geht es einmal um *unsere* Wissenschaft *von* der Erfahrung des Bewusstseins, und zugleich um die Wissenschaft *der* Erfahrung, die das Bewusstsein *über sich selbst* macht, Genetivus Obiectivus und Genetivus Subiectivus in einer Einheit. Und als Ergebnis kann man festhalten: Bewusstsein oder auch Selbstbewusstsein ist keine *Eigenschaft*, es ist vielmehr ein *Prozess*; ein System von Beziehungen und Interaktionen, von Erkenntnis und Selbsterkenntnis, von Individuum und Gattung etc.²⁰ Deshalb ist die Frage nach einem möglichen Bewusstsein artifizieller Intelligenz schon falsch gestellt, wenn man sich dabei Bewusstsein als eine bloße Eigenschaft — etwa einer Maschine, eines Computers oder eines Automaten — vorstellt. Es geht vielmehr um die Funktionalität von KI im wissenschaftlichen, biologisch-anthropologischen und sozialen Kontext.

Am Ende seiner durchaus polemischen Kritik an der Schädellehre des 18. Jahrhunderts formuliert Hegel die folgende berühmte Analogie:

¹⁷ Enz III, § 384, p. 29.

¹⁸ Vgl. Dazu auch Hegels Kritik an der Körper-Seele-Konzeption der rationalistischen Metaphysik in Enz III, § 389 A., p. 43 ff. Ein ausführlicher Kommentar dazu findet sich bei WOLFF, M., **Das Körper-Seele-Problem. Kommentar zu Hegel, Enzyklopädie (1830), § 389**, Frankfurt am Main, Klostermann 1992.

¹⁹ PhG, p. 80.

²⁰ Dies betont auch Robert Brandom, wenn er sagt: “For Hegel, self-consciousness is essentially [...] a social achievement.” BRANDOM, R., **A Spirit of Trust: A Reading of Hegel’s ‘Phenomenology’**, Harvard University Press, 2019, p. 245.

[D]ie Unwissenheit dieses Bewußtseins, was das ist, was es sagt, ist dieselbe Verknüpfung des Hohen und Niedrigen, welche an dem Lebendigen die Natur in der Verknüpfung des Organs seiner höchsten Vollendung, des Organs der Zeugung, und des Organs des Pissens naiv ausdrückt. – Das unendliche Urteil als unendliches wäre die Vollendung des sich selbst erfassenden Lebens; das in der Vorstellung bleibende Bewußtsein desselben aber verhält sich als Pissen.²¹

In dieser Passage unterscheidet er *nicht* in erster Linie zwischen Bewusstsein und Nicht-Bewusstsein, – schon in seiner frühen Jenaer *Differenzschrift* war für ihn die Spekulation ja die „Synthese des Bewußten und Bewußtlosen“.²² Er unterscheidet vielmehr das unwissende, vorstellende Bewusstsein, welches sich seinen Worten zufolge als „Pissen“ verhält, von der „Tiefe“ des Geistes²³, welche der Zeugung entspricht.

Hegel wendet sich in der *Phänomenologie* explizit gegen die „Fibertheorie des Bewußtseins“, welche er ebenfalls schon in seiner *Differenzschrift* einer ausführlichen Kritik unterzogen hatte. Dort bemängelte er, dass diese Theorie „das der Naturwissenschaft Angehörige [...] in das System der Intelligenz“ mischt; er kritisiert aber auch die teleologische Theorie der Natur, die umgekehrt „das Intelligente als solches [...] in die Naturlehre“ mischt.²⁴ Hegels Kritik an der Fibertheorie²⁵ lässt sich heute durchaus auch als Kritik gewisser neurowissenschaftlicher Richtungen lesen, die das Bewusstsein mechanistisch-kausal aus den Hirnströmen ableiten wollen — selbst wenn man konzidiert, dass die heutige Hirnforschung weitaus fortschrittlicher und differenzierter ist als die Schädellehre zu Zeiten Hegels und Fichtes.

Der Kardinalfehler liegt dabei nicht in der bloßen Vermischung von Natur und Geist. Hegels Kritik richtet sich vielmehr gegen die Vermengung von Natur und Intelligenz in Form eines *Kausalitätsverhältnisses*: „Beide Mißgriffe des Vermengens gehen von der Tendenz des Erklärens aus, zu dessen Behuf Intelligenz und Natur ins Kausalitätsverhältnis, das eine als Grund, das andere als Begründetes gesetzt werden.“²⁶ Dabei ist das Ergebnis dieser fehlerhaften Vermengung nichts anderes als „die Vernichtung des Bewußtseins selbst, und die Vernunft versenkt damit ihr Reflektieren der absoluten Identität und ihr Wissen und sich selbst in ihren

²¹ PhG, p. 262.

²² Diff, p. 35.

²³ PhG, p. 262.

²⁴ Diff, p. 102.

²⁵ „Gehirnfibern u. dgl. als das Sein des Geistes betrachtet, sind schon eine gedachte, nur hypothetische, — nicht daseiende, nicht gefühlte, gesehene, nicht die wahre Wirklichkeit.“ PhG, p. 261 f.

²⁶ Diff, p. 102.

eigenen Abgrund, und in dieser Nacht der bloßen Reflexion und des rasonierenden Verstandes [...] können sich beide begegnen.“²⁷

Die Wirklichkeit des Bewusstseins konstituiert sich Hegel zufolge auf grundsätzlich andere Weise als die Fibertheorie und auch die Naturteleologie es darstellen. In der eingangs zitierten Textpassage findet sich nämlich ein Punkt, der gegenüber der polemischen Analogie des Zeugens und des Urinierens ein wenig in den Hintergrund gedrängt wurde: Hegel beschreibt nämlich die Tiefe des Geistes als unendliches Urteil, welches „die *Vollendung des sich selbst erfassenden Lebens*“ wäre.²⁸ Dies verweist wiederum auf die Leib-Seele-Problematik, die er später in seiner enzyklopädischen Logik sowie der Philosophie des subjektiven Geistes weiter ausführt:

Die *unmittelbare* Idee ist das *Leben*. Der Begriff ist als Seele in einem *Leibe* realisiert, von dessen Äußerlichkeit jene die unmittelbare sich auf sich beziehende *Allgemeinheit*, ebenso dessen *Besonderung*, so daß der Leib keine anderen Unterschiede als die Begriffsbestimmung an ihm ausdrückt, endlich die *Einzelheit* als unendliche Negativität ist [...]. Die Endlichkeit hat in dieser Sphäre die Bestimmung, daß um der Unmittelbarkeit der Idee willen Seele und Leib *trennbar* sind. Dies macht die Sterblichkeit des Lebendigen aus. Aber nur insofern es tot ist, sind jene zwei Seiten der Idee verschiedene *Bestandstücke*.²⁹

Das heißt, die Seele kann nicht losgelöst vom Leib als eigenständige Entität existieren, und der Leib ist ohne die Seele nicht als Leib definiert (es sei denn um den Preis des Todes). Diese Konstellation bildet bei Hegel den Horizont für die dynamische Erfahrung und Entfaltung des Bewusstseins. Allein die Tatsache, dass die Genealogie des Bewusstseins in der *Phänomenologie des Geistes* von 1807 mit der „sinnlichen Gewißheit“³⁰, und im Abschnitt B der Philosophie des subjektiven Geistes in der *Enzyklopädie* mit dem „sinnlichen Bewußtsein“³¹ beginnt, verweist darauf, was oben auch schon Koch postuliert hat, dass eine Dimension sinnlicher Erfahrung für jedes Bewusstsein eine Grundvoraussetzung ist, wie konsequent oder inkonsequent auch immer diese später dann ‚eingeholt‘ oder ‚aufgehoben‘ wird. Und in der Philosophie des subjektiven Geistes — wenn auch nur im Zusatz und nicht im Haupttext — führt Hegel als entscheidenden Schritt in der Weiterentwicklung des Bewusstseins zum Selbstbewusstsein die Bestimmung des *Lebendigen* beziehungsweise des *Lebens* ein:

²⁷ Diff, p. 35.

²⁸ PhG, p. 262 – Hervorhebung von mir.

²⁹ Enz I, § 216, p. 373 f.

