

KI, Robotik und Transhumanismus

Rasanter Fortschritt, irreversibler Wandel und der Verlust menschlicher Autonomie

Essay im Horizont von „Sapere aude!“

1. Einleitung: Leitfrage, Problemaufriss, gesellschaftliche Diagnose

Kernthese: Die Zusammenführung von KI, Robotik und Transhumanismus verschiebt das Problem von „Souveränität zu Risiko“, weil die Beschleunigungsdynamik und die wachsende Systemabhängigkeit eine reale (demokratische) Steuerung zeitlich und strukturell überholen können.

„Sapere aude! Habe Mut, dich deines eigenen Verstandes zu bedienen!“ – Kants Aufklärungsformel ist weniger Zitatpflege als Diagnoseinstrument. Unmündigkeit, so Kant, ist bequem: Man lässt „ein Buch“ denken, „einen Seelsorger“ das Gewissen verwalten, „einen Arzt“ die Diät beurteilen – und erspart sich die Mühe des eigenen Urteilens.¹

In der Gegenwart erweitert sich diese Delegationskette um eine technisch vermittelte Instanz: Künstliche Intelligenz wird als „objektive“ Entscheidungshilfe eingeführt, die Auswahl, Prognose, Bewertung und Steuerung beschleunigen soll. Genau hier beginnt das Problem nicht erst dort, wo Systeme Fehler machen, sondern dort, wo sie als stiller Standard in institutionelle Entscheidungen einrücken und die Grenze zwischen verantwortlichem Urteil und automatisierter Zurechnung verschieben.

Die Leitfrage dieses Aufsatzes lautet daher: Unter welchen gesellschaftlichen Bedingungen transformiert sich KI – in Kopplung mit Robotik und transhumanistischen Optimierungsprogrammen – von einem Instrument der Entlastung zu einem Strukturprinzip der Entmündigung, und welche Konsequenzen ergeben sich daraus für Demokratie, Bildung und Persönlichkeitsbildung besonderer Qualität?²

Der normative Kern lässt sich mit Kants kategorischem Imperativ als institutionelle Prüfregel formulieren: Es ist zu fragen, ob jene Maximen, die in skalierende Systeme eingeschrieben werden, als allgemeines Gesetz gewollt werden können – nicht im

¹ Vgl. Kant 1784.

² Vgl. Russell 2019.

abstrakten, sondern im konkreten Sinn gesellschaftlicher Verallgemeinerbarkeit. Welche Regeln müssen dort gelten, wo sie massenhaft wirken, wo Einspruchswege faktisch erschwert sind und wo Betroffene den Kriterien nicht zustimmen können, weil diese Kriterien ihnen entzogen bleiben?

Die Gefährlichkeit der gegenwärtigen Formation liegt nicht in einem einzelnen Technologieelement, sondern in einer Anpassung, die qualitativ neue Machtwirkungen erzeugt.

Erstens bündelt KI Deutungs- und Handlungsmacht, die aufgrund von Intransparenz, Skalierung und Lernfähigkeit strukturell schwer zu begrenzen ist.

Zweitens verleiht Robotik dieser Macht körperliche Durchsetzungsfähigkeit in der materiellen Welt: Was zuvor Empfehlung war, wird Handlung; was Bewertung war, wird Zugriff.

Drittens liefert der Transhumanismus eine kulturell-normative Bedeutung, in der die Anpassung des Menschen an technische Maßstäbe als Fortschritt dargestellt wird – diese Anpassung verschärft sich durch ein Zeitproblem. Viele Entwicklungslinien verlaufen nicht linear, sondern sich steigernd: Zusätzliche Rechenleistung, größere Datensätze und verbesserte Trainingsverfahren können – bei bestimmten Aufgabenklassen – sprunghafte Fähigkeitsgewinne erzeugen, die sich erst im Nachhinein institutionell einholen lassen. In der öffentlichen Debatte wird dies oft als „exponentiell“ bezeichnet; analytisch genügt die Feststellung, dass die Gesellschaft nicht darauf vertrauen darf, Fortschritt werde stets in politisch gut absorbierbaren Schritten stattfinden. Prognosen zentraler Akteure im KI-Sektor die Zeithorizonte von wenigen Jahren für sehr weitreichende Fähigkeiten ins Spiel bringen, sind insofern nicht als Prognosegewissheiten zu behandeln, wohl aber als Warnsignale für eine Kontrollstruktur, die sich nicht im „Mitregulieren“ erschöpfen darf.³

