

Stefan Müller

Technologischer Fatalismus und das Ende der akademischen Bildung

Zusammenfassung

Unter der Prämisse der unvermeidlichen Entwicklung einer allgemeinen künstlichen Intelligenz (AGI) in den kommenden Jahren und den daraus folgenden tiefgreifenden gesellschaftlichen Veränderungen wird in dem Essay anhand von drei Szenarien eine zukünftige Ausprägung von akademischer Bildung beschrieben.

Das Fortbestehen akademischer Bildung in einer Welt mit AGI erscheint als sehr voraussetzungsreich und damit insgesamt fraglich. Inwiefern die veröffentlichten Annahmen von Technologieunternehmen als Grundlage für Entwicklungen im Hochschulwesen taugen, bleibt offen. Das weitgehende Ignorieren der Entwicklung von KI-Technologien in der Hochschulpolitik – im Sinne eines analogen Eskapismus – ist jedoch eine ebenso radikale wie längerfristig riskante Position.

1 Einleitung

„Wir neigen dazu, die Auswirkungen einer neuen Technologie auf kurze Sicht zu überschätzen, sie aber auf lange Sicht zu unterschätzen.“
(vermutlich Roy Amara)

Mit der Veröffentlichung von ChatGPT entbrannte eine breite gesellschaftliche Debatte über die Fähigkeiten generativer KI-Technologie und deren Auswirkungen auf gesellschaftliche Praktiken. Neben der Debatte hätte die Veröffentlichung der damals noch unzureichenden Version von ChatGPT den relevanten Akteur:innen in der Gesellschaft Zeit verschafft, Anpassungen zu entwickeln und zu erproben¹. Je-

¹ In dem Zusammenhang formulierte Sam Altman (CEO von OpenAI): „Societal changes that ChatGPT is going to cause or is causing. There's a big one going now about the impact of this on education, academic integrity, all of that. But starting these now where the stakes are still relatively low, rather than just putting out what the whole industry will have in a few years with no time for society to update, I think would be bad.“ (Altman 2023a)

doch blieb dies an den Hochschulen² ebenso wie in weiten Teilen der restlichen Gesellschaft über weite Strecken aus.

In der Hochschulpolitik lässt sich der Mangel an Anpassung unter anderem am Entwurf der sogenannten „Strategie der digitalen Transformation im Hochschulbereich“ (SMWK 2023) sowie an den Schwierigkeiten vieler Hochschulleitungen beobachten, eine fundierte Positionierung der Hochschulen in Bezug auf KI-Technologie zu entwickeln und Empfehlungen für deren Umgang zu geben (Solis 2023)³.

Demgegenüber zeichnen Berichte und Prognosen von Unternehmen, die KI-Technologie entwickeln, eine Parallelwelt, in der der technologische Fortschritt exponentiell voranschreitet. Neue emergente KI-Fähigkeiten⁴ sowie die Entwicklung einer allgemeinen KI (AGI) bzw. Superintelligenz⁵ werden als absehbar und unvermeidlich betrachtet. Dies scheint in der internen und externen Kommunikation ebenso Konsens zu sein, wie die dramatischen Auswirkungen auf alle Bereiche des Lebens, auf menschliche Aktivitäten und Gesellschaft (Sutskever, 2023). Es wird eine nahe Zukunft⁶ beschrieben, in der sich gesellschaftliche Strukturen, Lo-

² „Hochschule“ wird hier als Oberbegriff für Universitäten, medizinische Hochschulen, Musik- und Kunsthochschulen, Berufsakademien sowie Fachhochschulen genutzt. Dabei werden Hochschulen vordergründig in ihrer Funktion als Bildungseinrichtungen betrachtet.

³ Suleyman sieht das Meiden der Diskussion auf politischer Ebene in der deprimierenden Vorläufigkeit möglicher Antworten begründet. Es nennt dies „Pessimismus-Aversion“ (2023a). Möglicherweise setzen viele politische Entscheidungsträger auf eine Stagnation in der technologischen Entwicklung (Tegmark 2023a). Agüera y Arcas & Norvig sehen das Zögern, das Maß an Intelligenz in den aktuellen KI-Modell anzuerkennen, wie folgt begründet: eine gesunde Skepsis gegenüber Metriken für AGI, ein ideologisches Bekenntnis zu alternativen KI-Theorien oder KI-Techniken, eine Hingabe an den menschlichen (oder biologischen) Exzeptionalismus sowie Bedenken hinsichtlich der wirtschaftlichen Auswirkungen von AGI (2023).

⁴ Mit emergenten Fähigkeiten werden Fähigkeiten von KI-Systemen bezeichnet, auf die das KI-System nicht trainiert wurde, die jedoch ab einer gewissen Größe des KI-Systems unvorhergesehen auftreten. Eine Auflistung von „137 emergent abilities of large language models“ ist bei Wei zu finden (2022). Es wird angenommen, dass weitere emergente Fähigkeiten, wie das Planen über verschiedene Zeithorizonte, Zwischenschritte in der Entwicklung von AGI sein könnten (Suleyman, 2023b).

⁵ Mit allgemeiner KI (AGI) wird ein technisches System beschrieben, das generell klüger als Menschen ist (Altmann, 2023b) und weitgehend autonom wertvollere (ökonomische) Arbeit leistet als es Menschen in der Lage wären (OpenAI, 2023a). Forschende von Google DeepMind differenzieren AGI zwischen (meta-)kognitiven Fähigkeiten auf dem Niveau ungelerner Menschen (Emerging AGI) bis hin zu einem Übertreffen aller Menschen (Artificial Superintelligence) (Morris et al., 2023).