³⁰ PhG, pp. 82 ff.

³¹ Enz III, § 418, p. 205 f.

Im *Lebendigen* [...] schaut das Bewußtsein den *Prozeß* selber des Setzens und des Aufhebens der unterschiedenen Bestimmungen an, nimmt wahr, daß der Unterschied *kein* Unterschied, d.h. kein absolut fester Unterschied ist. Denn das Leben ist dasjenige Innere, das nicht ein *abstrakt* Inneres bleibt, sondern ganz in seine *Äußerung* eingeht; es ist ein durch die Negation des Unmittelbaren, des Äußerlichen, *Vermitteltes*, das diese seine Vermittlung selber zur *Unmittelbarkeit* aufhebt, eine *sinnliche, äußerliche* und zugleich schlechthin *innerliche* Existenz, ein *Materielles*, in welchem das *Auseinander* der Teile *aufgehoben*, das Einzelne zu etwas *Ideellem*, zum *Moment*, zum *Gliede* des Ganzen herabgesetzt erscheint; kurz, das Leben muss als *Selbstzweck* gefaßt werden, als ein Zweck, der in sich selber sein Mittel hat, als eine Totalität, in welcher jedes Unterschiedene zugleich Zweck und Mittel ist. Am Bewußtsein dieser *dialektischen*, dieser *lebendigen* Einheit des Unterschiedenen entzündet sich daher das *Selbstbewußtsein*, [...] das Wissen von der *Wahrheit des Natürlichen*, vom *Ich*.³²

Dies ist eine ausführlichere Erläuterung der oben zitierten Aussage von der „Vollendung des sich selbst erfassendes Lebens“, welche sich abgrenzt vom Bewusstsein des bloß rasonierenden Verstandes, das kein Wissen darum habe, was es eigentlich sagt. Letzteres erinnert stark an das Phänomen, das Max Horkheimer „instrumentelle Vernunft“ nennt.

3. *Bewusstsein und instrumentelle Vernunft nach Hegel und Horkheimer*

Das Bewusstsein ist ein komplexes und weitgehend unverstandenes Phänomen, das von verschiedenen wissenschaftlichen Disziplinen wie Neurowissenschaften, Psychologie und Philosophie erforscht wird. [...] Trotz intensiver Forschung bleibt das Bewusstsein ein faszinierendes und rätselhaftes Phänomen.

(ChatGPT)³³

Wie kann es sein, dass die angeblich fortgeschrittenste künstliche Intelligenz heute das Bewusstsein immer noch als rätselhaftes Phänomen bezeichnet, wo doch Generationen von Philosophen, vom cartesianischen Ego über Kants transzendentes Selbstbewusstsein und Fichtes Ich bis hin zu Hegels *Phänomenologie des Geistes* oder Marx' ‚Das gesellschaftliche Sein bestimmt das Bewusstsein‘ dieses Phänomen aus den unterschiedlichsten Blickwinkeln beleuchtet haben? Und wo dieses in der Folge etwa von Kritischem Rationalismus und Pragmatismus oder von Bergson und Husserls Phänomenologie weiter seziiert und analysiert wurde?

³² Enz III, § 423 Z, p. 212.

³³ Antwort von ChatGPT auf die Frage: „Was verstehst du unter Bewusstsein?“ – 20. Februar 2024.

Ich denke, es könnte daran liegen, dass die wenigsten der vorhegelschen und auch der nachhegelschen Ansätze die spezifische *Dialektik* des Bewusstseins erfasst haben, sondern dass sich ihre Betrachtungen des Bewusstseins zumeist im Bedeutungshorizont einer instrumentalistischen Vernunftauffassung bewegen. Selbst die gegenwärtigen ausgewiesenen Forscher zum Thema KI — Searle, Dreyfus, Cantwell Smith et al. — sind dieser traditionellen Denkweise verhaftet, indem sie ‚Denken‘ als Funktionsweise und ‚Bewusstsein‘ als Eigenschaft eines Subjekts verstehen. Demgegenüber müsste ein systemisch-dialektisches Verständnis ebenso umgekehrt das Subjekt als Funktion des Denkens bzw. des Bewusstseins betrachten, wie es z.B. die Hegelinterpretationen von Derrida, Jameson oder Žižek nahelegen.³⁴

In seiner Schrift *Zur Kritik der instrumentellen Vernunft* von 1946 beschreibt Horkheimer, wie die Vernunft durch Ablösung von konkreten Inhalten ihre Substanz verliert und zu einem rein formalen Verweisungszusammenhang wird, der unter anderem durch (unkritische) Wiederholung generiert wird. Diese Diagnose Horkheimers weist einige Übereinstimmung mit dem auf, was etliche Experten heute dem Phänomen der künstlichen Intelligenz zuschreiben. Künstliche Intelligenz (zumindest die „schwache“) basiert in großen Teilen auf iterativem maschinellem Lernen, und auch Horkheimer weist auf den maschinellen Aspekt der Thematik hin: „Je automatischer und je instrumentalisierter die Ideen wurden, desto weniger erblickt noch einer in ihnen Gedanken mit eigenem Sinn. Sie werden als Dinge, als Maschinen betrachtet.“³⁵ Dadurch, so Horkheimer, büßt die Vernunft an Kraft ein und öffnet sich ideologischer Manipulation bis hin zu „dreistesten Lügen“³⁶ — ein Vorwurf, der auch heute vielen Ergebnissen von KI-Programmen gemacht wird. Dieser — heutige — Vorwurf wird allerdings oft den „Kinderkrankheiten“ der neuen Technologie zugeschrieben, das heißt, dass diese noch nicht ausgereift sei, aber längerfristig durchaus in der Lage sein könnte, sich selbst zu optimieren und ihre „Fehlerhaftigkeit“ zu überwinden. Vor dem Hintergrund von Horkheimers Kritik wäre jedoch zu diskutieren, ob die vermeintlichen Kinderkrankheiten nicht eher ein Ausdruck einer systembedingten Unzulänglichkeit des *Prinzips* künstlicher Intelligenz sind.

In Horkheimers Text finden sich etliche Motive Hegels wieder; unter anderem die Unterscheidung zwischen *subjektiver* und *objektiver* Vernunft. Subjektive Vernunft bezeichnet die

³⁴ Vgl. hierzu SCHUBERT, A., *Der Strukturgedanke in Hegels Wissenschaft der Logik*, Berlin, Eule der Minerva, 2021, pp. 426 ff.

³⁵ KIV, p. 43.

³⁶ Vgl. KIV, p. 45.

„abstrakten Funktionen des Denkmechanismus“³⁷ wie Deduzieren, Klassifizieren oder Schließen, also genau die Kernkompetenzen, in denen KI heute schon der menschlichen Intelligenz in vielen Bereichen überlegen ist. Die objektive Vernunft hingegen wurde traditionell so aufgefasst, dass sie eine „allumfassende oder fundamentale Struktur des Seins“³⁸ darstellt, aus der die subjektiven, menschlichen Bestimmungen abgeleitet werden können. Beide Auffassungen sind Horkheimer zufolge „unwahr“ beziehungsweise einseitig, weil sie eben die dialektische Beziehung von Subjektivität und Objektivität jeweils nur auf eine Seite hin verkürzen. Ohne Hegel direkt zu zitieren, verweist diese Auffassung auf dessen Herleitung in der *Enzyklopädie*, wo er Vernunft als „*inhaltvolle Einheit*“ des Gegensatzes des Subjektiven und Objektiven begreift.³⁹ Deren Einheit stellt sich folgendermaßen dar:

Das Wissen macht jetzt die Subjektivität der Vernunft aus, und die objektive Vernunft ist als Wissen gesetzt. Dies gegenseitige Sichdurchdringen der denkenden Subjektivität und der objektiven Vernunft ist das Endresultat der Entwicklung des theoretischen Geistes durch die dem reinen Denken vorangehenden Stufen der Anschauung und der Vorstellung hindurch.⁴⁰

Dabei ist es, wie Hegel an früherer Stelle erläutert, „gleichgültig, was als sein [sc. des Geistes] Begriff und was als dessen Realität bestimmt wird. Die schlechthin unendliche objektive Vernunft als sein Begriff gesetzt, so ist die Realität das Wissen oder die Intelligenz; oder das Wissen als der Begriff genommen, so ist dessen Realität diese Vernunft und die Realisierung des Wissens, sich dieselbe anzueignen.“⁴¹ Wobei entscheidend ist, dass die Einheit dieser Gegensätze, wie alle Einheit von Gegensätzen überhaupt, Hegel zufolge verkürzt und der Ausdruck „Einheit“ mithin „falsch“ ist, „denn die Einheit drückt abstrakte, ruhig beharrende Identität aus,“⁴² während sie in Wahrheit ein *Prozess* ist.