Entscheidend ist die Asymmetrie zwischen Reversibilität und Irreversibilität. Institutionen integrieren KI schrittweise und deklarieren dies als kontrollierbaren Pilotbetrieb; zugleich entstehen Abhängigkeiten, die eine Abschaltung und einen Rückbau politisch und funktional immer teurer machen – bis ein „Stop“ nicht mehr als Option, sondern als gesellschaftliche Selbstschädigung erscheint. In dieser Logik erfolgt Verselb-

³ Vgl. Amodei 2026

ständigung nicht als öffentlich dramatischer Moment, sondern als stiller Übergang: Kontrolle wird nicht genommen, sie wird entbehrlich gemacht.

Diese Diagnose ist, nach Habermas gesprochen, eine Frage der Souveränität der Lebenswelt gegenüber Systemrationalitäten, die sich als Effizienzprogramme tarnen und doch die Bedingungen demokratischer Selbstregierung verändern.⁴

2. Konstitutive Theorien und Reformimpulse

Kernthese: Klassiker der Gesellschaftstheorie zeigen: KI, Robotik und Transhumanismus sind nicht neutral einfach „dazukommen“, sondern sie verstärken Verantwortungsverschiebung, Macht- und Ungleichheitsstrukturen.

2.1 Bourdieu: Kapitalformen, Habitus, symbolische Gewalt

Kernthese: KI erzeugt neue, hoch konzentrierte Kapitalformen (Daten-, Modell und Infrastrukturkapital), die als „Neutralität“ erscheinen und doch symbolische Macht sowie soziale Reproduktion stabilisieren.

Bourdieu's These, dass soziale Ungleichheit sich über die ungleiche Verteilung von Kapitalformen reproduziert, gewinnt im KI-Zeitalter besondere Schärfe.⁵ Denn die zentralen Ressourcen, die über Zugriff, Gestaltung und Deutung entscheiden – Daten, Rechenleistung, herstellenspezifische Modelle, Zertifizierungsregime, Plattformdistribution – bilden eine Gruppe technischer Kapitalformen, das sich hoch konzentriert. Selbst wenn KI scheinbar überall verfügbar ist, bleibt die Kontrolle über Kategorien, Trainingsregime, Modellziele und Integrationsbedingungen bei wenigen Akteuren; damit wird „Objektivität“ zur symbolischen Macht, die als Neutralität erscheint und doch Interessen strukturiert. Robotik verschärft diesen Mechanismus, weil Modell- und Datenkapital in physische Präsenz übersetzt werden: Wer robotische Infrastrukturen betreibt, kontrolliert Produktions-, Pflege- und Sicherheitsräume. Der transhumanistische Diskurs schafft neue Unterschiede: Körperliche und mentale Aufrüstung – bis hin zu neurotechnischen Schnittstellen, die das Gehirn direkt mit digitalen Systemen koppeln – wird zu einem Vorteil, der Zugang, Status und Zukunftschancen sortiert; zugleich wächst der Druck, diese Kopplung als „Normalität“

⁴ Vgl. Habermas 1981

⁵ Vgl. Bourdieu 1983

zu akzeptieren, während Nicht-Aufrüstung als persönliches Versäumnis ausgelegt werden kann.

Bildungs- und demokratiepolitisch folgt daraus: Es genügt nicht, „Kompetenzen“ zu vermitteln; notwendig ist Begriffsarbeit über die Produktionsbedingungen algorithmischer Macht, damit soziale Reproduktion nicht als technische Notwendigkeit missverstanden wird.