⁶ KI-Expert:innen gehen davon aus, dass die Entwicklung von AGI in den nächsten Jahrzehnten oder sogar in den nächsten Jahren gelingt (Roser, 2023). In den Prognosen von Technologieunternehmen werden Zeiträume von weniger als 10 Jahren genannt (Altman, Brockman & Sutskever, 2023; Tao, 2023; Leike & Sutskever, 2023). Suleyman

giken und Praktiken im gesamten Spektrum von Utopie⁷ bis Dystopie⁸ grundlegend verändern werden.

Die Veröffentlichungen der Technologieunternehmen dienen primär deren eigenen Interessen und nicht dem Bedienen eines allgemeinen, öffentlichen Erkenntnisinteresses. Es gab Fälle, in denen sich Ankündigungen als übertrieben herausstellten⁹ und relevante Ergebnisse zurückgehalten wurden¹⁰. Unabhängige, wissenschaftliche Publikationen zur Entwicklung, Beschaffenheit und Leistungsfähigkeit der „State of the Art“-KI-Systeme der führenden Technologieunternehmen liegen nicht vor. Öffentliche Forschungseinrichtungen verfügen nicht über die Ressourcen der Technologieunternehmen¹¹ und haben normalerweise keinen direkten Zugriff auf die (unveröffentlichten) KI-Systeme der Technologieunternehmen. Es werden aber

hält es für plausibel, dass innerhalb der nächsten 3 bis 5 Jahre KI-Technologie mit menschlichen Fähigkeiten in allen kognitiven Feldern konkurrieren könnte (2023a). Die abweichenden Definitionen von AGI führen auch zu sehr unterschiedlichen Zeithorizonten für das Entstehen von AGI. Agüera y Arcas & Norvig argumentieren, dass rückblickend die aktuellen KI-Modelle als die ersten Beispiele für AGI gelten werden (2023).

⁷ Altman erwartet, dass Arbeitskosten „gegen Null“ fallen, sobald eine ausreichend leistungsstarke KI in die Arbeitswelt einsteigt. Daraus würde eine enorme Steigerung des Lebensstandards unter der Bedingung folgen, dass es der Politik gelingt, eine entsprechende Ressourcenverteilung zu gewährleisten (Altman, 2021).

⁸ Bedenken hinsichtlich einiger der schwerwiegendsten Risiken der fortgeschrittenen KI-Technologie werden in offenen Briefen (z. B. Future of Life Institute, 2023) und Statements (z. B. Center for AI Safety, 2023) veröffentlicht. Viele zentrale Akteur:innen aus den Technologieunternehmen beteiligen sich an den Veröffentlichungen. Die Motive der Beteiligungen sind allerdings unklar. So könnten die offenen Briefe und Statements auch der Darstellung der Mächtigkeit und der politischen Relevanz der Technologieunternehmen dienen.

⁹ Die umfangliche Bereitstellung der im März 2023 angekündigten KI-gestützten Steuerung von Officeanwendungen durch „Microsoft 365 Copilot“ (Spataro, 2023) wurde ebenso wie die angekündigte Multimodalität von GTP-4 (OpenAI, 2023b) erst Ende September der Allgemeinheit zugänglich gemacht. In einem Demonstrationsvideo für Googles KI-Modell Gemini wurden Fähigkeiten suggeriert, über die das Modell gar nicht verfügt (Coldewey, 2023).

¹⁰ Beispielsweise hatte ein Team von Microsoft Research Monate vor der Veröffentlichung von ChatGPT (November 2022 auf Grundlage von GPT-3) bereits Zugriff auf eine frühe Version von GPT-4. Erst im März 2023 wurden Ergebnisse der Experimente veröffentlicht (Bubeck et al., 2023). Benaich et al. konstatieren, dass ehemals transparente Technologieunternehmen wie Google und OpenAI bezüglich der Forschungsergebnisse zu ihren generativen KI-Systemen eine „Kultur der Opazität“ eingeführt haben (2023).

¹¹ „Im Jahr 2022 gab es 32 bedeutende, von der Industrie produzierte Modelle für maschinelles Lernen, verglichen mit nur drei, die von der öffentlichen Wissenschaftseinrichtungen erstellt wurden. Der Aufbau hochmoderner KI-Systeme erfordert zunehmend große Mengen an Daten, Rechenleistung und Geld – Ressourcen, über die Branchenakteure im Vergleich zu gemeinnützigen Organisationen und der Wissenschaft von Natur aus in größeren Mengen verfügen“ (Maslej et al., 2023).

auch die konzeptionellen Ansätze der Technologieunternehmen hinterfragt. So herrscht keineswegs Einigkeit innerhalb der KI-Forschungsgemeinschaft, ob generative KI-Technologie eine Grundlage für die Entwicklung von AGI sein kann¹². Ebenso wird die Annahme eines transformativen Potenzials der KI-Systeme kritisch betrachtet (z. B. Ramani & Wang 2023). Den Veröffentlichungen von Technologieunternehmen ist daher mit angemessener Skepsis zu begegnen.

Der Zweck des Essays besteht darin, tiefere Einblicke in die Rolle und den Einfluss von Technologieunternehmen auf gesellschaftliche Bereiche zu gewähren und ein deutliches Bewusstsein für die Thematik an Hochschulen zu schaffen. Angesichts der zunehmenden Relevanz und des wachsenden Einflusses dieser Unternehmen, insbesondere im Hinblick auf ihr Bestreben, AGI zu entwickeln, ist eine kritische Auseinandersetzung mit den damit einhergehenden gesellschaftlichen Konsequenzen unerlässlich. Zudem zielt der Essay darauf ab, die Herausforderungen und Unwägbarkeiten einer Welt mit AGI zu konturieren. Während Analogien – wie Tegmarks Vergleich zwischen Kindern und Erwachsenen (2017) oder Moravecs Analogie des ansteigenden Meeresspiegels (1998) – nur begrenzt Aufschluss bieten, soll entlang von drei Szenarien die möglichen Auswirkungen von AGI auf die akademische Bildung greifbar gemacht werden. Obwohl die Szenarien nur vereinfachend und holzschnittartig gelingen, sollen sie dem Zweck dienen, provokative und diskursanregende Überlegungen auf einer Metaebene zu fördern. Es soll in einer dynamischen und von Unsicherheit geprägten Zeit der Diskurs darüber angeregt werden, wie Hochschulen auf die rasante Entwicklung von KI-Technologie reagieren könnten. Die Radikalität der Prämisse soll dem Diskurs um mögliche Anpassungen an Hochschulen (z. B. Amlung et al., 2023) unterstützen.