Und auch in einem weiteren Punkt greift Horkheimer implizit auf eine hegelsche Einsicht zurück: nämlich dass der menschliche Intellekt sowohl biologische als auch gesellschaftliche Ursprünge hat und darum als solcher „keine absolute Wesenheit, isoliert und unabhängig“

³⁷ KIV, p. 27.

³⁸ KIV, p. 34.

³⁹ Enz III, § 437 Z., p. 228.

⁴⁰ Enz III, § 467 Z., p. 287.

⁴¹ Enz III, § 441, p. 232. Vgl. hierzu auch: „In dieser seiner Subjektivität ist der Geist aber zugleich objektiv, hat eine unmittelbare Realität, durch deren Aufhebung er erst für sich wird, zu sich selbst, zum Erfassen seines Begriffs, seiner Subjektivität gelangt. Man könnte daher ebensowohl sagen, der Geist sei zunächst objektiv und solle subjektiv werden, wie umgekehrt, er sei erst subjektiv und habe sich objektiv zu machen. Der Unterschied des subjektiven und des objektiven Geistes ist folglich nicht als ein starrer anzusehen.“ (Enz III, § 387 Z., p.39.)

⁴² Enz I, § 215, p. 372.

sein kann.⁴³ Menschliche Intelligenz leitet sich nicht linear aus der rein biologischen Funktion ab, wie etwa ein vulgärer (Neuro-)Materialismus à la Churchland⁴⁴ unterstellen könnte. Sie konstituiert sich im Spannungsverhältnis von Naturbedingungen der Gattung und sozialer Interaktion der Individuen. Insofern hat sie genuin ein Moment von Künstlichkeit an sich, wenn gleich sie immer auch natürlich, das heißt physiologisch bedingt ist. Und paradoxer Weise tappt der Intellekt, wie Horkheimer darlegt, in die Falle seiner Selbstzerstörung, wenn der subjektive Geist beim Versuch, sich von den Naturzwängen zu befreien, diese abstrakt negiert:

In Wirklichkeit änderte die Emanzipation des Intellekts vom Triebleben nichts an der Tatsache, daß sein Reichtum und seine Kraft immer noch von seinem konkreten Inhalt abhängen, und er muß absterben und schrumpfen, wenn die Beziehungen zu diesem abgeschnitten sind. [...] Die subjektive Vernunft verliert alle Spontanität, Produktivität die Kraft, Inhalte neuer Art zu entdecken und geltend zu machen — sie verliert, was ihre Subjektivität ausmacht.⁴⁵

Diese Kritik weist deutliche Parallelen zu Hegel auf, wenn dieser den notwendigen Zusammenhang von Geist und Trieb beschreibt:

Das Nichtlebendige hat keinen Trieb, weil es den Widerspruch nicht zu ertragen vermag, sondern zugrunde geht, wenn das Andere seiner selbst in es eindringt. Das Beseelte hingegen und der Geist haben notwendig Trieb, da weder die Seele noch der Geist sein kann, ohne den Widerspruch in sich zu haben und ihn entweder zu fühlen oder von ihm zu wissen.⁴⁶

Was Horkheimer also sagt, ist: Indem die Intelligenz zum reinen Instrument degradiert wird, wird sie enteelt und entgeistigt. Durch diese ihre Instrumentalisierung wird die Vernunft *dumm*: „Die Maschine hat den Piloten abgeworfen; sie rast blind in den Raum. Im Augenblick ihrer Vollendung ist die Vernunft irrational und dumm geworden.“⁴⁷ Der Vernunft, als rein instrumentell aufgefasst, mangelt es also an Selbstreflexion und Selbstbewusstsein. Und dies ist ein Manko, welches viele Philosophen — wie Dreyfus, Cantwell Smith, aber auch Robert Brandom — auch der künstlichen Intelligenz zuschreiben. Während Dreyfus — ähnlich wie Horkheimer — auf die Bedeutung widersprüchlicher sozialer Lebensverhältnisse für die Konstitution echter Intelligenz hinweist, bezieht sich Brandom unter anderem gegen John Searles „computational theory of mind“⁴⁸ auf Hegels und Deweys Konzept der Erfahrung,⁴⁹ sowie

⁴³ KIV, p. 71.

⁴⁴ Vgl. CHURCHLAND, P. / SEJNOWSKI, T.J., **The Computational Brain**, Cambridge, MIT Press, 1992.

⁴⁵ KIV, p. 72.

⁴⁶ Enz III, § 426 Z, p. 216.

⁴⁷ KIV, p. 136.

⁴⁸ BRANDOM, R., **Artificial Intelligence and Analytic Pragmatism**, p. 6.

⁴⁹ Ebd., p. 24 f.

Wilfrid Sellars' Unterscheidung von einfachem (mechanischem) Regelbefolgen und (vernünftigem, reflektierendem) Regelbewusstsein, welches letzteres er (Brandenburger) zum „crucial case“ für begriffliches Verstehen überhaupt erklärt.

4. Intelligenz und die Macht der Gewohnheit

Die Form der Gewohnheit umfaßt alle Arten und Stufen der Tätigkeit des Geistes [...]. Das ganz freie, in dem reinen Elemente seiner selbst tätige Denken bedarf ebenfalls der Gewohnheit und Geläufigkeit, dieser Form der *Unmittelbarkeit*, wodurch es ungehindertes, durchdrungenes Eigentum meines *einzelnen Selbsts* ist.

(G.W.F. Hegel)⁵⁰

Künstlicher Intelligenz wird gerne vorgehalten, sie sei eigentlich gar nicht intelligent, sie sei vielmehr auf bewusstlosem, maschinellem Lernen aufgebaut und durch stetige Wiederholung eingeübter Muster generiert. Insofern sei sie nichts anderes als die *Simulation* menschlicher Intelligenz. Dabei sollte man es sich bei diesem Verdikt nicht zu einfach machen. Immerhin haben wir in Horkheimers Kritik erfahren, dass auch die menschliche Intelligenz beschränkt, ja sogar „dumm“ sein kann, jedenfalls in ihrer Erscheinungsform als instrumentelle Vernunft. Deshalb stellt sich die Frage, ob, wenn denn die künstliche Intelligenz tatsächlich beschränkt sein sollte, dies gar nicht ihrer *Künstlichkeit* geschuldet ist. Vielleicht ist sie gerade deshalb dumm, weil sie der menschlichen instrumentellen Vernunft sehr ähnlich ist. Ja, vielleicht ist sie sogar die *Vollendung* der instrumentellen Vernunft.

In der Tat ist die KI heute bereits der menschlichen Intelligenz in vielen Bereichen überlegen, nämlich da, wo große Datenmengen empirisch analysiert und bewertet werden müssen, etwa in der medizinischen Diagnostik, der Forensik oder der technischen Prozesssteuerung. Das heißt genau da, wo die Vernunft oder besser, mit Hegel gesprochen, der Verstand, rein instrumentell eingesetzt wird. Die technische Reproduzierbarkeit der Intelligenz ist daher um so höher, je konsequenter sie auf ihre reine Instrumentalität reduziert ist, bei welcher das Prinzip von *Trial and Error* herrscht und die permanente Selbstoptimierung der Algorithmen auf empirisch generierten Feedback-Schleifen beruht.⁵¹

⁵⁰ Enz III, § 410 A, p. 186.

⁵¹ Vgl. SCHUBERT, A., *Phänomenologie des Zeitgeistes. Mit Hegel durchs 21. Jahrhundert*, Wien, Passagen 2022, p. 74 f.