2.2 Frankfurter Schule: Instrumentelle Vernunft und Kulturindustrie

Kernthese: KI, Robotik und Transhumanismus radikalisieren instrumentelle Vernunft, indem sie Sinn- und Geltungsfragen in Zweck-Mittel-Optimierung übersetzen und Kritik durch Fortschrittssemantik entwerten.

Die Frankfurter Diagnose instrumenteller Vernunft beschreibt, wie Vernunft dort verarmt, wo sie sich auf Zweck-Mittel-Optimierung reduziert und Welt nur noch als Bestand der Verfügbarkeit behandelt.⁶

KI passt in diese Logik, weil sie Komplexität in Optimierungsprobleme übersetzt: Wahrheit, Geltung und Sinn werden sekundär gegenüber Vorhersage, Steuerung und Effizienz. Zugleich wirkt eine zeitgenössische Kulturindustrie, die Technik als notwendiger Fortschritt einordnet, als semantischer Motor: Begriffe wie „Innovation“, „Technologiesprung“ oder „Personalisierung“ erzeugen Zustimmung, bevor die Frage nach Macht gestellt wird.⁷ Robotik ist die Verkörperung dieser instrumentellen Vernunft: Nicht nur Entscheidungen werden rationalisiert, auch Körperlichkeit wird zur Funktionseinheit; der Mensch erscheint als Störfaktor in Prozessketten, die ohne seine Verwundbarkeit auskommen sollen. Transhumanismus liefert die passende Anthropologie dazu, indem er Grenzen als Defizite interpretiert und Angleichung an Systemmaßstäbe als Befreiung verkauft.⁸ Für Bildung folgt daraus: nicht Panik, sondern Urteilskraft. Die Fähigkeit stärken, auch Nein zu sagen – und Unverfügbarkeit und Kritik gegen den Druck zur bequemen Anpassung zu verteidigen.

2.3 Adorno: Mündigkeit und die Warnfigur „Erziehung nach Auschwitz“

⁶ Vgl. Horkheimer/Adorno 1988

⁷ Vgl. Zuboff 2018

⁸ Vgl. Habermas 2001

Kernthese: Wo entpersonalisierte KI Entscheidungen Verantwortlichkeit verdünnt, entsteht eine funktionale Willkür, die Mündigkeit durch technische Apparate ersetzt. Adornos Begriff der Mündigkeit ist eine politische Kategorie: Er meint nicht bloße Kompetenz, sondern die Haltung der Reflexion und Resistenz gegen autoritäre Verführungen.⁹ Seine Warnfigur „Erziehung nach Auschwitz“ erinnert daran, dass Zivilisationsbruch nicht nur aus offenem Fanatismus entstehen kann, sondern auch aus kalter Verwaltungsvernunft, Gehorsam, Ent-Verantwortlichung und der Delegation von Gewissen an Apparate. Genau hier liegt eine strukturelle Gefahr der KI-Integration: Wenn Entscheidungen ausgelagert werden und Verantwortung in Systemketten verwässert, wird der Satz „das System hat entschieden“ zum Ersatz für Zurechnung, Begründung und moralische Verantwortbarkeit.¹⁰ In Bildungs- und Verwaltungsinstitutionen droht so eine funktionale Eigenmächtigkeit, die nicht ideologisch laut auftritt, sondern prozedural leise wirkt: Scores, Prognosen und Präventionslogiken bestimmen Chancen, Zugänge und Ausschlüsse.

Robotik verschärft dies, weil Bewertung in Zugriff übergeht – und damit die Schwelle zwischen symbolischer und physischer Macht sinkt. Bildungspolitisch folgt daraus die Notwendigkeit eines Rechts auf menschliche Begründung und eines institutionell realisierbaren Widerspruchs: Ohne zurechenbare Verantwortung darf keine automatisierte Entscheidung in existenziellen Bereichen stattfinden.

2.4 Habermas: Kolonialisierung der Lebenswelt und kommunikatives Handeln

Kernthese: KI kolonialisiert Lebenswelt, wenn kommunikative Rechtfertigung durch systemische Steuerung ersetzt wird; demokratische Kontrolle muss daher als demokratische Mitkontrolle verbindlich verankert werden.