2 Technologischer Fatalismus

Mit dem Begriff des technologischen Fatalismus soll hier die Vorstellung beschrieben werden, dass die Entwicklung einer allgemeinen künstlichen Intelligenz in den kommenden Jahren unvermeidlich (fatal) ist und mit tiefgreifenden gesellschaftlichen Veränderungen einhergehen wird. Die technologische Revolution wird als unaufhaltsam angesehen (Altman, 2021) und die Nutzung der Technologie erscheint als unvermeidlich (Hendrycks, 2023).

Da der technologische Fatalismus maßgeblich von Technologieunternehmen vorangetrieben wird, werden auch deren Denkmuster und Weltanschauungen in den

¹² Mitchell & Krakauer untersuchten die aktuelle, hitzige Debatte in der KI-Forschungsgemeinschaft darüber, ob LLMs Sprache sowie die physischen und sozialen Situationen, die Sprache kodiert, in einem menschenähnlichen Sinne verstehen (2023). Im Verstehen kann eine Grundlage für logisches Ableiten, agentisches Verhalten und AGI gesehen werden.

gesellschaftlichen Diskurs übertragen. Es können die Denkmuster und Weltanschauungen im Kontext der Technologieunternehmen entlang von drei Ausprägungen skizziert werden:

- ein spieltheoretisches Verständnis von Menschen, Organisationen und Gesellschaften, das von Wettbewerb und Machtstreben geprägt ist¹³,
- Ansätze eines transhumanistischen Menschenbildes, in denen die Verbesserung des fehlerbehafteten Menschen mit Hilfe von Technologie propagiert wird¹⁴ und
- eine tendenziell kultische Verehrung von technologischem Fortschritt¹⁵.

Wenn unter der Annahme des technologischen Fatalismus die Zukunft der akademischen Bildung prognostiziert wird, ist es ebenso wichtig, die Denkmuster und Weltanschauungen der Technologieunternehmen zu berücksichtigen.

¹³ Tegmark beschrieb das kompetitive Umfeld der KI-Entwicklung als „Moloch“: eine Logik, nach der Akteure in einem Wettlauf ein Ziel anstreben, das sowohl die Chance auf enorme Gewinne als auch das Risiko der eigenen Zerstörung beinhaltet. Tegmark bemüht in dem Sinne eine Analogie, in der auf eine Klippe zugerast wird: “The closer to the cliff you go, the more money there is, (...) so you want to drive there very fast. But it’s not in anyone’s incentive that we go over the cliff, and it’s not like everybody’s in their own car. All the cars are connected together with a chain. So if anyone goes over, they’ll start dragging others down too“ (2023b). Technologieunternehmen wie OpenAI haben sich gänzlich der Entwicklung von AGI verschrieben: „Anything that doesn’t help with that [AGI] is out of scope“ (OpenAI, 2023c).

¹⁴ AGI würde durch die Steigerung von Wohlstand und dem Finden neuer wissenschaftlicher Erkenntnisse die Menschheit verbessern (Altmann, 2023b). Darüber hinaus würde eine synthetische Biologie einen „new dawn for humanity“ bedeuten (Suleyman, 2023a).

¹⁵ „Es wäre unangemessen, OpenAI als Kult zu bezeichnen, aber viele Führungskräfte des Unternehmens glauben, dass jemand, der nicht an die Ankunft der Allgemeinen Künstlichen Intelligenz (AGI) und deren Bedeutung als eines der größten Ereignisse in der Menschheitsgeschichte glaubt, sich wahrscheinlich nicht wohl bei OpenAI fühlen würde. Sie fragen sich, warum jemand, der nicht daran glaubt, dort arbeiten möchte. Es wird davon ausgegangen, dass die Belegschaft (...) sich aus Personen zusammensetzt, die diese Vision teilen“ (WIRED, 2023). Es wurde berichtet, wie der Chief Scientist von OpenAI Sutskever in Erwartung der Ankunft der allmächtigen Technologie begann, sich wie ein spiritueller Führer zu verhalten. Sein ständiger, enthusiastischer Refrain sei gewesen: „Feel the AGI“ (The Atlantic, 2023).

3 Szenarien zur akademischen Bildung im Zeitalter der Allgemeinen Künstlichen Intelligenz

„Es ist ganz wahr, was die Philosophie sagt, daß das Leben rückwärts verstanden werden muß. Aber darüber vergisst man den andern Satz, daß vorwärts gelebt werden muß.“
(Sören Aaby Kierkegaard)

Szenarien basieren auf ausgewählte Prämissen und angenommenen Bedingungen. Angesichts der Unsicherheit bezüglich der zukünftigen Entwicklung kann zumindest die Gültigkeit der Prämissen und angenommenen Bedingungen überprüft und die Folgerichtigkeit der Szenarien untersucht werden. Neben der Prämisse des „technologischen Fatalismus“ werden die Bedingungen „Institutionelle Trägheit“ und „Gesellschaftliche Basis“ (Bedarf an akademisch qualifizierten Fachkräften und Bereitstellung notwendiger Ressourcen) als relevante Variablen für eine zukünftige akademische Bildung gesetzt. Auf dieser Grundlage werden folgend drei Szenarien skizziert.