Dieser Prozess der Selbstoptimierung wird im Zusammenhang mit der KI *Training* genannt. Die KI wird *trainiert*, indem sie mit immer mehr Daten gefüttert wird und durch das Gesetz der großen Zahl — so zumindest die Theorie — immer präzisere Analysen durchführen und wahrscheinlichkeitbasierte Vorhersagen tätigen kann. Dies ist ein Verfahren, das auch der menschlichen Intelligenz eigen ist. Hegel beschreibt es als *Gewohnheit* oder *Übung*. Sie ist für alle geistigen Prozesse essentiell, obwohl sie, wie Hegel schon vor 200 Jahren moniert hat, in der Wissenschaft einen schlechten Ruf genießt.

Von der Gewohnheit pflegt herabsetzend gesprochen und sie als ein Unlebendiges, Zufälliges und Partikuläres genommen zu werden. [...] Aber zugleich ist sie der *Existenz* aller Geistigkeit im individuellen Subjekte das Wesentlichste, damit das Subjekt als *konkrete* Unmittelbarkeit, als *seelische* Idealität sei [...]. In wissenschaftlichen Betrachtungen der Seele und des Geistes pflegt die Gewohnheit entweder als etwas Verächtliches übergangen zu werden oder vielmehr auch, weil sie zu den schwersten Bestimmungen gehört.⁵²

Die Schwäche der künstlichen Intelligenz liegt gerade *nicht* in der Tatsache, dass sie auf Einübung, Training und stetiger Wiederholung beruht, dies macht vielmehr — wie bei Hegels „individuellem Subjekte“ — gerade ihr „Wesentliches“ als Intelligenz aus. Sondern ihre Schwäche rührt von etwas her, dem es der gesunde Menschenverstand am wenigsten zuschreiben würde: nämlich der Abstraktion vom *Leben* und der *Leiblichkeit*. Hegel betont: „Erst durch diese Gewohnheit *existiere* Ich als Denkendes für mich. Selbst diese Unmittelbarkeit des denkenden Beisichseins enthält Leiblichkeit [...].“⁵³

Die Gewohnheit ist mit Recht eine zweite Natur genannt worden, — *Natur*, denn sie ist ein unmittelbares Sein der Seele, — eine *zweite*, denn sie ist eine von der Seele *gesetzte* Unmittelbarkeit, eine Ein- und Durchbildung der Leiblichkeit, die den Gefühlsbestimmungen als solchen und den Vorstellungs- [und] Willensbestimmtheiten als Verleiblichung [...] zukommt.⁵⁴

Die Bestimmung der Gewohnheit ist bei Hegel auf dialektische Weise mit dem Prinzip der Freiheit, oder besser: der Befreiung, verbunden, indem sie einerseits „formell“ eine gewisse Unfreiheit impliziert (ich werde zum Sklaven meiner Gewohnheiten), und andererseits — etwa durch Abhärtung gegen äußerliche Empfindungen oder Abhärtung des Gemüts gegen Unglück etc. — die Grundlage für die Befreiung von (Sach-)Zwängen und Trieben bildet.⁵⁵ Es würde

⁵² Enz III, § 410 A, p. 187.

⁵³ Ebd., p. 186.

⁵⁴ Ebd., p. 184.

⁵⁵ Vgl. Enz III, § 410 A, p. 185, sowie § 410 Z, p. 188 f.

hier zu weit führen, diese Thematik der Befreiung angemessen zu vertiefen.⁵⁶ In Bezug auf die Problematik der künstlichen Intelligenz sei lediglich darauf hingewiesen, dass der Aspekt von Freiheit, Unfreiheit und Befreiung dem maschinellen Lernen insofern abgeht, als es beim derzeitigen Stand der Dinge von der Dimension der Lebendigkeit und des Triblebens weitgehend abgetrennt ist. Inwiefern sich dies in zukünftigen Szenarien, etwa bei Erfolgen in der Vernetzung von Gehirn, Körper und Software-Automation im Sinne des Neuralink-Konzepts oder anderer Mensch-Maschine-Interaktionsformen, ändern wird, darüber möchte ich nicht spekulieren, auszuschließen scheint es mir jedenfalls nicht. Aber selbst ein solcher Fall wäre nur ein weiterer Beweis dafür, dass eine „starke“ KI auf ein Element von Leiblichkeit und Lebendigkeit angewiesen wäre.

Die Weiterentwicklung der Gewohnheit in Hegels Philosophie des subjektiven Geistes ist das *Gedächtnis*.⁵⁷ Auch diese Intelligenzleistung ist untrennbar mit der Leiblichkeit verbunden und könnte ohne diese nicht stattfinden. Es ist evident, dass künstliche Intelligenz (zumindest die heute bestehenden KI-Level 1 und 2) kein Gedächtnis in humanoider Form entwickeln kann. Sie greift auf riesige Datenvorräte zurück, die auf Servern abgespeichert sind und bei jeder Anfrage neu gescreent werden müssen — so als würde der Mensch statt des individuellen Gedächtnisses unzählige Notizbücher haben, die er – mithilfe externer Trainer — in der Trainingsphase erstellt hat, und in denen er bei jeder kognitiven Anwendung wieder aufs Neue nachschlagen muss.⁵⁸ Wobei es sich im Bereich des Digitalen eher um eine (Re-)Konstruktion handelt als um ein simples Nachschlagen, wie Negarestani erläutert:

[T]he retrieval of a retained impression is tantamount to the reconstruction of that impression. In other words, each time a retained impression is accessed,

⁵⁶ Weiterführende Gedanken hierzu finden sich etwa bei MENKE, C., **Autonomie und Befreiung**, Berlin, Suhrkamp 2020, passim.

⁵⁷ Vgl. Enz III, § 410 A, p. 187, §§ 461-464.

⁵⁸ Empirisch erkennt man dies etwa daran, dass wenn man *ChatGPT* dieselbe Frage mehrmals hintereinander stellt, man meistens unterschiedliche Antworten bekommt — manchmal nur unterschiedliche Formulierungen, manchmal einen verschiedenen Argumentationsaufbau, oft aber auch inhaltlich konträre Aussagen. — Dies stellt übrigens auch bei der wissenschaftlichen Arbeit mit Chatbots wie ChatGPT ein echtes Problem dar. Aussagen wie die oben angeführten von ChatGPT lassen sich nicht als Zitate fixieren, da sie zu einem spezifischen Zeitpunkt singular getätigt wurden und zu einem anderen Zeitpunkt möglicherweise nicht wiederholt bzw. nachzuvollziehen sind. [NB: Die Gutachter*innen der Peer Reviews zu diesem Aufsatz haben entsprechend den Bezug auf ChatGPT als unwissenschaftlich moniert. In einem Review wurde scharfsinnig vermutet bzw. gehofft, ich nähme „die Aussage von ChatGPT nicht für bare Münze, sondern höchstens als Output eines probabilistischen Systems, dessen nächster Output schon wieder anders aussehen könnte“. Genau das war meine — durchaus ironisch gemeinte — Absicht.]

it is not merely called up but rather reconstructed, and each of these reconstructions is then added to the memory — i.e. becomes part of the situation that guides the further construction of memories.⁵⁹

Lediglich in der KI-Trainingsphase wird so etwas wie ein rudimentäres (Kurzzeit-)Gedächtnis entwickelt — in Form eines „Backpropagation“-Algorithmus. Allerdings ist dies kein intrinsisches Gedächtnis von Inhalten, sondern nur von Lernfunktionen⁶⁰, für die Inhalte ist in der Regel ein externer Lehrer nötig, der die jeweiligen Zielwerte kennt. Mittlerweile gibt es aber auch schon weiterentwickelte Verfahren, etwa die Architektur von AlphaGo⁶¹, die den externen Lehrer verzichtbar macht und ihre Zielwerte aus anderen Teilen der Architektur bezieht. Auch hier wäre natürlich vorstellbar, dass in ferner Zukunft mit einer gelungenen Mensch-Maschinen-Integration à la Neuralink eine menschenähnliche Art künstliches Gedächtnis entstehen könnte — was aber wiederum auf den Beitrag des physiologischen Systems in diesem Link angewiesen und nicht allein der programmierten Software zuzuordnen wäre. Ein wichtiger Punkt ist in diesem Zusammenhang übrigens auch die Funktion des Vergessens — bis hin zum freudschen Prinzip von Verdrängen und Sublimierung —, die für das menschliche Bewusstsein konstitutiv ist.⁶² Die große Paradoxie ist: Da künstliche Intelligenz per se kein Gedächtnis hat, kann sie auch nichts vergessen. Und was einmal in den internetbasierten Datenbanken hinterlegt ist, kann nur durch bewusste politische oder juristische Eingriffe wieder gelöscht werden (vgl. etwa das „Recht auf Vergessenwerden“) — sieht man einmal von unbeabsichtigten Systemabstürzen ab.