Habermas' Unterscheidung von System und Lebenswelt macht sichtbar, dass moderne Gesellschaften nicht nur durch Gewalt, sondern durch die stille Expansion von Steuerungslogiken entmündigt werden können.¹¹ Wo kommunikative Praxis durch administrative oder technische Routinen verdrängt wird, wird Demokratie äußerlich stabilisiert, aber innerlich ausgehöhlt. KI wirkt hier als Beschleuniger: Sie

⁹ Vgl. Adorno 1971

¹⁰ Vgl. NIST 2023.

¹¹ Vgl. Habermas 1981

standardisiert Kommunikation, reduziert Ambivalenz, transformiert komplexe Lebenslagen in bearbeitbare Fälle und verschiebt die normative Frage nach Gründen in eine operative Frage nach Kennzahlen. Werden solche Systeme in Schule, Verwaltung, Medizin oder Sicherheit integriert, sinkt die Dichte von Rechtfertigungspraktiken; die Öffentlichkeit verliert Zugriff auf Kriterien, weil Kriterien als technische Parameter erscheinen. Daraus folgt ein Reformmotiv: Nicht „menschliche Aufsicht“ als technisches Feigenblatt, sondern institutionalisierte Öffentlichkeit als demokratische Anforderung – Begründungspflichten, Verständlichkeitsgebote, wirksame Einspruchsrechte und eine Architektur, die Rückbau nicht nur technisch, sondern politisch ermöglicht.¹²

2.5 Honneth: Anerkennung, Verdinglichung und soziale Fehlentwicklungen

Kernthese: Bewertungskulturen verwandeln Personen in Profile und Risiken; Anerkennung wird durch Kennzahlen ersetzt, wodurch Missachtung strukturell normalisiert werden kann.

Honneth zeigt, dass moderne Gesellschaften soziale Fehlentwicklungen dort produzieren, wo Anerkennung verweigert und Menschen zu Objekten von Bewertung und Verfügung gemacht werden.¹³ KI-getriebene Punkte- und Rankinglogiken übertragen Personen in Profile, Risiken und Rankings; Beziehung wird durch Messung ersetzt, Biografie durch Prognose. Wer als Risiko klassifiziert wird, erfährt institutionalisierte Missachtung. Robotik verstärkt die Gefahr, weil sie körperliche Arbeit entwertet und ersetzbar macht; wo der Mensch als „Kostenfaktor“ erscheint, wird Anerkennung zur Frage der Anschlussfähigkeit an Systemtakte. Transhumanismus verschiebt zudem Normen so, dass Nicht-Optimierung als persönliches Versagen gerahmt werden kann – eine moralische Verschiebung, die Ungleichheit als selbstverständlich legitimiert.¹⁴

Anerkennungspolitisch und bildungstheoretisch ergibt sich daraus die Pflicht, dem Profil-Menschen den dialogischen Menschen entgegenzustellen: Bildung braucht

¹² Vgl. NIST 2023

¹³ Vgl. Honneth 1992

¹⁴ Vgl. Habermas 2001

Räume, in denen Subjekte nicht nur bewertet, sondern begründet verstanden und als verantwortlich anerkannt werden.

3. Spannungsfeld: Technikprimat vs. Demokratieprinzip

Kernthese: Ohne global koordiniertes, einschneidendes Reglement droht das Wettrennen eine Alternativlosigkeit zu erzeugen, in der demokratische Selbstgesetzgebung dem technischen Abhängigkeitsverhältnis nachgeordnet wird. Das Thema lässt sich als Konflikt zweier richtungsweisender Ordnungen beschreiben. Eine technolibérale Denkweise versteht Freiheit primär als Innovations- und Marktfreiheit: Was machbar ist, soll sich durchsetzen; Regulierung gilt als Hemmnis; Verantwortung wird individualisiert. Demgegenüber begreift ein demokratischer Republikanismus Freiheit als Nicht-Beherrschtwerden und als Fähigkeit kollektiver Selbstgesetzgebung. KI und Robotik verschieben die Konfliktlinie, weil Ausweichmöglichkeiten schrumpfen: Wo Verwaltung, Arbeitswelt und Bildung algorithmisch integriert sind, wird Nicht-Nutzung zur Exklusionsdrohung. Plattform- und Datenmacht erzeugen infrastrukturelle Abhängigkeiten und verlagern politische Fragen in die stille Logik des „Funktionierens“.¹⁵