		Gesellschaftliche Basis	
		gegeben	gegeben
Institutionelle Trägheit	hoch	Szenario 1 „Pro-Forma Hochschule“	Szenario 3 „Das Ende der akademischen Bildung“
	gering	Szenario 2 „Transformierte und transformierende Hochschule“	

Abbildung 1: Szenarien zur akademischen Bildung im Zeitalter der AGI

3.1 Szenario 1: „Die Pro-forma-Hochschule“

Das erste Szenario beschreibt eine „Pro-Forma-Hochschule“. Die Voraussetzungen für dieses Szenario sind eine Gesellschaft, die weiterhin akademische Bildung benötigt und trägt sowie eine ausgeprägte institutionelle Trägheit in den (öffentlichen) Hochschulen. Die Anpassungsfähigkeit der Hochschule wird durch systemimmanente Merkmale (Hochschule als „unvollständige Organisation“ (Brunsson & Sahlin-Andersson, 2000) oder „organisierte Anarchie“ (Cohen et al, 1972)) sowie durch externe Faktoren wie Finanzierung und politischer Wille eingeschränkt. Dadurch würde an überkommenen Praktiken, Normen und Strukturen festgehalten. Die Diskrepanz zwischen dem raschen externen technologischen Fortschritt („technologischer Fatalismus“) und der internen Trägheit der Hochschulen („institutioneller Fatalismus“) hemmt das Innovationspotenzial der Hochschulen und die Schaffung relevanter Praktiken in der akademischen Bildung.

Allerdings zeigt sich diese Trägheit auch in anderen Teilen der Gesellschaft. Die Vorstellung, dass ein Hochschulabschluss Voraussetzung für verantwortungsvolle Tätigkeiten in hoch angesehenen Berufen ist, bleibt weit verbreitet. Bildung wird immer noch eng mit Hochschulen verknüpft. Eine Abschaffung der Hochschule als Bildungseinrichtung ist daher politisch nicht durchsetzbar. Die Hochschule bleibt als Struktur bestehen: als „Pro-Forma-Hochschule“.

Lehr- und Prüfungspraxis der Gegenwart

Das Szenario zeigt sich gegenwärtig in Bemühungen, KI-Technologie zu verbieten oder zu ignorieren. Es wird auf die Fehleranfälligkeit oder Unreife der KI-Systeme hingewiesen. Mit solchen Argumenten werden Anpassungen in Lehre und Prüfung abgelehnt.

Studium der nahen Zukunft

Es wird an dem traditionellen Zweck der Hochschulbildung festgehalten: der Reproduktion von Wissen. Die Vorstellung, dass Studierende hauptsächlich „Wissen auf Vorrat“ erwerben sollen, entkoppelt das Studium zunehmend von der Arbeits- und Lebenswelt. Punktuelle Versuche, KI-Technologie inhaltlich relevant zu integrieren oder KI-gestütztes Lehren und Lernen zu ermöglichen, können nicht mit dem Tempo des technologischen Fortschritts mithalten und finden in der Hochschule wenig Resonanz. Technologieunternehmen bieten effizientere Möglichkeiten für den Erwerb akademischer Kompetenzen an als Hochschulen¹⁶. Das studentische Lernen findet zunehmend außerhalb der Hochschule statt. Die Hochschule vollzieht einen Wandel, von einer umfänglichen Bildungseinrichtung hin zu einer formalen Instanz für Leistungsnachweise und zur Vergabe von Abschlüssen.

Akademische Bildung im Zeitalter der AGI

Die Hochschule wird funktional überflüssig: (Aus-)Bildung kann durch KI-Systeme schneller, effizienter und flexibler ermöglicht werden als durch traditionelle akademische Einrichtungen. Hochschulen dienen für Menschen lediglich zur ritualisierten Distinktion.

3.2 Szenario 2: „Die transformierte und transformierende Hochschule“

Für das Szenario der „transformierten und transformierenden Hochschule“ ist eine agile Hochschule erforderlich, die sowohl eine Offenheit für technologische Entwicklungen als auch einen gewissen „wertebasierten Konservatismus“ aufweist.

¹⁶ Googles „Career Certificates“, Microsofts „Career Essentials in Generative AI“ und die von OpenAI angekündigte Academy können hierfür als Beispiele dienen.

Die Hochschule begegnet dem technologischen Fatalismus mit einem humanistischen Determinismus¹⁷. Ethische Grundsätze werden den Denkmustern und Weltanschauungen der Technologieunternehmen entgegengesetzt und transhumanistischen Vorstellungen wird entgegengetreten. Die Anwendung von KI-Technologie soll sich den menschlichen Bedürfnissen und der menschlichen Entfaltung unterordnen (Buyx, 2023). Ethische und soziale Fragen stehen im Mittelpunkt der Debatte über die gesellschaftliche Entwicklung. Mit kollektivistisch-sozialen Überzeugungen wird dem Machtstreben der Technologieunternehmen begegnet. Der Einsatz von KI-Technologie wird an soziale Themen wie gesellschaftliche Teilhabe und Ressourcenverteilung geknüpft. Kultischen Ansätzen wird mit Wissenschaftlichkeit und kritischem Denken entgegengetreten. Die Leistung der Hochschulen wird (weiterhin) als eine Investition in „natürliche Intelligenz“ bzw. in „blended intelligence“ (Eberhart et al., 2015) verstanden.