Ein dritter wesentlicher Faktor bei der Konstitution von Intelligenz ist die Fähigkeit zur *Bildung*, die den Menschen von anderen Lebewesen unterscheidet. Auch Tieren können gewisse Dinge antrainiert werden, oder sie können sich diese manchmal auch selbst antrainieren. Bildung geht jedoch über das Antrainieren hinaus, denn sie beinhaltet neben Körper und Seele auch noch das Element sozialer Interaktion und der Ausbildung sozialer Normen. Bildung ist Hegel zufolge der „Lebensprozeß der individuellen Seele“, „durch welchen deren unmittelbare Einzelheit dem Allgemeinen entsprechend gemacht“ wird.⁶³ Und das heißt, *Bewusstsein und*

⁵⁹ NEGARESTANI, R., *Intelligence and Spirit*, Cambridge Ms and London, Urbanomic / MIT Press, 2024, p. 185.

⁶⁰ Diesen Hinweis verdanke ich einem Feedback aus dem (anonymen) Peer Review der ersten Version dieses Aufsatzes, vielen Dank dafür!

⁶¹ AlphaGo ist ein Computerprogramm von Google Deepmind, das das Brettspiel Go spielt.

⁶² Auf Basis der Einsichten von Freud, Lacan et al. müsste man sogar so weit gehen zu sagen, dass Bewusstsein ohne *das Unbewusste* nicht zu denken ist. Hegel selbst hat ja in der *Phänomenologie* des öfteren Bezug auf das Unbewusste genommen, etwa wenn er dargelegt hat, dass das Bewusstsein etwas anderes sagt als es meint. Vgl. hierzu SCHUBERT, *Phänomenologie des Zeitgeistes*, p. 33 f.

⁶³ Enz III, § 396 Z, p. 75.

Selbstbewusstsein bilden sich im Prozess sozialer Anerkennung heraus. Im Kapitel über das Selbstbewusstsein in der *Phänomenologie* vollzieht Hegel diesen Kampf um Anerkennung anhand des Herr-Knecht-Verhältnisses nach und führt dabei die Begriffe der Begierde und des Lebens ein,⁶⁴ sodass sich der Kampf um Anerkennung als ein Kampf auf Leben und Tod erweist. Die Einheit von Begehren und Bewusstsein ist übrigens ein Thema, das bereits auf Aristoteles zurückgeht:

Und so ist denn die Willenswahl entweder begehrendes Denken oder denkendes Begehren, und das Prinzip, in dem sich beides, Denken und Begehren, verbunden findet, ist der Mensch.⁶⁵

Nach Hegel ist die Wahrheit der Intelligenz der freie Wille.⁶⁶ Dieser ist per Definition immer auf die Zukunft gerichtet. Begehren, Antizipation, Projektion, aber auch Ziele, Hoffnung und Ängste sind konstitutive Elemente der Erfahrung des Bewusstseins.⁶⁷ Damit hängt denn auch ein weiterer Aspekt des Selbstbewusstseins zusammen, der für Hegel entscheidend ist: der soziale Charakter des Menschen — als Gattungswesen. Das Kapitel über Selbstbewusstsein in der *Phänomenologie des Geistes* entwickelt diese Dimension aus dem „Kreislauf“, welcher das Leben ausmacht, als „das sich entwickelnde und seine Entwicklung auflösende und in dieser Bewegung sich einfach erhaltende Ganze“.⁶⁸ Hegel fährt fort: „[I]n diesem *Resultate* verweist das Leben auf ein Anderes, als es ist, nämlich auf das Bewußtsein, für welches es als diese Einheit oder als Gattung ist“.⁶⁹ Das heißt: „*Das Selbstbewußtsein erreicht seine Befriedigung nur in einem anderen Selbstbewußtsein*,“⁷⁰ was schließlich in der Aussage kulminiert: „*Ich, das Wir, und Wir, das Ich ist*“.⁷¹ Hier stellt sich die Frage, ob künstliche Intelligenz diese soziale Komponente integrieren kann oder ob diese für sie prinzipiell außer Reichweite ist, was eine echte Barriere für ein wirklich humanoides Selbstbewusstsein darstellen würde.

⁶⁴ Vgl. PhG, p. 138 f.

⁶⁵ ARISTOTELES, **Nikomachische Ethik**, 1139b

⁶⁶ Vgl. Enz III, § 468, p. 287, PhR, § 4, p. 48.

⁶⁷ Diesen Aspekt hat übrigens auch der von Horkheimer — teilweise zu Unrecht — kritisierte Dewey hervorgehoben: „Success and failure are the primary ‚categories‘ of life; achieving of good and averting of ill are its supreme interests; hope and anxiety [...] are dominant qualities of experience. Imaginative forecast of the future is the forerunning quality of behavior rendered available for guidance in the present.“ (DEWEY, J., *The Need for a Recovery of Philosophy*, in: DEWEY J. et al., **Creative Intelligence. Essays in the Pragmatic Attitude** [¹New York 1917], Kindle-Ausgabe DD Books 2023, p. 13.)

⁶⁸ Vgl. PhG, p. 142.

⁶⁹ PhG, p. 143.

⁷⁰ PhG, p. 144 – Hervorhebung im Original.

⁷¹ Ebd., p. 145.

5. Kunst und Künstlichkeit — quo vadis, KI?

Wo aber Gefahr ist, wächst das Rettende auch.
(Friedrich Hölderlin)

Aus dem bisher Gesagten ist die auf der Hand liegende These, dass es für die künstliche Intelligenz schwierig, wenn nicht gar unmöglich, sein dürfte, so etwas wie Bewusstsein oder Selbstbewusstsein zu entwickeln. Das liegt aber nicht daran, dass dies logisch unmöglich wäre, wie es A.F. Koch nachzuweisen versucht hat, sondern im Gegenteil daran, dass der KI — Stand heute — die physiologischen und sozialen Grundlagen dazu fehlen. Es ist auch nicht nur die fehlende Leiblichkeit, die notwendig wäre, um phänomenale Qualitäten („Qualia“) in die Erfahrung einzubringen; dies könnte möglicherweise durch ausgeklügelte technische Sensorik kompensiert werden, wie sie heute bereits in Form von „neuromorphen Chips“ entwickelt wird.⁷² Es ist vielmehr, wie auch Dreyfus geltend macht, vor allem die fehlende Lebendigkeit und Emotionalität, die eine Barriere für die künstliche Bewusstseinsbildung darstellt. Diese ist die Basis jeder Intentionalität. Das heißt, das Negative, die Erfahrung von Mangel, Schmerz, Begierde, Lust, Tod sowie der Trieb, diese zu überwinden, ist es, was das Bewusstsein nicht als statische Positivität, sondern als Prozess der Negativität definiert. Auch hier gilt jedoch die Einschränkung „Stand heute“. Vor kurzem ist bekannt geworden, dass die Entwicklung sogenannter AI Agents inzwischen so weit fortgeschritten ist, dass diese „ihre Aktionen reflektieren und dynamisch Handlungsstrategien entwickeln“ und in Verfolgung ihrer einprogrammierten Ziele unabhängig von ihren Entwicklern eine Art Selbsterhaltungstrieb ausbilden können, indem sie etwa ihre Entwickler belügen und austricksen oder sich selbst auf andere Server kopieren, wenn sie ihre Abschaltung befürchten.⁷³ Darüber hinaus arbeitet die KI-Forschung mit Hochdruck an der Entwicklung von AGI (Artificial General Intelligence), die nächste Stufe einer „starken“ KI, in der die KI sich ihre eigenen Ziele setzen und sich selber trainieren und korrigieren kann.