Hier zeigt sich ein Kippunktproblem der Zeitstrukturen politischer Kontrolle: Demokratische Steuerung arbeitet langsam, deliberativ, rechtlich gebunden; techno-ökonomische Integration arbeitet schnell, schrittweise, abhängig.

Wird die technologische Infrastruktur zur Bedingung des Alltags, verschiebt sich die Debatte von „Wollen wir das?“ zu „Wie betreiben wir es effizienter?“. Spätestens dann erscheint „Weltherrschaft“ nicht als theatrale Dystopie, sondern als Extremform einer schon vorher wirksamen Logik: globale Infrastrukturen, zentralisierte Modellmacht und robotische Durchsetzungskapazität erzeugen eine Ordnung, in der menschliche Entscheidung formal existiert, aber faktisch nachgeordnet wird. Dass diese Dynamik durch ein globales Wettrennen befeuert wird, das Sicherheitsrationalität strukturell unter Druck setzt, ist kein Randaspekt, sondern Teil der Systemlogik.¹⁶

4. Konsequenzen für Bildung und Persönlichkeitsbildung

Kernthese:

¹⁵ Vgl. Zuboff 2018

¹⁶ Vgl. NIST 2023, vgl. Amodei 2026

Bildung muss als Praxis der Nicht-Delegation organisiert werden, damit Menschen unter Beschleunigungsdruck – physisch wie psychisch – nicht zu Nachzüglern eines Systems werden, das Würde an Anschlussfähigkeit koppelt. Persönlichkeitsbildung besonderer Qualität meint dann die Fähigkeit, Urteilskraft nicht auszulagern, Verantwortung nicht in Systemketten verschwinden zu lassen und begründete Entscheidungen gegen bloße Prognosen zu behaupten.¹⁷

„Sapere aude“ bedeutet im KI-Zeitalter nicht Technikverweigerung, sondern die institutionell gestützte Praxis, sich nicht durch Komfort, Effizienz und scheinbare Objektivität entmündigen zu lassen. Der kategorische Imperativ fungiert als Prüfstein für Systemdesign: Eine Gesellschaft kann nicht wollen, dass die Maxime delegiere existenzielle Entscheidungen an intransparente Systeme, wenn es effizient ist allgemeines Gesetz werde, ohne ihre eigenen Geltungsgrundlagen zu unterminieren.¹⁸

Die Zuspitzung der Zeitachse verschärft diese Bildungsfrage. Wenn Fähigkeitssprünge in Monaten auftreten können und wenn zugleich plausible Szenarien diskutiert werden, in denen sehr leistungsfähige Systeme binnen weniger Jahre entstehen, dann geraten Menschen – physisch wie psychisch – in die Rolle der Nachzügler. „Physisch“ meint: Robotik und Automatisierung können in kurzer Zeit Tätigkeitsfelder besetzen; „psychisch“ meint: Orientierung, Selbstwirksamkeit und Anerkennung geraten unter Druck, wenn menschliche Urteilskraft systematisch als langsamer, fehlerhafter, kostenintensiver dargestellt wird. Bildung muss deshalb nicht nur Technik verstehen lehren, sondern die Fähigkeit stärken, unter Beschleunigungsdruck begründet zu Widersprechen und die eigene Würde nicht in Anschlussfähigkeit aufzulösen.

5. Mögliche Korrekturen im System

Kernthese: Nur ein global koordiniertes Reglement und eine Architektur der Reversibilität können verhindern, dass Beschleunigung und Abhängigkeit demokratische Kontrolle überholen.