Die Hochschule passt sich humanistisch determiniert den technologischen Entwicklungen an, behält allerdings gegenüber Technologieunternehmen die Hoheit und stellt sich einem Technologieoptimismus entgegen. Sie versteht sich als Gestalterin der Gesellschaft und wirkt so den Denkmustern und Weltanschauungen der Technologieunternehmen mit einem wertebasierten Konservatismus entgegen.

Lehr- und Prüfungspraxis der Gegenwart

Lehrende, Studierende und Akteur:innen in der Hochschulverwaltung erproben Variationen der Prüfungs- und Lehrpraxis, bei denen KI-Technologie inhaltlich und methodisch integriert wird. Dabei orientieren sie sich an fachlichen, wissenschaftlichen und ethischen Standards¹⁸ sowie an den individuellen Voraussetzungen der Akteur:innen.

Studium der nahen Zukunft

Alle Curricula der Hochschulen bereiten die Studierenden inhaltlich, methodisch und ethisch auf eine Arbeits- und Lebenswelt vor, die von KI-Technologie unterstützt wird¹⁹. Hochschulen fungieren als Orte, an denen KI-Technologie im gesellschaftlichen und ethischen Kontext verstanden wird. KI-Technologie wird aktiv in

¹⁷ Ein Ziel, dem sich bereits universitäre Einrichtungen wie das „Institute for Human-Centered AI“ an der Stanford University verschrieben haben.

¹⁸ So verschließt sich die Deutsche Forschungsgemeinschaft in einer Stellungnahme nicht den generativen Modellen für die Text- und Bilderstellung, fordert jedoch die Einhaltung etablierter Standards guter wissenschaftlicher Praxis (2023). Der Deutsche Ethikrat hat eine umfassende, multidisziplinäre Stellungnahme zu den ethischen Fragen des Verhältnisses von Mensch und Maschine erarbeitet (2023).

¹⁹ Es liegen Prognosen vor, dass bereits durch die Integration der gegenwärtigen KI-Systeme in den (Office-) Anwendungen ein Großteil der akademischen Tätigkeiten effizienter gestaltet werden können (Eloundou, et al., 2023). Best-Practice-Ansätze zur Anpassung der Curricula liegen vor (z. B. Übersicht in Mah & Torner, 2022).

die Gestaltung von Lehre und Lernen integriert, beispielsweise durch personalisierte Lernpläne und individuelle KI-Tutoren für Studierende. Hochschulen entwickeln in Kooperation KI-Werkzeuge für Lehr-, Lern- und Forschungsprozesse oder kooperieren mit Technologieunternehmen unter strikten, ethischen Vorgaben. Die Hochschulen sind zu Orten geworden, an denen kritisches Denken, ethische Abwägungen und kreative Fähigkeiten gefördert werden, die von KI noch nicht ersetzt werden können. Hochschulen betonen die Bedeutung sozialer Beziehungen und kollektiver Anstrengungen. Sie positionieren sich als „Denkfabriken“ für die aktuellen und zukünftigen Herausforderungen einer Welt mit AGI.

Akademische Bildung im Zeitalter der AGI

In einer Welt, in der AGI viele kognitive Aufgaben übernehmen kann, könnte die Notwendigkeit hochspezialisierter, akademischer Berufe erheblich abnehmen. Damit büßt die Hochschule ihre Funktion als Qualifikationsinstanz für Berufs- oder Wissenschaftsfelder weitgehend ein und übernimmt stattdessen eine breitere Bildungs- und Transformationsrolle. Die neue Rolle soll hier entlang von vier möglichen Ausgestaltungen skizziert. Die Ausgestaltung adressieren jeweils eine Reaktion der Hochschule auf gesellschaftliche Disruptionen im Zuge des Aufkommens von AGI. Eine Kombination der Ausgestaltungen ist ebenso denkbar.

Hochschule als Ort des menschlichen Logos: In einer AGI-Welt könnten neue Arten der Realitätsbewältigung entstehen. Es könnte (a) eine „Maschinenwirklichkeit“ entstehen. Bereits heute fällt es Menschen häufig schwer, menschengemachte Aussagen kriteriengeleitet zu hinterfragen, auf Belegbarkeit, logische Fundiertheit und zukünftige Folgen zu untersuchen. Eine von AGI geschaffene Wirklichkeitsbeschreibung („Maschinenwirklichkeit“) würde von Menschen noch schwerer überprüfbar sein. Gleichzeitig würden mit der „Maschinenwirklichkeit“ sehr funktionale Aussagen und Handlungen ermöglicht werden. Ebenso könnte (b) ein „digitaler Pantheismus“ oder ein „Technotheismus“ entstehen, in dem AGI von Menschen als göttlich oder als Manifestation des Göttlichen angesehen wird. Es könnte eine Religion entstehen, deren Glaubensinhalte direkt von einer übermenschlichen Instanz erzeugt wurden und das Göttlich für Menschen technisch-mittelbar zugänglich wäre.

Die Hochschule könnte als Sphäre des „menschlichen Logos“ dienen. Der „Maschinenwirklichkeit“ oder dem „Technotheismus“ könnte eine Wirklichkeitsbeschreibung und -bewältigung entgegengestellt werden, die auf einen „menschlichen Logos“, also auf für Menschen nachvollziehbare und begründbare Erklärungen basiert.

Hochschule als Low-Tech-Ort des zwischenmenschlichen Diskurses: In einer Welt mit AGI ist jede Kommunikation auf Distanz dem Verdacht ausgesetzt, eine „Maschinenkommunikation“ zu sein: adaptiv, effizient und unabhängig von physischen Grenzen, aber auch ein Ungleichgewicht an Macht, Wissen und Ressourcen.

Die Hochschule könnte die Funktion eines technologiesparsamen Ortes für direkte zwischenmenschliche Interaktion²⁰ sein. Menschen könnten an der Hochschule eine Entschleunigung und Alternative zu den Interaktionen in einer immersiven, virtuellen Welt finden.