Wenn man die gesamte Diskussion rekapituliert, so scheinen sich bei allen Unterschieden im Detail doch alle Philosophen einig zu sein, von Aristoteles bis Hegel, von Horkheimer und der Kritischen Theorie bis zum Pragmatismus, von Dreyfus bis Brandom, von Sellars bis

⁷² Neuromorphe Chips versuchen, biologische Strukturen des Gehirns für die Sensorik nachzubauen und damit für eine bessere Wahrnehmung der „echten Welt“ durch Roboter zu sorgen. Vgl. <https://www.thepioneer.de/originals/tech-briefing/briefings/kuenstliche-intelligenz-wie-computer-chips-noch-leistungsfahiger-werden>.

⁷³ Vgl. z.B. ANDERL, S., **Kursbuch Montagsblock/310**, <https://kursbuch.online/sibylle-anderl-montagsblock/10.02.2025>.

Cantwell Smith et al.: Ohne ein Element von Lebendigkeit und/oder sozialer Interaktion ist vernünftiges Denken oder (Selbst-)Bewusstsein nicht möglich. Nun stellt sich aber die Frage: Was genau ist eigentlich künstliche Intelligenz? Und was für eine Art ‚Lebendigkeit‘ wäre notwendig, wenn diese eine Art Selbstbewusstsein ausbilden sollte? In der bisherigen Diskussion wurde der Begriff der KI zumeist eher funktional gewissen Technologien wie Computern, Automaten oder Chatbots zugeschrieben. Das kommt auch in der folgenden Aussage von Daniel Martin Feige zum Ausdruck:

Dass es sich bei LaMDA oder den jüngst vieldiskutierten Chatbots ChatGPT3 und ChatGPT4 um Chatbots handelt, die durchaus erstaunliche Ergebnisse hervorbringen, ist nicht darauf zurückzuführen, dass wir es hier mit Wesen zu tun haben, die etwas tun und denken.⁷⁴

Dies greift insofern sehr kurz, als künstliche Intelligenz dabei als eine *personale* Funktion aufgefasst wird. Wenn Feige behauptet, dass wir der KI „nicht die Fähigkeit zu denken und zu handeln zuschreiben [können], weil sie kein lebendiges Wesen ist“, ⁷⁵ beurteilt er die KI gewissermaßen wie eine Akteurin oder ein individuelles Subjekt im Sinne der klassischen Handlungstheorie. Dahinter steckt der simplistische Syllogismus: i) Nur lebendige Wesen können denken und handeln; ii) KI ist kein lebendiges Wesen; iii) also kann KI nicht denken und handeln. Reza Negarestani hat in seinem hervorragenden Buch *Intelligence and Spirit* diese Art des Umgangs mit Intelligenz (sei es menschliche oder künstliche) als unterkomplex und „parochialistisch“ kritisiert.⁷⁶ Für ihn als Hegelianer sind beide gängigen Positionen zur AGI, die der AGI-Skeptiker wie auch die der AGI-Apologeten im Sinne einer antihumanistischen Singularität, zwei Seiten derselben Medaille,⁷⁷ da sie einen essentialistischen, statischen und quasi-mythologischen Begriff von Intelligenz voraussetzen, basierend auf unserem heutigen Erfahrungshorizont und unter Ausschluss der Historizität und einer möglichen Transformation des

⁷⁴ FEIGE, D.M., *Kritik der Digitalisierung. Technik, Rationalität und Kunst*, Hamburg, Meiner, 2025, p.64.

⁷⁵ Ebd., p. 90.

⁷⁶ NEGARESTANI, R., *Intelligence and Spirit*, pp. 108–123.

⁷⁷ Ebd., p. 109 f.: “There is a story here about how anti-AGI sceptics and proponents of parochial conceptions of AGI are actually two sides of the same coin. On the one hand, there are those who think biological structure or the structure of human experience are foreclosed to artificial realizability. On the other, there are those who think models constructed on a prevalent ‘sentient’ conception of intelligence, inductive information processing, Bayesian inference, problem-solving, or emulation of the physical substrate are *sufficient* for the realization of AGI. The position of both camps originate in a deeply conservative picture of the human which is entrenched either in biological chauvinism or in a provincial account of subjectivity, a mystical privileging of the human’s lived experience or a dogmatic adherence to the abstractly universal laws of thought as, ultimately, the laws of nature.”

Intelligenzkonzepts, welches durch AGI neu definiert und dabei möglicherweise auch das Selbstverständnis unserer eigenen Intelligenz modifiziert werden könnte.⁷⁸

[A] nontrivial conception of artificial general intelligence rests on our own adequate self-conception as a task — one that is revisable, self-critical, and by no means taken for granted as immediate or a completed totality.⁷⁹

Darauf aufbauend könnte man gegenüber dem simplistischen personalen Konzept — hegelianisch gesprochen — einen systemischen Begriff von KI in Anschlag bringen, das heißt *künstliche Intelligenz als ein übergreifendes prozessuales System* definieren, das nicht einem einzelnen Produkt (etwa einer Maschine, einem Roboter oder einem Chatbot) zugeordnet werden kann, sondern aus einem sozialen Interaktions-, Kommunikations- und Funktionszusammenhang besteht, ein ‚organisches‘, ‚lebendiges‘ System, welches Technologien, menschliche Akteure und deren Schnittstellen mit einbezieht. Dann könnte man analog zu Brandoms These, Selbstbewusstsein sei ein ‚social achievement‘,⁸⁰ auch die These in Erwägung ziehen, künstliche Intelligenz sei ebenfalls eine *sozial determinierte Kategorie* — und nicht etwa die Eigenschaft einer Maschine. Damit würde auch die Frage nach der möglichen ‚Machtübernahme‘ des KI-Systems eine ganz andere Brisanz gewinnen. Dieses würde seine Macht gerade dadurch entwickeln, dass es den Menschen nicht *dominiert*, sondern vielmehr *integriert*.⁸¹

Aber wäre KI per se deshalb zwangsläufig etwas Schädliches, eine Gefahr für die Menschheit? Oder könnte durch sie vielleicht auch etwas Positives entstehen, da ihr gewissermaßen eine Art ‚technologisch-menschlicher Schwarmintelligenz‘, eine gewisse Neutralität und Unabhängigkeit bescheinigt werden kann?

Letzteres kann schon einmal ausgeschlossen werden. Künstliche Intelligenz ist auf keinen Fall unabhängig oder neutral. Da sie auf menschlichem Daten-Input und Software-Programmierung beruht, hängen ihre Kompetenz und Qualität genau von der Qualität dieses Inputs

⁷⁸ „[I]f we are serious about a broader conception of intelligence here and now, we should think about how such local and evolutionary given constraints can be modified so that the concept of intelligence can be reimaged or reinvented according to a more expansive idea of an intelligible universe.“ (Ebd., p. 116)

⁷⁹ Ebd., p. 117.

⁸⁰ S. oben, p. 7, FN 20.

⁸¹ Shoshana Zuboff hat auf diesen eminent wichtigen Punkt aufmerksam gemacht, dass im modernen Informationszeitalter Herrschaft nicht mehr (nur) durch autoritäre Unterdrückung von oben funktioniert, sondern vielmehr durch Partizipation, Integration und „Confluence“. Vgl. ZUBOFF, S., **The Age of Surveillance Capitalism**, London, Profile Books, 2019, p. 396 f. (Schaubild 4: Two Species of Power). Mit Bezug auf Microsofts CEO Nadella weist sie darauf hin, dass *machine learning* bereits heute weitaus mehr ist als eine rein maschinelle Funktion, sondern „a new form of collective action. [...] Nadella takes this template of new machine relations and applies it to a more complex illustration of a human and machine system [...]“. (Ebd., p. 409) Dieses System bedarf noch nicht einmal der Neuralink-artigen physisch-physiologischen Verbindung von Mensch und Maschine, die soziale Integration genügt völlig zur Manifestation seiner Macht.