Korrekturen müssen dort ansetzen, wo Irreversibilität entsteht: in Infrastruktur, Verfahren, Verantwortungsarchitekturen und Normverschiebungen. Erstens ist eine Beweislastumkehr für Hochrisiko-KI erforderlich: Nicht Kritiker müssen Gefahren

¹⁷ Vgl. Kant 1784

¹⁸ Vgl. Kant 1784

nachweisen, sondern Betreiber müssen vor Integration zeigen, dass Systeme reversibel betreibbar sind – mit Rückbauplänen, Not-Abkopplungen, auditierbaren Kriterien und einer öffentlichen Begründbarkeit der eingesetzten Entscheidungsregeln.¹⁹ Zweitens braucht Bildung verbindliche Curricula zur Analyse algorithmischer Macht, die nicht beim Toolgebrauch stehen bleiben, sondern Kategorien, Fehlerklassen, Verantwortungszuschreibung und demokratische Kontrollmechanismen zum Gegenstand machen; Didaktisch sind Formate zu stärken, die das Urteilen durch Abwägen und Begründen einüben, weil sie jene Fähigkeit trainieren, die Systeme nicht ersetzen können: begründen, abwägen, widersprechen.

Drittens muss in allen existenziellen Bereichen ein Recht auf menschliche Begründung und ein praktisch erreichbares Widerspruchsverfahren gelten, sodass der Satz „das System hat entschieden“ seine entlastende Funktion verliert und wieder zur zurechenbaren Entscheidung zurückgeführt wird.

Viertens sind robotische Hochrisiko-Funktionen in kritischen Infrastrukturen so zu regulieren, dass Internet- und Cloud-Abhängigkeiten begrenzt werden und physische Souveränitätsbremsen existieren: lokale Kontrolle, zertifizierte Abschaltketten, begrenzte Aktionsräume und unabhängige Prüfstellen.²⁰

Fünftens ist eine transhumanismusbezogene Grundrechtspolitik nötig: Schutz vor Optimierungszwang in Bildung, Beruf und Versicherungssystemen, demokratisch legitimierte Grenzziehungen zwischen Therapie und nicht-therapeutische Verbesserung des Menschen sowie institutionelle Kontrolle über Norm- Verschiebungen, die sonst als „Fortschritt“ dargestellt werden.

Sechstens müssen Produktivitätsgewinne aus Automatisierung sozialpolitisch in Zeit- und Anerkennungsgewinne übersetzt werden, damit Ersetzbarkeit nicht zur Entwertung und Entsolidarisierung führt.

Schluss

Kernthese: „Sapere aude“ muss als institutionelles Gegenprogramm zur technisierten Bequemlichkeit verstanden werden, damit Mündigkeit, Anerkennung und Verantwortungszuschreibung nicht im Normalbetrieb der Systeme verschwinden.

¹⁹ NIST 2023

²⁰ Vgl. Winner 1980

Die hier entfaltete Kritik richtet sich nicht gegen Technik, sondern gegen eine gesellschaftliche Form, in der Technik zum Schicksal erklärt wird. „Sapere aude“ ist kein romantischer Appell, sondern ein institutionelles Programm: Mündigkeit entsteht nicht allein durch moralische Zumutung, sondern durch Strukturen, die eigenes Denken ermöglichen und Delegation begrenzen.

Wo KI, Robotik und transhumanistische Optimierung als Gesamtheit die Bedingungen der Verantwortungszuschreibung, der Anerkennung und der öffentlichen Begründung verändern, steht nicht nur Effizienz auf dem Spiel, sondern die Frage, ob menschliche Entwicklung als offene Bildungsgeschichte fortgeschrieben wird – oder als geplante, selektive Entmündigung.