Hochschule als Ort der Moderationsausbildung: In einer Welt mit AGI werden menschliche Expertise und menschliche kognitive Fähigkeiten in relevanten gesellschaftlichen Prozessen (z. B. in Wirtschaft und Kultur) kaum noch benötigt, da KI-Technologie in diesen Bereichen überlegen ist. Akademische Berufe und Tätigkeitsfelder könnten weitgehend verschwinden. Stattdessen werden Menschen benötigt, die mit Maschinen arbeiten und dabei auf Alignment²¹, Vertrauenswürdigkeit und Werteorientierung achten. Diese Menschen tragen dazu bei, eine Beziehung zwischen AGI und Menschen zu gestalten.

Hochschulen bereiten auf diese Tätigkeitsfelder vor. Es könnten Moderator:innen ausgebildet werden, die an der Schnittstelle zwischen Mensch und Maschine arbeiten können. Hochschulen forschen an Mechanismen zur Überprüfung der Vertrauenswürdigkeit und des Alignments von KI-Systemen. Es werden Methoden erprobt, mit denen die Entscheidungen und Prognosen von KI-Systemen erklärbar und transparent gemacht werden (z. B. durch Ansätze wie „Explainable Artificial Intelligence“).

Hochschule als Ort der Philosophie: Durch AGI wird das Selbstverständnis des Menschen als überlegene Lebensform in Frage gestellt. In einer Welt ohne AGI konnten Konzepte wie Bewusstsein, Intelligenz und Kreativität vom Menschen her gedacht werden: als anthropologische Differenz. Mit dem Aufkommen von AGI werden das Selbstverständnis des Menschen sowie Konzepte wie Bewusstsein, Intelligenz und Kreativität überdacht werden. Kants philosophische Fragen (Was kann ich wissen? Was soll ich tun? Was darf ich hoffen? Was ist der Mensch?) würden in einer Welt mit AGI neu verhandelt werden.

Hochschulen könnten als Orte dienen, an denen philosophische Diskussionen über die menschliche Sinnhaftigkeit und das menschliche Selbstbild in Anbetracht einer kognitiven Überlegenheit der Maschinen geführt und erforscht werden. Akademi-

²⁰ Reinmann nennt als ein „hochschuldidaktische Utopie“ die „Aufwertung der sozialen Beziehung“ in Studium und Lehre an Hochschulen als ein „Gegengewicht zur Mensch-Maschine-Relation“ (2023).

²¹ Mit „Alignment“ wird ein Zustand beschrieben, in dem sich KI-Entitäten im Einklang mit menschlichen Werten und Zielen befinden (Bostrom, 2014)

sche Bildung könnte ein Beitrag sein, sich in einer Welt mit AGI gesellschaftlich zu engagieren, Verantwortung zu übernehmen oder einen sinnstiftenden Lebensweg zu finden.

3.3 „Das Ende der akademischen Bildung“

Die vorherigen Szenarien 1 und 2 setzen voraus, dass eine Gesellschaft nicht nur einen Bedarf an akademisch qualifizierten Fachkräften hat, sondern auch die notwendigen Ressourcen bereitstellen kann. Diese Voraussetzungen könnten in einer Welt mit AGI vakant sein.

Gegenwärtig wird der Bedarf an akademischer Qualifikation eng mit komplexen Lebens- und Arbeitswelten verknüpft, in denen die kognitiven Fähigkeiten von Menschen eine zentrale Rolle spielen. Zukünftig könnte AGI auf unterschiedlichen Ebenen der Lebens- und Arbeitswelten wirksam werden – mit der Folge, dass der Bedarf an akademischer Qualifikation vergeht.

- Persönliche KI-Assistenten könnten Menschen in die Lage versetzen, funktionale Entscheidungen in den Lebens- und Arbeitswelten zu treffen. Damit blieben Menschen zwar in den zentralen Rollen, benötigen aber zur Ausfüllung dieser keine akademische Ausbildung mehr.
- AGI könnte in einer agentischen Form direkt zur Bearbeitung komplexer Probleme eingesetzt werden und Menschen aus den traditionellen Rollen verdrängen.
- AGI könnte – in einer konsequenten Verwirklichung transhumanistischer Ansätze – die physische und psychische Beschaffenheit des Menschen weiterentwickeln, was zu einer radikal neuen menschlichen Konstitution führen könnte, die sicher nicht auf akademische Bildung angewiesen wäre.

Neben dem Bedarf an akademischer Qualifikation ist ebenso ungewiss, ob und in welcher Form in einer Welt mit AGI eine Gesellschaft existieren kann, die Ressourcen für den Betrieb von Hochschulen bereitstellt. Es scheint zweifelhaft, dass ein Alignment bzw. eine Regulation von AGI erreicht werden kann²². Doch selbst wenn eine Regulation der größten KI-Modelle führender Technologieunternehmen gelänge, bliebe die Kontrolle über Open-Source-Modelle als Herausforderung bestehen. In dem Zusammenhang beschreiben Harari den Verlust der menschlichen

²² „Derzeit haben wir keine Lösung, um eine potenziell superintelligente KI zu steuern oder zu kontrollieren und zu verhindern, dass sie abtrünnig wird. Unsere aktuellen Techniken zur Ausrichtung der KI, wie beispielsweise das verstärkende Lernen aus menschlichem Feedback, basieren auf der Fähigkeit des Menschen, die KI zu überwachen. Aber Menschen werden nicht in der Lage sein, KI-Systeme, die viel intelligenter sind als wir, zuverlässig zu überwachen, und daher werden unsere derzeitigen Ausrichtungstechniken nicht zur Superintelligenz skalierbar sein. Wir brauchen neue wissenschaftliche und technische Durchbrüche.“ (Leike & Sutskever, 2023)

Kontrolle über die eigene Entwicklung, als das „Ende der menschlichen Geschichte“ (2023). Nach Ansicht von Barrat (2013) wäre AGI „Our Final Invention“. Neben einer nicht-kontrollierbaren AGI steht auch der Gebrauch von AGI zu militärischen, nachrichtendienstlichen oder terroristischen Zwecken ein Risiko für die Gesellschaft dar.