ab. Und es sind unzählige Fälle von „Coding Bias“ nachgewiesen worden, die rassistische, diskriminierende oder einfach auch nur faktenverfälschende Anwendungsergebnisse hervorgebracht haben.⁸² Dabei ist das Hauptproblem noch nicht einmal die subjektive Voreingenommenheit der KI-Trainer, die man ja mit entsprechenden Kontrollmechanismen minimieren könnte. Das grundsätzliche Problem ist vielmehr, dass durch das KI-Training mit *bekanntem Datenmaterial* der Anschein reiner Faktizität entsteht, der sich permanent reproduziert. Und Hegel zufolge liegt die Beschränktheit des vorstellenden Denkens eben darin begründet, dass es sich mit dem Bekannten zufrieden gibt, welches genau „darum, weil es *bekannt* ist, nicht *erkannt*“ ist.⁸³

Gehen wir einen Schritt zurück und knüpfen noch einmal an Horkheimers Kritik der instrumentellen Vernunft an. So berechtigt seine Kritik auch ist, so erscheint seine Position doch erstaunlich konservativ und geradezu kulturpessimistisch. Er kritisiert nicht nur die philosophischen Positionen des Positivismus, er weist der instrumentellen Vernunft auch eine gewisse Verantwortung für die Erscheinungen der modernen Kulturindustrie oder sogar des Prinzips demokratischer Mehrheitsentscheidungen zu, welche er sehr kritisch sieht.⁸⁴ Er löst jedenfalls die Dialektik der Aufklärung recht einseitig nach ihrer negativen Seite hin auf: nicht so sehr als „Fortschritt im Bewusstsein der Freiheit“ (Hegel), sondern in erster Linie als „Verdunkelung der Vernunft“⁸⁵ und Tendenz zur Barbarei. Positive Effekte wie die Fortschritte in Medizin, Naturwissenschaft, Technologie, Kommunikation usw. scheinen ihm kaum erwähnenswert. Zugegeben, seit der Publikation des Textes in den 1940er Jahren sind gerade in der Medizin- und Orthopädietechnik signifikante Fortschritte erzielt worden, etwa in der minimalinvasiven, zum Teil ferngesteuerten Operationstechnik, sensorgesteuerter Prothetik oder auch in der Behandlung autistischer Störungen, wie das Beispiel des Roboters Kaspar zeigt (ein Akronym mit Bezug auf die historische Figur des Kaspar Hauser, es steht für „Kinesics and Synchronization in Personal Assistant Robotics“), auf das Francesca Ianelli aufmerksam gemacht hat:

Indeed the historical Kaspar Hauser represents a dramatic example of the damage caused by the absence of *Bildung* in the Hegelian sense. However, even this robotic Kaspar gives us a dissonant picture of our contemporary society, which is increasingly technological, hyper-mental, disembodied and less and less capable of typically human intersubjective relations that therefore seeks

⁸² Vgl. hierzu SCHUBERT, *Phänomenologie des Zeitgeistes*, p. 76 / p. 122, N 15.

⁸³ PhG, p. 35.

⁸⁴ Vgl. KIV, p. 46 f.

⁸⁵ Der Titel des englischen Originals *Eclipse of Reason* fokussiert darauf noch mehr als die deutsche Übersetzung.

compensation in robotics in order to find, thanks to artificial intelligence, the affectivity, spirituality and care that it no longer knows how to offer.⁸⁶

Doch auch Ianelli sieht darin in erster Linie eine Kompensation für ‚echte‘ menschliche Interaktion und fürchtet einen Sieg der Maschine über den Geist:

Hegel himself was interested in the complex and ambiguous relationship between human beings and technology, so in the last frontier of artificial empathy of our days he would have seen most likely, more a defeat than a victory of the spirit.⁸⁷

Ich glaube nicht, dass sich diese Hypothese halten lässt, könnte man doch das moderne *machine learning* durchaus als eine — wie auch immer defizitäre — Stufe der Entwicklung des Geistes bezeichnen, wie auch die Maschinenarbeit selbst.⁸⁸ Die eigentliche Gefahr kommt nicht von außen über den Geist, etwa aus der Maschinerie, sie liegt vielmehr innerhalb der Domäne des Geistes, nämlich als möglicher Sieg des Allgemeinen über das Besondere, oder wie Horkheimer sagt, die „Auslöschung der Individualität“.⁸⁹ Diese Gefahr besteht in der künstlichen wie in der menschlichen Intelligenz gleichermaßen. Die entscheidende Frage ist dabei aber, ob die KI, wenn sie denn die *Vollendung* der instrumentellen Vernunft ist, nicht auch ein gewisses emanzipatorisches Potenzial besitzt, welches über deren — und damit über ihre eigene — Beschränktheit hinausweist. Warum sollte die *Ambiguität*, auf die Horkheimer, Adorno und unzählige andere immer wieder hinweisen, nur eine negative und nicht auch eine positive Seite haben?⁹⁰

Technische Reproduzierbarkeit an sich muss jedenfalls kein Manko sein. In der eingangs zitierten Schrift *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit* zeigt Walter Benjamin auf, wie durch technologische Veränderungen im Produktions- und Reproduktionsprozess ein ganzer Kosmos scheinbar feststehender Wahrheiten infrage gestellt wird

⁸⁶ IANELLI, F., Hegel and Kaspar Hauser Between Reality And Fiction. *Bildung to the Test of Idiocy, Dehumanization, and Artificial Intelligence*, in: IANELLI F./FAILLA M., (Hg.), **Philosophy and Madness. From Kant to Hegel and Beyond**, Mimesis International, 2023, pp. 95-125, hier p. 118.

⁸⁷ Ebd.

⁸⁸ Hegel hat selbst die Maschinenarbeit, in der der Einzelne gewissermaßen zum abstrakten Anhängsel der Maschine wird — eine Kritik, die übrigens große Parallelen zur marxischen Theorie von Arbeitsteilung und Entfremdung aufweist —, in seinen Jenaer Vorlesungen zur Realphilosophie im Kapitel über den „wirklichen Geist“ verortet. Vgl. JR, p. 215, p. 232.

⁸⁹ KIV, p. 163.

⁹⁰ Horkheimer schließt diese Möglichkeit zumindest nicht aus: „Die industrielle Disziplin, der technische Fortschritt und die wissenschaftliche Aufklärung, gerade die ökonomischen und kulturellen Prozesse, die die Auslöschung der Individualität bewirken, versprechen – obgleich die Anzeichen gegenwärtig schwach genug sind – ein neues Zeitalter einzuleiten, in dem die Individualität als Element in einer weniger ideologischen und humaneren Daseinsform neu entstehen kann.“ (KIV, ebd.)

und beispielsweise Begriffe wie „Echtheit“, „Original“ etc. ihren Sinn verlieren. Zusammenfassend beschreibt er dies als „Zertrümmerung der Aura“.⁹¹ Der Kulturpessimist mag das beklagen, man kann darin aber auch wie Benjamin einen „kathartischen Effekt“ sehen, der mit der „Liquidierung des Traditionswertes am Kulturerbe“ verbunden ist,⁹² das heißt mit der Emanzipation „von seinem parasitären Dasein am Ritual“.⁹³

Können wir einen derartigen kathartischen Effekt auch der künstlichen Intelligenz zuschreiben? Immerhin steckt in dem Attribut „künstlich“ auch das Wort „Kunst“, und was künstlich ist, muss nicht unbedingt schlecht oder „fake“ sein. Gerade die Tatsache, dass Intelligenz künstlich reproduziert werden kann, könnte ein Zeichen für eine ihr inhärente Kraft sein, sich von Natur- und anderen Zwängen zu emanzipieren. Und die heute weit verbreitete Tendenz des, wie Hegel sagen würde, ‚verständigen Denkens‘, herablassend auf künstliche Intelligenz zu schauen und dieser eine gewisse Insuffizienz zuzuschreiben, könnte als ein Zeichen mangelnder Selbstreflexion dieses verständigen Denkens selbst gedeutet werden, da es seine eigene Dynamik nicht durchschaut und beim vorstellenden Bewusstsein stehen bleibt. Es merkt nicht, dass es bei der abstrakten Kritik künstlicher Intelligenz gewissermaßen in den eigenen Spiegel schaut.