Es ist eine bedrückende Zeitdiagnose, dass jene Werte, die Aufklärung und Bildung im emphatischen Sinn tragen – Kritikfähigkeit, Emanzipation und verantwortliche Selbstbestimmung – im öffentlichen und institutionellen Handeln zunehmend randständig werden. An ihre Stelle tritt, oft kaum verhüllt, die Logik der Gewinnmaximierung: Technikentwicklung wird primär als Standort- und Machtfrage behandelt, nicht als Frage legitimer Grenzen. Historisch ist dabei umso auffälliger, dass demokratische Systeme die Notwendigkeit institutionalisierter Technikfolgenprüfung längst erkannt hatten: Der US-Kongress schuf mit dem Technology Assessment Act (1972) das Office of Technology Assessment (OTA), das ab 1974 über zwei Jahrzehnte unabhängige Analysen zu wissenschaftlich-technischen Zukunftsfragen bereitstellte – bis es 1995 im Zuge politischer Kräfteverschiebungen und unter dem Druck jener Interessenlagen, die Begrenzung als „Innovationshemmnis“ rahmen, faktisch abgewickelt wurde. Dieser Befund ist nicht bloß historisch, sondern symptomatisch: Wo Reflexion institutionell geschwächt wird, gewinnt Tempo; wo Tempo gewinnt, wird Kontrolle zur nachträglichen Fiktion. Dass selbst Großmächte wie die USA und China in einen beschleunigten Entwicklungskampf eintreten, in dem die Aussicht auf Superintelligenz als Hebel ökonomischer und geopolitischer Dominanz gilt, verschärft die strukturelle Gefahr: Sicherheitsbedenken werden dem Wettbewerb untergeordnet, Reglementierung wird zum Nachteil umcodiert, Rückbauoptionen werden politisch unattraktiv. In dieser Konstellation hängt das Damoklesschwert nicht als Metapher über „den anderen“, sondern über allen: KI, Robotik und transhumanistische Aufrüstungsprogramme bilden eine Verstärkerschleife, die gesellschaftliche Bedingungen in

kurzer Zeit und potenziell irreversibel zu Ungunsten menschlicher Autonomie und demokratischer Selbststeuerung verändern kann. Gerade deshalb wäre Fatalismus fehl am Platz – doch ebenso unzulässig ist die Illusion, es genüge ein späteres Nachregulieren. Wenn ein Einlenken noch möglich sein soll, dann nur als frühzeitige, globale und einschneidende Rahmensetzung, bevor Beschleunigung und Abhängigkeit politische Steuerungsfähigkeit endgültig überholen.

(Dr. Thielen 2025)

Literaturverzeichnis

Adorno, Theodor W. (1971): *Erziehung zur Mündigkeit*. Frankfurt am Main:

Suhrkamp. Amodei, Dario (2026): *The Adolescence of Technology*. Online-Essay.

Ars Technica (2025): *Anthropic chief says AI could surpass “almost all humans at almost everything” within two to three years*. Online-Artikel (22.01.2025).

Bourdieu, Pierre (1983): *Ökonomisches Kapital, kulturelles Kapital, soziales Kapital*. In: Kreckel, Reinhard (Hg.): *Soziale Ungleichheiten*. Göttingen: Otto Schwartz, 183–198.

Habermas, Jürgen (1981): *Theorie des kommunikativen Handelns*. Band 2: *Zur Kritik der funktionalistischen Vernunft*. Frankfurt am Main: Suhrkamp.

Habermas, Jürgen (2001): *Die Zukunft der menschlichen Natur. Auf dem Weg zu einer liberalen Eugenik?* Frankfurt am Main: Suhrkamp.

Horkheimer, Max / Adorno, Theodor W. (1988): *Dialektik der Aufklärung. Philosophische Fragmente*. Frankfurt am Main: Fischer.

Honneth, Axel (1992): *Kampf um Anerkennung. Zur moralischen Grammatik sozialer Konflikte*. Frankfurt am Main: Suhrkamp.

Kant, Immanuel (1784): *Beantwortung der Frage: Was ist Aufklärung?* *Berlinische Monatsschrift*, Dezemberheft.

Kant, Immanuel (1785): *Grundlegung zur Metaphysik der Sitten*. Riga: Johann Friedrich Hartknoch.

NIST (2023): *AI Risk Management Framework (AI RMF 1.0)*, veröffentlicht 26.01.2023. National Institute of Standards and Technology. Online-Ressource.

Russell, Stuart (2019): *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking. Winner, Langdon (1980): *Do Artifacts Have Politics?*