Selbst wenn eine Gesellschaft durch die Existenz oder den Missbrauch von AGI nicht gefährdet wäre, könnten die ökonomischen und sozialen Strukturen durch AGI in einer Weise transformiert werden (Altman, 2023c), die eine Ressourcenbereitstellung für Hochschulen unmöglich macht. Die Rolle von Erwerbsarbeit und die Bedeutung von Geld könnten sich grundlegend wandeln, während sich die Machtverhältnisse in der Gesellschaft neuformieren oder sich soziale Ungleichheiten zuspitzen. Ebenso könnte der soziale Zusammenhalt und das gesellschaftliche Vertrauen erodieren, insbesondere wenn es nicht mehr möglich ist, zwischen menschlicher und künstlicher Kommunikation zu unterscheiden.

In Anbetracht dessen erscheint es unwahrscheinlich, dass in einer Welt mit AGI eine Gesellschaft existiert, die einen Bedarf an akademisch ausgebildeten Menschen hat und Ressourcen für den Betrieb der Hochschulen bereitstellen kann. In dem Szenario 3 „Das Ende der akademischen Bildung“ gibt es keine tragende Grundlage für die Hochschule, unabhängig von der Anpassungsleistung der Hochschulen.

Literatur

- Agüera y Arcas, B. & Norvig, P. (2023). *Artificial General Intelligence Is Already Here*. Berggruen Institute. Online verfügbar: <https://www.noemamag.com/artificial-general-intelligence-is-already-here>
- Altman, S. (2021). *Moore's Law for Everything*. Online verfügbar: <https://moores.samaltman.com>.
- Altman, S. (2023a). *StrictlyVC in conversation with Sam Altman*. Interview mit Connie Loizos. Transkript online verfügbar: <https://www.lesswrong.com/posts/PTzsEQXkCfig9A6AS/transcript-of-sam-altman-s-interview-touching-on-ai-safety>.
- Altman, S. (2023b). *Planning for AGI and beyond*. OpenAI Blog. Online verfügbar: <https://openai.com/blog/planning-for-agi-and-beyond>.
- Altman, S. (2023c). *Sam Altman: OpenAI CEO on GPT-4, ChatGPT, and the Future of AI*. Lex Fridman Podcast #367. Online verfügbar: https://www.youtube.com/watch?v=L_Guz73e6fw.
- Altman, Brockman & Sutskever (2023). *Governance of superintelligence*. OpenAI Blog. Online verfügbar: <https://openai.com/blog/governance-of-superintelligence>.
- Amlung, F., Plankenbichler, P., Müller, S., Heinz, M., Rada, U., Freudenreich, R., Marquardt, J., Baier, J., Heitz, R., Grunert, C. (2023). *Hochschuldidaktische Perspektiven auf KI-Technologien*. Hochschuldidaktik Sachsen, Blog. Online verfügbar:

- bar: <https://www.hd-sachsen.de/veroeffentlichungen/blog/hochschuldidaktische-perspektiven-auf-ki-technologien>.
- Barrat, J. (2013). *Our Final Invention: Artificial Intelligence and the End of the Human Era*. Thomas Dunne Books.
- Benaich, N., Chalmers, A., Sebbouh, O. & Gurau, C. (2023). *State of AI Report 2023*. Online verfügbar: <https://www.stateof.ai>.
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Brunsson, N., & Sahlin-Andersson, K. (2000). *Constructing Organizations. The Example of Public Sector Reform*. *Organization Studies*, 21 (S. 721-746).
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., Zhang, Y. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. Microsoft Research. Online verfügbar: <https://arxiv.org/pdf/2303.12712.pdf>.
- Buyx, A. (2023). *Ethikrat: Künstliche Intelligenz darf menschliche Entfaltung nicht vermindern*. Deutscher Ethikrat. Online verfügbar: <https://www.ethikrat.org/mitteilungen/mitteilungen/2023/ethikrat-kuenstliche-intelligenz-darf-menschliche-entfaltung-nicht-vermindern>.
- Coldewey, D. (2023). *Google's best Gemini demo was faked*. Techcrunch. Online verfügbar: <https://techcrunch.com/2023/12/07/googles-best-gemini-demo-was-faked>.
- Center for AI Safety (2023). *Statement on AI Risk*. Online verfügbar: <https://www.safe.ai/statement-on-ai-risk>.
- Cohen, M. D., March, J. G., & Olsen, J. P. (1972). *A Garbage Can Model of Organizational Choice*. *Administrative Science Quarterly*, 17 (S. 1-25).
- Deutsche Forschungsgemeinschaft (2023). *KI, ChatGPT und die Wissenschaften – DFG formuliert Leitlinien für Umgang mit generativen Modellen zur Text- und Bilderstellung*. Information für die Wissenschaft Nr. 72. Online verfügbar: https://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_23_72/index.html.
- Deutscher Ethikrat (2023). *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz*. Stellungnahme. Online verfügbar: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
- Eberhart, R., Palmer, D., Kirschenbaum, M. (2015). *Beyond computational intelligence: blended intelligence*. Swarm/Human Blended Intelligence Workshop (SHBI), Cleveland, OH, USA, 2015 (S. 1-5).
- Eloundou, T., Manning, S., Mishkin, P., Rock, D. (2023). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*. Online verfügbar unter <https://arxiv.org/pdf/2303.10130.pdf>
- Future of Life Institute (2023). *Policymaking In The Pause*. Online verfügbar: <https://futureoflife.org/open-letter/pause-giant-ai-experiments>.
- Harari, Y. N. (2023). *What does the AI revolution mean for our future?* Mustafa Suleyman & Yuval Noah Harari im Interview mit The Economist. Online verfügbar: <https://www.youtube.com/watch?v=7JkPWHr7sTY>.
- Hendrycks, D. (2023). *Natural Selection Favors AIs over Humans*. Center for AI Safety. Online verfügbar: <https://arxiv.org/pdf/2303.16200.pdf>.