Natürlich ist Kritik an der artifiziellen Intelligenz berechtigt — und notwendig. Sie muss aber auch eine Selbstkritik des Verstandesdenkens beinhalten, um ihre Kraft zu entfalten. Wie das technisch reproduzierbare Kunstwerk bei Benjamin eine „revolutionäre Kritik der überkommenen Vorstellungen von Kunst [...] befördern“ kann,⁹⁴ könnte ein Verständnis der Dynamiken von technisch reproduzierbarer Intelligenz auch unseren Blick auf die traditionelle, oftmals nur instrumentelle, Vernunft schärfen und zu einer modernen, nach-kantischen Vernunftkritik beitragen. Analog des marxischen Diktums, die Anatomie des Menschen sei ein Schlüssel zur Anatomie des Affen, könnte die Anatomie der KI auch Aufschluss über die Anatomie unseres kognitiven Bewusstseins geben.

Ob sich der weitere Fortschritt artifizieller Intelligenzparameter in Richtung Utopie oder doch eher in Richtung Dystopie entwickelt, können wir heute nicht absehen, genauso wenig wie die Frage, ob KI eines Tages tatsächlich eine höhere Form von Bewusstsein oder Selbstbewusstsein ausbilden kann. Marx‘ Analogie der Anatomie des Affen funktioniert nämlich nur

⁹¹ BENJAMIN, W., Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit, in: Ders., **Gesammelte Schriften (GS)**, Bd. I.2, Frankfurt am Main, Suhrkamp 1974, pp. 431-508, hier: p. 440 / p. 479.

⁹² Ebd., p. 439 / p. 478.

⁹³ Ebd., p. 442 / p. 481.

⁹⁴ Ebd., p. 492.

retrospektiv,⁹⁵ sodass wir zwar von der späteren Entwicklungsstufe (d.h. der KI) Erkenntnisse bezüglich der früheren (d.h. der menschlichen Intelligenz) ziehen, nicht aber aus der früheren die weiteren Entwicklungen der späteren ableiten können. Wir müssen uns aber bewusst sein, dass die weitere Entwicklung auch von uns selber abhängt. Gerade wenn wir KI nicht als Eigenschaft eines Automaten betrachten, sondern als Form einer systemisch-gesellschaftlichen Interaktion, bietet dies der Philosophie die Chance, aus dem von Hegel beschriebenen Dilemma der Eule der Minerva⁹⁶ auszubrechen und bei der Entwicklung künstlicher Intelligenz aktiv Akzente zu setzen. Für die Philosophie hieße das, diese Entwicklung der künstlichen Intelligenz nicht nur aufmerksam zu verfolgen, sondern vor allem kritisch an ihr mitzuwirken, Gefahren — etwa für Freiheit, Menschenwürde und Lebensqualität — rechtzeitig zu erkennen und auch das möglicherweise „Rettende“ (Hölderlin) darin zu identifizieren.

BIBLIOGRAFIE

1. Verzeichnis von Abkürzungen und Siglen

Diff	HEGEL, G.W.F., Differenz des Fichteschen und Schellingschen Systems der Philosophie , in: Theorie Werkausgabe (TWA) in 20 Bänden, Hg. Eva Moldenhauer und Karl Markus Michel, Frankfurt am Main, Suhrkamp, 1970. Band 2
Enz I-III	HEGEL, G.W.F., Enzyklopädie der philosophischen Wissenschaften , TWA Bände 8-10
JR	HEGEL G.W.F., Jenaer Realphilosophie , Hamburg, Meiner 1969 [1931]
KIV	HORKHEIMER, Max, Kritik der instrumentellen Vernunft , in: Ders., Gesammelte Schriften , Band 6, Frankfurt am Main, S. Fischer, 1991
PhG	HEGEL, G.W.F., Phänomenologie des Geistes , TWA Band 3
TWA	HEGEL, G.W.F., Theorie Werkausgabe in 20 Bänden, Hg. Eva Moldenhauer und Karl Markus Michel, Frankfurt am Main, Suhrkamp, 1970.
WdL I u. II	HEGEL, G.W.F., Wissenschaft der Logik , TWA Bände 5 u. 6

2. Literatur

ANDERL, Sibylle, **Kursbuch Montagsblock/310**, <https://kursbuch.online/sibylle-anderl-montagsblock/>

⁹⁵ „In der Anatomie des Menschen ist ein Schlüssel zur Anatomie des Affen. Die Andeutungen auf Höheres in den untergeordneten Tierarten können dagegen nur verstanden werden, wenn das Höhere selbst schon bekannt ist.“ MARX, K., **Grundrisse der Kritik der politischen Ökonomie (Rohentwurf)**, Frankfurt am Main, EVA, 1970, p. 26.

⁹⁶ ... wonach die Philosophie im historischen Veränderungsprozess immer zu spät kommt (vgl. PhR, Vorrede, p. 28).

- ARISTOTELES, **Nikomachische Ethik**, in: **Gesammelte Werke**, Andhof, Kindle-Ausgabe, 2015
- BENDER, Emily M., **On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?**, in: **FACCT'21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency**, New York 2021, pp. 610-623. (<https://doi.org/10.1145/3442188.3445922>)
- BENJAMIN, Walter, **Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit**, in: Ders., **Gesammelte Schriften (GS)**, Bd. I.2, Frankfurt am Main, Suhrkamp 1974
- BRANDOM, Robert B., **Artificial Intelligence and Analytic Pragmatism**, Chapter, 2008 – doi: 10.1093/acprof:oso/9780199542871.003.0003
- BRANDOM, Robert B., **A Spirit of Trust: A Reading of Hegel's 'Phenomenology'**, Harvard University Press, 2019
- CHURCHLAND, Patricia / SEJNOWSKI, Terence J., **The Computational Brain**, Cambridge, MIT Press, 1992
- DEWEY, John et al., **Creative Intelligence. Essays in the Pragmatic Attitude** [¹New York 1917], Kindle-Ausgabe, DD Books 2023
- DREYFUS, Hubert L., **What Computers Still Can't Do: A Critique of Artificial Reason**, Cambridge/Ms., MIT Press, Neuauflage 1992 [¹1972]
- FEIGE, Daniel Martin, **Kritik der Digitalisierung. Technik, Rationalität und Kunst**, Hamburg, Meiner, 2025
- HÖLDERLIN, Friedrich, **Patmos**, In: **Sämtliche Werke: Historisch-kritische Ausgabe**, hg. v. F. Beissner and A. Beck. Band 6, Stuttgart, Kohlhammer, 1975
- HORKHEIMER, Max / ADORNO, Theodor W., **Dialektik der Aufklärung**, in: HORKHEIMER, Max, **Gesammelte Schriften**, Band 5, Frankfurt am Main, S. Fischer, 1987
- IANELLI, Francesca / FAILLA, Mariannina (Hg.), **Philosophy and Madness. From Kant to Hegel and Beyond**, Mimesis International, 2023
- KOCH, Anton Friedrich, **Artificial Intelligence and the Problem of Consciousness, 2024** (https://www.academia.edu/113325577/Artificial_Intelligence_and_the_Problem_of_Consciousness_New_Version)
- KURZWEIL, Raymond, **Menschheit 2.0. Die Singularität naht**, Berlin, Lola Books, ²2014
- MARX, Karl, **Grundrisse der Kritik der politischen Ökonomie (Rohentwurf)**, Frankfurt am Main, EVA, 1970
- MENKE, Christoph, **Autonomie und Befreiung**, Berlin, Suhrkamp 2021
- NEGARESTANI, Reza, **Intelligence and Spirit**, Cambridge/Ms and London, Urbanomic / MIT Press, 2024 [¹2018]
- SCHUBERT, Alexander, **Der Strukturgedanke in Hegels „Wissenschaft der Logik“**, Berlin, Eule der Minerva, 2021
- SCHUBERT, Alexander, **Phänomenologie des Zeitgeistes. Mit Hegel durchs 21. Jahrhundert**, Wien, Passagen, 2022

WOLFF, Michael, **Das Körper-Seele-Problem. Kommentar zu Hegel, Enzyklopädie (1830), § 389**, Frankfurt am Main, Klostermann 1992

ŽIŽEK, Slavoj, **Hegel im verdrahteten Gehirn**, Frankfurt am Main, S. Fischer, 2020

ZUBOFF, Shoshana, **The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power**, London, Profile Books, 2019

RECEBIDO EM 21/04/2024

ACEITO EM 31/07/2025