- Leike & Sutskever, S. (2023). *Introducing Superalignment*. OpenAI Blog. Online verfügbar: <https://openai.com/blog/introducing-superalignment>.
- Mah, D.-M. & Torner, C. (Hrsg.) (2022). *Anwendungsorientierte Hochschullehre zu Künstlicher Intelligenz*. Impulse aus dem Fellowship-Programm zur Integration von KI-Campus-Lernangeboten. Online verfügbar: https://ki-campus.org/sites/default/files/2023-02/2022_KI-Campus_Fellowship-Sammelband_Nr.2.pdf.
- Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J. & Perrault, R. (2023). *The AI Index 2023 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023. Online verfügbar: <https://aiindex.stanford.edu/report>.
- Mitchella, M. & Krakauer, D. C. (2023). *The debate over understanding in AI's large language models*. PNAS. Online verfügbar: <https://www.pnas.org/doi/10.1073/pnas.2215907120>
- Moravec, H. (1998). *When will computer hardware match the human brain?* Journal of Evolution and Technology. Vol. 1. Online verfügbar: <https://jetpress.org/volume1/moravec.htm>.
- Morris, M. R., Sohl-dickstein, J., Fiedel, N., Warkentin, T., Dafoe, A., Faust, A., Farabet C. & Legg, S. (2023). *Levels of AGI: Operationalizing Progress on the Path to AGI*. Google DeepMind. Online verfügbar: <https://arxiv.org/pdf/2311.02462.pdf>.
- OpenAI (2023a). *How should AI systems behave, and who should decide?* Online verfügbar unter <https://openai.com/blog/how-should-ai-systems-behave>.
- OpenAI (2023b). *GPT-4 Technical Report*. Online verfügbar: <https://arxiv.org/pdf/2303.08774.pdf>.
- OpenAI (2023c). *Core values*. Online verfügbar: <https://openai.com/careers>.
- Ramani, A. & Wang, Z. (2023). *Why transformative artificial intelligence is really, really hard to achieve*. The Gradient. Online verfügbar: <https://thegradient.pub/why-transformative-artificial-intelligence-is-really-really-hard-to-achieve>.
- Reinmann, G. (2023). *Deskilling durch Künstliche Intelligenz? Potenzielle Kompetenzverluste als Herausforderung für die Hochschuldidaktik*. Diskussionspapier Nr. 25. Berlin: Hochschulforum Digitalisierung. Online verfügbar: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_DP_25_Deskillung.pdf.
- Roser, M. (2023). *AI timelines: What do experts in artificial intelligence expect for the future?* Online verfügbar: <https://ourworldindata.org/ai-timelines>.
- Sutskever, I. (2023). *The Exciting, Perilous Journey Toward AGI*. TED. Online verfügbar: <https://www.youtube.com/watch?v=SEKGLj0bwAU>.
Campus.pdf.
- SMWK (2023). *Strategie der digitalen Transformation im Hochschulbereich*. Sächsischen Staatsministeriums für Wissenschaft, Kultur und Tourismus. Online verfügbar: <https://www.studieren.sachsen.de/digitale-transformation-im-hochschulbereich-4681.html>.

- Solis, T. (2023). *Die ChatGPT-Richtlinien der 100 größten deutschen Universitäten*. Scribbr. Online verfügbar: <https://www.scribbr.de/ki-tools-nutzen/chatgpt-universitaere-richtlinien>.
- Spataro, J. (2023). *Introducing Microsoft 365 Copilot – your copilot for work*. Official Microsoft Blog. Online verfügbar: <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work>.
- Suleyman, M. (2023a). *The Coming Wave. AI, synthetic biology and a new dawn for humanity*. Bodley Head.
- Suleyman, M. (2023b). *What does the AI revolution mean for our future?* Mustafa Suleyman & Yuval Noah Harari im Interview mit The Economist. Online verfügbar: <https://www.youtube.com/watch?v=7JkPWHr7sTY>.
- Tao, T. (2023). *Embracing Change and Resetting Expectations*. In: Eric Horvitz (ed.), *AI Anthology*. Online verfügbar: <https://unlocked.microsoft.com/ai-anthology/terence-cao>.
- Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. Allen Lane.
- Tegmark, M. (2023a). *The 'Don't Look Up' Thinking That Could Doom Us With AI*, TIME USA. Online verfügbar: <https://time.com/6273743/thinking-that-could-doom-us-with-ai>.
- Tegmark, M. (2023b). *The Case for Halting AI Development*, Lex Fridman Podcast #371. Online verfügbar: <https://www.youtube.com/watch?v=VcVfceTsD0A>.
- The Atlantic (2023). *Inside the Chaos at OpenAI*. Online verfügbar: <https://www.theatlantic.com/technology/archive/2023/11/sam-altman-open-ai-chatgpt-chaos/676050>.
- Wei, J. (2022). *137 emergent abilities of large language models*. Online verfügbar: <https://www.jasonwei.net/blog/emergence>.
- WIRED (2023). *What OpenAI Really Wants*. Online verfügbar: <https://www.wired.com/story/what-openai-really-wants>.